# Meritocracy and Its Discontents:
## Long-run Effects of Repeated School Admission Reforms

Mari Tanaka Yusuke Narita Chiaki Moriguchi*

October 17, 2020

**Abstract**

We study the short-run and long-run impacts of changing admissions systems in higher education. We take advantage of the world's first known implementation of nationally centralized admissions and its subsequent reversals in early twentieth-century Japan. This centralization was designed to make admissions more meritocratic, but our analysis shows that there was a sharp tradeoff between meritocracy and equal regional access to higher education and career advancement. Specifically, in the short run, the meritocratic centralization led students to make more inter-regional and ambitious applications. However, as high ability students were located disproportionately in urban areas, increased regional mobility caused urban applicants to crowd out rural applicants from higher education. Moreover, the impacts were persistent. Four decades later, compared to the decentralized admissions, the meritocratic centralization increases the number of urban-born elites (e.g., top income earners) relative to rural-born elites. We also find that the meritocratic centralization produced more top-ranking bureaucrats relative to the decentralized system.

*Keywords*: Elite Education, Market Design, Strategic Behavior, Regional Mobility, Universal Access, Persistent Effects

*Preliminary and Incomplete: Not for Circulation*

*Yeon-Koo Che, Pat Kline, Al Roth, Chris Walters, and seminar participants at Berkeley, Stanford, Rice, and the NBER provided helpful feedback. We thank Hidehiko Ichimura and Yasuyuki Sawada for sharing a part of the Japanese Personnel Inquiry Records data, and an extensive list of students for research assistance. We gratefully acknowledge financial support from JSPS KAKENHI Grant Number JP19K13719, Yale Council on East Asian Studies, Hitotsubashi University IER Joint Usage and Research Center Grant (FY 2018). Mari Tanaka: Department of Economics, Hitotsubashi University, email: mari.tanaka@r.hit-u.ac.jp. Yusuke Narita: Department of Economics, Cowles Foundation, and Y-RISE, Yale University, email: yusuke.narita@yale.edu. Chiaki Moriguchi: Institute of Economic Research, Hitotsubashi University, email: chiaki@ier.hit-u.ac.jp.

# 1 Introduction

College and school admission processes vary across time and places. For example, American college admissions use a decentralized system where each college makes its own admissions often based on opaque evaluation criteria. In contrast, China's public college admissions illustrate a regionally-integrated, single-application, and single-offer system using a transparent admission criterion. How do different admissions institutions affect students' behavior and future life?

This paper studies the short- and long-run consequences of making school admissions meritocratic and centralized. We do so by combining a series of natural experiments in history, newly assembled historical data, and economic theory. Our theoretical and empirical investigations reveal the pros and cons of centralized meritocratic admissions, especially a tradeoff between meritocracy and equal regional access to selective higher education and career achievements.

Our empirical setting is the first known transition from decentralized to nationally-centralized school admissions. At the end of the 19th century, to modernize its higher education system, the Japanese government set up elite national schools (high schools or colleges in today's system) that served as an exclusive entry point to the most prestigious tertiary education. These schools later produced many of the most influential members of the society, including several Prime Ministers, Nobel Laureates, and founders of global companies like Toyota. Acceptance into these schools was merit-based, using annual entrance examinations. Initially, the government let each school run its own exam and admissions based on exam scores, similar to many of today's decentralized K-12 and college admissions. The schools typically held exams on the same day so that each applicant could apply for only one school. Similar restrictions on the number of applications exist today in the college admission systems of Italy, Japan, Nigeria, and the UK.

At the turn of the 20th century, the government introduced a centralized system in order to improve the quality of incoming students. In the new system, applicants were allowed to rank multiple schools in the order of their preferences and take a single unified exam.[1] Given their preferences and exam scores, each applicant is assigned to a school (or none if unsuccessful) based on a computational algorithm. The algorithm was a mix of the so-called Immediate Acceptance (Boston) algorithm and Deferred Acceptance algorithm with a meritocracy principle that the only highest-achieving applicants can get into any school. To the best of our knowledge, this instance is the first recorded, nation-wide use of any matching

[1] As shown later, the defining feature of the centralized system is to allow applicants to list multiple schools, not the use of a single unified exam.

algorithm.[2] Furthermore, for reasons detailed below, the government later re-decentralized and re-centralized the system several times, producing multiple natural experiments for studying the consequences of the different systems.[3] We provide evidence that the reform timings are close to exogenous.

We exploit these bidirectional institutional changes to identify the impacts of the meritocratic centralization. We first use a stylized theoretical model to predict the impacts of centralization on application behavior and admissions. Consistent with the stated goal of centralization, we first confirm that the centralized system produces more meritocratic school seat allocations. Our model also predicts that centralization would cause applicants to apply to more selective schools and make more inter-regional applications, increasing regional mobility. These theoretical results guide our empirical analysis.

For constructing the dataset for our analysis, we newly digitalized several historical sources. From Government Gazettes and administrative documents by the Ministry of Education, we assemble data on applicants, their applications, and birth prefectures for 1898-1930. We combine this data with a complete administrative list of entrants by school, year, and birth prefecture, using the administrative Higher School Student Registers published annually by each school. To our knowledge, these documents have not been used in economics research.

We find that meritocratic centralization had large short-run effects on both application behavior and enrollment outcomes. First, consistent with the theoretical predictions, centralization caused stark strategic responses in application behavior. In particular, strategic incentives in the centralized system led both urban and rural applicants to more frequently rank the most selective school first.[4] Second, the centralized system caused a greater number of high-ability applicants from urban areas to be admitted to schools in rural areas, often after being rejected by their first choice schools. As a result, urban high-achievers crowded out rural applicants; the number of entrants to any national elite school coming from the urban area increases by about 10% during centralization.[5]

[2] The earliest known large-scale use of the Boston algorithm is the assignment of medical residents to hospitals in New York City in the 1920s (Roth, 1990). The oldest known national use of the Deferred Acceptance algorithm is the National Resident Matching Program (NRMP) in the 1950s (Roth, 1984). See Abdulkadiroğlu and Sönmez (2003) for the details of these algorithms in school admission contexts.

[3] The admission system was centralized in 1902, decentralized in 1908, re-centralized in 1917, re-decentralized in 1919, re-centralized in 1926, and re-decentralized in 1928, resulting in three periods of centralization (1902-07, 1917-18, and 1926-27).

[4] We use the nomenclature of "urban" and "rural" schools, but note that "rural" schools were located in regional cities rather than in the countryside.

[5] It is also empirically true that the centralized system made a greater number of rural applicants apply to and enter urban schools. The centralized system thus increased regional mobility across the country. But their net effects are such that urban high-achievers crowded out rural applicants.

Historical documents suggest that this distributional consequence upset rural schools and communities. Partly as a result of such rural discontents, the government went back and forth between decentralized and centralized systems, finally settling for a decentralized scheme. This series of bidirectional reforms enables us to identify the causal effects of meritocratic centralization more precisely than a usual, single policy change would.

Most importantly, we find that meritocratic centralization had lasting impacts on students' career outcomes. Since the centralized system is designed to be more meritocratic and high-achieving students are disproportionately located in urban areas, centralization is expected to let urban areas disproportionately gain school access relative to rural areas. Our short-run analysis confirms this expectation. This result motivates us to compare long-term career outcomes of urban- and rural-born individuals by each cohort's exposure to the centralized system.

We use this difference-in-differences approach to estimate the distributional effects of meritocratic centralization on urban- vs rural-born individuals. The career outcome data come from the two editions of the Japanese Personnel Inquiry Records (JPIR) published in 1934 and 1939, more than thirty years after the first episode of meritocratic centralization. The JPIR is an equivalent of Who's Who in Japan that provides a list of highly distinguished individuals (e.g., high-income earners, national medal recipients, high-ranking government officials) along with their personal information. We provide extensive investigations about the quality of this long-term outcome data.[6]

Our estimates suggest persistent effects of meritocratic centralization. Almost four decades later, relative to the decentralized system, the centralized system produced a greater number of top income earners, prestigious medal recipients, and other elite professionals who came from urban areas compared to rural areas. Quantitatively, the number of urban-born career elites increased by 10-20% for the cohorts exposed to the centralized admissions. We also obtain suggestive evidence that, in the long run, the centralized system increased the number of career elites residing in urban areas in their middle age relative to those residing in rural areas. The design of admission systems therefore affects the geographic origins and destinations of highly educated and skilled individuals, also known as "upper-tail human capital" (Mokyr, 2005). The geography of elites is an important determinant of economic growth and inequality (Glaeser, 2011; Moretti, 2012; Autor, 2019).

Finally, we find that the meritocratic centralization produced more elite bureaucrats

[6] We find that the data covers a large fraction of the national population of elites (e.g., 53% of the top 0.01% income earners) by comparing our data with national statistics. We also find no systematic variation in the sampling rates across prefectures, consistent with our assumption that sample selection bias is uncorrelated with the prefecture-cohort variation we use. Dell and Parsa (2019) also use the Japanese Personnel Inquiry Records in their analysis.

relative to the decentralized system. That is, the cohorts exposed to the centralized system were more likely to become top-raking government officers than the cohorts exposed to decentralized system.

In total, the impacts of meritocratic centralization highlight an equity-meritocracy tension, both in the short- and long-run. On the one hand, the centralized system achieved the goal of rewarding applicants with higher academic performance. On the other hand, this meritocracy came at the cost of urban applicants dominating rural applicants. This distributional effect turned out to be persistent after decades.

We focus on the scholastic and career elites on both substantive and pragmatic grounds. The individual holdings of earnings and prestigious executive, bureaucratic, cultural, and scientific positions are skewed and concentrated on only a tiny minority of the population. This tiny minority often have big influences on the economy and society. The social importance of top elites also resulted in relatively rich data, making it possible to trace back their careers, from humble beginnings as students to wide recognition and acclaim.

**Literature.** Our analysis sheds light on the causal impacts of selective admission systems, contributing to the literature on their effects on application behavior, regional mobility, and applicants' academic achievement and welfare (Abdulkadiroğlu et al., 2006, 2009, 2017; Calsamiglia et al., 2010; Avery et al., 2014; Pallais, 2015; Machado and Szerman, 2017; Hafalir et al., 2018; Carvalho et al., 2019; Grenet et al., 2019; Knight and Schiff, 2019).[7] While these prior studies focus on the short-run effects, we estimate the long-run effects by taking advantage of bidirectional, repeated policy changes in history. This use of bidirectional policy changes echoes other studies with similar identification strategies (Niederle and Roth, 2003; Redding and Sturm, 2008).

With its interest in long-run effects, this paper also relates to studies of the long-term effects of educational resources (Duflo, 2001; Currie and Moretti, 2003; Meghir and Palme, 2005; Oreopoulos, 2006; Pischke and Von Wachter, 2008). These studies focus on the effects of expanding resources (such as school constructions and compulsory education extensions), while we investigate the effects of changing resource allocation mechanisms given a fixed amount of resources. This zero-sum nature of school seat allocations induces an equity concern, sharing much in common with ongoing policy discussions on affirmative actions

[7]Other studies measure the effects of selective schools conditional on a particular admission system (Dale and Krueger, 2002; Altonji et al., 2012; Dobbie and Fryer, 2013; Hastings et al., 2013; Pop-Eleches and Urquiola, 2013; Deming et al., 2014; Lucas and Mbiti, 2014; Kirkeboen et al., 2016; Beuermann et al., 2018; Abdulkadiroglu et al., 2019; Zimmerman, 2019; Jia and Li, 2019). In addition to these empirical studies, several papers theoretically compare admission mechanisms with different degrees of centralization and choice based on their effects on application behavior and welfare (Haeringer and Klijn, 2009; Pathak and Sönmez, 2013; Che and Koh, 2016; Chen and Kesten, 2017; Hafalir et al., 2018; Shorrer, 2019).

and meritocratic college admissions (Arcidiacono and Lovenheim, 2016).[8] Finally, this paper belongs to the broad literature on the effects of expanding school choice and competition (Hoxby, 2007).

From a broader historical perspective, this paper relates to the literature that uses historical data to understand the emergence and evolution of resource allocation mechanisms (Greif, 1993; Kranton and Swamy, 2008; Börner and Hatfield, 2017; Donna and Espín-Sánchez, 2018) and to a limited literature that investigates the long-term effects of such mechanisms (Dell, 2010; Bleakley and Ferrie, 2014, 2016). Our analysis is also related to Bai and Jia (2016), who examine political consequences of the abolition of a meritocratic elite recruitment system (civil service exam) in early twentieth-century China. While they focus on the short-run effects on revolution participation, we study the long-run consequences of introducing a meritocratic admissions system on career trajectories.

The next section provides historical and institutional backgrounds. Section 3 develops theoretical predictions, which we test in Section 4 using data described in Section 4.1. Section 4 examines the short-term impacts of meritocratic centralization, while Section 5 analyzes their long-term impacts. Finally, Section 6 summarizes our findings, discusses their limitations, and outlines future directions.

# 2 Background

## 2.1 College Admissions around the World

One major trend in modern college admissions systems is a growing degree of centralization with transparent admission criteria. Today, over 30 countries use regionally- or nationally-integrated, single-application, and single-offer college admissions. Figure 1 depicts countries that adopt some centralized college admissions in dark red and countries without any centralized college admissions in light yellow, showing that centralized college admissions are used in all continents except North America. These systems have well-specified admission criteria, mixing meritocratic achievement elements (such as GPA and entrance exams), affirmative action and other priority considerations.

Before the turn of the 20th century, however, no country used such a centralized system (see Online Appendix Table 1).[9] Even today, many countries, including the U.S. and Canada, continue to use decentralized systems. Such decentralized schemes tend to come with less transparent criteria for ranking applicants, as illustrated by recent court cases

[8] See also Kamada and Kojima (2015) and Agarwal (2017) for discussions about regional inequality in other matching markets.

[9] The link is `https://www.scribd.com/document/437545135/Online-Appendix191018`

against American universities. Similar observations also apply to K-12 school admissions. How does the centralization of college and school admissions affect students' application behavior, enrollment outcomes, and future careers? Understanding the costs and benefits of meritocratic centralization is the goal of our paper.

## 2.2 Bidirectional Admissions Reforms in History

To evaluate the impacts of different school admissions systems, we take advantage of unique historical episodes in early twentieth-century Japan. After 250 years of the seclusion policy that ended with the arrival of US ships in 1853, to catch up with Western knowledge, science, and technologies, education reforms became a central part of modernization efforts by the Japanese government. In 1894, the government set up a new system of national higher education consisting of one Imperial University and five national Higher Schools. By 1908, the system was expanded to four Imperial Universities (Tokyo, Kyoto, Tohoku, and Kyushu) and eight National Higher Schools (First in Tokyo, Second in Sendai, Third in Kyoto, Fourth in Kanazawa, Fifth in Kumamoto, Sixth in Okayama, Seventh in Kagoshima, and Eighth in Nagoya, named after the order of establishment) in key locations across Japan, as shown in Appendix Figure A.1. Hereafter we refer to these eight National Higher Schools as Schools 1–8 for short.[10]

Schools 1–8 served as an exclusive entry point to Imperial Universities (the most prestigious tertiary education). Virtually all graduates of Schools 1–8 were admitted to these universities without further selection well into the 1920s. Imperial University graduates were also partially or wholly exempted from the Higher Civil Service Examinations and other selective national qualification exams to become high-ranking administrators, diplomats, judges, and physicians (Amano, 2007). As a result, entering Schools 1–8 was considered equivalent to a passport into the elite class. In fact, Schools 1–8 are known to have produced highly distinguished and influential individuals, including several Prime Ministers, Nobel Laureates, world-leading mathematicians, renowned novelists, and the founders of global companies like Toyota. To apply to these schools, one must be male aged 17 or older and have completed

[10] Schools 1-5 were established in 1894 and Schools 6, 7, and 8 were established in 1900, 1901, and 1908, respectively. Despite the growing demand for national higher education, due to fiscal constraints, the number of National Higher Schools remained constant until 1918. From 1918 to 1925, the number of National Higher Schools gradually increased from 8 to 25, but Schools 1–8 remained the most distinguished among all 25 schools. In addition to Schools 1–8, there was a quasi-national school, Yamaguchi Higher School, which was established in 1894, discontinued in 1904, and re-established in 1918. The number of higher education institutions increased after 1918, as the government permitted not only national but also local public and private higher schools and universities. In our empirical analyses, we control for the number of national higher schools as well as other characteristics of higher education institutions.

a five-year middle school.[11] As Schools 1–8 admitted fewer than 2,300 students each year throughout 1900–1930, they constituted less than 0.5% of the cohort of males aged 17.

The admission to Schools 1–8 was merit-based and determined by annual entrance exams. Initially, the government took a laissez-faire approach and let each school administer its own exam and admissions. Schools 1–8 typically held their exams on the same day so that each applicant could only apply to one school. Following the convention in the literature (Che and Koh, 2016; Hafalir et al., 2018), we call this system "decentralized admissions," "decentralized applications," or Dapp for short. The single choice aspect of Dapp captures an essential feature of decentralization, which incentivizes each applicant to self-select into an appropriate school by comparing the selectivity of schools with his own standing.

Among the eight schools, School 1 in Tokyo was considered by far the most prestigious due to its location in the capital and geographical proximity to Tokyo Imperial University (today's University of Tokyo). The next most prestigious was School 3 in Kyoto. By contrast, located in a remote southwest region, Schools 5 and 7 were considered the least prestigious among all schools. Consequently, the schools differed substantially in their popularity and selectiveness (Takeuchi, 2011, Chapter 2), as empirically confirmed in Appendix Section A.1. For example, the acceptance rate (i.e., the share of admitted applicants in all applicants) of School 1 (Tokyo) was always much lower than that of School 5. In fact, a large number of high-achieving students applying to School 1 (Tokyo) were rejected and had to give up advancing to an Imperial University or retake the exam in the subsequent year, while less popular schools were admitting lower-achieving students. For the government whose goal was to select the best and brightest and send them to Imperial Universities, the decentralized system seemed inefficient. According to the Education Minister, failing to admit a high ability student was "*a loss for the country*" (Yoshino, 2001b, p.24).

To solve this problem, in 1901, the schools agreed to unify their entrance exams to a single one, while maintaining decentralized admission decisions. Then, in 1902, the government launched a centralized admission system in which applicants were allowed to apply for multiple schools, rank them in order of their preferences, and take a unified exam at any school. Based on their exam scores and preferences, applicants were then assigned to a school (or no school if unsuccessful) by a well-specified computational algorithm announced ex ante. We call this system "centralized admissions," "centralized applications," or Capp for short. In proposing the centralized system, the Higher Education Committee stated that its purpose was to enroll students with "*superior academic ability*" in each school, placing a clear emphasis on meritocracy (Yoshino, 2001a, p.53).

[11] The eligibility was changed in 1919 to males aged 16 or older who have completed the fourth year of middle school.

The centralized system operated as follows. Each year, the Ministry of Education announced application procedures in April, three months before the exam, as a public notice in Government Gazette. With some simplification for expositional purpose, the assignment algorithm reads as follows (see Appendix Figure A.2 for a reprint of the original public notice in Japanese).[12]

(1) In the order of exam scores, select the same number of applicants as the sum of all schoolsâ capacities. If the score is tied, decide by lottery.

(2) For applicants selected in (1), in the order of exam scores, assign each applicant to the school of his first choice. If the score is tied, decide by lottery.

(3) For applicants for whom the school of his first choice is already filled as a result of the assignments in (2), in the order of exam scores, assign each applicant to the school of his second choice. If the score is tied, decide by lottery.

(4) For applicants for whom the school of his second choice is already filled as a result of the assignments in (3), then assign each applicant to the school of his third choice or below, repeating the same procedure as (3).

(5) If all the schools that an applicant has chosen are filled, then such applicants are not admitted to any schools.

(6) If there is any vacancy as a result the above assignments or due to an accident, then fill the vacancy by applying the above method to unadmitted applicants.

Written more than a century ago in natural language, the rule description was mathematically precise. Observe that the above method imposes meritocracy up front in which only top-scoring applicants were considered for admission regardless of their preferences (Step (1)). Equivalently, this step selects only applicants who would be admitted by any school under the Serial Dictatorship (Deferred Acceptance) algorithm, one of the most widely used algorithms in today's college and selective K-12 admissions. These applicants are then assigned to one of Schools 1–8 using the Immediate Acceptance (Boston) algorithm (Steps (2) to (4)). This algorithm is therefore a variant of the Immediate Acceptance algorithm with a meritocracy constraint, making it closer to the Serial Dictatorship (Deferred Acceptance)

[12] To be precise, each school was divided into several departments, such as law and literature, engineering, science, and medicine, and the assignments were made separately at the department level. For simplification, we assume away departments in presenting the basic assignment algorithm.

algorithm. To the best of our knowledge, this is the world's first recorded nation-wide use of an assignment algorithm.[13]

This institutional innovation was short-lived, however. Due to political and administrative reasons detailed below, the government switched back to Dapp (with a unified exam) in 1908. The government then continued to oscillate between decentralization and centralization, reintroducing Capp in 1917, moving back to Dapp (with a unified exam) in 1919, reinstituting Capp (with major modifications of allowing applicants to list at most two schools) in 1926, and finally settling down to Dapp (with separate exams) in 1928. In a space of thirty years, therefore, there were three periods of centralized admissions: first in 1902–1907, next in 1917–1918, and finally in 1926–1927.

According to historical studies, these repeated policy changes were the results of intense bargaining between the Ministry of Education, who pushed for centralization to advance meritocracy, and the Council of School Principals, who preferred decentralization to protect school autonomy and regional interests (Yoshino, 2001a,b; Takeuchi, 2011; Amano, 2017). We exploit this series of bidirectional policy changes to identify the impacts of centralization on the selection of students and their career outcomes.[14]

## 3 Theory

To guide our empirical investigation, we develop a model to predict the impacts of centralization on application behavior and assignment. We first confirm that centralized admissions (Capp) was indeed designed to make the school seat allocation more meritocratic compared to decentralized admissions (Dapp). Our model also has two predictions about application behavior. First, a greater number of applicants apply to the most popular school under Capp than under Dapp. Second, applicants make more inter-regional applications under Capp relative to Dapp, thus breaking the "local monopoly" of each school in its local area.

A school admission problem is $(S, I, q, (t_i)_{i\in I}, \succ)$ where $S = \{s_1, \ldots, s_m\}$ is the set of schools while $I = \{i_1, \ldots, i_n\}$ is the set of students. Motivated by our empirical setting, schools' common priority order over students is based on test scores $(t_i)_{i\in I} \in \mathbb{R}^n_+$ (the higher the better). Without loss of generality, sort students so that $t_{i_j} > t_{i_k}$ if $j < k$. We also assume that all students are acceptable for any school, which, in our institutional setting,

[13]See Appendix A.1 for the actual admission outcomes in 1917 under the above centralized algorithm.

[14]These historical episodes are well known among historians of Japanese education, who provide detailed institutional accounts (e.g., Yoshino, 2001a,b; Takeuchi, 2011; Amano, 2017). The preceding studies, however, are mostly descriptive and qualitative. An important exception is Miyake (1998, 1999), who examines regional variations in access to higher schools and compares the number of higher school students per population across prefectures. Building on these studies, we combine a formal model and quasi-experimental research design to identify the causal effects of admission reforms.

is true conditional on the pool of eligible applicants. A capacity vector is $q = (q_{s_1}, \ldots, q_{s_m})$ where $q_s$ is the number of students school $s$ can accommodate. The profile of student (strict) reported preferences is $\succ = (\succ_{i_1}, \ldots, \succ_{i_n})$ defined over $S \cup \{o\}$ where $o$ is the outside option. Let $P_i$ denote the set of all possible preference relations for student $i$. $P = \times_{i \in I} P_i$ is the set of all preference profiles. Let $\succ$, $\succ'$ and so on denote students' reported preference profiles.

The outcome of a school admission problem is a matching $\nu : I \to S \cup I$ where $\nu(i)$ means the school that admits student $i$ (or no assignment if $\nu(i) = i$) with the following properties.

- $\nu(i) \notin S \implies \nu(i) = i$ for every $i \in I$, and
- $\left|\nu^{-1}(s)\right| \leq q_s$ for every $s \in S$.

A mechanism is a systematic procedure that determines a matching for each reported preference profile. Formally, it is a function $\mu : P \to \mathcal{M}$ where $\mathcal{M}$ denotes the set of all matchings. Let $\mu_s(\succ)$ denote the set of students assigned to $s$ in mechanism $\mu$ for reported preference profile $\succ$. Let $\mu^C$ be the Capp mechanism introduced in Section 2.

We compare mechanisms with a thought experiment where the same set of applicants with the same true preferences and test scores participate in different mechanisms. Applicants may change their preference reports, depending on which mechanism they participate in. The set of schools and their capacities are assumed to stay constant. Index each school seat by $j = 1, \ldots, k \equiv \sum_{i \in S} q_i$. Let $t_{\mu(\succ)}(j)$ be the test score of the student assigned to seat $j$ under mechanism $\mu$ for preference profile $\succ$. $t_{\mu(\succ)}(j) = 0$ if no student is assigned to seat $j$. Let $F_{\mu(\succ)}$ be the cumulative distribution of test scores among assigned students under any mechanism $\mu$ for preference profile $\succ$, defined as

$$F_{\mu(\succ)}(t) = \frac{\left|\{j \in \{1, \ldots, k\} \mid t_{\mu(\succ)}(j) \leq t\}\right|}{k}$$

for all $t \in \mathbb{R}_+$.

As should be the case given the official goal of centralization, Capp is more meritocratic than any other mechanism, especially Dapp, in that Capp induces a first-order-stochastic-dominance improvement of the test score distribution among admittees.

**Proposition 1.** *For any school admission problem and any mechanism $\mu$, we have $F_{\mu^C(\succ)}(t) \leq F_{\mu(\succ')}(t)$ for all $t \in \mathbb{R}_+$ and $\succ, \succ' \in P$.*

This fact implies that the worst test score among assigned students under Capp is weakly better than that under any other mechanism, including Dapp. Proposition 4 in Appendix A further shows that in terms of the test score distribution, Capp is as meritocratic as the possibly most meritocratic mechanism, i.e., the Serial Dictatorship or Deferred Acceptance mechanism.

To derive additional predictions about applicant behavior, we need to impose more structures on the model. We consider a model with two schools $s_1$ and $s_2$ with capacities $q_1$ and $q_2$, respectively, and any number of applicants. Each applicant takes an action under each mechanism. Under Capp, for example, each applicant submits a preference list $\succ_i$. Under Dapp, each applicant applies to a school. The mechanism then uses these actions to obtain a matching. This procedure induces a strategic form game, $\langle I, (A_i)_{i\in I}, \succ^o\rangle$. The set of players is the set of applicants $I$. The action space of each applicant is $A_i$. Under Capp, this is the set of all possible preference relations $P_i$ over schools. Under Dapp, this is the set of schools $S = \{s_1, s_2\}$. The outcome is evaluated through the true preferences $\succ^o = (\succ^o_{i_1}, \ldots, \succ^o_{i_n})$.

Take any mechanism as given. Let $A_{-i}$ denote the set of possible strategy profiles for all applicants except applicant $i$. Let $i$ denote remaining unassigned. We define a *stochastic dominance* relation, denoted $sd(\succ^o_i)$, on the set of actions $A_i$ as follows: Upon enumerating $S \cup \{i\}$ from best to worst according to $\succ^o_i$, we define

$$a_i \, sd(\succ^o_i) \, a'_i \iff \sum_{l=1}^{t} p_{il}(a_i, a_{-i}) \geq \sum_{l=1}^{t} p_{il}(a'_i, a_{-i}) \text{ for all } t \text{ and } a_{-i} \in A_{-i}$$

where $p_{il}(a_i, a_{-i})$ is the probability that applicant $i$ gets assigned to the $l$-th best option in $S \cup \{i\}$ according to $\succ^o_i$ if he plays action $a_i$, given action profile $a_{-i}$ of other applicants. We say that strategy $a_i$ is a *dominant strategy* if we have $a_i \, sd(\succ^o_i) \, a'_i$ for all $a'_i \in A_i$. This notation allows us to obtain the following result.

**Proposition 2.** *Suppose that every applicant prefers $s_1$ over $s_2$ or every applicant prefers his local school over the other. Also assume that every applicant submits the true preference whenever it is a dominant strategy. Then the number of applicants who apply to the most popular school $s_1$ is weakly larger under centralized admissions than under decentralized admissions.*

Intuitively, Capp would cause applicants to give a shot at the most prestigious and selective school since Capp gives applicants a chance of acceptance by lower-choice schools after rejected by the first-choice school.

To obtain the final theoretical prediction, assume that each applicant lives in a school's local area. Let $n_j$ be the number of students from school $s_j$'s area. Assume the cardinal utility of applicant $i$ from school $s$ to be $U_{is} = U_s + V * 1\{i \text{ is from } s\text{âs area}\}$. Applicants cannot observe their test scores when submitting their preferences, which is the case in our empirical setting. Assume that each applicant believes that every applicant's test score is independent and identically distributed, i.e., $t \sim_{iid} F(t)$ for some distribution $F$. Define $p(n,q)$ as the probability of being one of the top $q$ applicants among $n$ applicants as per i.i.d test scores, i.e., $p(n,q) = \min\{\frac{q}{n}, 1\} * 1\{n > 0\}$.

As above, an admission mechanism induces a strategic form game $\langle I, (A_i)_{i \in I}, (U_i)_{i \in I} \rangle$. The set of players and the action space remain the same. The outcome is now evaluated accordingly to cardinal utility. Define $U_i(.)$ as the expected payoff of player $i$ at the application stage, i.e., $U_i(a_i, a_{-i}) = p(\bar{n}_{a_i}, q_{a_i}) * U_{ia_i}$ if he plays action $a_i$, given action profile $a_{-i}$ of other applicants, where $\bar{n}_a = \sum_{j \in I} 1\{a_j = a\}$. A strategy vector $a = (a_1, \ldots, a_n)$ is an *equilibrium* if for each applicant $i \in I$ and each strategy $a_i' \in A_i$, we have $U_i(a) \geq U_i(a_i', a_{-i})$. An equilibrium $(a_1, \ldots, a_n)$ is called a *symmetric equilibrium* if $a_i = a_j$ for all $i$ and $j$ from the same area. We make the following assumptions for the rest of this section:

**A1.** Applicants play a symmetric equilibrium, which is assumed to exist.

For a given mechanism and an equilibrium play, $w_j$ denotes the number of applicants assigned to school $s_j$ while $w_{jk}$ denotes the number of applicants assigned to school $s_j$ who come from school $s_k$'s area. Define the *proportion of assigned applicants assigned to their local school* as

$$\frac{w_{11} + w_{22}}{w_1 + w_2}.$$

**Proposition 3.** *Under assumptions A1, for sufficiently large $V$ or sufficiently large $|U_1 - U_2|$, the proportion of assigned applicants assigned to their local school is higher under Dapp than under Capp.*

Capp therefore reduces the number of local entrants born in the school's prefecture. Our empirical investigation starts with testing whether these theoretical predictions hold in the data.

# 4 Short-run Impacts

## 4.1 Data

To analyze short-run effects of centralization, we collect data on applications, enrollments, and other outcomes by digitalizing several administrative and non-administrative sources. First, we collect data on the number of applicants and their first choice schools for 1898-1930 from multiple sources: Government Gazettes for 1902; letters exchanged between the Ministry of Education and the Tokyo Imperial University for 1903 and 1904; *Middle School World* (*Chugaku Sekai*, Yoshino (2001a)) for 1907; the Investigation Records of Higher School

Entrance Examinations by the Ministry of Education for 1917, 1918 and 1927; and the Ministry of Education Yearbook for the other years (except 1905, 1906, and 1926 for which there are no data). For 1916 and 1917, we collect more detailed data on the number of applicants by their first-choice school, birth prefecture, and the prefecture of their middle school, based on the Investigation Records of Higher School Entrance Examinations. Birth prefecture is defined by the prefecture of legal domicile registered in the official family registry system. We include applicants born in all 47 prefectures (excluding colonies) and exclude foreign-born applicants.

Next, we newly collect data on the number of first-year students as a proxy for the number of entrants by school, year, and birth prefecture from 1898 to 1930, using the Higher School Student Registers published annually by Schools 1–8. We exclude foreign-born students and students born in colonies.

Third, we collect data on the number of middle school graduates by year, school type (public or private), and prefecture (defined by the location of middle school) from 1897 to 1930, using the Ministry of Education Yearbook. We use these data to control for the supply of potential applicants as well as the general education level. We also control for the numbers of national, public, and private higher schools by prefecture that were established in addition to Schools 1–8 starting in 1919, using the same source.

Finally, we compute a measure of the geographical mobility of applicants and entrants. Since the finest geographical unit of observation is prefecture, we define the distance between an applicant's birth prefecture and the school of his first choice by the direct (straight-line) distance between the capital of the birth prefecture and the capital of the prefecture in which the school was located. Similarly, the distance between an entrant's middle school and one of Schools 1–8 he was admitted to is defined by the direct distance between the two prefectural capitals determined by the prefectural locations of the middle school and the Schools 1–8. The distance data are from the Geospatial Information Authority of Japan. Descriptive statistics of main variables are summarized in Appendix Table A.2.

## 4.2 Strategic Responses by Applicants

As an immediate effect, switching back and forth between Capp and Dapp caused stark strategic responses in application behavior. Figure 2 shows that the three periods of Capp are associated with a sharp increase in the share of applicants who select the most selective School 1 as their first choice, as predicted by Proposition 2. To show this, the top panel of Table 1 reports the difference in the propensity of applicants to rank School 1 as their first choice between the two years, 1916 (under Dapp) and 1917 (under Capp), using the following regression:

$$Y_{it} = \alpha + \beta \times Centralized_t + \epsilon_{it},$$

where $Y_{it}$ is the indicator variable that takes 1 if applicant $i$ in year $t$ selects School 1 as his first choice. $Centralized_t$ is the indicator variable that takes 1 if year $t$ is under Capp. The first column of Table 1 shows that, at the national level, the share of applicants who rank School 1 first increased by 16 percentage points under Capp. This is about 64% increase compared to the mean of 25% under Dapp (reported as the estimate of the constant term $\alpha$). Next, to observe regional variations, we group applicants into school regions based on which of Schools 1–8 was closest to the applicant's middle school (see the map below Table 1) and run the same regression for each region. In all school regions, the share of applicants selecting School 1 rose substantially (by 11 to 19 percentage points) under Capp.

In total, Table 1 shows that the meritocratic centralization induced applicants around the nation to rank the most prestigious school in Tokyo first and to make more long-distance applications. As a result, the competition to enter School 1 became even more intense under the centralized system. Appendix Figure A.3 depicts changes in the competitiveness of Schools 1–8, measured by the ratio of the number of applicants who select the school as their first choice (hereafter first-choice applicants) to the number of entrants to the school. During the periods of centralized admissions, the ratio spiked at School 1 (Tokyo), increased modestly at School 3 (Kyoto), and declined sharply at the rest of the schools. For instance, at the second introduction of Capp in 1917, School 1 attracted 12 times more first-choice applicants (4,428 in total) than its capacity (361 seats). This implies that only a small fraction of the first-choice applicants were admitted to School 1, leaving hundreds of high-scoring applicants rejected by School 1.

## 4.3 Regional Mobility in Enrollment

Recall that under the meritocratic algorithm, first-choice applicants will be rejected if second-choice applicants with higher scores will apply. The centralized assignment rule allows high-scoring applicants from the urban area to be admitted to lower-choice schools, even after being rejected by their first choice. As a result, the centralized system is associated with a sharp and discontinuous increase in enrollment distance, especially in the first two periods of Capp.[15] Figure 3 shows this by plotting the average enrollment distance (i.e., the distance between an entrant's birth prefecture and the school he entered).

[15]The centralized mechanism used in the third period of Capp in 1926-27 was qualitatively different from that in the first and second periods. In the third period of Capp, schools were divided into two groups and applicants were allowed to choose and rank at most two schools (one school per group) in 1926-27.

This increase in regional mobility is also visible as a sharp reduction in the number of "local" entrants (defined by entrants who entered schools in their birth prefectures). We show this by estimating the following regression for each school $s$ separately:

$$\begin{aligned} Y_{pt} = \beta_1 \times Centralized_t \times 1\{\text{school } s \text{ is located in prefecture } p\} \\ + \beta_2 \times Centralized_t \times 1\{\text{school } s \text{ is 1-100 km away from prefecture } p\} \\ + \beta_3 \times Centralized_t \times 1\{\text{school } s \text{ is 101-300 km away from prefecture } p\} \\ + X_{pt} + \gamma_t + \gamma_p + \epsilon_{pt}, \end{aligned}$$

where $Y_{pt}$ is the number of entrants born in prefecture $p$ who entered school $s$ in year $t$. $Centralized_t$ is the indicator that is turned on if the system was centralized in year $t$. $1\{\text{school } s \text{ is 1-100 km away from prefecture } p\}$ is the indicator that takes 1 if school $s$ is not located in, but within 100 km from prefecture $p$. $X_{pt}$ controls for observable characteristics of prefecture $p$ and year $t$, including the number of middle school graduates from prefecture $p$ in year $t$ and the number of higher schools other than Schools 1–8 in prefecture $p$ in year $t$. $\gamma_t$ and $\gamma_p$ are year and prefecture fixed effects.

Capp reduces the number of local entrants born in the school's prefecture, as shown in Table 2. The coefficients of $Centralized_t \times 1\{\text{school } s \text{ is located in prefecture } p\}$ are significantly negative for all schools. Column 1 shows that the number of School 1 entrants born in Tokyo Prefecture declined by about 27%. Most affected was School 7 (where the number of local entrants declined by 49%), while least affected was School 8 (with a decline of 17%). Schools 4–7 experienced reductions in the number of entrants born not only from the school's prefecture but also from surrounding prefectures. In other words, centralization weakened the local monopoly power of each school by creating a national market for higher education, consistent with Proposition 3. These results are robust to whether or not to control for prefecture characteristics (results available upon request).

## 4.4 Meritocracy vs Equal Regional Access

As established above, centralization reduced the number of applicants who were admitted to their local schools. Then who gained more school seats under the centralized system? Figure 4a plots the change in the number of entrants to Schools 1–8 from Dapp to Capp by birth prefecture (where blue colors indicate decreases and red colors indicate increases). The figure shows that most of the western and northern prefectures lost school seats, while Tokyo Prefecture (around School 1) and its surrounding area gained school seats under Capp.

Figure 4b depicts the time evolution of the share of entrants to Schools 1–8 who were born in the Tokyo area defined as prefectures located within 100 km from Tokyo (see Appendix

Figure A.4 for its location). The share of Tokyo-area born entrants rose significantly during the years of centralization.

Table 3 compares the effects of Capp on Tokyo-area born entrants and other entrants, by estimating the following equation for each school $s$:

$$\begin{aligned} Y_{pt} = \beta_1 &\times Centralized_t \times 1\{\text{prefecture } p \text{ is Tokyo}\} \\ &+ \beta_2 \times Centralized_t \times 1\{\text{prefecture } p \text{ is 1-100 km away from Tokyo}\} \\ &+ \beta_3 \times Centralized_t \times 1\{\text{prefecture } p \text{ is 101-300 km away from Tokyo}\} \\ &+ \beta_4 \times Centralized_t \times 1\{\text{school } s \text{ is located in prefecture } p\} \\ &+ \beta_5 \times Centralized_t \times 1\{\text{school } s \text{ is 1-100 km away from prefecture } p\} \\ &+ \beta_6 \times Centralized_t \times 1\{\text{school } s \text{ is 101-300 km away from prefecture } p\} \\ &+ X_{pt} + \gamma_t + \gamma_p + \epsilon_{pt}, \end{aligned}$$

where $Y_{pt}$ is the number of entrants born in prefecture $p$ who entered school $s$ in year $t$. Column 1 of Table 3 shows that the number of Tokyo-area born students admitted to any of Schools 1–8 increased by about 10% from the average under Dapp. The school-by-school estimates in columns 2-9 reveal that this effect comes mainly from Tokyo-area born students entering less selective rural schools.[16] In other words, the net effect of Capp is such that the increased inter-regional applications caused high-achieving students in the Tokyo area to crowd out lower-achieving, rural-born students from their local schools.

Why did the Tokyo area gained more school seats under the centralized system? In Table 4, using the same specification as column 1 of Table 3, we replace the Tokyo area indicators by alternative explanatory variables, i.e., population, income, educational infrastructure, and preference for School 1, to explore this question. For ease of interpretation, each of these variables is standardized to have a mean of zero and a standard deviation of one.

We first examine the reasons why the Tokyo-area born applicants tended to outperform rural-born applicants in the centralized exam. First, as Tokyo prefecture's population was the largest among all prefectures, it might naturally have the largest pool of high-ability students. As shown in column 1, however, when we interact $Centralized_t$ with prefecture population, the coefficient is not statistically significant. Second, Tokyo prefecture's average income was the second highest (next to Osaka prefecture) during our data period. In general, we expect the richer households to have the greater capacity and willingness to invest in their children's education. Column 2 of Table 4 shows that one standard deviation increase in GDP per capita is associated with additional 9.3 entrants to Schools 1–8 under the centralized system,

[16] The results remain almost the same whether we control for observable prefecture characteristics or not (a table available upon request).

and this estimate is marginally significant. Third, in terms of educational infrastructure, Tokyo prefecture hosted the largest number of middle schools among all prefectures, including almost all private middle schools.[17] As a result, it had the greatest number of middle-school graduates who were eligible for higher education. Since students could (and did) enter middle schools outside their birth prefecture, in Column 3, we interact $Centralized_t$ with both (a) the number of middle-school graduates in the birth prefecture and (b) the number of middle-school graduates in the surrounding prefectures. The result is highly significant, indicating that one standard deviation increase in the number of middle-school graduates in the birth prefecture is associated with 4.4 more entrants to Schools 1–8.

Furthermore, to understand why the Tokyo area gained more school seats during the centralized system compared to the decentralized system, it is important to consider not only students' ability but also their preferences. In general, students had a preference for a local school for geographical and cultural proximity, and the local school for students born in the Tokyo area happened to be the most competitive School 1. As a result, it is likely that, holding the ability constant, during the decentralization, students born in the Tokyo area had a higher tendency to apply to School 1 and therefore fail to enter any higher school (even though they might have been able to enter Schools 2–8 had they applied). Column 4 of Table 4 is consistent with this hypothesis, indicating that one standard deviation increase in the share of applicants to School 1 during the decentralization is associated with 5.4 more entrants to Schools 1–8 during the centralization. In summary, our analysis suggests that the reasons why the Tokyo area gained more school seats under the centralized system was a combination of better educational infrastructure, higher income, and strong preference for School 1.

## 4.5 Political Economy of School Admission Reforms

The short-term impacts of centralization highlight a meritocracy-equity tradeoff. On the one hand, the centralized admissions made the school seat allocation more meritocratic, enabling high-ability students to enter one of the elite schools even if they failed at the most selective one. On the other hand, this meritocracy came at the expense of equal regional access to higher education, as high-achieving urban applicants dominated rural applicants.

This meritocracy-equity tradeoff was one of the main reasons why the government went back and forth between the centralized and decentralized systems. In this section, we briefly discuss why centralization was implemented three times (in 1902–07, 1917–18, and 1926–27) and why it was short-lived each time. Historical evidence indicates that the repeated policy

[17]For example, out of 281 middle schools (including national, public, and private) in Japan in 1906, 32 schools were located in Tokyo Prefecture.

changes were the results of intense bargaining between the Ministry of Education (MOE) and the Council of School Principals (CSP). The former pushed for centralization to advance meritocracy while the latter preferred decentralization to protect school autonomy and local interests of rural schools and communities (Yoshino, 2001a,b; Takeuchi, 2011; Amano, 2007, 2013).

In the course of centralizing the school admissions, the MOE repeatedly emphasized the importance of enrolling only the best and brightest to the national higher education system. The problem of the decentralized system was that the ability of admitted students varied widely across schools depending on their selectiveness.[18] The Minister of Education criticized the decentralized system as follows (*Education Times* No.1146, p.21, published in February 15, 1917):

> "*[Under the decentralized system] among applicants rejected by School 1 and School 3, which attract a large number of high ability applicants, there are many applicants whose academic performance is superior to that of applicants admitted to other rural schools. (...) Namely, hundreds of applicants with sufficiently high academic ability to enter rural schools are idly wasting another year [to retake the exam]. This is not only a pity for them, but also a loss for the country.*"

To maximize the quality of entrants, the MOE proposed to centralize admissions. They envisioned a system where all applicants take a single unified exam on the same date, all exam sheets are sent to a central exam committee and graded by a single person per question to ensure fairness, and applicants would be admitted in the order of their exam scores.[19]

The Council of School Principals was opposed to the idea of centralization, however. First of all, the principals deemed it as an intrusion on their power and autonomy.[20] Second, the CSP argued that the centralized system was disadvantageous to rural schools in both the quality of entrants and the quality of match between schools and entrants. According to the CSP, under the centralized system, urban schools were able to enroll all the best students, because applicants in all areas tended to rank urban schools as their first choice. As a result, rural schools lost the most talented students in their local areas who would have entered rural schools under decentralized admissions.[21] Moreover, after reviewing the

[18] *Education Times* (*Kyouiku Jiron*) No. 609, p.40, March 15, 1902; No.610, p.29, March 25, 1902; No.1141, pp.17-18, December 25, 1916; No.1146, pp.21, February 15, 1917; No.1151, pp.12-13, April 5, 1917.

[19] *Education Times* No. 609, p.40, March 15, 1902; No.610, p.29, March 25, 1902; No.1141, pp.17-18, December 25, 1916.

[20] *Education Times* No. 1143, p.21, January 15, 1917; Yoshino (2001b), p.30.

[21] "A Proposal Regarding a Revision of the Higher School Entrance Examination Rules" by the CSP submitted to the MOE in 1906, reprinted in *Compendium of Higher Schools, Volume 3: Education (Kyusei Koutou Gakkou Zensho: Kyouiku-hen)*, pp.605-607.

admission results of 1917 (the first year of the second centralization period), the CSP found that, in rural prefectures, the number of middle-school graduates admitted not only to their local higher school, but also to any higher schools, declined considerably compared to the previous three years of decentralized admissions.[22]

The CSP further complained that under the centralized system, rural schools must admit a sizable number of reluctant and unmotivated students who came to the school as a fallback option.[23]

> "*Students who entered the school of their second choice or below can never dispel a thought that they had to enter that school because of their exam results. As a result, they are unmotivated to study and have no loyalty to their school. Especially, in those schools that enroll many students who chose the school as their fourth or fifth choice, these students often have adverse effects on the general quality of education.*"

This was upsetting to rural schools as well as rural communities, as they typically donated land and other resources when inviting a higher school to their prefectures (Takeuchi, 2011, p.56 and pp.106-107).[24]

Reflecting on these issues underlying the unusual series of centralization and its abolitions, a noted historian writes as follows (Takeuchi, 2011, p.121):

> "*Urban applicants 'overwhelm' rural applicants by applying for rural schools as fallback options. Urban applicants rob rural applicants of opportunities that were once open to them. This ruins the meaning of building national higher schools across the nation.*"

Prompted by this equity-meritocracy tradeoff, the government oscillated between decentralized and centralized systems, finally settling down to the decentralized system in 1928.

[22] Investigative Records of Higher School Entrance Examinations in 1917, p.40, published by the MOE; *Education Times* No. 1190, p.18, May 5, 1918.

[23] "A Proposal Regarding a Revision of the Higher School Entrance Examination Rules" by the CSP submitted to the MOE in 1906, reprinted in *Compendium of Higher Schools, Volume 3: Education (Kyusei Koutou Gakkou Zensho: Kyouiku-hen)*, pp.605-607.

[24] Finally, the administrative cost of implementing the centralized system was always a serious concern. Both MOE officials and the school principals repeatedly pointed out the difficulty of grading thousands of exam sheets by a small number of people in a short period of time and assign these applicants to schools according to the algorithm. Certainly, time and labor costs of implementing the centralized admissions in the absence of modern computers and photocopying technologies was high (*Education Times* No. 609, p.40; No.1146, p.21; No. 1148, pp.16-17; No. 1149, p.17, March 15, 1917; Takeuchi, 2011, pp.118-119).

## 4.6 Other Institutional Changes

We discuss potential threats to our empirical analysis, especially whether changes in other institutional factors could explain our short-run results. Our analysis takes the timing of the reforms as exogenous, which raises a few concerns. The first concern is that if there were simultaneous reforms in middle schools, it could affect application behavior. Second, if there were capacity changes at Schools 1–8 that were correlated with the admission reforms, it could influence application behavior and enrollment outcomes. The third concern is that if the capacity of School 1 increased relative to the capacity of other schools with the admission reforms, this could explain our findings on application behavior.

We investigate these concerns and confirm that time-series changes in the number of middle school graduates, the total number of entrants to Schools 1–8, and the share of entrants to School 1 in all entrants are not correlated with centralization periods (columns 1-3 in Appendix Table A.4). In columns 4 and 5, we also verify that the number of applicants as well as the level of competitiveness (measured by the number of entrants divided by the number of applicants) do not move systematically with introductions of Capp. In addition, if the probability of unsuccessful applicants retaking the exam in subsequent years changes with the admission reforms, this may also affect our results. As shown in column 6, however, we find that the average age of entrants does not change with the introductions of centralization.

A potential concern with the above robustness analysis is that the insignificant results in Appendix Table A.4 may be due to a small sample size (the number of observations is around 30). Yet, using the same empirical specification, we find that centralization is significantly correlated with our main outcome variables (the share of applicants to School 1, the enrollment distance, and the share of entrants who were born in the Tokyo area), as shown in columns 7-9 of Appendix Table A.4. Taken together, these results suggest that our findings are unlikely to be driven by institutional changes other than the school admission reforms.

Finally, the centralization reform introduced not only the meritocratic assignment algorithm, but also the unified entrance exam that applicants could take at any school locations. As such, the estimated impacts of centralization may be confounded by the unification of entrance exams and more flexible exam location choices. To investigate this issue, we analyze how key outcomes change from 1900 to 1901, during which the government also introduced a single entrance exam that applicants were allowed to take anywhere while the assignment method remained unchanged (i.e., decentralized). Figures 2 and 3 show that this institutional change from 1900 to 1901 induced little changes in application and enrollment patterns. The estimated impacts of centralization are therefore likely due to the meritocratic assignment algorithm rather than the changes in exam contents and locations.

# 5 Long-run Impacts

To assess long-run effects of the meritocratic centralization, we provide two sets of empirical analysis. First, using prefecture-cohort level data compiled from Who's Who publications, we employ a difference-in-differences strategy and compare career outcomes of urban- and rural-born individuals across birth cohorts that differed in their exposure to the centralized admissions. Our analysis shows that the centralization greatly increased the number of career elites born in urban areas relative to those born in rural areas, indicating that the admission reforms had major impacts on the regional composition of elites.

Second, using the data compiled from the list of individuals who passed the Higher Civil Service Examinations, we compare the number of civil officials across cohorts entering higher schools that differed in their exposure to the centralized admissions. We find that the centralized admissions increased the number of officials who were promoted to top ranks, indicating that the centralization likely improved the quality of career elites.

## 5.1 Urban-Rural Disparity in Producing Career Elites

### JPIR Data

To analyze long-run effects of the centralization on students' career outcomes, we first use the Japanese Personnel Inquiry Records (JPIR) published in 1939 as our main data source. The JPIR is an equivalent of Who's Who, which compiles a highly selective list of distinguished individuals such as high-income earners, national medal recipients, top business managers, elite professionals, high-ranking politicians, bureaucrats, and military personnel.[25] In total, the 1939 JPIR lists 55,742 individuals or 0.15% of the adult Japanese population of that time. In selecting these individuals, the JPIR uses a variety of sources, including the directory of banks and companies, the government personnel directory, the directory of Japanese notables, and the directory of industrial associations board members.[26]

To capture the effects of the first period of the centralized admission system in 1902–1907, we use the cohorts born in 1880–1894, who turned 17 years old (the age eligible for application) in 1897–1911. The cohorts born in 1880–1894 were 45 to 59 years old in 1939.[27]

[25]The JPIR also lists the imperial and peerage families, but they are excluded from our data as our analysis focuses on career elites.

[26]The directory of banks and companies includes a list of all directors of banks and companies whose capital is 300,000 yen or above. The government personnel directory provides a complete list of politicians, military personnels, and civil servants in national and local governments, including Imperial University professors. The directory of Japanese notables includes high tax payers defined by individuals who paid more than 50 yen of income tax or more than 80 yen of business tax.

[27]The average life expectancy at age 20 for males born in 1880–1900 was about 40 years.

The number of individuals listed in the JPIR in each of these cohorts is about 1,800. We use the following information from the JPIR data for each individual: full name, birth date, birth prefecture,[28] prefecture of residence, final education,[29] occupational titles and positions, the name of the employer (if applicable), the medal for merit and the court rank awarded (if any), and the amount of national income tax and business tax paid.

We define the following (mutually non-exclusive) groups of elites as subsets of JPIR-listed individuals: (1) top 0.01% and 0.05% income earners,[30] (2) medal recipients (individuals who received either the medal of the Fifth Order of Merit or above, or the court rank of the Junior Fifth Rank or above, excluding military personnels),[31] (3) professionals (individuals whose occupation is either physician, engineer, lawyer, or scholar), (4) professors at Imperial Universities (individuals whose occupation is either professor or associate professor at one of the Imperial Universities), and (5) managers (individuals employed in a private sector with a positive amount of income or business tax payment). Descriptive statistics in Appendix Table A.2 show that these categories are highly selective groups of career elites. For instance, the average number of the top 0.05% income earners per cohort per prefecture is fewer than 5 and the total number for the whole country per cohort is just 230. These categories encompass social, economic, political, and cultural definitions of career elites, or "upper-tail human capital" of the society (Mokyr, 2005).

We use this data to count the number of elites in each group by birth prefecture and birth cohort. These counts allow us to conduct a difference-in-differences analysis that compares long-term career outcomes of urban- and rural-born individuals by each cohorts exposure to the centralized system. Descriptive statistics of main variables are summarized in Appendix Table A.2.

[28] The JPIR obtained information about birth date and prefecture from the official family registry system administered by local governments.

[29] As final education is typically a university or its equivalent, there is no information about higher school.

[30] The threshold income tax payments for the top 0.01% and 0.05% income earners are 9,967 yen and 2,385 yen, respectively. For example, 9,967 yen of income tax payment is equivalent to around 50,000 yen of taxable income, which is well over 50 times the estimated mean household income in 1936 (Yazawa, 2004). The top 0.01% income group earned about 3% of national income in the 1930s, indicating a high degree of income concentration at the top of income distribution comparable to that of the U.S. during the same period (Moriguchi and Saez, 2006, 2008).

[31] In the Japanese honor system, the medals for merit and the court ranks were conferred on individuals in recognition of their exceptional public service or distinguished merit. The medals had 8 grades from the First Order of Merit (the highest honor) to the Eighth Order of Merit (the lowest honor), and the court ranks had 16 ranks from Senior First Rank (the highest) to Junior Eighth Rank (the lowest). The highest orders and ranks were awarded mostly to top-ranking military officers, bureaucrats, and politicians, but a small number of private individuals such as top corporate executives received the Fourth and Fifth Orders of Merit (Ogawa, 2009).

### Assessing the Coverage and Bias of the Data

Since our JPIR data is not exhaustive administrative data, we are concerned about potential sample selection bias. For top income earners and Imperial University professors, we can compute the exact sampling rates by comparing the number of individuals in our data against complete counts reported in government statistics. We find that the sampling rates are decent even by modern standards: 53% and 39% for the top 0.01% and 0.05% income earners, respectively, and 70% for Imperial University professors. Consistent with the nature of our data, which lists only distinguished individuals, the sampling rates increase with the income level (see Appendix Figure A.5).

Sample selection bias becomes a problem for our difference-in-differences analysis only if the difference in sampling rates between urban and rural areas changes with cohorts' exposure to the centralized admission system. To assess this possibility, we examine the prefecture-level sampling rates for top income earners. As Appendix Figure A.6 shows, the number of high income earners in our data and the complete count from tax statistics are highly correlated at the prefecture level, with similar sampling rates across prefectures. This result provides further support for the quality of our data. Even so, one potential concern is that Imperial University graduates might have a higher likelihood of being sampled by our JPIR data even after controlling for the income level. However, we find no positive correlation between the sampling rates of top income earners and the numbers of Imperial University graduates across prefectures (see Appendix Table A.5). This series of findings suggests that possible sample selection bias in the long-term data is unlikely to drive our empirical results.

Finally, we collect and control for a set of time-varying prefecture characteristics. To control for demographic changes, we collect prefecture-level birth populations for the cohorts born in 1886–1894 from the population census and estimate birth populations for the cohorts born in 1880-1885 using age-specific population data available in 1876-1894. To control for local economic conditions, we take prefecture-level manufacturing GDP estimates in 1874, 1890, 1909, and 1925 from Tangjun et al. (2009) and interpolate them linearly for each prefecture. To control for changes in middle schools, we collect the number of middle school graduates in each prefecture in the year when the cohort became age 16.

### Difference-in-Differences Analysis

We estimate the long-run impacts of the centralized admissions (Capp), by conducting a difference-in-differences analysis by birth cohorts and birth areas. The key idea behind our empirical strategy is that applicants born in the Tokyo area should experience a greater gain in entering Schools 1–8 under Capp relative to Dapp, since the centralized system is designed to be more meritocratic and high-achieving students are disproportionately located in urban areas. Figure 4 and Table 3 confirm this expectation. We exploit this differential gain in school access to compare the career outcomes of individuals born inside and outside the Tokyo area by the cohort's exposure to Capp. If admission to Schools 1–8 increases one's chance of becoming a career elite, we should observe a greater number of elites born inside the Tokyo area for the cohorts exposed to Capp. We estimate a difference-in-differences specification as follows:

$$Y_{pt} = \beta \times Centralized_t \times Urban_p + \gamma_p + \gamma_t + \epsilon_{pt},$$

where $Y_{pt}$ is the number of elites born in cohort $t$ and prefecture $p$. $Centralized_t$ is a measure of cohort $t$'s exposure to Capp, which is, in the baseline specification, a binary variable that takes 1 if cohort $t$ turned 17 during Capp (1902–1907). $Urban_p$ is the indicator variable that takes 1 if prefecture $p$ is in the Tokyo area. The prefecture fixed effects $\gamma_p$ capture any systematic difference in career outcomes across prefectures that do not vary across cohorts. The cohort fixed effects $\gamma_t$ control for common shocks that affect career outcomes in all prefectures as well as secular time trends. To allow for serial correlation of $\epsilon_{pt}$ within prefecture over time, we cluster the standard errors at the prefecture level in our baseline specification.[32] In addition, we report the results of clustering at cohort level, which are estimated by wild cluster bootstrap (Cameron and Miller, 2015; Roodman et al., 2019) due to the small number of clusters (15 cohorts).

The above regression defines $Centralized_t$ to be a binary indicator, as the simplest proxy for the intensity of exposure to Capp. In reality, however, a nontrivial number of unsuccessful applicants retook the exam at age 18 and beyond.[33] As a result, the cohorts who turned age 17 in 1899–1901 were partially and increasingly exposed to Capp (as they might have taken the exam in 1902), the cohorts who turned age 17 in 1902–1904 were fully exposed to Capp, and the cohorts who turned age 17 in 1905–1907 were partially and decreasingly exposed to Capp (as they might have taken the exam in 1908), and the intensity of exposure drops to zero for the cohorts who turned age 17 in 1908. For this reason, we explicitly incorporate Capp exposure in visual results and also provide a robustness check by dropping the cohorts who were also partially exposed to Dapp below.

[32]Bertrand et al. (2004) evaluate approaches to deal with serial correlation within each cross-sectional unit in panel data. They suggest that clustering the standard errors on each cross-section unit performs well in settings with 50 or more cross-section units, as in our setting.

[33]According to the limited data available on the Government Gazette in 1903, out of all Schools 1–8 entrants in 1903, 63% graduated middle school in the same year, 29% graduated in the previous year, 6% graduated two years before, and 1% graduated three years before.

We first check whether the number of Imperial University graduates born inside the Tokyo area increased for the cohorts exposed to Capp. Since all Schools 1–8 graduates were automatically admitted to an Imperial University during this period, the areas that produced more Schools 1–8 entrants should produce more Imperial University graduates. Figure 5 (a) compares the average number of Imperial University graduates who were born in prefectures inside and outside the Tokyo area by cohorts (represented by their birth year plus 17 on the horizontal axis). In these and subsequent plots, we color cohorts according to their intensity of exposure to Capp as described above. Figure 5 (b) confirms that the urban-rural difference in the number of Imperial University graduates rises as the intensity of exposure to Capp increases. The difference then falls after the end of Capp in 1908. Column 1 in Table 5 shows that the estimate of $\beta$ in the above regression is positive and statistically significant.

Our main results are presented in Figure 5 (c)–(h) and Table 5 columns 2–7. Figure 5 (c)–(h) show difference-in-differences plots that compare the number of the top 0.05% income earners, professionals (physicians, engineers, lawyers, and scholars), and medal recipients (the Fifth Order of Merit or the Junior Fifth Rank and above) who were born inside and outside the Tokyo area by the cohort's exposure to Capp. Across all elite categories, the plots show that the difference between the Tokyo area and the rest grows as the intensity of exposure to Capp increases, and then drops sharply after the end of Capp in 1908.

Table 5 columns 2–7 show that the long-run effects of Capp are economically and statistically significant. Panel A controls only for cohort and prefecture fixed effects. Panel B additionally controls for time- and cohort-varying prefecture characteristics (i.e., cohort birth population, the number of primary schools, the number of middle school graduates, prefecture-level manufacturing GDP). The coefficients fall slightly in magnitude after adding control variables, but remain sizable. For the cohorts exposed to Capp, the number of career elites born inside the Tokyo area (compared to those born outside the Tokyo area) increases by 36% for the top 0.01% income earners, 24% for the top 0.05% income earners, 20% for managers, and 12% for professionals, 40% for Imperial University professors, and 35% for medal recipients (in Panel B).[34]

Panel C shows that the effects are symmetric with respect to the direction of the admission reforms, i.e., the change from Dapp to Capp and the change from Capp to Dapp produce quantitatively similar effects of the opposite sign. These results suggest that almost four decades after its implementation, Capp had lasting effects on the career trajectories of students.

[34] For professionals, as shown in Appendix Table A.6, even if we look at each occupation separately (i.e., scholars, physicians and lawyers, engineers), the results remain similar, but with lower precision due to smaller sample size. The last column of Appendix Table A.6 further shows that a similar result holds for high-ranking government officials.

In Panel D, we replace the centralization dummy by the cohort's intensity of exposure to Capp (the same values we use to color Figure 5).[35] The results remain qualitatively the same as the baseline results.

The above results are robust to alternative specifications. First, the analysis in Panel D assumes that the cohort's intensity to exposure to Capp is exogenously determined and the same across years, which may be a strong assumption. However, even when we drop the cohorts who are heavily exposed to both Capp and Dapp (i.e., cohorts who became age 17 in 1901 and 1907) from the sample, we still find qualitatively the same results with higher statistical significance (see Appendix Table A.7).

Second, we test if the assumption of parallel pre-event trends holds. Appendix Table A.8 verifies that the differences in pre-event trends between the areas of comparison are small and statistically insignificant for all of our outcome variables.

Another potential threat to our identification strategy is that there may be some age-specific trends in the number of elites that covary with the cohort-region variation we use. Specifically, the number of observations in the 1939 long-term data peaks at around the cohort who were 51 years old in 1939 (corresponding to the cohort who turned age 17 in 1905) and gradually falls for younger and older cohorts, suggesting that there are certain ages at which individuals are more likely to be listed in the long-term data. Such age effects may generate different trends in the number of elites born in the Tokyo and other areas, due, for example, to differences in population size across these areas. To address this concern, we use the earlier edition of the JPIR published in 1934, construct the prefecture-cohort level data for the same cohorts used in our main analysis (but observed 5 years earlier), and conduct similar regression analyses. The results in Appendix Table A.9 confirm that our key results remain qualitatively the same even when we use the 1934 JPIR data.

Finally, we conduct placebo tests to examine if the results are driven by other factors such as the sample selection of the long-term data or changes in cohort populations. Table 6 column 5 confirms that the urban-rural difference in the cohort's birth populations do not change significantly with the cohort's exposure to Capp. As an additional placebo test, we also look at unrelated career outcomes. Among the elites listed in the long-term data, we expect that landlords (defined as individuals whose occupational titles includes landlord, but excluding managers and professionals) are least likely to be affected by the introduction of Capp as receiving higher education was not a typical pathway to becoming a landlord. As shown in Table 6 column 4, the estimated effect of Capp on the number of landlords is small

[35] The intensity is defined as 0.01, 0.06, and 0.29 for the cohort who became age 17 in 1899, 1900, and 1901, respectively. The intensity is 1 for the cohort who became age 17 in 1902, 1903, and 1904. The intensity is 0.99, 0.94, and 0.71 for the cohort who turned age 17 in 1905, 1906, and 1907, respectively. For the younger cohorts, the intensity is 0. These values are determined based on the Government Gazette in 1903.

and statistically insignificant.

### Understanding the Mechanism

We now explore potential mechanisms through which centralization affect career outcomes. First, we test if the centralization-induced increase in inter-regional mobility in the short-run boosted the geographical mobility of elites in the long run. Surprisingly, it did not: The urban-rural difference in the fraction of elites whose prefectures of residence differ from their birth prefectures did not significantly increase under Capp, as shown in Table 6 columns 1 and 2. We find similar results when we use the distance between an elite's birth prefecture and his prefecture of residence as an alternative measure of long-run mobility. This result suggests that, even though a greater number of students born in the Tokyo area entered rural schools under Capp, most of them might have returned to the Tokyo area when pursuing their careers.

We also test whether the centralization affected the urban-rural gap in the *quality* (as opposed to quantity) of Schools 1–8 entrants. As a quality measure, we use the ratio of the number of Imperial University graduates listed in the JPIR data to the total number of Schools 1–8 entrants when the cohort became age 17, assuming that the higher quality of entrants would result in a larger fraction of them listed in the JPIR in their adulthood. We hypothesize that, as the quantity of urban-born entrants relative to rural-born entrants increased greatly under Capp, their relative quality might have declined. The estimated coefficients in Table 6 column 3 are negative but small and insignificant, indicating that the greater urban-rural difference in the quantity of Schools 1–8 entrants under Capp was not associated with a significant decline in the quality of urban-born entrants relative to rural-born entrants.

### Geographical Destinations of Career Elites

Having established that Capp affected the geographic *origins* of highly educated elites, we now ask how it affected their geographic *destinations*. While the former is about regional inequality in educational opportunities, the latter is about regional inequality in the supply of highly skilled human capital, which potentially affects both regional and aggregate economic growth and inequality. If the greater number of Tokyo-area born students admitted to rural schools under Capp returned to the Tokyo area eventually for their subsequent careers, we should observe a greater number of elites living in the Tokyo area for the cohorts exposed to Capp. To test this hypothesis, we redefine the outcome variables by changing the prefecture ($p$) from birth prefecture to prefecture of residence and estimate the equation with the same specification.

Table 7 shows large positive effects of Capp on the urban-rural gap in the number of elite residents, although some of the coefficients come with large standard errors and are not statistically significant. For the cohorts exposed to Capp relative to Dapp, the number of elites living in the Tokyo area in their middle age (compared to those living outside the Tokyo area) increases by 22% for Imperial University graduates, 17% for professionals, and 28% for medal recipients (Panel B). These results suggest that the centralized system likely intensified the concentration of career elites in urban areas relative to rural areas in the long run.

## 5.2 The Quality of Career Elites

In the first analysis, using a difference-in-differences framework, we examined the distributional consequence of the centralized admissions. In the second analysis, we turn to its efficiency implication and explore whether the meritocratic centralization improved the quality of career elites in the long run. To do so, we focus on a specific group of elites, i.e., higher civil officials, for which we have the best available data. In the following analysis, we investigate whether cohorts exposed to Capp produced a greater number of top-ranking civil officials compared to cohorts exposed to Dapp.

#### Higher Civil Officials Data

Our main data source is the list of individuals who passed the Higher Civil Service Examinations (HCSE) and their biographical information compiled by Hata (1981). The HCSE were highly selective national qualification exams held annually from 1894 to 1947.[36] We digitized the information of all individuals who passed the administrative division of the HCSE in 1894–1941, including their full name, education, year of university graduation, year of passing the exam, starting position, final position, year of retirement, and other notable positions held. Because education includes not only final but also the second to final education, unlike the JPIR data, we observe both university and higher school (if applicable) in the HCSE data.

In the Japanese bureaucracy system, the higher civil service refers to the top ten ranks of national government offices in the administrative, judicial, and diplomatic divisions. Within the higher civil service, the top three ranks were distinctively called "imperial appointees" in the prewar period. The first rank consisted of minister level positions, and the second and

[36] The 1893 ordinance required all individuals to pass the HCSE for appointment in the administrative division of higher civil service with some exceptions for special appointments (Spaulding, 1967, Chapter 25; Shimizu, 2019, Chapter 5).

third ranks consisted of vice minister level positions such as vice minister, director general, bureau chief, and prefectural governor.[37] In the following analysis, we define "top-ranking officials" as higher civil officials who were internally promoted to reach one of the top three ranks by the end of their career.[38]

To identify each individual's exposure to the centralized admissions (Capp), we must find out in which year each individual had taken the entrance exam and entered a higher school (or failed and entered an alternative school). However, since we only observe the year of university graduation, we estimate "the year of entering a higher school or its equivalent" (which we call "cohort" in the following analysis) as follows. First, for top-ranking officials who are graduates of Schools 1–8, we find the exact year by searching each individual's full name in the list of first-year students in the Student Registers of Schools 1–8.[39] For these officials, the number of years taken from entering a higher school to graduating from an imperial university ranged from 6 to 10 with the average of 6.6 years.[40] Second, for the rest of individuals who are graduates of Schools 1–8 but not top-ranking officials, we assume that the distribution of the number of years taken from entering a higher school to graduating from an imperial university is the same as that of the top-ranking officials (who are Schools 1–8 graduates and) who graduated from an imperial university in the same year.[41] Third, for individuals who are not Schools 1–8 graduates, we simply assume that the year of entering a higher school or its equivalent is the year of university graduation minus 6.[42]

Next, to create cohort level data, we count the number of individuals who passed the administrative division of the HCSE (hereafter "exam passers") by the year of entering a higher school or its equivalent. We also count the number of top-ranking officials defined

[37]The correspondence between civil service positions and their ranks is reported in the government personnel directory published in 1939 and other years.

[38]Precisely speaking, we define "top-ranking officials" as higher civil officials (a) whose final position was in the top three ranks excluding postwar governorship and (b) whose final and notable positions do not include positions in the first rank (i.e., minister-level positions). We exclude postwar governors from top-ranking officials because starting in 1947 governors were no longer internally promoted but selected by direct election. We further exclude higher civil officials who were appointed to any minister-level positions because these positions were filled by political appointments and not by internal promotion. See Spaulding (1967) and Shimizu (2019) for the development of a meritocratic system of internal promotion.

[39]The exact match was found for 699 out of 733 individuals with a matching rate of 95.1%.

[40]See Appendix Figure A.7 for more information. Because higher school and imperial university were both three-year programs, in principle it would take 6 years to complete both programs. However, it was fairly common for students to repeat the same year, especially in the earlier period when there was a system of holding back students if they failed the year-end exams.

[41]To check the validity of this assumption, for two representative years (1914 and 1922), we find the exact year of entering a higher school for all individuals who graduated from Schools 1–8 and compare the average number of years taken from entering a higher school to graduating from an Imperial University between the top-ranking officials and the rest. As Appendix Figure A.7 shows, the difference is not significantly different in both 1914 and 1922.

[42]We provide a robustness check to this assumption in Appendix Table A.10.

above by cohort. Out of 6,255 exam passers in our dataset, 3,490 individuals or 55.8% are Schools 1–8 graduates, 4,767 individuals or 76.2% are Imperial University graduates, and 982 individuals or 15.7% are top-ranking officials.[43] Among 982 top-ranking officials in our dataset, 701 officials or 71.4% are Schools 1–8 graduates, and 891 officials or 90.7% are Imperial University graduates.[44] Descriptive statistics of main variables are summarized in Appendix Table A.2.

**Empirical Analysis of Higher Civil Service Exams Passers**

Before providing a long-run analysis, we first examine the impacts of the centralization on the number of individuals who passed the highly selective HCSE as an intermediate outcome. We expect that, compared to the decentralized system, the centralized system (which effectively selected top-scoring students and assign them to higher schools) would increase the average quality of students who entered Schools 1–8, which in turn would improve the likelihood of Schools 1–8 graduates to pass the HCSE.

To test this hypothesis, we divide exam passers into three mutually exclusive subgroups: (a) those who graduated from School 1, (b) those who graduated from Schools 2–8, and (c) those who are not Schools 1–8 graduates. For each subgroup, we count the number of exam passers by cohort. We estimate the following equation for the entire group and for each subgroup:

$$Y_t = \theta Centralized_t + \xi_1 X_t + \xi_2 Trend_t + \xi_2 Trend_t^2 + \omega_t,$$

where $Y_t$ is the number of exam passers in a given group in cohort $t$ (defined by the year of entering a higher school or its equivalent), and $Centralized_t$ is the indicator that takes 1 if cohort $t$ entered a higher school or its equivalent during Capp. For a subgroup regression, we control for the total number of exam passers in cohort $t$ (denoted by $X_t$). We also control for a quadratic time trend where $Trend_t$ is the number of years since 1897.

The regression results are presented in Table 8 Panel A columns 1–4. Columns 1 and 2 indicate that Capp did not have statistically significant effect on the total number of exam passers or the number of exam passers who graduated from School 1. By contrast, column 3 shows that Capp increased the number of exam passers who graduated from Schools 2–8 by 23% (compared to the mean of 64 for Dapp cohorts), while column 4 indicates that Capp

[43] The share of Imperial University graduates is greater than that of Schools 1–8 graduates because the number of higher schools increased from 8 to 25 in 1919–1930.

[44] The HCSE consisted of preliminary and main exams, and because Imperial University law graduates were the only group exempted from the administrative preliminary exams from 1894 to 1922, they had substantial advantages in passing the HCSE (Hata, 1981, pp.663-666; Spaulding, 1967, Chapter 12). This exemption, however, does not explain why Imperial University graduates had such a high share in the top-ranking officials.

reduced the number of exam passers who are not Schools 1–8 graduates by 13% (compared to the mean of 92 for Dapp cohorts). These changes are statistically significant at the 5% level. In other words, the centralization had little impact on the total number of exam passers, but had a major impact on the composition of exam passers, and its positive effect was concentrated on Schools 2–8 graduates. This finding is consistent with the result of our short-run analysis that the centralization led to a large increase in Tokyo-area born entrants to Schools 2–8 (Table 3 columns 3–9), which likely improved the academic standing of these schools.

One may argue, however, that a greater number of Schools 2–8 graduates exposed to Capp were able to pass the HCSE, not because their ability was higher, but because they obtained better educational qualification or better alumni connections. For example, it is possible that a greater number of Schools 2–8 graduates exposed to Capp entered the most prestigious Tokyo Imperial University. To examine this possibility, we restrict our sample to exam passers who graduated from Schools 2–8 and Tokyo Imperial University. As column 5 of Table 8 Panel A shows, even within this narrowly defined subgroup, the number of exam passers is 35% greater for Capp cohorts compared to Dapp cohorts, and this difference is statistically significant at the 1% level.

**Empirical Analysis of Top-Ranking Officials**

Finally, we analyze the long-run impact of the centralization on the career outcome of higher civil officials. Our outcome variable is the number of officials who were internally promoted to the top three ranks by the end of their career. As before, we divide top-ranking officials into three mutually exclusive subgroups: (a) those who graduated from School 1, (b) those who graduated from Schools 2–8, and (c) those who are not Schools 1–8 graduates, and count the number of top-ranking officials by cohort in each subgroup. We use the same specification as above and run a regression for the entire group and for each subgroup. For all regressions, we control for the total number of exam passers. The results are presented in Table 8 Panel B columns 1–4.

Importantly, column 1 shows that the total number of top-ranking officials increased by 15% for Capp cohorts compared to Dapp cohorts, and this difference is highly statistically significant. One potential threat to our identification is a possibility that the number of available top-ranking positions happened to have increased during the periods of Capp. However, we argue that even if this was the case, it is not likely to affect our results, since our cohort is defined by the year of entering a higher school or its equivalent, and not by the year of becoming top-ranking officials. Namely, as long as individuals in a given cohort were not promoted to a top-ranking position in the same year, a potential correlation between the number of top-ranking positions and the lagged periods of centralized admissions does not bias our results. In Appendix Figure A.8, we show that this condition largely holds.[45]

To explore the mechanisms behind the result of Table 8 Panel B column 1, we move to the results of columns 2–4. Columns 2 and 3 indicate that Capp had a small, negative, and insignificant effect on the number of top-ranking officials who graduated from School 1, but had a large, positive, and significant effect on the number of top-ranking officials who graduated from Schools 2–8 (see Figure 6 for a visual presentation of this result). According to column 3, the number of top-ranking officials who graduated from Schools 2–8 increased by 47% for Capp cohorts compared to Dapp cohorts, and the difference is highly statistically significant. This result is consistent with our analysis of the exam passers in Table 8 Panel A columns 2–3. Unlike Panel A column 4, however, Panel B column 4 shows that Capp had no significant negative effect on the number of top-ranking officials who did not graduate from Schools 1–8.

From these observations, we can conclude that Capp had a positive effect on the total number of top-ranking officials (column 1) because Capp's positive effect on Schools 2–8 graduates was so large (column 3) that it dominated Capp's small and negative effect on those who did not graduate from Schools 1–8 (column 4). It is important to note that this result is inconsistent with the selection hypothesis which claims that the role of the centralized admissions is simply to select a fixed number of high ability students from a pool of applicants and send them to receive national higher education (i.e., higher school and imperial university), but national higher education itself does not give students any added value. Under this selection hypothesis, Capp would produce a greater number of top-ranking officials from Schools 1–8 graduates, but such effect would be precisely offset by a smaller number of top-ranking officials from non-Schools 1–8 graduates so that the total number of top-ranking officials would be constant.

Then what are the mechanisms through which the meritocratic centralization increased the total number of top-ranking officials? There are four main hypotheses: (1) matching, (2) peer effects, (3) connections, and (4) signaling.

The matching hypothesis states that the higher quality of match between students and

[45]We first randomly selected two cohorts exposed to Capp (1903 and 1916) and two cohorts exposed to Dapp (1913 and 1922). Then for all top-ranking officials who graduated from Schools 1–8 in these cohorts, we searched the years in which these officials were appointed to their first top-ranking positions. Using online searches to find biographical information for each official, we obtained necessary information for 82% of these officials. As shown in Appendix Figure A.8, within each cohort, the number of years taken from entering a higher school to the appointment for the first top-ranking position varied widely from 20 to 30 years. This within-cohort variation is a sum of the two variations, the variation in the number of years taken from entering a higher school to passing the exams and the variation in the number of years taken from passing the exams to becoming a top-ranking official. Therefore, it is unlikely that the result of column 1 is driven by a greater number of available top-ranking positions that coincided with the periods of Capp.

schools would result in greater human capital. If higher ability students gained more from national higher education than from private higher education (due to higher quality of teachers and more demanding curriculum, for example), then by assigning top-scoring students to Schools 1–8, the meritocratic centralization would produce a greater number of upper-end human capital. The peer effect hypothesis claims that students benefit more from having higher ability peers. If that is the case, by gathering top-scoring students in Schools 1–8, the meritocratic centralization would produce positive learning externality among these students. Lastly, national higher education may not improve students' ability or human capital per se, but students may benefit from gaining connections with powerful alumni or simply from obtaining better educational qualification (that signals their high ability), which may improve their prospects of getting promoted to top-ranking positions.

Although we cannot distinguish matching effects and peer effects in our data, we can test if the connections or signaling was an important channel. In column 5 of Table 8 Panel B, we reexamine the result of column 3 by controlling for the number of exam passers from the same schools in the same cohort (instead of the total number of exam passers). The coefficient of Capp becomes smaller, but remains highly statistically significant. In column 6, we further restrict our sample to top-ranking officials who graduated from Schools 2–8 and Tokyo Imperial University. Even within this narrowly specified subgroups with common connections and educational qualification, the coefficient of Capp is positive and statistically significant. In summary, our analyses using HCSE data indicate that the meritocratic centralization improved the quality of civil officials and produced a greater number of top-ranking officials.

# 6 Conclusion

The design of school admissions persistently impacts the geography of career elites. We reveal this fact by looking at the world's first recorded use of nationally centralized admissions and its subsequent abolitions in early twentieth-century Japan. While centralization was designed to make the school seat allocations more meritocratic, there turns out to be a tradeoff between meritocracy and equal regional access to higher education and career success. In line with a theoretical prediction, the meritocratic centralization led students to apply to more selective schools and make more inter-regional applications. As high ability students were concentrated in urban areas, however, centralization caused urban applicants to crowd out rural applicants from advancing to higher education.

Most importantly, these impacts persisted in the long run: Several decades later, the meritocratic centralization increased the number of high income earners, medal recipients, and other elite professionals born in urban areas relative to those born in rural areas. Moreover, we show that the cohorts exposed to the centralized system were more likely to become top-ranking government officers than the cohorts exposed to decentralized system.

Though our study uses the admission reforms unique to Japan, the implications of our study might be relevant for other contexts. For instance, distributional consequences of centralized meritocratic admissions may be a reason why many countries continue to use seemingly inefficient decentralized college admissions. Methodologically, the use of natural experiments in history may be also valuable for studying the long-run effects of market designs in other areas, such as housing, labor, and health markets.

It is the multiple bidirectional policy changes in history that allow us to measure the long-run effects. The disadvantage of using historical events, however, is the limited availability of data. The ideal way to alleviate the data concerns would be to use modern administrative data. For example, one may imagine linking administrative tax return data and school district data to measure the long-run effects of school choice reforms in the past few decades. Such an effort would be a fruitful complement to our historical study.

# References

**Abdulkadiroglu, Atila, Joshua D Angrist, Yusuke Narita, and Parag A Pathak**, "Breaking Ties: Regression Discontinuity Design Meets Market Design," 2019.

**Abdulkadiroğlu, Atila and Tayfun Sönmez**, "School Choice: A Mechanism Design Approach," *American Economic Review*, 2003, *93*, 729–747.

__ **, Nikhil Agarwal, and Parag A. Pathak**, "The Welfare Consequences of Coordinated School Assignment: Evidence from the NYC High School Match," *American Economic Review*, 2017.

__ **, Parag A. Pathak, and Alvin E. Roth**, "Strategy-proofness versus Efficiency in Matching with Indifferences: Redesigning the New York City High School Match," *American Economic Review*, 2009, *99(5)*, 1954–1978.

__ **, Parag Pathak, Alvin E Roth, and Tayfun Sonmez**, "Changing the Boston school choice mechanism," Technical Report, National Bureau of Economic Research 2006.

**Agarwal, Nikhil**, "Policy Analysis in Matching Markets," *American Economic Review*, 2017, *107* (5), 246–50.

**Altonji, Joseph G, Erica Blom, and Costas Meghir**, "Heterogeneity in Human Capital Investments: High School Curriculum, College Major, and Careers," *Annual Review of Economics*, 2012, *4* (1), 185–223.

**Amano, Ikuo**, *Social History of Examination (In Japanese)*, Heibon-sha, 2007.

__ , *The Era of Higher Education, Vol.2 (In Japanese)*, Chukou Sousho, 2013.

__ , *Imperial Universities: Elite Factory in Modern Japan (In Japanese)*, Chuou Kouron Shinsha, 2017.

**Arcidiacono, Peter and Michael Lovenheim**, "Affirmative Action and the Quality-fit Trade-off," *Journal of Economic Literature*, 2016, *54* (1), 3–51.

**Autor, David**, "Work of the Past, Work of the Future," in "Richard T. Ely Lecture given at Annual Meeting of American Economics Association, Atlanta, GA, January," Vol. 4 2019.

**Avery, Christopher, Soohyung Lee, and Alvin E Roth**, "College Admissions as Non-Price Competition: The Case of South Korea," 2014. Working Paper.

**Bai, Ying and Ruixue Jia**, "Elite Recruitment and Political Stability: the Impact of the Abolition of China's Civil Service Exam," *Econometrica*, 2016, *84* (2), 677–733.

**Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan**, "How Much Should We Trust Differences-in-differences Estimates?," *Quarterly Journal of Economics*, 2004, *119* (1), 249–275.

**Beuermann, Diether, C Kirabo Jackson et al.**, "Do Parents Know Best?: The Short and Long-Run Effects of Attending The Schools that Parents Prefer," 2018. NBER Working Paper No. 24920.

**Bleakley, Hoyt and Joseph Ferrie**, "Land Openings on the Georgia Frontier and the Coase Theorem in the Short and Long Run," 2014. Working Paper.

— **and** — , "Shocking Behavior: Random Wealth in Antebellum Georgia and Human Capital Across Generations," *Quarterly Journal of Economics*, 2016, *131* (3), 1455–1495.

**Börner, Lars and John William Hatfield**, "The Design of Debt-Clearing Markets: Clearinghouse Mechanisms in Preindustrial Europe," *Journal of Political Economy*, 2017, *125* (6), 1991–2037.

**Calsamiglia, Caterina, Guillaume Haeringer, and Flip Kljin**, "Constrained School Choice: An Experimental Study," *American Economic Review*, 2010, *100(4)*, 1860–74.

**Cameron, A Colin and Douglas L Miller**, "A practitioner's guide to cluster-robust inference," *Journal of Human Resources*, 2015, *50* (2), 317–372.

**Carvalho, José Raimundo, Thierry Magnac, and Qizhou Xiong**, "College Choice, Selection and Allocation Mechanisms: A Structural Empirical Analysis," *Quantitative Economics*, 2019.

**Che, Yeon-Koo and Youngwoo Koh**, "Decentralized College Admissions," *Journal of Political Economy*, 2016, *124* (5), 1295–1338.

**Chen, Yan and Onur Kesten**, "Chinese College Admissions and School Choice Reforms: A Theoretical Analysis," *Journal of Political Economy*, 2017, *125* (1), 99–139.

**Currie, Janet and Enrico Moretti**, "Mother's Education and the Intergenerational Transmission of Human Capital: Evidence from College Openings," *Quarterly Journal of Economics*, 2003, *118* (4), 1495–1532.

**Dale, Stacy Berg and Alan B Krueger**, "Estimating the Payoff to Attending a More Selective College: An Application of Selection on Observables and Unobservables," *Quarterly Journal of Economics*, 2002, *117* (4), 1491–1527.

**Dell, Melissa**, "The Persistent Effects of Peru's Mining Mita," *Econometrica*, 2010, *78* (6), 1863–1903.

— **and Sahar Parsa**, "Managerial Talent and Economic Performance: Evidence from Discontinuities in Douglas MacArthur's Economic Purge," 2019. Working Paper.

**Deming, David, Justine Hastings, Thomas Kane, and Douglas Staiger**, "School Choice, School Quality and Postsecondary Attainment," *American Economic Review*, 2014, *104(3)*, 991–1013.

**Dobbie, Will and Roland G. Fryer**, "The Impact of Attending a School with High-Achieving Peers: Evidence from the New York City Exam Schools," *American Economic Journal: Applied Economics*, 2013.

**Donna, Javier and José-Antonio Espín-Sánchez**, "The Illiquidity of Water Markets: Efficient Institutions for Water Allocation in Southeastern Spain," *Available at SSRN 2667654*, 2018.

**Duflo, Esther**, "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment," *American Economic Review*, 2001, *91* (4), 795–813.

**Glaeser, Edward**, *Triumph of the City*, Pan, 2011.

**Greif, Avner**, "Contract Enforceability and Economic Institutions in Early Trade: The Maghribi Traders' Coalition," *American Economic Review*, 1993, pp. 525–548.

**Grenet, Julien, Yinghua He, and Dorothea Kübler**, "Decentralizing Centralized Matching Markets: Implications from Early Offers in University Admissions," 2019. Working Paper.

**Haeringer, Guillaume and Flip Klijn**, "Constrained School Choice," *Journal of Economic Theory*, 2009, *144*, 1921–1947.

**Hafalir, Isa, Rustamdjan Hakimov, Dorothea Kuebler, and Morimitsu Kurino**, "College Admissions with Entrance Exams: Centralized versus Decentralized," *Journal of Economic Theory*, 2018.

**Hastings, Justine, Christopher Neilson, and Seth Zimmerman**, "Are Some Degrees Worth More than Others? Evidence from College Admission Cutoffs in Chile," 2013. Working Paper.

**Hata, Ikuhiko**, *The System, Organization, and Personnel of Bureaucracy in Prewar Japan (In Japanese)*, Tokyo Daigaku Shuppankai, 1981.

**Hoxby, Caroline M**, *The Economics of School Choice*, University of Chicago Press, 2007.

**Jia, Ruixue and Hongbin Li**, "Designated Elitism," 2019. Working Paper.

**Kamada, Yuichiro and Fuhito Kojima**, "Efficient Matching under Distributional Constraints: Theory and Applications," *American Economic Review*, 2015, *105* (1), 67–99.

**Kirkeboen, Lars, Edwin Leuven, and Magne Mogstad**, "Field of Study, Earnings, and Self-Selection," *Quarterly Journal of Economics*, 2016, *131*, 1551–1552.

**Knight, Brian and Nathan Schiff**, "Reducing Frictions in College Admissions: Evidence from the Common Application," 2019.

**Kranton, Rachel and Anand V Swamy**, "Contracts, Hold-up, and Exports: Textiles and Opium in Colonial India," *American Economic Review*, 2008, *98* (3), 967–89.

**Lucas, Adrienne and Isaac Mbiti**, "Effects of School Quality on Student Achievement: Discontinuity Evidence from Kenya," *American Economic Journal: Applied Economics*, 2014, *6(3)*, 234–263.

**Machado, Cecilia and Christiane Szerman**, "Centralized Admissions and the Student-College Match," 2017. Working Paper.

**Meghir, Costas and Marten Palme**, "Educaitonal Reform, Ability, and Family Background," *American Economic Review*, 2005, *95(1)*, 414–424.

**Miyake, Sono**, "Regional Difference in Access to Higher Schools during the Late Meiji Era (in Japanese)," *Socio-Cultural Studies of Education*, 1998, *5*, 35–52.

__ , "Regional Difference in Attaining High Academic Career: From an Analysis of Higher School Enrollment in Taisho Era (in Japanese)," *Bulletin of the Graduate School of Education at Kyoto University*, 1999.

**Mokyr, Joel**, "Long-Term Economic Growth and the History of Technology," in "Handbook of economic growth," Vol. 1, Elsevier, 2005, pp. 1113–1180.

**Moretti, Enrico**, *The New Geography of Jobs*, Houghton Mifflin Harcourt, 2012.

**Moriguchi, Chiaki and Emmanuel Saez**, "The Evolution of Income Concentration in Japan, 1886-2002: Evidence from Income Tax Statistics," 2006. NBER Working Paper 12558.

— **and** — , "The Evolution of Income Concentration in Japan, 1886–2005: Evidence from Income Tax Statistics," *Review of Economics and Statistics*, 2008, *90* (4), 713–734.

**Niederle, Muriel and Alvin E Roth**, "Unraveling Reduces Mobility in a Labor Market: Gastroenterology with and without a Centralized Match," *Journal of Political Economy*, 2003, *111* (6), 1342–1352.

**Ogawa, Kenji**, *Sociology of the Medals (in Japanese)*, Shoyo Shobo; Kyoto, 2009.

**Oreopoulos, Philip**, "Estimating Average and Local Average Treatment Effects of Education When Compulsory Schooling Laws Really Matter," *American Economic Review*, 2006, *96* (1), 152–175.

**Pallais, Amanda**, "Small Differences That Matter: Mistakes in Applying to College," *Journal of Labor Economics*, 2015, *33* (2), 493–520.

**Pathak, Parag A. and Tayfun Sönmez**, "School Admissions Reform in Chicago and England: Comparing Mechanisms by Their Vulnerability to Manipulation," *American Economic Review*, 2013, *103(1)*, 80–106.

**Pischke, Jörn-Steffen and Till Von Wachter**, "Zero Returns to Compulsory Schooling in Germany: Evidence and Interpretation," *Review of Economics and Statistics*, 2008, *90* (3), 592–598.

**Pop-Eleches, Cristian and Miguel Urquiola**, "Going to a Better School: Effects and Behavioral Responses," *American Economic Review*, 2013, *103(4)*, 1289–1324.

**Redding, Stephen J and Daniel M Sturm**, "The Costs of Remoteness: Evidence from German Division and Reunification," *American Economic Review*, 2008, *98* (5), 1766–97.

**Roodman, David, Morten Ørregaard Nielsen, James G MacKinnon, and Matthew D Webb**, "Fast and wild: Bootstrap inference in Stata using boottest," *The Stata Journal*, 2019, *19* (1), 4–60.

**Roth, Alvin E.**, "The Evolution of the Labor Market for Medical Interns and Residents: A Case Study in Game Theory," *Journal of Political Economy*, 1984, *92*, 991–1016.

— , "New Physicians: A Natural Experiment in Market Organization," *Science*, 1990, *250*, 1524–1528.

**Shimizu, Yuichiro**, *The Origins of the Modern Japanese Bureaucracy*, Bloomsbury Academic, 2019.

**Shorrer, Ran**, "Simultaneous Search: Beyond Independent Successes," 2019. Working Paper.

**Spaulding, Robert**, *Imperial Japan's Higher Civil Service Examinations*, Princeton University Press, 1967.

**Takeuchi, Hiroshi**, *The Glory and Baffle of Educational Elites (in Japanese)*, Koudansha, 2011.

**Tangjun, Yuan, Settsu Tokihiko, Jean-Pascal Bassino, and Kyoji Fukao**, "Gross Prefectural Domestic Product and Industrial Structure in Pre-war Japan [in Japanese]," 2009. Working Paper.

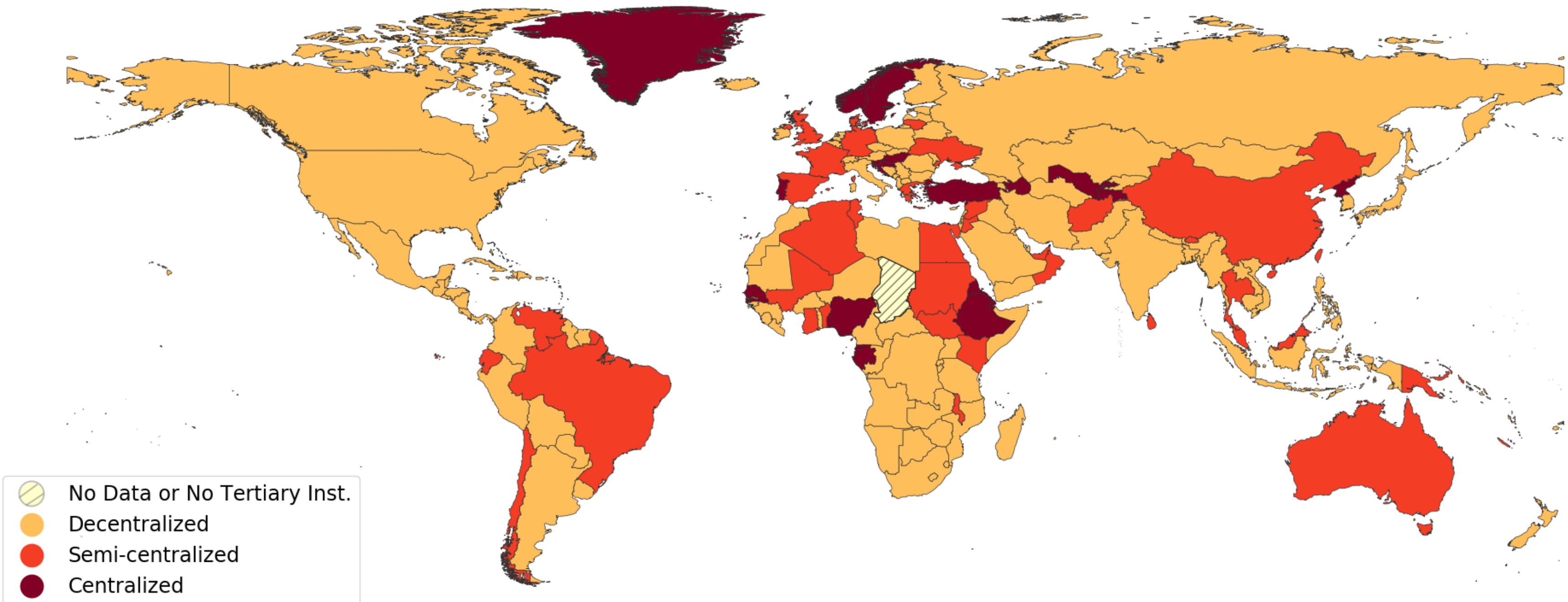
**Yazawa, Hirotake**, *Income Distribution and Family Economics in Modern Japan (in Japanese)*, Tokyo Book Center, 2004.

**Yoshino, Takehiro**, "Entrance Examination of Higher Schools in the Early Taisho Period (in Japanese)," *Keio University Studies in Sociology, Psychology, and Education*, 2001, (53), p.19–31.

__ , "Entrance Examination of Higher Schools in the Meiji Period (in Japanese)," *Keio University Studies in Sociology, Psychology, and Education*, 2001, (52), p.51–62.

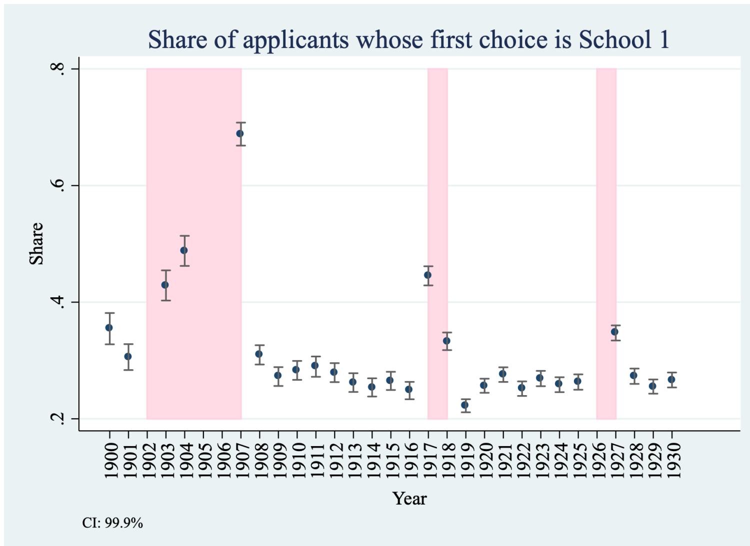
**Zimmerman, Seth D**, "Elite Colleges and Upward Mobility to Top Jobs and Top Incomes," *American Economic Review*, 2019, *109* (1), 1–47.

Figure 1: College Admissions around the World Today




*Notes*: This figure summarizes each country and territory's college admission system today. Dark, red color (e.g. Norway): Regionally- or nationally-centralized college admissions, where a single-application, single-offer assignment algorithm (well-defined rule) is used to make admissions to both public and private universities. Medium, orange color (e.g. Brazil): Semi-centralized, defined as either (1) there is a centralized system, but not all universities (e.g. private universities) are included in the single-application, single-offer system or (2) students submit a single application and receive multiple offers. Light, orange color (e.g. U.S.A.): Decentralized college admissions, where each college defines its own admissions standards and rules. Yellow with diagonal lines (e.g. Chad): Not enough information available or if the country or territory does not have tertiary institutions. See Section 2 for discussions about this figure.

Figure 2: Centralization Caused Applicants to Apply More Aggressively: First Look




*Notes*: This figure shows the time evolution of the share of applicants who selected the most prestigious School 1 (Tokyo) as their first choice. Colored years (1902–07, 1917–18, and 1926–27) indicate the three periods of the centralized school admission system. No data are available for 1902, 1905, 1906, and 1926. Bars show the 99.9 percent confidence intervals. See Section 4.2 for discussions about this figure.

Table 1: Centralization Caused Applicants Across the Country to Apply More Aggressively

| Dependent var | Select School 1 as First Choice | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Centralized | 0.159*** | 0.192*** | 0.151*** | 0.146*** | 0.128 | 0.168*** | 0.180*** | 0.166*** | 0.114*** |
| | (0.0106) | (0.00924) | (0.0329) | (0.0232) | (0.0646) | (0.0245) | (0.0336) | (0.0136) | (0.00786) |
| Constant | 0.248*** | 0.494*** | 0.169*** | 0.0892*** | 0.178** | 0.107*** | 0.184*** | 0.0813** | 0.127* |
| | (0.0717) | (0.0437) | (0.0357) | (0.0162) | (0.0373) | (0.0185) | (0.0218) | (0.00991) | (0.0508) |
| Sample region | All | S1 Region | S2 Region | S3 Region | S4 Region | S5 Region | S6 Region | S7 Region | S8 Region |
| Observations | 20,913 | 6,505 | 2,555 | 3,248 | 1,266 | 2,730 | 2,276 | 615 | 1,718 |
| Dependent var | Application Distance | | | | | | | | |
| Centralized | -2.534 | -92.88*** | 10.95 | 2.080 | -15.74 | 128.0*** | 46.52** | 145.4** | -25.57 |
| | (23.22) | (2.888) | (24.65) | (5.482) | (22.92) | (23.11) | (13.91) | (21.27) | (18.64) |
| Constant | 226.2*** | 231.7*** | 289.7*** | 158.8*** | 166.7* | 252.6*** | 294.1*** | 218.0* | 154.2* |
| | (15.74) | (16.43) | (79.51) | (28.11) | (56.94) | (42.52) | (51.54) | (70.94) | (48.89) |
| Sample region | All | S1 Region | S2 Region | S3 Region | S4 Region | S5 Region | S6 Region | S7 Region | S8 Region |
| Observations | 20,913 | 6,505 | 2,555 | 3,248 | 1,266 | 2,730 | 2,276 | 615 | 1,718 |

*Notes*: In the first panel, we estimate the effects of centralization on the propensity of an applicant to select the most prestigious and selective school (School 1) as his first choice, using the applicant-level data in 1916 (under the decentralized system) and 1917 (under the centralized system). The prefecture-level application data are available only for these two years. We group applicants into "school regions" based on which school (among Schools 1–8) is nearest to the applicant's middle school in 1916, where "nearest" is defined by the distance between the prefectural capitals. The following map shows the locations of the eight school regions. The second panel measures the effects on the application distance between an applicant's first-choice school and middle school. Standard errors are clustered at the prefecture level. ***, **, and * mean significance at the 1%, 5%, and 10% levels, respectively. See Section 4.2 for discussions about this table.

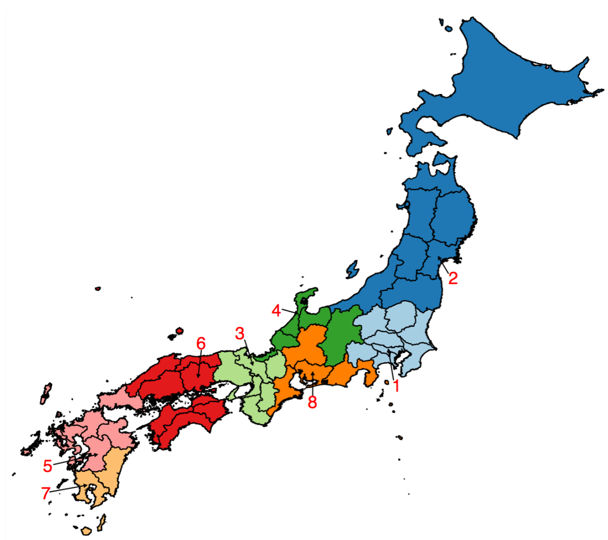
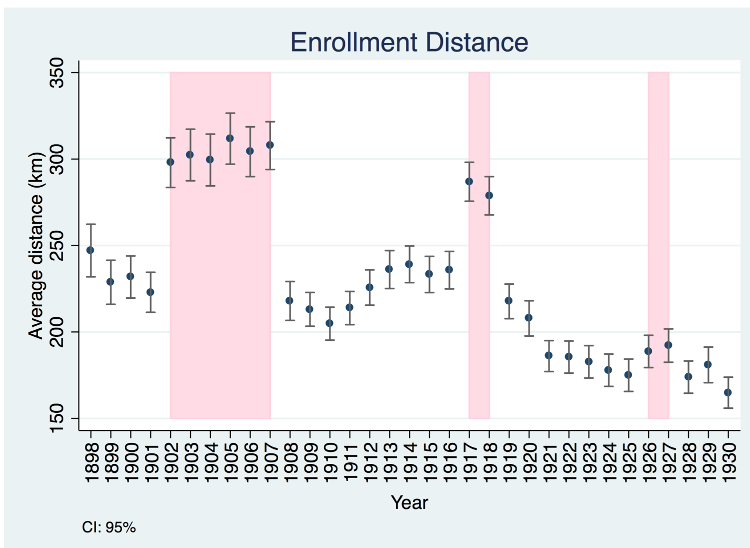


Figure 3: Centralization Increased Regional Mobility in Enrollment: First Look




*Notes*: This figure shows the time evolution of the average enrollment distance, i.e., the distance between an entrant's birth prefecture and the prefecture of the school he entered (measured by the direct distance between the two prefectural capitals). Colored years indicate the three periods of the centralized school admission system. Bars show the 95 percent confidence intervals. See Section 4.3 for discussions about this figure.

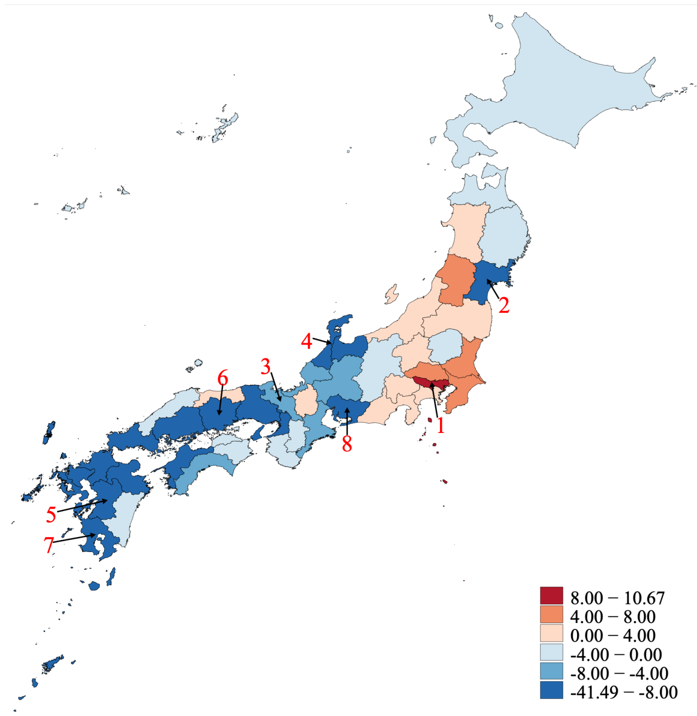
Table 2: Centralization Broke Local Monopoly and Increased Regional Mobility across the Country

| Dependent variable = No. of entrants to: | (1) Sch. 1 | (2) Sch. 2 | (3) Sch. 3 | (4) Sch. 4 | (5) Sch. 5 | (6) Sch. 6 | (7) Sch. 7 | (8) Sch. 8 |
|---|---|---|---|---|---|---|---|---|
| Centralized x Born in school's prefecture | -26.08 | -17.92 | -15.69 | -23.50 | -28.95 | -23.53 | -45.35 | -12.72 |
| | (0.30)*** | (0.15)*** | (0.30)*** | (0.25)*** | (0.21)*** | (0.44)*** | (0.30)*** | (0.65)*** |
| | [8.40]*** | [7.75]** | [8.57]* | [7.91]*** | [7.60]*** | [12.34]* | [13.07]*** | [15.09] |
| Centralized x Born near school's prefecture (1-100 km) | 0.34 | -3.10 | -4.16 | -9.43 | -11.83 | -2.90 | -1.91 | 1.07 |
| | (0.67) | (2.54) | (2.01)** | (3.15)*** | (3.02)*** | (1.30)** | (0.21)*** | (0.86) |
| | [0.53] | [1.42]** | [1.00]*** | [2.65]*** | [3.18]*** | [1.02]*** | [2.55] | [2.70] |
| Centralized x Born near school's prefecture (100-300 km) | 1.19 | -0.08 | -0.31 | -0.19 | -3.23 | -2.17 | -3.17 | 1.27 |
| | (0.54)** | (0.67) | (0.58) | (0.54) | (0.93)*** | (0.81)** | (0.82)*** | (1.02) |
| | [0.66]* | [0.50] | [0.38] | [0.41] | [0.81]*** | [0.64]*** | [1.40]** | [0.64]* |
| Observations | 1,410 | 1,410 | 1,363 | 1,410 | 1,363 | 1,410 | 1,269 | 1,034 |
| Prefecture FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Mean dep var | 7.94 | 5.53 | 6.21 | 5.68 | 6.30 | 5.23 | 5.06 | 5.73 |
| Mean dep var (School's pref dur. Dapp) | 104.3 | 62.15 | 56.05 | 60.30 | 74 | 76.35 | 92.78 | 76.50 |
| Mean dep var (Within 1-100km dur. Dapp) | 9.192 | 21.05 | 17.80 | 27 | 34.73 | 8.367 | 8.389 | 15.63 |

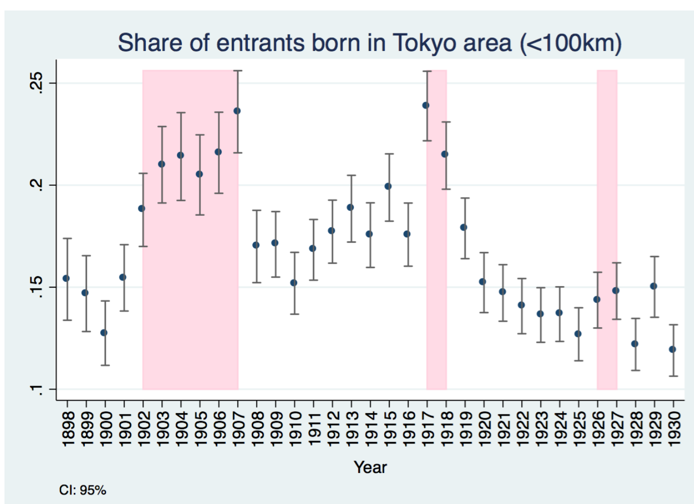
*Notes*: Using the prefecture-year level data in 1900–1930, we define the dependent variable as the number of entrants who were born in the prefecture and entered the school indicated in the column in each year. "Born in school's prefecture" takes 1 if the school indicated in the column is located in the entrant's birth prefecture. "Born near school's prefecture (1-100 km)" takes 1 if the school indicated in the column is not located in, but within 100 km from the entrant's birth prefecture (measured by the distance between the two prefectural capitals). "Born near school's prefecture (100-300 km)" takes 1 if the school indicated in the column is between 100 km and 300 km from the entrant's birth prefecture. We control for year fixed effects, prefecture fixed effects, the number of middle school graduates in the prefecture, and the number of higher schools other than Schools 1–8 in the prefecture. "Mean dep var" shows the mean of the dependent variable during decentralization for all prefecture-year observations. "Mean dep var (school's pref dur. Dapp)" shows the mean number of entrants to the school during decentralization, restricted to those born in the prefecture where the school is located. "Mean dep var (within 1-100 km dur. Dapp)" shows the mean number of entrants to the school during decentralization, restricted to those born in the prefectures within 100 km (excluding the prefecture where the school is located). For Schools 7 and 8, we drop the years in which they held an early exam (School 7 in 1908–1910 and School 8 in 1908). Standard errors reported in parentheses are clustered at the prefecture level. Standard errors reported in square brackets are clustered at the year level. ***, **, and * mean significance at the 1%, 5%, and 10% levels, respectively. See Section 4.3 for discussions about this table.

Figure 4: Which Regions Win from Centralization? First Look

(a) Where Did Entrants Increase during Centralization?




(b) Centralization Increased Tokyo Area-born Entrants to Schools 1–8




*Notes*: Panel (a) estimates and plots the prefecture-specific coefficient $\beta_p$ in $\#entrants_{pt} = \beta_p Centralized_t + \alpha_p X_{pt} + e_{pt}$, using the 1900-1930 data for each prefecture $p$, where $\#entrants_{pt}$ is the number of entrants in year $t$ who were born in prefecture $p$ and $X_{pt}$ is the number of schools other than Schools 1–8 in prefecture $p$ in year $t$. Panel (b) uses the entrant-level data from 1898 to 1930 to show the time evolution of the fraction of entrants to Schools 1–8 who were born in the Tokyo area defined as a set of prefectures that are within 100 km from Tokyo (see Appendix Figure A.4 for a map). Bars show the 95 percent confidence intervals. See Section 4.4 for discussions about this figure.

Table 3: Which Regions Win from Centralization?

| Dependent variable = No. of entrants to: | (1)<br>All schools | (2)<br>Sch. 1 | (3)<br>Sch. 2 | (4)<br>Sch. 3 | (5)<br>Sch. 4 | (6)<br>Sch. 5 | (7)<br>Sch. 6 | (8)<br>Sch. 7 | (9)<br>Sch. 8 |
|---|---|---|---|---|---|---|---|---|---|
| Centralized x Born in Tokyo prefecture | 27.79 | | 1.52 | 3.99 | 6.66 | 3.69 | 11.50 | 5.69 | 19.19 |
| | (4.76)*** | | (0.19)*** | (0.38)*** | (0.38)*** | (0.20)*** | (0.31)*** | (0.27)*** | (0.82)*** |
| | [19.35] | | [3.20] | [3.01] | [2.41]*** | [1.72]** | [1.71]*** | [3.04]* | [3.61]*** |
| Centralized x Born near Tokyo prefecture (1-100 km) | 12.83 | | 0.58 | 1.04 | 2.00 | 0.19 | 0.94 | 0.31 | 0.48 |
| | (2.29)*** | | (0.62) | (0.34)*** | (0.45)*** | (0.35) | (0.41)** | (0.47) | (0.71) |
| | [3.63]*** | | [0.49] | [0.26]*** | [0.63]*** | [0.34] | [0.29]*** | [0.42] | [0.36] |
| Centralized x Born near Tokyo prefecture (100-300 km) | 5.41 | | 1.05 | 1.77 | -0.13 | 0.55 | 1.44 | 0.89 | 0.23 |
| | (2.44)** | | (0.48)** | (0.78)** | (0.90) | (0.34) | (0.59)** | (0.48)* | (0.49) |
| | [2.49]** | | [0.40]** | [0.45]*** | [0.86] | [0.36] | [0.36]*** | [0.28]*** | [0.69] |
| Centralized x Born in school's prefecture | -18.33 | -26.08 | -17.67 | -15.09 | -23.15 | -28.66 | -22.46 | -47.44 | -12.85 |
| | (4.26)*** | (0.30)*** | (0.17)*** | (0.35)*** | (0.92)*** | (0.21)*** | (0.28)*** | (0.34)*** | (0.85)*** |
| | [7.51]** | [8.40]*** | [7.72]** | [8.57]* | [8.43]** | [7.52]*** | [12.23]* | [13.13]*** | [14.61] |
| Centralized x Born near school's prefecture (1-100 km) | -2.88 | 0.34 | -3.89 | -3.86 | -9.15 | -11.54 | -1.84 | -1.72 | 1.05 |
| | (2.34) | (0.67) | (2.58) | (1.91)** | (3.14)*** | (3.03)*** | (1.26) | (0.29)*** | (0.89) |
| | [1.60]* | [0.53] | [1.45]** | [1.01]*** | [2.82]*** | [3.16]*** | [0.89]** | [2.52] | [2.53] |
| Centralized x Born near school's prefecture (100-300 km) | -5.79 | 1.19 | -0.32 | -0.47 | -0.63 | -2.94 | -1.29 | -2.79 | 0.26 |
| | (2.89)* | (0.54)** | (0.74) | (0.53) | (0.43) | (0.93)*** | (0.67)* | (0.86)*** | (0.39) |
| | [3.41] | [0.66]* | [0.52] | [0.36] | [0.49] | [0.74]*** | [0.55]** | [1.34]** | [0.61] |
| Observations | 1,410 | 1,410 | 1,410 | 1,363 | 1,410 | 1,363 | 1,410 | 1,222 | 1,034 |
| Year FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Pref FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Mean dep var | 45.43 | 7.942 | 5.531 | 6.214 | 5.678 | 6.300 | 5.228 | 5.001 | 5.725 |
| Mean dep var (Tokyo pref dur. Dapp) | 200.4 | 104.3 | 27.10 | 10.63 | 14.45 | 5.750 | 9.200 | 12 | 20.33 |
| Mean dep var (Within 1-100km from Tokyo pref. dur. Dapp) | 26.43 | 9.192 | 6.725 | 1.211 | 2.867 | 0.758 | 1.242 | 1.598 | 3.269 |

*Notes*: This table uses the prefecture-year level data in 1900–1930. In column (1), the dependent variable is the number of students from birth prefecture $p$ who entered one of Schools 1–8 in year $t$. In columns (2)–(10), the dependent variable is the number of students from birth prefecture $p$ who entered the school indicated in the column in year $t$. For Schools 7 and 8, we drop the years in which they held an early exam (School 7 in 1908–1910 and School 8 in 1908). We control for prefecture fixed effects, year fixed effects, the number of middle school graduates in the prefecture and the number of higher schools other than Schools 1–8 in the prefecture. "Born in Tokyo prefecture" takes 1 if the entrant's birth prefecture is Tokyo prefecture. "Born near Tokyo prefecture (1-100 km)" takes 1 if the entrant's birth prefecture is not Tokyo prefecture but within 100 km from Tokyo prefecture. "Born near Tokyo prefecture (100-300 km)" takes 1 if the entrant's birth prefecture is between 100 km and 300 km from Tokyo prefecture. The definitions of "Born in school's prefecture," "Born near school's prefecture (1-100 km)," and "Born near school's prefecture (100-300 km) are the same as in Table 2." "Mean dep var" shows the mean of the dependent variable during decentralization for all prefecture-year observations. "Mean dep var (Tokyo pref dur. Dapp)" shows the mean number of entrants to the school during decentralization, restricted to those born in Tokyo prefecture. "Mean dep var (within 1-100 km from Tokyo pref. dur. Dapp)" shows the mean number of entrants to the school during decentralization, restricted to those born in the prefectures within 100 km from Tokyo prefecture (excluding Tokyo prefecture). Standard errors reported in parentheses are clustered at the prefecture level, and standard errors reported in square brackets are clustered at the year level. ***, **, and * mean significance at the 1%, 5%, and 10% levels, respectively. See Section 4.4 for discussions about this table.

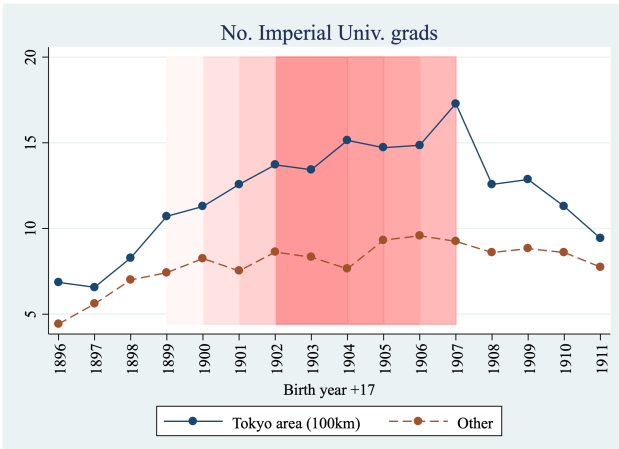
Table 4: Why Does the Urban Area Win?

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | Entrants to Schools 1-8 | | | |
| Centralized × Population in prefecture | 2.63<br>(3.48)<br>[2.07] | | | |
| Centralized × GDP per capita in prefecture | | 9.27<br>(5.61)<br>[4.66]* | | |
| Centralized × middle-school graduates in prefecture | | | 4.42<br>(1.22)***<br>[2.80] | |
| Centralized × middle-school graduates in nearby prefectures (1-100 km) | | | 1.71<br>(0.48)***<br>[0.73]** | |
| Centralized × Share of applicants to School 1 (dur. Dapp) | | | | 5.40<br>(0.78)***<br>[1.67]*** |
| Population in prefecture | 10.09<br>(6.67)<br>[3.17]*** | | | |
| GDP per capita in prefecture | | 12.31<br>(8.20)<br>[3.17]*** | | |
| Middle-school graduates in prefecture | 11.98<br>(3.85)***<br>[2.35]*** | 12.18<br>(4.87)**<br>[1.55]*** | 15.92<br>(3.14)***<br>[2.21]*** | 17.10<br>(3.15)***<br>[2.26]*** |
| Middle-school graduates in nearby prefectures (1-100 km) | | | -0.40<br>(0.72)<br>[0.26] | |
| Observations | 1,410 | 1,410 | 1,410 | 1,410 |
| Year FE | Yes | Yes | Yes | Yes |
| Pref FE | Yes | Yes | Yes | Yes |
| Mean dep var | 45.43 | 45.43 | 45.43 | 45.43 |

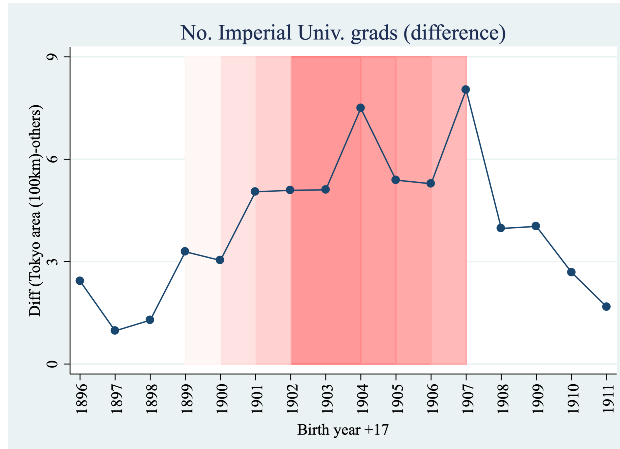
*Notes*: This table uses the prefecture-year level data in 1900–1930. The dependent variable is the number of students from birth prefecture $p$ who entered one of Schools 1–8 in year $t$. "Population in prefecture" is the number of population in prefecture $p$ in year $t$. "GDP per capita in prefecture" is gross value-added per capita in prefecture $p$ in year $t$. "Middle-school graduates in prefecture" is the number of students who graduated from middle schools in prefecture $p$ in year $t$. "Middle-school graduates in nearby prefectures" is the number of students who graduated from middle schools in the prefectures within 100 km from prefecture $p$ (excluding prefecture $p$) in year $t$. "Share of applicants to School 1 (dur. Dapp)" is the share of applicants to School 1 among all applicants to Schools 1–8 in prefecture $p$ under the decentralized system in 1916 (the only year for which the information is available). All variables interacted with "Centralized" are standardized to have a mean of 0 and a standard deviation of 1. We control for year fixed effects, prefecture fixed effects, and the number of higher schools other than Schools 1–8 in prefecture $p$ in year $t$. We also control for "Born in school's prefecture", "Born near school's prefecture (1-100 km)", and "Born near school's prefecture (100-300 km)" as in Table 3. Standard errors reported in parentheses are clustered at the prefecture level, and standard errors reported in square brackets are clustered at the year level. ***, **, and * mean significance at the 1%, 5%, and 10% levels, respectively.

Figure 5: Long-run Impacts of Centralization: Geographical Origins of Career Elites

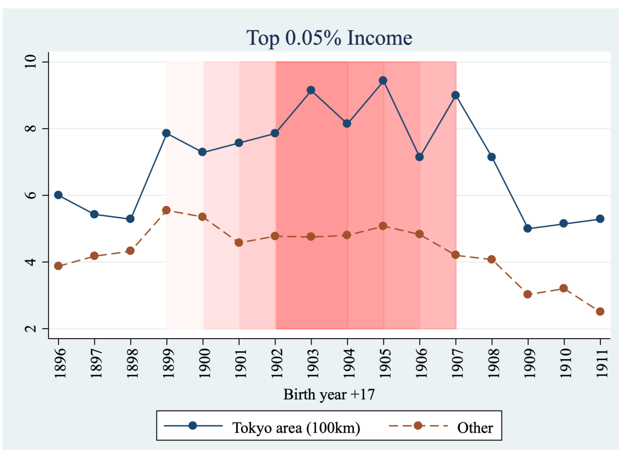
(a)



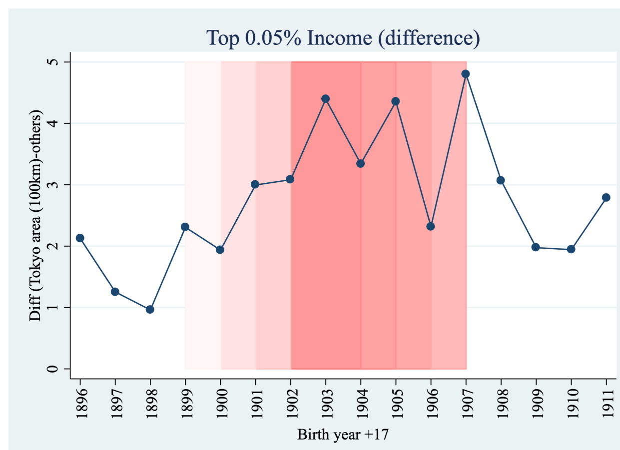

(b)




(c)



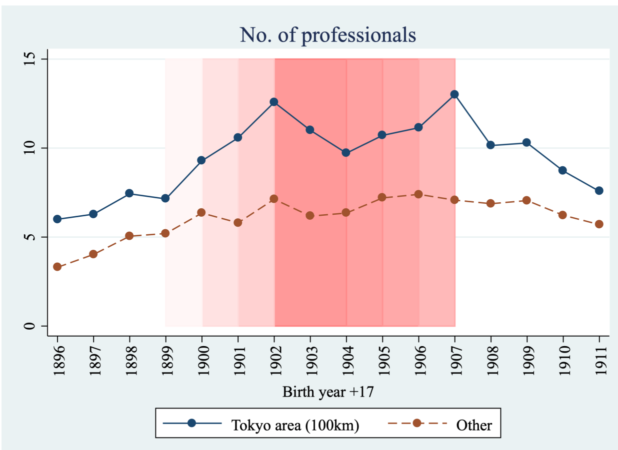

(d)



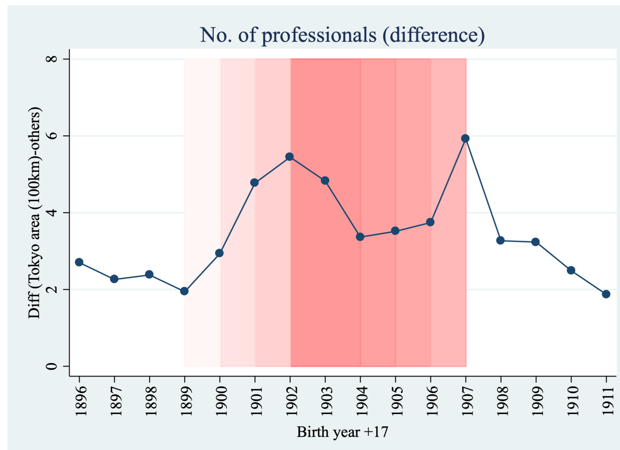

*Notes*: This figure shows difference-in-differences plots that compare the average number of elites born in prefectures inside and outside the Tokyo area by cohorts. The plots are based on the prefecture-cohort level data from the JPIR in 1939, counting the number of elites born in the prefecture by birth cohorts between 1879-1894. The vertical axis shows the number of individuals in each of the above categories of elites who were born in the indicated area in the indicated birth cohort. The cohorts are colored according to their intensity of exposure to the first period of centralized admissions (Capp) in 1902–1907, where the darker color indicates the higher intensity of exposure. The intensity gradually increases from the cohort who turned age 17 in 1899 as some unsuccessful applicants might have retaken the exam in 1902 under Capp. The intensity reaches the highest level for the cohorts who turned age 17 during 1902–1904 and declines from the cohort who turned age 17 in 1904 as some might have retaken the exam in 1908 under Dapp. The intensity drops to zero for the cohort who turned age 17 in 1908 as they had no opportunity to take the exam under Capp. See Section 5.1 for discussions about this figure.

Figure 6: Long-run Impacts of Centralization: Geographical Origins of Career Elites (Continued)

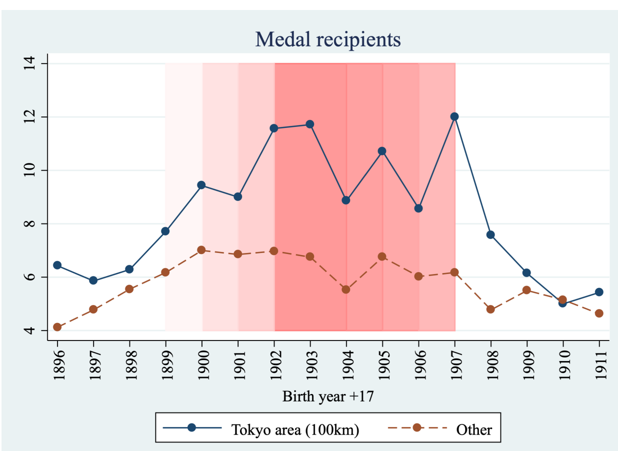
(e)




(f)




(g)




(h)

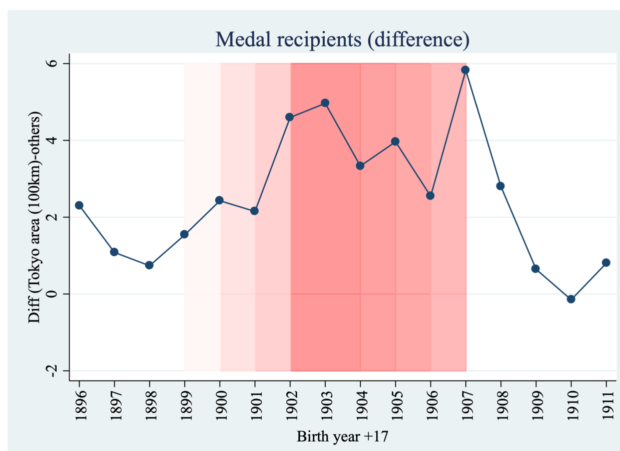


Table 5: Long-run Impacts of Centralization: Difference-in-Differences Estimates

| VARIABLES | (1) Imperial Univ. grads | (2) Top 0.01% income earners | (3) Top 0.05% income earners | (4) Managers | (5) Professionals | (6) Imperial Univ. professors | (7) Medal recipients |
|---|---|---|---|---|---|---|---|
| | | | | A. Baseline Specification | | | |
| Age 17 during Centralization | 3.18 | 0.62 | 1.58 | 3.38 | 1.67 | 0.46 | 2.87 |
| × Tokyo area (<100 km) | (0.027)** | (0.015)** | (0.053)* | (0.084)* | (0.039)** | (0.046)** | (0.010)*** |
| | [0.000]*** | [0.026]** | [0.007]*** | [0.002]*** | [0.009]*** | [0.116] | [0.000]*** |
| | | | | B. With Control Variables | | | |
| Age 17 during Centralization | 2.14 | 0.55 | 1.47 | 2.76 | 1.02 | 0.40 | 2.44 |
| × Tokyo area (<100 km) | (0.000)*** | (0.020)** | (0.030)** | (0.042)** | (0.062)* | (0.042)** | (0.002)*** |
| | [0.002]*** | [0.005]*** | [0.005]*** | [0.002]*** | [0.124] | [0.127] | [0.001]*** |
| | | | | C. Bidirectional Specification with Control Variables | | | |
| Age≤17 in 1902 | 1.76 | 0.73 | 1.72 | 3.40 | 0.60 | 0.29 | 1.97 |
| × Tokyo area (<100 km) | (0.068)* | (0.065)* | (0.003)*** | (0.074)* | (0.326) | (0.141) | (0.001)*** |
| | [0.005]*** | [0.033]** | [0.008]*** | [0.001]*** | [0.345] | [0.246] | [0.001]*** |
| Age≤17 in 1908 | -2.58 | -0.34 | -1.18 | -2.02 | -1.49 | -0.53 | -2.98 |
| × Tokyo area (<100 km) | (0.001)*** | (0.022)** | (0.221) | (0.024)** | (0.079)* | (0.023)** | (0.004)*** |
| | [0.012]** | [0.153] | [0.055]* | [0.034]** | [0.043]** | [0.059]* | [0.009]* |
| | | | | D. Centralization Exposure with Control Variables | | | |
| Cohort's exposure to centralization | 2.23 | 0.64 | 1.53 | 2.92 | 1.10 | 0.46 | 2.52 |
| × Tokyo area (<100 km) | (0.000)*** | (0.008)*** | (0.024)** | (0.020)** | (0.079)* | (0.053)* | (0.001)*** |
| | [0.001]*** | [0.007]*** | [0.003]*** | [0]*** | [0.084]* | [0.118] | [0.001]*** |
| Observations | 705 | 705 | 705 | 705 | 705 | 705 | 705 |
| Birth cohort FE | YES | YES | YES | YES | YES | YES | YES |
| Birth prefecture FE | YES | YES | YES | YES | YES | YES | YES |
| Mean dep var | 8.77 | 1.24 | 4.76 | 9.92 | 6.76 | 0.81 | 6.28 |
| Mean dep var (Tokyo area during Dapp) | 10.62 | 1.512 | 6.22 | 13.59 | 8.60 | 0.87 | 6.94 |

*Notes*: This table shows difference-in-differences estimates of the long-run effects of the centralized admission system on the geographical origins of career elites. The estimates are based on the prefecture-cohort level data from the JPIR in 1939, counting the number of elites born in each prefecture and each cohort between 1880-1894. "Age 17 during Centralization" is the indicator variable that takes 1 if the cohort became age 17 (main application age) during Capp in 1902-1907. "Age≤17 in 1902" (or "Age≤17 in 1908") is the indicator variable that takes 1 if the cohort turned 17 years old in 1902 (or 1908) or later. "Mean dep var" shows the mean of the dependent variable for all prefecture-cohort observations. "Mean dep var (Tokyo area during Dapp)" shows the mean of the dependent variable in the Tokyo area during decentralization. In panels B, C, and D, we control for time- and cohort-varying prefecture characteristics, i.e., the number of primary schools in the prefecture in the year when the cohort turned eligible age, the number of middle school graduates in the prefecture in the year when the cohort turned age 17, log of manufacturing GDP of the prefecture when the cohort turned age 20, and birth population of the cohort in the prefecture. P-values based on standard errors clustered at prefecture level are in parentheses. Wild cluster bootstrap p-values based on standard errors clustered at cohort level are in square brackets. See Section 5.1 for discussions about this table.

Table 6: Long-run Impacts of Centralization: Pathways and Placebo Tests

| VARIABLES | (1) Pathway: Fraction moved in the long-run | (2) Pathway: Distance moved in the long-run | (3) Pathway: Imperial Univ. grads. divided by school entrants | (4) Placebo: Landlords | (5) Placebo: Population |
|---|---|---|---|---|---|
| | | | A. Baseline Specification | | |
| Age 17 during Centralization × Tokyo area (<100 km) | -0.01 | -4.81 | -0.01 | 0.13 | 0.34 |
| | (0.449) | (0.471) | (0.190) | (0.650) | (0.270) |
| | [0.418] | [0.734] | [0.488] | [0.825] | [0.238] |
| | | | B. Adding Control Variables | | |
| Age 17 during centralization × Tokyo area (<100 km) | -0.02 | -8.54 | -0.01 | 0.21 | 0.12 |
| | (0.368) | (0.405) | (0.290) | (0.532) | (0.745) |
| | [0.304] | [0.505] | [0.601] | [0.633] | [0.685] |
| Observations | 705 | 705 | 703 | 705 | 705 |
| Birth cohort FE | YES | YES | YES | YES | YES |
| Birth prefecture FE | YES | YES | YES | YES | YES |
| Mean dep var | 0.23 | 89.73 | 0.038 | 1.22 | 11.67 |
| Mean dep var (Tokyo area during Dapp) | 0.24 | 24.07 | 0.039 | 3.41 | 13.18 |

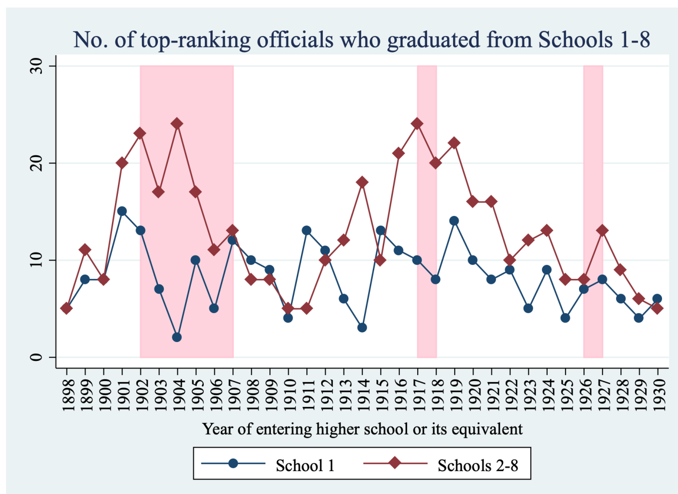
*Notes*: This table shows difference-in-differences estimates to explore pathways of the long-run effects and to provide placebo tests. The estimates are based on the prefecture-cohort level data from the JPIR in 1939, counting the number of elites born in each prefecture and each cohort between 1880-1894. In (1), "Fraction moved" is defined as the fraction of individuals whose prefecture of residence is different from his birth prefecture. In (2), "Distance moved" is defined as the average distance between the birth prefecture and the prefecture of residence of individuals. In (3), "Imperial University graduates divided by School Entrants" is defined by the number of Imperial University graduates divided by the total number of entrants to Schools 1–8 in the year when the cohort became age 17. This variable is a measure of the quality of Schools 1–8 entrants. In (4), "Landlords" is defined as the number of individuals whose occupational titles include landlord, but excluding managers and professionals. In (5), "Population" is the cohort's birth population in the birth prefecture. "Age 17 during Centralization" is the indicator variable that takes 1 if the cohort became age 17 during Capp in 1902-1907. "Mean dep var" shows the mean of the dependent variable for all prefecture-cohort observations. "Mean dep var (Tokyo area during Dapp)" shows the mean of the dependent variable in the Tokyo area during decentralization. In panel B, we control for time- and cohort-varying prefecture characteristics, i.e., the number of primary schools in the prefecture in the year when the cohort turned eligible age, the number of middle school graduates in the prefecture in the year when the cohort turned age 17, log of manufacturing GDP of the prefecture when the cohort turned age 20, and birth population of the cohort in the prefecture (except for column (5)). P-values based on standard errors clustered at prefecture level are in parentheses. Wild cluster bootstrap p-values based on standard errors clustered at cohort level are in square brackets. ***, **, and * mean significance at the 1%, 5%, and 10% levels, respectively. See Section 5.1 for discussions about this table.

Table 7: Long-run Impacts of Centralization: Destinations of Career Elites

| VARIABLES | (1) Imperial Univ. grads | (2) Top 0.01% income earners | (3) Top 0.05% income earners | (4) Managers | (5) Professionals | (6) Imperial Univ. professors | (7) Medal recipients |
|---|---|---|---|---|---|---|---|
| | | | | A. Baseline Specification | | | |
| Age 17 during centralization × Tokyo area (<100 km) | 4.20 | 0.48 | 1.86 | 6.81 | 2.55 | 0.82 | 3.12 |
| | (0.060)* | (0.143) | (0.284) | (0.175) | (0.069)* | (0.132) | (0.057)* |
| | [0.000]*** | [0.142] | [0.058]* | [0.001]*** | [0.001]*** | [0.040]** | [0.003]*** |
| | | | | B. Adding Control Variables | | | |
| Age 17 during centralization × Tokyo area (<100 km) | 3.17 | 0.69 | 2.08 | 5.92 | 1.83 | 0.54 | 2.73 |
| | (0.014)** | (0.230) | (0.279) | (0.145) | (0.029)** | (0.031)** | (0.037)** |
| | [0.007]*** | [0.004]*** | [0.015]** | [0.002]*** | [0.024]** | [0.087]* | [0.008]*** |
| Observations | 705 | 705 | 705 | 705 | 705 | 705 | 705 |
| Birth cohort FE | YES | YES | YES | YES | YES | YES | YES |
| Prefecture of residence FE | YES | YES | YES | YES | YES | YES | YES |
| Mean dep var | 8.30 | 1.22 | 4.65 | 12.75 | 6.43 | 0.70 | 5.94 |
| Mean dep var (Tokyo area during Dapp) | 14.73 | 2.46 | 8.95 | 25.60 | 10.70 | 1.10 | 9.79 |

*Notes*: This table shows difference-in-differences estimates of the long-run effects of the centralized admission system on the geographical destinations of career elites. The estimates are based on the prefecture-cohort level data from the JPIR in 1939, counting the number of elites who resided in the prefecture in 1939 by birth cohorts between 1880 and 1894. Unlike the previous tables, all outcome variables are measured at the level of prefecture of residence and birth cohort. In panel B, we control for time- and cohort-varying prefecture characteristics, i.e., the number of primary schools in the prefecture in the year when the cohort turned eligible age, the number of middle school graduates in the prefecture in the year when the cohort turned age 17, log of manufacturing GDP of the prefecture when the cohort turned age 20, and birth population of the cohort in the prefecture. P-values based on standard errors clustered at prefecture level are in parentheses. Wild cluster bootstrap p-values based on standard errors clustered at cohort level are in square brackets. See Section 5.1 for discussions about this table.

Figure 6: Long-run Impacts of Centralization: Top-ranking Officials




*Notes*: This figure plots the number of top-ranking officials who graduated from School 1 and the number of top-ranking officials who graduated from Schools 2–8 by cohort (defined by the year of entering a higher school). The plots are based on the data from the complete list of individuals who passed the administrative division of the Higher Civil Service Exams (HCSE) in 1894-1941 and their biographical information. The number of top-ranking officials who graduated from School 1 in cohort $t$ is the number of individuals who entered School 1 in year $t$, passed the administrative division of the HCSE, and were promoted to the top three ranks of higher civil service by the end of their lifetime. Colored cohorts are those who entered Schools 1–8 during the centralized admission system (Capp) in 1902–07, 1917–18, and 1926–27.

Table 8: Long-run Impacts of Centralization: Higher Civil Officials

(a) Passers of the Higher Civil Service Exams

| VARIABLES | (1) Exam passers | (2) Exam passers graduated from School 1 | (3) Exam passers graduated from Schools 2–8 | (4) Exam passers not graduated from Schools 1–8 | (5) Exam passers graduated from Schools 2–8 and Tokyo Imperial Univ. |
|---|---|---|---|---|---|
| Centralized | 5.96 | -2.81 | 14.93** | -12.11** | 18.70*** |
| | (23.59) | (2.25) | (5.68) | (5.75) | (6.32) |
| Higher Civil Service Exam passers | | 0.10*** | 0.35*** | 0.55*** | 0.26*** |
| | | (0.02) | (0.07) | (0.08) | (0.06) |
| Observations | 33 | 33 | 33 | 33 | 33 |
| Mean dep var | 189.6 | 37.29 | 68.47 | 83.80 | 58.88 |
| Mean dep var (Dapp) | 194.2 | 37.96 | 64.09 | 92.17 | 53.22 |

(b) Top-Ranking Officials

| VARIABLES | (1) Top-ranking officials | (2) Top-ranking officials graduated from School 1 | (3) Top-ranking officials graduated from Schools 2–8 | (4) Top-ranking officials not graduated from Schools 1–8 | (5) Top-ranking officials graduated from Schools 2–8 | (6) Top-ranking officials graduated from Schools 2–8 and Tokyo Imperial Univ. |
|---|---|---|---|---|---|---|
| Centralized | 4.19*** | -0.48 | 5.22*** | -0.55 | 2.72*** | 1.76** |
| | (1.40) | (1.04) | (1.30) | (1.02) | (0.67) | (0.81) |
| Higher Civil Service Exam passers | 0.10*** | 0.02* | 0.05*** | 0.03*** | | |
| | (0.02) | (0.01) | (0.02) | (0.01) | | |
| Exam passers graduated from Schools 2-8 | | | | | 0.17*** | |
| | | | | | (0.02) | |
| Exam passers graduated from Schools 2-8 and Tokyo Imperial Univ. | | | | | | 0.19*** |
| | | | | | | (0.02) |
| Observations | 33 | 33 | 33 | 33 | 33 | 33 |
| Mean dep var | 29.77 | 8.273 | 12.97 | 8.528 | 12.97 | 11.82 |
| Mean dep var (Dapp) | 28.66 | 8.304 | 11.22 | 9.138 | 11.22 | 10 |

*Notes*: Panel (a) shows OLS estimates of the effects of the centralized admissions (Capp) on the number of individuals who passed the administrative division of the Higher Civil Service Exams (administrative HCSE). Panel (b) shows OLS estimates of the effects of Capp on the number of top-ranking civil officials. The estimates are based on the cohort level data, 1898–1930, where cohort is defined by the year of entering a higher school or its equivalent. The data are compiled from the list of individuals who passed the administrative HCSE in 1894–1941 and their biographical information. "Higher Civil Service Exam passers" or "Exam passers" is the number of individuals in cohort $t$ who passed the administrative HCSE. "Top-ranking officials" is the number of top-ranking officials in cohort $t$ (i.e., the number of individuals who entered a higher school or its equivalent in year $t$, passed the administrative HCSE, and were promoted to the top 3 ranks of higher civil service in their lifetime). "Centralized" is the indicator variable that takes 1 if cohort $t$ entered a higher school or its equivalent during Capp in 1902–07, 1917–18, and 1926–27. "Mean dep var (Dapp)" is the mean of the dependent variable for the cohorts who entered a higher school or its equivalent during the decentralized admissions (Dapp). In all regressions, we control for quadratic time trends. Newey-West standard errors with the maximum lag order of 3 are shown in the parentheses. ***, **, and * mean significance at the 1%, 5%, and 10% levels, respectively. See Section 5.2 for discussions about this table.

# A Appendix

## A.1 Looking at Centralized Assignment

Appendix Table A.1 presents the number of admitted applicants (to the Departments of Law and Literature in each school) and their exam scores by their school preference order. Observe that, in the prestigious Schools 1 (Tokyo) and 3 (Kyoto), all seats were filled with applicants who ranked these schools first. Both the maximum and minimum exam scores of School 1 entrants were the highest among all schools, confirming that School 1 was the most selective, followed by School 3, School 4, and School 2, in that order. By contrast, Schools 5 and 7 admitted a sizable number of students who ranked the school third or lower, because they did not have a sufficient number of high-scoring applicants who placed these schools at the top of their preferences. These observations confirm the selectivity hierarchy among schools, which plays a key role in our analysis.

Students who were admitted to the school of their third choice or below are not necessarily low ability students. For example, the highest-score entrant to School 7 (with the score of 450) was the applicant admitted to his third choice after failing to enter Schools 1 and 3 by a narrow margin. This verifies the fact that schools received reluctant and unmotivated students who came to the schools as a fallback option, as claimed by local interest parties (see Section 4.5).

## A.2 Additional Theoretical Results

Let $\mu^I$ be the mechanism that selects a matching based on the following Student-Proposing Deferred Acceptance or Serial Dictatorship algorithm.

- Step 1. Each student $i$ proposes to her most-preferred school. Each school $s$ holds top $q_s$ students and rejects the rest. If less than $q_s$ students proposed, then it holds all the students that proposed to $s$.

- Step $k$. Any student who was rejected at step $k-1$ makes a new proposal to his most-preferred school that has not yet rejected him. If no acceptable choices remain, she makes no proposal. Each school holds its most-preferred $q_s$ students to date and rejects the rest. If less than $q_s$ students proposed, then it holds all the students who proposed to $s$.

- The algorithm terminates when there are no more rejections. Each student is assigned to the school that holds her in the last step.

Motivated by the fact that Schools 1–8 are prestigious public schools with no significant competitors, we assume that every student prefers Schools 1–8 over the outside option.

**Assumption 1.** $s \succ_i o$ for all $i \in I$ and $s \in S$.

Under this assumption, $\mu^I$ and Capp are partially equivalent in the following sense.

**Proposition 4.** *For any school choice problem with Assumption 1,* $\cup_{s\in S}\mu_s^C(\succ) = \cup_{s\in S}\mu_s^I(\succ')$ *for all* $\succ, \succ' \in P$.

This result says that the same students are assigned to schools under Capp and DA, which is the most meritocratic mechanism we can design. This result holds regardless of applicant behavior.

## A.3 Proofs

*Proof of Proposition 1.* As mentioned in step 1 of Capp, school seats are assigned to applicants $i_1, \dots, i_k$ under $\mu^C$, i.e., $\cup_{s\in S}\mu_s^C(\succ) = \{i_1, \dots, i_k\}$ and $\cup_{j\in\{1,\dots,k\}}\{t_{\mu^C(\succ)}(j)\} = \{t_{i_1}, \dots, t_{i_k}\}$. Let $\cup_{s\in S}\mu_s(\succ') = \{i_{j_1}, \dots, i_{j_l}\}$ with $l \le k$, $j_1 < \dots < j_l$ and $\{j_1, \dots, j_l\} \subseteq \{1, \dots, n\}$. This gives $\cup_{j\in\{1,\dots,k\}}\{t_{\mu(\succ')}(j)\} = \{t_{i_{j_1}}, \dots, t_{i_{j_l}}, \cup_{i=1}^{k-l}\{0\}\}$. Since $t_{i_1} \ge t_{i_{j_1}}$,..., $t_{i_l} \ge t_{i_{j_l}}$, we have that $\left|\{j \in \{1, \dots, k\} \mid t_{\mu^C(\succ)}(j) \le t\}\right| \le \left|\{j \in \{1, \dots, k\} \mid t_{\mu(\succ')}(j) \le t\}\right|$ so that

$$F_{\mu^C(\succ)}(t) = \frac{\left|\{j \in \{1, \dots, k\} \mid t_{\mu^C(\succ)}(j) \le t\}\right|}{k} \le \frac{\left|\{j \in \{1, \dots, k\} \mid t_{\mu(\succ')}(j) \le t\}\right|}{k} = F_{\mu(\succ')}(t).$$

Therefore, we have that $F_{\mu^C(\succ)}(t) \le F_{\mu(\succ')}(t)$ for all $t \in \mathbb{R}_+$ and $\succ, \succ' \in P$. □

*Proof of Proposition 2.* The proposition follows from a lemma below.

**Lemma 1.** *(a) Under Capp, submitting the true preference is a dominant strategy.*

*(b) Under Dapp, there is no dominant strategy.*

*Proof of Lemma 1.* Suppose, without loss of generality, for applicant $i$ $s_1 \succ_i^o s_2 \succ_i^o i$ where $i$ denotes remaining unassigned.

**Part (a).** Under Capp, applicant $i$ has four strategies available: reporting $s_1$ as first choice and $s_2$ as second, denoted $a_i$ $(=\succ_i^o)$; reporting $s_2$ as first choice and $s_2$ as second choice, denoted $a_i'$; and reporting a single school as top choice, either $s_1$ or $s_2$. Fix any $a_{-i} \in A_{-i}$. We have to show that reporting $a_i$ is a dominant strategy for applicant $i$.

Notice that reporting a single school as top choice is not a dominant strategy since it is dominated by reporting that school as first choice and the other school as second because

$$\sum_{k=1}^{2} p_{ik}(a_i, a_{-i}) > \sum_{k=1}^{2} p_{ik}(s_1, a_{-i}) \quad \text{and} \quad \sum_{k=1}^{2} p_{ik}(a_i', a_{-i}) > \sum_{k=1}^{2} p_{ik}(s_2, a_{-i})$$

Now we show that $a_i$ dominates $a_i'$. First, only top $k$ students are assigned a school, that implies if a student is unassigned under $\mu^C(a_i, a_{-i})$, he would be unassigned under $\mu^C(a_i', a_{-i})$ as well i.e. $p_{i3}(a_i, a_{-i}) = p_{i3}(a_i', a_{-i})$. Therefore,

$$\sum_{k=1}^{2} p_{ik}(a_i, a_{-i}) = \sum_{k=1}^{2} p_{ik}(a_i', a_{-i})$$

.

Second, if the student gets $s_2$ by reporting $s_1$ as first choice, it is clear that he cannot get $s_1$ by reporting $s_2$ as first choice because in that case he would be assigned $s_2$ in the second step of Capp. Therefore,

$$p_{i1}(a_i, a_{-i}) \geq p_{i1}(a_i', a_{-i})$$

.

Therefore we have that $a_i$ is a dominant strategy.

**Part (b).** Under Dapp, applicant $i$ has two strategies available: applying to $s_1$, denoted $a_i$, and applying to $s_2$, denoted $a_i'$. Fix any any $a_{-i} \in A_{-i}$. Notice that $p_{i2}(a_i, a_{-i}) = p_{i1}(a_i', a_{-i}) = 0$.

$a_i$ is not a dominant strategy since in the case $a_{-i}$ is such that all students apply to $s_1$, (note that applicant $i$ is one of the top $q_1$ students with a positive probability) we have that,

$$\sum_{k=1}^{2} p_{ik}(a_i, a_{-i}) = p_{i1}(a_i, a_{-i}) < 1 = p_{i2}(a_i', a_{-i}) = \sum_{k=1}^{2} p_{ik}(a_i', a_{-i})$$

.

$a_i'$ is not a dominant strategy either since in the case $a_{-i}$ is such that all students apply to $s_2$, $p_{i1}(a_i, a_{-i}) = 1$ and therefore, $a_i \, sd(\succ_i^o) \, a_i'$. □

*Proof of Proposition 3.*

**Lemma 2.** *For sufficiently large $V$, all applicants apply to their local schools in any symmetric equilibrium under Dapp.*

*Proof of Lemma 2.* First, we show that for sufficiently large $V$, none of the following symmetric equilibrium survive: (i) all applicants apply to $s_i$ (for $i = 1, 2$), and (ii) applicants from $s_1$'s area apply to $s_2$ while those from $s_2$'s area apply to $s_1$.

**Case (i).** Applicants from school $j$'s area apply to $s_i$ if $p(n_i + n_j, q_i) * U_i \geq U_j + V$. For $V > (p(n_i + n_j, q_i) * U_i) - U_j$, therefore, all applicants applying to $s_i$ (for $i = 1, 2$) cannot be a symmetric equilibrium.
**Case (ii).** Suppose applicants from $s_1$'s area apply to $s_2$ while those from $s_2$'s area apply to $s_1$. It must be that the case that, for applicants from $s_1$'s area: $p(n_1, q_2) * U_2 \geq p(n_2 + 1, q_1) * (U_1 + V)$. While for applicants from $s_2$'s area: $p(n_2, q_1) * U_1 \geq p(n_1 + 1, q_2) * (U_2 + V)$. For sufficiently large $V$, this cannot be a symmetric equilibrium.

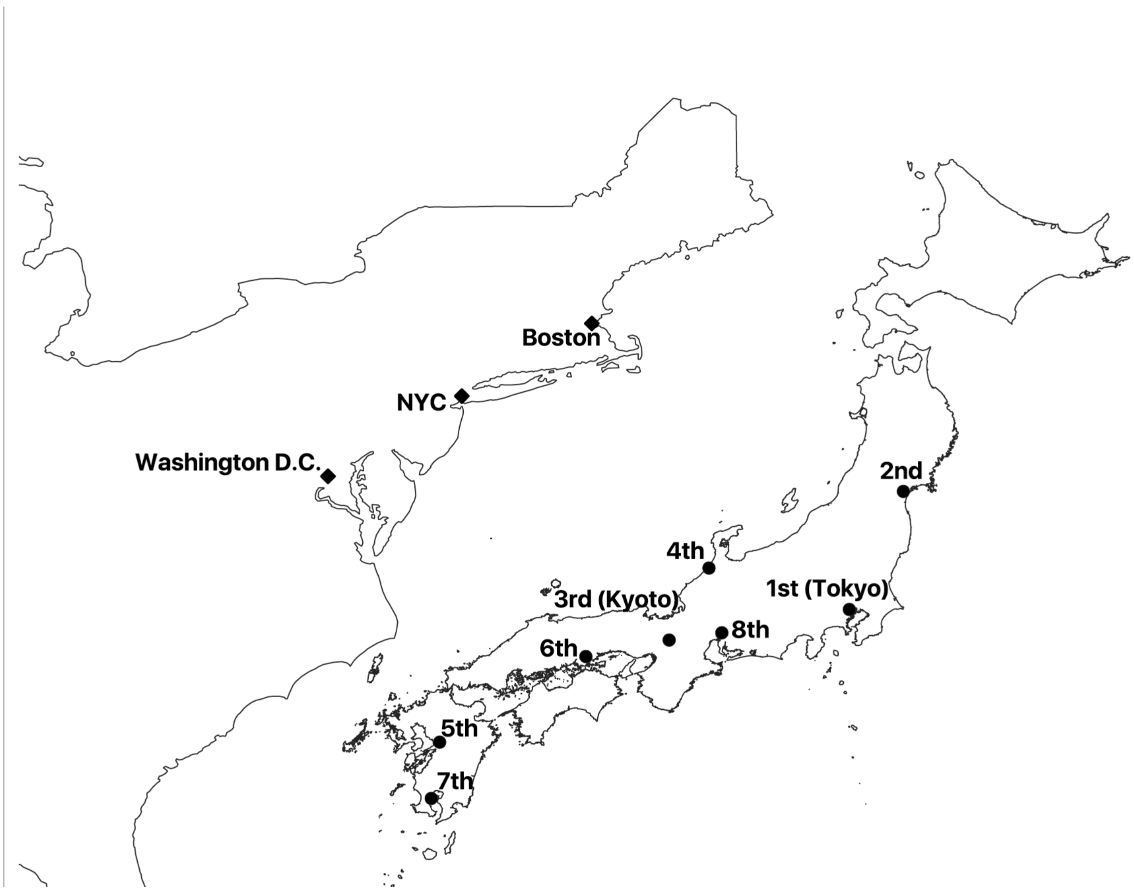
Now we show that, for large enough $V$, all students applying to their local schools is indeed a symmetric equilibrium. For applicants from school 1's area to apply to $s_1$, it must be the case that $p(n_1, q_1) * (U_1 + V) \geq p(n_2 + 1, q_2) * U_2$. For applicants from school 2's area to apply to $s_2$, $p(n_2, q_2) * (U_2 + V) \geq p(n_1 + 1, q_1) * U_1$ must hold. Since the left hand sides of both the inequalities are increasing in $V$, the equilibrium conditions hold for sufficiently large $V$. □

From Lemma 2, under assumption A1, we know that under Dapp applicants apply to their locals schools. Therefore, the expected proportion of assigned applicants assigned to their local school under Dapp is 1 (the highest).

*Proof of Proposition 4.* As mentioned in step 1 of Capp, school seats are assigned to applicants $i_1, \ldots, i_k$ under $\mu^C$, i.e., $\cup_{s \in S} \mu^C_s(\succ) = \{i_1, \ldots, i_k\}$ for all $\succ \in P$. Under assumption 1, any student $i_{k'}$ with $k' > k$ will be rejected at some step of the Student-Proposing Deferred Acceptance Algorithm. Assumption 1 therefore implies that the top $k$ students are assigned to some school under $\mu^I$, i.e., $\cup_{s \in S} \mu^I_s(\succ') = \{i_1, \ldots, i_k\}$ for all $\succ' \in P$. Therefore, $\cup_{s \in S} \mu^C_s(\succ) = \cup_{s \in S} \mu^I_s(\succ') = k$ for all $\succ, \succ' \in P$. □

## A.4 Additional Tables and Figures

Figure A.1: Map of Schools 1–8 in Japan (with the U.S. East Coast in Comparison)




*Notes*: This figure shows the locations of Schools 1–8 and compares their geographical distribution to the US East Coast in the same scale unit. See Section 2 for discussions about this figure.

Figure A.2: Centralized Assignment Rule

第六條　入學ヲ許可スヘキ者ハ左ノ方法ニ依リ之ヲ定ム

一、高等學校ヲ通シ各部每ニ其ノ部ニ入學セシムヘキ人員ノ總數ト同數ノ人員ヲ試驗ノ成績順ニ依リ選出ス

二、前號ノ場合ニ於テ試驗成績相同シキトキハ抽籤ニ依ル

三、前號ニ依リ選出セル人員ニ就キ試驗ノ成績順ニ依リ第一部又ハ第二部ノ志望者ニ在リテハ本人ノ指定スル第一ノ志望類ニ基キ第一ノ志望學校ニ第三部ノ志望者ニ在リテハ本人ノ指定スル第一ノ志望學校ニ配當ス

四、前號ノ場合ニ於テ試驗成績相同シキトキハ抽籤ニ依ル

五、第三號及第四號ニ依リ配當ノ結果本人ノ指定スル第一ノ志望學校已ニ滿員トナリタル場合ニ於テハ更ニ成績順ニ依リ本人ノ指定スル第二ノ志望學校ニ配當ス

六、前號ノ場合ニ於テ試驗ノ成績相同シキトキハ抽籤ニ依ル

七、第五號及第六號ニ依リ配當ノ結果本人ノ指定スル第二ノ志望學校已ニ滿員トナリタル場合ニ於テハ更ニ其ノ第三以下ノ志望學校ニ就キ第五號及第六號ニ準シ配當ス

八、第一部又ハ第二部ノ志望者ニ在リテハ本人ノ指定スル第一ノ志望類カ志望學校ニ於テ悉ク滿員トナリタルトキハ更ニ本人ノ指定スル第二以下ノ志望順中缺員アルモノニ之ヲ配當ス其ノ方法ハ第三號乃至第七號ニ準ス

九、本人ノ志望スル類及學校悉ク滿員トナリタルトキハ入學スルコトヲ得サルモノトス

前項ニ依リ配當ノ結果又ハ事故ノ爲メ入學者ニ缺員ヲ生シタルトキハ入學スルコトヲ得サリシ者ニ就キ更ニ前項ノ方法ニ依リ之ヲ補塡ス

*Notes*: This figure is a reprint of the assignment algorithm of the centralized admission system stated in the Ordinance of the Ministry of Education No. 4 published in *Government Gazette* No. 1419, pp.580-581, on April 27, 1917. See Section 2 for an English translation and discussions.

Table A.1: Admission Outcomes of the Centralized Assignment Algorithm

Exam Scores of Entrants in 1917 Under Centralized Admission System

| School Name | School 1 | School 2 | School 3 | School 4 | School 5 | School 6 | School 7 | School 8 |
|---|---|---|---|---|---|---|---|---|
| Location | Tokyo | Sendai | Kyoto | Kanazawa | Kumamoto | Okayama | Kagoshima | Nagoya |
| Total no. of entrants | 77 | 29 | 38 | 22 | 68 | 36 | 37 | 64 |
| Entrants Admitted to their 1st Choice | | | | | | | | |
| No. of entrants | 77 | 14 | 38 | 18 | 23 | 18 | 6 | 18 |
| Max exam score | 548 | 462 | 521 | 496 | 471 | 456 | 415 | 455 |
| Min exam score | 451 | 374 | 404 | 364 | 363 | 364 | 364 | 363 |
| Entrants Admitted to their 2nd Choice | | | | | | | | |
| No. of entrants | | 15 | | 4 | 30 | 18 | 8 | 46 |
| Max exam score | | 450 | | 450 | 438 | 433 | 449 | 450 |
| Min exam score | | 442 | | 421 | 362 | 369 | 372 | 363 |
| Entrants Admitted to their 3rd Choice | | | | | | | | |
| No. of entrants | | | | | 15 | | 3 | |
| Max exam score | | | | | 450 | | 450 | |
| Min exam score | | | | | 393 | | 407 | |
| Entrants Admitted to their 4th Choice | | | | | | | | |
| No. of entrants | | | | | | | 9 | |
| Max exam score | | | | | | | 400 | |
| Min exam score | | | | | | | 366 | |
| Entrants Admitted to their 5th Choice | | | | | | | | |
| No. of entrants | | | | | | | 11 | |
| Max exam score | | | | | | | 444 | |
| Min exam score | | | | | | | 369 | |

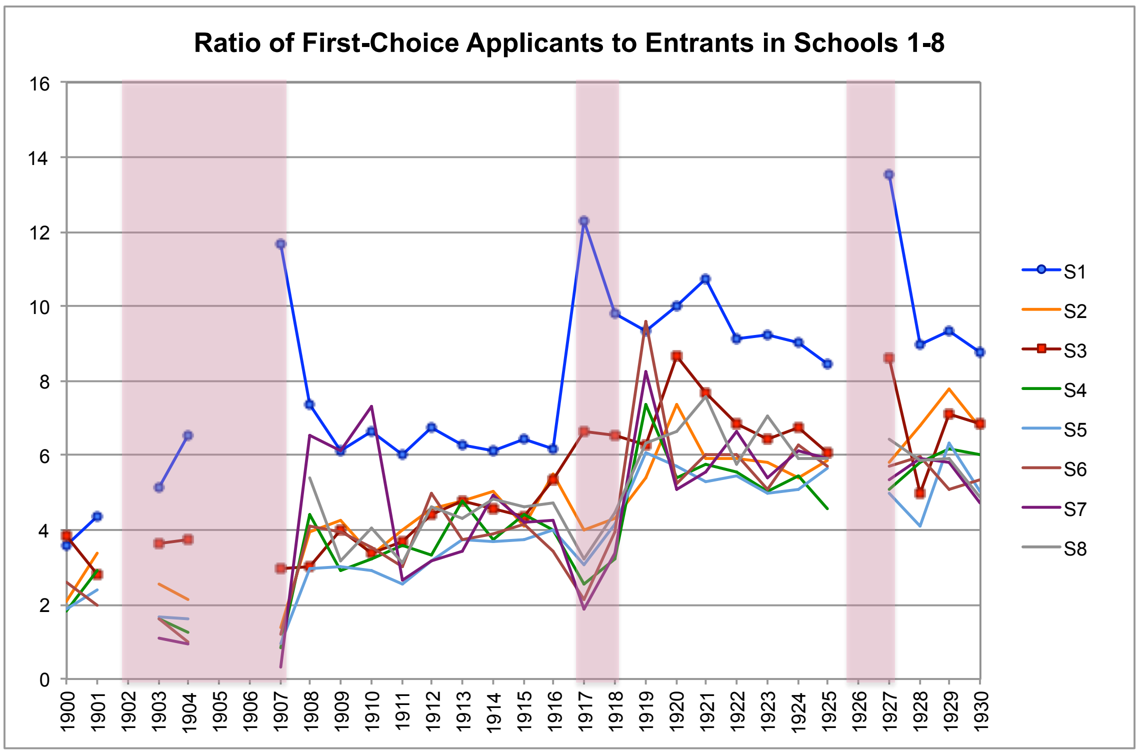
*Notes*: This figure shows admission outcomes for the Departments of Law and Literature in Schools 1–8 in 1917 under the centralized assignment algorithm. See Section A.1 for discussions about this figure.

Table A.2: Summary Statistics

| Variable | Mean | Std. Dev. | Median | N |
|---|---|---|---|---|
| **Year level data on short-run outcomes, 1900–1930** | | | | |
| No. of applicants to Schools 1–8 | 10613 | 4221 | 9997 | 28 |
| Share of applicants choosing School 1 as their first choice | 0.314 | 0.097 | 0.274 | 28 |
| No. of entrants to Schools 1–8 | 1821 | 227 | 1919 | 31 |
| **Applicant level data on short-run outcomes, 1916–1917** | | | | |
| Distance between middle school prefecture and the first-choice school (km) | 224.88 | 272.03 | 117 | 20913 |
| Applying to School 1 as first choice | 0.33 | 0.47 | 0 | 20913 |
| **Entrant level data on short-run outcomes, 1900–1930** | | | | |
| Distance between birth prefecture and the school entered (km) | 226.8 | 258.65 | 139 | 66193 |
| Entering the nearest school from birth prefecture | 0.49 | 0.5 | 0 | 66193 |
| Born in Tokyo prefecture | 0.09 | 0.29 | 0 | 66193 |
| Born in the Tokyo area (7 prefectures within 100 km from Tokyo) | 0.17 | 0.38 | 0 | 66193 |
| **Prefecture-year level data on short-run outcomes, 1900-1930** | | | | |
| No. of entrants to Schools 1-8 | 45.06 | 37.45 | 34 | 1469 |
| No. of entrants to School 1 | 7.88 | 14.11 | 5 | 1469 |
| No. of entrants to School 2 | 5.5 | 10.4 | 2 | 1469 |
| No. of entrants to School 3 | 6.19 | 10.34 | 3 | 1421 |
| No. of entrants to School 4 | 5.64 | 9.91 | 3 | 1469 |
| No. of entrants to School 5 | 6.27 | 14.2 | 1 | 1422 |
| No. of entrants to School 6 | 5.19 | 11.8 | 2 | 1421 |
| No. of entrants to School 7 | 5.03 | 12.99 | 2 | 1328 |
| No. of entrants to School 8 | 5.67 | 12.45 | 2 | 1093 |
| No. of public middle school graduates | 415.38 | 350.01 | 299 | 1410 |
| No. of private middle school graduates | 118.49 | 435.12 | 0 | 1410 |
| No. of national higher schools other than Schools 1-8 | 0.13 | 0.43 | 0 | 1469 |
| **Prefecture-cohort level data on long-run outcomes, JPIR-listed individuals born in 1880–1894** | | | | |
| No. of all Imperial University graduates | 8.77 | 7.81 | 7 | 705 |
| No. of individuals earning top 0.01% level of income | 1.24 | 2.06 | 1 | 705 |
| No. of individuals earning top 0.05% level of income | 4.76 | 6.97 | 3 | 705 |
| No. of managers in private sector paying a positive amount of tax | 13.25 | 16.34 | 7 | 705 |
| No. of scholars, physicians, lawyers, and engineers | 6.76 | 6.11 | 5 | 705 |
| No. of Imperial University professors | 0.81 | 1.18 | 0 | 705 |
| No. of civilians receiving medal of the Order of Fifth Class and above | 6.28 | 5.03 | 5 | 705 |
| **Cohort level data on long-run outcomes, government officials entering higher school or equivalent in 1898-1930** | | | | |
| No. of the passers of the Higher Civil Service Exams | 189.6 | 84.3 | 159.7 | 33 |
| No. of top-ranking officials (internally promoted to top three ranks) | 29.8 | 8.4 | 29.1 | 33 |
| No. of top-ranking officials graduated from School 1 | 8.3 | 3.4 | 8 | 33 |
| No. of top-ranking officials graduated from Schools 2–8 | 13.0 | 6.0 | 12 | 33 |
| No. of top-ranking officials not graduated from Schools 1–8 | 8.5 | 5.7 | 6.8 | 33 |

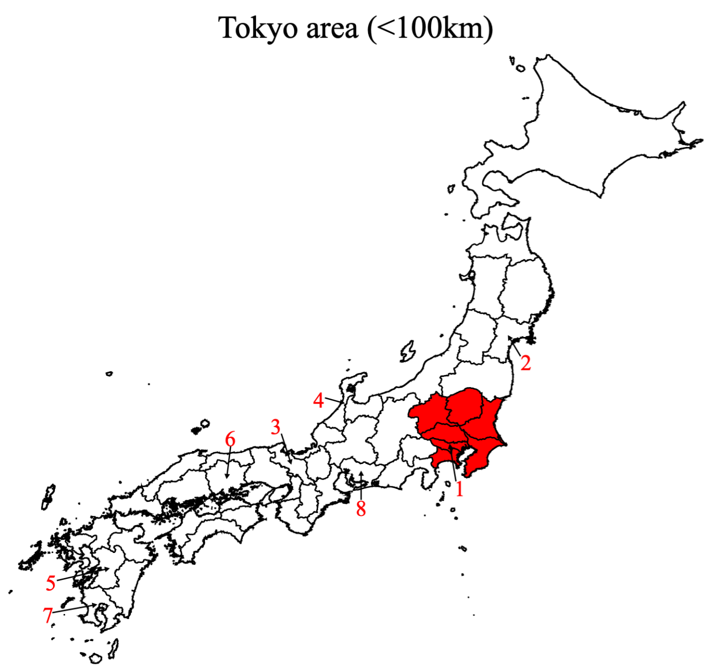
*Notes*: This table provides summary statistics of main variables used in the empirical analyses. For the empirical analyses, see Sections 4 and 5.

Figure A.3: Changes in the Competitiveness of Schools 1–8




*Notes*: This figure shows the changes in the competitiveness of each school (measured by the ratio of the number of applicants who rank the school first to the number of entrants to the school) from 1900 to 1930. No data are available for 1902, 1905, 1906, and 1926. Colored years (1902-07, 1917-18, 1926-27) indicate the periods of the centralized system, while other years used the decentralized system. School 7 in 1901, 1908, 1909, and 1910, and School 8 in 1908 held their exams on different dates from other schools due to special circumstances, attracting a high number of applicants in these years. See Section 4.2 for discussions about this figure.

Figure A.4: Definition of the Tokyo Area




*Notes*: This figure shows the Tokyo area (in the red color) defined as prefectures that are within 100 km from Tokyo (Tokyo, Chiba, Kanagawa, Saitama, Ibaraki, Tochigi, and Gunma prefectures). See Section 4.3 for discussions about this figure.

Table A.3: Which Regions Win from Centralization? Analysis at Middle School Level

| No. entrants to: | (1)<br>Schools 1–6<br>(Unbalanced panel) | (2)<br>Schools 1–6<br>(Balanced panel) | (3)<br>Schools 1<br>(Unbalanced panel) | (4)<br>Schools 1<br>(Balanced panel) | (5)<br>Schools 2–6<br>(Unbalanced panel) | (6)<br>Schools 2–6<br>(Balanced panel) |
|---|---|---|---|---|---|---|
| Centralized × | 2.92** | 3.40* | -2.32*** | -2.71** | 3.82*** | 4.38*** |
| Middle School in Tokyo area (< 100km) | (0.045) | (0.081) | (0.006) | (0.013) | (0.000) | (0.002) |
| Observations | 825 | 380 | 825 | 380 | 825 | 380 |
| Mean dep var | 9.275 | 13.89 | 1.903 | 3.087 | 5.743 | 8.542 |

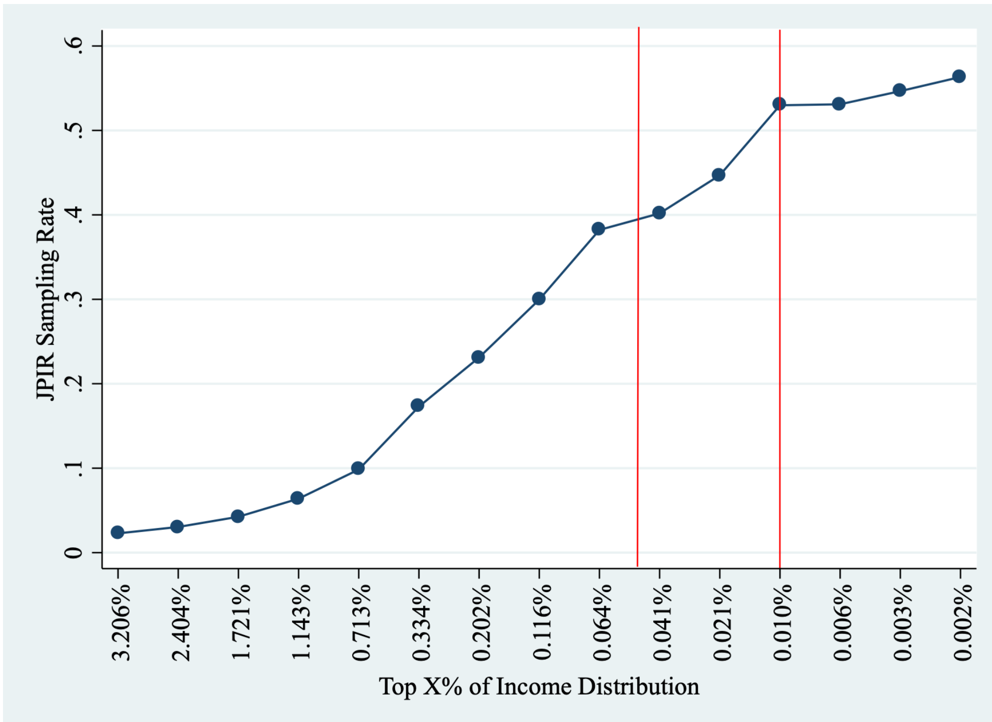
*Notes*: This table uses the panel data of the number of entrants to the higher schools by middle schools by year from 1900 to 1905. Columns indicating "Unbalanced panel" use all of the middle schools from which at least one student entered one of the higher schools at least once during the sample period. Columns indicating "Balanced panel" use the sample of the middle schools from which at least one student entered one of the higher schools every year during the sample period. P-values based on the standard errors clustered at middle school level are shown in the parentheses.

Table A.4: Testing Exogeneity of Centralization

| VARIABLES | (1) No. middle school graduates | (2) No. entrants to Schools 1–7 | (3) Share of entrants to School 1 | (4) No. applicants to Schools 1–7 | (5) Ratio of entrants to applicants | (6) Mean age of entrants | (7) Government expenditure for Higher Schools and Imperial Univ. | (8) Share of applicants to School 1 | (9) Enrollment distance | (10) Share of entrants born in Tokyo area |
|---|---|---|---|---|---|---|---|---|---|---|
| Centralized | 1.351 | 20.55 | -0.00131 | -0.190 | 0.0185 | 0.116 | 1.212 | 0.195*** | 49.36*** | 0.0371*** |
| | (1.647) | (23.50) | (0.00335) | (0.726) | (0.0159) | (0.0741) | (1.871) | (0.0524) | (14.77) | (0.0115) |
| No. of middle school graduates | | -1.248 | -0.000618 | -0.134* | 0.00180 | 0.0549*** | -0.468* | -0.00335 | 1.990** | 0.00167** |
| | | (2.596) | (0.000459) | (0.0674) | (0.00107) | (0.0165) | (0.234) | (0.00202) | (0.832) | (0.000762) |
| Law Department | | 315.9*** | -0.00512 | 1.145 | -0.0481* | 0.437* | 2.271 | 0.218** | -38.41 | -0.0120 |
| | | (42.27) | (0.00799) | (1.706) | (0.0256) | (0.219) | (2.827) | (0.0793) | (24.44) | (0.0215) |
| Observations | 31 | 31 | 31 | 27 | 27 | 26 | 31 | 27 | 31 | 31 |
| Mean dep var | 23.42 | 1821 | 0.196 | 9.790 | 0.214 | 19.03 | 13.56 | 0.313 | 228.9 | 0.174 |

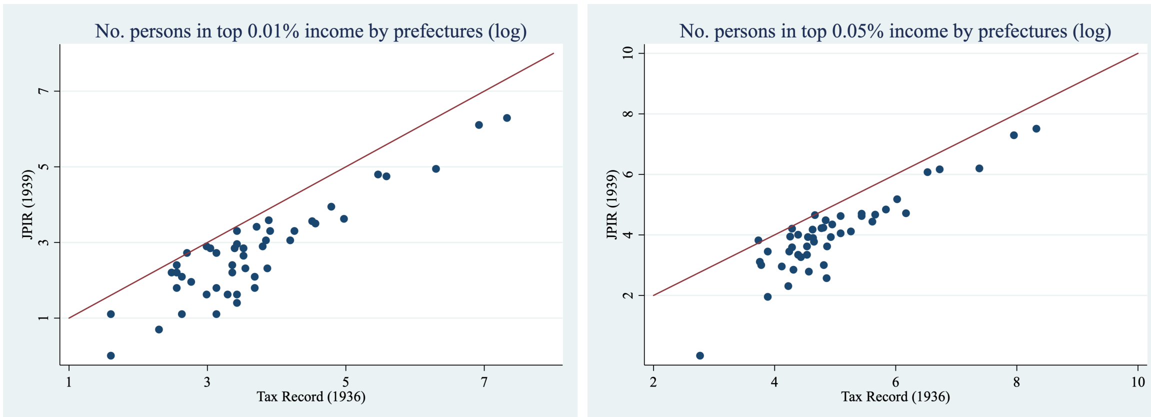
*Notes*: Columns 1–6 test if important institutional variables are correlated with the timing of centralization using year-level data. Columns 7–9 examine the robustness of our main short-run outcomes using year-level data. All numbers are at the national-level from 1900 to 1930. We focus on Schools 1-7 when calculating the number of entrants, the share of entrants to School 1, the number of applicants, the entrants to applicants ratio, and the share of applicants to School 1. The numbers of middle school graduates and applicants are denominated by 1,000. "Law Department" is an indicator variable that takes 1 in 1907 and afterwards to control for the creation of the Law Department in each School 1-8 that increased school capacity. In all regressions, quadratic time trends (i.e. trend and trend squared, where the trend is defined by "year - 1899") are controlled. Newey-West standard errors with the maximum lag order of 3 are shown in the parentheses. See Section 4.6 for discussions about this table.

Figure A.5: Sampling Rates of High Income Earners in the JPIR




*Notes*: This figure plots the sampling rate of the high income earners in JPIR (1939) by the income level expressed as a top percentile of the national income distribution. The sampling rates and the top income percentiles are computed from income tax statistics in the Tax Bureau Yearbook. The vertical lines indicate the top 0.05% and top 0.01%. See Section **??** for discussions about this figure.

Figure A.6: High Income Earners in JPIR vs Income Tax Statistics across Prefectures




*Notes*: These figures compare the number of high income earners in each prefecture listed in the JPIR (1939) and the number of all high income earners in each prefecture reported in the tax statistics in the Tax Bureau Yearbook (1936). The vertical axis is log of the number of individuals in JPIR (1939) who earned more than 50,000 yen taxable income (or the top 0.01% income group) or 18,000 yen taxable income (corresponding to the top 0.05% income group) in 1938. The horizontal axis is log of the number of individuals in tax statistics who earned more than 30,000 yen taxable income (corresponding to the top 0.013% income group) or 10,000 yen taxable income (corresponding to the top 0.08% income group) in 1936 (the closest year to 1938 for which prefecture-level data are available). See Section **??** for discussions about these figures.

Table A.5: Correlations between Prefecture-level Sampling Rates and Outcome Variables

| VARIABLES | Top 0.01% income earners | Top 0.05% income earners | Top 0.01% income earners | Top 0.05% income earners |
|---|---|---|---|---|
| Entrants to Schools 1–8 | -0.000035 | 0.00005 | | |
| | (0.000059) | (0.000038) | | |
| Imperial Univ. grads | | | -0.000052 | 0.000029 |
| | | | (0.000032) | (0.00002) |
| Observations | 47 | 47 | 47 | 47 |
| R-squared | 0.003 | 0.032 | 0.006 | 0.009 |
| Mean dep var | 0.444 | 0.218 | 0.444 | 0.218 |

*Notes*: This table shows the results of regressing the sampling rates of the JPIR (1939) on our outcome variables using prefecture-level data. "Top 0.01% income earners" is the sampling rate of the top 0.01% income earners defined by the number of individuals with more than 50,000 yen of taxable income in 1938 divided by the complete count of the number of individuals with more than 30,000 yen of taxable income in 1936. "Top 0.05% income earners" is the sampling rate of the top 0.05% income earners defined by the number of individuals with more than 18,000 yen of taxable income in 1938 divided by the complete count of the number of individuals with more than 10,000 yen of taxable income in 1936. "Entrants to Schools 1–8 " is the number of entrants to Schools 1–8 during 1900–1911 who were born in the prefecture (mean=590 and SD=383). "Imperial Univ. grads" is the number of individuals residing in the prefecture in 1938 who graduated from one of the Imperial Universities (mean=224 and SD=349). See Section **??** for discussions about this table.

Table A.6: Long-run Impacts: Professionals and Government Elites

| VARIABLES | Professionals: Scholars | Professionals: Physicians & Lawyers | Professionals: Engineers | Government |
|---|---|---|---|---|
| Age 17 during Centralization × Tokyo area (<100 km) | 0.88 | 0.58 | 0.79 | 1.02 |
| | (0.022)** | (0.094)* | (0.026)** | (0.010)*** |
| | [0.099]* | [0.094]* | [0.052]* | [0.025]** |
| Observations | 705 | 705 | 705 | 705 |
| Birth cohort FE | YES | YES | YES | YES |
| Birth prefecture FE | YES | YES | YES | YES |
| Mean dep var | 3.587 | 2.959 | 2.401 | 3.373 |
| Mean dep var (Tokyo area during Dapp) | 4.746 | 3.603 | 3.016 | 3.937 |

*Notes*: This table shows difference-in-differences estimates of the long-run effects of the centralized admission system. The estimates are based on the birth-prefecture-cohort level data from the JPIR in 1939, which includes cohorts who were born in 1880-1894 and turned age 17 (main application age) in 1897-1911. All regressions control for birth-prefecture fixed effects and cohort fixed effects. All outcome variables below are measured at the prefecture-cohort level. "Scholars," "Physicians & Lawyers," and "Engineer" are defined as the number of individuals whose occupation is scholar, physician or lawyer, and engineer, respectively. "Government" is the number of individuals who work at the central government either as an officer or a politician. "Age 17 during Centralization" takes 1 if the cohort turned 17 years old during 1902–1907, and takes 0 otherwise. "Mean dep var" shows the mean of the dependent variable for all prefecture-cohort observations. "Mean dep var (Tokyo area)" shows the mean of the dependent variable in the Tokyo area during decentralization. P-values based on standard errors clustered at prefecture level are in parentheses. Wild cluster bootstrap p-values based on standard errors clustered at cohort level are in square brackets. See Section **??** for discussions about this table.

Table A.7: Long-run Impacts: Difference-in-Differences Estimates excluding Cohorts who Turned Age 17 in 1901 or 1907

| VARIABLES | Imperial Univ. grads | Top 0.01% income earners | Top 0.05% income earners | Managers | Professionals | Imperial Univ. professors | Medal recipients |
|---|---|---|---|---|---|---|---|
| | | | | A. Baseline Specification | | | |
| Age 17 during Centralization | 3.05 | 0.69 | 1.47 | 3.33 | 1.63 | 0.52 | 2.65 |
| × Tokyo area (<100 km) | (0.003)*** | (0.006)*** | (0.041)** | (0.052)* | (0.008)*** | (0.081)* | (0.007)*** |
| | [0.001]*** | [0.042]** | [0.010]** | [0.006]*** | [0.001]*** | [0.153] | [0.002]*** |
| | | | | B. Adding Control Variables | | | |
| Age 17 during Centralization | 1.85 | 0.59 | 1.31 | 2.55 | 0.87 | 0.45 | 2.14 |
| × Tokyo area (<100 km) | (0.010)*** | (0.011)** | (0.010)*** | (0.010)** | (0.261) | (0.085)* | (0.000)*** |
| | [0.011]** | [0.016]** | [0.007]*** | [0.008]*** | [0.193] | [0.142] | [0.008]*** |
| Observations | 611 | 611 | 611 | 611 | 611 | 611 | 611 |
| Birth cohort FE | YES | YES | YES | YES | YES | YES | YES |
| Birth prefecture FE | YES | YES | YES | YES | YES | YES | YES |
| Mean dep var | 8.684 | 1.229 | 4.727 | 9.830 | 6.682 | 0.750 | 6.149 |
| Mean dep var (Tokyo area during Dapp) | 10.38 | 1.411 | 6.054 | 13.25 | 8.357 | 0.857 | 6.679 |

*Notes*: In this table, we repeat the same analysis in Table 5, but excluding the cohorts who turned age 17 (main application age) in 1901 or 1907 from the sample as these cohorts were exposed to both Capp and Dapp. All the variables are defined in the same way as in Table 5. P-values based on standard errors clustered at prefecture level are in parentheses. Wild cluster bootstrap p-values based on standard errors clustered at cohort level are in square brackets. See Section **??** for discussions about this table.

Table A.8: Long-run Impacts: Pre-event Trends Are Parallel

| VARIABLES | Imperial Univ. grads | Top 0.01% income earners | Top 0.05% income earners | Managers | Professionals | Imperial Univ. professors | Medal recipients |
|---|---|---|---|---|---|---|---|
| | | | | A. Baseline Specification | | | |
| Tokyo area (< 100 km) | 0.353 | -0.092 | 0.068 | 0.026 | 0.015 | 0.047 | 0.071 |
| × Time trend | (0.744) | (0.516) | (0.847) | (0.964) | (0.981) | (0.544) | (0.917) |
| | [0.250] | [0.812] | [0.562] | [0.938] | [0.938] | [0.875] | [0.750] |
| | | | | B. Adding Control Variables | | | |
| Tokyo area (< 100 km) | -0.500 | -0.044 | -0.103 | -0.184 | -0.575 | 0.036 | -0.471 |
| × Time trend | (0.144) | (0.692) | (0.565) | (0.703) | (0.018)** | (0.724) | (0.249) |
| | [0.375] | [0.500] | [0.375] | [0.562] | [0.062]* | [0.750] | [0.375] |
| Observations | 235 | 235 | 235 | 235 | 235 | 235 | 235 |
| Birth cohort FE | YES | YES | YES | YES | YES | YES | YES |
| Birth prefecture FE | YES | YES | YES | YES | YES | YES | YES |
| Mean dep var | 6.91 | 1.42 | 4.93 | 8.68 | 5.17 | 0.57 | 5.79 |
| Mean dep var (Tokyo area during Dapp) | 9.15 | 1.52 | 6.61 | 12.49 | 7.52 | 0.73 | 7.39 |

*Notes*: This table tests if there are differences in pre-event trends between urban and rural areas in the difference-in-differences analysis in Table 5. The estimates are based on the birth-prefecture-cohort level data compiled from the JPIR in 1939, which includes cohorts born in 1874-1883 who turned age 17 (main application age) in 1891-1900. This table runs the following regression:

$$Y_{pt} = \beta \times Timetrend_t \times Urban_p + \alpha_p + \alpha_t + \epsilon_{pt},$$

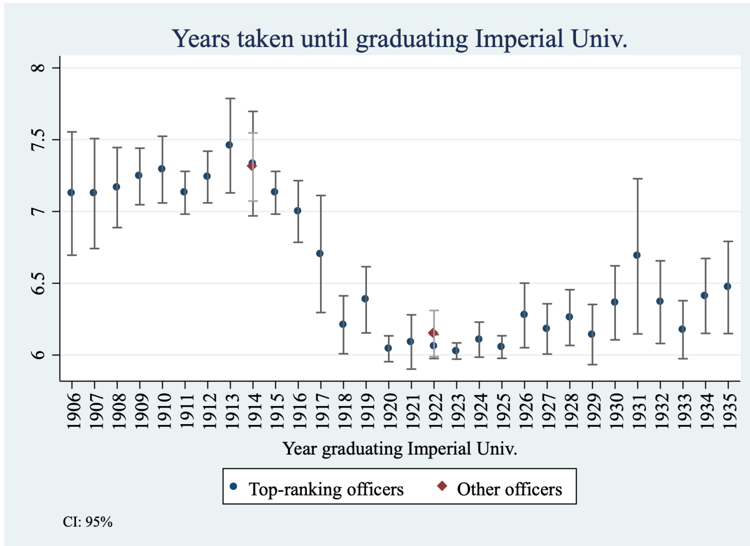
where $Timetrend_t$ is defined as the cohort's birth year minus 1870 (the linear time trend). All the other variables are defined in the same way as in Table 5. P-values based on standard errors clustered at prefecture level are in parentheses. Wild cluster bootstrap p-values based on standard errors clustered at cohort level are in square brackets. See Section **??** for discussions about this table.

Table A.9: Long-run Impacts: Using the JPIR in 1934

| VARIABLES | (1) Imperial Univ. grads | (2) Top 0.01% income earners | (3) Top 0.05% income earners | (4) Managers | (5) Professionals | (6) Imperial Univ. professors | (7) Medal recipients |
|---|---|---|---|---|---|---|---|
| Age 17 during Centralization × Tokyo area (<100 km) | 1.72 | 0.44 | 1.62 | 1.30 | 1.18 | 0.15 | 1.42 |
| | (0.011)** | (0.069)* | (0.028)** | (0.003)*** | (0.044)** | (0.386) | (0.072)* |
| | [0.001]*** | [0.002]*** | [0.004]*** | [0.008]*** | [0.018]** | [0.490] | [0.003]*** |
| Observations | 705 | 705 | 705 | 705 | 705 | 705 | 705 |
| Birth cohort FE | YES | YES | YES | YES | YES | YES | YES |
| Birth prefecture FE | YES | YES | YES | YES | YES | YES | YES |
| Mean dep var | 4.79 | 0.81 | 3.64 | 3.22 | 3.28 | 0.47 | 3.93 |
| Mean dep var (Tokyo area during Dapp) | 5.68 | 1.11 | 6.18 | 5.06 | 4.18 | 0.59 | 4.52 |

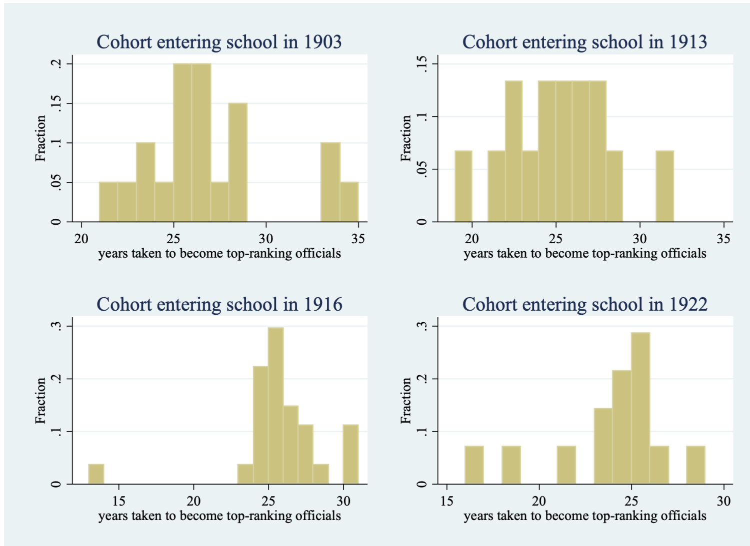
*Notes*: In this table, we repeat the same analysis as in Table 5 Panel B, but alternatively using the JPIR data published in 1934. In the 1934 JPIR, we observe the cohorts born in 1880–1894 when they are 40 to 54 years old. Sampling rates in the 1934 JPIR for top income earners are similar to those in the 1939 JPIR: 51% and 40% for top 0.01% and top 0.05% income earners. All the variables are defined in the same way as in Table 5. P-values based on standard errors clustered at prefecture level are in parentheses. Wild cluster bootstrap p-values based on standard errors clustered at cohort level are in square brackets. See Section **??** for discussions about this table.

Figure A.7: Years Taken from Entering Higher School to University Graduation




*Notes*: This figure plots the average number of years taken from entering a higher school to graduating from an imperial university by cohort (where cohort is defined by the year of university graduation) for two groups of individuals who passed the administrative division of the Higher Civil Service Exams. The first group is top-ranking officials who graduated from Schools 1–8 (shown in black round markers), and the second group is non-top-ranking officials who graduated from Schools 1–8 (shown in red diamond markers for two representative years, 1914 and 1922, only). For both groups, we find the exact year of higher school entrance for each individual using the Student Registers. The figure shows that the average number of years is not significantly different between the two groups. See Section 5.2 for discussions about this figure.

Figure A.8: Years Taken to Become Top-ranking Officials




*Notes*: This figure plots the distribution of the number of years taken from entering a higher school to becoming a top-ranking official for representative cohorts. We focus on top-ranking official who graduated from Schools 1–8 and four randomly selected cohorts (the cohorts entering a higher school in 1903, 1913, 1916, and 1922). For each top-ranking official in each cohort, we look for his biographical information by online searches to find the year of appointment to his first top-ranking position (with a success rate of 82%). The figure shows that, for all cohorts, there is a reasonably large within-cohort variation in the number of years taken to be promoted to a top-ranking position. See Section 5.2 for discussions about this figure.

Table A.10: Long-run Impacts on Top-Ranking Officials: Alternative Estimation of the Year of Entering Higher School or its Equivalent

| VARIABLES | (1) Top-ranking officials | (2) Top-ranking officials graduated from School 1 | (3) Top-ranking officials graduated from Schools 2-8 | (4) Top-ranking officials not graduated from Schools 1-8 | (5) Top-ranking officials graduated from Schools 2-8 | (6) Top-ranking officials graduated from Schools 2-8 and Tokyo Imperial Univ. |
|---|---|---|---|---|---|---|
| Centralized | 4.15*** | -0.54 | 5.05*** | -0.36 | 2.69*** | 1.76** |
| | (1.28) | (1.02) | (1.31) | (0.94) | (0.68) | (0.81) |
| Higher Civil Service Exam passers | 0.10*** | 0.02** | 0.06*** | 0.03*** | | |
| | (0.02) | (0.01) | (0.01) | (0.01) | | |
| Exam passers graduated from Schools 2-8 | | | | | 0.17*** | |
| | | | | | (0.02) | |
| Exam passers graduated from Schools 2-8 and Tokyo Imperial Univ. | | | | | | 0.19*** |
| | | | | | | (0.02) |
| Observations | 33 | 33 | 33 | 33 | 33 | 33 |
| Mean dep var | 29.77 | 8.273 | 12.97 | 8.528 | 12.97 | 11.82 |
| Mean dep var (Dapp) | 28.57 | 8.304 | 11.22 | 9.051 | 11.22 | 10 |

*Notes*: This table shows OLS estimates of the long-run effects of the centralized admission system on the number of top-ranking government officials. Definitions of the variables and the specifications are the same as in Table 8, except that the computation of the year of entering a higher school or its equivalent for individuals who did not graduate from Schools 1–8 was changed from "year of graduating the final education - 6" to "year of graduating the final education - 7". The results are qualitatively the same as the results in Table 8. See Section 5.2 for discussions about this table.