ARE REGIONAL QUOTAS FAIR? SIMULATING MERIT-BASED COLLEGE ADMISSIONS USING UNIQUE STUDENT-LEVEL DATA FROM CHINA

Tianshu Guo

Prashant Loyalka, Stanford University

Xiaoyang Ye, University of Michigan

February 2018

Abstract

This paper examines whether and to what degree the existence of regional quotas creates merit-based inequalities in college admissions in China. Using unique student-level data from China, we estimate the degree to which a student's region of origin affects his or her college admissions outcomes; and simulate how the regional composition of student body would change at colleges and elite colleges if admissions were solely based on merit. The study finds that (1) under the current regional quota system, college admissions opportunities are strongly biased towards students from more developed regions; (2) switching to a solely merit-based admissions process would significantly change the current regional student body composition; (3) analyses controlling for non-academic skills are highly consistent with analyses controlling for academic skills; and (4) a major contributing factor of the existing regional inequality is that four-year college and elite college distribution in China are highly correlated with economic development.

Keywords

Higher education; Regional quota; College opportunity; Regional inequality; College distribution

1 Introduction

A relatively new and unexplored topic in market design is the imposition of regional quotas in college admissions. When regional quotas are imposed, stable and fair matches that existed before, may no longer exist (Goto et al., 2016, 2017; Kamada & Kojima, 2015, 2017). This is consistent with public opinion as well as descriptive studies that claim that regional quotas for college admissions are not merit-based and are unfair (Espenshade et al, 2004; Ayalon & Yogev, 2005; Fu, 2013).

Debates about the efficiency and equity of regional quotas in college admissions are especially fierce developing countries such as in China (Li, 2010; Liu et al., 2014). A constant refrain from families and the Chinese media is that students from developed regions like Beijing and Shanghai are assigned more spots in colleges and elite colleges than students from less developed regions, even after conditioning on academic ability (Yang, 2006). On the surface, the claim appears to be valid as the central government assigns a higher proportion of college and elite college seats to more developed regions than less developed regions. The central government, however, argues that developed regions deserve more spots precisely because their students (because of higher levels of household wealth and access to quality schooling) are more prepared academically by the time they apply for college (Liu et al., 2014).

Unfortunately, the justifications that policymakers in China give for assigning regional quotas have not been carefully explored. Each region or province in China has its own college entrance exam (which is the sole measure of a student's qualifications during the application and admissions process), and each region separately determines the content, administration, and grading of the college entrance exam (Chen & Kesten, 2017). Absent a universal standard by which student ability can be compared, it is impossible to compare the qualifications of students from different regions and thereby determine whether regional quotas in college admissions are indeed merit-based and equitable.

In this paper, we examine whether and to what degree the existence of regional quotas creates merit-based inequalities in college admissions in China. To fulfill this purpose, we (a) estimate the degree to which a student's region of origin affects his or her admissions outcomes; and (b) simulate how the regional composition of the student body would change at colleges and elite colleges if admissions were solely based on merit. We conduct a series of parametric and semi-parametric models to control for the standardized measures of universal ability.

We analyze unique data that we collected on a national random sample of close to 2,000 college freshmen from 36 universities in six provinces in China. The data include students' college entrance exam scores, that are particular to a province, as well as a second set of student achievement scores on essentially concurrent standardized assessments that are universal across provinces.

Our findings are stark. First, we find that admissions opportunities are heavily biased towards students from more developed regions. Holding academic achievement constant, a student from Beijing can attend a college that is ranked 100 places (out of 400 tier 1 and tier 2 public colleges) higher than a student from Shaanxi and more than 80 places higher than a student from Sichuan or Henan. Second, according to simulations, switching to a merit-based admissions process from the one based on regional quotas would markedly change the regional balance in admissions outcomes. In solely merit-based admissions, the regional compositions for students from Henan would be more than doubled compared to current levels and increase by 60% for students from Sichuan; the composition for Beijing students would shrink by 40%; and those for students from Guangdong and Shaanxi would be significantly lower to below 30% and

10% of current levels. Third, analyses on critical thinking skills find similar results as analyses controlling for academic skills. This means that a major reason for the government assigning unequal regional quotas outside of academic merit, which is that some other types of non-academic skills are higher for students from developed areas such as Beijing, does not hold either. Finally, we suggest that a major contributing factor of the current regional merit-based admissions inequalities is that college and elite college distributions in China are unbalanced. Four-year college and elite college distributions highly concentrate in provinces with higher levels of economic development, and these provinces may vie for higher quotas for their students.

The rest of the paper proceeds as follows. Section 2 introduces the background on elite selections through examinations in China and the current college admissions system. Section 3 introduces the analysis data and quantitative analysis strategies. In Section 4, we summarize all analyses results with robustness checks and extension analysis, and we discuss about potential limitations the study has. In Section 5, we link the regional inequality issue to the broader topic of college distributions. Section 6 concludes.

2 Background

2.1 The Examination System in China and Reforms of the Gaokao System

China has a long history of selecting candidates for the state bureaucracy and other positions through examinations. The imperial examination system in Imperial China, called *Keju*, has been established in the Sui dynasty and broadly expanded since the mid-Tang dynasty, with the purpose of evaluating and selecting candidates based on merit. As the oldest high stakes examination system, Keju has been the major path to becoming civil officials for people without family or political connections for more than 1200 years (Elman, 2009).

Gaokao, the National Higher Education Entrance Examination in modern China, is the largest high stakes examination system in the world. It is an extension of the Keju system in Imperial China and is used now to select the best potential candidates for higher education in China. As with the goal of the Keju system, Gaokao is designed to test and select participants for merit. Gaokao was created in 1952, canceled in 1966 for the ten years of the Cultural Revolution, and resumed in 1977 (China Examinations, 2006).

The first Gaokao after the Cultural Revolution took place in 1977. Due to the lack of a unified high school curriculum at that time, each region (province, municipality, and autonomous region) in China separately determined the content, administration, and grading of the examination, as well as the final college admission process (China Examinations, 2006). In the next year, the Ministry of Education (MOE) designed and published the official guidelines and review materials for Gaokao. From 1978 and until 1985, the Gaokao examination questions were created by the MOE, and all regions in China used the same set of questions (China Examinations, 2006). Gaokao were hold annually, and the test dates were fixed and the same across all regions.

In 1980, Heilongjiang and 7 other provinces implemented the pre-Gaokao screening examination, which selected relatively better performing students, as many as 3 times to 5 times of the total college capacity for that year, from all potential college applicants. This pre-Gaokao selection process reduced the workload for Gaokao administration and grading, and 5 more provinces implemented this process in 1981 (China Examinations, 2006).

In 1982, the MOE set up rules for Gaokao bonus points. Extra points were given to students with exceptional performance in high school, student athletes, applicants with extraordinary accomplishments in arts and other areas, and other applicants with certain backgrounds or special talent. The MOE also allowed higher education institutions to lower the admission cutoff scores for minority students and students applying to certain major such as agriculture, forestry, and geology (China Examinations, 2006).

In 1983, the MOE proposed the directional recruitment program, which combined college admissions and post-baccalaureate career choices. Students applying for the program might only choose among limited majors and had to work in specific industries after graduation, but might be admitted on a lower score (China Examinations, 2006). In 1984, the MOE started to allow direct admission of outstanding students. Students through direct admissions would not have to take the Gaokao to enter colleges (China Examinations, 2006).

In 1985, China had a big reform of the Gaokao system. The MOE re-designed Gaokao and made it a standardized test for college admissions, ensuring that Gaokao questions, administrations, grading, and interpretations were all designed and performed in a standard manner. The newly designed Gaokao was implemented in Guangdong in 1985 and more provinces in later years (China Examinations, 2006). In 1989, the standardized English subject examination was used in all regions in China.

At the same time as Gaokao became standardized, some regions started to separately determine the content of Gaokao based on their own high school curriculums. In 1988, 1999 and 2002, Shanghai, Guangdong and Beijing started to use their own Gaokao questions, respectively, and nine more regions followed later in 2004 (China Examinations, 2006). In terms of Gaokao content, started in 1994, the number of subjects being tested in Gaokao reduced from seven subjects to five with three required and two self-chosen subjects, and finally took the form of "3+X" with three compulsory subjects of Chinese literature, Mathematics, and English language, and one additional examination either in social-science-oriented area or natural-science-oriented area based on students' preferences. In 2002, the "3+X" system was implemented in all regions in China (China Examinations, 2006).

In addition to improvements of the Gaokao examination itself, the MOE took other actions to better the Gaokao system. For example, the MOE started to offer financial aids and students loans in 1993 to help students in need. In addition, started in 1999, Beijing began to operate the spring Gaokao. Another important improvement in the Gaokao system is the implementation of the Independent Freshman Admission Program (IFAP). The IFAP allowed higher education institutions to give priority admission to qualified students, up to 5% of the targeted freshman admission quota. From 2003 to 2006, the number of institutions participated in the IFAP increased from 22 to 53 (China Examinations, 2006).

2.2 Regional Quota and Region Inequality in the Higher Education System in China

The Gaokao system has a critical role in determining economic opportunity and social mobility in China. As a result, the fairness of the current system has always been the focus of policymakers in education as well as the public. Among all issues, lots of attention and complaints are on regional quotas, which is to distribute different quotas in different regions, in the Gaokao system.

Regional Quota in the Keju System

The regional quotas did not originate from the Gaokao system. Rather, it already existed in the Keju system in Imperial China. In the Ming dynasty, quotas were distributed to the southern, northern, and central regions at the ratios of 55:35:10 (Shen, 2012). In the Qing dynasty, the state bureaucracy regulated that the Keju recruitment quotas within each region should be determined based on the total number of applicants in the region. In addition, from the Song dynasty to the Qing dynasty, the quota for the capital city has always been the highest among all regions (Shen, 2012).

College Admission Planning in the Gaokao System

Regional quotas in the current Gaokao system are determined jointly by the MOE, the National Development and Reform Commission (NDRC), provincial education departments, and provincial development and reform commissions.

Gaokao takes place every year in June. In September of the previous year, the MOE coordinates with the NDRC and designs the general guidelines for college admissions for the next year. Provincial education departments and development and reform commissions report the provincial college admission plans to the MOE by the end of December. Provincial college admission plans should be based on the ratio of planned versus true college enrollments in the region from last year, provincial population growth, educational investments, college qualities, and other criteria (Kang, 2005). Then, the MOE and NDRC create the draft of a more detailed college admissions plan for the coming Gaokao based on provincial plans. In March of the year of Gaokao, the detailed plan is assessed by the National People's Congress (NPC) and finalized in April. By the end of May, higher education institutions should finalize their own college admissions plans for Gaokao in June (Kang, 2005).

Regional Quota and Regional Inequality in the Gaokao System

In terms of college admission quotas assigned to each region, the MOE stated in the 1977 general guidelines of college admission that higher quotas may be distributed to areas with more applicants. Later in 1987, the MOE recommends that higher quotas may be assigned to regions with more applicants and regions with applicants of higher academic ability. In later years, the MOE further suggests higher levels of quota to be distributed to central and western regions in China where educational resources are relatively limited (The Ministry of Education, 1977, 1987, 2006).

The regional quotas policy in the current Gaokao system has always been questioned for that it leads to unfairness and regional inequalities in college admissions. A constant refrain from families and the Chinese media is that students from developed regions like Beijing and Shanghai are assigned more spots in colleges and elite colleges than students from less developed regions, even after conditioning on academic ability (Yang, 2006). According to Li (2007), regional differentials in college admission cutoff scores are huge. For example, in 1999, the cutoff score for elite colleges for social-science-oriented students was 466 for Beijing students and 556 for Hunan students. The cutoff scores across different regions were as large as more than 100 Gaokao points. In addition, college enrollment rates also vary a lot across different regions. As stated by Wang (2006), in 2003, the college enrollment rates in developed regions were around 50%, while those in some provinces in western China were only at around 11% - 12%.

These huge admission score and college opportunity differentials across regions cause a series of social issues, such as the Gaokao migration (Fan, 2006; Feng, 2006; Liu et al., 2006; Zhao, 2006). Students from areas with low college admission quotas transfer to developed regions in the year of their Gaokao in order to gain better college opportunities.

Arguments for and Against Regional Quota

As argued by many researchers, as well as the public, regional quotas make the current Gaokao system unfair (Feng, 2006; Li, 2007; Wang, 2006; Yang, 2007; Zhao, 2006). Students from more developed regions have more and better college opportunities. Many researchers suggest that college admission cutoff scores should be the same across all regions (Li, 2007; Liu et al., 2006; Zhao, 2006).

On the other hand, others argue that developed regions deserve more spots precisely because their students are more prepared academically by the time they apply for college (Liu et al., 2014; Yan, 2006). As commented by college admission directors at Peking University and Tsinghua University, students from more developed regions generally have higher abilities that are not tested in Gaokao but important in college studies, such as in oral English skills, computer science related skills, and social skills (Yan, 2006).

In addition, another reason for the existence of regional quotas in the current system is that colleges aim to have higher employment rates for their graduates. Since students from more developed regions may gain job opportunities more easily in these regions, colleges have an incentive to admit more students from these regions for higher alumni employment rates after graduation. (Yan, 2006).

Moreover, provincial governments tend to give admission priorities and more post-baccalaureate career opportunities to local students (Li, 2007; Song, 2007). Students from less developed regions and attending colleges in a more developed region are likely to stay and work in the more developed region after graduation. Provincial governments may not want too many career opportunities to be taken by these students from other provinces (Li, 2007; Song, 2007). As a result, colleges in a province may admit more students within the same province. Since top colleges relatively concentrate in more economically developed regions such as Beijing and Shanghai, local students from these areas are more likely to be admitted by these elite colleges in their provinces. Therefore, the existing inequality favoring students from more developed regions could be a result from regional inequalities in educational investment as well as college quantity and quality.

Gaokao Comparability Across Regions

Although many studies and news articles focus on the Gaokao admission cutoff score differences across regions, one thing they ignore is the comparability of the Gaokao examinations across regions. In other words, the same Gaokao score in different regions may indicate different levels of academic ability of students. Especially as more regions now independently design their own Gaokao questions, absolute Gaokao scores are hardly comparable across regions. Therefore, a 100-point higher admission cutoff score in a region does not necessarily suggest that the marginal student being admitted from that region has higher academic ability even though with higher Gaokao scores. Lower Gaokao scores in a region could be due to more difficult exam questions, stricter grading standards for the region, or other reasons. Absent a universal standard by which students' academic achievements can be compared, it is impossible to assess and compare the qualifications of students from different regions and thereby determine whether regional quotas in college admissions are indeed merit-based and equitable.

Behind Regional Inequality

Studies on higher education in China all agree that educational opportunities in China increase substantially over time at all levels (Fang and Feng, 2005; Li, 2010; Wu and Zhang, 2010; Yang, 2006). However, studies also suggest geographical and other types of inequalities in students' access to higher education (Fang and Feng, 2005; Li, 2010; Li 2017; Liu, 2015; Liu et al., 2014; Liu and Zhang, 2006; Wu and Zhang, 2010; Treiman, 2013; Yang, 2006; Yeung, 2013). For example, Liu (2015) examines the geographical stratification in access to higher education in China. The study finds that the current decentralized governance favors students in the eastern area and limits the opportunities of students from western and central provinces. In addition, Li (2017) and Liu and Zhang (2006) argue that college admissions in China are subject to hometown-bias, which widens the educational opportunities inequalities between rich and poor regions.

3 Empirical Strategy

In the study, we examine the existence and degree of merit-based inequalities among provinces under the current regional quotas system in college admissions in China.

3.1 Data and Universal Academic Ability Measures

We collect and analyze unique data on a national random sample of 9,286 college students from 36 universities in the six provinces of China – Beijing, Sichuan, Shandong, Guangdong, Henan, and Shaanxi. The data includes students' total and sectional scores in the college entrance exam, performances on the Super-Tests, the province, city, and county at which students took the college entrance exam, the colleges and majors that students enrolled in, as well as students' demographic information and family background measures. The college entrance exam scores are particular to the province where students took the exam, while the Super-Tests scores are a measure of students' academic achievement that are universal across all provinces.

The sample contains 5,140 college freshmen who took the college entrance exam in the year 2015 and 4,146 college students in their junior year who took the exam in the year 2013. For our main analyses on regional inequalities in college admissions in China, we focus on first-year college students and the relationship between their college entrance exam scores and Super-

Tests performances. As introduced earlier, in the college entrance exam, students choose to participate in either the social-science-oriented (i.e. Wen-Ke) exam or the natural-science-oriented (i.e. Li-Ke) exam. Since more than 97.5% of the college freshmen sample (4,941 out of 5,140) are in natural-science-oriented area, we particularly focus on this group of students in our analyses.

Students' performances on three sections of the Super-Tests are analyzed, including the Mathematics section, the Physics section, and the Critical Thinking section. As our analyses focus specifically on natural-science-oriented students, we use the mathematics and physics scores as measures of students' academic ability on mathematics and science in our main analyses and the critical thinking scores as robustness check. There are 2,445 first-year natural-science-oriented student observations with valid Super-Tests scores in both the mathematics and the physics sections, as well as a valid total score on the college entrance exam.

Finally, in our analyses, we focus on college admissions and college opportunity inequalities among students who were originally from the above six provinces of interest upon high school graduation. Therefore, only students who participated in the college entrance exam in the six provinces are included in the analyses. The final analysis sample for our main analyses on mathematics and science ability consists of 1,825 observations, and the sample for the critical thinking skills consists of 900 observations.

3.2 Econometrics

The analytic approach in this paper is to use our unique data to link students' province specific college entrance exam scores and their universally comparable Super-Tests performances. By establishing this relationship, we make students in different provinces comparable in terms of their academic ability. With this standardized measure of universal ability, we examine whether and to what degree the existence of regional quotas creates merit-based inequalities in college admissions in China.

Relationship Between Gaokao Score and Percentile Rank Across Provinces

First, we examine the basic relationship between students' college entrance exam scores and their provincial percentile rank across different provinces. We use both the Ordinary Least Squares (OLS) model and a partial linear model in our analyses, with province fixed effects. The basic specification for the OLS regression in our analyses is as below:

$$Y_i = \alpha_0 + \alpha_1 Math_i + \alpha_2 Physics_i + \sum_{j=1}^{5} \beta_j \, Province_{ij} + \epsilon_i$$

where Y_i represents the outcome variable of student *i*, which, in this analysis, is the percentile rank of student *i* in the province where the student took the exam. $Math_i$ and $Physics_i$ represent student *i*'s Super-Tests scores on the mathematics section and physics sections, respectively. $Province_{ij}$ are the province dummy variables for each student *i*, indicating province fixed effects in the model. Since there are six provinces of interest, we take Beijing as the control province and include the other five provinces – Sichuan, Shandong, Guangdong, Henan, and Shaanxi – in the model as dummy variables. The coefficients of interests in all analyses are the five β_j's, indicating differences across provinces.

Regional Inequality on Elite College Opportunities

Our first main research question is to estimate regional inequality, if any, on the chances of getting into an elite college for students with the same academic ability but taking the college entrance exam in different provinces. We simply categorize all the 36 colleges in our sample as either elite or non-elite colleges based on their overall quality and ranking. Using both the logit and probit models with province fixed effects, we estimate regional differences in terms of the probability of a student in a province getting into an elite college. In other words, we examine whether living in a particular province advantage or disadvantage students with same academic abilities in terms of elite college opportunities. We apply the same specification as above, where we have Y_i indicate whether student *i* is currently enrolled in an elite college. Again, Beijing is left out as the control group.

Regional Inequality on College Quality

Next, we examine whether there exist regional differences in the best college students could have gotten into with their scores in their provinces. Even though college quality is a broad and relatively subjective term, here we use the national rank of each college as an approximation for the college's overall quality. Therefore, this analysis compares the regional differences in the highest possible college ranking a student in a particular province could have attained. In other words, we estimate the regional inequality on the highest possible college ranking or best possible college quality for students with similar academic ability but taking the college entrance exam in different provinces.

In the regression specification for this analysis, Y_i represents the highest possible college ranking that student *i* could have attained with his or her college entrance exam score in his or her province. Both linear and partial linear models are used, with province fixed effects comparing all other provinces to Beijing.

Merit-Based Regional Student Body Compositions

Finally, the last part of our main analysis is a simulation on the regional compositions of freshmen student body at colleges and elite colleges under the ideal situation where college admissions are solely based on merit. There would be no merit-based regional inequalities of any kind in this ideal situation. In this analysis, instead of looking only at our sample data, we consider the whole populations participating in the college entrance exam in the six provinces and estimate academic ability for the six exam populations.

To accomplish the simulation for ideal situation, we first separately estimate the relationship between students' the college entrance exam scores and Super-Tests scores in each of the six provinces. Then, based on these relationships, we estimate the academic ability for every student in the six provinces based on their college entrance exam performances. With the estimated universal measure on ability, we rank every student in the six provinces on ability from high to low and assign them into colleges based on their relative positions among all students and college capacities.

4 Results

4.1 Main Analysis

In the main analysis, we control for students' academic ability on mathematics and science, which are measured by their scores on the mathematics and physics sections in the Super-Tests, and compare the differences on multiple college admission outcome variables across provinces. In all regressions, we leave out Beijing as the control group and include the other five province dummy variables – Sichuan, Shandong, Guangdong, Henan, and Shaanxi – in the formula.

Relationship Between Gaokao Score and Percentile Rank Across Provinces

Table 2 Column 1 and 2 summarize the results for regional percentile rank differences when controlling for mathematics and science ability and province fixed effects on the 1,825 valid observations under OLS and partial linear model, respectively. As expected, the coefficients on Super-Tests scores are both positive and statistically significant, indicating that higher academic ability are associated with higher percentile rank in each province.

The coefficients of interest are the five province dummy coefficients that indicate the existence and levels of regional differences across provinces. In the OLS regression, four out of the five province dummy coefficients are statistically significant, indicating that for students with the same academic ability, their percentile ranks would be statistically different if they take the college entrance exam in different provinces. The partial linear model gives similar results with three province dummy coefficients being statistically significant. The results suggest that not only the absolute college entrance exam scores but also province-wide percentile ranks are not comparable across provinces. Furthermore, the results also imply potential differences on student ability distributions across provinces.

Regional Inequality on Elite College Opportunities

Table 3 Column 1 and Table 4 Column 1 report the regression results for elite college opportunities with ability controls and province fixed effects under the logit model and profit model, respectively. The two models give similar predictions. Again, as expected, both of the academic ability measures are positive and statistically significant in the regressions, which suggest that students with higher academic ability have higher chances of getting into elite colleges.

For coefficients on province dummy variables, under both the logit and probit models, all five coefficients are negative. In the logit model, four out of five coefficients are statistically significant, except for the one on province Shaanxi. In the probit model, all five coefficients are statistically significant. The negative and statistically significant province dummy coefficients suggest that compared to students in Beijing, students with the same academic ability in other provinces have statistically significantly lower chances of getting into elite colleges in the current system.

This analysis has important implications. The results prove the existence of significant regional inequality in terms of elite college opportunities in the current college admissions system in China. Among the six provinces included in our analysis, students from Beijing have huge and statistically significant advantages over students from other areas, even after controlling for academic ability.

Regional Inequality on College Quality

Table 5 Column 1 and 2 describe the regression results for regional differences on the highest possible college ranks students could have attained, under linear and partial linear models. The coefficients on Super-Tests scores are both negative and statistically significant, meaning that students with higher academic ability will be able to get into colleges with better ranks.

The coefficients on province dummy variables represent the regional differences on the highest possible college ranks that students with same academic achievement in different provinces could have attained. Based on the results under the OLS model, when holding academic achievement constant, a student from Beijing can attend a college that is ranked nearly 100 places higher than a student from Shaanxi, about 85 places higher than a student from Sichuan or Henan, and more than 30 places higher than a student from Shandong. The regression shows huge and significant regional inequalities favoring students from Beijing. The partial linear model generally agrees with the OLS model besides that it shows that students from Guangdong benefit even more under the current unfair system. Both the linear and partial linear models suggest that students from Sichuan, Henan, and Shaanxi are significantly disadvantaged under the current system.

Consistent with the previous analysis results on elite college opportunities, this analysis also confirms that regional unfairness exists and is huge under the current regional quota policy and college admissions system. Specifically, students from Beijing in general have significant advantages in college admissions. Students from Beijing are able to go to a college with much higher national rank compared to students with the same academic ability from Sichuan, Henan, or Shaanxi.

Merit-Based Regional Student Body Compositions

Table 6 shows simulation results for the ideal scenario of solely merit-based admissions without restrictions by regional quotas. According to simulations, in the ideal situation with no regional college admissions inequality, the regional balance in admissions outcomes would change remarkably. Specifically, under solely merit-based admissions, the regional composition for students from Henan would be more than doubled compared to current level, and that for students from Sichuan would increase by around 60%. On the other hand, the composition for students from Beijing would shrink by nearly 40% compared to current level, and those for students from Guangdong and Shaanxi would be significantly lower to around 35% and below 10% of current levels, respectively.

The simulation results also confirm the current advantage of Beijing students and the disadvantage of students from Sichuan and Henan under current regional-quota-based admissions system. However, another thing to observe from the ideal case regional composition is the percentage of the provincial exam population that are admitted. The number of students taking the college entrance exam in different provinces can be very different. For example, in year 2015, there are 38,387 students taking the exam in Beijing, and in Henan, there are 365,686 students, which is almost ten times the Beijing exam population. Considering in terms of percentage of the exam population in each province, Beijing has the highest percentage that would be admitted by top 400 colleges even under complete merit-based admissions. The percentages for Sichuan, Shandong, and Henan would all be increased compared to current levels, while those for Guangdong and Shaanxi would be lower than current levels. From the percentage perspective, the high percentage of Beijing in the no-inequality situation implies that academic ability of Beijing students might be higher in general, at least among students competing for top universities.

In conclusion, the current disproportion of regional college admission quotas results from a combination of both systematic unfairness and potential existing ability inequality on secondary education across provinces. Earlier analyses show that regional inequality exists such that Beijing students have higher chances of getting into an elite university or universities of higher rank. The simulation also confirms the current advantage of Beijing students, but at the same time shows that Beijing students getting high quotas is at least partially because Beijing students have comparatively higher academic ability on average, at least among top students. However, even after taking into consideration the potential ability differences across provinces, there still exists huge systematic regional unfairness in the current college admissions system in China, which requires attention by policy makers.

4.2 Robustness Checks

Critical Thinking as Academic Ability Measures

Both mathematics and physics are directly tested, in separate sections, on the college entrance exam for natural-science-oriented students. Therefore, academic ability on these two subjects are highly correlated with students' college entrance exam performances, at least on the two relevant sections, and potentially academic achievement in college as well. However, there are other ability that are important in college studies but are not directly tested on the college entrance exam. For example, critical thinking skill is an essential in postsecondary studies. As robustness check, we use the critical thinking section on the Super-Tests as non-academic ability measures and repeat all the regressions.

The sample of students with valid college entrance exam scores and Super-Test critical thinking scores contains 900 observations. Table 7 summarizes regression results for the highest possible college rank analysis using the critical thinking sample. The results for this analysis and other repeated analyses, when controlling for critical thinking skills, are highly consistent with our main analyses results, which controls for academic ability on mathematics and science, both in terms of coefficient magnitudes and the levels of statistical significance. The only difference is that in the regressions for elite college opportunities, the disadvantages of students from Sichuan and Guangdong are no longer statistically significant. Shandong and Henan are the two provinces where students are consistently disadvantaged in terms of elite college opportunities when either academic skills or non-academic skills are controlled. The high consistency between the two sets of regressions also confirm the existence of significant regional inequalities in college admissions under the current system. In addition, it suggests that the government's argument that students from more developed provinces have higher non-academic skills may not be valid. In other words, the reason that government gives for the existing regional quota differences may not hold.

Sub Sample Analyses

For each of our analyses, we run regressions using the selected sample described in the earlier section on data, containing 1,825 students for mathematics and science analyses and 900 students for the critical thinking analysis. In addition, we repeat all analyses using three other sets of subsamples. Table 1 summarizes the selection criteria for all four samples used in the analyses.

As described in Table 2, Sample 1 is the original sample. Sample 2 contains only student observations whose college admissions decisions are purely based on their performances on the college entrance exam. As introduced earlier, students may get into colleges through other channels such as IFAP, direct admissions (i.e. Bao-Song), targeted admissions (i.e. Ding-Xiang Zhao-Sheng), and student athletes. Students being admitted through these alternative programs may generally have lower exam scores compared to normal admission standards, or their exam scores and Super-Tests scores may have different relationships. In sample 2, we exclude these observations and include only students admitted purely through the college entrance exam to better capture elite college opportunities and the relationship between provincial exam scores and our universal Super-Tests scores. Sample 2 contains 1,789 observations for the mathematics and science analyses and 883 observations for the critical thinking sample.

Sample 3 and sample 4 are designed to minimize the impact of potential measurement and input errors in the data sample. As mentioned earlier, the sample data contains students' total scores on the college entrance exam as well as their separate sectional scores on the exam. If all data was reported and input correctly, the separate sectional scores should sum up to the total exam score for each student. However, due to potential measurement error, input error, or survey bias discouraging students reporting their true exam performances, there are about 400 observations whose sectional exam scores do not sum up to their total scores, with score differences ranging from 0.5 point to hundreds of points. To minimize the impact of measurement errors and possible bias, we include only observations with less than 50 points in absolute score differences between reported total exam scores and summed up sectional scores. This leads to a sample of 1,735 observations for the mathematics and science sample and a sample of 854 observations for the critical thinking in sample 3. Sample 4 is selected by applying the same score-difference restrictions on sample 2, leaving 1,706 and 840 observations for the two sets of analyses, respectively.

In general, regression results on these sub samples are highly consistent with sample 1, and in some cases, regional inequalities are more statistically significant under sub sample analyses.

4.3 Extension – Regional Quota Equalization by Proportion

Both regression analyses and the simulation results show that the current college admissions system in China, and the regional quota system specifically, are associated with significant regional inequalities. However, analyses results do not suggest alternative quota distribution policies that may potentially lead to improvements. As an extension, we test the effectiveness of a simple quota distribution model which distributes college admissions quotas to each province purely based on and proportional to the size of the exam populations within each province.

Table 11 gives comparisons on regional admission cutoff scores, regional admission cutoff percentile ranks, and average regional academic ability estimates under the current quota distribution scenario, the no-inequality scenario based on simulation, and the proportional regional quota distribution scenario.

As expected, under the no-inequality scenario from the simulation, there is no significant ability differences between provinces. Specifically, the maximum ability difference between provinces is less than 0.5 point in terms of ability measures. Under the current regional quota system, the maximum ability difference among the six provinces is 13.6 points, with the highest ability province being Henan with an average academic ability of 48.9 points and the lowest ability province being Shaanxi with an average academic ability of 35.3 points. Under the proposed proportional regional quota distribution, the ability difference decreases by about 2 points and improves to 11.8 points. The province with the highest estimated academic ability in this case is Beijing, with an ability level of 48.0 points. The province with the lowest ability is again Shaanxi, which has 36.2 points in ability measures.

This simply comparison gives several implications. First, the current merit-inequality between provinces could be improved, even though not by a huge degree, by simply assigning regional quotas proportionally to each province based on the size of exam populations each year. Second, if quotas were to distributed proportionally, there would still exist merit-inequality between province, but to a lower degree. However, these new regional inequalities no longer give advantages to students from Beijing. On the contrary, Beijing students would be disadvantaged and admitted with a higher standard compared to students from other provinces under the proportional quota system. This again implies that students from Beijing may in general have higher academic achievement by high school graduation. Therefore, in a completely fair system, it may still be true that comparatively more quota, in proportional terms, should be assigned to students from Beijing, but definitely not as high as current quota level.

4.4 Limitations

This study provides valuable new evidence proving the existence of regional inequalities in college admissions in China, but at the same time has some limitations mainly due to the data being used. First, we use Gaokao score distribution tables published by each province every year after Gaokao grading. The score distribution tables provide detailed information on Gaokao scores and the number of people at each score in the province. We use this information to calculate the percentile rank corresponding to each score. Provincial Gaokao populations are also obtained based on these data. However, since score distributions are separately published by each province, there are some inconsistency in the score format between provinces. For example, Guangdong arranges all scores into 5-point score ranges (i.e. 550-554 points; 555-559 points), while most of the other provinces use 1-point per category. Using score ranges as categories may lead to inaccuracy when calculating the percentile ranks and estimated academic ability corresponding to a specific score. In addition, more students are included in each category when score ranges are used. This may cause inaccuracy in the simulation results, because when ranking ability and selecting students under simulation, most likely a whole category will either be all admitted or all rejected in the simulated scenario. However, if each category includes too many students, ranking and selections results may not be accurate.

Second, in one of our main analyses, we calculate the highest possible college ranks that students could have attained, where we use the national ranking of top 400 colleges in China and the admission cutoff scores of each college. However, in the national ranking college lists, there are some colleges that are specialized in certain areas. For example, the medical school of Peking University ranks the 12 among all colleges, but specializes only in medical science. This and similar colleges with potentially limited major selections usually have much lower admission scores compared to colleges with similar ranks, which may lead to inaccuracy for the highest possible ranks to be calculated.

Finally, the study may be potentially limited by the relatively small regression sample size for some provinces. When relating Gaokao scores to academic ability measures, we separately determine the relationship in each province. Most of the provinces have at least more than 200 observations, but Beijing has only less than 60 observations in the sample. This is not necessarily a sample selection issue, but rather due to the small Gaokao population size in Beijing. This limited sample size for Beijing may cause the estimated ability measures biasing towards the selected sample and create inaccuracy in ability comparisons.

5 Discussion

In the study, we confirm that the regional inequalities exist and are huge in college admissions in China. In addition, we take one step further to explore potential factors contributing to these inequalities beyond the regional quota policy. As argued by earlier literature, college admissions in China are affected by hometown-bias, which favors in-province students compared to students coming from other provinces. As there has always been arguments stating that colleges in China are disproportionally distributed and concentrate in more economically developed regions, we examine and compare the relationship between economic development and college distributions in China and in the U.S.

Data

In comparing college and elite college distributions in the U.S. and in China, we use school-level data for the two countries. For U.S. college distribution, we analyze the data from the Integrated Postsecondary Education Data System (IEPDS) on all 7,537 U.S. colleges reported for the year 2016. For the college distribution in China, the main data source we use is the official complete college list published by the Ministry of Education (MOE) for the year 2017, containing all 2,914 governmental-approved higher education institutions in mainland China. GDP and regional population data are based on relevant statistics by the Bureau of Economic Analysis and the National Bureau of Statistics of China for the U.S. and China, respectively.

Comparison Results

Table 12 summarizes the results for college and elite college distributions and their relationships to economic development, measured by GDP per capita. We separately compare the distributions of all colleges, four-year colleges, and elite colleges in the U.S. and in China.

In the U.S., the number of colleges and elite colleges in a state is not related to the level of economic development in that state. In China, the total number of higher education institutions in a province is not statistically significantly related to GDPs. However, there is a significant positive correlation between the number of colleges offering Bachelor’s degrees (i.e. Ben-Ke) in a province and the province’s GDP per capita. There is even a stronger and statistically significantly relationship between the number of elite colleges, ranking top 100 among all colleges in China, and GDP levels.

Table 13 adds state or provincial populations in addition to GDP per capita. In the U.S., the number of colleges and elite colleges increases with the state’s population and the impact of population on the number of colleges are consistently statistically significant. In China, the total number of higher education institutions and the number of colleges offering Bachelor’s degrees in a province increase with both GDP and population. However, the number of elite colleges in a province is statistically significantly related to GDP levels, but relate to population to a much lower degree and significance.

In conclusion, colleges and elite colleges distributions in the U.S. are not related to economic development and mainly depend on state population levels. However, in China, both the number of regular Bachelor-degree-offering colleges and the number of elite colleges in a province statistically significantly increase with the level of economic development in the province. Provincial population matters for college distributions in general, but not quite for elite college distributions. These facts partially explain our earlier observations that students from less developed regions in China are subject to limitations and disadvantages in terms of colleges and elite colleges opportunities.

6 Conclusions

In the study, we analyze unique student-level data in China and provide new evidence on the existence of significant unfairness and regional merit-based inequalities under the regional quotas system in college admissions in China. Since each province in China has its own college entrance exam, we create a universal standard, the Super-Tests, to compare student ability across provinces. The study compares college opportunities in the six provinces in China – Beijing, Sichuan, Shandong, Guangdong, Henan, and Shaanxi.

The study suggests the following main conclusions. First, college opportunities, especially elite college opportunities, are significantly biased towards students from Beijing and disadvantage students from Sichuan, Henan, and Shaanxi. Controlling for academic achievement, Beijing students have significantly higher chances of being admitted into elite colleges. In addition, a student from Beijing could attend a college that is ranked much higher compared to another student with the same academic ability but from other provinces – about 100 places higher compared to Shaanxi, 85 places higher compared to Sichuan and Henan, and more than 30 places higher compared to Shandong.

Second, in solely merit-based admissions, regional compositions of admitted student body would change remarkably. In this ideal situation with no merit-based regional admissions inequalities, the regional compositions of students from Henan and Sichuan would increase significantly by about 125% and 60%, respectively. The seats to students from Beijing would shrink by about 40%. The regional compositions of students from Guangdong and Shaanxi would be even lower.

Third, analyses controlling for non-academic skills, in our case critical thinking skills, give similar conclusions as analyses controlling for academic skills in mathematics and science. This suggests against the government's argument that some types of non-academic skills are higher for students from developed provinces, and therefore against the reason that the government gives for distributing higher quotas to developed regions.

Fourth, in the pure merit-based admissions, comparatively a higher percentage of Beijing students would be admitted by top 400 colleges in China. This implies potential student ability differences and high school education quality disparities across provinces, at least among high academic achieving students. The proportional quota system analysis results also confirm that complete proportional quota distribution would disadvantage Beijing students. Therefore, the current regional quota differences are results of both the unfairness embedded in the current admissions system and the potential ability inequality across provinces.

Finally, the study confirms that four-year college and elite college distributions in China are significantly associated with economic development. There are more colleges and elite colleges in provinces with higher levels of GDP per capita. This is very likely to be a major factor contributing to the existing regional inequalities. In less developed province there are fewer good colleges, so high ability students need to apply for out-of-province colleges in more developed provinces. At the same time, governments in more developed provinces try to reserve college opportunities and future career opportunities for in-province students, and therefore limit the number of out-of-province students coming for colleges, making it even harder for students in less developed regions to get into good colleges.

References

Ayalon, H. and Yogev, A. 2005. "Field of Study and Students' Stratification in an Expanded System of Higher Education: The Case of Israel." *European Sociological Review*, 21:3, 227-241.

Chen, Y., & Kesten, O. (2017). Chinese college admissions and school choice reforms: A theoretical analysis. *Journal of Political Economy, 125*(1), 99-139.

China Examinations (2006). "30 Years of Gaokao."

Dong, Y. (2012). "Wo Guo Difang Gaoxiao Zhaosheng Jihua Diqu Fenpei Moxing Zhibiao Tixi Yanjiu" [A study of regional college admissions planning and regional quota distribution system in China]. *Journal of East China Normal University (Educational Sciences)*, 2012/12, Vol. 30, No. 4.

Elman, Benjamin. (2009). "Civil Service Examinations (Keju)." Berkeshire Encyclopedia of China, Great Barrington, MA: Berkshire, pp. 405-410.

Espenshade, T. J. and Chung, C. Y. 2010. "Standardized Admission Tests, College Performance, and Campus Diversity." *Office of Population Research, Princeton University*.

Espenshade, T. J., Chung, C. Y. and Walling, J. L. 2004. "Admission Preferences for Minority Students, Athletes, and Legacies at Elite Universities." *Social Science Quarterly*, 85:5, 1422-1446.

Espenshade, T. J. and Chung, C. Y. 2005. "The Opportunity Costs of Admission Preferences at Elite Universities." *Social Science Quarterly*, 86:2, 293-305.

Fan, B. (2006). "Shehuixue Shijiao Xia de 'Gaokao Yimin' Wenti" [Gaokao Immigration from a Sociology Perspective]. *China Examinations.*

Fu, Y. (2013). "China's unfair college admissions system." *The Atlantic*, 19.

Fang, C. and Feng X. 2005. "Jieceng Chayi Yu Jiaoyu Huode: Yixiang Guanyu Jiaoyu Fenliu de Shiyan Yanjiu" [How Distinction of Social Stratum Affects the Attainment of Education: An Analysis of Split-Flows of Education]. *Tsinghua Journal of Education*, 5, 22-30.

Feng, X. (2006). "Taolun Gaokao 'Fei Kaoshi Jingzheng' Yu Jiaoyu Gongping" [A Discussion on Non-Exam-Related Competition in Gaokao and Educational Equality. *China Examinations*.

Goto, M., Iwasaki, A., Kawasaki, Y., Kurata, R., Yasuda, Y., & Yokoo, M. (2016). Strategyproof matching with regional minimum and maximum quotas. *Artificial intelligence*, 235, 40-57.

Goto, M., Kojima, F., Kurata, R., Tamura, A., & Yokoo, M. (2017). Designing matching mechanisms under general distributional constraints. *American Economic Journal: Microeconomics, forthcoming*

Kamada, Y., & Kojima, F. (2015). Efficient Matching under Distributional Constraints: Theory and Applications. *American Economic Review, 105*(1), 67-99.

Kamada, Y., & Kojima, F. (2017). Recent developments in matching with constraints. *American Economic Review, 107*(5), 200-204.

Kang, N. (2005). *Zhongguo Jingji Zhuanxing Zhong Gaodeng Jiaoyu Ziyuan Peizhi de Zhidu Chuangxin* [System Improvements of Higher Education Resource Allocations Under Economic Reforms in China]. Educational Science Publishing House. Vol 1. 285-286.

Li, C. (2010). "Gaodeng Jiaoyu Kuozhang Yu Jiaoyu Jihui Bupingdeng" [Expansion of Higher Education and Inequality in Opportunity of Education]. *Sociological Studies,* 3, 82-113.

Li, L. (2007). "Jiaoyu Gongzheng Shiye Zhong de Gaokao Luqu Zhidu Gaige" [College Admissions Reforms and Educational Equality]. *Hubei Social Science,* 2007(9).

Li, T (2017). Financial Decentralization and Geographical Stratification of Access to Higher Education in China: The Case of Shanghai, *Chinese Sociological Review*, 49:3, 212-238, DOI: 10.1080/21620555.2016.1271701.

Liu, Y. 2015. "Geographical Stratification and the Role of the State in Access to Higher Education in Contemporary China." *International Journal of Educational Development*, 44: 108–17. doi:10.1016/j.ijedudev.2015.08.003

Liu, C., Lv, J., and Cheng, W. (2006). "'Gaokao Yimin' Xianxiang de Renshi Yu Sikao" [Understanding and Thoughts on Gaokao Immigration]. *China Examinations*.

Liu, L., Wagner, W., Sonnenberg, B., Wu, X. and Trautwein, U. (2014). "Independent Freshman Admission and Educational Inequality in the Access to Elite Higher Education." *Chinese Sociological Review*, 46:4, 41-67.

Liu, Y., and Q. Zhang. 2006. "Regional Discrimination in China's Higher Education, Endogenous Growth and Welfare Analysis." *China Economic Quarterly* (Jing Ji Xue Ji Kan) 2006(4): 1023–40.

Lin, L. (2007). By-province College Admissions Quotas Are Established According to a System of Indicators.

Shen, B. (2012). "Keju Shidai de 'Yidi Gaokao' He 'Gaokao Yimin'" [Out-of-Province Gaokao and Gaokao Immigration during Keju]. *Nanfang Metropolis Daily*, 2012-4-1. AA29.

Song, X. (2007). "Lixing Duidai Gaokao Gaige" [A Rational View on Gaokao Reform]. *Guangming Daily*, 2007-4-18.

The Ministry of Education (1977). "Guanyu Yi-Jiu-Qi-Qi Nian Gaodeng Xuexiao Zhaosheng Gongzuo de Yijian" [Comments on 1977 College Admissions Planning]. 1977-9-30.

The Ministry of Education (1987). "Putong Gaodeng Xuexiao Zhaosheng Zanxing Tiaoli" [Temporary Guidelines for College Admissions]. 1987-4-21.

The Ministry of Education (2006). "2006 Nian Putong Gaodeng Xuexiao Zhaosheng Laiyuan Jihua Guanli Shixing Banfa" [Trial Guidlines for 2006 College Admissions Planning and Administration].

Wang, X. (2006). "Gaokao Fensheng Zizhu Mingti de Libi Fenxi" [Pros and Cons of Independent College Entrance Exams in Different Provinces]. *China Examinations*.

Wang, X., Liu, C., Zhang, L. and Luo. R. 2013. "College Is a Rich, Han, Urban, Male Club: Research Notes from a Census Survey of Four Tier One Colleges in China." *China Quarterly*, 214, 456-470.

Wu, X., and Z. Zhang. 2010. "Changes in Educational Inequality in China, 1990–2005: Evidence from the Population Census Data." *Research in Sociology of Education*, 17: 123–52. doi:10.1108/s1479-3539(2010)0000017007.

Treiman, D. J. 2013. "Trends in Educational Attainment in China." Chinese Sociological Review 45(3): 3–25. doi:10.2753/csa2162-0555450301.

Yan, X. (2006). "Wo Guo Gaoxiao Zhaosheng Kihua Zhidu Yanjiu – Jiyu Jiaoyu Gongping de Shijiao" [A study of College Admissions Planning in China from the Perspective of Educational Equality]. Unpublished manuscript.

Yang, D. (2006). "Access to Higher Education: Widening Social Class Disparities." *Tsinghua Journal of Education*, 27:1, 19-26.

Yeung, W.-J. 2013. "Higher Education Expansion and Social Stratification in China." *Chinese Sociological Review*, 45(4): 54–80. doi:10.2753/csa2162- 0555450403.

Zhao, Y. (2006). "Ruhe Dujue Gaokao Yimin Shixian Jiaoyu Gongping" [How to Avoid Gaokao Immigration and Achieve Educational Equality]. *China Examinations*.

Zhou, L. (2006). "Jianchi Gaokao Zhidu Shi Zhi Gengjia Kexue Yu Wanshan" [Proceed with the Gaokao System and Make It Better]. *China Examinations*.

Table 1 Sample Construction

Math and Science Samples	1	2	3	4
Natural-science-oriented students only	Yes	Yes	Yes	Yes
Admission purely based on Gaokao performance (excluding IFAP, direct admission etc.)	No	Yes	No	Yes
Input/measurement error on CEE score within 50 points	No	No	Yes	Yes
N	1,825	1,789	1,735	1,706

Critical Thinking Samples	1	2	3	4
Natural-science-oriented students only	Yes	Yes	Yes	Yes
Admission purely based on Gaokao performance (excluding IFAP, direct admission etc.)	No	Yes	No	Yes
Input/measurement error on CEE score within 50 points	No	No	Yes	Yes
N	900	883	854	840

Table 2 Math and Science Analysis – Gaokao Regional Percentile Rank vs. Academic Ability

Percentile Rank	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Super-test math (z-score)	6.639***	-	6.604***	-	6.510***	-	6.459***	-
Super-test physics (z-score)	4.899***	-	4.731***	-	4.860***	-	4.797***	-
vs. Beijing								
Sichuan	9.711***	13.797***	10.501***	14.450***	10.320***	13.066***	10.528***	13.836***
Shandong	1.003	4.289	1.551	4.474	1.358	5.048	1.439	4.995
Guangdong	15.605***	13.488***	16.193***	14.378***	15.573***	15.578***	15.644***	16.269***
Henan	4.178*	2.762	5.109**	4.111	4.511**	4.928	4.886**	5.931*
Shaanxi	15.426***	18.288***	16.081***	18.040***	16.210***	19.760***	16.129***	19.339***
Sample	1	1	2	2	3	3	4	4
Model	Linear	Partial linear	Linear	Partial linear	Linear	Partial linear	Linear	Partial linear
N	1,825	1,824	1,789	1,788	1,735	1,734	1,706	1705
R-squared	0.3470	0.0635	0.3458	0.0634	0.3630	0.0745	0.3590	0.0739

Table 3 Math and Science Analysis – College Type vs. Academic Ability (Logit Model)

Elite College	(1)	(2)	(3)	(4)
Super-test math (z-score)	1.475***	1.454***	1.473***	1.457***
Super-test physics (z-score)	1.239***	1.195***	1.305***	1.250***
vs. Beijing				
Sichuan	-1.480**	-1.440**	-1.402**	-1.378**
Shandong	-2.822***	-3.035***	-2.840***	-2.996***
Guangdong	-0.816*	-0.794*	-0.803*	-0.803*
Henan	-2.823***	-2.750***	-2.904***	-2.850***
Shaanxi	-0.525	-0.545	-0.524	-0.568
Sample	1	2	3	4
N	1,825	1,789	1,735	1,706
R-squared	0.3539	0.3452	0.3693	0.3627

Table 4 Math and Science Analysis – College Type vs. Academic Ability (Probit Model)

Elite College	(1)	(2)	(3)	(4)
Super-test math (z-score)	0.671***	0.660***	0.666***	0.661***
Super-test physics (z-score)	0.607***	0.586***	0.646***	0.620***
vs. Beijing				
Sichuan	-0.865***	-0.834***	-0.825***	-0.808***
Shandong	-1.606***	-1.690***	-1.622***	-1.683***
Guangdong	-0.523*	-0.504*	-0.527*	-0.523*
Henan	-1.656***	-1.610***	-1.700***	-1.669***
Shaanxi	-0.423*	-0.422*	-0.411	-0.426*
Sample	1	2	3	4
N	1,825	1,789	1,735	1,706
R-squared	0.3353	0.3267	0.3505	0.3442

Table 5 Math and Science Analysis – Highest Possible College Quality/Rank vs. Academic Ability

College Rank	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Super-test math (z-score)	-23.891***	-	-24.062***	-	-23.799***	-	-23.896***	-
Super-test physics (z-score)	-15.755***	-	-15.817***	-	-16.294***	-	-16.373***	-
vs. Beijing								
Sichuan	84.580***	57.083***	83.898***	54.983***	84.933***	62.485***	84.923***	60.203***
Shandong	31.672***	5.091	31.694***	5.353	32.986***	7.938	33.389***	8.798
Guangdong	-0.410	-43.120*	-0.470	-44.381*	0.539	-37.914*	1.043	-39.191*
Henan	85.177***	56.131***	84.173***	50.595**	86.207***	61.136***	85.886***	56.270***
Shaanxi	97.749***	55.745***	99.709***	60.361***	98.228***	55.440***	101.049***	60.771***
Sample	1	1	2	2	3	3	4	4
Model	Linear	Partial linear	Linear	Partial linear	Linear	Partial linear	Linear	Partial linear
N	1,695	1,694	1,661	1,660	1,618	1,617	1,590	1,589
R-squared	0.3982	0.0653	0.3989	0.650	0.3964	0.0710	0.3996	0.0706

Table 6 Simulation Results

	Beijing	Sichuan	Shandong	Guangdong	Henan	Shaanxi	Total Quota
Ranking: top 100							
Current	6,295 (16.4%)	10,030 (3.4%)	10,573 (3.9%)	15,385 (4.8%)	9,425 (2.6%)	6,512 (3.4%)	58,220
Ideal	3,656 (9.5%)	17,343 (6.0%)	12,073 (4.4%)	33 (0.01%)	25,115 (6.9%)	0 (0.0%)	
Ranking: top 200							
Current	12,119 (31.6%)	16,149 (5.5%)	23,440 (8.6%)	24,151 (7.5%)	21,755 (6.0%)	16,711 (8.7%)	114,325
Ideal	6,269 (16.3%)	31,977 (11.0%)	29,169 (10.7%)	1,906 (0.6%)	44,499 (12.2%)	505 (0.3%)	
Ranking: top 300							
Current	15,439 (40.2%)	25,230 (8.6%)	42,302 (15.5%)	42,172 (13.1%)	28,400 (7.8%)	27,529 (14.4%)	181,072
Ideal	8,540 (22.2%)	48,423 (16.6%)	48,652 (17.8%)	8,466 (2.6%)	65,220 (17.8%)	1,771 (0.9%)	
Ranking: top 400							
Current	16,583 (43.2%)	38,280 (13.1%)	58,905 (21.5%)	49,209 (15.3%)	35,766 (9.8%)	35,489 (18.6%)	234,232
Ideal	10,155 (26.5%)	60,454 (20.8%)	63,198 (23.1%)	17,477 (5.4%)	80,151 (21.9%)	2,797 (1.5%)	
Population	38,387	291,322	273,640	321,439	365,686	191,557	

*Numbers in the parenthesis represent the percentiles of the last admitted student in each section in each province. For example, the numbers in row 2 column 1 means that in the ideal case, 1) 3,656 Beijing students should be admitted by the top 100 universities, and 2) the 3,656 students are the top 9.5% of all students in Beijing.

Table 7 Critical Thinking Analysis – Highest Possible College Quality/Rank vs. Academic Ability

College Rank	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Super-test Critical Thinking (z-score)	-9.393***	-	-9.324***	-	-9.394***	-	-9.007***	-
vs. Beijing								
Sichuan	59.300***	57.085*	58.201***	56.160*	60.289***	56.137*	58.023***	61.285*
Shandong	20.597	2.642	18.472	0.898	21.062*	6.718	19.126	12.598
Guangdong	-14.145	-15.899	-16.349	-13.505	-14.780	-13.237	-16.815	-0.983
Henan	72.522***	59.256*	69.955***	60.513*	73.300***	60.894*	70.971***	74.268*
Shaanxi	110.387***	97.597***	110.010***	95.634***	110.811***	98.373***	110.729***	110.427***
Sample	1	1	2	2	3	3	4	4
Model	Linear	Partial linear	Linear	Partial linear	Linear	Partial linear	Linear	Partial linear
N	840	839	827	826	803	802	791	790
R-squared	0.3343	0.0478	0.3372	0.0462	0.3317	0.0482	0.3323	0.0531

Table 8 Regional Quota Proportion Equalization

	Beijing	Sichuan	Shandong	Guangdong	Henan	Shaanxi
Gaokao population	38,387	291,322	273,640	321,439	365,686	191,557
Current						
Top 400	16,583 (43.2%)	38,280 (13.1%)	58,905 (21.5%)	49,209 (15.3%)	35,766 (9.8%)	35,489 (18.6%)
Cutoff score	552	542	572	577	555	502
Cutoff percentile	56.62	86.84	78.46	83.50	90.12	81.10
Ability	37.90	46.87	44.39	40.34	48.91	35.30
Simulation						
Top 400	10,155 (26.5%)	60,454 (20.8%)	63,198 (23.1%)	17,477 (5.4%)	80,151 (21.9%)	2,797 (1.5%)
Cutoff score	598	519	568	617	511	618
Cutoff percentile	73.55	79.25	76.90	94.21	78.08	98.54
Ability	43.92	44.01	43.97	43.88	43.91	44.34
Proportional Quota						
Top 400	6,067 (15.8%)	46,043 (15.8%)	43,248 (15.8%)	50,803 (15.8%)	57,796 (15.8%)	30,275 (15.8)
Cutoff score	629	533	587	577	530	513
Cutoff percentile	84.02	84.06	83.96	83.50	83.91	84.15
Ability	47.98	45.75	45.97	40.34	46.07	36.16

Table 9 The Number of Colleges vs. GDP Per Capita

Number of Colleges in Region	(1) U.S.	(2) China	(3) U.S.	(4) China	(5) U.S.	(6) China
GDP per capita (in thousands)	-0.068	0.422	0.098	0.284*	0.033	0.117***
Sample	All Colleges		Four-year Colleges (Ben-Ke)		Elite Colleges	
N (Colleges)	7,537	2,914	3,097	1,244	119	95
N (Regions)	51	31	51	31	51	31
R-squared	0.0001	0.0583	0.0012	0.1236	0.0407	0.5454

Table 10 The Number of Colleges vs. GDP and Population

Number of Colleges in Region	(1) U.S.	(2) China	(3) U.S.	(4) China	(5) U.S.	(6) China
GDP per capita	-0.054	0.427**	0.103	0.286***	0.033*	0.117***
Population (in millions)	0.019***	1.402***	0.008***	0.566***	0.0004***	0.036*
Sample	All Colleges		Four-year Colleges (Ben-Ke)		Elite Colleges	
Region F.E.	No		No		No	
N (Colleges)	7,537	2,914	3,097	1,244	119	95
N (Regions)	51	31	51	31	51	31
R-squared	0.9353	0.8365	0.8718	0.7184	0.6251	0.6064