To what extent are Contextualised Admission policies in China just and fair? And how could they be improved?

TAN, YIYI

To what extent are Contextualised Admission policies in China just and fair? And how could they be improved?

Yiyi Tan

Thesis submitted in partial fulfilment of the qualification of

Doctor of Philosophy in Education

School of Education

Durham University

May 2023

Abstract

Contextualised admission (CA) to higher education (HE) is a policy attempt to reduce the stratification of student intakes, in widespread use internationally. This study examined whether the use of residence in impoverished provinces, a rural hukou or a minority ethnicity in China could reliably, steadily and accurately identify disadvantaged students for contextualised admission purposes. And it examined where there are better indicators that could be used instead, or as well.

This study involved a structured review of existing evidence on the disparities in HE participation in terms of popularly discussed indicators in China. It then employed secondary data analysis including administrative data, universities' data, and two nationally representative cohort datasets with more than 10,000 cases (Chinese General Social Survey (CGSS) and China Family Panel Study (CFPS)) to answer research questions. Finally, in order to bring this data up to date and obtain some further information, a cross-sectional survey of 800 middle school students was conducted asking about their plans for higher education.

The findings of the structured review and descriptive analysis of CGSS and CFPS confirmed that there are important disparities in university participation between different provinces in China, especially in terms of hukou (residential registration) status, and family socio-economic background. According to CGSS and CFPS, the currently used indicators also include living in an under-developed province, a rural hukou and having a minority ethnicity. Not all of these indicators are necessarily

appropriate to be used for contextualised admissions.

For example, group-level or area-based indicators such as ethnicity, or province or urban/rural residence will exclude some truly disadvantaged students, and falsely include some clearly advantaged students. The analysis of CGSS/CFPS showed that both false identification of disadvantage, and non-identification of disadvantage, were relatively common.

The results of CGSS and CFPS analyses also showed that other potential indicators could be problematic. They provided no evidence that the sex of the student and their month of birth are clearly associated with disadvantage. Other indicators such as social class and parental occupation are hard to define and hard for administrators to verify. Furthermore, some indicators such as having non-party member parents cover a very large sub-set of the population, and school-level indicators have the same deficits as area ones – they mistake the individual for their peers.

The most promising single indicator, according to the findings, might be parental educational credentials. Students with parents who have only completed compulsory education have more difficulty in accessing HE, and they are more likely to be part of a less advantaged family. Parental education information is usually officially verifiable and this information is accessible to higher education institutions (HEIs), making it safe and relatively easy to use.

The survey results illustrate the key point that current contextualised admission policies in China only focus on National College Entrance Examination (NCEE) candidates, and so ignore around half of students in each age cohort. This means that the policies are only concerned with students who are already on track to university. There is little or no widening participation on offer for the large group who do not take NCEE and are largely invisible. This group is stratified by sex, age-in-year, and other contextual variables, as well as by parental education, socio-economic status (SES), and province.

The study has several suggested implications. For non-NCEE students, alternative policies are needed leading to a possible HE future for them, or else they need to have at least the chance to be part of contextualised admissions. For all, removal of artificial barriers such as the restrictions associated with hukou status would be a major step towards a fairer system, less predicated on accidents of birth. For contextualised admissions, authorities need to move away from reliance on group or area-based indicators, and towards use of accessible, verifiable indicators of genuine disadvantage.

Key words: Contextualised Admissions, HE participation, China, secondary data analysis, provinces, Hukou, ethnicity, family SES

Table of Contents

Declaration

I declare that this thesis is my own work and has not previously been submitted elsewhere for any other qualification or degree. The initial findings of how fair the province-based quota policy is have been published in Imagining Better Education: Conference Proceedings (peer-reviewed) (Tan, 2021). The findings of CFPS analysis have been published in Making your doctoral research project ambitious: Developing large-scale studies with real-world impact (Tan, 2022, ed by Siddiqui & Gorard).

Acknowledgement

I would like to thank my supervisors, Associate Professor Nadia Siddiqui and Professor Stephen Gorard, for giving me guidance and support in my PhD. They helped me as supervisors and encouraged like friends. Many thanks also go to Professor Beng Huat See for her ongoing help.

I would like to thank my family. They helped me physically and emotionally during this journey. They always kindly encourage and comfort me when I was upset.

I would like to thank my friends, Yuchong Yang, who took good care of me when I was ill and accompanied with me when I was sad; Bohua Deng, Jialiang Feng, Yuqing Cheng, Yi Feng, who always stayed with me and raised me up; Ruopu Zhai, we supported each other during lock-down; Wenqing Chen, Jiajia Yan, whom I usually discussed academic issues with; Weiqi Zou, Lizi Yi, who kept supporting me although in China.

I hope you all happy.

Part 1 Background

Chapter 1 Introduction

1.1 Research Background

Education is highly valued by scholars, politicians and educators as it is regarded as "a key driver of economic and social success for individuals, employers and nations" (Machin, 2006, p7). Education can improve upward social mobility for disadvantaged individuals and those in poverty by enhancing their chances of getting better occupations and earning a higher income (Chen, Zheng & Xu, 2019; Hu & Hibel, 2014; Su, Xiao & Hu, 2019). However, education might also reinforce social disadvantages as educational opportunities, experiences and outcomes tend to be stratified by individuals' backgrounds (Yang, 2020). Therefore, there is a need for equalised education.

As an important stage in education, Higher Education (HE) can also improve social mobility. For instance, according to Zhao and Wang (2020), the data for 2012, 2013 and 2015 in a nationally representative large-scale social survey, the China General Social Survey (CGSS), display a trend that the chances of upward social mobility are much higher for those in China who have attended HE than for those who have not. Zhang and Liu (2019) also find a 'sheepskin' effect of HE in China. To achieve equity it is crucial to equalise opportunities for access to HE. Equity here does not mean treating people in exactly the same way but helping them get access to HE if they want it regardless of their sex, ethnicity, language and family socioeconomic status (SES) (Salmi & Bassett, 2014).

Internationally, there has been expansion in HE in the past half century, but more student places available in HE does not necessarily mean an improvement in equity. On the contrary, some evidence reveals that socio-economically advantaged students are preferred in allocating these new places. For instance, in Chile from 1990 to 2009 participation in tertiary education by the least socio-economically advantaged quintile only increased from 4.5% to 19.9%, while that by the most socio-economically advantaged quintile increased from 39.7%

to 82% (Salmi & Bassett, 2014). In Brazil, despite an increase in the number of undergraduates in the population from 1.5 million in 1992 to more than 3.8 million in 2003, most of the new places in both state and private universities were taken by students from higher SES backgrounds (McCowan, 2007). A similar pattern has also been witnessed in the UK. Weedon (2014) also highlights the under-representation in HE of students from lower SES backgrounds. Even if they can attend HE, these disadvantaged groups are more likely to attend less-prestigious new universities or colleges. Worse still, not only participation but also willingness to apply for HE is strongly associated with individuals' family backgrounds (Vignoles & Murray, 2016).

Selection bias is even more apparent in the admissions to prestigious universities. Social class patterns in the intakes of older universities in the UK hardly changed between 1960 and 1995 (Boliver, 2011). Russell Group universities in the UK failed to widen access by applicants from less privileged social backgrounds with the Opening Doors policy (Boliver, 2013a; Boliver, 2013b; Boliver, 2015; Boliver & Crawford et al., 2017). Hemsley-Brown (2015) analysed UK survey data on 10,723 university students and found that students who attended private schools, i.e. who are relatively socio-economically advantaged, are more likely to attend prestigious Russell Group universities than their counterparts who attended state schools. Even if members of disadvantaged groups such as ethnic minorities try to get places in prestigious universities they are less likely to receive an offer (Boliver, 2016). Therefore, seven times more students in the most advantaged quintile access HE than those in the most disadvantaged two quintiles (Milburn, 2012).

China is similar. In order to meet the demand for more professional labour, the Chinese government started to expand HE enrolment so that the HE intake in 1999 was 47.4% higher than the figure in 1998 (Ngok, 2008). Since then, the number of HE students has continued to increase (Wang & Liu, 2011). The gross HE enrolment rate increased from 10.5% of candidates in the National College Entrance Examination (NCEE) in the late 1990s to 23% in 2006 (Yang, D. P., 2007). However, this fast development seems to have been at the expense of equity (Zha, 2011). Although it has been claimed that HE inequality in China has decreased (Deng & Fu, 2020; Fang & Feng, 2020; Liu, 2006; Zhang, Huan & Li, 2007), gaps in HE participation between advantaged groups and disadvantaged groups, such as urban and rural residents, are still remarkable. According to Guo, Song and Chen (2019), more educational opportunities produced by the Compulsory Education Law and HE expansion in China did not automatically transform into social or educational mobility for everyone but only benefited urban residents. Fang and Feng (2020) looked at 35,400 cases in CGSS 2003, 2005, 2006, 2008, 2010 and 2012 data and performed a historical change analysis. They also found gaps in educational attainment between rural and urban students in 70 birth cohorts from the 1920s to the 1990s.

Governments in various countries have introduced policies to narrow HE participation gaps and improve HE equity by supporting disadvantaged students. In the UK, contextualised admission (CA) policies have been implemented to widen participation by less advantaged groups. Some universities in the US have also implemented similar policies, named Affirmative Action, which favour minority races, first-generation in HE students or those from low-income families. In China, there are also admission policies functioning as CA. They treat students from distant and under-developed provinces, rural residents and minority ethnic group members as target groups and try to support these students during the transition to HE.

However, regardless of the context, these policies have been subject to many doubts and criticisms regarding their reasonableness, purpose, accuracy and effectiveness. One of the criticisms regards inappropriateness of contextual indicators such as POLAR (non-HE participation of local areas) in the British context and race in the US context (Frisancho & Krishna, 2016; Trent et al., 2003). It is claimed that they erroneously focus on target groups that are not necessarily disadvantaged (Boliver, Gorard & Siddiqui, 2019; Jerrim, 2021).

In China, although extra support for disadvantaged students has been implemented for more than half a century and abundant studies have explored disparities in HE attendance, there are few studies discussing whether the contextual indicators currently employed to identify disadvantaged students are appropriate. The appropriateness of contextual indicators is crucial in CA policy implementation. Only when the indicators accurately and reliably identify the correct target groups can CA policies be just, reasonable and effective. This might not be a sufficient condition but it is a necessary condition. Therefore, researchers need to make efforts to evaluate currently used CA indicators and identify possible better ones to assist the role of gatekeepers in successful CA policies. This thesis examines the quality of currently used CA indicators in China and discusses other possible alternative indicators in order to identify the best ones.

1.2 Research Questions

Good indicators must clearly reveal the disadvantaged. Therefore, this study aims to discover the extent to which HE equity in China has improved or, on the contrary, that to which HE inequality in China between different groups has been aggravated. If there is no difference in HE attendance between groups with different characteristics it might be meaningless to regard these characteristics as CA indicators. Furthermore, this study evaluates the quality of currently used contextual indicators in Chinese CA policies, and to find out if there are more accurate, reliable and accessible indicators to identify disadvantaged students in HE. Therefore, it will answer these research questions in the following chapters:

- What are the disparities in HE participation revealed by the official indicators used in contextualised admission policies in China – province, hukou status, ethnicity?
- What disparities in HE participation are revealed by other indicators such as sex, parental education and occupation?
- To what extent do each of these indicators accurately identify a disadvantaged group of students and only those students?

- Do the indicators lead to better identification of disadvantaged students than those used currently?
- To what extent does an accumulative disadvantage exist in the transition to HE?

1.3 Overview

Chapter 2 introduces the key concepts involved in this study. Chapter 3 describes what a contextualised admission (CA) policy is and discusses its underlying principles. Chapter 4 introduces the CA policies implemented in China currently. Chapter 5 is a short literature review on how CA policies, and especially the quality of their indicators, are evaluated in other contexts such as the UK.

Part 2 contains Chapters 6, 7 and 8, which present the research design, the datasets employed and the indices and formulas used in this study.

Part 3 contains the results. Chapter 9 provides a structured review which provide a snapshot of HE participation gaps between groups with different characteristics. Chapter 10 evaluates the indicator of hukou province. Chapters 11 to 17 evaluate the indicators of hukou status, ethnicity and other alternative ones based on analyses of the Chinese General Social Survey (CGSS) and the China Family Panel Study (CFPS). Chapters 18 and 19 present the findings of an original primary survey conducted by the author. Chapter 20 discusses the findings and concludes the whole thesis.

After considering previous evidence in the structured review and obtaining updated evidence from analyses of large-scale secondary datasets and survey data, this study reveals the disparities in HE participation existing between provinces, hukou status, ethnic groups, family socio-economic statuses and school backgrounds. In short, students from competitive or less developed provinces, those with rural hukou, those belonging to minority ethnicities, those living in less affluent families and those studying in non-selective schools are less likely to access HE. All of these groups are under-represented in HE participation. However, there are no apparent gaps in HE participation between the sex and birth-month groups age in year.

However, this does not necessarily mean that all of these indicators are high quality identifiers of disadvantaged students. The currently used hukou province, rural hukou status and minority ethnicity indicators are subject to an ecological fallacy, meaning that privileged students might be mistakenly supported by CA policies while truly disadvantaged students might be excluded from the target group. Therefore, it might be safer and fairer to use these indicators combined with other indicators when targeting disadvantaged HE applicants. However, such combined indicators cannot deal with the problem of 'false negatives.'

On the other hand, some popularly discussed indicators are also inappropriate for official use in CA policies. Sex groups and birth-month groups, for example, do not reveal noteworthy gaps in HE participation. In other words, they do not clearly identify disadvantage. Family SES-related indicators are associated with HE participation, but some of them are still problematic when selecting disadvantaged students. For instance, non-party member parents represent a huge proportion of the population, and parents' occupations and workplaces do not necessarily indicate their income or financial situation. Language ability and the language used for daily communication, apart from the potential problems of inaccuracy and measurement errors, are vulnerable to deliberate concealment and cheating if they are used as CA indicators. School-level indicators are also unsatisfactory as they are likely to be the results of previous stratification, for example non-selective schools and classes and lower cognitive ability scores and test scores are associated with disadvantaged backgrounds.

The most promising indicator may be parents' educational achievement. Students whose parents only completed compulsory education are remarkably under-represented in HE and this indicator of less-educated parents is closely related to other potentially disadvantaged characteristics. Parental educational qualifications significantly contribute to predicting students' HE participation. Importantly, this information is reliable and easily accessible as higher education institutions (HEIs) can obtain it from an official website.

Another critical finding is the exclusion of disadvantaged students in earlier education stages. Current CA policies in China only support National College Entrance Examination (NCEE) candidates, who are likely to be ones who 'have beaten' many other competitors at earlier stage of education. However, those who have been excluded in earlier education stages tend to be really or more disadvantaged. This accumulative disadvantage makes them lose in the competition for entry to academic high schools and these students have almost no chance to enter HE. This missing group represents around half of the corresponding age cohort but is invisible to current CA indicators.

This is just a summary of some main findings in this study. More specific analyses and detailed discussion are provided in the following chapters.

Chapter 2 Introduction to Key Concepts

This chapter explains some key concepts in this thesis. They are ‘higher education (HE),’ ‘National College Entrance Examination’ (NCEE, ‘Gaokao’ in Chinese), ‘Hukou’ and ‘minority ethnicity.’ These concepts are not necessarily controversial or inconsistent in terms of their definitions or mearnings and neither are they particularly sophisticated. They are introduced either because they are specific to the Chinese context, such as hukou, NCEE and minority ethnicity or pertinent to the purpose of this study, such as higher education (HE). Therefore, it is helpful to clarify them before answering the research questions.

2.1 Higher Education

Higher education (HE) is one of the most important concepts in this thesis. This section first shows how higher education is defined in this study and then introduces the hierarchy and distribution of higher education institutions (HEIs) in China. Finally, there is a comparison of HEIs between provinces.

2.1.1 The Definition of Higher Education in this Study

Higher education (HE) usually refers to post-secondary or tertiary education, and includes various categories such as university education, vocational college education and adult higher education. This is almost the same in China. HE in China is officially defined as formal education activity after secondary education which aims to produce highly skilled and professional members of society who can participate in the advancement and practice of knowledge (Higher Education Law of the People’s Republic of China, 2018). Chinese HE includes:

1) Degree courses such as undergraduate education, vocational college education and postgraduate education; and

2) Non-degree courses such as undergraduate courses, spare-time education, open education and distance education.

HE in this thesis is distinguished from other types of post-secondary education. Specifically, instead of including broader forms of post-secondary education, HE here only indicates degree-based undergraduate and postgraduate university education. HE in this study does not cover the Top-up project (a project in which some graduates from vocational colleges can take a test to get a place to study for one or two more years in academic universities), part-time undergraduate education, vocational education or non-degree courses. These programmes are regarded here as post-secondary education but not HE.

The reasons for this deliberate selection are, first, that the Top-up project, part-time undergraduate education and non-degree courses do not necessarily select and accept their students after the NCEE but often many years later. For instance, Top-up project students are likely to be ones who did not perform well in the NCEE and opted for university through the vocational pathway. After two or three years of study in vocational colleges, these students have a chance to get places in universities if they pass a Top-up examination. Furthermore, those attending part-time undergraduate education and non-degree courses are more likely to be non-traditional mature students who also did not go into HE immediately after leaving high school for various reasons.

Because this study is about equity in HE admissions based on selection through the NCEE, it might not be able to clearly understand disparities in HE admission processes if non-traditional students are included. However, what needs to be clarified here is that this does not mean that these non-traditional students are not important in research interested in HE equality and equity. On the contrary, they are included in the discussion on educational equity but for the sake of clarity they are best treated as a separate category instead of vaguely being regarded as ‘HE participants.’

The second reason for this clarification is that full-time undergraduate education in China, on the one hand, and vocational education, on the other, are strongly distinguished in terminology, qualifications, application processes, prestige and other factors (Jia & Ericson, 2017). For instance, part-time undergraduate education can only grant students a bachelor's degree with an additional 'part-time' label. Students in the Top-up Project who complete an academic degree from a university can only obtain an educational qualification with a 'Top-up Project' label, which also indicates that the qualification differs from those from full-time undergraduate education. Students in vocational college education and non-degree courses cannot have a bachelor's degree as their institutions are not qualified to award it at all.

In addition, many Chinese students prefer formal full-time undergraduate education, which might be an imperceptible effect of long-standing meritocracy. Typically, students who get good results in the NCEE are unwilling to participate in part-time undergraduate education, vocational education or non-degree courses. It is reasonable to say that the competition to attend HE in China is actually competition to attend a full-time undergraduate university ('benke' in Chinese), and for most students the competition to attend a full-time undergraduate university in China is a competition to get a place in an elite university (Lai et al., 2016; Liu & Li, 2014). Furthermore, students believe that a degree from a high-ranking university will be helpful for their careers (Liu, X. H., 2014). Many employers, including the government, prefer graduates from universities (Zhang & Liu, 2019). For instance, SANY Group (三一集团 in Chinese) and ByteDance clearly set a bachelor's degree or master's or even a Ph.D. degree as an essential criterion for applicants. College rankings are also found to be closely associated with the chance to work in a managerial position (Hu & Vargas, 2015). This implies that individuals without university qualifications find it more challenging to achieve upward social mobility. Therefore, formal full-time undergraduate education is differentiated from other post-secondary education and is labelled HE in this study.

Postgraduate education in China requires a bachelor's degree as the fundamental requirement to be able to apply. That is to say that only when someone obtains a bachelor's degree from formal undergraduate education can they be eligible to apply for postgraduate education. For the above reasons, this study only refers to full-time undergraduate and postgraduate education when it uses the term HE. The study investigates whether access and pathways to these two education categories are fairly accessible in the context of current education policies.

According to their specialised status that they can award a HE degree, the HEIs in this study are degree-awarding bodies that can grant a bachelor's degree. These include prestigious HEIs such as those in the Double First-Class project (DFC)/World Double-First project (WDF), Project 985 and Project 211, and general universities.

This restricted focus does not suggest that other courses and institutions are valueless. Neither does it seek to defend the status quo. Descriptions of hierarchy, desirability and prestige used in HE discourses refer to how the situation is generally perceived and practised in China. The study does not advocate hierarchy and it might be better to have a more open HE system valuing students equally whatever their age and sex, for example.

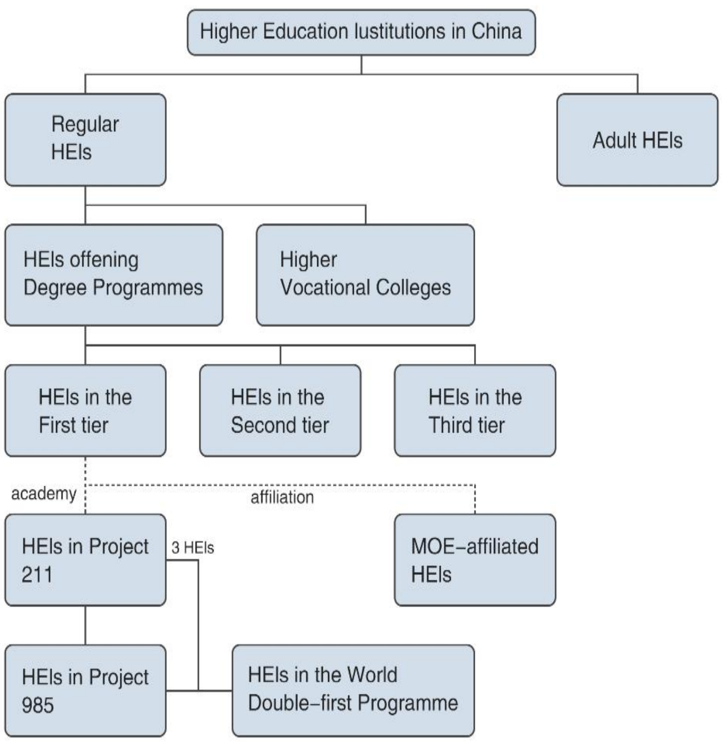
2.1.2 The Hierarchy of HEIs in China

There is a strict hierarchy of HEIs in China. According to data on the website of the Ministry of Education (MOE, 2020), there were 2,740 undergraduate HEIs and 265 Adult HEIs in mainland China on 30 June 2020. Among the 2,740 undergraduate HEIs, there were 1,272 HEIs developing students' academic ability which can offer bachelor's degree programmes. These are called regular HEIs in this study. The remaining 1,468 HEIs are vocational colleges, which focus more on skill development, and they are not part of this study.

Furthermore, the 1,272 regular HEIs are ranked in several tiers and HEIs in different tiers admit their students in a succession of admission rounds. The first tier contains the well-known top universities and some provincial selective HEIs. These HEIs select their students in the first wave of the admission process and they skim off students who score in the first tier NCEE threshold (yibenxian in Chinese) (Yang, Wang & Chen, 2019; Zhao & Guo, 2002). The second-tier (and third-tier in some provinces) HEIs are considered less prestigious but are still able to award degrees to graduates, and they can select their students in the second (or sometimes the third) wave of the admission process (Yang, Wang & Chen, 2019).

What needs attention here is that the two/three tiers are not rigorously distinguished. The tier classification is not consistent because there is no definite regulation for HEIs to indicate which tier they are in, and this is more related to the waves in which universities can select their students in different provinces. Some HEIs, for example, might only be allowed to accept applications from NCEE candidates in the second wave in some provinces, but they can accept applications in the first wave in other provinces. It would be difficult to classify the tiers of these HEIs precisely.

Figure 2.1.1 The hierarchy of HEIs in China



It would therefore be ambiguous if this study used the tier as the criterion for prestigious universities. In order to avoid confusion as much as possible, instead of defining elite HEIs as first-tier HEIs, the study will employ a less changeable criterion and consider Ministry of Education (MOE)-affiliated HEIs, Project 211, Project 985 and DFC Project universities as elite HEIs (see Figure 2.1.1). A detailed introduction to these elite groups is provided later.

Affiliation hierarchy.

Elite HEIs are more likely to be affiliated with the central government or its departments. For instance, 124 of the 1,272 HEIs are governed by Central Ministries or other agencies, including the Ministry of Education, the Ministry of Public Security of the People's Republic of China, the Chinese Academy of Sciences, the Overseas Chinese Affairs Office of the State Council, the Ministry of Industry and Information Technology of the People's Republic of China, the National Ethnic Affairs Commission of the People's Republic of China, the Ministry of Transport of the People's Republic of China (MOT), the Ministry of Justice of the People's Republic of China, the Civil Aviation Administration of China, and the Chinese Academy of Social Science. Many of them are classified as first tier. They are regarded as prestigious universities in China, especially those (76) under the direct control of the MOE, which are universally accepted as the best in China.

These elite HEIs usually get more funding not only from the local government but also from the central government (Chen & Peng, 2018). As a corollary, the support from the central government approximately equals the support from taxpayers nationwide. This indicates that, theoretically, these HEIs are supposed to accept students on a nationwide basis rather than being biased towards local or regional enrolment.

Most of the other HEIs are administered by provincial governments and are mainly supported by them in terms of finance, infrastructure and other factors (Sun & Barrientos, 2009; Wang, 2001). These local HEIs, especially those in poor provinces, have substantially fewer resources than their 124 central-government-affiliated counterparts. The elite group that receive greater financial resources promote 'competitiveness' and become even more prestigious. This imbalance in resources among HEIs can partly explain why the HEIs controlled by the central government, especially those affiliated with the MOE, are the most attractive to Chinese students.

Academic Hierarchy.

Regarding academic excellence, HEIs can be classified in different projects. One of the oldest but most influential of these is Project 211, which was started in 1993 and now includes 115 universities. All Project 211 universities are uncontroversially in the first tier (MOE, 2020; Qin & Buchanan, 2019; Zhao & Guo, 2002). They are selective universities and enjoy some privileges in the allocation of important resources such as financial support.

Five years after Project 211, in 1998, the MOE decided to focus more resources on certain universities to promote the development of 'world-class' universities. It selected 39 Project 211 universities to constitute Project 985. These 39 HEIs were given higher priority for resource allocation and are regarded as the highest-quality universities.

Later in 2015, the Chinese government came up with Double First-Class (DFC)//World-Double First (WDF) Project universities and disciplines. The purposes of this project are to establish first-class universities and develop first-class disciplines. The universities included in this project were named in 2017. They include all 39 universities in Project 985 and three new ones, namely Xinjiang University (in Xinjiang), Yunnan University (in Yunnan), and Zhengzhou University (in Henan). All three new elite universities are in underdeveloped areas.

Within the DFC project, there is also a hierarchy: 36 universities are classified as Type A and they are recognised as the most prestigious in China; and 6 universities are evaluated as Type B, a slightly lower level than Type A but still considered elite universities.

Except for the three youngest elite universities, all the other DFC universities are directly governed and controlled by central government ministries. This affiliation is essential for them to be perceived as having high prestige. In China, the more elite HEIs are, the more closely they associate with the central government, and vice versa.

As it is the highest classification, this study will employ the DFC universities as representatives of prestigious universities in its analysis of intakes and participation in prestigious HEIs.

2.1.3 The Geographical Distribution of HEIs in China

HEIs, and especially prestigious HEIs, are unevenly distributed across China (Gou, 2006; Wang, L. X, 2019; Wei & Liu, 2015). For example, as is shown in Table 2.1.1, Jiangsu province has the most HEIs that can offer degree programmes to their students, with 78. The provinces with more than 60 degree-awarding HEIs are Beijing (67), Hebei (61), Liaoning (65), Zhejiang (60), Shandong (70), Hubei (68) and Guangdong (67). All these are either economically developed areas (such as Beijing, Zhejiang and Guangdong) or areas with huge populations (such as Hebei, Shandong, Hubei and Guangdong).

On the other hand, areas which are more remote and/or economically disadvantaged – such as Qinghai (4), the Tibet Autonomous Region (hereafter Tibet) (4), Hainan (8), the Ningxia Hui Autonomous Region (hereafter Ningxia) (8), the Inner Mongolia Autonomous Region (hereafter Inner Mongolia) (17) and the Xinjiang Uygur Autonomous Region (hereafter Xinjiang) (19) – tend to have fewer regular HEIs, which refer to universities that could offer Degree programme.

When it comes to the prestigious HEIs groups within regular HEIs, the distributions between provinces are even more unbalanced. Beijing has the most, with 24 MOE-affiliated HEIs. Shanghai is in second place with 8 MOE-affiliated HEIs. Jiangsu and Hubei both have 7 HEIs directly controlled by the MOE. These are followed by Shaanxi with 5, Sichuan with 4 and Shandong with 3. Six provinces (Tianjin, Chongqing, Liaoning, Jilin, Hunan and Guangdong) have only 2 MOE-affiliated HEIs and six provinces (Hebei, Heilongjiang, Zhejiang, Anhui, Fujian and Gansu) only 1.

There are 12 regions with no MOE-affiliated HEIs at all. Therefore, the HEIs in these areas might be disadvantaged in terms of resources such as research funding and/or investments. These 12 regions are Shanxi, Jiangxi, Henan, Hainan, Guizhou, Yunnan, Qinghai and five autonomous regions.

Table 2.1.1 The distribution of different categories of HEIs in China

	DFC HEIs	985 Project	211 Project	MOE-affiliated	HEIs offering Degree Programme
Beijing	8	8	26	24	67
Shanghai	4	4	10	8	40
Hunan	3	3	4	2	52
Shaanxi	3	3	8	5	57
Tianjin	2	2	3	2	30
Liaoning	2	2	4	2	65
Jiangsu	2	2	11	7	78
Shandong	2	2	3	3	70
Hubei	2	2	7	7	68
Guangdong	2	2	4	2	67
Sichuan	2	2	5	4	53
Chongqing	1	1	2	2	26
Jilin	1	1	3	2	37
Heilongjiang	1	1	4	1	39
Zhejiang	1	1	1	1	60
Anhui	1	1	3	1	46
Fujian	1	1	2	1	39
Henan	1	0	1	0	57
Yunnan	1	0	1	0	32
Gansu	1	1	1	1	22
Xinjiang	1	0	2	0	19
Hebei	0	0	1	1	61
Shanxi	0	0	1	0	34
Jiangxi	0		1	0	45
Hainan	0	0	1	0	8
Guizhou	0	0	1	0	29
Qinghai	0	0	1	0	4
Inner Mongolia	0	0	1	0	17
Guangxi	0	0	1	0	38
Tibet	0	0	1	0	4
Ningxia	0	0	1	0	8
Overall	42	39	115	0	1,272

Source: MOE

Note: the provinces are listed in the order of the numbers of DFC HEIs

Regarding the distribution of prestigious academically excellent HEIs – DFC, Project 985 and Project 211 universities – from Table 2.2.1 it is clear that the capital, Beijing, has more DFC Project universities than any other region (8). This number is twice that of Shanghai, which has the second most DFC universities (4). Both cities are rich in economic, cultural and educational resources. Hunan and Shaanxi rank in third place with three DFC universities each. However, more than half the provinces (19 out of 31) fall behind with only one or none. These provinces are Chongqing, Jilin, Heilongjiang, Anhui, Fujian, Jiangxi, Henan, Yunnan, Gansu, Hebei, Shanxi, Hainan, Guizhou and Qinghai and five autonomous regions.

The distribution of Project 985 HEIs is similar but a little worse than that of DFC project HEIs. Three more regions are particularly disadvantaged with no prestigious HEIs: Yunnan, Henan and Xinjiang.

The distribution of Project 211 HEIs is slightly different to the other two categories. Beijing still occupies the most advantaged place, with 26 Project 211 universities, followed by Jiangsu with 11 and Shanghai with 10. Shaanxi and Hubei rank fourth and fifth with 8 and 7 Project 211 HEIs respectively. These are followed by Sichuan (5) and Liaoning, Heilongjiang, Hunan and Guangdong (4 each). Tianjin, Jilin, Anhui and Shandong each have 3 Project 211 HEIs and Chongqing, Fujian and Xinjiang each have 2. The remaining 14 provinces have only one Project 211 university. It is worth noting that the number of Project 211 HEIs in Beijing (26) is more than the sum of those in the bottom 16 provinces (20). The first five provinces, namely Beijing, Jiangsu, Shanghai, Shaanxi and Hubei have half (62) the total Project 211 HEIs (115), while 26 other regions only have 53 in total.

To sum up, the distribution of HEIs in China, especially the high-quality ones, is uneven. The most advantaged areas in terms of HEIs are Jiangsu, Beijing, Shandong, Hebei, Liaoning, Zhejiang, Hubei and Guangdong. The most advantaged areas in terms of high-quality HEIs are Beijing, Jiangsu, Shanghai, Hubei and Shaanxi. There are quite a few overlaps between these two groups.

As for the poorest areas, Qinghai, Hainan and four autonomous regions have the fewest HEIs. In addition to them, Gansu, Yunnan, Guizhou, Henan, Jiangxi, Fujian, Anhui, Zhejiang, Heilongjiang, Shanxi and Hebei have the lowest numbers of prestigious HEIs.

Although these differences suggest that the distribution of HEIs between provinces is unbalanced, it is too early to conclude that the differences are unjust, because the above comparison does not take HE applicants or potential HE students into account. Therefore, in the next section a more careful analysis shows how many HEIs there are for every ten thousand NCEE candidates in each region.

2.1.4 Numbers of HEIs and numbers of NCEE candidates

Table 2.1.2 compares numbers of HEIs for every ten thousand NCEE candidates. Beijing, Shanghai and Tianjin are still advantaged, not only in terms of regular HEIs (see the fifth column) but also in terms of prestigious ones (see the first four columns). There are respectively 11.36, 5.4 and 8 regular HEIs for every ten thousand NCEE candidates in Beijing, Tianjin and Shanghai. These are followed by Liaoning, Jilin and Jiangsu, with respectively 2.66, 2.27 and 2.3. The most disadvantaged areas, where the problem of HEI shortage is so severe that every ten thousand candidates have less than 1 regular HEI, are Anhui, Henan, Guangdong, Sichuan, Guizhou, Gansu, Qinghai, Inner Mongolia and Guangxi.

The gaps are even more prominent regarding prestigious HEIs. For instance, in Beijing, every ten thousand NCEE candidates have 1.36 DFC universities and 1.36 Project 985 universities, 4.41 Project 211 universities and 4.07 MOE-affiliated universities. The corresponding figures in Tianjin are 0.36, 0.54 and 0.36 and in Shanghai 0.8, 2 and 1.6.

However, except for these three regions, all the other provinces have less than 0.1 DFC universities and Project 985 universities for every ten thousand NCEE candidates. A few of them have little more than 0.1 Project 211 universities and MOE-affiliated HEIs, but many provinces do not have any. The most disadvantaged group in terms of universities per ten thousand NCEE candidates remains stable: Hebei, Shanxi, Zhejiang, Anhui, Jiangxi, Shandong, Henan, Hunan, Guangdong, Guizhou, Hainan, Yunnan, Gansu, Qinghai, Inner Mongolia, Guangxi and Ningxia.

These results repeat the finding that provinces with large numbers of NCEE candidates or with the less-developed economies are more likely to be disadvantaged in terms of HEIs. Both regular and prestigious HEIs are unevenly distributed between provinces. Beijing, Shanghai and Tianjin are always the most privileged areas. Theoretically, candidates in these three areas enjoy many more HE resources than their counterparts in other areas, who are up against more competition to win a place in a HEI.

This comparison is somewhat problematic. First, it is based on the assumption that HEIs only accept applicants from the province where they are located, which is not true. However, although DFC, Project 985, Project 211 and MOE-affiliated HEIs are theoretically supposed to admit students from across the nation, their intakes have always been locally biased (see Chapter 10). Therefore, the richer the places where prestigious HEIs are located, the more of their places will be assigned to the region.

On the other hand, as was previously mentioned, lower-tier HEIs, whose sponsors are more likely to be the local government and local companies, enrol most of their students from the locality. Therefore, it is natural that the more regular HEIs there are in a province, the more places will be allocated to students from the province.

Another problem with the comparison is that it does not consider admission quotas, infrastructure resources, teaching-equipment resources, teacher quality and academic ability, and other HEI issues. An elite university enrolling about 10,000 students every year will make a different contribution to local HE admissions than one which only enrols around 6,000 students. These differences in student intakes among universities exist and should not be ignored.

Table 2.1.2 Numbers of HEIs for every ten thousand NCEE candidates

	DFC universities	985 Project	211 Project	MOE-affliated	HEIs offering Degree Programme	NCEE candidates (ten thousand)
Beijing	1.36	1.36	4.41	4.07	11.36	5.9
Shanghai	0.80	0.80	2.00	1.60	8.00	5
Tianjin	0.36	0.36	0.54	0.36	5.36	5.6
Shaanxi	0.09	0.09	0.25	0.15	1.75	32.59
Liaoning	0.08	0.08	0.16	0.08	2.66	24.4
Jilin	0.06	0.06	0.18	0.12	2.27	16.27
Hunan	0.06	0.06	0.08	0.04	1.04	49.9
Jiangsu	0.06	0.06	0.32	0.21	2.30	33.9
Hubei	0.05	0.05	0.18	0.18	1.77	38.4
Heilongjiang	0.05	0.05	0.20	0.05	1.91	20.4
Fujian	0.05	0.05	0.10	0.05	1.88	20.78
Chongqing	0.04	0.04	0.08	0.08	0.98	26.4
Gansu	0.04	0.04	0.04	0.04	0.82	26.68
Shandong	0.04	0.04	0.05	0.05	1.25	55.99
Zhejiang	0.03	0.03	0.03	0.03	1.88	32
Yunnan	0.03	0.00	0.03	0.00	1.00	32
Sichuan	0.03	0.03	0.08	0.06	0.82	65
Guangdong	0.03	0.03	0.05	0.03	0.87	76.8
Anhui	0.02	0.02	0.06	0.02	0.90	51.3
Henan	0.01	0.00	0.01	0.00	0.57	100
Average	0.00	0.01				
Hebei	0.00	0.00	0.02	0.02	1.09	55.96
Shanxi	0.00	0.00	0.03	0.00	1.08	31.4
Jiangxi	0.00	0.00	0.02	0.00	1.07	42.1
Hainan	0.00	0.00	0.17	0.00	1.33	6
Guizhou	0.00	0.00	0.02	0.00	0.63	45.87
Qinghai	0.00	0.00	0.18	0.00	0.70	5.7
Inner Mongolia	0.00	0.00	0.05	0.00	0.85	19.9
Guangxi	0.00	0.00	0.02	0.00	0.83	46
Ningxia	0.00	0.00	0.14	0.00	1.12	7.17
Tibet						——
Xinjiang						——

Note: the provinces are listed in the order of the numbers of DFC HEIs theoretically-enjoyed by every ten thousand NCEE candidates; the numbers of NCEE candidates are data for 2019; the figure of 0.00 for some provinces does not necessarily mean 0 but might be smaller than 0.005.

Finally, the comparison would be improved if it took the size of the population of 18-year-old students into consideration instead of just NCEE candidates. The NCEE is a vital examination for Chinese students to transit to HE at the end of their high school education, and it will be introduced in the next section. The population of 18-year-old students are people who theoretically would be NCEE candidates or HE applicants if they did not abandon education because of various disadvantages. The group of NCEE candidates only contains candidates who have passed the selective high school entrance examination and had high school education. They are already winners before they apply for university places. Therefore, the focus on this group is risky as it ignores the more disadvantaged students. The latter unquestionably should not be ignored, although data limitations make it necessary here. If these potentially disadvantaged students were taken into account in the comparison, the differences between numbers of HEIs per ten thousand 'candidates' might be even more notable, as economically-developed areas might have higher transition rates to next-step education because their basic education is better developed (Xiang et al., 2020; Yang, 2018).

2.2 The National College Entrance Examination (NCEE)

The last section provided some background on HEIs in China and how they are distributed. This section introduces the most crucial means for Chinese students to get a place in HE. It is called the National College Entrance Examination (hereafter NCEE) or 'Gaokao' in Chinese. The NCEE aims to overcome regional differences in the performance of students and give them all an equal chance of entering the HE system. Exam-based selection of students for HE is also implemented in many other countries, although there is no evidence that it increases fairness or equality or whether it is just an additional barrier for students who aspire to university qualifications, and a superficial screening method. However, this issue is not the focus of this study.

NCEE, as the name indicates, is an official examination to evaluate whether students are qualified to attend HE and what kind of HE they deserve. It is usually held on the 7th and 8th of June with '3+X' tests in Chinese, Mathematics, English and an optional combined subject. Students in their last year of high school who want to take the NCEE have to register six months beforehand, in November or December of the previous year. The raw NCEE scores are often available at the end of June and then students can start to apply for university places according to their scores. There are two (three in some provinces) waves of HE admission, as was explained in Section 2.1. Only when the more elite admission wave is finished can the less prestigious HEIs start to enrol students.

Although the NCEE is not the only means of participating in HE, it is not overtly biased by students' backgrounds and is a commonly adopted examination. Unlike interviews, essays or other forms of examination, it is a standardised test and has more objective outcomes, which are relatively easily evaluated. More importantly, a test using interview performance might have some disadvantages for students from less privileged backgrounds, as they are more likely to perform less well and have less 'guanxi' (a Chinese social concept that personal relationships based on exchanges of favours are important in many social activities) than their more advantaged counterparts in interviews (Tsang, 2013). In fact, students submit their NCEE answers anonymously. It is not possible to know which test paper has been done by whom. In addition, the evaluation process is also blind. Answers to multiple choice questions are automatically evaluated by computers, as NCEE primarily employs an e-rater. As for subjective questions, examiners are only responsible for one question and will not know which question they will evaluate until the NCEE papers are submitted. This double blindness gives the NCEE a "veil of ignorance" (Rawls, 1999), which is somewhat helpful in reducing potential discrimination. Therefore, some authors highly praise the NCEE as a tool to select qualified HE applicants and promote social equity and mobility (Liu & Wu, 2006).

Other ways of entering HE such as recommendations ('baosong' in Chinese) are more background-biased and unequal. 'Baosong' is useful for students who show outstanding talent in sports or STEM subjects, including Mathematics, Physics and Computer Science, and have won prizes in international competitions in these subjects. These students can be accepted by universities without them taking the NCEE. However, even if their talents are outstanding and valuable, they are likely to be associated with good family backgrounds. There is some evidence that wealthier families and better-educated parents are more likely to be able to afford training in these areas and would be willing to pay for it (Wang, 2010). That is to say, the reward for winning a prize in a competition might be the reward for the opportunity to receive specialised training in a valued subject, and so actually for a better family background. If so, it may be even more unequal and unjust than test-based selection.

There are various requirements for NCEE registration. The first is that students must abide by the Constitution and other laws in China and the second is having gained graduation certification from a high school. The third requirement is important for this study. It is that students have to register for the NCEE in their hukou location. The hukou location is one of the critical elements in the hukou system, which is a system for managing people in China. More details about hukou are provided in the next section. This requirement limits students' mobility between provinces. Otherwise, they might not be eligible to register for NCEE.

After taking the NCEE, although many students start their applications with their raw NCEE scores, some such as rural hukou students and ethnic minority students can gain extra credits according to the extra credits policy and 'three special plans' (see Chapter 4). Rural or urban hukou status is another key element of hukou and it will be explained in detail together with hukou location in Section 2.3. Furthermore, the consequences of ethnicity in China might be somewhat different from those in other contexts. Detailed information on this is given in Section 2.4.

2.3 Hukou

Hukou, or huji, is the name of the household registration system in China. It is an administrative tool to help the Chinese government manage people. It is such a powerful system that it would not be an exaggeration to say that it has permeated into Chinese people's daily lives, including in the field of education. There are two essential elements of hukou. The first is hukou location and the other is hukou classification or hukou status. Both are closely associated with education.

2.3.1 Hukou location

Hukou location refers to one's registered permanent residence, i.e. which village/town, which prefecture/district in which city, and in which province. This location information does not prevent people from going to or living in other provinces, but it can still be a barrier against internal migration for people in China (Zhang, H. F., 2017). One of the reasons is the social welfare system changes from place to place. Quite a few provinces only provide public services such as medical insurance and free compulsory education for people with that hukou location. Even if people with a hukou location in another province live and work in these provinces, they cannot enjoy the same social welfare as locally-registered people. In addition, whenever people living in other areas need legal certification for any reason, such as renewing their ID cards or changing their names, they have to go back to their hukou residence to deal with these issues.

For HE, the most critical information in the hukou location is the registered province because of the province-based quota admission policy (see Chapter 4). Students in China who want to enter HE and take the NCEE are only allowed to register in their official hukou location province. When these students get their NCEE scores and are going to apply for HEIs, they are also only allowed to use the enrolment quotas in their own hukou province (MOE, 2020b, 2021b). For instance, students whose hukou location is Hebei can only register for the NCEE and apply for universities in Hebei rather than in Beijing or Tianjin, even though Hebei is very close to the latter two municipalities and has a much larger population of candidates and smaller HE admission quotas.

Because of this requirement, students' mobility before the NCEE is restricted to a certain extent. Despite this restriction, some students try to change their hukou location from a competitive province such as Henan, Hebei, Hunan and Hubei to an advantaged province such as Beijing and Tianjin. They are called 'NCEE migrants' ('gaokao yimin' in Chinese) (Liu, 2020; Luo, 2020). In order to reduce the number of, or more ambitiously eliminate, NCEE migrants, the Chinese government has implemented various policies. One of these policies requires students to take the NCEE in the province where they finish at least their 3-year high school education. If they only attend the last year of high school education in their new hukou province, they cannot register for the NCEE. They have to attend high school in that province from the first year. However, a transition to a high school also requires a local hukou. Therefore, this policy seems able to reduce the number of NCEE migrants and they have indeed become fewer (or less visible) in the past few years, but it is still hard to evaluate the effects of this policy.

On the other hand, the government has loosened the restrictions for students with other hukou locations so that students can register and take the NCEE in other provinces nowadays. However, there are many additional certifications needed to achieve 'yidi gaokao' (students taking the NCEE in places not where their hukou location is). For example, if migrant students want to take the NCEE in Shanghai, their parents must have living in Shanghai certification and get more than 120 points for this certification. The criteria for the points include age, educational attainment, tax-paying, a professional title or a technical post, marriage and occupation. Most of these criteria are advantage-favoured, which means that students with parents working as migrant-workers or peddlers still find it difficult to obtain yidi gaokao. They might have no choice but to return to their hukou province if they want to take the NCEE.

2.3.2 Hukou status

The second key element in hukou is hukou classification, or hukou status as it is referred to by Golley and Kong (2018). The categories are agriculture hukou, non-agriculture hukou and collective hukou. The collective hukou status is rare and is for people in the armed forces.

This status is largely ignored in the rest of this thesis.

At the time of the founding of the People's Republic of China (hereafter PRC), in order to manage people, the government classified them according to their way of making a living. People mainly engaged in agriculture were labelled agriculture hukou, also known as rural hukou, and people working in business, industry or other kinds of non-agricultural work were labelled non-agriculture hukou, or urban hukou (Fan, 2013; Wang, 2014). People's offspring generally inherit this Hukou status. Although they are from a different era and seemingly obsolete, these classifications still matter and still restrict people's choices.

Although they might have done in the past, rural and urban status do not necessarily refer to rural or urban areas of residence nowadays, but rural hukou people are indeed more likely to live in rural areas. Neither does this status prevent internal migration or rural/urban people moving to urban/rural areas. People with rural hukou can therefore move to urban areas. Even if they do, however, it will not mean that they have changed their hukou status to urban, and they cannot enjoy the same social welfare as urban residents. Differentiated policies and social welfare for rural hukou and urban hukou, which are underlying or sometimes even straightforward discrimination against rural hukou people, might confine their mobility (Xu & Montgomery, 2021).

An example of this differentiated social welfare is compulsory education (Li, Wang & Han, 2020). Some migrant workers with rural hukou work and temporarily live in urban cities, but many of them have to leave their children in their rural hometowns. As the children do not have local urban hukou, they cannot have free compulsory education in cities like their urban counterparts (Heckman & Yi, 2012). They have no choice but to stay in their hometown, usually with their grandparents. The topic of 'left-behind children' or 'hometown-leave children' has been hotly debated in China (Deng, 2021; Yao, 2021).

Nowadays, the restrictions on migrant workers' children attending compulsory education in urban cities are getting a little looser, especially in metropolitan areas such as Beijing, Shanghai and Shenzhen. The children can receive compulsory education there but they are likely to be segregated in schools specialised for them (Jiang, 2017; Lu & Zhou, 2013; Ngok, 2007). Furthermore, if these children have different hukou locations, they still have to return to their hometown province to attend high school if they plan to apply for HE. It has been claimed that this is a form of structural inequality (Ling, 2017).

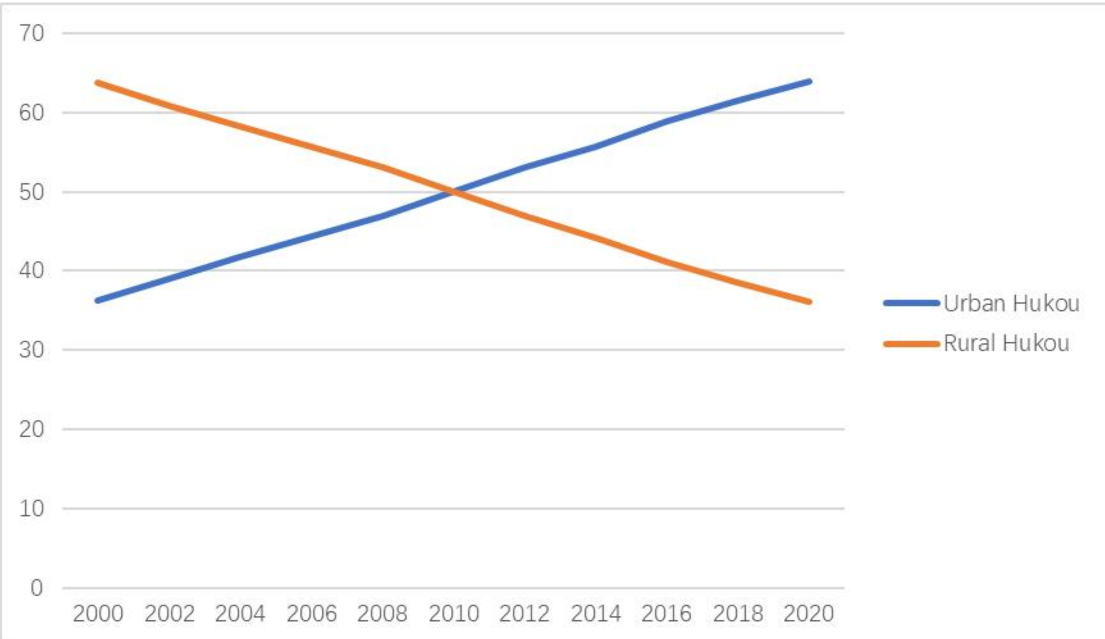
To participate in HE, the restrictions by hukou status are not as strict as those by hukou location as there are no more distinctions in the provincial admission quotas between rural hukou students and urban hukou students. In other words, whatever their hukou status students can be enrolled in universities as long as their scores in the NCEE reach the entry requirement.

However, for rural hukou students, neither of the choices above, ‘left in hometown’ or ‘studying in segregated schools in cities,’ seem promising of good-quality compulsory education (Chen & Feng, 2013; Gu & Yeung, 2020; Xu & Wu, 2022). This worse access to good-quality early education is likely to be associated with less satisfactory results in the NCEE later on and more difficulty in attending HE (Chen & Feng, 2013; Gu & Yeung, 2020; Xu & Wu, 2022). Rural hukou students tend to be less competitive than their urban hukou counterparts and many researchers have found gaps in HE participation between them (Qian & Smyth, 2008; Wu, 2012; Xiang et al., 2020; Zhang, H. F, 2017). Even worse, it can be challenging for rural hukou students to get into high school.

For people disadvantaged with a rural hukou it is not impossible to change their hukou status, but it is not easy. Some possible ways are the government deciding on an urbanisation process and attending HE (Wu & Zheng, 2018). The former refers to the process of the government deciding to change an originally rural area into an urban area, which will automatically change the hukou identity for all the people living there into urban hukou. The urbanisation process has been much employed in the past two decades, resulting in a change in the composition of the population from 63% being rural hukou and 37% being urban hukou in 2000 to only 38% being rural hukou population and 62% urban hukou in 2020 (see Figure 2.3.1). Attending HE allows people to change their hukou to that of the place where their HEI is located, and many HEIs in China are located in urban cities.

However, the urbanisation process is completely determined by the government. More importantly, it does not make much difference even if the status is changed from 'rural' to 'urban' in this way because the gaps in economic and educational resources do not disappear. Therefore, although it is good that the government has started to abolish the distinction between rural hukou and urban hukou by combining them into residence hukou ('jumin hukou') and there is an impressive reverse in the distribution of the two groups, more effort is still needed to reduce the gaps between originally rural hukou people and originally urban hukou people. Although rural people can migrate to urban cities once they have urban hukou more freely, given the higher cost of living, those who can afford to do this will be the more advantaged families, thus creating a new inequality.

Figure 2.3.1 Percentages of urban hukou residents and rural hukou residents, 2000-2020



Source: National Statistics Bureau

The second way to change hukou status is related to the topic of this thesis. Students can only change their hukou status to where their HEI is located once they successfully go to this HEI. This study will explore the equity of this premise.

2.4 Minority Ethnicity

Minority ethnicity refers to all the ethnic groups except for the ethnic majority in a multi-ethnic country. For instance, the UK government recognises as ethnic minorities all ethnic groups, such as Asian and black minorities, except the white British group (GOV.UK, 2021).

In the Chinese context, however, minority ethnicity has little relationship with race. It only refers to the 55 non-Han minority ethnic groups. According to the seventh population census published in 2021 (NSB), the size of the ethnic minority population is 125,470,000, 8.89% of the total population.

Although the 55 non-Han ethnicities can all be called minority ethnicities, when it comes to HE admission they have very different levels of disadvantage. For instance, these minority ethnic groups live in various parts of China. Some live mixed with the Han majority and other minority ethnic groups in urban cities. Other minority ethnic people constitute small colonies such as in the five autonomous regions – Tibet, Guangxi, Xinjiang, Inner Mongolia and Ningxia – and in autonomous counties in Yunnan, Guizhou, Sichuan, Gansu, Heilongjiang, Liaoning, Jilin, Hunan, Hubei, Hainan, Taiwan and Chongqing. The above provinces with large minority ethnic populations are often not advanced in economic and educational resources. Therefore, the students there might find it more challenging to receive good-quality education and win a place in a HEI than those living in mixed urban cities.

In addition, language barriers might result in problems (Jiao, 2020). Some minority ethnic groups have their own oral and written language, such as the Tibetans, but some do not or only have a different accent from Mandarin. The former groups will be more disadvantaged in written examinations in Mandarin than the latter ones. Although some examinations and textbooks are written in minority ethnic languages such as Mongolian, Tibetan or Uygur, mainstream education, or at least mainstream high-quality education, is in Mandarin. Therefore, besides the development of their living areas, language ability and fluency in Mandarin might be another barrier against receiving better education for minority ethnic students .

Considering their difficulties in getting access to better education, the Chinese government has implemented several policies which are beneficial to minority ethnic students to encourage and help them to attend HE. These policies include regarding minority ethnic students as a disadvantaged group and adding extra credits to their raw scores in NCEE when they apply for HE. More details of this policy will be provided in Chapter 4.

2.5 Discussion and Conclusion

This chapter has clarified some concepts that are crucial in this study: higher education (HE), hukou and minority ethnicity. In summary, HE in this study refers to formal full-time undergraduate and postgraduate education, not vocational college education, part-time undergraduate education or non-degree courses. 'HEIs' in this study only means regular universities eligible to grant bachelor's degrees instead of covering a more comprehensive range of vocational colleges, independent colleges or adult colleges.

Hukou is a particular Chinese system connected to residence. It is closely related to HE enrolment and participation. Hukou location information restricts the places where students can register for the NCEE and apply for HE, and hukou status, which creates long-term disadvantages for rural hukou people, leads to inequalities in HE.

Ethnicity in this study is not related to race but just the 56 recognised ethnicities in China. The majority of Chinese people are of Han ethnicity and the other 55 ethnicities are called minority ethnicities. There are differences in the levels of disadvantage of different minority ethnic groups.

In order to reduce misunderstanding, these key concepts have been explained at the very beginning of this thesis. Some of them might be further explained when they are mentioned in the following chapters. There are other crucial concepts that need clarifying, one of which is that of contextualised admission policies. However, because CA policies are at the core of this study and deserve deeper discussion, they will be introduced in a separate chapter. The next chapters, therefore, introduce and justify contextualised admission policies and describe how they are employed in China and other countries worldwide.

Chapter 3 Contextualised Admission Policies

The previous chapter clarified some key concepts in this study. This chapter introduces the most important term, namely contextualised admission (CA) policies, and discusses how they are employed in reality. Furthermore, instead of arguing in favour of CA policies in detail, this study assumes that they are a reasonable short-term measure, and simply explains the philosophical principles underlying them.

3.1 An Introduction to Contextualised Admission Policies

Although they may be given various names, contextualised admission policies are policies that take into account students' background information in university application and admission procedures (Coulson et al., 2017). They provide some additional help or support for recognised disadvantaged students. At least two things need to be clarified here. The first is which group(s) of disadvantaged students are actually concerned. In other words, what indicators are used to identify disadvantaged subjects? The other is what kind of help or support is provided.

In order to clarify the first, it is important to understand what contextual data is. According to Gorard, Boliver, Siddiqui, Banerjee and Morris (2017), contextual data indicate disadvantages that students suffer from such as physical problems, language barriers and disabilities. These complex disadvantaged data are often translated into one-dimensional quantitative flags or markers (Coulson, Garforth, Payne & Wastell, 2017).

The contextual data can be different in different contexts. For instance, in the UK, there are three main types of indicators: individual- and family-level indicators; area- or neighbourhood-level indicators; and school-level indicators (Boliver & Gorard et al., 2017; Boliver, Gorard & Siddiqui, 2015; Gorard et al., 2017). Indicators in the first group include free school meal (FSM) eligibility, special educational needs (SEN), parent/carer educational qualifications, family income/tax credits and English as an additional language (EAL). The neighbourhood-level indicators include POLAR, IDACI, IMD, SIMD and ACORN[1]. The school-level indicators include school type, performance at KS4 (age 16) and KS5 (age 18), percentages of students with FSM eligibility and percentages of HE participation. Furthermore, some scholars have identified another category of contextual data: participation in outreach programmes conducted by universities themselves (Boliver & Crawford et al., 2017; Weedon, 2014). This indicator has been widely used by Sutton Trust 30 HEIs.

There has been increasing implementation of CA policies in universities in the UK. Of 68 UK universities involved in a 2015 survey, 84% of them confirmed they implemented CA policies, compared with 37% in 2012 (Sundorph, Vasilev & Coiffait, 2017). The University of Edinburgh was the pioneer that first officially implemented a CA policy (Mountford-Zimdars, Moore & Graham, 2016). It was followed by several Scottish universities. Now nearly all Scottish HEIs have implemented such policies (Boliver, Gorard & Siddiqui, 2017; Gorard et al., 2017; Sundorph, Vasilev & Coiffait, 2017). Boliver & Gorard et al. (2017) summarise the indicators used by Scottish HEIs when they make decisions in the admission process. There is a wide variety of indicators chosen by various HEIs. For example the University of Dundee School of Medicine gives its applicants an opportunity, known as an 'adversity statement,' to self-claim an 'adverse circumstance,' including poverty or a less favourable educational environment, that they think HEIs should take into account (Owen, Anderson & Dowell, 2018).

Some HEIs in England, including the London School of Economic and Political Science, Durham University, the University of Northampton and the Open University, have also implemented CA policies (Butcher, Corfield & Rose-Adams, 2012; Sundorph, Vasilev & Coiffait, 2017). Boliver and Crawford et al. (2017) show how CA policies are implemented in Sutton Trust 30 universities. They use the cases of the University of Bristol and Newcastle University to exemplify how universities use contextualised indicators to identify their

[1] POLAR = HE non-participation of local areas; IDACI = Income Deprivation Affecting Children Index; IMD = Index of Multiple Deprivation; SIMD = Scottish Index of Multiple Deprivation; ACORN = A Classification Of Residential Neighbourhoods.

targeted disadvantaged students. The University of Bristol selects its CA-eligible students using school type, POLAR information, participation in outreach programmes provided by the University of Bristol and spending time in care. Newcastle university checks students' school type, FSM eligibility, time in care, first-generation HE in a family with no one working in a professional occupation, deprived areas, and family environment as indicators to evaluate eligibility for CA.

Furthermore, the University of Nottingham uses the contextual criteria 'living in an area with low HE participation rates', 'refugee status' and 'spent more than three months in care' to help identify disadvantaged students (University of Nottingham, 2023). The university of Manchester (2023) also implements a CA policy, which gives contextual offer to students who 'live in a disadvantaged area or an area with low progression into HE', 'study in the school/college that has performed less well than the national average', 'spent more than three months in care' and 'refugee status'. The University of Cambridge (2023), the university of Liverpool (2023) and the University of Surrey (2023) are also employ several types of indicators: the individual-level indicators, the area-based indicators and the school-level indicators.

Jerrim (2021) describes a rarely used indicator: Output Area Classification (OAC). This indicator, which combines students' home postcode and various other data provided by schools, the government and students themselves, has recently been used by Cambridge University. Many other HEIs, such as the University of Manchester, have also implemented CA policies and even conducted research to examine their effectiveness (Mullen, 2011).

The UK is not unique in implementing CA policies. For example, similar policies are named 'affirmative action' rather than 'contextualised admission' in the US (Dougherty & Callender, 2017). Since the 1960s, HEIs in the US have positively discriminated in favour of minority racial groups such as African Americans and Hispanics, and females (Allen et al., 2002; Downing et al., 2002; Garrison-Wade & Lewis, 2004; Orlans, 1992; Tierney, 1997).

For example, the University of North Carolina at Chapel Hill, one of the prestigious public universities in the US, admits African American and Hispanic students with lower SAT scores (Arcidiacono, Lovenheim & Zhu, 2015). Princeton University gives some support to students who are the first generation to attend HE in their family and from low-income families (first-generation low income – FLI) (Princeton University, 2022a). There is even an event called 'FLI is FLY Week' in Princeton (Princeton University, 2022b). There is a similar programme called U@Uni academy in Australia, which aims to help students with lower socio-economic backgrounds enter HE (Dodd, Ellis & Singh, 2020).

Not only English-speaking countries but also countries in other regions have implemented complementary policies to attenuate educational inequality. India, for example, regards scheduled castes (SC) and scheduled tribes (ST) as disadvantaged groups, so they are the beneficiaries of affirmative action in Indian HE (Arcidiacono, Lovenheim & Zhu, 2015; Frisancho & Krishna, 2016).

A second important element in contextualised admission policies, although not the focus of this study, is support for potentially disadvantaged students. The extra support in Scottish HEIs includes a lower entry requirement such as one or more grades lower, more consideration for an admission offer and an adjusted offer (Boliver & Gorard et al., 2017; Mountford-Zimdars & Moore, 2020). The support provided by Sutton Trust 30 universities, which is summarised in Boliver & Crawford et al. (2017), is somewhat similar to that provided by Scottish universities. In addition, some HEIs offer additional programmes for disadvantaged students. For example, the Faculty of Medicine at the University of Southampton has a BM6 programme, which provides an additional year, called Year 0, for students from less privileged backgrounds to help them become familiar with university culture and life (Curtis et al., 2014). Some universities in the US provide more diverse assistance, including extensive outreach programmes in high schools located in less-advantaged districts, free airline tickets for disadvantaged students if they have difficulty in paying for transport to attend recruitment interviews, and directly adding more points to their

SAT scores (Downing et al., 2002).

Despite different names, indicators and kinds of help, these policies are similar in their essential respects. They can all be viewed as positive discrimination in favour of potentially disadvantaged students (Gorard, Boliver & Siddiqui, 2017). Their aim, or at least one of their aims, is to help HEIs identify potential applicants with lower socioeconomic backgrounds or from less-favoured groups when they make admission decisions and then provide these applicants with assistance to improve their opportunities for HE participation (Garrison-Wade & Lewis, 2004; Mountford-Zimdars, Moore & Graham, 2016). By increasing these disadvantaged students' chances of participating in HE, these policies aim to mitigate injustice and improve educational equity.

The next section elaborates a little further on the principles underlying CA policies and explain why CA policies are viewed as just.

3.2 The Philosophical Basis for Contextualised Admission Policies

Contextualised admission policies are controversial. However, it is not the aim of this study to discuss the philosophical rationale for this positive discrimination practice. Instead, this section will only explain why in this study these policies are regarded as just on the basis of their underlying principles.

Admittedly, equity is a complicated and controversial term in social science. This term and its near synonyms, equality, justice and fairness, have been debated for a long time. Philosophers, politicians and social scientists have presented a variety of opinions on why equity is important, in which respects people should and should not be equal and, fundamentally, what equity means.

In this study, equity is regarded as a synonym of justice because equity in education is closely associated with distributive justice. For instance, one of the primary resources to be distributed is the opportunity to attend HE. On the other hand, equity in this study cannot simply be interpreted as equality, which refers to indiscriminate treatment of every individual. Such impartial fairness would imply the same examination, the same entry requirement or the same selection process for HE for students from different backgrounds, which may be praised as equality of opportunities or procedural justice but is actually somewhat lacking in distributive justice.

The concept of distributive justice has been controversial since Aristotle proposed it. Aristotle supported distribution based on teleology, holding that distribution of any resources or rights should take their purposes into account and assign them to those who deserve them (Barker, 1948). The only criterion to evaluate whether a person deserves the resource is his "contribution to the function of that society" （Barker, 1948, p.129). For instance, the only person who deserves a better flute is the one who is superior in playing the flute, regardless of whether this person is great in any other way. Therefore, according to Aristotle's reasoning, HE opportunities or resources should be assigned to those who deserve them.

However, several questions need to be answered before this position can be accepted. First, one of the most important questions is what are the purposes of universities? If we cannot determine what they are, how can we judge who deserves HE? On the other hand, even if we assume that we know the purposes of HE and hold that they are, for example, to promote individual socialisation, to improve social development and to achieve the common good, can we really find objective criteria to measure them? Finding someone who deserves HE for these purposes is not as easy as finding the best flute player.

In addition, and possibly most crucially, Aristotle's thoughts partly justify extra honours for wealthy people who can make contributions to the community, like a family that can provide their children with several flutes and can afford the best flute tutor to teach their children so they can play very well (Sandel, 2010). Is it really just if they are given better flutes? Then, is it really just if the opportunity to go to university is only, or mostly, given to students with abundant resources?

Besides proportionate equality, teleology evaluates justice as whether a person fits his/her role or his/her properties. This argument has been much criticised because it ignores individuals' liberty and rights (Sandel, 2010). These limitations are undoubtedly closely associated with Aristotle’s historical background, but this does not necessarily mean that they can be free from criticism because of their long history. To sum up, the concept of equity, or distributive justice, based on teleology is not sufficiently tenable and it does not seem appropriate in contemporary society.

Another well-known philosophy that has been prominent in the English-speaking world for centuries is utilitarianism. Utilitarianism does not support distribution based on one's contribution but holds that just distribution should be good for the most people or the long-term interests of society (Mill, 2014). In the HE domain, this may be interpreted as meaning that HE places should be allocated considering the interests of most people. There are, however, some problems that need to be clarified regarding this.

First, what are the interests of most people in HE participation? Furthermore, how can we confirm and examine them? More importantly, even if we could be sure what the greatest interests of most people are and we are confident that we can satisfy them, how about the rest of society? Is it really just to pursue the greatest happiness at the expense of, or at least ignoring, the interests of the minority? Therefore, utilitarianism has also been criticised for neglecting individual rights and liberty (Callinicos, 2000). Although Mill’s proposals mitigated conflicts between Benthamism and individual rights to a certain extent, they are still vulnerable to this kind of criticism, as Mill held that protecting individual rights promotes social development, which makes respect for individual rights more contingent (Sandel, 2010).

Egalitarian liberalism partly modifies the problem of utilitarianism in this respect. It holds that individuals' rights and liberty should be respected, and emphasises the need to amend social and economic inequalities and offer everyone opportunities to succeed. According to Rawls (1999, p54), "All social values … are to be distributed equally unless an unequal distribution of any, or all, of these values is to everyone's advantage." "Everyone's advantage" here does not mean the aggregate advantage of everyone but the advantage of each individual. For instance, some originally disadvantaged individuals are more likely to remain disadvantaged when resources are distributed. Rawls (1999) calls these people the least advantaged. The least advantaged should never be ignored, even though they might be the minority. Only when this group is going to be taken care of can unequal distribution be forgiven. This is to say that extra support and help in the distribution or re-distribution of resources and social values should be reasonably assigned to the disadvantaged. In the HE domain, then, it would be just to give extra support to disadvantaged students so that they can access HE.

However, there are some counterarguments. First, neo-liberal defenders might argue that everyone has the same right and freedom to compete for a HE place in the free HE market, and nobody will be forced to not fight for a place. The failure of those who do not get a place in the HE market is a consequence of their free choice and lack of merit rather than of inequality. Therefore, there is no need, not even an ethical one, for policy to give them additional support (Callinicos, 2000).

Another counterargument is that because HE selection, and especially that without an interview but only with a written examination with anonymous codes like the NCEE in China, is similar to "the veil of ignorance" (Rawls, 1999), it is just and reasonable to select people who are talented and diligent and exclude those who are not eligible according to the standards set under the veil (Liu, F. T., 2013).

These arguments seem plausible, but actually they are weak. While it is true that theoretically the HE market is equally accessible by every individual, in reality to what extent do students who abandon HE make this decision out of their true free will instead of being restricted by other factors or issues? This does not necessarily mean that they are forced to leave education by someone or something. They might encounter difficulties, such as poverty, which make them have to abandon it. If so, for them abandoning education is less a free choice and more a necessity.

If they do not choose to give up but are excluded by standardised tests because they have unsatisfactory scores, should they be blamed, and do their counterparts with higher test scores really deserve their HE places? There have been many studies and theories that reveal the relationship between educational achievement and family resources (Bathmaker, Ingram & Waller, 2013; Coleman, 1966; Coleman, 1968; Mountford-Zimdars, Moore & Graham, 2016; Nahai, 2013): the richer family resources are, the more high-quality educational resources the family might possess (Xu, Song & Liu, 2018) and the more likely students are to get good outcomes in education and cognitive tests, even if they are of similar ability to their counterparts from less-advantaged families (Cardak & Ryan, 2006; Cui, Liu & Zhao, 2019; Duan, Guan & Bu, 2018; Huang, Xie & Xu, 2015; Qi & Wu, 2020; Yao & Ye, 2018). Therefore, equal test scores do not necessarily imply individuals are equal in potential, so it would be appropriate to take contextual factors into consideration when making any decisions on the basis of formal educational outcomes (Schwartz, 2004).

Not only educational achievements but also intentions to work hard and educational aspirations are associated with family socioeconomic status (Gil-Flores, Padilla-Carmona & Suárez-Ortega, 2011; Gorard, See & Davies, 2011; Liu, Zhang & Li, 2014; Liu, Zhang & Li, 2015; McCulloch, 2017; Sheng, 2014). Although Li and Xie (2020) claim that the dependence of individuals' educational expectations on family SES background in East Asian countries is not as strong as in Western countries, there is much evidence of the close relationship between these two factors (Fang et al., 2020; Jian & Peng, 2019; Liu, Jiang & Chen, 2020; Wei & Ma, 2017; Yang, Yao & Zhang, 2016). Not only social luck such as family socioeconomic status but also natural luck, including birth order (Booth & Kee, 2009; Harkonen, 2014) and birth month (Alton & Massey, 1998; Strom, 2004), which seem more random, have also been claimed to be to a certain extent associated with educational outcomes. That is, there might already be some inequality in educational outcomes before selection by so-called 'merit', so selection which seems to be based on talent and diligent high scores is more likely to be actually based on family SES or other characteristics (Gorard et al., 2007).

Therefore, even though opportunities to aspire for, apply for and attend HE may seem equally distributed, opportunities to become talented, develop merit and be eligible to stand behind the veil of selection for HE and be able to afford the financial cost of HE are already unequal. This will never be a simple issue about equality of HE opportunities but a more complicated result of education justice. As Rawls states,

> *Since inequality of birth and natural endowment are undeserved, these inequalities are to be somehow compensated for. Thus, the principle holds that in order to treat all persons equally, to provide genuine equality of opportunity, society must give more attention to those with fewer native assets and to those born in less favourable social positions* (Rawls, 1999, p86).

In addition, regarding the talents and abilities that are rewarded in selection, Rawls (1999) questions extra rewards for them even if they are really innate. He stresses that the "difference principle" is justified that nobody should be rewarded because of independently defined advantages, as nobody deserves better natural abilities and talents and, more importantly, living in a society that rewards our strengths is also not a result of merit. An example of this might be changes in the concept of talent in China's HE selection. In the examination system, talent refers to the ability to get high scores in the secondary academic curriculum. In contrast, in the recommendation system, talent is regarded as one's ethical abilities, character and personality, and even, with some bias towards workers and peasants, social status in the 1970s

(Bratton, 1979). Talent in the former context has changed little, but ‘talent’ for recommendation involves outstanding abilities in STEM subjects or sports and no longer prefers the descendants of workers and peasants (see Chapter 4).

The differences in the concepts of talent in Chinese HE selection support Rawls’s position to a certain extent. The question is then how can it really be fair and just to reward a socially constructed advantage? If it is not fair, why do disadvantaged individuals, regardless of which factors they are disadvantaged in, not deserve additional help to achieve social equity and justice? Callinicos summarises the reason for egalitarian liberalists seeking to improve equity: "*People should not suffer the consequences of disadvantages for which they are not responsible, whether these disadvantages derive from the distribution of productive resources or the incidence of natural talents*" (Callinicos, 2000, p. 60).

Therefore, the argument for free choice and emphasising talents and diligence is not established. It seems that compared with the free and open HE market with equal opportunities that neo-liberalism pursues, Rawls’s (1999) “difference principle” appeals more in essence. If the absolute equality of education with the same selection standards is insisted on in order to select for HE by merits, it could be far from genuine justice. If so, it is possible to expect more exclusion of disadvantaged groups from HE. To improve equity or at least reduce inequality (Coleman, 1975), it is reasonable to implement more complementary CA policies for disadvantaged groups, as equity in education is the means to the end of equality and equality of life chances (Wong & Nicotera, 2004), not the end itself.

In addition, according to Gorard, Smith & Benadusi (2010, p65), equity is more than a “state, quality, or ideal of being impartial, just and fair” but it helps explain the underlying logic when one judges something to be fair or unfair. Therefore, instead of being a universal principle, equity, which is changeable and flexible, is more like a supplement to common law for true fairness. Its connotation changes both in different situations and according to different individuals’ feelings and values.

For example, it might be considered unacceptable for teachers to give extra guidance to male or female students, but it is more acceptable when this help is given to students with learning difficulties or special educational needs (SEN) (Gorard, Smith & Benadusi, 2010). This type of extra support is more likely to be viewed as equity in education. On the other hand, while parents of SEN students think this extra help is what they deserve, parents of non-SEN students might still argue that it is unfair to pay more attention to SEN students and require completely equal treatment.

Such different assumptions by parents concerning extra help for particular sex groups or SEN groups imply that equity is different in different contexts. It is difficult, if not impossible, to satisfy everyone and reach so-called absolute equality. More importantly, absolute equality is far from true justice, as was explained above. Moreover, these differences also exist in people's feelings and perspectives on racism, standards, equal opportunities and in other equity-related fields (Orlans, 1992).

Gorard, Smith & Benadusi (2010) argue that equity is not and might never become a universal or relative value but it is an attempt to understand the reasons we use when evaluating policies, propositions and behaviours. Equity will only be effective in certain situations and from particular viewpoints and it is not possible to satisfy everyone.

3.3 Discussion and Conclusion

This chapter has introduced contextualised admission policies and shown how they are implemented in some countries. In order to widen participation, universities implement contextualised admission policies to provide their target groups with additional support to help them get access to HE. Although the names vary, contextualised admission policies have been implemented in various countries. In addition, the target groups of CA policies, which depend on which contextual data are employed, are different in different contexts.

This study does not intend to debate the reasonableness of CA policies but directly claims that these policies, which are a kind of positive discrimination, are right to provide the disadvantaged with additional support, such as lower entry requirements. What needs a little more explanation is that, first, there is no intention to argue that improving equity is the only purpose of CA policies. It is recognised that these policies have other purposes such as "compensation, correction and diversity" (Tierney, 1997). However, as this study focuses on the equity problem it has not discussed other purposes. Second, it does not necessarily mean that the importance of equal opportunities and procedural justice should be denied. On the contrary, both are vital conditions for equity and the resulting justice. However, they are somewhat superficial if we only pay attention to them, and then it might be difficult to avoid greater inequalities. Therefore, CA policies, which as Steinberg (1995, p 165) argues to pursue equity more than "as a right or a theory but as a fact and a result," would be more appropriate in this respect.

The next chapter will explain the HE admission process and HE admission policies in China. Furthermore, it will discuss the CA policies currently implemented in the Chinese context.

Chapter 4 Contextualised Admission Policies in China

This chapter discusses contextualised admission policies in Chinese HE. The chapter explains the policy objectives and regulations. A brief critique of the implementation of the policies and its limitations follows.

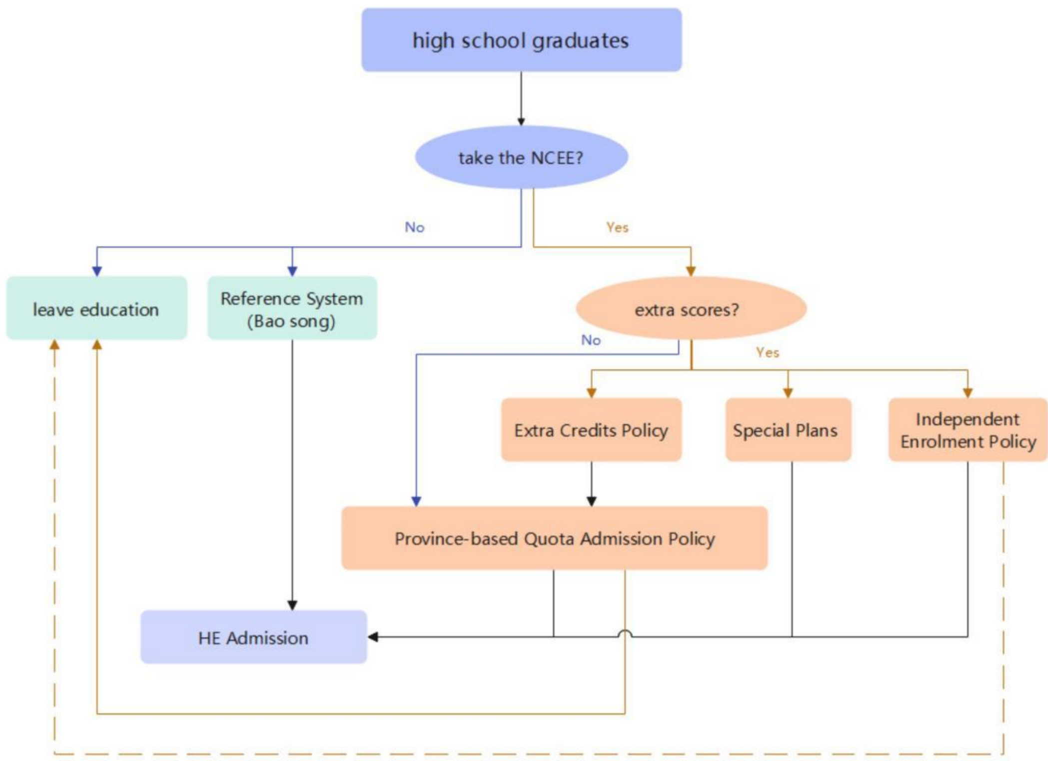
4.1 The HE Admission Process in China

Although the admission systems in China may not be as diversified or as flexible as those in some other countries, there are several routes to access HE. Before going further into Chinese CA policies, it is necessary to briefly explain these routes.

Figure 4.1 roughly summarises the common routes for Chinese students to enter HE. At the beginning of the application process, high school students need to decide whether or not they are going to take the National College Entrance Examination (NCEE). There are several alternatives to taking the traditional NCEE route. For example, they might be disadvantaged in educational achievement or their financial situation so they have to abandon the idea of HE. On the other hand, an increasing number of students from China go abroad for HE and this second group is likely to come from socioeconomically advantaged families in which the parents have the means to invest in their children's education. The two above groups are aggregated in the 'leaving education' group in Figure 4.1, because there will be no relevant admission information in the Chinese HE system.

Some high school students do not take the NCEE because they intend to enter HE on the basis of talent and extraordinary potential. These students are outstandingly talented in subject-specific skills or performance-based disciplines such as art and sport. Universities, mostly prestigious ones, offer them places because of their extraordinary potential. This process is called the 'Reference System' or 'Recommendation System' (Bao song in Chinese).

Figure 4.1 The process of admission to HE in China



The Reference System is a meritocracy-based system, which aims to be an improvement on the standardised NCEE test to select talented students in broader areas such as sports, language and morality (Cheng, 2020; Davey et al., 2007; Wang, 2010; Wang, 2011). However, due to increasing criticisms, the Reference System criteria have become stricter. There are only five remaining possibilities for students to get admission to HE through the Reference System without taking the NCEE (MOE, 2020a; 2021a). They are as follows:

1) to be selected for the National Olympia team for high school students in Physics, Chemistry, Biology, Computing or Mathematics because of talent in these fields, and take part in the International Science Olympia;

2) to win an 'Outstanding Student' award at the provincial level;

3) to be a retired athlete who has won one of the first three prizes in a domestic sport

competition, one of the first six prizes in Asian sports competitions, one of the first eight prizes in international sports competitions or who has won a ‘Master Sportsman’ award at the international level;

4) to be selected from certain Foreign Language High Schools as a gifted student at learning foreign languages (but the major at university for these students is usually restricted to one kind of foreign language);

5) to be a student with a parent who was in the police and died on duty (but only Public Security universities or colleges which aim to train qualified police officers will admit these students without NCEE scores).

Most high school students take the NCEE if they intend to enter HE. After finishing the examination and receiving the results in June, students can start to apply for HE online if they are not eligible for any other extra scores. However, if they are, they can first add the extra scores to their raw NCEE scores and then start their applications.

There are three main ways to get extra scores. These are called the Extra Credits Policy, Three Special Plans and the Independent Enrolment Policy. The first two are complementary but are different and will be introduced in detail in the following sections, while the last one is more talent-oriented.

The Independent Enrolment Policy, also called the Independent Recruitment Policy, allows universities, usually elite ones, to administer another independent examination to identify talented students (Wang, 2010; Wang, 2013). NCEE candidates interested in these universities can apply to take part in these examinations. As the examinations are more meritocracy-oriented than the NCEE and the timings of these examinations usually overlap, it takes plenty of time and/or money to prepare for them and it is hard for students to take them in several different universities (You & Hu, 2013). If students pass these examinations they can be admitted with lower NCEE scores than the university’s normal admission threshold

(Wang, 2013). Although there are still certain quotas for universities in the Independent Enrolment Policy, the quotas in the Province-based Quota Admission Policy are not affected.

Regardless of whether students get extra points or not, most of them have to apply for universities through the Province-based Quota Admission Policy, which is the most important admission policy in China. It will be described in Section 4.2. If students fail to be included in a quota in the Province-based Quota Admission Policy, they might have to leave education or repeat the last year of high school and then take another NCEE the following year.

In summary, admission policies in China have various purposes and contents. Some aim to improve equity in education while others pay more attention to academic talent and productivity. This study does not intend to criticise or challenge the reasonableness of meritocratic selection such as in the Recommendation system or in the Independent Enrolment Policy. However, patterns of HE admissions show that there are underlying disparities in terms of family backgrounds, minority groups, residence and sex behind the selection procedure. These disparities have been discussed in the literature (Jia & Ericson, 2017; Wang, 2010; Wu, 2017). Nevertheless, because the main focus of this study is on how far the admission policies aiming to improve equity in education achieve their aims and how they can be modified to make greater contributions, the following sections mainly focus on the Province-based Quota Admission Policy, the Extra Credits Policy and the Three Special Plans.

4.2 The Province-based Quota Admission Policy

The Province-based Quota Admission Policy is the primary admission policy implemented in China. It allocates quotas of HE admission places to every province. Students then have to apply for HE in their hukou province. They are not allowed to take up quotas allocated to other provinces.

This province-based quota allocation has a long history. In the Ming and Qing dynasties, government officials were selected by means of examinations, but people in the south usually did much better than those in the north (Zhou, 2012). In order to make the government as inclusive as possible and encourage people in the north and other educationally less-developed areas to study hard, the emperors decided to distribute quotas for different areas.

HE itself was developed after the establishment of the PRC government. For nearly three decades, the government paid the tuition fees and the living expenses of HE students (Li & Min, 2001). In other words, HEIs were fully state-funded at that time. The government set admission quotas for HEIs and HEIs could not violate these. Due to the imbalanced development of different areas, the MOE of the new PRC government distributed some quotas to under-developed provinces to improve equality. Furthermore, the MOE stated that all the HEIs in provinces were required to submit their admission plans to the MOE and MOE would check these plans and confirm or reject them according to the number of candidates in the province (Zhou, 2012). After the 1970s, with the re-introduction of the NCEE, a more complete Province-based Quota Admission Policy was gradually framed. A report on college and university enrolments in 1977 (MOE, 1977) stated that HEIs were required to admit their students according to plans set by the government.

After the expansion of HE in China in the 1980s and 1990s, no more national plans were implemented for students who intended to enter HE or to support HE studies by paying tuition fees and living expenses. However, the province-based quotas were kept. HEIs are still expected to follow the quota plans set together by the government and HEIs themselves (MOE, 2020b; 2021b). HEIs have few extra admission quotas for certain provinces and they cannot transfer unfilled quotas from one province to another province either.

It remains unclear exactly how these admission quotas for each province are decided (Pan et al., 2010) and the Province-based Quota Admission Policy has been criticised for inequality. For example, the development of HE and basic education varies from province to province, but the quotas in the Province-based Quota Admission Policy are somewhat positively correlated with the education resources, especially in HE, that provinces possess. Therefore, the quota allocations are very unbalanced among provinces. The advantaged provinces are more likely to get more HE admission quotas, even in prestigious HEIs. Wang (2019) and Cheng (2020) call this accumulated advantage in privileged provinces “the Matthew Effect” in HE.

This study will evaluate whether the Province-based Quota Admission Policy is as equal as it intends to be and then discuss the extent to which the indicator that is used to identify the disadvantage of provinces is appropriate.

4.3 Extra Credits Policy

The Extra Credits Policy (jiafen policy) refers to projects that directly add points to raw NCEE scores or give one some priority in admission. After taking the NCEE, students get their raw score, one of the most widely-accepted certificates of talent and intelligence and the most significant entry pass to HE, about 20 days later. Most students then begin the application process according to their raw scores and rankings in their hukou province, but some have the chance to get additional ‘bonuses’ according to the Extra Credits Policy. These additional bonuses include priority admission, admission with a lower entry requirement and extra points added to raw NCEE scores (Zeng, 2018).

The Extra Credits Policy was first implemented in the 1950s (Li & Yang, 2011; Peng & Jin, 2015; Wang, 2018; Zeng, 2018). At that time, the MOE asked HEIs to prioritise applications by HE applicants

1) who had been industrial workers for more than three years;

2) who had been army cadres for more than three years;

3) who were overseas Chinese students; and

4) who belonged to fraternal minority ethnic groups.

The children of workers and peasants were the primary beneficiaries of this policy until 1978. Family backgrounds were not as indicative of privilege as before and were replaced with “students who came from rural areas or areas with fewer education resources” in 1978 and “students who were capable of sports, science and mathematics or were highly evaluated in morality” in the 1980s (SEC, 1987). Increasingly more students from rural and remote areas were able to attend HE with a lower requirement.

Since the first decade of the 21st century, however, the extra score system has become a more controversial public issue. Some politicians and scholars have criticised the violation of equity in implementing the Extra Score Policy, especially when aimed at improving efficiency (Guo, 2019; Peng & Jin, 2015). Therefore, the MOE and other government ministries required all provincial education departments to adjust or restrict the extra bonus points. Furthermore, in 2014, the MOE, along with the National Ethnic Affairs Commission, the Ministry of Public Security, the General Administration of Sport and the China Association for Science and Technology, published ‘Suggestions on Further Reducing and Standardising the Extra Scores in the NCEE,’ which announced that five categories of qualifications for extra scores would be cancelled at the national level from 1 January 2015 (MOE, NEAC, MPS & GASCAST, 2014). In 2018 and 2019, the MOE further restricted not only the qualifications for getting additional scores but also the size of scores that could be given. In 2020 and 2021, the latest regulations on HEI admission programmes were issued, which gave the newest information about the Extra Score Policy (MOE, 2020b; 2021b). According to the updated regulation, the following criteria remain valid and can result in up to 20 extra points:

1) extra scores for students who are the children of martyrs;

2) extra scores for students who belong to minority groups;

3) extra scores for students who are the children of overseas Chinese or who come back to China as overseas Chinese or are from Taiwan Province;

4) extra scores for students who are retired soldiers gaining second-class merit or granted an honorary title.

And the following criterion is valid for no more than 10 additional points:

1) extra scores for students who are retired soldiers and choose to seek employment by themselves

It is clear that not all these criteria aim at improving HE participation by disadvantaged students. Some of them, such as additional points for merit-worthy retired soldiers' children, are more like rewards, and some, such as extra credits for overseas students, serve as attractions. However, the extra credits for minority ethnic students take into consideration their disadvantages. Minority ethnicities have been regarded as less developed and more disadvantaged than the Han majority ethnicity for a long time, and they have been given extra scores in the transition to HE since the1950s (Chen, 2019).

There are some other criteria not for more points but for priority admission. However, these criteria are more reward-oriented such as priority admission for students with a parent who has worked in the armed forces and won second-class merit or higher in normal times or third-class merit or higher in wartime.

As this study mainly focuses on indicators of disadvantage, it does not evaluate other reward-oriented or attraction-oriented indicators but only those which have the purpose of identifying disadvantage. Therefore, the focused indicator in the Extra Credits Policy that will be further evaluated and discussed in this study is minority ethnicity.

4.4 Three Special Plans

Apart from the Province-based Quota Admission policy and the Extra Score policy, there are three special admission plan policies. They aim to improve HE participation for potentially disadvantaged students with rural hukou or from impoverished areas. These special plans include National Special Plans, Local Special Plans and the Special Programme for Colleges and Universities. The information on the Three Special Plans below is summarised from the website named 'Yang guang gaokao' (2022), the official MOE-affiliated website focusing on HE transitions.

4.4.1 National Special Plans

National Special Plans is a CA project to provide extra help during admission to HE for students who come from officially-designated 'poor prefectures,' or the four prefectures in the south of Xinjiang. Specifically, students who plan to apply for a quota in the National Special Plans are required to meet all three of the following criteria:

1) be qualified to take the NCEE in that year;

2) have valid hukou information that shows a) that the students' hukou places have been targeted prefectures for at least three successive years; b) that students' parents or statutory guardians' hukou places are targeted prefectures;

3) have registered in and studied at high schools in the targeted prefectures where their hukou places are for at least three successive years.

Students who meet these criteria can ask the Admission Department of the Local Education and Examination Agency for an application form. After filling out and submitting the application form, the students' information is checked to ensure that they really satisfy the criteria. Without the application or the examination, students are not allowed to join this project.

The HEIs that the MOE asks to issue National Special Plans are almost all nationally or provincially prestigious universities. Therefore, they provide students living in remote and less-developed areas with a good chance of getting access to elite HEIs. However, this does not necessarily mean that all students who take part in the National Special Plans are admitted by universities. Instead, although they do not need to compete for a place in HE with most of the other NCEE candidates in the same province through the Province-based Quota Policy, these students compete for a quota in the HEIs listed in the National Special Plans with other students qualified to take part in this project, which, however, is certainly a smaller group.

As the quotas in National Special Plans are differentiated from the ones in the Province-based Quota Policy, the admission wave for HE applicants to National Special Plans is not the first, second or third (if the province issue three waves) but one before all these, which is called the ‘beforehand wave’ (tiqianpi in Chinese). Usually, students do not need to take any other examinations but just the NCEE to acquire offers from universities once their qualification for the special plan has been confirmed. The students in this project are also ranked according to their NCEE scores, and they are then selected by prestigious universities according to their rankings. However, the threshold is typically not much lower (up to 20 points) than the general entry requirement for general applicants.

4.4.2 Local Special Plans

Local Special Plans is another CA policy project. Unlike National Special Plans, Local Special Plans aim to help rural hukou students in specific areas access good-quality HE. Therefore, there is one more criterion to apply for this project, which is rural hukou.

Furthermore, the universities that implement this project are provincially elite universities, which are less prestigious, although still selective, than those in the National Special Plans. These universities are asked to reserve at least as many as 3% of their regular quota places for applicants to Local Special Plans.

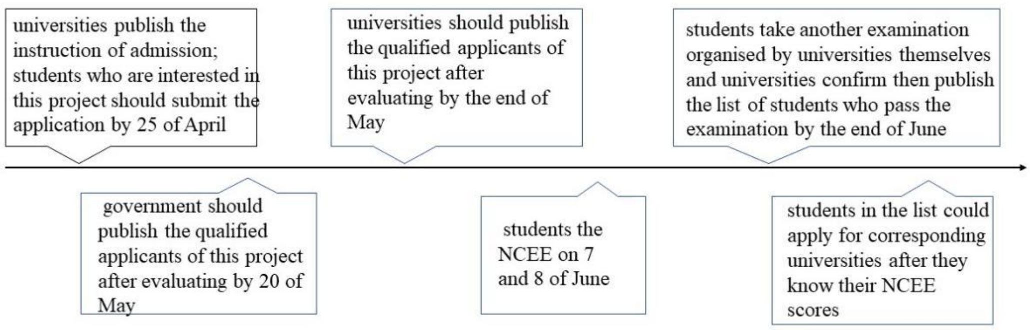
Similarly to National Special Plans, there is usually no additional examination when applying to Local Special Plans but just evaluation of the criteria for applicants. The admission wave for this project is also before the normal two (or three) waves.

4.4.3 Special Programme for Colleges and Universities

The Special Programme for Colleges and Universities is also a project for rural hukou students. Outstanding diligent rural hukou students from remote, impoverished or ethnic prefectures can apply for this project, but they also need to be evaluated. The evaluation criteria are similar to those for the other two special plans, but there is one more criterion: the hukou status of students and their parents must be rural. For the provinces/cities/districts where there is no distinction between rural hukou and urban hukou, students are required to differentiate between rural areas and urban areas according to official documents about area distributions published by the National Bureau of Statistics in order to confirm their backgrounds.

The HEIs that implement the Special Programme for Colleges and Universities are almost the same as those affiliated with the MOE. Some exceptions also enjoy high prestige. Appendix 1 provides a list of HEIs participating in this project in 2021. These universities are asked to reserve at least 2% of their regular admission quotas for undergraduates to accept qualified applicants from this project.

Figure 4.4.1 The process of implementing the Special Programme for Colleges and Universities



Unlike the first two projects, the Special Programme for Colleges and Universities requires applicants to take another examination in addition to the NCEE. Because of this, students are expected to apply to only a few universities to save time, costs and energy. The general process for this project is shown in Figure 4.4.1.

The admission wave for the Special Programme for Colleges and Universities is also before all the regular admission waves.

The differences between the three special plans are briefly listed in Table 4.4.1. These special plans have no fixed standard for the extra support, such as 10 or 20 points (but the upper limit is 20). Instead, they just reduce the number of competitors for disadvantaged students to a certain extent, as these students do not need to compete with all the other NCEE candidates in their provinces but only with other officially confirmed 'disadvantaged' students.

Table 4.4.1 The differences between the Three Special Plans

	National Special Plans	Local Special Plans	Special Programme for Colleges and Universities
Indicator	From officially designated poor prefectures	Rural hukou; from certain prefectures	Rural hukou; from certain prefectures
HEIs involved in this	Nationally selective	Provincially selective	MOE-affiliated universities

project	universities	universities	and other elite HEIs
Additional test	×	×	√

Because of the lack of sufficient valid prefecture data in datasets, this study is not able to do a robust examination of this indicator. Therefore, this study primarily evaluates the accuracy and reliability of the rural hukou indicator in the following chapters.

4.5 Discussion and Conclusion

This chapter has described admission policies implemented in China. It has shown how HE applicants can get an offer from HEIs through various routes and who Chinese CA policies aim to provide with additional support during the admission process. The chapter has presented a detailed discussion of the Province-based Quota Admission Policy, which aims to improve HE participation by students from under-developed provinces and in which the Hukou provincial status indicator becomes most relevant. This is further investigated in Chapter 10. Nowadays, other officially used CA indicators include minority ethnicities, rural hukou and poor prefectures. Because of data limits this study will focus on evaluating and discussing minority ethnicities and rural hukou as CA indicators.

The next chapter is a literature review. It is a snapshot of how CA policies, especially the indicators implemented in CA policies in other countries, are evaluated in research studies, including their advantages and shortcomings.

Chapter 5 Literature Review

This chapter summarises literature and research findings evaluating contextualised admission policies. Drawing on existing literature, the chapter focuses on the quality of the contextualised indicators in current CA policies. The discussion is a critical analysis of the accuracy, reliability and trustworthiness of CA indicators in identifying targeted students and of whether they are suitable for implementation in CA policies. It needs to be remembered that this brief literature review does not attempt to cover all of the countries in the world but is somewhat UK-centric because studies relevant to CA policies in some contexts do not answer the questions that this study tries to answer, and in some contexts studies involving CA policies are scarce.

5.1 Criticisms of Contextualised Admission Policies

Contextualised admission policies have been widely critiqued. One of the most frequently discussed concerns regards the effectiveness of these kinds of policy, including whether they have led to visible improvements in opportunities for HE participation by disadvantaged minorities. For instance, there are evaluations of affirmative actions in the US that try to explore whether the share of minority groups in HE has increased or not (Allen et al., 2002; Arcidiacono, Lovenheim & Zhu, 2015; Tierney, 1997).

Another criticism centres on chaos in the implementation of CA policies. As Downing et al. (2002) state, there is no consistent or standard way to implement affirmative action in the US. Some studies draw attention to the transparency of implemented CA policies. The difficulty universities have in accessing CA-relevant information makes potentially eligible students less likely to get a place in HE (Banerjee, 2018; Boliver & Crawford et al., 2017).

However, before discussing the issues mentioned above regarding contextualised admission policies, it is more important to discuss whether the groups targeted by these policies are really disadvantaged. In other words, it is important be clear whether the indicators employed in contextualised admission policies accurately and reliably identify disadvantaged students who need and deserve additional help. If the contextualised indicators mistakenly identify targets, such as advantaged students in minority groups, it could lead to more severe inequality in education and then also in social mobility (Howson, Cohen & Viola, 2022). Such advantaged students, despite belonging to a minority, might not only be undeserving beneficiaries of extra support but could also be barriers against disadvantaged students in non-minority groups getting access to HE. This is called the "creamy layer problem" in India (Frisancho & Krishna, 2016) and "false positives" in the UK (Boliver, Gorard & Siddiqui, 2015; Gorard et al., 2017).

Due to the different dimensions of social stratification in different societies, various contextual indicators are employed. For instance, the indicators in the UK focus on social class, taking into consideration the underlying relationships between socioeconomic status and educational achievement, while for historical reasons affirmative action indicators in the US pay more attention to race and sex issues (Dougherty & Callender, 2020).

As was shown in Chapter 4, in China contextual indicators are more likely to be aimed at redressing inequalities in HE participation related to location, ethnicity and hukou status. However, these indicators might be poor at identifying disadvantage. Some biographical indicators, for example, are not necessarily associated with disadvantage. On the other hand, some indicators have been employed for several decades, and the problems that led to these indicators being adopted might no longer be so serious. If so, supporting the targets identified by these indicators would be an obvious mistake.

Therefore, before implementing them, contextualised indicators should be evaluated to measure the extent to which they are able to identify disadvantaged students. The next section reviews literature evaluating contextualised indicators.

5.2 Limitations of the Indicators Currently Used in Contextualised Admission Policies

CA policies have been widely implemented in a variety of forms worldwide. The commendable intention is to identify and help the disadvantaged and minorities access HE. In order to achieve this aim, a first vital step is to accurately and correctly identify the targeted groups. CA policies can only be effective when the contextualised indicators they employ are effective (Boliver, Gorard & Siddiqui, 2017). Therefore, it is important to employ good contextualised indicators.

Schwartz (2004) reports the features of good-quality contextualised indicators: they are relevant, accurate and able to provide HE applicants with the chance to achieve their potential. Boliver & Gorard et al. (2017, p3) argue that good contextualised indicators need to at least be valid and reliable. They define 'valid' as appropriate and precise, and 'reliable' as trustworthy, accurate and consistent. Gorard et al. (2017) also hold that contextualised indicators must be accurate, reliable, accessible, complete, appropriate and trustworthy. However, many widely used indicators have been criticised for having limitations and drawbacks.

The fundamental quality of indicators employed in CA policies is an important issue. They need to be rigorous enough to identify the disadvantaged groups targeted. As the CA policies of few HEIs are based on research evidence (Banerjee, 2018), the indicators currently used in CA policies might be far from meeting this requirement.

For instance, while they admit the potential of CA policies to widen participation in HE, Boliver, Gorard & Siddiqui (2015) analyse secondary datasets and question the quality and effectiveness of indicators employed in the UK. They suggest that many individual-level indicators in CA policies in the UK have low validity because they employ unverified self-declared information such as being first generation in HE, parents having less-favoured occupations and low family income. This self-reported information is not reliable. Harrison

(2017) supports this argument. Harrison admits that it is possible for occupation data to indicate social class, but he is concerned about missing data, inaccuracy and the impossibility of coding self-reported information. Furthermore, although it is true that minimal parental educational qualifications might be strongly associated with other socio-economic characteristics, this indicator could identify a population that is too large (Jerrim, 2021).

However, even officially verified indicators, including Free School Meal (FSM) eligibility, Special Educational Needs (SEN) and Education Maintenance Allowance (EMA) eligibility, can also suffer from mis-targeting, although they have clearer definitions, greater reliability and more convenient access. Harrison (2017) points out that although FSM eligibility is a primary government-published indicator, it is not reliable except for students aged around 18. FSM eligibility, and even EverFSM6, which refers to "pupils on roll in January 2012 that are known to have been eligible for free school meals (FSM) on any pupil level census in the last six years (DfE, 2012), for example, is criticised as being unable to identify the "hidden poor" or previously poor (Treadaway, 2014). In addition, even though FSM eligibility has a moderate correlation (0.44) with low family income (Jerrim, 2021), it is still regarded as an unsatisfactory indicator because HEIs cannot access individual-level FSM data as these are regarded as sensitive information according to the General Data Protection Regulation. Llie, Sutherland and Vignoles (2017) conclude that FSM eligibility is inferior to parents' occupation or educational qualifications in predicting students' educational attainment. Finally, changeable criteria for FSM eligibility and confusing approaches to reporting FSM eligibility make this indicator contentious (Pickering, 2019).

Furthermore, although some indicators such as maturity, suffering from a chronic illness, sex and being a refugee are reliable and valid to a certain extent, there is a lack of evidence showing whether they are really associated with disadvantage (Gorard et al., 2017). Ethnicity, or race, a popular indicator, is not only self-reported and ambiguously defined but there is also a high risk of false positives. Although Trent et al. (2003) advocate utilising racial characteristics in affirmative action because there are significant gaps between white and minority groups in HE enrolment, there is evidence that additional support might be given to the wrong individuals. According to Frisancho and Krishna (2016), black students who receive support to participate in HE are usually from wealthier families, making the support less meaningful. Therefore, there have been appeals for changes in affirmative action holding that that targets would be better selected based on their economic backgrounds rather than race.

In addition, some individual-level indicators are only available for students who successfully transit to KS5 and who apply for HE but not for the more general student body (Gorard et al., 2019; Mountford-Zimdars & Moore, 2020).

On the other hand, aggregated indicators such as neighbourhood-level and school-level ones are even more problematic. POLAR and TUNDRA (tracking underrepresentation by area) do not have close correlations with low family income, which might mean that they are not related to less-advantaged socio-economic status (Jerrim, 2021). Jerrim praises ACORN as the best area-level indicator because of its higher correlation with low income and fewer false negative and false positive biases. However, fewer does not mean none, and all these area-based indicators are vulnerable to the ‘ecological fallacy’ and so wrongly help students from higher SES backgrounds or exclude students from lower SES backgrounds (Boliver, Gorard & Siddiqui, 2015; 2022; Riddell, Blackburn & Minty, 2013; Weedon, 2014). For instance, POLAR and ACORN wrongly identify nearly 90% and 75% respectively of students living in so-called disadvantaged areas who might not be genuinely disadvantaged (Boliver, Gorard & Siddiqui, 2019). Furthermore, they also false-negatively exclude a large number of truly disadvantaged students, such as BAME students, students living in a rented home, school leavers and students who do not live in lower participation areas (Boliver, Gorard & Siddiqui, 2019; Jerrim, 2021; Pickering, 2019; Sundorph, Vasilev & Coiffait, 2017). The doubtable validity of POLAR in accurately selecting students from low socio-economic backgrounds is also emphasised by Harrison (2017). In addition, POLAR is problematic as it uses previous patterns of transitions to HE to predict the current situation (Mountford-

Zimdars & Moore, 2020).

IDACI, as introduced in Chapter 3, another area-level indicator, is also far from reliable at identifying disadvantaged students as according to an analysis of secondary data many, even most, disadvantaged students do not live in the areas expected (Boliver, Gorard & Siddiqui, 2019; Gorard, Siddiqui & Boliver, 2017). This is also the case of SIMD 20 and SIMD 40 (see Chapter 3), two indicators used in Scotland (Weedon, 2014). Riddell, Blackburn & Minty (2013) suggest that neighbourhood-based indicators should be used in combination with other indicators such as NS-SEC and weakly performing schools. However, Weedon (2014) questions the accuracy, validity and completeness of NS-SEC information.

School-level indicators also have some problems, although they are more accessible than individual-level and area-level indicators. One of the most contentious school-level indicators is state schools. Despite solid support for contextualised admissions favouring students studying at state schools due to their high accessibility, this indicator does not necessarily refer to low socio-economic status (Weedon, 2014). Lasselle, McDougall-Bagnall and Smith (2014) also critique the use of school types in CA policies. Instead, they approve the use of school performance as an indicator to identify target students. A proponent of school-level indicators, Mullen (2011), holds that it is reasonable for HEIs to accept students from lower performing schools rather than ones from higher performing schools when there are two applicants with identical scores. He justifies this position by claiming that the former might have more potential to succeed academically. However, this justification might not be entirely acceptable. First, schools might be a factor in students' educational outcomes but they cannot be the only factor. Family background might be more associated with educational achievement. More importantly, school type and relative school contexts are not necessarily good predictors of disadvantage at the individual student level (Gorard, Siddiqui & Boliver, 2017).

A second limitation is missing data in CA indicator documents. It is hard to imagine that there is no data missing at all in the information collected. Even very high-quality official datasets suffer from missing cases or missing values for some variables (Gorard, 2021).

Some sensitive information, including family income and parental educational qualifications, which might be promising indicators if they are accurate and reliable enough, have too much missing data, so they cannot safely be used as CA indicators (Gorard et al., 2017). Moreover, Boliver, Gorard & Siddiqui (2015) even claim that FSM eligibility, one of the most promising indicators, which is clearly defined, somewhat good quality and easily accessible in official information, also has serious missing data. These missing data compromise the quality of FSM eligibility as a CA indicator. Worse still, the largest group of answers about ethnic origin is 'missing or unknown.' This missing data problem exists not only in self-reported individual indicators but also in aggregated ones such as school-level indicators, which some people regard as reliable.

Importantly, missing information needs careful attention regardless of whether the reason for it being missing is participants refusing to give information, vague or unclear recorded responses or loss of responses during data cleaning (Siddiqui, Boliver & Gorard, 2019). On the contrary, missing data such as missing FSM eligibility are likely to be missing for the disadvantaged group. Gorard, Siddiqui & Boliver (2017) look further into missing data in the National Pupil Database (NPD) and in Higher Education Statistics Agency (HESA) data and find that missing data such as missing FSM eligibility records, missing IDACI information, missing school type information and missing parental education information are closely associated with disadvantage so that it would be unjust to ignore this group.

A third limitation of contextualised indicators, or CA policies themselves, is that some of them are weak at identifying really disadvantaged individuals. According to Gorard et al. (2006), students in the UK who apply for HE are very likely to come from advantaged social classes so it is not surprising that advantaged groups are over-represented in HE participation. Furthermore, individual decisions on HE applications and participation are made much earlier than when current CA policies can cover them. One of the key factors influencing these decisions is prior attainment, which is also stratified by background information. Students without satisfactory prior attainment are more likely to leave education after the compulsory phase and even many of those who continue to KS5 can take other educational or vocational routes than participating in HE (Gorard, Siddiqui & Boliver, 2017). It can be said that choosing HE participation and previously continuing to KS5, which seems to be based on objective prior educational attainment, is also stratified by students' backgrounds (Gorard et al., 2017).

However, the current CA policies in the UK only concern HE applicants. The larger group of students who leave education at KS5 and do not apply for HE are not included in the process (Gorard et al., 2019). This will already lead to mismatched subjects. As Gorard et al. (2017; 2019) state, "contextualised admission focuses on those missing access by a few grades or points, not on the really challenged and disadvantaged in society at all."

5.3 Discussion and Conclusion

This chapter has briefly reviewed literature evaluating currently employed contextualised indicators, primarily in the UK. Instead of discussing the effectiveness or the implementation of contextualised admission policies in different HEIs, the short review has focused more on the quality of the contextualised indicators assessed in studies. CA policies can only be reasonable and effective when the indicators employed have high accuracy, reliability and availability.

The review suggests that many indicators might be less than satisfactory. First, some individual-level indicators, which should be the most accurate in identifying disadvantaged students, are not reliable enough as they are usually self-reported and unverified. Employing these indicators might lead to confusion or even unfairness.

On the other hand, even if we choose verified indicators such as FSM eligibility, they are also problematic. For example, FSM eligibility suffers from omission of previous or hidden poverty.

Furthermore, although they might be somewhat reliable and accessible, area-based and school-level indicators are also problematic. They are very vulnerable to ecological fallacies and can mis-target potential beneficiaries.

Fourth, some indicators are not necessarily associated with disadvantage in HE participation. The additional support provided to groups identified by these indicators will risk creating new inequities.

Another problem with contextualised indicators is missing data. Almost every indicator suffers from this shortcoming. The missing data cannot be viewed as missing at random because they are not, and there is even a close link between data being missing and disadvantage. However, missing data cannot be viewed as an indicator of disadvantaged groups in case it is a result of deliberate concealment.

Finally, most of the indicators discussed are only available for students who have successfully transited to KS5 and applied to HEIs. This group of students are likely to be the minority. In other words, current contextualised indicators can easily miss the majority of more disadvantaged or truly disadvantaged students.

To sum up, it is worth evaluating whether the currently used contextualised indicators and the popularly discussed indicators sufficiently identify the disadvantaged students. According to the review, in the English context the answer might not be ‘yes.’

Studies that evaluate the quality of contextualised indicators in the Chinese context are still not common, but they are worth conducting. Therefore, this study will try to answer the following research questions:

- What are the disparities in HE participation sin terms of the official indicators used in contextualised admission policies in China – province, hukou status, ethnicity?
- And in terms of other indicators such as sex, parental education and occupation?
- To what extent do each of these indicators accurately identify a disadvantaged group of students, and only those students?
- Do the indicators lead to fairer identification of disadvantaged students?
- To what extent does an accumulative disadvantage exist in transitions to HE?

The next part explains the methodology employed. It includes a description of the methods utilised in this study, an evaluation of the data sets employed and an introduction to the indices and formulas used.

Part 2 Methodology

Chapter 6 Introduction to the Methods Employed

This chapter presents the methods employed in this study in order to answer the research questions on disparities in admission patterns and access by disadvantaged students. These were a structured review, secondary data analysis and a supplementary small-scale cross-sectional survey, as Table 6.1 shows. This introduction explains how these methods were applied and why they were chosen.

Table 6.1 Summary of research design, data sources and methods of analysis

Research question	Research design	Data sources	Methods of analysis
What are the disparities in HE participation in China in terms of the official indicators used in contextualised admission policies – province, hukou status, ethnicity?	Structured review; Secondary data analysis.	Literature from four databases; CGSS; CFPS; Data from government website; Data from university website.	Percentages; Gorard segregation index; odds ratio.
And in terms of other indicators such as sex, parental education and occupation?	Structured review; Secondary data analysis.	Literature from four databases; CGSS; CFPS.	Percentages; cross-tabs; odds ratio.
To what extent do each of these indicators accurately identify a disadvantaged group of students, and only those students?	Secondary data analysis	CGSS; CFPS;	Cross-tabs; logistic regression; odds ratio.
Do the indicators lead to fairer identification of disadvantaged students?	Secondary data analysis; Cross-sectional survey.	CGSS; CFPS; Primary survey data.	Cross-tabs; logistic regression; odds ratio; Cohen's effect size.
To what extent does an accumulative disadvantage exist in transitions to HE?	Secondary data analysis.	CGSS; CFPS; Primary survey data.	Cross-tabs; logistic regression; odds ratio; Cohen's effect size.

6.1 Structured Review

A structured literature review was conducted to obtain a snapshot of the potential answers to the research questions about disparities in HE participation according to different indicators. This review collected documents as systematically and extensively as possible and evaluated them as objectively and critically as possible.

There were two stages in the search process. A primary search for research materials was implemented. This was concentrated in the first year of the Ph.D. This was followed by periodical supplementary search activities in order to make the search as inclusive as possible. The search process covered a wide range of contextual indicators of disadvantage such as hukou province, hukou status and ethnicity, and non-officially acknowledged but popularly discussed ones such as sex, parental education level, parental occupation and family income. These indicators were intended to help "build the learning 'trajectory' of individuals," which is related to their participation in HE or their decisions as to whether participating in HE was an option or not (Gorard et al., 2006).

The academic literature in four online databases was searched: China National Knowledge Infrastructure (CNKI), Google Scholar, ScienceDirect and ProQuest. CNKI is one of the most popular academic databases in China. It was included in the search process because this study investigates the Chinese context. Google Scholar covers millions of research documents in various academic databases, making it is easy to access potentially relevant literature. ScienceDirect and ProQuest were also involved in order to make the search for documents, especially dissertations and Ph.D. theses, as inclusive as possible.

The search syntax was revised for the different databases because some search terms would be too long for some databases. Furthermore, there would be a risk of missing some relevant literature if some words were deleted, so two search processes using two groups of search terms were conducted. The search syntax used was as follows:

Google Scholar and ProQuest:

((“higher”OR“tertiary”)AND“education”OR“universit”)AND(“equalit*”OR“fair*” OR“injustice”OR“unjust”OR“privileg*”OR“disadvantage*”OR“stratif*”OR“segregation*”OR“disparit*”)AND(“admission*”OR“enrol*”OR“participat*”)AND(“China”)*

(("higher"OR"tertiary")AND"education"OR"universit")AND("equalit*"OR"fair*"OR"injustice")AND(“enrol*”OR“admission*”)AND(“China”OR”Chinese”)AND(“CGSS”OR“CFPS”OR“CGS”)*

ScienceDirect (maximum 8 Boolean connectors):

(("higher"OR"tertiary")AND"education")AND("equalit"OR"fair")AND("admission"OR"enrol"OR"participat")AND("China”)

(("higher"OR"tertiary")AND"education"OR"universit")AND(“China”OR”Chinese”)AND(“CGSS”OR“CFPS”OR“CGS”)

CNKI:

(“中国”)AND(“高等教育”OR“大学”)AND(“录取”OR“入学”OR“招生”)AND(“公平”OR“差异”)

(“教育”)AND(“CGSS”)

(“教育”)AND(“CFPS”)

(“教育”)AND(“CGS”)

There were two other search conditions. 1) Sources had to be published from 2000 onwards, because HE policies were very different before this. Before the 1990s HEIs were likely to select their target students based on their social characteristics (‘chengfen’ in Chinese), while after the 1990s they started to select students from a meritocratic perspective. 2) they had to be written in English or Chinese. The results of the structured review will be presented in Chapter 9.

All of the collected documents were screened through title, key words, abstract at the first step. Then an extraction for more relevant documents have been conducted by reading the full texts. Finally, the left relevant documents are evaluated in terms of research quality and reliability with the help of 'sieve' deeloped by Gorard (2021). These studies were rated from 0 padlocks (the lowest quality) to 4 padlocks (the highest quality). The rating was based on five criteria: the research design (whether the design is able to answer the research questions), the scale of study (how large the sample size is), missing data (how many missing data the studies have), the quality of data obtained (how good the measurements for key variables are) and other threats to validity (e.g. conflict of interest). If the collected documents have not reported any issues among these five, they would be rated as of low quality. The ratings take no account of what the results and conclusions are like. Only studies rated with 2 padlocks (moderately trustworthy) and above are selected and introduced in Chapter 9.

The results of the studies in the final selection were classified and synthesised according to the indicators employed. These included hukou province, ethnicity, hukou status, sex, family SES including parents' education qualifications, parents' occupations, parents' political party memberships and family income, and other characteristics such as the number of siblings.

6.2 Secondary Data Analysis

Assessments of HE admissions based on secondary datasets clarified understanding of the clustering and segregation of HE participation in China. Analyses of secondary datasets could answer the research questions on the performance of officially acknowledged or potential contextualised indicators. The findings could provide further evidence of CA policies and their impact on HE access.

This study mainly used the following datasets:

1) Administrative datasets collected and published by official branches of the Chinese government;

2) University-level data on the student intakes of prestigious HEIs in China;

3) Social survey datasets, which provided micro-level information for the analyses.

The research questions on the feasibility and rigour of CA policies could yield better and higher quality results if individual-level information from the Population Census could be obtained and analysed. In the case of China, permission to access micro-level raw data from the Population Census is not possible. Therefore, there was no choice but to employ some substitute data. The shortcomings of these substitutes should be kept in mind. They will be evaluated in the next section.

Administrative data collected officially by the government exist at the aggregate level in the form of the Population Census, the Mini Census and the Micro Census. These census surveys were conducted by the National Bureau of Statistics (hereafter NBS) and the results are published on its website. These data were used to describe patterns in education participation in China in recent decades.

Analyses of national and provincial data were conducted, including numbers of NCEE candidates from 2016 to 2019 and numbers of primary/middle/senior high school students in selected age cohorts. The data series are publicly available on the websites of the provincial Education Bureaus, the Ministry of Education of the People's Republic of China and Sina Education.

Second, institution-level data, such as the admission quota plans of DFC universities allocated to each province from 2016 to 2019, were collected. These quota plans for various university subjects are usually published on the official websites of DFC universities a few months before the NCEE. Aggregate admission quotas for each prestigious university were manually counted by the author.

Finally, individual-level data were employed to explore the relationships between education participation and individual characteristics, which is a helpful way to avoid ecological fallacies. Two nationally representative social surveys with reasonably large sample sizes (over 10,000), namely the Chinese General Social Survey (CGSS) and the China Family Panel Study (CFPS), were analysed. The CGSS data were downloaded from the Chinese National Survey Data Archive, and those of the CFPS were obtained from Peking University Open Research Data.

Although these secondary datasets have several limitations, which will be evaluated in detail in the next chapter, they are still much better than ones that researchers can collect by themselves. The two datasets have much larger sample sizes and more informative variables, as they were collected and published by either government or authoritative organisations. Therefore, this study mainly conducted secondary data analysis rather than collecting primary data. Descriptive analysis and regression analysis are conducted and the detailed analysis plan are introduced in Chapter 7. However, as these datasets are somewhat outdated, the study also conducted a cross-sectional survey of middle school students. This primary data collection was for two reasons: to overcome variable limitations in the existing datasets and to hear some voices from recent reality. The survey is introduced in the next section.

6.3 Cross-sectional Survey

As was mentioned above, apart from analysing existing secondary data, a survey was also conducted with a sample of middle school students in order to investigate potential contextualised indicators. 18 classes in 8 middle schools were involved in the survey, although 12 classes in 6 schools were expected to participate at the beginning. As Table 6.3.1 shows, one of these middle schools was in Changsha, the capital city of Hunan province; one was in Huaihua, a city in the west of Hunan; four were in Sangzhi, an impoverished prefecture of Zhangjiajie city in Hunan; one was in Hengshui, a city in Hebei province, and one was in Lishui, a city in Zhejiang province.

Both Hunan and Hebei are competitive provinces with prominent NCEE candidate populations, but the former is richer in prestigious HEIs (see Chapter 2). Changsha, the capital of Hunan, enjoys the richest primary, secondary and HE institutions in Hunan. Huaihua is a mediumly developed city located in the southwest of Hunan. Sangzhi is a poor and less-developed prefecture in the northwest of Hunan which is affiliated with Zhangjiajie. Compared with Changsha, both Huaihua and Sangzhi, and particularly the latter, are more limited in high-quality education, as they have fewer selective schools and lower-qualified teachers.

Table 6.3.1 The distribution of samples in places

Places		Sample size (N)
Hunan	Changsha	38
	Sangzhi	500
	Huaihua	194
Hebei	Hengshui	65
Zhejiang	Lishui	5
Total (N)		802

Hebei is a province in the north of China with a larger student population but less prestigious HEIs than Hunan. Hengshui is a city located in the southeast of Hebei province, but it is not the capital. Although one of the best high schools in China is in Hengshui – Hengshui High School of Hebei – Hengshui does not have other outstanding properties in terms of educational resources.

Zhejiang is one of the wealthiest provinces in China, but it lacks prestigious HEIs, as Chapter 2 showed. There are only a few DFC, Project 985, Project 211 and MOE-affiliated universities in Zhejiang province, while its neighbours, Jiangsu and Shanghai, enjoy many more elite HEIs. Furthermore, every 10,000 NCEE candidates in Zhejiang compete for places in 0.03 DFC universities, whereas the corresponding figures in Jiangsu and Shanghai are 0.06

and 0.8 respectively. It might be possible that students from this province have relatively affluent family backgrounds because of the well-developed economy in Zhejiang compared with their counterparts living in less developed provinces. However, they still suffer from fierce competition for HEI places, as they have to compete for a quota place in DFC universities with students from their provinces. However, the cases collected in this province are too few. If more samples could be collected from Zhejiang, we might be able to see some interesting findings in this province.

All the students involved in the survey were in the third (last) year of middle school when they answered the questionnaire. The questionnaire was delivered in either a paper version or as an online link. It might be better to use just one version. However, given that some students in less-developed places might not have internet access to complete the questionnaire, it was delivered in both ways. Specifically, students in four middle schools in Sangzhi and those in one school in Huaihua answered paper questionnaires, while the rest used the online version. The paper questionnaire was delivered on regular school days so that a maximum number of students could participate, while the online survey option was available for teachers to use as an additional way to increase student participation.

The questions in the two versions are the same. The only difference is that in a question asking how much three factors affected students' plans after leaving middle school and at age 18 the order of the three factors is different. Nevertheless, this did not affect the analysis as the answers were carefully recorded.

The paper version was delivered to 700 students and 694 questionnaires were returned, a response rate of 99%. 110 responses were received from the online group. After deleting some repeated responses (two online interviewees mistakenly submitted their answers twice), there were 802 respondents to the survey. The distribution of their characteristics is reported in Chapter 7. Chinese and English versions of the questionnaire can be found in Appendix 2.

Apart from the questionnaire, there were also follow-up interviews for 12 students. These interviews were designed to collect students' opinions on the fairness of currently used CA indicators in Chinese higher education admissions.

The survey was not foreseen in the initial proposal but was added as a follow-up study in the third year of this Ph.D. project. The reason for conducting this additional survey was related to some preliminary findings from the analysis of CFPS secondary data. It was clear that inequalities in HE participation start long before the HE admission process and the NCEE. These inequalities become visible from the end of compulsory education (middle school education in China). Therefore, it is possible that more disadvantaged students who had the potential and aspiration to attend HE but needed more academic support abandoned education at this stage and did not even get to participate in the NCEE. If so, not only the indicators of potential talent in disadvantaged groups but also the currently implemented CA policies could be problematic and less accurately identify targets.

On the other hand, this study tries to put an eye in the reality rather than only in secondary datasets which might be published some years ago. Therefore, a survey was designed to help understand transition patterns and educational exclusion and inclusion at an early stage. The data from this survey are useful in that they provide information on this educational stage and possible links with HE entry, but as the sample is not considerable and non-representative robust conclusions cannot be drawn. The survey only indicates potential gaps in students' aspirations for HE and high school education, and their expectations of academic achievement.

Furthermore, the survey also collected information on the students' perceptions of official CA indicators, namely rural hukou and minority ethnicity. The findings show differences in students' perceptions and awareness, and the extent to which they see contextual indicators as fair. To sum up, this primary survey was not intended to provide any definitive conclusions. The analysis plan for this survey is described in the next chapter.

6.4 Discussion and Conclusion

This chapter has introduced the primary methods employed in this study to answer the research questions: a structured review, secondary data analysis and a small-scale survey of targeted groups. As secondary data analysis was the most important way to answer the research questions and it would be unrealistic to assume the datasets employed are perfect, it is essential to discuss their qualities. Therefore, the next chapter will evaluate the datasets used in this study.

Chapter 7 Introduction to the Datasets

This chapter looks at the three categories of data collected: government administrative data, institutional data and survey data. Fundamental information about them is first provided and then their strengths and shortcomings for the purposes of this study are assessed.

7.1 Administrative Data

7.1.1 Population Census/Micro Census/Mini Census

Basic introduction.

The Population Census survey is conducted each decade using standardised means of collecting and recording information. Although it was first undertaken in 1953, the Population Census was not formally used by the Chinese government as a systematic survey until 1994. Later it was decided to carry out the census in years ending in 0 (NBS, 2021). There have been seven waves of the Population Census: 1953, 1964, 1982, 1990, 2000, 2010 and 2020.

The Population Census surveys citizens living in mainland China, which includes four municipalities directly under the control of the central government (hereafter municipalities), five ethnic minority autonomous regions and 22 provinces, but excludes Hong Kong, Macao and Taiwan. The census collects information such as citizens' age, sex, education level, hukou identity, hukou province, occupation and other personal information which is not only useful for governing but valuable for social science research. Some critical information can be calculated from the raw data such as the population size, density, fertility, mortality, natural population growth rate, household sizes, sex composition, and age composition. The results are aggregated and summarised at the national or provincial level and then published as *Chinese Census Data* on the NBS website.

The Micro Census is also conducted by the NBS. This is a sample survey of 1% of the population conducted once a decade, always in years ending with 5. The sampling process uses multi-level and multi-stage proportional cohort sampling methods and views the whole country as the first unit and provinces as the second one.

The Mini Census, another sample survey conducted by the NBS, is carried out annually except for the years when the Population Census or the Micro Census is conducted. The Mini Census uses similar sampling methods to the Micro Census, and its sample size is around 1% of the total population.

Both the Micro Census and the Mini Census collect demographic information similar to that collected in the Population Census. The results are analysed and summarised at the macro provincial or nationwide level. The results of these smaller-scale censuses are published in the *China Statistics Yearbook* on the NBS website.

Limitations of census data.

Despite the potentially high quality of periodic administrative datasets, for the purposes of this study there are some challenges in using Population Census, Micro Census and Mini Census data. First, the response rates are not published along with the data. Although these censuses are conducted by the government and so are more likely to be responded to as completely as possible than other national surveys made by non-government institutions, it is still unrealistic to assume they are complete.

Second, and more importantly, all the census publications only display macro-level rather than micro-level data. This means that definitive arguments based on the collective information from censuses cannot be made if we wish to avoid ecological fallacies. The published national- and provincial-level data might only be safe to be used in descriptive analyses. For example, some researchers use these census data to compare changes in the annual gross student intakes of all HEIs in China (Zheng & Sun, 2017). Therefore, this study does not conduct any causal or correlational analysis but only a descriptive analysis to demonstrate the clustering and tendencies in education participation. Nevertheless, it should be noted that for social science researchers, this lack of micro-individual-level data reduces the usefulness and value of these censuses.

To sum up, the Population Census, the Micro Census and the Mini Census might be high-quality administrative datasets, but the lack of individual-level data is detrimental to their utility in some social science research, or at least for this study. Therefore, this study will not use these censuses as the main data source but only as supplements on occasion when there is a need for some description or evidence of nationwide macro-situations or changes, such as tendencies in the development of HE and high school education.

7.1.2 Other Government-published Data

Basic introduction.

Other important aggregated data are the numbers of NCEE candidates in each province. NCEE candidates take the NCEE on the 7th and 8th of June. As was explained in Chapter 4, only when students complete their applications to take the NCEE can they get permission to take part in this important examination. The numbers of NCEE candidates are collected by the government and then published at the provincial level.

One thing needs some explanation. The numbers of NCEE candidates reported in each province include both fresh high school graduates and students who failed in the previous NCEE and had to repeat the last year of high school and take another NCEE. Therefore, the number of NCEE candidates can be greater than that of high school graduates in some years.

Other aggregated administrative data are the student intakes of primary schools/middle schools/high schools (high schools here only refers to ordinary academic high schools attended after compulsory education which provide their students with academic knowledge rather than practical skills, while high school education in China officially includes ordinary academic high schools, vocational high schools, technical high schools, adult high schools and ordinary specialised high schools).

Limitations of aggregated data.

There are some limitations of the above aggregated data that might require more attention. First, there are some missing data on NCEE candidates in some provinces for some years. For example, there is no information about NCEE candidates in Tibet and Xinjiang in 2019. The complete list of numbers of NCEE candidates can be found in Appendix 3.

In addition to missing data, the successfully collected data are also far from perfect as they do not indicate students who actually took the NCEE in the end but only those who applied to take this examination. It is hard to imagine that the numbers of these two groups of students are completely the same. Some students might not actually take the NCEE even though they registered.

Furthermore, because not all information on NCEE candidates could be found on official government websites some was collected from Sina Education (2020). This data might not be as accurate or authoritative as the primary/middle/high school student intakes published by the MOE.

More importantly, both the numbers of NCEE candidates and the student intakes are province-level data, which cannot be used to support micro-level arguments. This should be kept in mind when analysing, reporting and concluding based on these data.

7.2 Institutional Data

Basic introduction.

The institutional data employed in this study are mainly the admission quota plans of elite HEIs in China. Elite HEIs here refers to the 42 HEIs in the Double First Class (DFC) project. According to the province-based quota admission policy, the government and HEIs decide together how many students these elite universities should, can and have to admit from each province. These decisions for different university subjects, such as Law, English, Finance, Applied Mathematics, Medicine and Politics, are often published on the official websites of the HEIs in the summer. They indicate the enrolment plans of these elite HEIs at the provincial level.

As mentioned above, these admission quota plans can be checked and collected from the DFC universities' websites, which announce important admission information and news. These admission plans involve students majoring in Liberal Arts (wenke in Chinese), Science (like in Chinese) and Art (yishusheng in Chinese) but exclude those who enjoy the privilege of being admitted through other programmes such as the Independent Recruitment Programme and those who are directly enrolled by universities without taking the NCEE (baosong in Chinese). In addition, because since 2017 high school students in Shanghai and Zhejiang have not had to choose only Liberal Arts or Science majors but have been able to choose separate subjects that they prefer (or are good at) to take the examination in, the 2017, 2018 and 2019 data for these two regions are sums of all the subjects.

Besides the admission quota plans of prestigious HEIs, the actual student intakes of all regular HEIs in every province were also collected. Here again regular HEIs are to be understood as HEIs that are qualified to offer bachelor's degree programmes but not independent colleges, adult colleges or vocational colleges. The data on provincial admission numbers were retrieved from official websites of provincial governments, provincial educational departments and Sina Education.

Limitations of institutional data.

One of the most serious limitations of these institutional data is missing data. The collection of admission quota plans for DFC HEIs was intended to cover all 42 HEIs from 2016 to 2019. However, not surprisingly, some DFC HEIs, such as Tsinghua University, did not publish their admission plans online, so it was not possible to obtain that information. In addition, despite the aim to collect data from 2016 to 2019 in reality data are not always available for all four years. For instance, Peking University only published its admission plan in 2016, and Remin University of China only in 2019. Furthermore, the regular HEI student intake data also lack enrolment data for some provinces for some years. More detailed information on this problem is provided in Appendix 4.

Second, these institutional data are not as accurate or reliable as government administrative data. For example, admission quota plans only reflect government admission plans rather than the actual intakes of DFC HEIs in the year in question. It is true that Chinese HEIs are highly political. Most of them are administered by the government by means of imperative government documents such as notifications, plans and regulations. All of the DFC HEI admission plans are required to be sent to the Ministry of Education for government approval before they are published. Otherwise, they cannot be implemented (MOE, 2021). During the implementation the HEIs are also asked to not violate the plans. However, it is still not true that there is no difference between the numbers in reality and in the plans. In fact, according to the regulation, HEIs are allowed to prepare extra reserve quota plans, but the number of these extra quotas must not exceed 1% of the number in the original approved plans.

Furthermore, as few DFC universities publish their admission quotas as total numbers, the author manually added the figures. Although this was done very carefully, it cannot be assumed that no mistakes were made during the collection and calculation.

The student intakes to regular HEIs are also problematic, because of inaccuracy. While they do indicate actual enrolments instead of a plan, this does not mean that they represent the real numbers of students who finally register and attend HEIs. Instead, the figures only show how many students theoretically get places according to their scores in the NCEE. The figures ignore the possibility that some students are not satisfied with their admission results so they refuse their offers, repeat the last year of high school and then take another NCEE the following year. Besides this limitation, not all data on student intakes can be collected from government websites, which again casts doubt about their accuracy.

Moreover, as both admission quota plans and student intakes are aggregate information rather than individual-level information, it is not possible to identify advantaged and disadvantaged individuals from the data regardless of which indicators are used to distinguish them. Hence, analysis of these data can only focus on the macro-level, such as differences between provinces and between geographical areas.

7.3 Social Survey Data

7.3.1 Chinese General Social Survey (CGSS)

Basic introduction.

The Chinese General Social Survey (CGSS) is the first comprehensive academic nationwide survey project in China. It has been conducted by the National Survey Research Centre (NSRC) at Renmin University of China since 2003. In 2008 CGSS finished its first phase, which consisted of five separate survey waves, and in 2010 it started its second phase, which also consists of five separate survey waves.

Despite some slight differences among the three sampling designs in 2003-2006 and the 2008 and 2010 waves, CGSS mainly adopted multi-stage stratification PPS random sampling in 22 provinces, 4 municipalities and 5 minority ethnicity autonomous regions in China (Tibet was only included in the CGSS after 2010), excluding Hong Kong, Macau and Taiwan (CGSS, 2021). From 2003 to 2006, the CGSS selected 125 counties/districts as its primary sampling units (PSU), and then from them selected 500 towns/streets. Within these towns/streets, 1,000 village communities (cunweihui in Chinese)/neighbourhood communities (juweihui in Chinese) were chosen, and finally around 10,000 people were randomly picked from these communities. In the 2008 survey, the sample size shrank to 6,000 people from 600 village communities/neighbourhood communities, 300 towns/streets and 100 counties/districts. However, at the beginning of the second phase of the CGSS in 2010, the sample size was enlarged to 12,000 cases from 480 village communities/neighbourhood communities, 100 counties/districts and 5 metropolitan areas, namely Beijing, Shanghai, Guangzhou, Shenzhen and Tianjin. The CGSS sampling process is reasonably random, and the sample size is adequate.

As for the group targeted, the CGSS includes citizens aged 18 to 69 as potential respondents, with few exceptions. That is to say, the dataset does not cover students who are being educated in primary schools, junior high schools or even senior high schools, as in China 18 is always the theoretical age for students to complete senior high school and pursue higher education.

The CGSS response rates are not bad. Even with the most complete data, surveys can still miss about 12% of their cases (Gorard & See, 2013). The response rates in 2012 and 2013 were reasonable: 71.5% and 72.17% respectively (CNSDA, 2021). See Table 7.3.1. However, the response rates for the surveys conducted in 2015 and 2017 have not been published.

Table 7.3.1 CGSS Response Rates

CGSS Wave	2012	2013	2015	2017
Response Rate	71.5%	72.17%	——	——

Apart from sampling and response rates, another crucial element is the variables. Although the main focus of the CGSS is not education but the intrinsic relationship between social structure and people's quality of life (CGSS, 2021), it does inquire about many variables which are informative for educational research. For example, the CGSS asks respondents about their sex, hukou, ethnicity, age, residence, SES, parental education level and parental occupations. Variables which are often examined by educational researchers or used as indicators of disadvantage.

Limitations of the CGSS.

Sample ages

For the purpose of this study, there are some limitations of the CGSS datasets that need to be discussed before further analyses. First, the CGSS samples are aged from 18 to 69. As explained earlier, students in primary, junior high and senior high schools are excluded. This might be less of a problem for studies focusing on educational returns, such as social or employment returns from higher education, but it is less helpful for studies which examine inequalities before or during the HE transition. Disadvantages behind education inequalities might be accumulated since the early years of education. These disadvantages not only tend to be negatively associated with students' educational outcomes, which are essential criteria for HE admission, but might also increase the probability of students abandoning HE or even high school education. The limited age group in the CGSS might not be able to reveal this earlier disadvantage.

Missing data

The second problem is missing data. First, the response rate. Admittedly, as a nationwide survey the CGSS response rates are not very low, although they are not outstanding either, only slightly higher than 70% in 2012 and 2013. However, the 30% missing samples should be addressed, as 30% is considerably larger than many disadvantaged groups found in research (Gorard & See, 2013). Worse still, the response rates of the 2015 and 2017 CGSS have not been published (CNSDA, 2021).

Table 7.3.2 Percentages of missing data for some essential variables in the CGSS (%)

	2012	2013	2015	2017
Sex	0	0	0	0
Ethnicity	0.1	0.1	0.2	0
Education Attainment	0.1	0.0	0.2	0
Hukou Status	0.0	0.0	0.0	0
Father's Education	4	3.2	7.6	6.6
Father's Job Status	1	1.2	3.8	1.8
Father's Workplace	8.3	8.9	12.9	11.8
Mother's Education	2.5	1.8	6.2	4
Mother's Job Status	0.6	0.7	3.7	1.4
Mother's Workplace	18.9	20.9	22.6	23

Note: 0 means that there is no missing value in this variable; 0.0 means that the proportion of missing value in this variable is smaller than 0.01.

On the other hand, the proportions of missing values in the CGSS of essential variables analysed in this study are reasonable (see Table 7.3.2). Most commonly used variables such as sex, ethnicity, education attainment and hukou status have less than 1% of values missing. Nevertheless, the SES-related variables such as parents' workplaces, which means whether the parents worked for a state-owned business, the government, a public institution, a private enterprise or were self-employed, have much higher missing rates, up to 23%. This more considerable number of missing values for variables could lead to potential biases in results.

Variable deficiencies

Third, there are some limitations in the CGSS variables. The most remarkable shortage, for instance, is a lack of school-level data. In the CGSS, there is little information about school types or education outcomes. Analysts cannot know if the interviewees attended a public school, a private school or a migrated-child-segregated school and neither can they know how well they did at school. The only information on the interviewees' education is their latest educational qualifications and when they obtained it.

Moreover, there is also nothing about the level of prestige of the HEIs which the interviewees attended. That is, it remains unclear whether a person went to a DFC university, a lower-level first-tier university, a second-tier university or a third-tier university. This information is significant as disadvantaged students might be more likely to lag behind in the intakes of prestigious HEIs (Lucas, 2001).

In addition, as the CGSS is not a longitudinal study some crucial information about the interviewees is only a snapshot of the past rather than tracing them. A typical example is family economic status. There is a question about family income in the CGSS, but it seems that the questionnaire designers were only interested in the family income of respondents in the year before they completed the questionnaire. For respondents who were in their thirties, their answers on family income in the previous year cannot be regarded as indicating the family's financial situation in the period when they were eligible to go into HE. Therefore, other information indicating the family's economic situation is needed. Some substitutes are questions about social class, such as "Compared to other people living in the same area, at which economic level do you think your family is located?" and "At which economic level do you think your family is/was located now/ ten years ago/ when you were 14 years old/will be ten years in the future?" These questions seem more time-relevant, but they might be more open to bias than ones asking directly about family income. Respondents only needed to estimate the amount when they were asked about income, but for the questions comparing their family's economic level with that of other people they had to know not only their family's economic situation but also that of others, which could be more complicated.

Another problematic variable in the CGSS is language ability. There are questions asking "How do you evaluate your ability to speak Mandarin?" and "How do you evaluate your ability to understand Mandarin?" However, language abilities are not easy to measure. Even the results of official language examinations can be inaccurate, let alone self-evaluations. It will be much harder for the result to be objective than that of the traditionally used indicator for language, namely the first language. Note that this does not necessarily mean that an indicator of the first language is better than one of language ability at identifying disadvantaged students.

The variables indicating parental occupations are also problematic. The relevant questions are "Where does your father/mother work?" rather than "What is your father/mother's job?" Although the latter might also be misunderstood, the former could be more troublesome as workplaces are not necessarily an accurate indicator of occupation. For example, if a respondent's parent works as a cleaner in a government agency, they might answer "My parent's workplace is a government agency," which might lead them to be mistakenly classified as advantaged, as there is no information about family income in the

CGSS. On the other hand, it is also possible that individuals are mistakenly viewed as disadvantaged when they answer that their parent is ‘self-employed,’ but he/she actually runs a company by himself/herself.

Finally, the ‘sex’ variable is only binary in the CGSS, which might hide potentially disadvantaged transsex or other individuals. For this reason, the term of “sex”, rather than more sociologically-related term of “gender”, is used in this study.

In conclusion, although the CGSS has some weaknesses concerning the group targeted and missing data and variables, it is still a valuable dataset for this study because of its large sample size and informative variables. However, this does not mean that the aforementioned drawbacks can be neglected. In this study the 2012, 2013, 2015 and 2017 CGSSs are merged and simple descriptive analysis and regression analysis are employed. The results are provided in Chapters 11, 12, 13 and 14.

CGSS Analysis.

Sampling and selection

First, the 2012, 2013, 2015 and 2017 CGSS data were first merged into one file, which produced a more extensive dataset with 46,753 cases. As the present study only focuses on the period after HE expansion in China, i.e. after around 1999, cases who theoretically should have attended HE in or after 1999 were selected for further analysis. Given that the usual age that students go into HE in China is 18 or 19, the remaining cases were those born in and after 1980. After deleting the unselected cases, a new dataset with 10,089 cases remained. All the CGSS analyses were conducted with this group of cases.

Descriptive analysis

The first analysis of the CGSS cases was descriptive. In order to obtain a simple snapshot of disparities in HE participation, different characteristics of these cases were examined, including both officially recognised indicators, such as hukou status and ethnicity, and popularly used indicators such as sex, social class, parental educational attainment, parental workplace, language ability and birth month (Table 7.3.3).

However, first it is necessary to explain why the 'disability' variable was not examined. Although disability and special educational needs (SEN) are regarded as important contextual indicators in some countries, disabled students in China usually study in 'Special Education' schools rather than in regular schools (MOE, 1998). In other words, these students are in a different education system to the one that this study investigates. Although disabled students might be more disadvantaged and it is unfair to structurally exclude them from the regular education system, this problem is beyond the scope of the present thesis.

Table 7.3.3 The CGSS variables examined

	Variables examined
Currently used indicator group	hukou status, ethnicity
Biographical indicator group	sex, birth month
Family-level indicator group	parental educational qualifications, parents' workplaces, social class
Others	ability to understand Mandarin, ability to speak Mandarin

Cross-tabulation was then used to explore the relationships among these indicators, which might reveal double or even triple disadvantages.

Furthermore, as the CGSS is a cross-sectional survey, there is no information on the trajectories of the cases' education histories. Therefore, the relationship between the above indicators and the cases' education levels was investigated.

Regression analysis

Besides the descriptive analysis, regression analyses of the CGSS data were performed. Because the outcome variable, whether the individual went into HE or not, is binary and many predictors are categorical (Gorard, 2021), binary logistic regression was used here. However, as the distribution of values in the outcome variable is very unbalanced with proportions of nearly 20:80 (2,097 HE participants and 7,986 non-HE participants; 6 cases lacking this variable were deleted), it would be hard to increase the prediction percentages for such an unbalanced base model. Therefore, a series of logistic regressions were run on the original smaller group of cases (2,097 HE participants) and successive randomly-chosen subgroups from the larger group of cases (2,097 randomly chosen from the 7,986 non-HE participants) with the same number of cases as the smaller group. The resulting base model thus had proportions of 50:50 and 4,194 cases. This process was run ten times.

In each model, there were 3 blocks of predictors (Table 7.3.4). Block 1 in the first model primarily contained hukou status, ethnicity and sex; the second block contained family-level indicators such as parental education, parental workplaces and social class; block 3 added the abilities to understand and speak Mandarin (the complete outputs of the ten regression estimates of Model 1 are provided in Appendix 10).

Table 7.3.4 The predictors in each block in CGSS Regression Model 1 and CGSS Regression Model 2

	Model 1	Model 2
Outcome	HE participation or not	
Base	-	-
Block 1	sex, ethnicity, hukou status	parents' educational qualifications, mother's workplace, father's workplace, social class
Block 2	parents' educational qualifications, mother's workplace, father's workplace, social class	ability to understand Mandarin, ability to speak Mandarin
Block 3	ability to understand Mandarin, ability to speak Mandarin	sex, ethnicity, hukou status

In Model 2, in order to see whether family-level indicators contribute more to correctly predicting percentages, these indicators (parental education qualifications, parents' workplaces and social class information) were put in the regression first. This is because compared to hukou status, ethnicity and sex, family-level indicators are somewhat less likely to be subject to an 'ecological fallacy.' If these predictors can make significant differences, they might be more appropriate indicators to identify the disadvantaged. Hukou status, ethnicity and sex were moved into the last block to see how the percentage variations change (the complete outputs of the ten regression estimates of Model 2 are provided in Appendix

10).

Furthermore, because some values only involve very few cases, which might distort the results of regressions, they were re-coded as other valid values of the variable. For example, in the mother's workplace variable, only one interviewee answered 'army' so this answer was changed into the 'party/government agency' value. This change was based on the observation that the respondent with an 'army' mother was more likely to go into HE, and so were those with a mother working for a 'party/government agency.' In addition, cases with missing hukou status, ethnicity or language abilities were also changed to the value with the most cases in these variables.

Both Model 1 and Model 2 primarily employed the enter method because the final regression results needed to be the averages of ten basic regressions. It would be problematic if there were some exclusions of predictors as a result of employing the forward stepwise (conditional) method. However, the forward stepwise (conditional) method was used just as a weak reminder of which indicators might be less significant in predicting HE participation.

7.3.2 China Family Panel Study (CFPS)

Basic Introduction.

The China Family Panel Study (CFPS) was a nationally representative longitudinal study conducted by the Social Science Survey at Peking University (Xie & Hu, 2014). The CFPS was an interdisciplinary survey with a large sample size for general purposes. After a baseline survey in 2010, there were five other follow-up survey waves in 2011, 2012, 2014, 2016 and 2018.

CFPS employed a multistage probability proportional sampling method (PPS) in which the primary sampling unit (PSU) was the administrative district/prefecture, the second-stage sampling unit was the neighbourhood community/administrative village and the final sampling unit was the household. In the CFPS the household, an important concept referring to "an economically independent dwelling unit" (Xie & Hu, 2014, p.9) which is partly viewed as a substitute for the family in China, was considered the basic sampling unit. All members with Chinese nationality (excluding Hong Kong, Macao and Taiwan) who were economically interdependent immediate relatives and who were not relatives but were economically related and had been living in the household for at least three months were interviewed.

During the sampling process, rather than classifying cases as rural hukou residents and urban hukou residents beforehand, the CFPS considered the whole household to be an entity and gathered information about their hukou identity. The sampling process covered 25 provinces/municipalities/autonomous regions in mainland China, which represent around 94.5% of the Chinese population. The CFPS claimed to be nationally representative.

The target CFPS sample size was 16,000 households, of which half were planned to be selected from five excessively sampled 'large' provinces – Shanghai, Liaoning, Henan, Gansu and Guangdong – and the other half were planned to be selected from 20 other regions. In the end, the CFPS completed interviews with 42,590 individuals in 14,960 households, 33,600 adults and 8,990 children (under 16). As the final sampling unit was the household, the CFPS response rate was 93.5%.

As an interdisciplinary social survey, the CFPS has a rich range of variables in order to provide up-to-date timely information for social science research. Some of its variables which are useful for this study are individual characteristics such as sex, ethnic group, hukou status, birth month and cognitive ability and family characteristics such as parents' political party memberships, parents' occupation categories, parents' education levels, family education aspirations, annual family income and annual education expenditure. Partly making up for the limitations of the CGSS, CFPS also collected rich school-level information, such as school type, class type and scores in tests. These rich variables might be able to provide a broader and deeper understanding of the relationship between these characteristics and HE participation or educational attainment.

Limitations of the CFPS.

Missing data

The CFPS also has some limitations. First, as has already been mentioned, missing data can never be ignored as it is unlikely that they are missing by chance. Table 7.3.5. and Table 7.3.6 show the numbers of missing cases and the missing response rates for some significant variables for this study. Note that this discussion of missing cases and responses is based on this study's finally selected sample of 4,881 cases, instead of the original sample of 42,590 cases. The selection process will be described in detail in the next section.

As Table 7.3.5 shows, compared with the selected baseline sample of 4,881 cases, nearly a quarter of the cases could not be traced in the second and third CFPS waves. In the fourth wave in 2016, 35% of the original cases were missing. Finally, in 2018 only half the cases were left. Instead of directly deleting the missing cases, they were analysed as below.

Table 7.3.5 Missing cases in each CFPS wave

	2010	2012	2014	2016	2018
Sample size	4,881	3,731	3,552	3,171	2,489
Missing cases	-	1,150	1,329	1,710	2,392
Missing rate	-	23.6%	27.2%	35%	49%

Figure 7.3.1 and Figure 7.3.2 show how missing cases with different Hukou status and parents' educational qualifications are distributed. It is not difficult to see that missing cases tend to have rural hukou and have more poorly educated parents. This tendency might indicate that missing data are not randomly missing but instead in a pattern that missing cases are closely associated with disadvantages.

However, attention needs to be paid to cases which missed four waves, which means that they only completed the baseline survey. They are more likely to have urban hukou and have well-educated parents. This group shows an opposite trend to the groups that are missing in one, two or three waves. It might be that this seemingly privileged group had gone abroad and were hard to get in touch with. More comparisons of other characteristics of missing cases in the CFPS are made in Appendix 11.

Another kind of missing data is missing values. Table 7.3.6 lists the percentages of missing values of some crucial variables analysed in this study. The percentages only reflect the missing rates of these variables in the baseline survey, as these were used in further analyses. Basic information such as sex, ethnicity and hukou status are rarely missing. These are relatively clearly defined. However, the variables for parents' occupations, family income, school type and class type, which might be more private, have lower response rates.

Figure 7.3.1 The percentages of cases dropping out in successive CFPS waves by hukou status

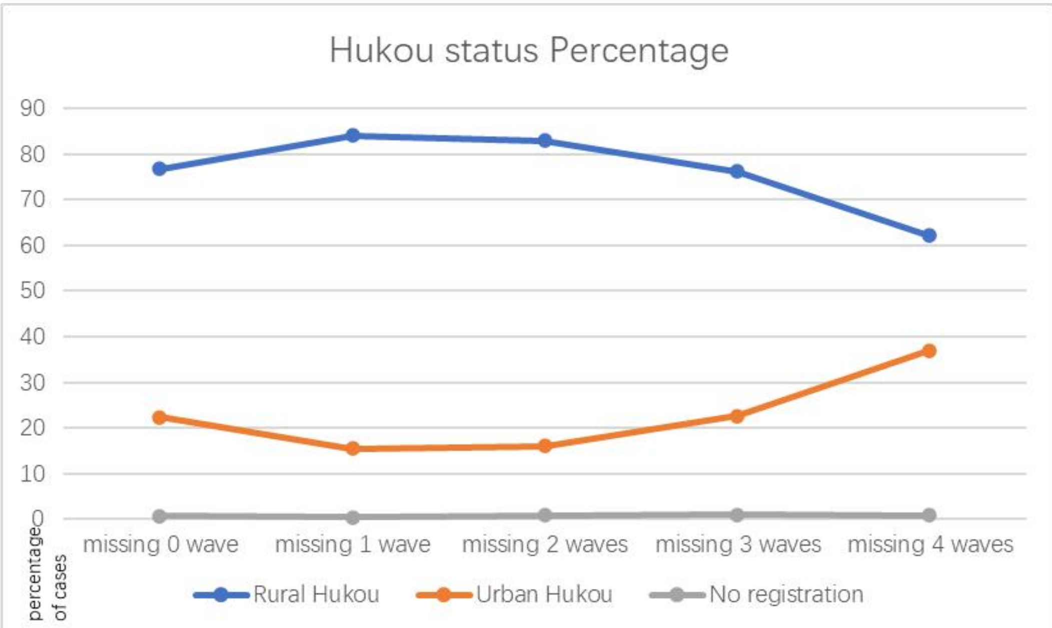
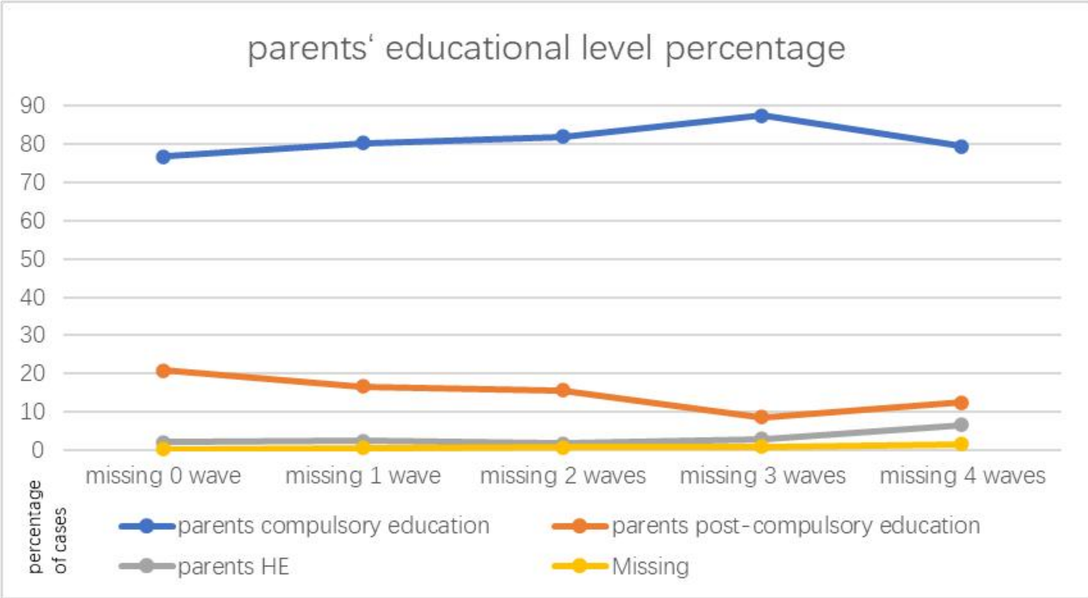


Figure 7.3.2 The percentages of cases dropping out in successive CFPS waves by parents' education levels



Again, missing data can never be assumed to be random, and they deserve more attention. These missing values will be analysed in Chapter 16 instead of being neglected and directly deleted.

Table 7.3.6 Percentages of missing responses for essential variables in the CFPS baseline survey

	2010
Sex	0
Birth month	0.43
Ethnicity	0.32
Hukou Status	0.14
Living place	0
Hukou province	0.31
Father’s education qualification	3.13
Mother’s education qualification	2.7
Father’s occupation	16.84
Mother’s occupation	24.44
Fathers’ political party memberships	6
Mother’s political party membership	6.02
Family income	45.5
Annual education expenditure	8.28
Daily language (family)	1.31
Daily language (school)	10.45
School type	14.38
Class type	14.61

Cognitive word test score	2.13
Cognitive mathematics test score	2.13
Grade of Chinese language test in last exam	20.3
Grade of math in last exam	21.7

Variable deficiencies

Second, despite there being a rich range of variables, there are some limitations. For instance, there is no mention of their scores in the Senior High School Entrance Examination or in the NCEE. Rather than varying from school to school these scores are standardised and so are relatively comparable at least at the province level, although they might not be nationally comparable. More importantly, the scores in these two examinations are very important, as they are fundamental factors in the transitions to high school and HE. Therefore, scores in the two examinations could be more useful than the binary classification of participating or not. As CFPS was a longitudinal survey and was conducted for a decade, it should have been able to gather this information. But there is no such information in the dataset, and the lack of these data somewhat reduces the advantages of the CFPS.

Besides these limitations, the values for sex are binary as in the CGSS, which might create bias against other possible sex groups. On the other hand, the daily language for communication, which might be easier to measure than self-evaluated language ability, can also be problematic. For example, the daily language might not indicate ability at writing or reading Mandarin.

In summary, the CFPS is a good quality social survey with reasonable sample sizes and abundant variables. Although there are some limitations, such as missing cases and missing values, it still is a valuable resource for this study and other social science research.

Preparation for the Analysis.

Sample selection

It is necessary to provide some brief information about the sample selection process. As has been mentioned, the CFPS data consist of four separate datasets updated in 2011, 2012, 2014, 2016 and 2018 after the baseline survey in 2010: datasets on children, parents, families and communities. The analyses in this study employ the first three of these datasets.

First, the child and adult datasets for each of the years were merged into one and the new larger dataset had 42,590 cases. These cases were then matched with their Family ID in the family dataset in order to access the family economic information of individuals.

As the target group in the present study is interviewees aged 10 to 18 in the baseline survey, the cases in this age range were selected. This age range was decided on to ensure that by the fifth survey wave in 2018, all the selected cases could theoretically have been admitted to HEIs so that differences in HE participation between a theoretical estimation and reality could be seen. After this first selection, the age-relevant sample contained 4,881 cases.

The next step was to clean the sample. The cases for whom information on HE participation was successfully collected were selected as the final sample for analysis, consisting of 2,490 cases.

Descriptive analysis

The descriptive analysis of the CFPS data is similar to that of the CGSS data. First, comparisons of HE participation were made between different indicators. Then, the relationships between these indicators were further explored. More indicators were examined than in the CGSS data, as the variables are more abundant in the CFPS and hopefully more accurately defined and measurable. The variables chosen were categorised in four groups: currently used indicators, biographical indicators, family-level indicators and school-level indicators. The data for these indicators were mostly chosen from the 2010 baseline survey. Table 7.3.7 provides detailed information about the indicators in each group.

Table 7.3.7 The variables in each indicator group

	Variables included
Currently-used indicator group	hukou status, hukou province, ethnicity

Biographical indicator group	sex, birth month
Family-level indicator group	parental education qualifications, parents' occupations, parents' political party memberships, family income (numerical), education expenditure (numerical), communication language at home
School-level indicator group	school type, class type, Chinese test score (numerical), mathematics test score (numerical), cognitive word test score (numerical), cognitive mathematics test score (numerical), communication language at school

Regression analysis

A series of binary logistic regression analyses were also carried out with the CFPS cases. Due to the unbalanced distribution of outcomes, with 610 HE participants and 1880 HE non-participants, the same method employed for the CGSS data was used. Finally, there were ten new sub-data sets with 1,220 cases (610 HE participants and 610 HE non-participants) in each.

Three models predicting HE participation each consisting of four blocks of indicators were analysed (Table 7.3.8). In Model 1 Block 1 contained biographical indicators: sex, hukou status, ethnicity and birth month; Block 2 contained family-level indicators: parents' education qualifications, parents' occupations, parents' political party memberships, communication language at home, annual family income and annual education expenditure; Block 3 contained school-level indicators which are not present in the CGSS data such as school type, class type, communication language at school and cognitive test scores in word and mathematics; and Block 4 focused on neighbourhood indicators: living area and Hukou province area.

Table 7.3.8 The predictors in each block in CFPS Regression Models 1, 2 and 3

	Model 1	Model 2	Model 3
Outcome	HE participation		
Base	-		
Block 1	sex, hukou, ethnicity, birth month	parents' educational qualifications, father's occupation, mother's occupation, parents' political party memberships, communication	school type, class type, communication language at school, word cognitive test score, maths cognitive test score

		language at home, annual family income, annual education expenditure	
Block 2	parents' education qualifications, father's occupation, mother's occupation, parents' political party memberships, communication language at home, annual family income, annual education expenditure	sex, hukou, ethnicity, birth month	sex, hukou, ethnicity, birth month
Block 3	school type, class type, communication language at school, word cognitive test score, maths cognitive test score	school type, class type, communication language at school, word cognitive test score, maths cognitive test score	parents' educational qualifications, father's occupation, mother's occupation, parents' political party memberships, communication language at home, annual family income, annual education expenditure
Block 4	living area, hukou province area	living area, hukou province area	living area, hukou province area

Table 7.3.9 The predictors in each block in the CFPS regression models predicting high school participation

	Model 1	Model 2	Model 3
Outcome	High school education participation		
Base	-		
Block 1	sex, hukou, ethnicity, birth month	school type, class type, communication language at schools, word cognitive test score, maths cognitive test score	parents' educational qualifications, father's occupation, mother's occupation, parents' political party memberships, communication language at home, annual family income, annual education expenditure

Block 2	parents' educational qualifications, father's occupation, mother's occupation, parents' political party memberships, communication language at home, annual family income, annual educational expenditure	parents' educational qualifications, father's occupation, mother's occupation, parents' political party memberships, communication language at home, annual family income, annual education expenditure	sex, hukou, ethnicity, birth month
Block 3	school type, class type, communication language at school, word cognitive test score, maths cognitive test score	sex, hukou, ethnicity, birth month	school type, class type, communication language at schools, word cognitive test score, maths cognitive test score
Block 4	living area, hukou province area	living area, hukou province area	living area, hukou province area

In Model 2, family-level indicators were brought forward to the first block as the aim was to see the extent to which these indicators are essential for HE participation. In Model 3, school-level indicators were moved to Block 1.

These regression analyses also employed the enter method, as the results needed to be averaged from the ten estimates.

As the CFPS is a longitudinal dataset, previous education experiences also needed to be explored. Therefore, in order to see the importance of high school participation for HE participation, a binary variable indicating whether or not the student attended ordinary high school was added to the regression. Furthermore, another regression analysis predicting high school education participation was also conducted (see Table 7.3.9).

As the outcome distribution here was more balanced, with only 55.7 percent predicted correctly by the base model, a simple binary logistic regression could directly analyse the sample. The orders of adding the predictors are summarised in Table 7.3.9. The method employed was the forward stepwise (conditional) method.

7.4 Primary Survey Data

7.4.1 Sample Characteristics

The sample size achieved in the primary survey was 804, and 802 valid cases could be included in the analysis after cleaning the data. In order to avoid tiny values in the analysis as much as possible, 4 cases who chose option 3 ‘other’ in the question on hukou status were randomly re-coded as rural hukou (2 cases) and urban hukou (2 cases). Similarly, 2 cases who answered ‘other’ as their ethnicity type were re-coded as han ethnicity (1 case) and a minority ethnicity (1 case).

Table 7.4.1 shows the characteristics of the final sample. The personal characteristics asked about in the survey were sex, ethnicity, hukou status, parental education qualifications and parental occupations. The four characteristics other than sex have either been used as official indicators of disadvantage or are somewhat associated with disadvantaged groups. Therefore, they were investigated in the survey.

Table 7.4.1 Characteristics of the survey sample

Characteristics		Size (N)	Proportion (%)
Sex	Male	412	51.4
	Female	390	48.6
Ethnicity	Han ethnicity	304	37.9
	Minority ethnicity	498	62.1
Hukou status	Urban hukou	127	15.8
	Rural hukou	675	84.2
Highest qualification of parents	Bachelor’s degree or above	70	8.7
	Vocational college qualification	58	7.2
	High school qualification or equivalent	223	27.8
	Compulsory education qualification or below	422	52.6
	Parental education qualification missing	29	3.6
Parents’ occupations	Working for a government department/institution	51	6.3
	Professional worker such as teacher, lawyer, doctor etc.	31	3.9

	Skilled worker such as gardener etc.	413	51.5
	Agriculture-related worker	73	9.1
	Staying at home and not employed	32	4.0
	Self-employed as street peddler or similar	40	5.0
	Self-employed running a company or similar	96	12.0
	Other/missing	66	8.2
Total		802	1

There is a reasonable distribution between the sexes. However, the distributions of ethnicities and hukou statuses are somewhat unexpected. Compared with the 7th Population Census (NBS, 2021) result, minority ethnicities and rural hukou cases are over-represented in the survey sample as they account for 62% and 84% respectively (while the corresponding percentages in the 7th Population Census results are 8.89% and 36.11%). These over-representations might be caused by the areas involved in the survey, as many of them are inhabited by minority ethnic groups and/or are rural areas.

Nevertheless, this does not necessarily mean that the sample in this survey is problematic. On the contrary, the over-representation might be helpful in some respects. First, there was no intention to generalise the survey results to all middle school students in China, which would be impossible. The purpose of conducting the survey was to access more up-to-date student voices about their plans for after completing compulsory education and after being 18 years old, and about their opinions on the currently-used contextualised admission indicators. There was no ambition to generalise the results to all students in China. Therefore, it is not a serious problem that the sample in the survey is not nationally representative. On the other hand, rural hukou students and minority ethnicity students are officially labelled disadvantaged but have not been paid particular attention in previous large-scale surveys. This survey focused more on them and allowed us to know their thoughts.

More than half of the students in the sample reported that their parents' highest educational qualification was only compulsory education (primary and middle school completed). Only 28% reported that their parents obtained the qualification necessary for high school education and only 7% and 9% of them reported that their parents achieved the necessary qualification for vocational college education and HE respectively. It seems that the students interviewed in the survey lack good home education backgrounds. However, caution is needed when interpreting parental information because students' knowledge about their parents' qualifications might be limited or inaccurate.

As the students reported, half of them had parents who worked as skilled workers. Around 15% of them have parents working as self-employed company managers or self-employed street peddlers. 10% of them stated that their parents did agricultural work and nearly 5% of them reported that their parents worked for government agencies or institutions. This latter group is generally regarded in China as having advantaged parental occupations. It is necessary to clarify that self-employed parents working as street peddlers are different from self-employed ones who manage a company because the former will be less financially stable and profitable. Again, students' knowledge and perceptions of their parent's jobs could be different from reality. The analysis and interpretation of findings based on students' reported information about their parents will be dealt with carefully.

To sum up, the predominant groups in the sample are traditionally regarded disadvantaged ones: rural hukou, minority ethnicities, those with less-educated parents and lower family SES.

7.4.2 Descriptive Analysis

A simple descriptive analysis was conducted to see whether there were differences in terms of aspirations for post-compulsory education among the students involved in the survey and how these differences are distributed by student characteristics such as sex, ethnicity, hukou status, parental educational qualifications, parents' occupations and students' self-evaluations of their school work. As there is little information in the secondary datasets about the prestige levels of the HEIs that students attended, this survey asked what kind of HEIs and high schools the middle school students wanted to go to in order to obtain a picture of this.

Besides aspirations for post-compulsory education, the factors that the students thought affected their education plans were also explored. Effect size was used to standardise the differences.

In addition, the students' understandings of the fairness of currently employed CA indicators were compared. The results show the extent to which these middle school students, who might be less advantaged, think currently employed CA indicators fairly and effectively select people in need.

7.4.3 Regression Analysis

Regression analyses were used to predict the students' aspirations for high school education and for HE.

Model 1 predicted HE aspirations. As Table 7.4.2 shows, Block 1 used sex, ethnicity and hukou status as predictors to explain the outcome of HE aspirations or not. Although they were self-reported in the survey, these three indicators are theoretically officially recorded information in huji registrations (government records including date of birth, hukou status, marriage status, ethnicity, sex and family address). Unlike other self-reported characteristics, these three have comparatively clear legal definitions.

Block 2 contained the highest parental educational qualification and parents' occupations. These two variables have been shown to have notable associations with children's educational aspirations but they are less reliable and mainly self-reported rather than officially verifiable.

Block 3 added students' self-evaluations of their school performance. Although self-evaluations might not accurately estimate educational achievements, and it is highly possible that they are associated with variables in the first two blocks, they are still regarded as predictors in this model because of the potential strong link between self-evaluation and educational aspirations. However, as self-evaluation is easily affected by other factors, it is added in Block 3.

Block 4 is the binary variable of whether students reported high school education aspirations. Models 2, 3 and 4 changed the order of blocks.

Table 7.4.2 The predictors in each block in the survey data regression models (HE aspirations)

	Model 1	Model 2	Model 3	Model 4
Outcome	HE aspiration			
Base	-	-		
Block 1	sex, ethnicity, hukou status	parental education qualifications, parents' occupations	parental education qualifications, parents' occupations	High school aspiration or not
Block 2	parental education qualifications, parents' occupations	sex, ethnicity, hukou status	sex, ethnicity, hukou status	parental education qualifications, parents' occupations
Block 3	self-evaluation of school performance	self-evaluation of school performance	self-evaluation of school performance	sex, ethnicity, hukou status
Block 4			High school aspiration or not	self-evaluation of school performance

Because the values in the outcome variable distribute evenly, there is no need for additional treatments. In order to keep the model as simple and parsimonious as possible and for it to be clear which variable(s) is less important, it was run with the forward stepwise (conditional) method entering the variables in each block. However, the enter method was also employed to get the coefficients of all the variables so that the gaps between all the characteristics could be known.

Another regression analysis predicted the students' aspirations for high school (Table 7.4.3). Because the distributions of the values in the outcome variable are unbalanced, with 76% of the students reporting that they wanted high school education and 24% of students not wanting, the regression was conducted ten times with new sub-data sets. The way the new sub-data sets were obtained was the same as that explained in the last few sections. The smaller group that did not show an aspiration for high school education contained 193 students (24%), while the larger group that reported an aspiration for high school education contained 609 students (76%). A comparison group was created of 193 cases randomly selected from the larger group to match the smaller one, satisfying the basic condition for a binary logistic regression with proportions of 50:50. This randomisation was repeated ten times, and then ten sub-data sets with 386 individual cases were created and analysed.

Table 7.4.3 The predictors in each block in the survey data regression models (high school aspirations)

	Model 1	Model 2
Outcome	High school aspiration	
Base	-	-
Block 1	sex, ethnicity, hukou status	parental education qualifications, parents' occupations
Block 2	parental education qualifications, parents' occupations	sex, ethnicity, hukou status
Block 3	self-evaluation of school performance	self-evaluation of school performance

The variables included in this regression model are similar to those in Table 7.4.2. As the final result was the average of ten regressions of randomised sub-data sets, it was necessary to cover all the scores. Therefore, the results obtained via the enter method rather than the forward stepwise (conditional) method will be provided.

7.5 Discussion and Conclusion

This chapter has provided basic information on and the limitations of the datasets in this study and how they were used. In short, administrative government aggregated data was mainly used to describe tendencies in the development of HE. Institutional university-level data were analysed to explore disparities in HE enrolment among provinces. Individual-level social survey data was used in more detailed analysis to examine the quality of the indicators currently used in Chinese contextualised admission policies and of popularly discussed potential indicators. Finally, the data obtained in a primary survey were used as supplementary data to try to get an idea about recent reality.

The next chapter introduces the indices and formulas used in the analyses in this study.

Chapter 8 Introduction to the Indices and Formulas

This chapter introduces the important indices used in the study. These are the Admission Rates Index, the Gorard Segregation Index, the Admission Opportunity Index and the “effect” size. They are used to compare disparities in HE admissions between provinces and between groups with different characteristics, such as hukou status and ethnic identity.

All of these indices are well known and widely employed by researchers. This chapter does not explain how they work but only why they have been chosen to be used. In short, the first three indices are used to compare differences in HE participation between provinces, while the effect size is used to demonstrate differences between groups with fewer categorises, such as sex groups and hukou status groups in survey datasets.

8.1 The Admission Rates Index

One of the most popular and widely used ways to examine HE equity is to compare HE admission rates for different groups (Hamnett, Hua & Liang, 2019; Zhang & Liu, 2019). Admission rates are simple to understand and easily calculated and the rates are more comparable (Cao & Zhang, 2016) than raw enrolment numbers. This study also employs this simple index to examine inequalities in HE participation and to answer the research questions of how much HE participation disparities exist in terms of provincial and geographical areas, ethnic groups, sex groups and other characteristics.

The following is the formula for the admission rate index.

$$ARi = Ai/Ci,$$

where:

ARi is the admission rate for group i (e.g. Hubei province, a minority ethnic group or male students, etc.);

Ai is the number of students from group i admitted by the universities studied (e.g. DFC HEIs or regular HEIs);

Ci is the population size of group i (e.g. the number of NCEE candidates in Hubei province, the overall size of a minority ethnic group or the number of male students, etc.).

Due to the de-centralisation of HEIs, they have to diversify their funding sources and they have started to absorb investments by local governments and local enterprises (Wang, X., 2001). In return, they allocate larger quotas for local students (Gao, 2017; Ji & Zhu, 2011; Yao, 2008; Wang & Zhang, 2014). However, unlike province-affiliated HEIs, most DFC universities are affiliated with ministries of the central government and get a large amount of their funding from this source, which means all taxpayers nationwide, every year. They are therefore expected to enrol students fairly from across the nation rather than having locally biased intakes (Liu & Li, 2014). It would be less acceptable for DFC universities to have extremely localised admissions. Therefore, this study explores whether there was a localised tendency in the admission plans of DFC universities from 2016 to 2019, and furthermore how far this local bias went. The localisation level is calculated as follows:

$$Lqi = Aqi/Aq,$$

where:

Lqi is the percentage of local admissions in university q in province i;

Aqi is the number of students from province i admitted by university q;

Aq is the number of university q's total admissions.

After collecting provincial and institutional data, the provincial admission rates of elite and regular HEIs were calculated using these formulas. These are compared in Chapter 10. The degree of localisation in admissions to DFC universities was also calculated.

However, it is important to keep in mind that the admission rate index is a somewhat poor indicator of equity of HE admissions, especially in provincial terms as the index does not take into account the broader population, including the total numbers of HE admissions and NCEE candidates, which to a certain extent undermines its rigor.

8.2 The Gorard Segregation Index

Because of the limitations of the admission rate index, a more rigorous index is required to reveal differences in HE admissions between provinces. For this purpose, the Gorard Segregation Index (GS index) was used.

Segregation means an uneven distribution of different student characteristics such as race, social class or first language between organisational units (Gorard & Taylor, 2002). For example, racial minority students in the US (Kenty-Drane, 2009) and FSM or SEN students in England (Gorard et al., 2013) are often found to be segregated between schools. In this study, segregation is an uneven distribution of HE admissions between different provinces in China.

In addition to the Gorard Segregation (GS) index, another index for calculating segregation levels is the Dissimilarity index. Both are popularly used in research, but the latter "represents the proportion of one group or other that would have to move, if there were no segregation", while the former "indicates the exact proportion of disadvantaged pupils who would have to move schools for there to be no segregation" (Bartholo & Costa, 2016, p.505). Furthermore, the GS index overcomes the limitations of inadequate composition invariance and less stability depending on the size of the groups (Gorard, 2007; Taylor, Gorard & Fitz, 2000). The GS index is calculated using the following formula:

$$GS=0.5*(\sum|Fi/F - Ni/N|)$$

The GS index has been widely used to reveal the segregation level between or among schools (Bartholo, 2013; Bartholo & Costa, 2014; Siddiqui, 2017). The GS index is a suitable index to explore unevenness in the distribution of potentially disadvantaged students/groups from some regions in a group of units (Gorard & See, 2013). Therefore, this study employs the GS index to compare disparities in HE participation between provinces. However, as this study ignores the absolute sign in the calculation of the GS Index, aiming to establish whether students from different provinces are over- or under-represented in targeted HEIs or in HE in general, the sum sign is no need. Therefore, the study actually employs GS Ratio and the formula is:

$$\text{GS Ratio (HE participation in China)} = 0.5*(Ai/A - Ci/C),$$

where:

Ai is the number of students in the admission quota plan of one university, for DFC universities, or the number of students who really participate in HE, for regular universities, in province i;

A is the total number of students in the admission quota plans of all DFC universities, or the total number of students who really participate in regular universities;

Ci is the number of NCEE candidates in a certain year in province i;

C is the total number of NCEE candidates in this year in China.

An additional explanation is needed here. When the GS Ratio is zero, it indicates that there is no segregation of students from the province in HE/the university, and educational resources are evenly distributed. In other words, HE equality and equity are possibly achieved, or at least are more likely to be achieved, among provinces in HE/the university. However, where there is a positive or negative segregation trend it indicates over-representation or under-representation, which means students from these provinces should be moved away or be accepted more by HE/the university in order to realise HE equity. Furthermore, when the size of the GS Ratio is further from zero, the segregation level and so education inequality is more serious.

8.3 The Admission Opportunity Rate Index

A third index employed to explore disparities in HE admissions between provinces is called the Admission Opportunity Rate index. Because compulsory education in China is only primary school and middle school education, competition for HE actually occurs much earlier than the NCEE. It starts at least at the point in which students complete compulsory education and transit to high school education. Those who fail to enter high school can hardly be accepted into HE. That is, it would be problematic if only NCEE candidates are taken into account, as they are already the winners of the semi-final rather than competitors at the beginning. Therefore, the Admission Opportunity Rate index, which takes earlier education experiences into consideration, is required.

The ideal figure to indicate all the competitors for HE in China is the size of the 18-year-old population in the year studied, as 18 is the typical age to start HE in China, but this information is not published. A substitute is needed. Some scholars use the number of graduates from primary schools (Zhang & Li, 2019) or from middle schools (Wang & Du, 2013) but these are not accurate enough due to ignorance of completion rates. On the other hand, these calculations could not provide information about the educational trajectory that help find the transition gaps among provinces. In this study, both the enrolment rates and graduation rates of high schools, middle schools and primary schools were collected, and all six rates combined with the admission rate of HEIs are used to calculate the Admission Opportunity Rate index. The formula for this is:

$$AOi=ARPi \times GRPi \times ARMi \times GRMi \times ARHi \times NRHi \times ARUi,$$

where:

AOi is the admission opportunity rate index of students in province i;

ARPi is the admission rate of primary schools in province i;

GRPi is the graduation rates of primary schools in province i;

ARMi is the admission rate of middle schools in province i;

GRMi is the graduation rate of middle schools in province i;

ARHi is the admission rates of high schools in province i;

NRHi is the rate of high school students applying to take the NCEE in province i;

ARUi is the admission rate of HEIs in province i.

As was mentioned above, attending high school is important because students have to complete high school education to be eligible to enter HE. Ignoring this would be unfair. On the other hand, although primary school and middle school education are compulsory, which might be one of the reasons why the enrolment rates of these two levels of education are not markedly different between provinces, they are included in case there are errors or biases due to exclusion. For example, some well-developed provinces might have a perfect enrolment of children in compulsory education, while some poorly-developed provinces might show a much worse picture with fewer children enrolled in primary and middle schools and more children dropping out before graduation. If primary and middle school graduates are not included in the calculation, drop-outs from these groups will be ignored.

However, it must be admitted that there is an inappropriate assumption in the formula that there is no mobility of students between groups. Although it is rare for there to be large-scale student mobility in education in China due to hukou restrictions, there will still be some mobility. For instance, some migrant workers' children go to Beijing, Shanghai and Guangdong, three of the most developed areas in China, with their working parents without local hukou and are accepted into compulsory education there. However, many of them who want to go to high school and take the NCEE are more likely to return to their hometown provinces as one of the conditions for taking the NCEE is registering in one's hukou province (MOE, 2021). This kind of mobility might result in an underestimation of enrolments in high schools in migratory provinces such as Beijing, Shanghai and Guangdong. At the same time, it can also lead to over-estimating enrolments in high schools in immigratory provinces, i.e. migrants' children's hometown provinces.

The target years selected to examine provincial disparities in HE admission in this study are 2016, 2017, 2018 and 2019, and the corresponding years when these groups started high school, graduated from middle school, started middle school, graduated from primary school and started primary school are listed in Table 8.3.1.

Table 8.3.1 The theoretical education schedule of the targeted sample groups

Starting HE	2016	2017	2018	2019
Taking NCEE	2016	2017	2018	2019
Starting high school	2013 (missing)	2014	2015	2016
Graduating from middle school	2013	2014	2015	2016
Starting middle school	2010 (missing)	2011	2012	2013
Graduating from primary school	2010	2011	2012	2013
Starting primary school	2004	2005	2006	2007

However, because the enrolment rates of middle schools in 2010 and those of high schools in 2013 could not be found, it is impossible to finish comparing the data for 2016. Therefore, this study only completed the analysis with the Admission Opportunity Rates index for 2017, 2018 and 2019. The full results can be found in Chapter 10.

8.4 Effect size

Although its name is somewhat misleading, effect size does not refer to any ‘effect’ or causation (Gorard, 2013; Gorard, 2021). Instead, it is a scaled measure to reveal the differences between two groups. There are two kinds of effect sizes: simple, or absolute, effect size; and standardised effect size (Baguley, 2009; Berben et al., 2012; Gorard, 2013; Gorard, 2021; Sullivan & Feinn, 2012). The former is mainly used in studies with the same measurements and the same scale, and the latter serves in studies that involve measurements without intrinsic meanings or with different scales (Sullivan & Feinn, 2012). This study primarily employs standardised effect sizes in order to make the results more comparable.

Cohen’s d

One of the commonly used effect sizes is Cohens’s d. The formula to calculate it is as follows (Berben et al., 2012; Gorard, 2021):

Effect size = (M1i-M2i)/SDpooled

where:

M1i is Group 1's mean score for the compared variable i;

M2i is Group 2's mean score for the compared variable i;

SDpooled is the pooled standard deviations for Group 1 and Group 2.

It can be seen from the above formula that this effect size not only takes into consideration the scale of differences, which is the gap between M1i and M2i, but also the scale of variation, which is SDpooled,. The differences between Group 1 and Group 2 are thus standardised.

What needs attention is that results analysed using standardised effect sizes do not necessarily mean that they are completely comparable to other studies with other measures, but they do, or hopefully they do, provide a rough impression of how substantial the results actually are (Gorard, 2021).

In this study effect sizes are primarily employed in the analysis of the survey results. As there are some questions in the survey that ask students to use 1 to 10 to quantify the extent to which they agree that potential factors affect their plans for after they leave compulsory education and after they are 18 years old, and some that ask them to evaluate the extent to which they agree that the CA indicators currently used in China improve the equality and equity of HE, it is appropriate to compare the scores using effect sizes. When the comparison is only between two groups, such as females and males or Han ethnicity and minority ethnicity, it is easy to calculate the effect sizes. When the comparisons involve more than two groups, such as parental education qualifications, including HE, vocational college education or equivalent, high school education or equivalent and compulsory education, the largest group, such as the group of parents with compulsory education, is chosen as the reference group. The other groups are then compared with the reference group rather than with each other.

Odds ratio

The odds ratio is another kind of effect size. It is a little different and it is mainly used to compare two categorical variables with binary outcomes (Sullivan & Feinn, 2012). Specifically, it compares “the ratio of the odds of an event occurring for one group” with “the odds for a second group” to show disparities between the two groups (Gorard, 2013, p175). According to Gorard (2013; 2021), the odds ratio can be calculated using the following formula:

$$\text{Odds ratio}=(a/c)/(b/d),$$

where (see Table 8.4.1):

a is the percentage with variable A in Group 1;

b is the percentage with variable B in Group 1;

c is the percentage with variable A in Group 2;

d is the percentage with variable B in Group 2.

Table 8.4.1 Calculating the odds ratio

	Variable A	Variable B
Group 1	a	b
Group 2	c	d

In this study the odds ratio is primarily used in the descriptive analyses of CGSS and CFPS data. The groups compared include sex groups, ethnic groups and hukou groups, and the variables compared are the outcomes of participating in HE and high school education. These outcomes are binary, so it is appropriate to use the odds ratio to clarify the importance of the results and make the results more comparable.

8.5 Discussion and Conclusion

This chapter has briefly introduced the formulae used in this study: the Admission Rate Index, the Gorard Segregation Index, the Admission Opportunity Index and effect size. The first three were primarily used for comparisons of HE admissions and the admissions of prestigious HEIs (DFC universities) among provinces. All of them are suitable for comparing a list of groups. The Admission Rate Index is one of the most popularly used indices to reveal disparities, but it ignores the wider population. Therefore, the Gorard Segregation Index was employed to overcome this problem. Furthermore, in order to consider theoretically potential HE participants as completely as possible, the Admission Opportunity Index, which takes earlier education experiences into account, was used.

The effect size (*Cohen's d* and the odds ratio) was also used in the analysis of CGSS, CFPS and primary survey data to reveal disparities between two groups with different characteristics.

The next chapter is the beginning of the results part of this thesis. Chapter 9 provides the results of the structured review. It summarises the arguments in the literature reviewed and evaluates the quality of the research described. Chapter 10 examines the hukou province indicator. Chapters 11 to 17 give the results of the analysis of the two large-scale social survey datasets and Chapters 18 and 19 provide the results of the primary survey.

Part 3 Results

Chapter 9 Structured Review Results

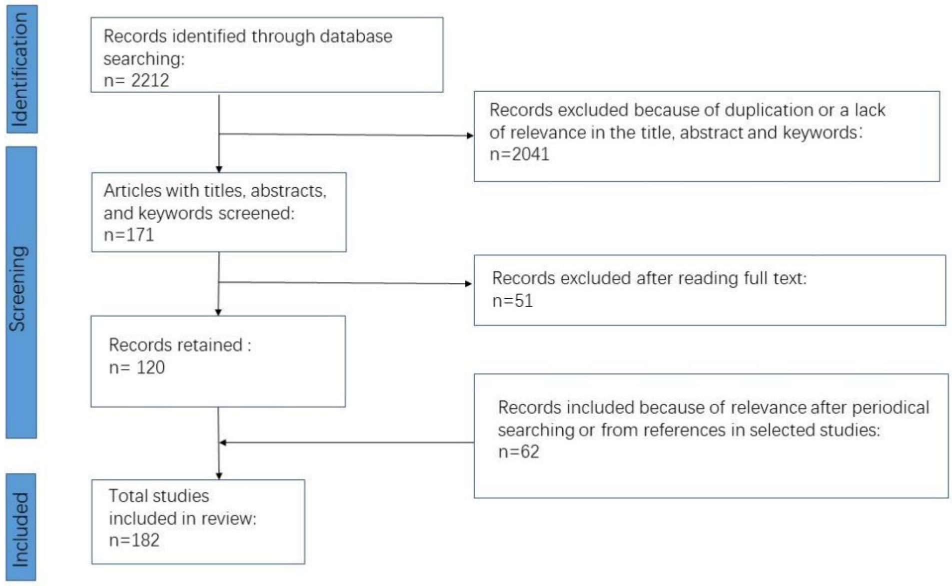
This chapter provides the results of the structured review, looking at HE participation disparities according to different indicators found in previous studies. The first section gives a brief introduction to the structured review process and then Section 9.2 summarises the findings and conclusions of the studies collected. These findings are categorised by the indicators employed. This is followed by a section evaluating the quality of these previous studies, including their shortcomings and limitations, and how the present study aims to improve on these limitations.

9.1 The Results of the Structured Review

As Figure 9.1.1 shows, 2,212 documents were identified in the main search process. After a simple selection scanning the titles, keywords and abstracts of these 2,212 documents, 2,041 documents that were irrelevant or duplicated were excluded and 171 were left. These 171 studies were skim read and 51 less relevant documents were excluded, leaving 120. A supplementary periodical search and some 'snowball' studies referenced in the selected studies produced 62 more studies. A total of 182 studies were included in the final review.

It should be noted that this was not a review preferring robust randomised controlled trials or high-quality studies but a structured review to first take relevant topics and themes in previous studies into consideration. However, this does not necessarily mean that relevant studies are cited in the review without them being evaluated. On the contrary, their qualities are objectively judged by using Gorard's (2021) sieve. The judging criteria have been introduced in Chapter 7. Only studies rated with 2 padlocks and above are introduced in Section 9.2 and they are synthesised by indicator groups. Besides the ratings, a summarised evaluation of the limitations of these studies is presented in Section 9.3.

Figure 9.1.1 Flowchart of the numbers of studies in each stage of the structured review



9.2 Summaries of the Documents

9.2.1 Currently-used Contextualised Indicators

Province

Provincial disparity

There are 61 documents discussing provincial disparities in HE participation. Table 9.2.1 lists their quality ratings and summarises their conclusions. The 37 documents rated 2 and above are discussed in the review.

The studies rated as 0 or 1 padlocks often have weak research design but strong definitive conclusions (E, 2008; Gu, 2011; Tan, 2013; Yang, X., 2010) or make conclusions from a small sample size (Bao, 2011; Liu, X., 2011; Wei & Liu, 2015; Zhe, 2016). These disadvantages are detrimental for research quality and are not rare in the collected studies. Section 9.3 makes effort to evaluate the studies, so the reasons why studies rated as low-quality would not be explained in the sub-sections of 9.2.1 or 9.2.2 in order to avoid repetition.

Table 9.2.1 Security ratings of studies showing disparities between provinces/regions

Security rating	Disparities in HE admission between provinces/regions	No disparities in HE admission between provinces/regions
0	12	-
1	12	-
2	28	-
3	9	-
4	-	-

Jacob (2006) conducted a primary cross-sectional survey to explore patterns of HE participation in ten universities. He found geographical barriers were one of the obstacles to equity in education in the early 21st century. Furthermore, by analysing secondary dataset of the 2006 CGSS, Zhang (2015) concluded that people living in the east of China, which is the wealthiest part of the country, are more likely to attend HE than people living in central and western China. The same is found by Zhu (2011), employing 2005 CGSS, and Li, J. X. (2016), employing 2010 CGSS. However, Li, J. X.'s analysis selected 1,812 out of 12,000 cases but did not give reasons for this, which undermines the quality of the study to a certain extent. The advantages of the east of China and the disadvantages of the west of China in HE have also been found by Wang (2019), Jiang (2020) and Zheng (2021).

Not only the HE participation but also the average length of schooling of people is also different by provinces. Wang, H. (2016) examined cases in 2008 CGSS data and she found that people living in Beijing, Tianjin and Shanghai have been educated for 3.9 years longer than those living in rural areas. This finding is echoed in Wei, Y. M. (2016)'s study of 2013 CGSS data.

However, one of the problems when exploring HE participation disparities between provinces by analysing survey datasets such as the CGSS and the CFPS is that sometimes there are few cases in some provinces. This lack of cases in certain provinces restricts comparisons of HE admissions to eastern, central and western areas rather than the province level, in order to avoid extreme values. Therefore, instead of survey datasets many studies choose to examine annual provincial admission data.

By collecting and analysing the numbers of NCEE candidates, HE admission rates, admission quotas and the numbers of HEIs in different provinces, many studies found inequality in HE admissions among provinces (Gao, 2017; Ji & Zhu, 2011; Shang, 2018; Wang, H. X., 2010; Zhang & Zhang, 2015). They found that some provinces such as municipalities (Beijing, Shanghai and Tianjin) were usually the most advantaged in HE admission, or even admissions of prestigious HEIs, while others including Xinjiang, Qinghai, Ningxia, Henan and Guizhou kept staying in the disadvantaged position. Gao (2017) claimed that three municipalities are "very easily and very likely to access elite HE groups." Gao gives the example of Beijing, the admission rate of which was the highest from 2011 to 2015. It was 5.52 times higher than that of Sichuan in 2011 and 4.29 times higher than that of Tibet in 2015.

One of the limitations of the above-mentioned studies is that they only focused on some HEIs or provinces. Such incomplete data might be vulnerable to bias. Therefore, studies with more complete admission data are rated higher. For example, Pan et al. (2010) used admission data for all the HEIs that were affiliated with the central government from 2004 to 2008 to get a more complete picture. They found that the admission opportunity indices in some provinces always exceeded the national average of 15.93%. They classifies provinces in three groups: the first group contains Beijing, Tianjin and Shanghai, where there were the best-developed HEIs; the second group includes Jiangsu and Hunan among others, which had rich fundamental education and HE resources; and the last group includes Tibet, Xinjiang, Hainan and Ningxia, which might benefit from preferential admission policies.

Liu, N. N. (2020) collected the admission data of prestigious HEIs, Project 211 universities, and ranked provinces by their admission rates. The top five provinces were Beijing, Tianjin, Ningxia, Shanghai and Qinghai and the last five were Hunan, Guangxi, Henan, Guangdong and Yunnan, most of which have huge populations. Furthermore, the most advantaged provinces had Project 211 university admission rates nearly four times those of the most disadvantaged provinces, although this gap is gradually narrowing.

Liu, C. (2015) analyses survey and macro-level provincial data, and compares regions' GDP, numbers of HEIs, numbers of college entrants, rates of college entry, the ratio of HEI numbers to high school graduates, the student-teacher ratio in primary schools, government expenditure per capita and the unemployment rate. She confirms a HE admission disparity between regions. However, some data, such as the numbers of college entrants and the rates of college entry, analysed in her study are less appropriate, because the figures do not indicate how many college entrants the province produced but how many college entrants the province absorbed. In other words, provinces that possess rich HE resources such as more HEIs have higher numbers of college entrants but this does not necessarily mean that the college entrants were from the provinces.

The studies by Pan et al, Liu, N. N. and Liu, C. together give a fairly complete picture of student intakes by HEIs affiliated with the central government, but they only pay attention to NCEE candidates and HE participants and ignore the possibly larger group of non-HE participants.

Therefore, some studies employed a broader view of their target groups. Liu (2007), Wang and Du (2013), Zhang and Li (2019) and Cao and Zhang (2016) respectively used the population aged 18 to 22, graduates from middle schools and graduates from primary schools (the last two studies) in their calculations. All four studies found disparities in HE admissions and elite HE admissions among provinces. Wang and Du further concluded that the more prestigious HEIs are, the larger are gaps in their admissions between provinces. It seems clear that Chinese HE entrance system prefers students who live in provinces with more universities and lower HE thresholds (Hamnett, Shen & Liang, 2019).

Li (2013) used the example of Henan as a case study to explore provincial disparity. She collected Henan's admission data for half a century. Although HE admission rates in Henan have risen since 1977, they are still low compared with other provinces. The rate was even the lowest in China in 2011 and 2012.

In short, according to the previous studies, there are disparities in HE admissions among provinces/regions. The eastern coastal areas are found to be advantaged in HE than the middle and the west of China. In specific, municipalities including Beijing, Tianjin and Shanghai are always in the most privileged position, whereas provinces suffering resource shortages and large population such as Tibet and Henan, respectively, are always in the disadvantaged group. This study collects provincial admission quotas from 2016 to 2019 both for prestigious DFC universities and for regular HEIs and employs three indices to compare HE admissions among provinces. The results are presented in Chapter 10.

Localised admissions

Besides disparities, another problem in HE admissions between provinces is the preference of elite HEIs, usually central-government-affiliated HEIs, for local students. As Table 9.2.2 shows, five papers rated 2 or higher discuss this issue. Pan et al. (2010) explored the admission quotas of HEIs affiliated with the central government from 2004 to 2008 and they found a localised tendency in the admissions of all the HEIs they examined, with an average local admission rate of 34.6%.

Table 9.2.2 Security ratings of studies showing localised admission tendency

Security rating	Localised admission tendency	No localised admission tendency
0	6	-
1	2	-
2	5	-
3	-	-
4	-	-

This finding is repeated by Zhou, B. H. (2018). After examining the local intakes of several central-government-affiliated HEIs, Zhou, B. H. found that the higher the quality of central-government-affiliated HEIs were, the higher their local enrolment rates would be.

Ji and Zhu (2011) compared the admission rates of municipalities and some provinces with large populations in 2009. They found inequality among the admissions of Project 985 universities in that locally biased admissions made provinces that were rich in prestigious universities also rich in HE places. This localised preference remained unchanged in 2012 (Zhang & Zhang, 2015). Gao (2017) analysed admission rates from 2011 to 2015 and also criticised high local biases in the admission rates of Peking university and Tsinghua university.

There seems to be a preference for localised HE admission. However, it can be seen from Table 9.2.2 that not many of the studies reporting this are of high quality. Most of the documents collected are rated 0 or 1. These are low-quality reviews such as ones re-reporting results from other studies without explaining or evaluating the HEIs' selection processes. The present study aims to provide more robust and up-to-date evidence on this issue. It examines localised admissions of DFC universities in Chapter 10.

Hukou status

63 documents in the review discuss HE participation disparities between different hukou status. As Table 9.2.3 shows, most of the documents collected reveal a rural-urban gap in HE admissions.

Table 9.2.3 Security ratings of studies showing disparities between urban hukou and rural hukou

Security rating	Disparities in HE admissions between hukou status - rural students disadvantaged	Disparities in HE admissions between hukou status - urban students disadvantaged	No disparities in HE admissions between hukou status
0	2	-	-
1	17	-	-

2	28	-	1
3	15	-	-
4	-	-	-

No matter employing national-level data such as census, or delivering a primary survey, an overall increase in HE participation in both urban and rural residents have been found after HE expansion in China (Li & Min, 2001; Li et al., 2011; Wu & Zhang, 2010; Zhang, 2009). With some exceptions that thought decreasingly smaller gaps in HE participation between hukou status in the recent younger generation (Fang & Fang, 2018; Wu, 2019; Wu, Zhang & Hu, 2021) and a finding that the HE participation rate of rural students is close that of urban students by analysing the China's Move to Mass Higher Education (CMMHE) dataset (Liu, J., 2011), many studies in the review found that the rural-urban gap in HE participation has widened rather than narrowed (Li & Min, 2001; Li et al., 2011; Wu & Zhang, 2010; Zhang, 2009). Rural hukou students need to make much more effort to access HE than their urban hukou counterparts (Zhang, 2009).

The similar results are obtained by analysing secondary social survey datasets. Zhu (2011) analysed 2005 CGSS data and explores the greater possibility for urban hukou students to attend HEIs. This advantage of urban hukouers is also presented among the cases in 2006 CGSS data (Fang & Feng, 2018; Li, 2011; Jiang, 2015; Tam & Jiang, 2015; Tang, Huang & Liu, 2014; Zhang, 2015), in 2008 CGSS data (Tang, 2016; Wang, Y. P., 2021; Yeung, 2013; Zhang & Chen, 2014), in 2010 CGSS data (Li, J. X., 2016; Li, 2015; Ou & Hou, 2019), in 2013 CGSS data (Wei, H. L., 2016; Zheng & Sun, 2017), in 2015 CGSS data (Deng & Fu, 2020; Pang, 2019; Zhu, Xu & Wang, 2018; Wu, Yan & Zhang, 2020; Zhou, F., 2018) and in 2017 CGSS data (Wu, Zhang & Hu, 2021). Some studies merged the CGSS datasets for several years in order to obtain a longer-term picture, but they did not find surprisingly different results from the previous studies (Xu, 2017; Yeung, 2013; Zheng & Sun, 2017).

Li (2015) employed 2010 CGSS data to make comparisons of participation in regular HE and adult HE between urban hukouers and rural hukouers. Despite an increase in HE admissions, rural hukouers still lag behind in regular HE attendance. On the other hand, he was surprised that adult HE, which had a lower entry threshold and provides practical skill learning, dis not attract many rural hukouers but showed a decrease in their enrolment. Li advised not insisting on increasing participation in regular HE by rural hukouers but encouraging rural hukou students to attend adult HE. However, as was mentioned in Chapter 2, in Chinese society adult HE is less valued than regular HE, so it would be less reasonable and less ethical to suggest that disadvantaged groups should select less-valued education paths rather than focusing on helping them compete for more-valued ones.

The CFPS is another popular social survey dataset. By analysing CFPS data, Li (2019), Pan and Wu (2020), Sun and Yan (2015), Xu, W. Q. (2021) and Zheng (2021) all concluded that rural students were disadvantaged in education. Rural hukou students encounter several barriers in their transitions to middle school and high school (Li, 2019; Pan & Wu, 2020), let alone to HE or prestigious universities. For example, rural hukou students accounted for less than 20% of the students at Tsinghua university, one of the top universities in China, from 2011 to 2019 (Zheng, 2021). Zheng also compared the ratios of urban hukou and rural hukou high school students, who are potential HE participants. The corresponding figures for 2000 are 86.86% (urban) and 13.14% (rural), and for 2010 are 93.28% (urban) and 6.71% (rural) in 2010. Rural hukou students seem disadvantaged from high school already. However, these ratios ignore the distributions of wider population in different hukou status, which might have some biases.

The under-representation of rural hukou students in HE or their lower average education levels are also revealed by Li and Min (2001), Guo, Song and Chen (2019), and Zhang (2017), analysing the Urban Household Survey database of China, the 2013 Chinese Household Income Project (CHIP) and the 2013 China Health and Retirement Longitudinal Study (CHARLS), respectively.

Li (2014) took the example of Henan province as a case study. His data analysis and review of previous studies led him to conclude that rural students in Henan are disadvantaged in education, and that policies aimed at helping rural students are irresponsive, ineffective and unreasonable.

So far, hukou differences seem one of the principal criminals of HE inequality. Yang, Huang and Liu (2014) and Li (2011) considered the urban-rural division to be the best indicator of education inequality, even better than occupational prestige, age, sex and regional gaps, after analysing secondary social survey data. Wen (2022) found that the disadvantage in HE participation by rural hukou students is associated with them having less favourable family backgrounds. Rural hukou students mostly come from lower SES backgrounds with less cultural capital and poorer educational resources.

There remain differences in HE participation not only by hukou status but also by living area. Jia and Ericson (2017) conducted a survey of 1,028 people and they found that the closer students live to cities the greater opportunities they have to attend HEIs and elite HEIs. This finding is supported by Xu (2021), by analysing the CFPS.

However, Treiman (2012) and Zhang, H. F. (2017) argued that hukou status is a more crucial factor than living area in individuals' educational attainment. Treiman (2012) compared two survey datasets, one with 6,090 cases and the other with 3,000 cases, and found that after 1985 people who lived in cities but have rural hukou did not have any advantage over people living in rural areas who had rural hukou. These urban rural hukou people mostly lived on the edges of cities. They seemed to enjoy more resources such as better equipped schools than rural areas, but they should not be regarded as having similar experiences to dwellers in large cities.

Zhang, H. F. made a similar case by analysing Rural Urban Migration in China (RUMiC) data, a longitudinal survey (with 15,000 cases in the baseline survey) focusing on rural-urban migration in China. He classified students as urban students (urban hukou and living in an urban area), rural students (rural hukou and living in a rural area), rural-hukou migrated students (rural hukou and living in an urban area) and new urban-hukou migrated students (who have successfully changed into urban hukou and live in an urban area). Zhang, H. F. found that rural students and rural-hukou migrated students were remarkably disadvantaged in education and had worse educational outcomes and lower transition rates. Even if students with rural hukou study in schools located in cities, their initial disadvantage seemed to change little. This is echoed by Zhang, Li and Xue (2015). Zhang, H. F. suggested that it is not a disadvantaged family background but hukou segregation that was the main reason for the inadequate schooling of rural-hukou migrated students.

In summary, according to previous studies rural hukou students are highly likely to be in a disadvantaged position in HE participation. The studies arguing smaller or even no disparities between hukou status were problematic in their improper variable and restricted target groups, which are discussed in detail in Section 9.3. In addition, living in rural areas might also be a disadvantage but it seems less detrimental than hukou segregation.

Minority ethnicity

As Table 9.2.4 shows, 22 of the documents collected investigate HE equity between ethnic groups. Nine documents found that there was a gap in HE admissions between Han ethnicity and minority ethnicity students, while thirteen documents did not find minority ethnicity students are disadvantaged in HE anymore. Furthermore, only five of fifteen good-quality studies (rated 2 and above) found ethnic gaps in HE admissions. It seems that although minority ethnicity has been used as an official CA indicator for a long time, it is not widely acknowledged as a good indicator any longer.

Table 9.2.4 Security ratings of studies showing disparities between Han ethnicity and minority ethnicity

Security rating	Disparities in HE admissions between ethnic groups - minority ethnicity disadvantaged	Disparities in HE admissions between ethnic groups - Han ethnicity disadvantaged	No disparities in HE admissions between ethnic groups

0	1	-	1
1	3	-	2
2	4	-	2
3	1	-	8
4	-	-	-

Specifically, after analysing CFPS data, Zhang (2014) found a disadvantage of minority ethnic students not only in HE admissions but also in earlier periods of education such as primary school. The most disadvantaged group is minority ethnicity students who lived in the west of China and were not able to understand Mandarin. The disadvantages of minority ethnicity in HE are also supported by Ou and Hou (2019), with the analysis for the CFPS.

In addition, there are some studies that collected primary data. Jia and Ericson (2017) surveyed 1,028 students and interviewed 54. They asked students for their NCEE scores and college admission information and found that in their sample minority ethnicity students are under-represented in regular HE but over-represented in prestigious HEIs such as Project 211 and Project 985 universities. However, the sample in this survey only includes HE participants, making it vulnerable to sampling bias.

Xie and Liu (2019) surveyed 848 respondents from Guizhou University for Nationalities, Southwest University for Nationalities, Yunnan University for Nationalities and Guangxi University for Nationalities, and Deng, Seepho and Lian (2020) interviewed 74 students from Yunnan. Both find that minority ethnicity students are more disadvantaged from the beginning of their education. This disadvantage continues into HE.

Moreover, in Xie and Liu's sample, 62.4% of the students agreed that the CA policy favouring minority ethnicity students helps these students in their educational transitions, and 60.9% of the students thought the policy is fair. However, due to the over-representative (nearly 50%) of minority ethnicity students in Xie and Liu's sample (given the national proportion is around 1:9), the aforementioned results may be problematic in generalisability.

Deng, Seepho and Lian also found that the CA policy benefits many minority ethnicity students, as it provides these students with more chances to participate in HE. Nevertheless, they argued that the extra scores in the CA policy for minority ethnicity students cannot solve the problem of minority ethnicity being a disadvantage in education. They called for a more comprehensive set of CA policies including more investment in basic education and in pedagogy.

However, the primary surveys introduced above almost focused on the HE students rather than the wider population of potential HE participants. This sampling bias somewhat weakens the robustness of these studies. On the other hand, many high-quality studies have not found noteworthy gaps in HE participation between ethnic groups and minority ethnicity is not necessarily disadvantaged in HE by analysing 2008 CGSS data (Tang, 2016; Zhang & Chen, 2013; 2014), 2015 CGSS data (Wen, 2022; Wu, 2019; Wu, Yan & Zhang, 2020), merged CGSS files (Wang, Y. P., 2021; Yeung, 2013). Nevertheless, Wen (2022) found minority ethnicity is a disadvantage in vocational education.

To sum up, minority ethnicity groups were disadvantaged in HE participation in the past but the gaps between minority and Han ethnicity admissions have shrunk. Nowadays, it might not be appropriate to regard all minority ethnicity students as in need without having more information on their socio-economic status. However, several studies state that minority ethnicity students have an accumulated disadvantage in the earlier stages of education, which is important. Therefore, this accumulated disadvantage will be examined in the present study.

9.2.2 Popularly Discussed Contextualised Indicators

Sex

There are 51 studies addressing sex disparities in HE participation. 18 studies find that males are advantaged in HE participation and 20 studies conclude that there are no such gaps. Furthermore, 13 studies investigate gaps within each sex group such as rural and urban hukou females and rural and urban hukou males. The studies rated 2 or higher mostly use secondary data analysis or are cross-sectional studies.

Li (2006) conducted a survey of 14,345 cases who were born between 1980 and 1986. She found an apparent sex difference according to the prestige level of HEIs. Males are largely over-represented in Project 985 and Project 211 universities. This finding is supported by Jacob's survey (2006).

Table 9.2.5 Security ratings of studies showing disparities between males and females

Security rating	Disparities in HE admission between sex groups - female disadvantaged	Disparities in HE admission between sex groups - male disadvantaged	Disparities in HE admission within sex groups	No disparities in HE admission between sex groups
0	-	-	-	1
1	2	-	1	3
2	14	3	8	4
3	2	2	4	7
4	-	-	-	-

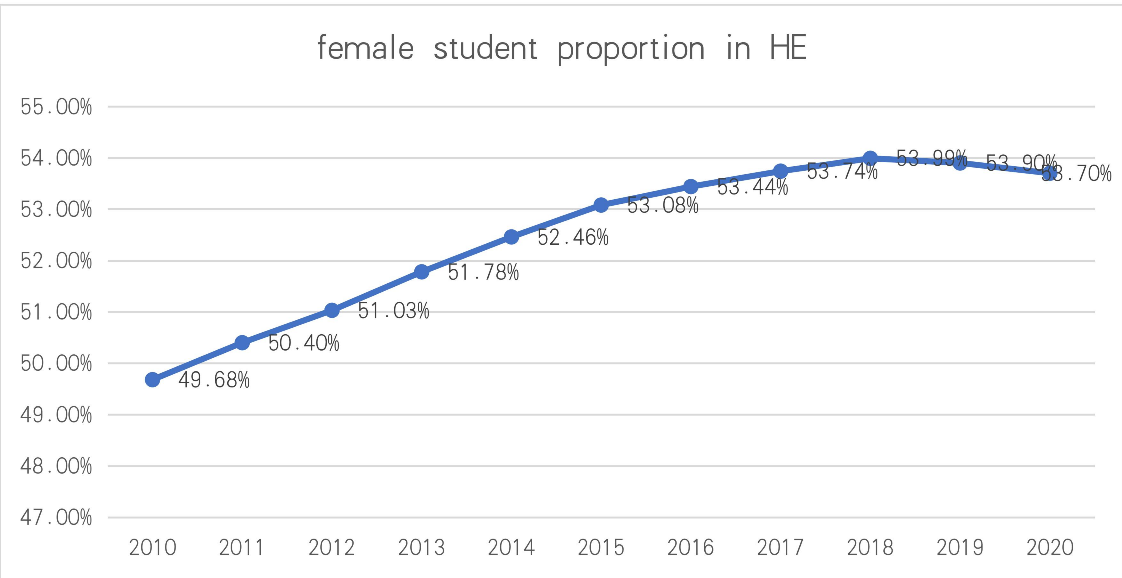
On the other hand, by employing secondary data, Li (2002) found that females were disadvantaged in HE admissions and they were more likely to go to vocational colleges. This disadvantage of females are also found by Zhu (2011), Zhang (2015), Li (2015), Pang (2019), Wei, H. L. (2016) and Zheng and Sun (2017). All the mentioned studies analysed CGSS data.

Liu, J. (2011), Ou and Hou (2019) and Zhao (2021) analysed CFPS and CMMHE data and agreed that males are advantaged in HE participation. Zhao found parents were less willing to invest in educating female children. Hannum, An and Cherng (2011) studied the case of Gansu, which is one of the most under-developed provinces in China, by investigating data from the Gansu Survey of Children and Families. They also found a slight disadvantage of females in HE participation.

Cheng (2020) provided a more updated comparison by examining admission data from 2014 to 2018 for a university in Nanjing, the capital of Jiangsu province. In the sample of 26,169 university students, males were over-represented at this university during the five years studied. In follow-up interviews, some interviewees suggested that females from under-developed areas were not usually expected to receive longer education, and females from families with more than one child might suffered from less attention being given to their education.

However, other studies found no significant gap in HE participation between sex groups after analysing datasets (Jia & Ericson, 2017; Li, J. X., 2016; Lu, Deng & Guo, 2016; Song, Liu & Wang, 2019; Wei, 2013b; Wen, 2022; Xu, 2021; Zhang & Chen, 2013; Zhang & Chen, 2014; Zhang, K., 2017; Zhou, F., 2018;). Some studies found a reversal in HE admissions between males and females. In other words, they find female students have become advantaged over their male counterparts (Du, Ni & Huang, 2018; Fang & Feng, 2018; Wu, Yan & Zhang, 2020; Wu & Zhang, 2010; Yeung, 2013). According to official data from MOE (2022), female students in HE have been over-represented than male students since 2011.

Figure 9.2.1 The proportions of female students in HE in China



Source: MOE

These different findings are not surprising. First, males and females might attend universities with different levels of prestige, with males tending to be over-represented in more elite HEIs (Li, 2006). Information on the prestige of universities attended cannot be found in social survey datasets, but studies that conducted surveys in universities focus more on elite universities. This could explain why some studies find inequality in HE enrolment between sex groups, while other studies analysing social survey datasets do not.

Second, there might be some complex issues within each sex group. For example, by investigating Peking university admission data from 1978 to 2005 Liu and Wang (2009) found a decrease in sex disparities among the student population but they also found that rural hukou female students were extremely disadvantaged, as they only constituted 4% of the student population. Liu and Wang feared that disadvantaged rural females were usually ignored when they are included in the general female group. The under-representation of rural hukou female students in HE and in prestigious HE is also revealed in Xie, Wang and Chen's (2010) and Guo, Tsang and Ding's (2006) survey for university students.

The disadvantage of rural females is also supported by other studies employing secondary data analysis. Based on the analyses of CGSS data, Wu et al., (2020), Zhu, Xu and Wang (2018) and Yang, Huang and Liu (2014) found that sex disparities among rural hukouers in western China are much greater than those among urban hukouers in eastern China. Lei et al. (2017), Wu and Huang (2015) and Zheng (2021), three studies that analyse CFPS data, also identified the above gaps. Although urban female students have increased their HE participation, rural female students have only improved the decrease in their illiteracy rate.

Furthermore, there are also disparities between minority ethnicity females and Han females. According to Xu (2017), who explores 2010-2013 CGSS data, minority ethnicity females are the most disadvantaged in HE participation. This finding is echoed by Wu et al., (2020), Zhu, Xu and Wang (2018) and Yang, Huang and Liu (2014).

In addition, Female students from higher SES family backgrounds or with better educated parents are more likely to enter HE than females from lower SES backgrounds and males (Luo, Guo & Li, 2021; Wang, Y., 2021).

On the other hand, there are some differences in university majors between sex groups. In Jacob's (2006), and Guo, Tsang and Ding's (2010) samples, females were more likely to study liberal arts or social sciences, while males were more likely to choose science or engineering. Males also tended to have higher starting salaries after graduating.

Therefore, sex disparities in HE appear to have become a complex issue not only involving sex differences but also other characteristics. Females might be disadvantaged not because of their sex but because of other factors such as rural hukou, a minority ethnicity and poverty.

Family socio-economic status

There are 92 documents collected in the review exploring disparities in HE or in education between different family socio-economic status, which can be classified in four groups: family income, SES/social class, parents' education qualifications and parents' occupations.

As indicated in Table 9.2.6, family income is an important indicator. Wang et al.'s (2011) cross-sectional study involving 20,253 university students and 1,177 high school students, and Li's (2006) study involving 14,345 university students both found that students from low-income families were greatly under-represented in HE participation compared to those from middle-income and high-income families.

Table 9.2.6 Security ratings of studies showing disparities in family income

Security rating	Negative relationship with family income	Positive relationship with family income	Not relevant to family income
0	-	2	-
1	-	4	-
2	-	7	-
3	-	2	2
4	-	-	-

Secondary datasets, including the CGSS, the CFPS, the Urban Household Survey database of China and the CMMHE, provide similar results that HE participation is positive associated with individuals’ family income (Jin & Li, 2022; Li & Min, 2001; Liu, J. (2011); Yang, 2020; Zhang & Chen, 2013; Zheng, 2021; Zheng & Sun, 2017). Students from low-income families are increasingly disadvantaged in HE participation (Zheng & Sun, 2017).

However, after analysing CFPS data, Li and Lu (2015) and Zhao (2021) found no significant influence of family income on HE participation. But the significance test is not necessarily meaningful in explaining the results of regression analysis.

Family SES/social class is also often considered (Table 9.2.7). Zhang (2009) regarded SES as a priority determinant in HE admissions because he found in survey samples that there were SES disparities among different HEIs and majors. Furthermore, he found that students from higher SES families generally obtain higher NCEE scores than ones from lower SES families. Zhang therefore feared that SES disparities make the “everyone is equal in front of NCEE scores” admission principle less fair.

Table 9.2.7 Security ratings of studies showing disparities in SES/social class

Security rating	Negative relationship with SES/social class	Positive relationship with SES/social class	Not relevant to SES/social class
0	-	-	-
1	-	-	-
2	-	4	2
3	-	1	-

4	-	-	-

Wu (2017) investigates 2009 Beijing College Students Survey data. He concludes that there is a close association between family backgrounds and the prestige of universities that students attend. Students from the upper-middle and upper classes are more likely to attend Project 211 universities, and they are also favoured by the recommendation system (baosong). This over-representation of students from higher SES backgrounds in prestigious HE is found by Ding et al. (2021), by analysing China Educational Finance Statistics – Postsecondary (CEFS-Postsecondary) and National Survey on College Graduates' Employment (NSCGE) data.

The analyses for the CGSS obtained similar results. HE expansion has not improved but instead exacerbated SES gaps in HE participation and it favours students from advantaged family background (Wang, Y. P., 2021; Wei, 2013a; Xu & Fang, 2020). On the other hand, students from lower SES background tend to be more disadvantaged in primary and middle school rather than in HE (Tang, 2016). This might be because entry into HE is a result of previous selection, so there seem to be fewer SES gaps.

However, Jia & Ericson (2017) conducted a survey involving 1,028 university students. They found that although students' SES backgrounds were associated with HE access they were not associated with the prestige of the HEIs that they accessed. Nevertheless, this different conclusion might be attributable to sample selection bias.

Third, parental educational qualifications are the most widely recognised indicator and most of studies in this review are rated as 2 and above (Table 9.2.8). In both Li's (2006) and Li and Min's (2001) survey samples, HE attendance is positively associated with parents' education levels. Students whose parents have bachelor's degrees are remarkably more advantaged in HE, or prestigious HE such as Project 985 universities and Project 211 universities, than ones whose parents only have high school qualifications or below. The importance of parental education qualifications to children's education is also found by Li (2007), analysing 2004 China College Student Survey data.

Table 9.2.8 Security ratings of studies showing disparities in parental education qualifications

Security rating	Negative relationship with parental education	Positive relationship with parental education	Not relevant to parental education
0	-	2	-
1	-	2	-
2	-	24	-
3	-	10	-
4	-	-	-

Furthermore, according to analyses of CGSS data, students' HE participation and educational attainment are closely associated with parents' education qualifications. For example, Zhang and Chen (2013) analyse 2008 CGSS data and find that the higher parental education qualifications are, the greater opportunity students have to attend HE. The role of parental education can even diminish sex disparities (Zhang & Chen, 2013) and regional disparities (Li, J. X., 2016) in HE participation, as males from families with parents who completed middle school education or above are not more advantaged. According to Fang and Feng (2018), who analyse a large merged dataset containing 2003, 2005, 2006, 2008, 2010, 2012 and 2013 CGSS data, this association between parental education and children's opportunities to attend HE has existed for a long time.

This result has been repeatedly supported by other studies exploring 2005 CGSS (Zhu, 2011), 2008 CGSS (Tang, 2016; Zhang & Chen, 2014), 2010 CGSS (Li, 2015), 2013 CGSS (Wei, H. L., 2016; Zhou et al., 2018) and 2015 CGSS (Pang, 2019; Wu, Yan & Zhang, 2020; Xu & Fang, 2020; Zhou, F., 2018; Zhou, et al., 2018) data and merged CGSS datasets (Yan & Wang, 2012; Yeung, 2013; Zhu, Xu & Wang, 2018). Some studies only examine the father's (Zhang, 2015) or the mother's education qualifications (Song, Liu & Wang, 2019; Xu, W. Q., 2017; Yang, 2020) but they have similar findings of an association between the education levels of the older and the younger generations.

A similar pattern is also found in results of analyses of the CFPS (Li & Lu, 2015; Pan & Wu, 2020; Zhao, 2021; Zheng, 2021), Urban Household Survey database of China (Li & Min, 2001) and CMMHE (Liu, J., 2011) data. College-educated parents affect not only HE enrolment and elite HE enrolment but also the field of study (Liu, J., 2011).

In addition, Xu, D. T. (2021) found strong intergenerational mobility in education in the CFPS. This intergenerational educational mobility is also found by Guo, Song and Chen (2019). They analysed 2013 CHIP data and found that there was stronger mobility of children's education when the parents' education levels are in the middle rather than the highest or lowest quintile.

Table 9.2.9 presents the ratings for studies discussing parents' occupations, many of which are rated as medium-quality or above. There seems a widely acknowledged positive correlation between parents' occupations and their children's education levels based on the analyses for some well-known survey data such as the CGSS, the CMMHE and the CFPS (Fang & Feng, 2018; Jin & Li, 2022; Li, 2015; Liu, J. 2011; Li & Lu, 2015; Wang, H., 2016).

Nevertheless, there are also some different findings. For instance, Zheng (2021) examined CFPS and census data and reached the conclusion that individuals' education attainment is positively correlated with their mothers' occupations but not with their fathers'. Pan and Wu (2020) also explored CFPS data to compare students' educational attainment in the transition to middle school, to high school and to HE. They found that students whose parents have a higher occupational status are advantaged in the transition to middle school and high school but there is no great difference in the transition to HE. Again, this fewer differences in the transition to HE may be attributable to previous selections. Furthermore, by analysing 2006 CGSS, 2010-2013 CGSS and 2015 CGSS data, Xu (2017), Yang (2020), Zhang (2015) and Zhou, F. (2018) found that parental occupations have no significant impact on children's HE enrolment because of non-significant p-values.

Table 9.2.9 Security ratings of studies showing disparities in parental occupations

Security rating	Negative relationship with parental occupations	Positive relationship with parental occupations	Not relevant to parental occupations
0	-	1	-
1	-	5	-
2	-	14	3
3	-	7	-
4	-	-	-

To sum up, previous studies generally agree that family socio-economic status, including family income, SES/social class, parents' education levels and parents' occupations, is associated with noteworthy disparities in HE admissions.

School indicators

Eight documents investigated the role of school-level characteristics and they are all rated 2 or higher. First, according to Table 9.2.10, the continuation to academic high school education rather than vocational high schools have been emphasised as the key to equity in further education on the basis of the results from 2005 CGSS and 2008 CGSS (Deng & Fu, 2020; Tang, 2016).

Table 9.2.10 Security ratings of studies showing disparities between academic high schools and vocational high schools

Security rating	Disparities in HE admissions between high school types	No disparities in HE admissions between high school types
0	-	-
1	-	-
2	1	-
3	1	-
4	-	-

In addition, school type is also crucial (Table 9.2.11). Jia and Ericson (2017) made a survey involving 1,028 students in Nanjing and found relationships between high school type and HE participation. They found that students from key high schools were more likely to go to prestigious universities. This result is repeated by Tang (2016), exploring 2008 CGSS data, and Wu (2017), exploring Beijing College Students Panel Survey. Lai et al. (2016) confirmed the importance of high school ranking in the prestige and location of universities that students were admitted by conducting a panel survey of 11 high schools, which covered 989 students.

Educational attainment is also important in students' educational destination (Table 9.2.12). After analysing 14,345 cases born between 1980 and 1986, Li (2006) found that students whose educational outcomes were in top ranking high schools were more likely to access HE, or elite HE. Wang et al. (2011) conducted a survey of 20,253 HE students and 1,177 high school students. They found that low-income high school students performed as well as their high-income counterparts in their schoolwork, so the barriers to HE exist before students attended high school.

Table 9.2.11 Security ratings of studies showing disparities between key high schools and regular high schools

Security rating	Disparities in HE admissions between high school types	No disparities in HE admissions between high school types
0	-	-
1	-	-
2	4	-
3	-	-
4	-	-

Table 9.2.12 Security ratings of studies showing disparities in high school achievement

Security rating	Disparities in HE admissions between high school achievement	No disparities in HE admissions between high school achievement
0	-	-
1	-	-
2	1	-
3	1	-

4	-	-

School-level indicators, such as school type and schoolwork achievement are related to gaps in HE admissions. However, there are at least three shortcomings of school-level indicators, or of studies above. The first is that it seems tautology when discussing the relationships between HE participation and, no matter school type or school work achievement, as the requirement for attending HE, prestigious HE, high schools or key high schools is good schoolwork achievement. Second, school-level indicators, as group-level indicators, could not avoid ecological fallacy. It would be untenable to assume that all students in key high schools are advantaged and all students not in key high schools are disadvantaged. Finally, school type, school outcomes and even the transition to post-compulsory education may already be the result of selection and be stratified by other characteristics such as SES. It might need more considerations about the extent that the school-level indicators are fair to be used as an indicator for the disadvantaged.

9.3 Limitations of Previous Studies

This section summarises the limitations of the documents collected in the review. This study aims to improve on them.

9.3.1 Inappropriate Research Design

A first problem is that the research design in some studies is inappropriate. For example, Jacob (2006) conducted a large-scale survey of ten HEIs to explore how fair their admission processes are. However, HEIs can only know about HE entrants, who are just one small group involved in the issue of HE equality and equity. Focusing research on HEIs will bias results as it only pays attention to the winners. What we really want to know is why non-HE participants do not go to HE rather than why HE participants do. Therefore, the larger group of non-HE participants should not be ignored. Jacob (2006) is not an exception as in fact many studies focusing on HE equality only focus on HE entrants and some even only investigate one university case (Cheng, 2020; Dong, 2009; Liu, 2019; Liu, X. H., 2014; Liu, Wang & Yang, 2012; Song, 2009; Xie, Wang & Chen, 2010; Yang, D.P., 2010).

Furthermore, some studies use existing literature to find answers to research questions but they fail to do a structured review (Gao, 2018; Li & Wang, 2017; Liu, 2013; Lu, 2018; Ma, 2012; Sun & Barrientos, 2009; Tan, 2013; Tang, 2011; Tu & Wu, 2012; Wang, L., 2011; Wang, Y., 2018; Xu, 2007; Xu, 2016; Yang, D.P., 2010; Yao, 2008). Few of them explain where the literature was found, what kind of search syntax was used, or how the studies were selected and evaluated. These might make readers doubt the quality and trustworthiness of these studies. Worse still, some studies do not have a clear research design (Bie & Zhu, 2003; Guo, 2010; Liang, 2010; Ma, 2012; Yang, X., 2010; Zhang, 2016).

Another problem is where an associative research design is used to make causal claims. There should be a clear distinction between causation and association, but many studies fail to clarify this and conclude with causal terms such as 'effect,' 'improve,' 'reduce' and 'influence' without the necessary justification (Gorard, 2013). For example, Fan (2014) concludes that family economic capital, social capital and cultural capital influence an individual's education attainment. But this causal claim is only based on a categorical regression analysis of CGSS data. Likewise, Hu (2012) employs a probit regression analysis but claims that intra-national migration impacts high school attendance. So do Hao, Hu and Lo (2014). Similarly, Fang and Feng (2020), Liu (2014) and Wu et al. (2020) conclude there is a causal relationship between SES and HE attendance or HE attainment by merely using simple regression analysis, and use the terms 'impact,' 'influence,' 'effect,' 'reduce,' and 'improve'. Yeung (2013) even points out that his study was not designed to investigate causation but he still uses strong causal terms in his conclusion. The studies mentioned here are not exceptions but are examples of a common problem (see also Fang & Feng, 2018; Hao, Hu & Lo, 2014; Wu et al. 2020; Xu, J. Y., 2017; Ye, 2015; Zheng & Sun 2017; Zhou, 2018; Zhou et al. 2018).

According to Gorard (2013), these definitive causal terms, which should be precisely distinguished from terms that only represent an association, should only be used when there is a robust causal research design such as a randomised controlled trial (RCT) or an instrumental variables (IV) or regression discontinuity approach (Gorard and Cook, 2007). In addition, complicated but plausible statistical techniques are also not certificates of causation and cannot make up for a weak research design. But in reality their functions are widely overestimated by researchers, who incorrectly claim causation by introducing fancy statistical modelling "with passive or cross-sectional datasets" (Gorard, 2013 p. 60,). To sum up, causation is different to correlation and researchers should be more cautious about using strong causal terms. Persuasive and reliable causal conclusions need to be supported by a robust research design.

As was explained in the previous chapters, this study employs secondary data analysis as the main design, and conducts a supplementary survey. It does not try to make any causal arguments.

9.3.2 Inappropriate Variables

Most of the studies collected choose secondary data analysis to explore the issue of HE equity. Some of the variables they analyse need some improvement. First, when discussing provincial HE disparities, many studies only take NCEE candidates into account in comparisons (Zhang & Zhang, 2015; Liu, 2011; Gao, 2017; Wang & Zhang, 2014; Yang, 2014; Wang, 2019; Li, 2013). As has been mentioned, NCEE candidates are very different to the age group that theoretically can enter HE and are just a very small part of this group. Some studies try to overcome this limitation by using the number of high school graduates (Wang, H. X., 2010; Liu, Y., 2015), primary school students (Liu, N. N., 2020) or even the local population (Qiao, 2007) in their calculations. However, these might not be satisfactory either. High school graduates are similar to NCEE candidates as they are already winners in the semi-final, and numbers of middle school and primary school students neglect possible migrations in three or nine years. The size of the overall local population is even worse as it does not take into account demographic structures in different provinces at all.

In order to focus on the targets as accurately as possible, this study first employs NCEE candidate numbers to calculate admission rates and Gorard segregation indices in each province and then uses both enrolment and graduation numbers of primary school students, middle school students and high school students and NCEE candidate numbers to calculate the admission opportunity indices for students in different provinces.

Furthermore, most studies that examine provincial disparities in prestigious HEIs only involve a few universities. Bao (2011), for example, examined four elite universities, namely Tsinghua University, Peking University, Zhejiang University and Nanjing University, to see how their admission quotas are allocated in the east, centre and west of China. Yang (2014) examined more universities but still only nine. This new study uses data from all 42 DFC universities, rather than collecting admission quotas from a few universities, and evaluates whether this prestigious group distributes their places fairly.

Another problem is a misuse of information on an individual's hukou status. Hukou status has been a popularly debated variable in HE transitions in China for a long time. But what should be pointed out is that in order to reasonably and correctly compare differences in HE access between urban and rural students, it is crucial to use individuals' hukou status before the NCEE, in other words before HE, in comparisons. This is because getting a place in HE is a classic way for the rural population in China to change their hukou status from rural to urban (Xu & Fang, 2020). If there is no effective restriction on the time when individuals got their urban hukou status there is a high risk of bias. In the CGSS, for example, the question directly asking about hukou status only asks for respondents' current hukou status rather than that before taking the NCEE. These data have been directly analysed by some researchers without any appropriate treatment (Hu & Vargas, 2015; Li, 2011; Liu, 2014; Wei, H. L., 2016; Zhang, 2015; Zhou, 2018). They might improperly count more HE participants as urban hukou holders and overestimate the advantage of urban students in HE competition.

Some studies take these biases into account and try to come up with improvements to this core variable. For instance, instead of using current hukou status Li (2015) and Fang and Feng (2018, 2020) choose the Hukou status of the interviewees at birth in their analysis. Moreover, Zhang and Chen (2013) view the answer to the question “What was your hukou status when you were 14 years old?” as the respondent’s Hukou status before HE. So do Deng and Fu (2020), Wu and Huang (2015), Yeung (2013), Zhao and Wang (2020) and Zhu, Xu and Wang (2018). Nevertheless, these treatments still risk biases. As the common age to participate in HE in China is 18, there is a long time for students, especially rural students, who intend to change their hukou status to achieve this aim. If a student had rural hukou when they were born or even until they were 14 years old but obtained urban hukou at the age of 17, it might result in an overestimation of HE admissions for rural students when the natal hukou status or 14-year-old hukou status is employed in the analysis.

The utilisation of information about hukou status has received more appropriate treatment in some studies. Tam & Jiang (2015) and Ye (2015) compared two variables: “age” and “the year when you obtained your urban hukou” and calculate a new variable representing the hukou status of respondents at age 18. This method is also used by Hao, Hu & Lo (2014), Song, Liu and Wang (2019), Xu and Fang (2020) and Wu et al., (2020). On the other hand, Ning (2018) abandons the usual binary division of hukou statuses and introduced a tertiary division between “always rural hukou,” “always urban hukou” and “transformation from rural hukou to urban hukou,” which moderates possible bias.

In order to obtain the Hukou information as accurately as possible, this new study makes some adjustments to the raw ‘hukou status’ variable in the CGSS. This variable only refers to the current status, which would result in an overestimation of urban hukouers’ HE participation and an underestimation of rural hukouers’ HE participation. Therefore, this study computes a new variable called ‘original hukou status’ and it is calculated using two variables existing in the CGSS, namely “the year you got urban hukou” and “the year you complete the highest education level (including the current one in progress).”

First, 2,097 HE participants in the CGSS were selected and a new variable was computed which equals “the year you complete the highest education level (including the current one in progress)” minus four (the common length of Chinese undergraduate education). This calculation was made to find the year the student started HE, as HE is one of the most important and common ways to change Hukou status so the status before HE is crucial. Then, the newly computed starting HE year variable was combined with "the year you complete the highest education level (including the current one in progress)" for non-HE participants. This became a new variable called ‘the year you complete the highest education level (below HE) or enter HE.’ Finally, if the year when people got urban hukou is earlier than the year in which they obtained their highest education qualification or went to HE, they were classified as ‘urban hukou’ residents. Otherwise, they were labelled ‘rural hukou’ residents. These re-coded values are put in a newly computed variable called “original hukou status”. The analysis involving hukou in the CGSS is based on this variable.

The CFPS is a longitudinal study so it is easier to trace students’ hukou status before they went into HE. Fewer adjustments were therefore needed.

9.3.3 Misuse of P Values

Some studies choose reasonable designs and variables but they seem to encounter pitfalls with statistical derivatives. The p-values, from significance tests, for example, are widely used in presenting findings in studies. However, there are some problems with this. First, the p-value or significance test is only valid and worth discussing when the sample is randomised, which while not impossible is very rare (Gorard, 2021). It cannot be used with completely population data, convenience samples or opportunity samples, or where there are missing cases or answers. Therefore, due to non-responses, attrition and other missing data, regarding the CGSS, the CFPS and other datasets it is not appropriate to discuss this statistical derivative because there has been no randomisation.

The second problem is that the function of the p-value is often incorrectly understood. Several studies use the p-value or significance test to predict the population, or specifically to see to what extent the results gained from the cases examined can represent the population. However, the p-value or significance test does not work like this. "A low probability of the difference observed in the sample" does not necessarily equal "a high probability of a difference in the population" but could refer to "a high, middling or low probability for the population" (Gorard, See & Davies, 2011, p.181). Therefore, although they are popularly employed, even if there is an ideal dataset with a completely randomised sample and with no missing data, the p-value, significance tests and other sampling theory derivatives are problematically used in many studies.

This study will not use statistical derivatives such as the p-value or significance tests when presenting the results, because the datasets are non-random surveys

9.3.4 Ignoring Missing Data

Another problem is missing data. As has been previously mentioned, ideal datasets with no missing data are rare. Missing data should not be ignored because they are unlikely to be randomly missing. Instead, the missing cases in longitudinal studies are likely to be the disadvantaged ones, and are often the group targeted by the research. On the other hand, in cross-sectional studies, where it might seem that there is less need to worry about missing cases, there may be problems of missing data in answers to sensitive questions. According to Siddiqui, Boliver & Gorard (2019), for instance, children with low SES backgrounds are more likely to drop out of studies, and questions about family income might receive fewer responses.

However, few of the studies collected in the structured review pay attention to missing data. Many previous studies directly delete cases with missing data (Deng & Fu, 2020; Du, Ni & Huang, 2018; Jin & Li, 2022; Li & Lu, 2015; Ou & Hou, 2019; Pang, 2019; Wu & Huang, 2015; Wu et al., 2020; Zhang, K., 2017; Zhang & Chen, 2013; Zhou, Li & Cui, 2018; Zhou et al., 2018) or do not even mention them (Bai, 2018; Li, 2015; Sun & Yan, 2015; Tao, Wang, H., 2016; Wang & Zhang, 2017; Wang & Zhao, 2020; Wei, 2013a; Wei, H. L., 2016; Wen, 2022; Wu, 2019; Xu, J. Y.,2017; Yang, 2020; Zhang, 2015).

The present study regards 'missing data' as a category in the variables (e.g. the hukou status variable can take the values 'rural hukou,' 'urban hukou' and 'missing hukou status') in the analysis of CGSS and CFPS data.

9.3.5 Ignorance of the Relationships Between Indicators

Although previous studies evaluated a great number of indicators, they rarely examine the relationships between potential indicators. However, this is necessary as it is possible that people identified with several indicators of disadvantage are likely to be double or even triple disadvantaged. For instance, minority ethnicity students might not necessarily be disadvantaged in HE participation, but a minority ethnicity student with rural hukou living in a rural area and from a low SES family is highly likely to be a valid target of CA policies.

Therefore, this study will use cross-tabulations to explore underlying relationships between the indicators discussed. The aim is, on the one hand, to reveal the possibility of an ecological fallacy for some indicators and, on the other hand, to find associations between indicators.

9.3.6 Outdated Data

Some studies evaluating province-based admission policy quotas collected somewhat outdated data on quota plans which are a decade old (Wang, 2019; Wei & Liu, 2015; Zhang & Zhang, 2015) or even two decades old (Bao, 2011; Ji & Zhu, 2011; Liu, X., 2011; Luo, 2011; Yang, 2014). Although they are valuable in explaining HE disparities in provinces, it would be more interesting and relevant if a more recent picture could be known.

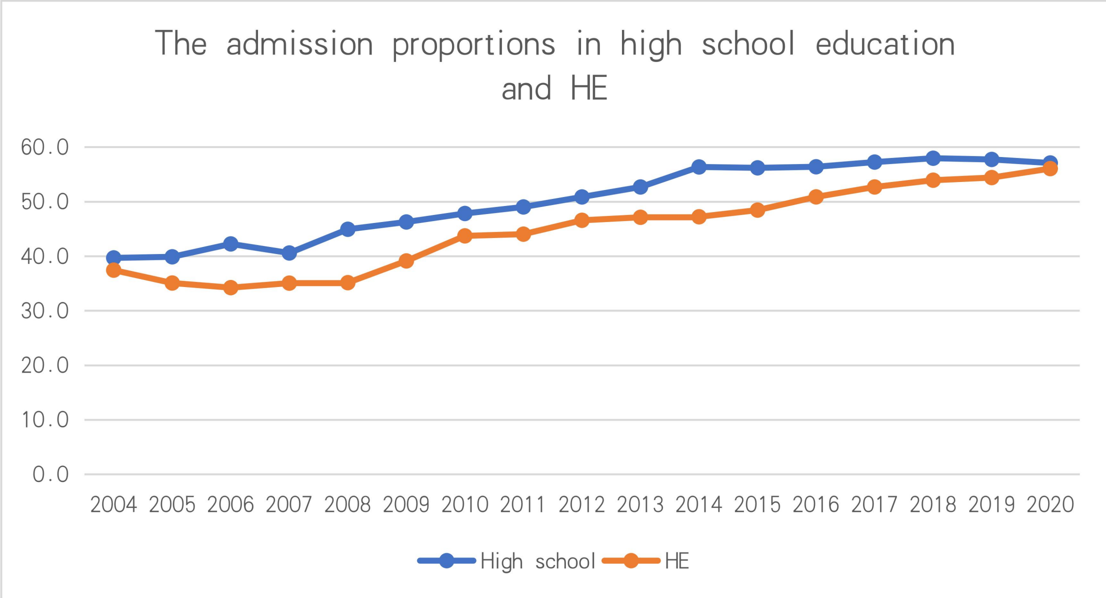
Therefore, this survey collects provincial quota plans of DFC universities, provincial HE enrolment numbers, and provincial NCEE candidate numbers from 2016 to 2019.

9.3.7 Ignorance of Earlier Gaps in HE Participation

It is well known that students have to continue into high school and complete high school education before they are able to apply for HE. That is to say, the NCEE is not the only or the first selection of students in HE admissions. Students who were not selected for high school education cannot participate in HE. However, little previous research that focuses on HE enrolments pays attention to this earlier gap.

According to Figure 9.3.1, high school admission rates, which are the result of high school enrolment numbers divided by middle school graduation numbers, and HE admission rates, which are the result of HE enrolment numbers divided by high school graduation numbers, both show a slow increase over 17 years. The line representing the HE admission rate increases from around 40% to nearly 60%. This seems a good achievement. However, the actual HE admission opportunity for students is far lower than this. Even if we assume that there is no omission of students except for two selections – students that complete compulsory education including primary school education and middle school education and students who successfully transit to high schools and finish high school education – the real change in the proportion of HE admissions is only from 14% (2006) to 35% (2020). Considering that the assumption is unlikely, the figure may be even lower.

Figure 9.3.1 Proportions of high school and HE admissions from 2004 to 2020 in China



Source: NBS

Note: HE enrolment numbers before 2004 do not distinguish between four-year academic universities and three-year vocational colleges, so data before 2004 are not presented.

Furthermore, students who finally successfully enter HEIs, especially prestigious HEIs, are likely to be ones who not only went to high school but also who have gone to key high schools or selective high schools (Ye, 2015). It may be more difficult for students who leave academic education after compulsory education and accept technical vocational education to go to university (Ling, 2015).

Although it might not be easy to know the corresponding figures for every step in the education process, we at least know the gaps between the two official selections. Therefore, it would be inappropriate to ignore those that left academic education before high school education, who are often the disadvantaged ones (Wang & Guo, 2019), when exploring equality and equity in HE participation.

This study reveals the gaps between the two selections using empirical evidence and finds whether there are accumulative disadvantages.

9.4 Discussion and Conclusion

This chapter has first summarised the findings of previous studies. According to these studies, there are disparities in HE participation between hukou statuses and different provinces. For instance, rural hukou students and students from provinces with large populations or under-developed economies or education are always under-represented in HEIs and prestigious HEIs. Nevertheless, there seem to be no remarkable gaps between Han majority and minority ethnicities, which had previously been a long-lasting historical problem.

As for other indicators, family socio-economic status, including parents' education qualifications, parents' occupations, parents' political party memberships and the family financial situation, is widely acknowledged as relevant to HE participation and educational attainment. School-level indicators are also regarded as contributing to HE attendance but they are likely to be both intermediaries of HE inequality and results of stratified backgrounds. Many researchers claim that sex disparities have been largely narrowed, but it would be dangerous to ignore inequalities within each sex. For example, females with rural hukou from poor families and belonging to minority ethnic groups might still suffer from fewer educational opportunities, although their urban counterparts might be found to exceed males in HE enrolment numbers. Furthermore, according to Effectively Maintained Inequality hypothesis, differences between sexes in the quality of HEIs attended and the majors studied might deserve more attention .

Despite many useful conclusions, some limitations are mentioned of the studies collected in this structured review. Some employ less-appropriate research designs, which might fail to logically support the conclusions. Some use reasonable research designs but are over-ambitious and make causal claims on the basis of correlational designs. Furthermore, some studies analyse large-scale survey data but select problematic variables. Some are obsessed with the p-value, which is actually not appropriate in presenting their results. Some studies aiming to evaluate HE inequality between provinces use somewhat outdated data.

Most studies in the review directly ignore missing data and the relationships between disadvantage indicators. Ignorance of these two issues might lead to bias in the results. More importantly, few studies recognise the significance of transitioning to high school education and explore this in depth.

Therefore, this new study aims to avoid these limitations as much as possible, demonstrate the clustering of HE participation by different indicators and identify good-quality indicators of disadvantage. The next chapter evaluates the province indicator.

Chapter 10 Evaluation of the Province Indicator

As was explained in Chapter 4, the province-based quota admission policy is the primary admission policy in HE enrolment in China. It regards each province in China as an admission unit and allocates different admission quotas. This policy was first implemented in order to improve HE participation by students from economically disadvantaged provinces (Qin & Buchanan, 2019), but it has been criticised for a long time for being unsatisfactorily implemented, as the structured review in Chapter 9 showed. However, many of the criticisms in Chapter 9 are based on relatively outdated or inadequate data, so more recent evidence is needed on how much this policy has achieved its aim or how fair it has been.

This chapter answers the research questions of what disparities in HE participation are revealed by the province indicator and the extent to which it is reasonable to use province as a CA indicator. The chapter first reveals how HE participation and elite university participation varies from province to province by means of analyses using three different indices: admission rates, Gorard segregation index and admission opportunity rates. The admission quotas of 22 DFC universities were collected for 2016, 25 for 2017, 31 for 2018 and 34 for 2019. Information on the data from which DFC university admission quotas were found is listed in Appendix 4. As for the admissions of regular HEIs, data on 17 provinces for 2016, 13 provinces for 2017, 24 provinces for 2018 and 20 provinces for 2019 were collected. The detailed admissions of provinces each year are listed in Appendix 5. At the end of this chapter, the quality of the province indicator is discussed.

10.1 Comparisons of HE Admission Rates in Provinces

The Admission Rates Index is one of the most easily understandable and commonly used indices to compare admission differences between groups. In this section this index is employed to investigate disparities between provinces in the proportions of students entering elite universities and regular universities.

10.1.1 Provincial Admission Rates of DFC Universities

Figures 10.1.1, 10.1.2, 10.1.3 and 10.1.4 plot the admission rates for Chinese provinces by DFC universities in the years 2016-2019. The Y-axis shows the proportion of university admissions for the provinces by DFC universities. Besides the columns in the figures representing these proportions, there are two lines which provide relatively constant references to make comparisons between the provinces more intuitive.

The line is the average proportion of DFC university admissions for individual provinces. It is calculated as follows:

$$\text{Average Admission Rate} = (\textstyle\sum_{x=1}^{x=i} ARx)i \ ,$$

where:

ARx are the admission rates;

i is the number of provinces.

Figure 10.1.1 shows the admission rates of DFC universities for provinces in 2016. Shanghai, Jilin, Beijing, Tianjin, Fujian, Qinghai, Liaoning, Hainan and Ningxia have admission rates above the average, Guangdong, Shaanxi, Henan and Hubei are closer to the average and Shandong, Hunan, Jiangsu, Xinjiang, Hebei, Shanxi, Inner Mongolia, Zhejiang, Guangxi, Jiangxi, Chongqing, Anhui, Yunnan, Gansu, Tibet and Guizhou are more disadvantaged and the rates for these areas are much lower than the average reference line.

Figure 10.1.2 shows that in 2017 Chongqing and Heilongjiang were added to the group of provinces advantaged in 2016. On the other hand, there were few changes to the disadvantaged group.

Figure 10.1.1 Admission rates of DFC universities for provinces in 2016

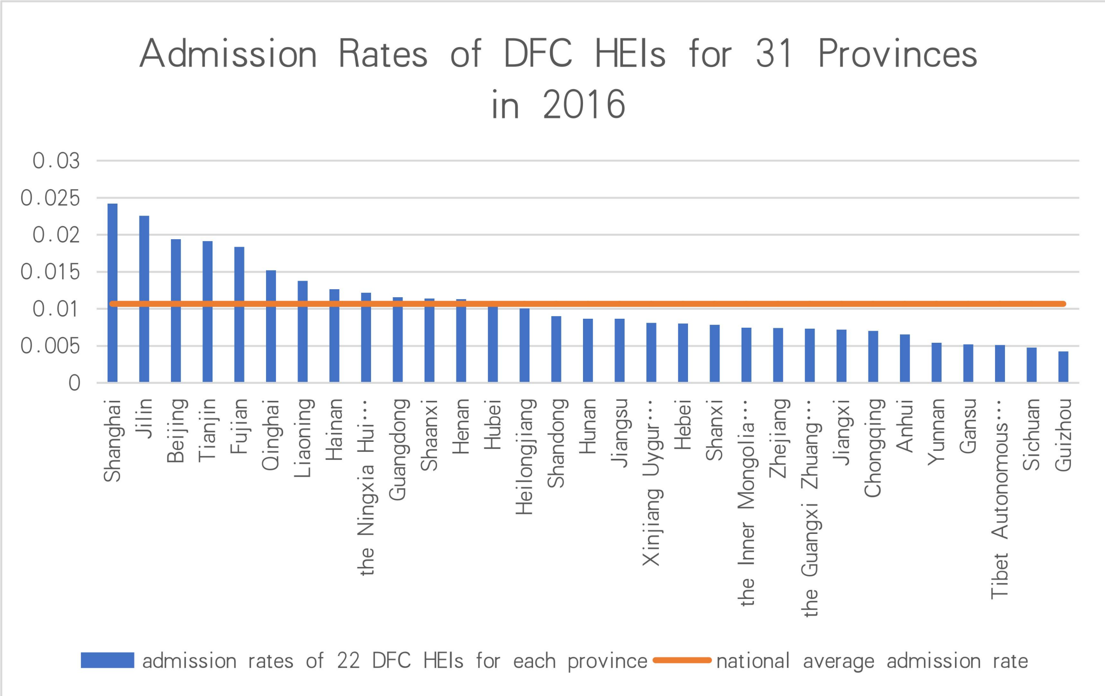


Figure 10.1.2 Admission rates of DFC universities for provinces in 2017

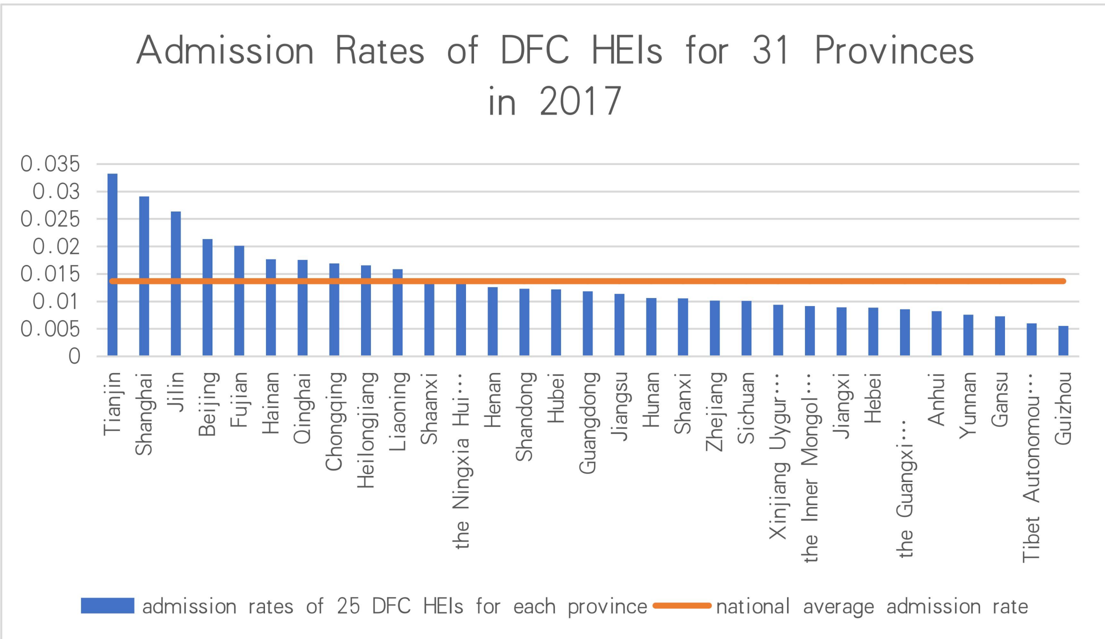


Figure 10.1.3 and Figure 10.1.4 display the admission rates of DFC universities for 31 provinces in 2018 and for 29 provinces in 2019 (data from Tibet and Xinjiang are missing) respectively. Both figures echo the patterns of advantaged and disadvantaged provinces in

2016 and 2017.

Figure 10.1.3 Admission rates of DFC universities for provinces in 2018

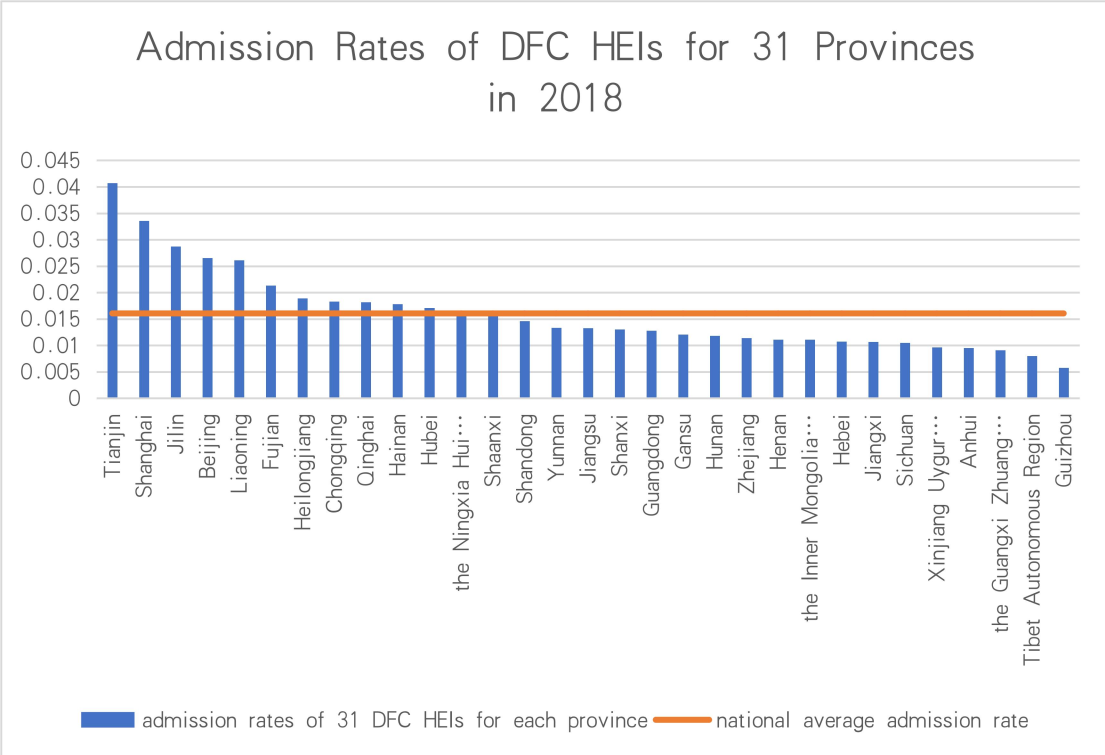


Figure 10.1.5 clusters the admission rates of DFC universities for each province from 2016 to 2019 in order to make the trends more visible. There is a stable advantage of some provinces, Tianjin, Beijing and Shanghai, and there are some long-term disadvantaged provinces such as Guizhou, Guangxi and Anhui. As the figure shows, the pattern of prestigious university admission quotas among the provinces remains mostly the same over the four years. As the completeness of the DFC university admission quota plans collected is different for 2016-17 and 2018-19, however, there are some ‘strange’ increases and decreases in the admission quotas for certain provinces. The reasons for these abnormal ups and downs are discussed at the end of this section.

To sum up, there are some gaps in the provincial admission rates of DFC universities. First, the admission rates for Beijing, Tianjin and Shanghai are always much higher than those for other provinces and nearly twice the average. This echoes the findings of Liu (2014), Yang (2014) and Zhang and Li (2019) when discussing elite HEI enrolments that these three are a group of " absolutely superior regions " (Zhang & Li, 2019). Beijing, Tianjin and Shanghai are in eastern China, the best developed region with more abundant economic and educational resources. They might have invested more resources in HE development and have earned more places in HEIs, even selective HEIs. Students from these three municipalities benefit from the number of places and are more likely to win in the competition for HE.

Figure 10.1.4 Admission rates of DFC universities for provinces in 2019

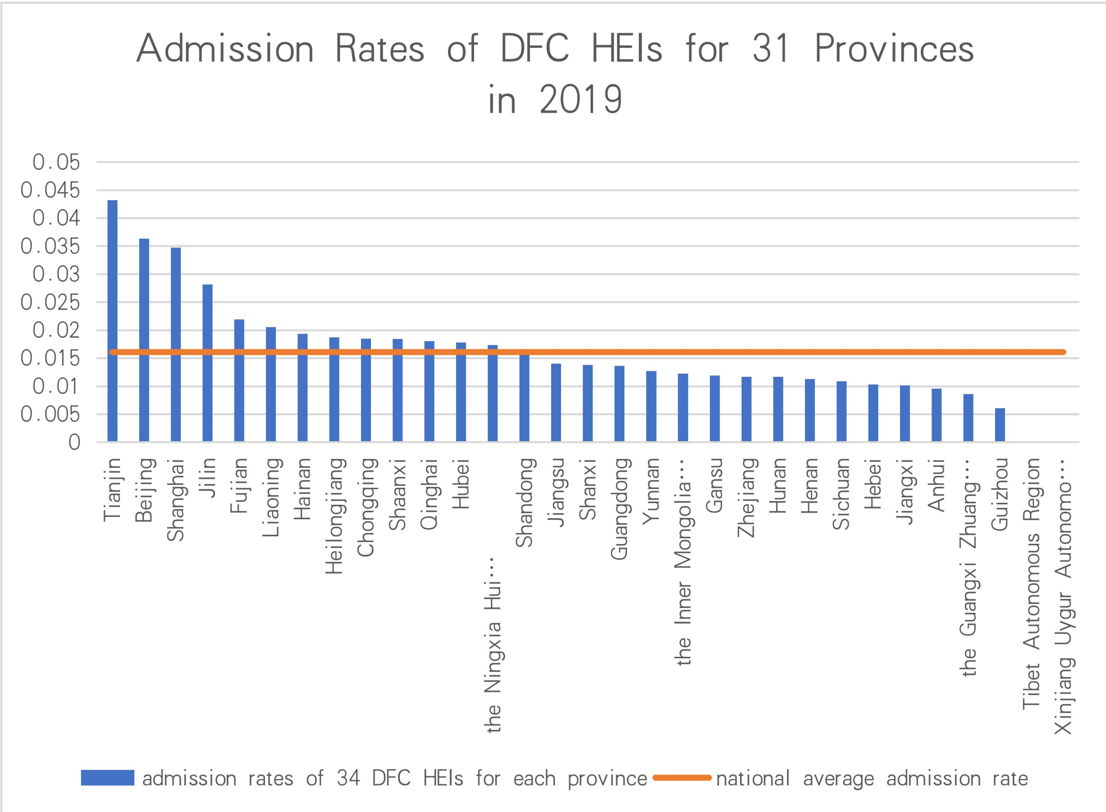
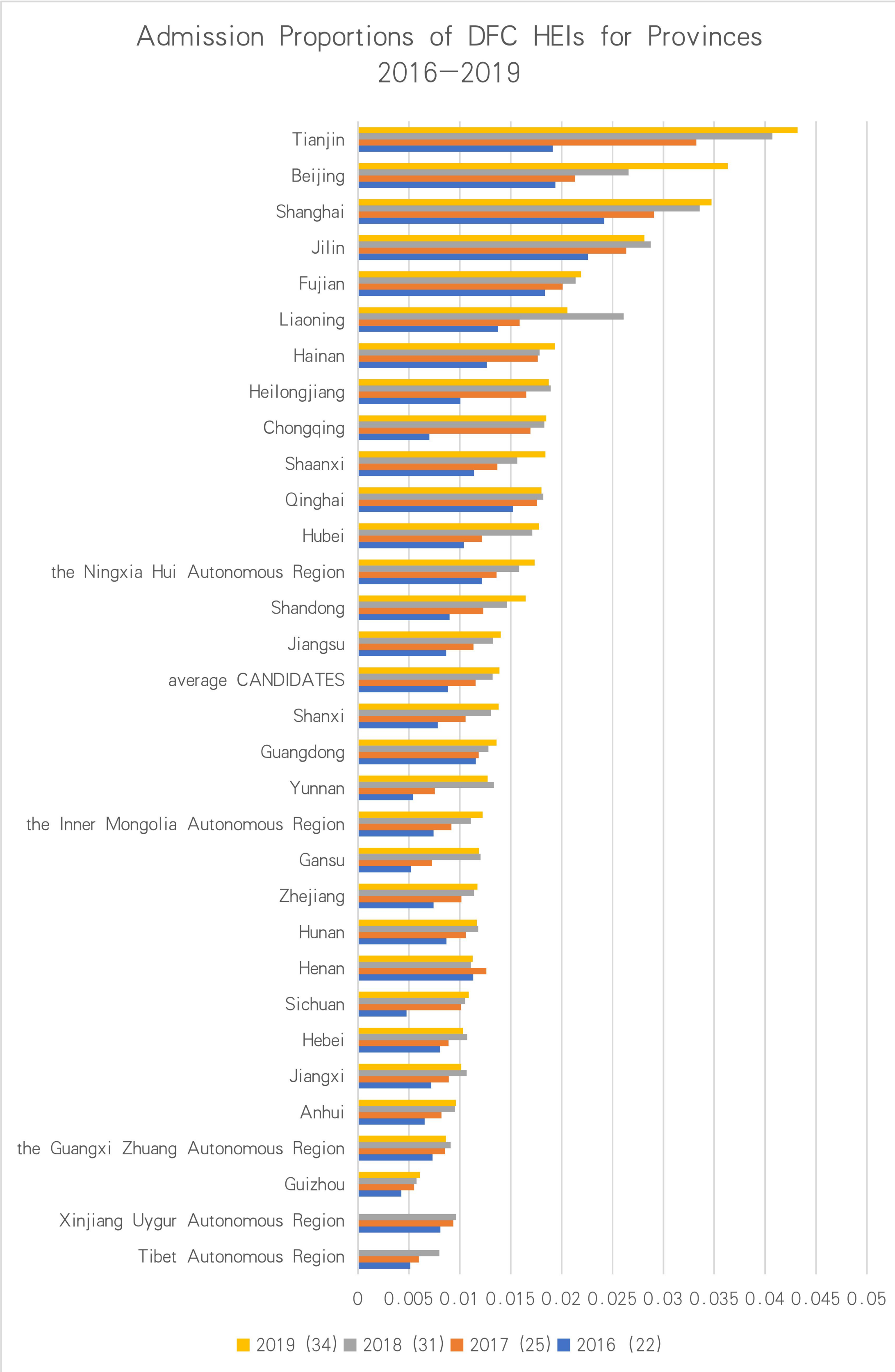


Figure 10.1.5 Admission Rates of DFC universities for provinces from 2016 to 2019



Note: the provinces are listed in the order of the 2019 hierarchy; the figures in the brackets refer to the numbers of

DFC HEIs successfully collected in the corresponding year

Apart from the above three municipalities, the figures also reveal another advantaged province: Jilin. Jilin presents extremely high proportions of admissions to DFC universities, sometimes even exceeding those of the three traditionally privileged municipalities. This high admission rate for Jilin might mainly be attributable to Jilin University, a DFC university located in Jilin. Jilin University not only enrols a large population of freshers every year but also distributes its admission quotas in a highly localised way (see Table 10.1.1 below). As a result, the admission rates of DFC universities for Jilin were outstanding during the years analysed. On the other hand, missing data on the admissions of other DFC universities might lead to under-estimations of the admissions in other provinces. The details will be discussed later.

Figure 10.1.6 Map of Chinese provinces

Second, despite the considerable gaps in elite universities' admission proportions between absolutely superior regions and other regions, there were still some relatively advantaged provinces where the admission rates surpassed the two reference lines. These regions were Hainan, Fujian, Liaoning, Chongqing and Qinghai. It is not easy to identify why these provinces would be given more quotas by DFC universities. They are not located in the same geographical area or nearly. Hainan is in the south of China and Fujian in the southeast (see Figure 10.1.6). Liaoning is in the northeast while Chongqing and Qinghai are in the western centre. Besides, they are also different in numbers of HEIs and many of them even lack prestigious HEIs, such as Hainan, Qinghai and Fujian (see Chapter 2). Their small advantages might result from incomplete data, but on the other hand there might be more complicated reasons for them.

Third, in contrast, some provinces are always lagging behind in the competition for places in elite HEIs. Some of these are in eastern China, such as Hebei and Zhejiang, and some are located in central China, such as Anhui, Jiangxi and Hunan. Nevertheless, most of the disadvantaged provinces are in western China: Sichuan, Guizhou, Yunnan, Gansu, Inner Mongolia, Guangxi, Tibet and Xinjiang. The admission rates of DFC universities for all these provinces fell behind the average in 2016, 2017, 2018 and 2019. Apart from Hunan, the provinces in this disadvantaged group seem to suffer from a lack of prestigious HEIs. Every 10 thousand NCEE candidates in these provinces have only one or even less than one DFC university in their province, as was shown in Chapter 2.

However, the findings should be dealt with cautiously because of missing data. Some DFC universities do not publish their admission quota plans on their website, and this missing data risks misunderstanding and bias. For example, as mentioned above, Jilin had an outstandingly high elite university admission rate in the four years studied, but the reason might be missing data. The admission quota plans of Jilin University were all collected, while the data for other universities are missing for some years. Due to localisation in the distribution of admission quota plans (Table 10.1.1), in the calculation the complete data for

Jilin University directly increased the number of students from Jilin that elite universities successfully accepted.

Table 10.1.1 Localised Admission Percentages of DFC universities (%)

Province	DFC HEI	2016	2017	2018	2019
Beijing	Peking University	16.19			
	Tsinghua University				
	Renmin University				10.22
	Beijing Institution of Technology	7.36	7.12	7.17	7.19
	Beihang University				9.35
	China Agriculture University			5.1	5.24
	Beijing Normal University	5.86	5.84	5.82	5.36
	Minzu University of China				2.99
Tianjin	Nankai University		19.18	18.96	19.53
	Tianjin University				
Liaoning	Dalian University of Technology			26.17	24.83
	Northeastern University	25.08	20.55	19.9	20.61
Jilin	Jilin University	23.68	23.38	23.87	23.61
Heilongjiang	Harbin Institution of Technology		24.63	27.25	26.89
Shanghai	Fudan University		8.57	9.69	9.5
	Tongji University	10.19	9.78	13.12	15.16
	Shanghai Jiaotong University	11.05	9.04	8.67	8.19
	East China Normal University	22.17	23.14	23.9	23.29
Jiangsu	Nanjing University				
	Southeast University	21.07	19.56	19.36	19.33
Zhejiang	Zhejiang University				
Anhui	University of Science and Technology of China	15.74	15.18	15.12	10.08
Fujian	Xiamen University	31.51	30.29	31.74	30.71
Shandong	Shandong University	36.51	36.84	35.88	34.83
	Ocean University of China	29.91	29.85	29.94	29.79
Hubei	Wuhan University	26.29	25.48	26.58	25.77
	Huazhong University of Science and Technology			28.17	27.35
Hunan	Zhongnan University	10.5	9.84	10.13	10.35
	Hunan University	11.51	12.7	12.64	12.85
Guangdong	Zhongshan University	55.04	54.69	54.59	53.4
	South China University of Technology	57.24	56.2	56.85	55.09
Sichuan	Sichuan University		30.34	30.12	28.8
	University of Electronic Science and Technology of China				

Chongqing	Chongqing University		30.4	30.79	29.85
Shaanxi	Xi'an Jiaotong University		30.24	29.33	29.07
	Northwestern polytechnical University	25.12			24.93
	Northewest A & F University	26.82	27.79	27.89	27.75
Gansu	Lanzhou University			26.86	18.88
Henan	Zhengzhou University	52.44	55.59	50.79	51.43
Xinjiang	Xinjiang University				
Yunnan	Yunnan University			40.96	37.56

On the other hand, missing information about other DFC universities not in Jilin might also mislead people. The lack of data on the admission quota plans of other DFC universities does not increase the admission rates of other provinces where they are located, which might make an indirect contribution to the notable advantage of Jilin in the analysis. In short, the missing data are likely to unintentionally include more students in Jilin as successful enrolees in elite universities while excluding more from other provinces. The advantages in admission rates of Jilin students by DFC universities might therefore be an overestimate.

Another example of a possible error due to missing data is Xinjiang. The Xinjiang Uygur Autonomous Region was disadvantaged in DFC university admissions from 2016 to 2018 (the data for 2019 are missing). This long-term disadvantage may have been for several reasons such as fewer educational resources, an underdeveloped economy or even HE admission discrimination, but it also might be a missing data error. The admission data for Xinjiang University, a DFC university located in Xinjiang are not known for any of the years studied. If these data had been collected, the admission rates for Xinjiang in these years might have been higher. This does not mean that Xinjiang would no longer belong to the disadvantaged group of provinces if the admission data for Xinjiang university were known. Instead, the data would be powerful enough to warn us that students from Xinjiang are at least in a disadvantaged position in the admission plans of DFC universities. They might deserve more attention and additional help in the competition for places in elite HEIs. More complete data are therefore important.

Finally, this section will simply compare the localisation of admissions in DFC universities as is shown in Table 10.1.1. It would be fair to state that many DFC universities show a tendency to admit more students from the provinces where they are located, although most of them keep the local admission rate under 30% as they are required to do by regulation (Qu, 2009). What might deserve more attention is that although it seems DFC universities in Beijing do not show extremely highly localised admissions compared to other universities, in fact students from Beijing still benefit a lot because the number of DFC universities in Beijing is the largest, nearly a quarter of the sum of DFC universities in China.

However, there are four exceptions to this localisation, namely China Agriculture University in Beijing, Beijing Normal University in Beijing, Minzu University of China in Beijing and Shanghai Jiaotong University in Shanghai. China Agriculture University allocated the most places to Shandong, with 22.18% and 21.99% in 2018 and 2019 respectively, rather than Beijing. The reason might be that Shandong province is one of the most advanced agricultural provinces in China. This province provides vegetables and cereals to the whole nation. This agricultural orientation not only makes Shandong need more professional and qualified labourers in the agriculture sector but also means that there are many farmlands and croplands for students who study agriculture to practise.

Another exception is that Beijing Normal University does not show a strong localised or particular provincial enrolment trend. Its highest admission rates are for Shandong, Henan and Sichuan, with a little more than 7%.

A third exception is Minzu University of China. Although its admission plans allocate places relatively evenly among the provinces, it did favour several particular ones in 2019, such as Jilin (5.61%), Inner Mongolia (5.98%), Guangxi (5.71%), Xinjiang (5.71%), Yunnan (5.13%), Hunan (4.91%), Sichuan (4.75%) and Guizhou (4.75%), all of which have higher quotas than the 2.99% for Beijing. A possible reason might be that Minzu University of China is one of the universities for nationalities, which are primarily for minority ethnicity students. The provinces mentioned above have large proportions of minority ethnicity populations. In other words, there are many targeted students in these provinces. Thus, it seems reasonable to give these provinces higher admission quotas.

Last, the admission rates of Shanghai Jiaotong University for Shanghai were slightly lower for Jiangsu in 2016 (11.05% vs. 11.52%), 2017 (9.04% vs. 11.12%), 2018 (8.67% vs. 11.06%) and 2019 (8.19% vs. 10.67%). However, the reason for this exception is not apparent. Jiangsu produces a large number of NCEE candidates and is close to Shanghai, so it seems reasonable for it to have more places in a selective university in Shanghai. In addition, Zhejiang is ultimately the same as Jiangsu in these two characteristics. Nevertheless, Zhejiang is not privileged in the admission rates of Shanghai Jiaotong University. Therefore, geography and numbers of candidates might not be the explanation for this, and other reasons such as government cooperation might need to be considered. The de-localised admissions of four DFC universities are presented in Appendix 6.

10.1.2 Provincial Admission Rates of All Regular HEIs

It would be superficial if the comparison only involved prestigious HEIs, as it might be argued that elite HEIs are intrinsically for meritocracy not equity, for rewarding talent not helping the disadvantaged. The claim is not necessarily valid, but it is worth exploring the student intakes in regular universities.

Figures 10.1.7, 10.1.8, 10.1.9 and 10.1.10 show the provincial admission rates of regular HEIs in China instead of differentiating between elite ones and less-prestigious ones. The formula used here is the same as that used in the last section, but Ai here is the number of students who are successfully accepted by regular HEIs. The X-axis tells us which province the column represents and the Y-axis is still the specific proportion. The line is the average rate of regular HEIs admissions in the provinces.

Figure 10.1.7 and Figure 10.1.8 show the admission rates of regular HEIs in 2016 and 2017. Beijing is notably advantaged in both years, with 66.7% in 2016 and 66.9% in 2017.

Fujian, Hainan, Zhejiang, Hubei and Inner Mongolia also show some advantages in 2016. Jiangsu, Hebei and Inner Mongolia show small advantages in 2017. In both years, Jiangxi, Shandong, Henan, Hunan, Guangdong, Qinghai and Guangxi were in disadvantaged positions, although at different levels.

Figure 10.1.7 Provincial Admission Rates of Regular HEIs in 2016

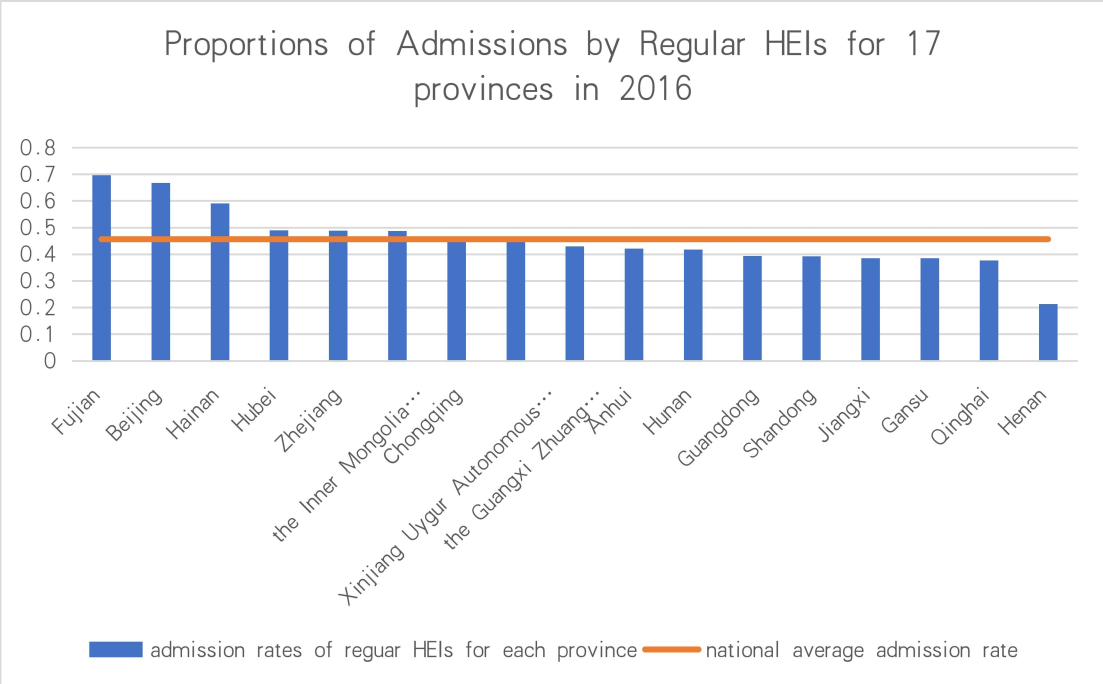
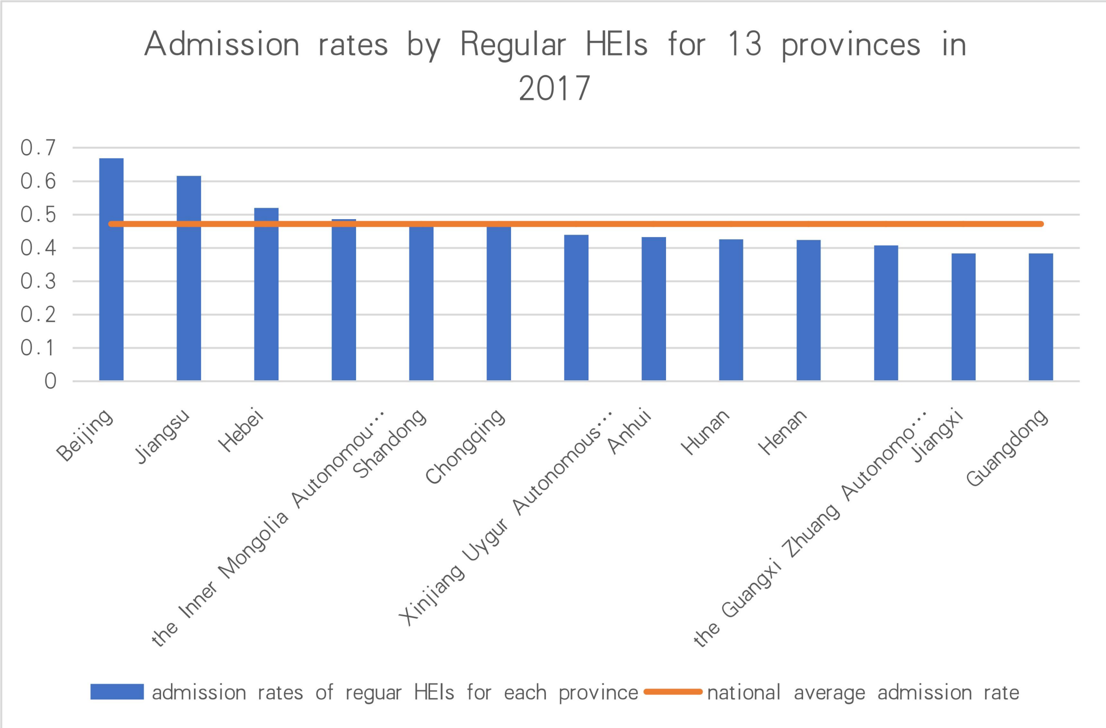


Figure 10.1.8 Provincial Admission Rates of Regular HEIs in 2017



As Figure 10.1.9 shows, in 2018 Liaoning, Tianjin, Beijing, Jiangsu, Fujian and Heilongjiang all exceeded the average to a large extent, while Henan, Ningxia, Yunnan, Gansu, Sichuan, Shanxi, Guangdong, Guangxi, Hunan, Jiangxi, Xinjiang, Anhui, Qinghai, Shandong and Chongqing were disadvantaged and below the average. Figure 10.1.10 shows that there was a similar pattern in 2019. Both the advantaged and disadvantaged groups hardly changed.

Figure 10.1.9 Provincial Admission Rates of Regular HEIs in 2018

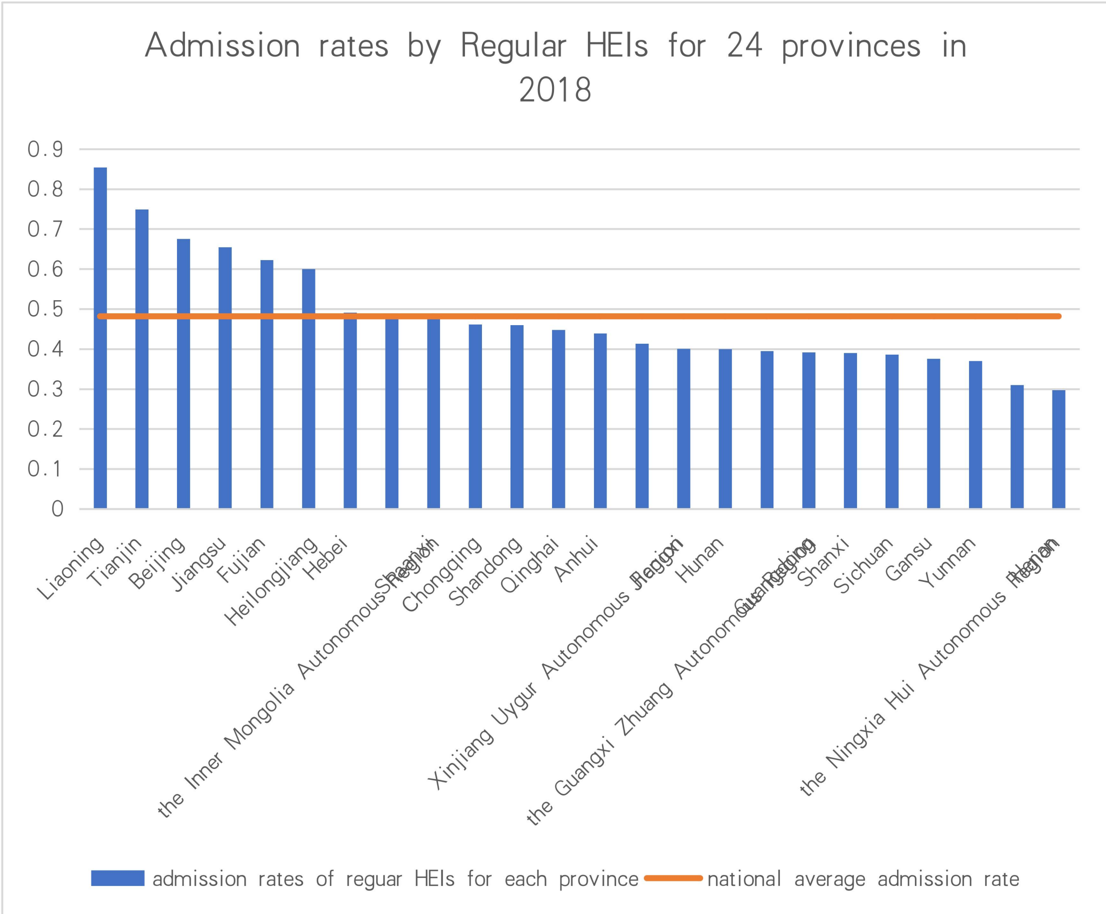
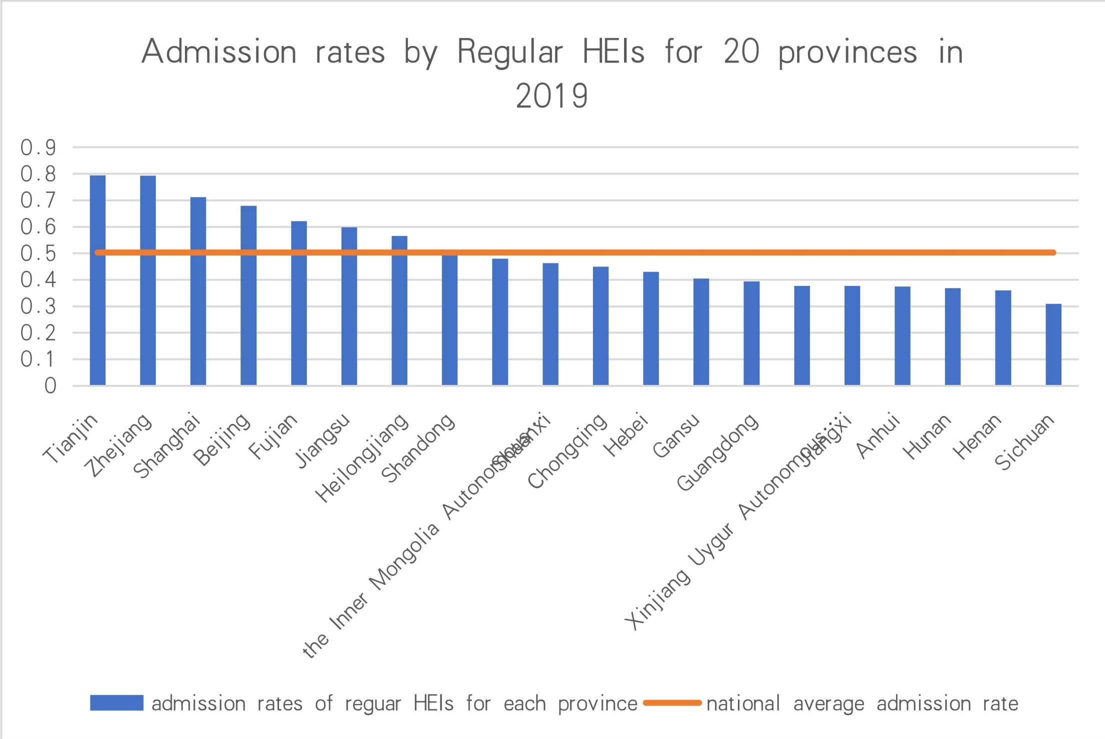


Figure 10.1.11 clusters the HE admission rates for provinces from 2016 to 2019. There do not seem to be great differences in the HE admissions in these years. The advantaged provinces and disadvantaged provinces remain almost the same.

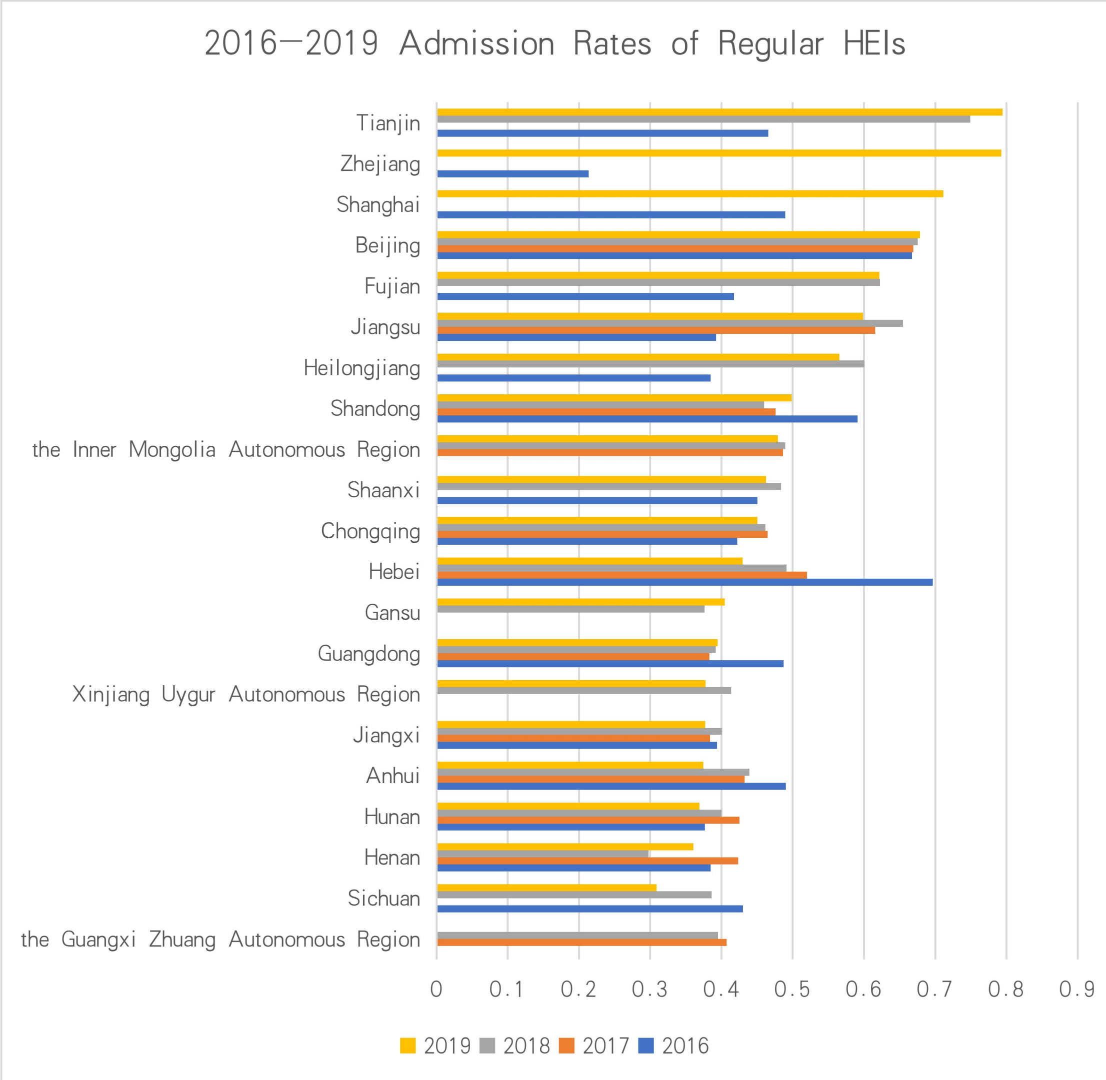
Figure 10.1.10 Provincial Admission Rates of Regular HEIs in 2019



In short, when comparing the provincial admission rates of regular HEIs from 2016 to 2019 it is easy to find some advantaged provinces. Despite a certain lack of data, the following provinces always lead in the HE enrolment competition as long as their data appear in the figures: Beijing (2016, 2017, 2018 and 2019), Fujian (2016, 2018 and 2019), Hainan (2016), Jiangsu (2017, 2018 and 2019), Tianjin (2018 and 2019), Liaoning (2018), Heilongjiang (2018 and 2019), and Shanghai (2019).

In addition, unlike their privileged counterparts, some provinces suffer from low proportions of admissions for a long time. Jiangxi, Henan, Hunan and Guangdong, for instance, are disadvantaged in all four years. Anhui, Gansu and Guangxi also show low admission rates by regular HEIs in three years. Furthermore, although the data on the enrolments by regular HEIs for Sichuan, Yunnan and Qinghai were only collected for one or two years, they are also below the average.

Figure 10.1.11 Provincial Admission Rates of Regular HEIs from 2016 to 2019



Note: only provinces that have data for two years or more are included; the provinces are listed in the order of the 2019 hierarchy

10.1.3 Conclusion

Both when comparing the admission rates of elite universities and when taking a broader view of all regular HEIs, provincial disparities in HE admissions clearly exist.

However, despite its popularity, the Admission Rates Index is somewhat problematic and limited in comparing gaps. Localisation and admission rate calculations only involve the numbers of NCEE candidates and successful enrolees from certain provinces. They ignore the wider population, such as the total number of NCEE candidates in China, which partly compromises the consistency of comparability of the data. A more precise comparison index is needed.

In short, the admission rate and localisation data disclose some problems, but not enough to make reliable comparisons of provincial disparities in HE enrolment. Therefore, a better calculation and a better index are required. The following sections will use the Gorard Segregation Index to calculate segregation indices for every province and make more persuasive comparisons.

10.2 Comparisons of HE Admissions for Provinces using the Gorard Segregation Index

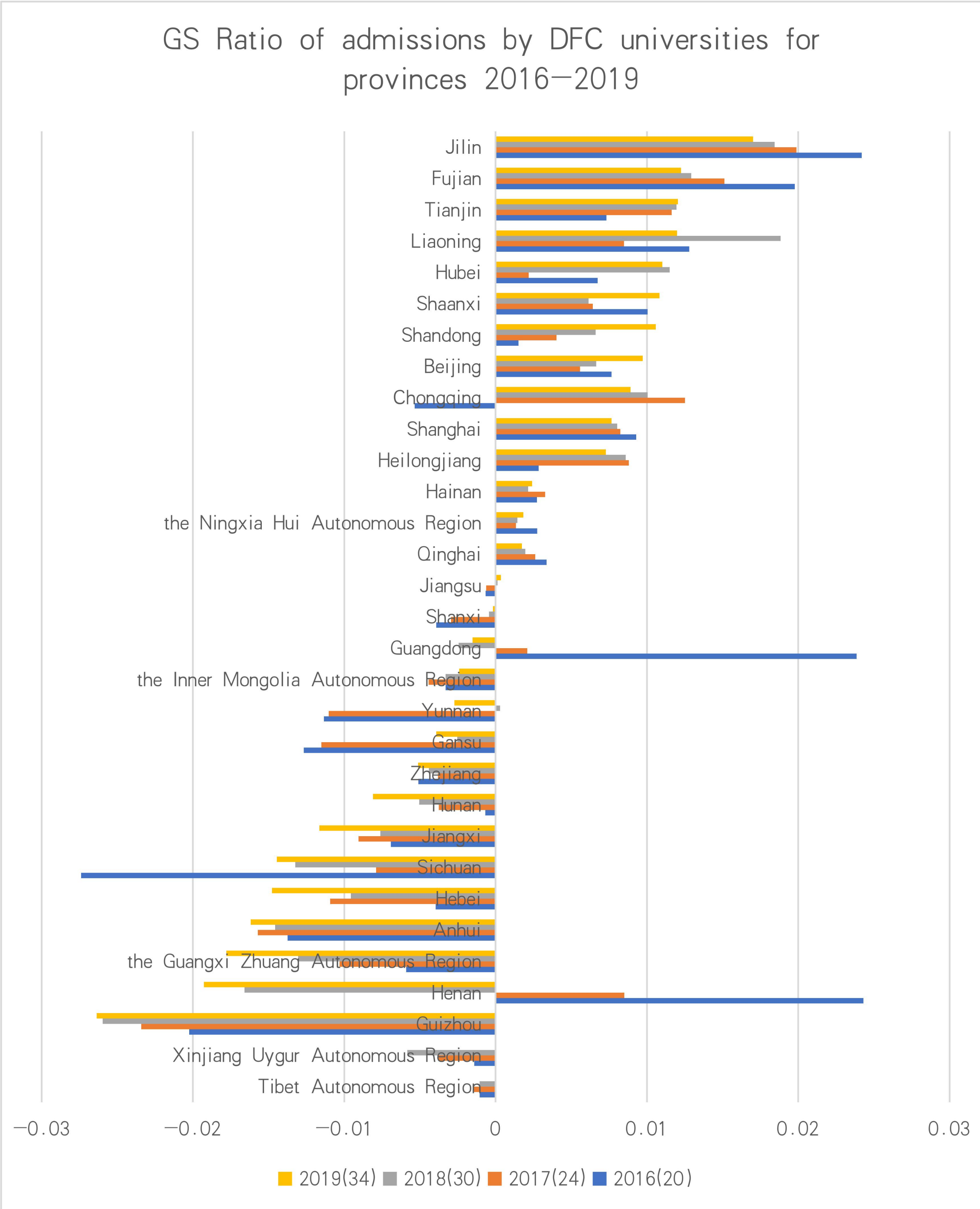
This section employs the Gorard Segregation Index (or GS Ratio) to explore where there are provincial disparities in both prestigious HE admissions and regular HE admissions.

10.2.1 Provincial Segregation of DFC Universities

Figure 10.2.1 shows the GS ratio for the admission quotas of DFC universities for the provinces. According to the figure, Jilin, Fujian, Tianjin, Liaoning, Hubei, Shaanxi, Shandong, Beijing, Shanghai and Heilongjiang are always remarkably advantaged in admissions by prestigious universities. Chongqing, Hainan, Ningxia, and Qinghai, despite some under-representation or lower over-representation, can also be regarded as winners in the competition for places in DFC universities.

In contrast, Inner Mongolia, Yunnan, Gansu, Zhejiang, Hunan, Jiangxi, Sichuan, Hebei, Anhui, Guangxi, Guizhou, Xinjiang and Tibet remain under-represented. Most of these also showed disadvantages when compared using Admission Rates. The separate results for the GS ratio each year are provided in Appendix 7.

Figure 10.2.1 GS ratio of admissions by DFC universities for provinces from 2016 to 2019



Note: the figures in brackets indicate how many DFC universities' data were collected for the corresponding year. The provinces are listed in the order of the 2019 hierarchy.

Despite a generally stable tendency, there are some noteworthy exceptions. First, Henan province showed over-representation in 2016 and 2017 but under-representation in 2018 and

2019. This fluctuation could be an error caused by limited data. There is a DFC university called Zhengzhou University in Henan province, the admission quota plans for 2016-2019 of which were completely collected. They show a highly localised trend (Table 10.1.1). Then, the more data on admission quota plans of other DFC universities are missing, such as the situations in 2016 (22 universities collected) and in 2017 (25 universities collected), the more Henan may be over-advantaged in its GS Ratio for HE admissions. However, given that Henan is a highly competitive province with a large number of NCEE candidates and few elite HEIs, these over-representations may be false.

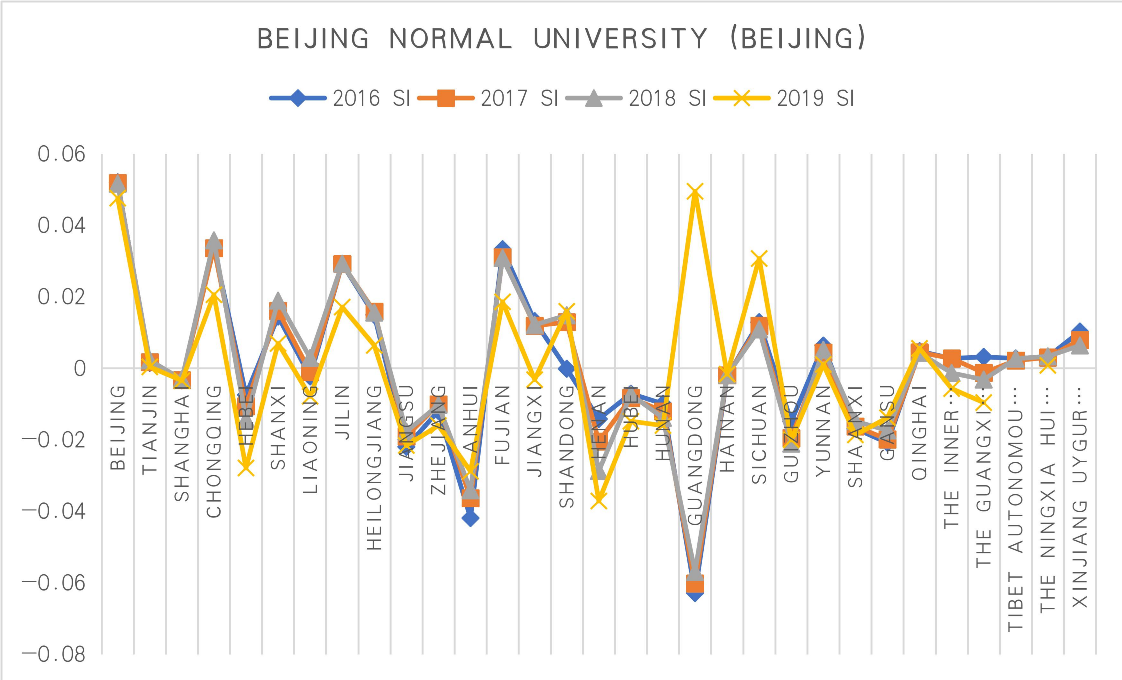
Another example is Guangdong, which shows a similar GS Ratio pattern to Henan. Its over-representation in the first two years is also possibly false because of missing data. That is, Guangdong is one of the best-developed provinces and is completely different to Henan but might also be disadvantaged in admissions by DFC universities. If this disadvantage is true, the reasons might be complicated. As one of the first batch of ‘open-door’ provinces, since the 20th century Guangdong has attracted a large population from neighbouring provinces such as Guangxi, Hunan, Jiangxi and Fujian. This mobility made the population of Guangdong quickly increase in the 1980s and 1990s. This increase then gradually permeated the population of NCEE candidates in the 21st century. In other words, Guangdong has a large population of NCEE candidates. However, the local government in Guangdong might not have invested many resources, or adequate resources, in HE development, or more specifically in elite HE development, to match the considerable size of its population of candidates, so Guangdong's elite HE capacity is limited. The limited number of places in HE and the pressure from the large number of applicants make universities in Guangdong, especially prestigious DFC universities, show highly localised enrolments (Table 10.1.1). But this localised admission does not make things better. As DFC universities in Guangdong do not allocate large quotas to other provinces, DFC universities in other provinces do not allocate many places to Guangdong in return. In the end, places in prestigious HEIs are still insufficient for NCEE candidates in Guangdong, who therefore show surprising under-representation.

There are some other interesting findings. For instance, when it comes to HEIs the data for which have been successfully collected for at least two years, it is easy to find that there are few changes in admission plans each year. In other words, the GS Ratio show almost the same patterns for the four years. However, there still are several institutions which present different trends, which might be worth paying more attention to.

The first exception is Beijing Normal University (Figure 10.2.2). In 2016, 2017, 2018 and 2019, it can be seen that there are opposite GS ratios of admission quota plans for Guangdong province. In the first three years, students from Guangdong province were greatly under-represented in this university but in 2019 they became the most over-represented group. This inconsistency might be attributed to the fact that a branch of Beijing Normal University in Zhuhai City, Guangdong province stopped accepting new students. This branch is called Beijing Normal University (Zhuhai) and it was established by Beijing Normal University in collaboration with the Zhuhai municipal government. The two institutions admitted their students separately before 2019 but in order to respond to government policy, Beijing Normal University (Zhuhai) decreased its enrolment numbers year by year and stopped accepting any new applicants in 2021. Then, Beijing Normal University might have taken some responsibility for the student intakes in its branch. Furthermore, Beijing Normal University (Zhuhai) might prefer to enrol more local students because of a localised admission preference. Therefore, in 2019, Beijing Normal University included more students from Guangdong than it used to. This assumption might provide a possible explanation for the polarised trend in the GS Ratio of student intakes by Beijing Normal University for Guangdong.

The second example is Tongji University (Figure 10.2.3) in Shanghai. Although the figures for most provinces remain stable during the period studied, those for Shanghai and Fujian fluctuate. In 2016 and 2017 students from Fujian were the second most over-represented group at Tongji University but in 2018 and 2019 Fujian had no particular advantage in admissions anymore. On the contrary, the positive segregation level for Shanghai suddenly increased to a large extent in the last two years. That is, more local students were admitted by Tongji University in 2018 and 2019. However, the reasons for these two contemporaneous changes are not clear. More research might be needed to clarify them.

Figure 10.2.2 GS Ratio of admissions by Beijing Normal University for provinces from 2016 to 2019



Another example is East Normal University (Figure 10.2.4), which is also located in Shanghai. According to the graph, Jiangsu province sharply dropped to being under-represented in 2019, while it had been over-represented in the first three years studied. On the other hand, Anhui province also had an unusual fluctuation from initially being over-represented to a slight under-representation in the second year and then back to the original position in 2018 and 2019. These erratic movements might be the result of missing data or cooperation between local governments/DFC universities. For instance, Shanghai was over-represented in admissions by the University of Science and Technology of China in Anhui in 2019, as Figure 10.2.5 shows. This might be an agreed exchange of quotas in DFC universities between Anhui and Shanghai.

Figure 10.2.3 GS Ratio of admissions by Tongji University for provinces from 2016 to 2019

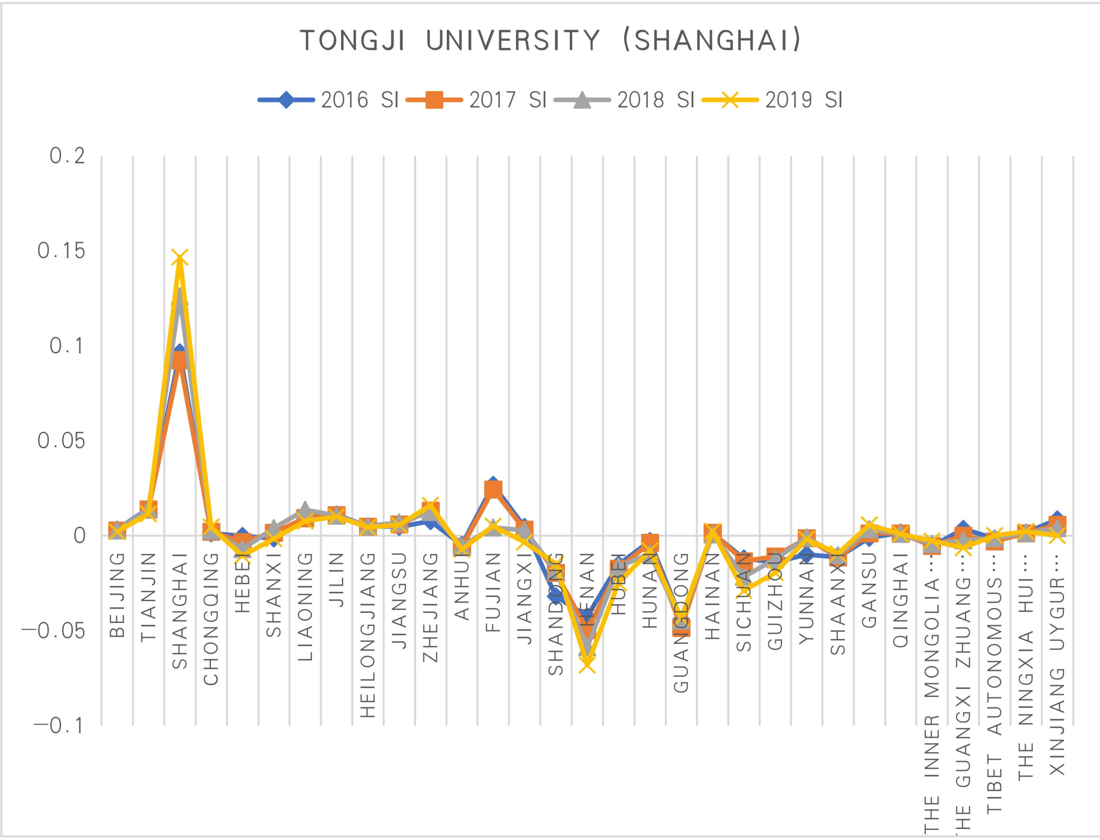
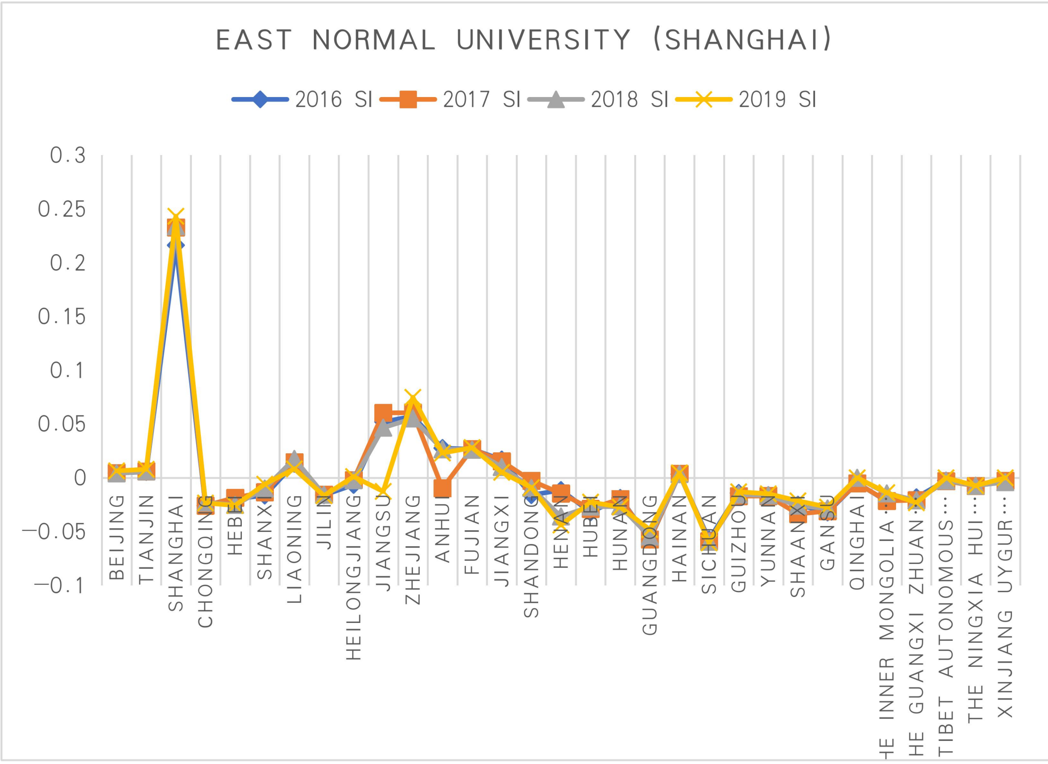
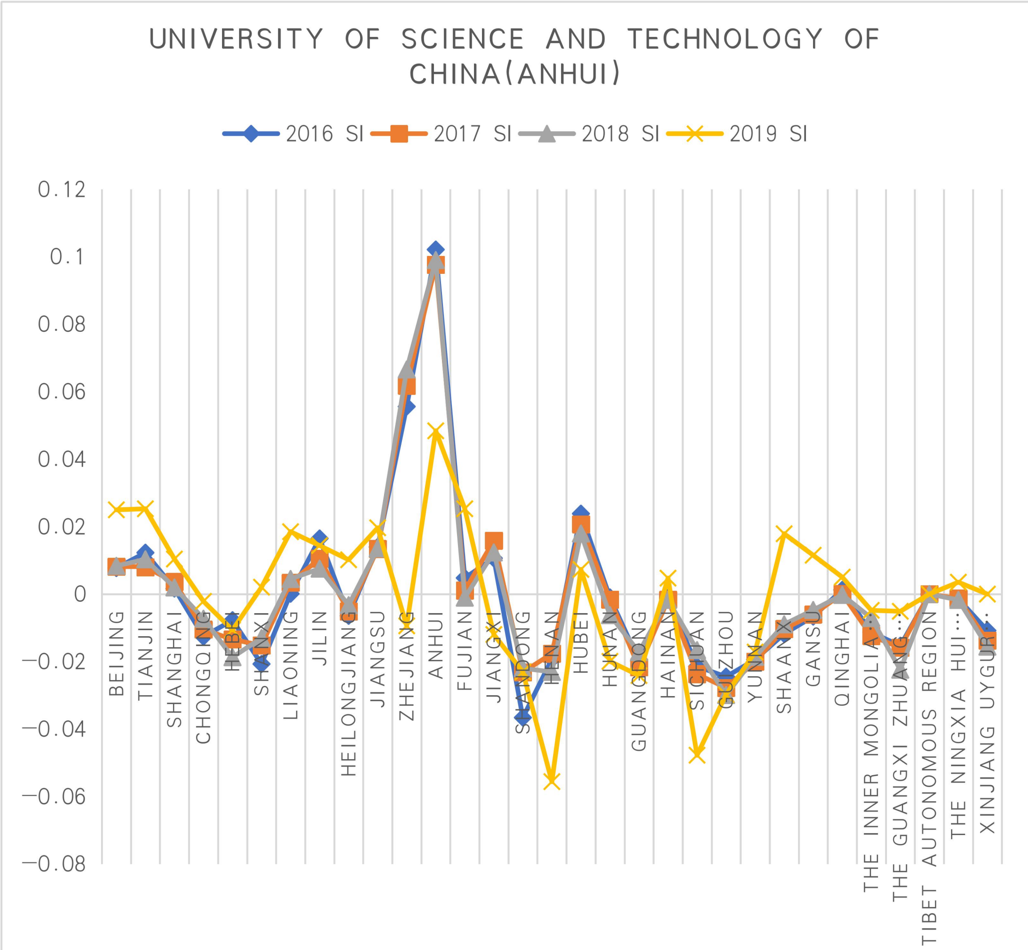


Figure 10.2.4 GS Ratio of admissions by East Normal University for provinces from 2016 to 2019



A more complicated example is the University of Science and Technology of China (Figure 10.2.5) in Anhui province. In the first three years, the segregation indices for every region do not change much. However, in 2019 the changes are notable. First, Beijing, Tianjin, Shanghai and Fujian became more positively segregated. Second, Shanxi, Liaoning, Heilongjiang, Shaanxi and Gansu became a little over-represented after previously being under-represented. Third, students from Zhejiang, Jiangxi and Hunan started to show under-representation in 2019. In addition, Anhui and Hubei became less over-represented, while Henan and Sichuan became more negatively segregated. It is not common to see such massive changes in the admission quota plans of DFC universities. They might be results of new agreements between universities and the government, or side effects of missing data.

Figure 10.2.5 GS Ratio of admissions by the University of Science and Technology of China for provinces from 2016 to 2019



Besides similar admission patterns, almost every elite university reported has a localised admission plan. They tend to allocate more places for local students. This result is not surprising and it largely echoes the findings in Section 10.1 (Table 10.1.1).

10.2.2 Provincial Segregation of All Regular HEIs

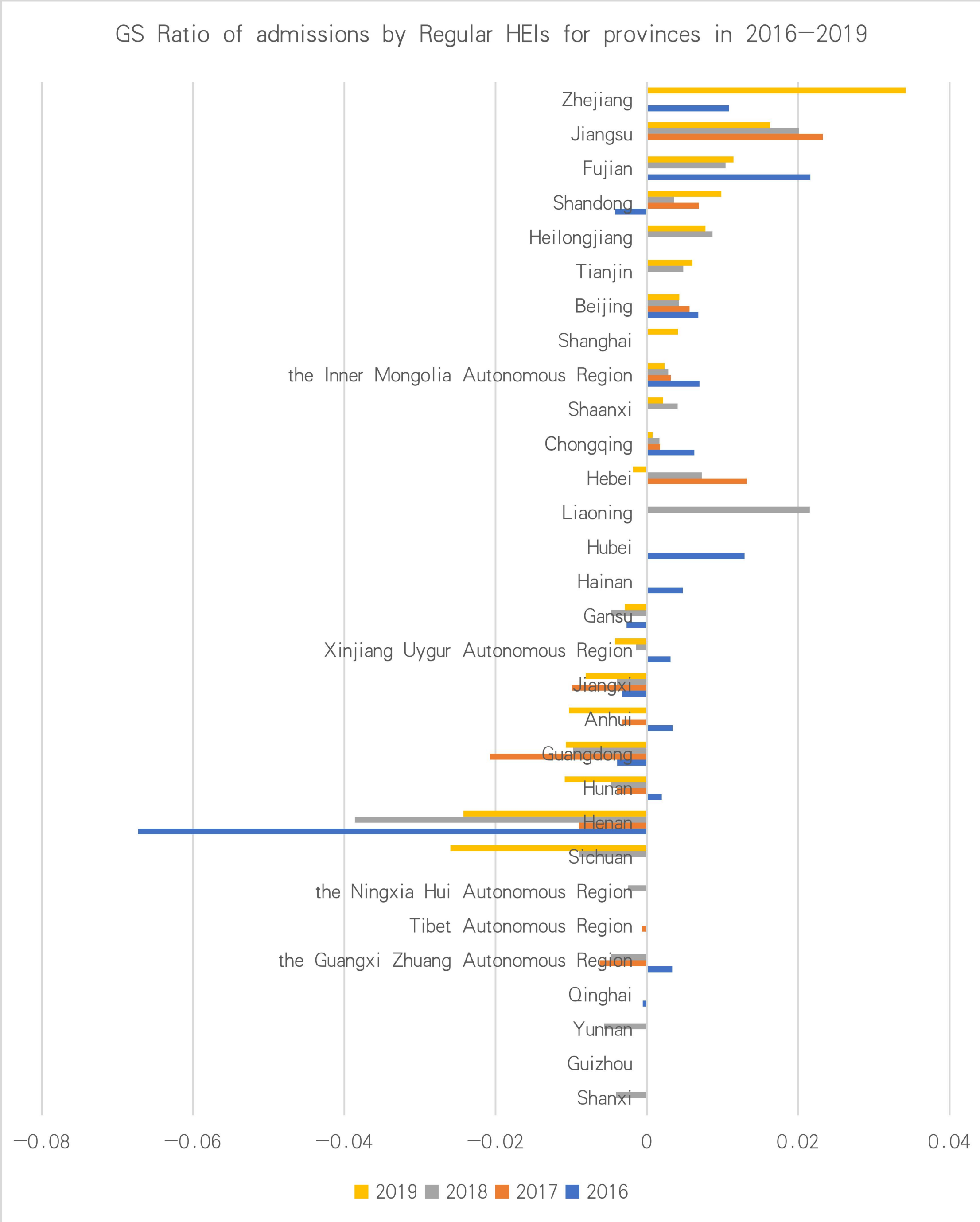
The results would be incomplete if only elite universities were investigated. Therefore, the GS ratios of admissions by regular HEIs in the four years studied are displayed in Figure 10.2.6.

Overall, the GS ratios of regular HEI admissions for provinces present a somewhat different picture from those using Admission Rates. According to the figure below, students from Beijing, Tianjin and Shanghai do not show the highest over-representation in HE admission anymore. The GS ratios of these widely acknowledged privileged areas are a little higher than zero.

Instead, students from Jiangsu, Zhejiang and Fujian, three provinces on the east coast, are the most over-represented in regular HEIs admissions. Apart from these provinces, the advantaged group also includes Liaoning, Heilongjiang, Shandong, Inner Mongolia, Chongqing and Shaanxi. However, the over-representations of these provinces are very small.

According to Figure 10.2.6, the disadvantaged group of provinces shows few changes compared with those discussed in the last few sections. Henan is much more under-represented in all four years than Guangdong, Sichuan, Jiangxi and Hunan, which are under-represented to a lesser extent. Other provinces including Gansu, Yunnan, Hebei and Xinjiang are also negatively segregated. The separate yearly results are listed in Appendix 8.

Figure 10.2.6 GS Ratio of admissions by all regular HEIs for provinces from 2016 to 2019



Note: the provinces are listed in the order of the 2019 hierarchy

To sum up, the analyses employing GS ratios of admission by all regular HEIs for provinces do not thoroughly overthrow the conclusions made in the last few sections. Despite the different levels of over-representation, the members of advantaged groups have mostly remained the same. These are, unsurprisingly, usually economically advantaged provinces. Furthermore, this consistency of membership is also maintained when it comes to disadvantaged groups. Provinces in remote central areas such as Yunnan and Sichuan and those with large populations or shortages of resources, such as Henan, continue to be disadvantaged.

10.2.3 Conclusion

Overall, the results for GS ratios illustrate a very similar pattern to that illustrated by the Admission Rates Index. That is, even if the wider population is taken into account, provincial disparities in the admissions of both prestigious universities and regular HEIs exist. The advantaged provinces and the disadvantaged ones usually remain the same. Therefore, it seems an appropriate choice to regard province as a criterion for CA policies. However, it does not necessarily so. Despite provincial disparities in HE admissions revealed by two indices, provinces are problematic in vulnerable to ecological fallacy, which will be discussed in Section 10.4.

The analyses by GS Ratio still have some limitations, not only because of missing data but also because of biased samples. The calculations in the last few sections only include NCEE candidates, which, as was argued in Chapter 6, greatly risk missing more disadvantaged student groups. Even though some studies use the number of graduates from high schools rather than that of NCEE candidates as the denominator (Wang, 2010), it does not much improve the results because competition for HE in China begins at the end of compulsory education. Therefore, the next section employs the Admission Opportunities Index, which takes enrolment rates and graduation rates in primary, middle and high schools into account, to more deeply explore HE equity.

10.3 Comparisons of HE Admission Opportunities for Provinces using the Admission Opportunities Index

As mentioned previously, it is important not to overlook students who leave education in earlier stages when investigating HE equity. Otherwise, a large group of potentially disadvantaged students might be missed. Therefore, this section explores the educational trajectories of students in provinces using the Admission Opportunities Index to identify potential gaps in education participation.

Tables 10.3.1, 10.3.2 and 10.3.3 demonstrate the admission rates and graduation rates in primary schools, middle schools and high schools (NCEE applicant rates) for each province. The last several columns present the admission probabilities for attending a DFC university and attending HE, respectively, for students from each province. The provinces are ranked according to their admission opportunity indices for prestigious HE.

Table 10.3.1 The percentages of students in each education period in provinces

	primary school admission rates in 2005	primary school graduation rates in 2011	middle school admission rates in 2011	middle school graduation rates in 2014	high school admission rates in 2014	NCEE application rates in 2017	DFC HEIs admission rates in 2017	AOI for DFC HEIs in 2017	admission rates for regular HEIs in 2017	AOI for regular HEIs in 2017
Tianjin	99.97	107.91	97.64	96.13	68.89	104.20	3.32	2.42		
Beijing	99.91	143.24	98.92	89.56	61.27	109.78	2.13	1.82	66.91	57.06
Jilin	99.35	105.25	102.38	90.61	62.91	108.09	2.63	1.74		
Shanghai	100.00	125.14	88.77	79.35	57.38	94.52	2.91	1.39		
Fujian	99.79	103.68	97.69	94.41	60.80	90.22	2.01	1.05		
Shaanxi	99.20	106.58	98.71	88.81	68.70	114.01	1.37	0.99		
Heilongjiang	98.47	98.77	99.76	80.43	67.36	103.52	1.65	0.90		
Qinghai	97.21	89.81	95.39	80.65	63.53	118.11	1.76	0.89		
Liaoning	99.74	99.29	99.92	89.02	62.27	99.81	1.59	0.87		
Hubei	99.64	102.82	106.14	77.05	61.89	122.59	1.22	0.78		
Ningxia	99.04	86.43	91.04	88.89	58.00	130.81	1.36	0.72		
Inner Mongolia	99.44	98.46	100.67	88.70	68.23	127.11	0.92	0.70	48.63	36.87
Shandong	99.86	102.45	99.03	93.97	56.26	104.26	1.23	0.69	47.62	26.59
Henan	99.65	98.92	96.43	70.94	56.24	134.25	1.26	0.64	42.35	21.57
Shanxi	99.42	96.17	96.49	88.10	56.96	124.02	1.06	0.61		
Hunan	99.03	102.71	100.64	88.77	56.02	112.39	1.06	0.61	42.53	24.34
Zhejiang	99.99	105.76	96.52	92.16	54.65	115.73	1.02	0.60		
Jiangsu	99.79	104.91	99.98	94.61	52.03	103.22	1.14	0.60	61.56	32.74
Hainan	99.79	72.14	95.22	87.89	51.59	100.18	1.77	0.55		
Guangdong	99.68	98.67	95.13	89.75	50.38	108.64	1.19	0.54	38.31	17.61

Hebei	99.70	101.78	99.24	82.85	63.45	114.10	0.89	0.54	51.99	31.41
Sichuan	97.17	95.41	101.54	86.76	54.22	117.10	1.01	0.52		
Xinjiang	98.70	96.81	99.48	91.99	57.99	105.64	0.94	0.50	43.93	23.53
Anhui	99.54	94.38	98.90	87.57	55.62	134.54	0.82	0.50	43.24	26.33
Jiangxi	99.01	99.76	101.17	81.59	57.37	115.40	0.89	0.48	38.37	20.71
Gansu	98.87	92.29	95.40	86.48	59.17	136.86	0.73	0.44		
Guangxi	99.07	95.64	98.01	89.83	49.24	119.83	0.86	0.42	40.74	20.05
Tibet	96.33	88.30	92.34	93.95	43.91	154.98	0.60	0.30		
Yunnan	96.30	99.41	94.25	82.61	47.23	109.47	0.76	0.29		
Guizhou	98.33	100.10	96.47	90.38	51.78	118.02	0.55	0.29		
Average	99.04	100.18	97.78	87.16	57.46	115.42	1.29	0.72	43.55	26.06
SD	1.04	11.88	3.51	5.78	6.35	13.95	0.68	0.48	9.00	10.71

Table 10.3.2 The percentages of students in each education period in provinces

	primary school admission rates in 2006	primary school graduation rates in 2012	middle school admission rates in 2012	middle school graduation rates in 2015	high school admission rates in 2015	NCEE application rates in 2018	DFC HEIs admission rates in 2018	AOI for DFC HEIs in 2018	admission rates for regular HEIs in 2018	AOI for regular HEIs in 2018
Tianjin	99.38	105.36	96.88	100.14	64.11	102.23	4.07	2.71	74.93	49.89
Beijing	99.96	149.79	98.72	85.82	61.12	111.11	2.66	2.29	67.56	58.20
Jilin	99.22	97.79	98.27	88.90	67.12	111.17	2.87	1.82		
Shanghai	100.00	119.14	90.73	80.23	56.64	93.63	3.36	1.54		
Chongqing	99.92	92.74	101.25	95.53	61.09	125.88	1.83	1.26	46.15	31.81
Liaoning	99.77	95.55	99.49	91.29	62.20	88.18	2.61	1.24	85.41	40.55
Hubei	99.49	90.12	99.44	90.31	60.38	134.24	1.71	1.12	0.00	0.00
Heilongjiang	98.89	104.02	99.71	77.27	67.78	105.19	1.89	1.07	60.05	33.93
Fujian	99.84	94.88	97.60	95.09	59.40	93.14	2.14	1.04	62.25	30.28
Shaanxi	99.36	92.42	91.73	92.22	69.65	120.70	1.57	1.02	48.35	31.57
Ningxia	99.27	89.68	92.76	88.92	57.18	136.81	1.58	0.91	31.00	17.81
Qinghai	97.05	79.65	90.62	89.02	61.22	125.00	1.82	0.87	44.80	21.38
Inner Mongolia	99.73	93.17	99.84	91.71	66.73	131.40	1.11	0.83	48.95	36.52
Shandong	99.96	99.05	95.80	97.47	55.86	106.92	1.46	0.81	45.96	25.37
Shanxi	99.63	104.27	84.51	92.45	58.09	122.79	1.30	0.75	39.00	22.58
Jiangsu	99.86	101.19	99.26	95.59	52.20	103.76	1.33	0.69	65.46	33.99
Hunan	99.53	98.03	96.40	94.25	54.34	118.80	1.18	0.68	40.00	22.89
Gansu	98.89	82.77	94.28	88.41	60.00	136.91	1.20	0.68	37.61	21.08
Zhejiang	99.99	101.59	94.85	93.68	54.34	117.74	1.14	0.66		
Hebei	99.41	99.43	97.69	88.97	58.93	119.30	1.07	0.65	49.10	29.65
Guangdong	99.72	96.35	93.52	92.19	51.39	114.09	1.28	0.62	39.18	19.03
Henan	99.86	96.38	92.79	78.16	54.99	144.72	1.11	0.62	29.75	16.53
Hainan	99.81	73.81	93.14	90.17	53.31	102.44	1.78	0.60		

Xinjiang	99.16	94.97	97.94	94.63	63.09	110.50	0.97	0.59	41.37	25.17
Jiangxi	99.64	95.21	97.88	84.85	57.57	118.60	1.07	0.57	40.04	21.54
Anhui	99.72	84.41	95.63	92.97	56.68	137.35	0.95	0.56	43.90	25.58
Sichuan	97.09	86.96	98.66	89.18	55.59	126.53	1.05	0.55	38.63	20.19
Yunnan	96.57	95.07	93.30	85.59	47.56	109.29	1.34	0.51	37.00	14.10
Guangxi	99.25	88.05	97.42	93.84	49.46	128.87	0.91	0.46	39.50	20.11
Tibet	96.54	89.79	91.37	92.36	48.90	129.08	0.80	0.37		
Guizhou	98.62	96.48	96.28	95.97	48.90	128.09	0.58	0.32		
Average	99.20	96.26	95.73	90.95	57.19	118.23	1.60	0.91	45.55	26.27
SD	1.00	13.03	3.65	5.30	5.80	14.29	0.79	0.54	16.79	11.76

Table 10.3.3 The percentages of students in each education period in provinces

	primary school admission rates in 2007	primary school graduation rates in 2013	middle school admission rates in 2013	middle school graduation rates in 2016	high school admission rates in 2016	NCEE application rates in 2019	DFC HEIs admission rates in 2019	AOI for DFC HEIs in 2019	admission rates for regular HEIs in 2019	AOI for regular HEIs in 2019
Tianjin	99.66	99.65	99.65	99.00	63.69	103.51	4.32	2.79	79.46	51.33
Beijing	100.00	102.38	95.44	81.01	61.89	110.28	3.63	1.96	67.87	36.67
Jilin	99.60	88.78	92.58	94.32	67.72	116.63	2.81	1.72		
Shanghai	100.00	122.27	89.44	76.34	57.82	94.16	3.47	1.58	71.14	32.34
Chongqing	99.96	93.58	101.13	96.65	62.58	132.20	1.85	1.40	45.02	34.05
Shaanxi	99.38	88.16	95.81	95.28	70.35	126.71	1.84	1.31	46.23	32.96
Liaoning	99.87	92.41	98.11	96.90	61.08	115.09	2.06	1.27		
Ningxia	99.64	94.79	95.09	90.50	54.39	150.31	1.74	1.15		
Hubei	99.79	80.12	99.67	91.89	61.82	138.58	1.78	1.12		
Heilongjiang	98.64	97.03	84.55	98.85	67.52	109.50	1.88	1.11	56.54	33.45
Fujian	99.93	95.04	96.89	97.16	57.88	95.72	2.19	1.09	62.18	30.79
Qinghai	98.60	82.84	91.86	90.92	61.09	133.80	1.80	1.01		
Inner Mongolia	99.71	89.50	97.77	96.55	66.89	135.37	1.23	0.93	47.92	36.56
Hainan	99.75	98.58	93.83	92.55	52.92	104.71	1.94	0.92		
Shandong	99.93	92.68	96.46	100.06	55.95	100.38	1.65	0.83	49.85	25.02
Shanxi	99.54	89.15	86.53	99.51	59.49	128.43	1.38	0.81		
Jiangsu	99.55	98.08	96.75	99.58	51.66	106.54	1.40	0.73	59.88	31.00
Guangdong	99.78	95.25	94.86	93.33	53.03	119.38	1.36	0.73	39.45	21.01
Gansu	98.94	85.06	90.21	93.45	61.91	137.95	1.19	0.72	40.43	24.49
Zhejiang	100.00	98.52	94.82	94.16	53.66	123.60	1.17	0.68	79.28	46.25
Hunan	99.83	89.30	99.48	96.53	53.24	126.68	1.17	0.67	36.89	21.30
Hebei	99.48	94.91	93.27	98.83	55.87	129.33	1.03	0.65	42.97	27.03
Sichuan	98.86	81.88	99.08	94.81	57.14	136.55	1.09	0.65	30.89	18.33
Henan	99.94	89.77	83.72	94.04	53.69	143.82	1.13	0.62	36.06	19.67
Jiangxi	99.83	89.14	93.12	94.06	57.47	127.50	1.01	0.58	37.72	21.55
Anhui	99.81	75.78	99.65	95.21	57.75	142.98	0.96	0.57	37.43	22.18

Yunnan	97.59	94.59	95.54	88.04	48.00	111.03	1.27	0.53		
Guangxi	99.08	92.65	98.29	94.33	52.11	135.89	0.86	0.52		
Guizhou	98.57	98.41	98.99	96.64	49.50	133.89	0.61	0.37		
Average	99.49	92.77	94.92	94.16	58.21	123.12	1.72	1.00	50.91	29.79
SD	0.58	8.45	4.55	5.22	5.72	15.42	0.87	0.52	15.09	8.96

First, there are few differences in the admission rates of primary schools among provinces. All provinces have high admission rates for primary school education over the course of three years.

However, the provincial gaps become more significant in the completion rates of primary schools. Generally, well-developed provinces such as Shanghai, Tianjin and Beijing still show high rates, which even exceed 100%. These over-100% completion rates indicate that the number of graduates from primary schools in these provinces is greater than the number of students initially enrolled in primary schools. On the other hand, under-developed provinces including Qinghai, Ningxia and Hainan, and populous provinces such as Hubei and Anhui, have much lower completion rates for primary schools. That may due to internal migration, as some migrant workers may take their children with them to metropolitan rather than leaving them in their hometown with fewer resources or larger populations when the children reach the age of compulsory education. However, their children tend to study in segregated schools for migrant workers' children and have to return to their hometown provinces for high school education and NCEE, as discussed in Chapter 2.

Furthermore, there are also some disparities in the admission rates and completion rates of middle school education, but the advantaged and disadvantaged group do not differ significantly compared to the previous ones. One noteworthiness is that Shanghai presents very low admission and completion rates for middle school. This may be attributed to an earlier selection by the education system or, another speculation, that local students in Shanghai are more likely to come from affluent families and tend to study abroad at a young age.

The admission rates for high school education experience a cliff-like decline for all provinces, with almost half of students leaving, or having to leave, education at this stage. Nevertheless, some provinces still show privileges. For instance, Tianjin, Shaanxi, Jilin and Heilongjiang, which also represent for the advantaged group in the previous two indices in Sections 10.1 and 10.2, have higher admission rates of high school education over the three-year study period. In contrast, Yunnan, Guangxi, Tibet and Guizhou present lower rates, which means that students from these provinces are less likely to transit to high schools, let alone HE.

The application rates for NCEE in each province appear high. In some provinces, the corresponding rates exceed 100%, where may be a group of students who repeated the last year of high school education and took the NCEE for a second (third, or several) time. The following sections will discuss the Admission Opportunities Indices in elite universities and in regular HE.

10.3.1 Provincial Admission Opportunities in DFC universities

According to Tables 10.3.1, 10.3.2 and 10.3.3, each year, the advantaged provinces with the Admission Opportunity Indices above average and the disadvantaged provinces with the indices consistently below average remain almost the same.

This clustering of provinces echoes the findings in the last two sections. Beijing, Tianjin, Jilin and Shanghai, for instance, are still strongly advantaged in admissions to DFC universities even when earlier educational experiences are taken into account. On the other hand, not surprisingly, Hebei, Anhui, Jiangxi, Hubei, Guangdong, Sichuan, Guizhou, Yunnan, Gansu, Guangxi, Tibet and Xinjiang still belong to the disadvantaged group.

The gap between the most advantaged province and the least advantaged one when using the Admission Opportunities Index is even more significant than when using the Admission Rate Index and the GS Ratio. For example, the Admission Opportunity Index of DFC universities for Tianjin (2.42%, 2.71% and 2.79% in 2017, 2018 and 2019 respectively) is nearly eight times higher than that for Guizhou (0.29%, 0.32% and 0.37%). The corresponding figure is about six to seven when using Admission Rates. This gap somewhat indicates that disparities in HE admissions do not start in the transition to HE but from earlier stages of education, and if early differences are neglected, the results might be biased and over- or under-estimated.

10.3.2 Provincial Admission Opportunities in regular HEIs

The results of using the Admission Opportunities Index to investigate admission opportunities in regular HEIs for each province in 2017, 2018 and 2019 are listed in the last column of the three tables. The traditionally advantaged provinces such as Beijing and Tianjin still show more advantage than other provinces. Furthermore, Jiangsu, Zhejiang, Inner Mongolia, Heilongjiang, Shaanxi, Chongqing and Fujian have higher admission opportunities than the average and become winners in the competition for HE admissions when taking into account primary school intakes. Most of these provinces have an abundance of regular HEIs (see Chapter 2).

Some provinces, however, always have below-average Admission Opportunity Indices, such as Guangdong, Hunan, Henan, Guangxi, Jiangxi, Sichuan, Anhui, Ningxia, Gansu, Shanxi and Yunnan. These gaps mean that students, even if the earlier education periods are taken into account, in these provinces face more competition and find it more challenging to enter HE.

10.3.3 Conclusion

To sum up, the analysis using the Admission Opportunities Index generally echoes the findings when using the Admission Rate Index and the Gorard Segregation Index. Beijing and Tianjin are usually in the most advantaged positions and are followed by Fujian, Jilin and

Heilongjiang in admissions to DFC universities, and by Jiangsu and Zhejiang in admissions to regular HEIs. On the other hand, some provinces such as Henan, Hunan, Guangdong, Sichuan and Guangxi are usually disadvantaged. These provinces either have large populations of students or have fewer economic and educational resources.

Using the Admission Opportunity Index even more significant gaps in HE admission opportunities between provinces are found, which might mean that differences in earlier stages of education can potentially affect equity in HE. Some provinces such as Tianjin and Jilin show privileges in admissions and completions almost in all stages of education, whereas some provinces such as Yunnan, Guizhou and Tibet are usually in the disadvantaged position wherever the educational period is. Furthermore, the huge decline in the transition to high school education may warn us that nearly half of students, who might be more disadvantaged, are ignored by studies which only focus on NCEE candidates and current CA policies. To explore HE equity it is necessary to get a broader view.

However, there are still some problems with the use of this index. First, there are again some missing data. For instance, data could not be found for Jilin, Hubei, Hainan, Guizhou and Tibet for any of the three years but they might be important. In the sections using the other two indices Jilin and Hainan were substantially over-represented, while Guizhou and Tibet were potentially disadvantaged. It would be helpful to see whether these situations were permanent or temporary.

There is another problem. As was discussed in Chapter 8, calculation of the Admission Opportunities Index takes enrolment and graduation in all stages of education into account but it is not able to reveal the real admission opportunity. The reason is that this index assumes that there is no mobility during education, which is highly unlikely. Although some drop-outs, school changes and stops in education for other reasons can be dealt with as they are reflected in graduation rates, the mobility of students at the end of compulsory education could lead to bias. For instance, the migrant students might go to metropolises with their working parents.

They study in specialised schools for migrant children in cities until they graduate from middle school. However, according to hukou policy (Chapter 2) once they leave middle school they have to go back to their home provinces to attend high school if they want to take the NCEE. Therefore, graduations from middle schools in metropolises such as Beijing, Shanghai and Guangzhou include this migrant group, but enrolments in high schools do not. This could lead to underestimation of the enrolment rates of local students in high schools in these well-developed areas and overestimation of these rates in the home provinces of these migrant students. These incorrect estimations might negatively affect calculation of the Admission Opportunities Index.

Despite clear disparities between provinces in HE admissions, province is not necessarily appropriate to be a CA indicator. More detailed evaluation is discussed in the next section.

10.4 Discussion and Conclusion

In conclusion, based on the results of using three different indices it seems reasonable to answer the research question that there are disparities in HE admissions between provinces. The provinces advantaged in admissions by both elite and regular HEIs are generally the same. Similarly, the disadvantaged ones are also consistent. Although there are some missing data which might corrupt these findings, the fact that provinces always stay in the same groups in the period studied indicates that they are quite robust. The results almost overlap with findings in previous studies (which were reported in Chapter 9), which might indicate that these provincial disparities have existed for a long time. There is a need for a re-allocation of places among provinces in both prestigious and regular universities.

Two provinces deserve more discussion: Jiangsu and Zhejiang. All three indices show that both provinces are disadvantaged in DFC university admissions but noticeably advantaged in admissions by regular HEIs. There might be two reasons for this inconsistency. The first reason might be that compulsory education and local HE are well developed. The local governments in Jiangsu and Zhejiang are rich as these two provinces are the most developed in China and they have invested many resources in not only HE but also basic education. Therefore, education in Jiangsu and Zhejiang is well developed. Fewer students might drop out of compulsory education and more places might be provided for local students in HE. Nevertheless, 'HE' here refers more to local universities which are affiliated with the local government. This assumption is partly supported by larger numbers of HEIs in these two provinces to a certain extent (See Table 2.1.2 in Chapter 2). Therefore, when comparing the admissions of regular HEIs, these two provinces usually show an advantage.

Second, however, the central government or its departments might restrict investment and support for these rich provinces as they might want to use the funding to help other provinces in need. Fewer investments mean weaker control by central government, which may seem to mean freedom, but elite universities in China are almost all controlled by the central government and its departments. Therefore, although the local universities might be rich in several kinds of resources, they can hardly be labelled top universities, such as DFC universities. Jiangsu and Zhejiang therefore only have a few DFC universities and low admission quotas in DFC universities. This might be why they seem to be disadvantaged in admissions to prestigious universities.

Although it would seem reasonable to conclude that there are inequalities in HE admissions between provinces, this does not necessarily mean that 'province' is suitable to be used as an indicator of disadvantage in CA policies. This is because no matter what index is used the analysis still focuses on macro province-level indicators. These indicators are somewhat similar to area-based CA indicators such as POLAR used in the UK, but they are even worse as a province in China is much larger and more complicated than an area indicated by POLAR. These macro-level indicators can easily result in 'ecological fallacies' and mistakenly target advantaged groups or exclude those in need (Boliver et al., 2022).

It might be acceptable to use 'province' as one of the indicators during the admission process as there are currently inequalities between provinces not only in education but also in many other factors such as economics, social welfare and medicine. These imbalanced developments between provinces might make it challenging to suddenly abandon province-based quota admission policies. On the other hand, some provinces always remain in disadvantaged positions. They lag behind in HE enrolment competition starting with early education and deserve more places in HE. They deserve a more equity-oriented province-based quota admission policy to provide larger quotas for them in HE.

However, it is dangerous to use provinces separately to identify the disadvantaged. Students from the same province are unlikely to be at the same level of disadvantage (Chen & Gao, 2019). If province is used alone as a CA indicator, students from wealthy families in a poor province could be admitted to HEIs, even top ones, with a low entry requirement. It might be a good idea to regard province as the basic unit of admission, and then use other micro-level indicators to find the truly disadvantaged students in each province.

It is necessary to find more accurate and reliable individual-level indicators. Therefore, the individual data in some large-scale comprehensive social surveys, including the CGSS and the CFPS were analysed to help understand the patterns of HE participation by different groups and evaluate the quality of potential indicators. The results of the analysis are presented in the following seven chapters.

Chapter 11 Exploring the Clustering of HE Participation by Different CGSS Indicators

Chapter 10 looked at the 'hukou province' indicator and showed what provincial equality and equity in HE enrolments look like. The next four chapters examine two other indicators used in three Special Plans and extra credit policies: students' hukou status and ethnicity. These chapters are based on analyses of CGSS data to investigate: 1) whether there are disparities in HE participation between hukou status and ethnicity groups; and 2) the extent to which these officially recognised indicators accurately identify genuine disadvantages. In addition, the four chapters also examine how other popularly discussed indicators reveal gaps in HE admissions and discuss the extent to which they could be useful CA indicators.

This chapter primarily examines the clustering of HE participation by both currently employed and non-officially-recognised indicators in the CGSS individual-based dataset. Chapters 12 and 13 explore the relationships between potential indicators of disadvantage, and predict HE admissions of individuals if these indicators are employed, evaluating their appropriateness as CA indicators. Chapter 14 examines the educational attainment of CGSS cases in more detail in order to examine the possible accumulative disadvantages in earlier stages of education.

11.1 Overall HE Participation of CGSS Cases

As was explained in Chapter 7, the CGSS is a nationally representative cross-sectional social survey with a reasonable sample size. After the selection that was made in Chapter 7 a new sub-dataset of 10,089 cases was obtained. Within these remaining 10,089 cases, there are 2,097 HE participants (21%), 7,986 non-HE participants (79%), and six cases missing this information (Table 11.1.1).

The first group of variables considered were ones officially employed as indicators in

Chinese CA policies: hukou status and ethnicity. The second and third groups of variables were the individual-level and family-level indicators shown in Table 11.1.2

Table 11.1.1 Distribution of cases by HE participation

	Percentage	Total (N)
HE	21	2,097
Non-HE	79	7,986
Missing		6
Total	100	10,089

Table 11.1.2 The examined indicators based on CGSS

Currently used	Hukou status, ethnicity
Individual-level	sex, birth month, language ability
Family-level	Social class, parental education qualifications, parental workplaces

It should be remembered that all the information provided in the survey is self-reported. Besides the difficulty in verifying such information, the other limitations of the indicators which were identified in Chapter 7 were vague definitions, an inaccurate measure of SES and a lack of information on disabilities and special educational needs. Despite these limitations, with its large sample size CGSS data can still contribute to identifying HE participation disparities and finding possibly better CA indicators.

11.2 Clustering of HE Participation by Currently-employed Indicators

This section will examine the two officially used CA indicators to see how HE participation is patterned by these characteristics.

11.2.1 Hukou Status

A widely used comparison is the percentage distributions of rural hukou holders and urban hukou holders, as Table 11.2.1 shows. It seems that rural hukou students are more disadvantaged in HE participation than their urban counterparts because of the 2,097 HE participants only 32% have rural hukou and 67% have urban hukou.

Table 11.2.1 Percentages of HE participants with different hukou status

	HE
Urban hukou	67
Rural hukou	32
Missing hukou	1

However, conclusions cannot be drawn based on just these figures (Guo et al., 2010; Wang & Liu, 2011; Wang, 2010). This simple comparison ignores a more complete picture also covering non-HE participants, who are very important when exploring HE equity (Gorard, Siddiqui & Boliver, 2017). For example, the smaller percentage of rural hukou holders among HE participants might not necessarily mean anything because it might just be a result of there being a smaller percentage of rural hukou holders in the whole population. Therefore, more information is needed. To draw a reasonable conclusion, at least the data in Table 11.2.2 are required. Once we know that among 2,097 HE participants 67% are urban hukou students and 32% are rural hukou students and that among 7,986 non-HE participants 71% are rural hukou students and 29% are urban hukou students we can conclude that rural hukou students are more disadvantaged.

Table 11.2.2 Percentages of HE participants and non-HE participants with different hukou status

	HE	non-HE
Urban hukou	67	29
Rural hukou	32	71
Missing hukou	1	

There is a better way of comparing the clustering of HE participation (see Table 11.2.3). There are 6,331 rural hukou residents and 3,734 urban hukou residents in the final sample accounting for 63% and 37%, respectively. As Table 11.2.3 shows, 11% of the rural hukou students could access HE, whereas 38% of the urban hukou students entered HE. Within the urban hukou resident group around 62% did not go into HE while this figure is 89% for rural hukouers. Rural hukou residents are less likely to go into HE than urban ones.

Regarding the odds ratio, students with urban hukou are 5.06 times as likely to enter HE than their counterparts with rural hukou. This gap in HE access characterises rural hukou residents as disadvantaged.

Table 11.2.3 Percentages of HE participation by students' hukou status

	Total (N)	HE (%)	non-HE (%)	Odds ratio
Urban hukou	3,734	38	62	5.06
Rural hukou	6,331	11	89	
Missing hukou	24	-	-	

A further 24 cases in the sample lack hukou information. This is a small number which will not have much effect on the percentages. Therefore, they are not discussed in detail.

The percentages in Table 11.2.3 are easier to understand and compare. The difference in HE participation between the groups is clear. Therefore, this form of comparison will be employed in subsequent analyses in this thesis.

11.2.2 Minority Ethnicity

Ethnicity is another contextual indicator used in current CA policies. Students with minority ethnicity status are eligible for extra help in their transition to HE. In the CGSS data there are 9,176 cases with Han majority ethnicity and 902 cases with minority ethnicities, accounting for 91% and 9% respectively. The proportion of minority ethnicity cases is similar to that in the seventh population census (9%). Therefore, the distribution of ethnic groups in the CGSS sample is quite representative.

As Table 11.2.4 shows, 21% of the individuals in the Han ethnic group attended universities, and the corresponding figure for minority ethnicity groups is only 15%. The odds ratio shows that Han students are 1.5 times as likely to participate in HE than minority ethnicity students. It might be fair to claim there is a gap in HE participation between Han majority ethnicity and other minority ethnicities. However, this gap seems to be narrower than the one between Hukou status.

Table 11.2.4 Percentages of HE participation by students' ethnicity

	Total (N)	HE (%)	non-HE (%)	Odds ratio
Han ethnicity	9,176	21	79	1.51
Minority ethnicity	902	15	85	
Missing ethnicity	11	-	-	

There are 11 missing cases, which is a small number compared to the size of the sample so they are ignored in the comparison of percentages of HE participation.

11.3 Clustering of HE Participation by Other Potential Indicators

This section examines some popularly discussed indicators: sex, social class, parents' education, parents' workplaces, language ability and birth month. Comparisons are made to investigate whether there are HE admission disparities according to these characteristics.

11.3.1 Sex

Besides the officially employed indicators, there are other indicators that are not used in CA policies in China but have been widely claimed to identify disadvantages. sex is one of these.

As Table 11.3.1 shows, in the CGSS data 20% of females and a slightly higher percentage, 21%, of male individuals went into HE. Females are a little under-represented. According to the odds ratio of 1.09 between sex groups, HE participation by the sex groups seems to be relatively equitable.

Table 11.3.1 Percentages of HE participation by students' sex

	Total (N)	HE (%)	non-HE (%)	Odds ratio
Female	5,275	20	80	1.09
Male	4,814	21	79	

It should again be stressed that sex is recorded as a binary category in official data and social surveys in China. This produces limited information and does not do justice to groups such as transsex people, who are likely to be disadvantaged and under-represented in official statistics.

11.3.2 Social Class

Socioeconomic status (SES) is another important issue. In general, although its meaning is controversial, SES should at least involve household income, educational attainment and occupation. However, the CGSS dataset lacks information on individuals' household income at the time they went to university. Therefore, this section uses information recorded in response to the question "Which social class do you think your family was in compared to your contemporaries when you were 14 years old?" The answers range from 1, the lowest social class, to 10, the highest. In this analysis, the first three social classes are combined into a new group named 'Working class,' classes 4, 5, 6 and 7 are grouped together and named 'Middle class' and the last three make an 'Upper class' group. These three social class groups contain 4,798, 4,934 and 277 individuals and account for 48%, 49% and 3% of the sample respectively.

As Table 11.3.2 shows, working class, middle class and upper class HE entrants respectively constitute 15%, 26% and 21% of the total CGSS sample. As the average proportion of the sample that participated in HE is 21%, it seems that only working class individuals are under-represented in HE. Students from the middle class or upper class are over- or at least reasonably represented. The middle-class group is even more advantaged than the upper-class one. This unique advantage might be caused by errors in misunderstanding, recording or editing in the data collection process, as 'social class' is an ambiguous and complicated variable. However, even if there was no such error in the data collection, this unexpected gap is still possible. There are two problems in the question on social class: "Which social class do you think your family was in compared to your contemporaries when you were 14 years old?" The first problem could be students having inaccurate memories of their lives when they were "14 years old." More importantly, this question will not elicit an objective assessment of one's social class because of the confusing use of the phrase "compared to your contemporaries". In addition, the cases reporting in upper class are much fewer than those in other two class groups. This small sample size might also be a reason for the more advantage of middle class.

The odds ratio here is compared by advantaged group, in other words, the non-working-class group, which includes both the middle class and upper class, and the less-advantaged group— the working-class group. The former is 1.96 times as likely to be admitted to HE than the latter.

Table 11.3.2 Percentages of HE participation by students' social class

	Total (N)	HE (%)	non-HE (%)	Odds ratio
Working class	4,798	15	85	
Middle class	4,934	26	74	1.96
Upper class	277	21	79	
Missing Social Class data	80	16	84	

What deserves attention is missing information on social class because 80 respondents did not respond to the question on social class. The group lacking this information were under-represented in HE admissions. Therefore, information on the size of the disadvantaged group could be biased or even missed if the group that lacks social class information is neglected. Nevertheless, 80 is still a relatively small number. There is more discussion about missing data in Chapter 12.

11.3.3 Parental Education

Parental education is a significant element in SES, which it is related to children's educational outcomes and attainments. ‘The first generation in HE’ has even been used as an indicator of disadvantage in admission by some prestigious universities, including Princeton University in the US (Princeton University, 2022a). In the 2012, 2013, 2015 and 2017 CGSS datasets there is information on parents' highest education levels. This information is recorded as reported by parents. In this study parental education is re-categorised in three groups:

1) at least one of the parents obtained a bachelor's degree or above;
2) at least one of the parents obtained a post-compulsory education qualification, including from a vocational middle school (zhongzhuan in Chinese), a technical school (jixiao in Chinese), a high school and a vocational high school (zhigaoin Chinese); and
3) both parents only obtained a compulsory education qualification, including from middle school, primary school and ‘sishu’ (an old-style private school) or below.

Rather than the more commonly used binary classification, this trichotomous one is employed because the births of the parents of the cases in the final dataset were generally in the 1950s. Theoretically, they would have gone into HE in the 1970s. At this time all levels of education and especially HE were not advanced and to a large extent were not accessible. It was not common for many people born in the 1950s to go into HE. There could be a risk of bias if the binary categories of ‘Parental HE or Not’ were used. Therefore, a more specific classification was chosen.

After this re-categorisation, 829 interviewees had a father or mother with a bachelor’s degree, 5,917 had a parent who finished post-compulsory education and 3,343 had parents who were only basically educated or not educated at all. These groups respectively account for 8%, 59% and 33% of the sample.

Table 11.3.3 shows that the percentage of students with well-educated parents who were admitted into HE (62%) is remarkably higher than those of the other two groups. 35% of those with parents with post-compulsory education qualifications and 12% of those whose parents only finished middle or primary school education (or equivalent) or did not experience any formal education at all entered HE. Students whose parents are only basically educated are remarkably disadvantaged in HE admissions.

Table 11.3.3 Percentages of HE participation by students' parental education

	Total (N)	HE (%)	non-HE (%)	Odds ratio
Compulsory education or below	6,596	12	88	
Post-compulsory education	3,062	35	65	4.42
Higher education	337	62	38	
Missing parental education data	94	16	81	

The odds ratio result echoes this disadvantage. The advantaged group in HE participation that includes individuals with parents who gained post-compulsory education qualifications and those with parents who gained a HE degree are 4.42 times as likely to go into HE than their counterparts whose parents only completed compulsory education or below. The difference is expected to be greater if the comparison groups are "at least one parent complete higher education" and "both parents only complete compulsory education or below".

The last row of 94 cases without parental education information shows under-representation, with 16% of them being HE participants (around 3% missing HE participation information), compared to the average of 21%. Again, this group with missing parental education data should not be directly deleted but require consideration.

11.3.4 Parental Workplaces

Another important indicator is parental occupational status and category, which can be partly seen as an indicator of family income. However, as was explained in Chapter 7, the CGSS lacks direct information on parents' occupations. This section, therefore, treats workplace as a proxy for occupation and explores the different patterns of HE participation for various workplaces.

As Table 11.3.4 shows, CGSS interviewees whose father worked in a "social community/committee" or "other" place or was "self-employed" are disadvantaged in HE participation. Only 14%, 13% and 13% respectively of the people in these groups were admitted to HE. Some other groups are over-represented in HE and the most advantaged are children of party or government members. These over-represented groups are collapsed together to one advantaged group when calculated the odds ratio as introduced before, and so are the under-represented three groups, which obtained a disadvantaged group for comparing. The advantaged group is 3.78 times as likely to attend HE than the disadvantaged one.

Table 11.3.4 Percentages of HE participation by students' fathers' workplaces

	Total (N)	HE (%)	non-HE (%)	Odds ratio
Party/government agency	276	46	54	3.78
Company	1,929	34	66	
Institution	719	40	61	
Army	23	35	65	
Social community/committee	97	14	86	
Self-employed	5,930	13	87	
Other	261	13	86	
Father's workplace data missing	854	24	76	

Table 11.3.5 provides percentages of HE participation by students' mother's workplaces. It shows a pattern similar to that in Table 11.3.4. Individuals whose mothers were "self-employed" or worked in an "other" place are less likely to enter HE, with only 12% and 14% of HE participation respectively. Nevertheless, those with a mother working in a "social community/committee" are not in a disadvantaged position, with 29% HE participation, but are among the privileged along with people whose mother worked in a "party/government agency," a "company" or an "institution." As aforementioned, the odds ratio is calculated by collapseing groups. The privileged group, which includes all those who are over-represented in HE participation such as students whose mother works in party/government agency, company, institutions, army and social committee, is 4.77 times as likely to enter HE than the disadvantaged group. The disadvantaged group includes students whose mother works in "other" places or works as "self-employed".

Two more things need to be discussed. The first is the last row of missing data in each table. Family SES-related information is often sensitive or not known so there is a large amount of missing data. People not reporting their father's or mother's workplace have a small advantage of 3% or 4% more HE participation than the average. People who lack this information might not necessarily be disadvantaged. The same pattern has been found in other studies and contexts (Gorard et al., 2007).

Table 11.3.5 Percentages of HE participation by students' mothers' workplaces

	Total (N)	HE (%)	non-HE (%)	Odds ratio
Party/government agency	104	55	45	4.77
Company	1,346	36	64	
Institution	531	50	50	
Army	1	100	0	
Social community/committee	66	29	71	
Self-employed	5,759	12	88	
Other	238	14	85	
Mother's workplace data missing	2,044	25	75	

The other thing is that that the mother's workplace is more important to, or more closely associated with, children's HE participation than the father's. As the two tables show, people whose mothers worked in a "party/government agency" (55%), a "company" (36%) or an "institution" (50%) have higher HE participation percentages than those whose fathers worked in the same places (46%, 34% and 39%). Furthermore, the odds ratio when comparing mothers' workplaces is higher than that when comparing fathers' workplaces (4.77 vs 3.78). Although more analysis is required, if these gaps are trustworthy, some previous findings that the father's occupation is more important than the mother's in children's education might be challenged (Pan and Wu, 2020; Wei, 2013b; Wu, 2019; Zhang, 2015).

11.3.5 Language Ability

Language does not seem to play an important role, and is not a popular topic, in HE equity in China. Nevertheless, in other countries the effect of language, such as English as an additional language (EAL), has been investigated (Gorard, Siddiqui & Boliver, 2017; Hurst, 2015; Strand, Malmberg & Hall, 2015). Although it is unusual for students whose first language is Japanese, English or French to take the NCEE test and apply to HEIs in China, language ability can still be a problem in HE transitions. As was explained in Chapter 2, some minority ethnic groups, including the Zangs and the Uygurs, have languages which are very different to Mandarin. For these groups, Mandarin is like an additional or second language, as it is for international students. Potential difficulties in using Mandarin could be a barrier to HE entry as Mandarin is the official language in most education and tests in China. Therefore, it is necessary to probe further into this possible indicator.

As Table 11.3.6 shows, only individuals who self-evaluate as 'very good' at understanding Mandarin are over-represented in HE participation, at 30%, while the other four groups in the table are all under-represented. The advantaged group is as 2.78 times as likely to go into HE than the disadvantaged group, which includes four groups showing under-representations in HE participation.

Table 11.3.6 Percentages of HE participation by students' ability to understand Mandarin

	Total (N)	HE (%)	non-HE (%)	Odds ratio
cannot understand at all	49	2	98	2.78
bad	193	5	94	
not bad	1,884	6	94	
good	3,594	18	82	
very good	4,365	30	70	
Ability to understand Mandarin missing	4	-	-	

Table 11.3.7 Percentages of HE participation by students' ability to speak Mandarin

	Total (N)	HE (%)	Non-HE (%)	Odds ratio
cannot speak at all	122	2	98	4.08
bad	613	4	96	
not bad	3,184	10	90	
good	3,283	22	78	

very good	2,883	36	64	
Ability to speak missing	4	-	-	

Table 11.3.7 shows that individuals who consider their ability to speak Mandarin to be ‘very good’ are remarkably advantaged in HE participation, at 36%, and those claiming that they are ‘good’ at speaking Mandarin are slightly advantaged, at 22%. Those who evaluate their spoken Mandarin as ‘not bad,’ ‘bad’ or ‘cannot speak at all’ are disadvantaged in HE admissions, especially the last two groups. The advantaged group, including students who are good at or very good at speaking Mandarin, is 4.08 times as likely to enter HE than the disadvantaged one, collapsed by the other three groups.

11.3.6 Birth Month

The birth month, like language ability, might also be a possible indicator of disadvantage, although it has not received much attention in the Chinese context. According to Gorard, Siddiqui and Boliver (2017) and Gorard (2021), however, children born in the summer, which are the younger and less mature group in a traditional class, might be less advantaged in continuing their education after compulsory education and so in HE participation. Therefore, in this section I explore whether there is a disparity in HE participation between summer-born pupils and autumn-born pupils in China.

The cut-off date for attending primary school education in China is 31 August, which means children born before this date can legally start primary school in September when they are aged six, while those born after this day have to wait until the following year. Here pupils born between 1 March and 31 August are labelled summer-born children and those born between 1 September and 28 (29) February autumn-born children. The former group in a class are always younger than the latter.

Only the CGSS surveys in 2012, 2013 and 2015 collected information on birth months, so a new merged file with these variables was created. The same selection criteria were used.

After deleting unselected cases, 6,964 individuals were left, 1,320 of which entered HE and 5,638 did not, accounting for 19% and 81% of the sample respectively. Six cases lack HE information.

There were 3,381 younger summer-born cases and 3,435 more mature autumn-born cases in the new data file (see Table 11.3.8). Table 11.3.8 shows that there is no particular gap between the two groups in terms of HE participation.

Table 11.3.8 Percentages of HE participation by students' birth months

	Total (N)	Missing HE (%)	HE (%)	non-HE (%)	Odds ratio
summer-born children	3,381	-	19	81	1
autumn-born children	3,435	-	19	81	
missing birth month	148	2	19	79	

148 cases lack birth month information and they have nearly the same pattern of HE attendance as the non-missing groups. Students' birth months do not seem to be associated with disadvantage.

11.4 Discussion and Conclusion

This chapter has compared percentages of HE participation according to several indicators and given the odds ratio between the disadvantaged and the advantaged group. According to the analyses, there is a remarkable disparity in HE participation between different hukou status. Individuals with rural hukou are disadvantaged with under-representation in HE admissions, which confirms the findings in previous studies (Li & Min, 2001; Sun & Barrientos, 2009; Zhang et al., 2015). The other currently employed indicator of minority ethnicity shows a smaller disadvantage in HE participation.

As for other potential indicators, individuals with self-employed parents or parents working in "other" places, with parents who only received compulsory education or below and who see themselves as working class are also under-represented in HE participation. They are more disadvantaged than their counterparts with parents who have prestigious workplaces or better-educated parents or are from higher social classes. The gaps between these groups are even larger than between Han and minority ethnicities. In addition, Mandarin language ability also indicates some HE participation gaps, as individuals with good (at speaking) or very good (at speaking and understanding) abilities in Mandarin are considerably more advantaged than those who are bad at Mandarin or cannot understand/speak it at all.

However, there are no prominent disparities between sexs or by birth month. Not only males and females but also summer-born and autumn-born children have reasonably equitable participation in HE.

The next chapter looks further into missing data and the relationships between all the indicators in order to investigate whether there are double disadvantages in these indicators and the extent to which these indicators risk being “false positives” and “false negatives” (Boliver, Gorard & Siddiqui, 2019).

Chapter 12 Relationships Between CGSS Indicators

This chapter looks at missing data for each indicator and investigates whether there are some underlying disadvantages. It also examines the relationships between potential indicators of disadvantage as some of them might be linked. Rural hukou students, for example, might be more likely to come from working-class families with less-educated parents. Such double or even triple disadvantages might help identify appropriate CA indicators.

There are at least two ways to discover relationships between indicators. The first is to explore the proportions of other potential indicators of disadvantage in each characteristic to be examined. For example, as Table 12.1.1 shows, the proportion of working-class people among rural hukou holders is 55%, while that among urban hukou holders is 35%. The other is to see how a characteristic is distributed in other indicator groups. In the group of people stating that they are working class, 73% are rural hukou holders and 27% are urban hukou holders (see Appendix 9). There are no huge differences between the results from the two types of comparison. In the following analyses the first type is mainly used.

12.1 Relationships Between Indicators that are Currently Employed and Others

This section explores relationships between the CA indicators that are currently used in China and other potential indicators of disadvantage to see whether there are any double or triple disadvantages.

12.1.1 Hukou Status

As was discussed in Chapter 11, rural hukou students are disproportionately under-represented in HE participation. To a certain extent this under-representation might be explained by Table 12.1.1. As the table shows, higher percentages of rural hukou residents state that they are working class (55%), that they come from a family with less-educated (80%) or self-employed (75% for father; 74% for mother) parents and that they cannot speak (63%) or understand (47%) Mandarin very well.

Table 12.1.1 Percentages of rural and urban hukou individuals with potential disadvantages

	Rural hukou	Urban hukou	Missing hukou
Minority ethnicity	10	7	13
Working class	55	35	42
Parents' education: compulsory education or below	80	41	54
Self-employed (father)	75	31	46
Other workplace (father)	3	1	0
Self-employed (mother)	74	28	25
Other workplace (mother)	3	1	8
Not good at understanding of Mandarin	47	25	33
Not very good at speaking Mandarin	63	46	50
Total (N)	6,331	3,734	24

These relatively close links between rural hukou and other indicators of disadvantage might indicate that rural hukou could be used as a proxy for individual disadvantage. However, what should be noted is that only using rural hukou as the indicator of disadvantage in HE admissions unfairly ignores 35% of urban hukou students who are working class, 31% and 28% of urban hukou students who have self-employed parents, and 41% of urban hukou students whose parents only received basic education or below. All of these urban hukou students are possibly disadvantaged in HE admissions. On the other hand, only using rural hukou might also create 'false positive' results for rural hukou holders who are upper class, have well-educated parents and have parents with good jobs.

As Hukou status is essential administrative information, only 24 of the 10,089 respondents lack this information. These individuals with missing hukou status do not show an obvious disadvantage. Therefore, missing hukou status might not be suitable as an indicator of disadvantage, but the sample size is too small to obtain a reliable result.

12.1.2 Minority Ethnicity

Minority ethnicity, which is another important indicator, does not reveal a remarkable disparity in HE participation, as hukou status did in the last chapter. Nevertheless, it can be seen in Table 12.1.2 that minority ethnicity groups tend to be more closely linked to other possible indicators of disadvantage than the Han majority ethnicity group. For example, minority ethnicity groups contain a higher proportion of individuals with rural hukou (72%; Han 62%), people who state they are working class (53%; Han 47%), people whose parents completed at most middle school education (70%; Han 65%) or have no formal employer (father: 66%; Han 58%, mother: 68%; Han 56%) and people with difficulty in speaking (50%; Han 38%) and understanding (66%; Han 56%) Mandarin.

Table 12.1.2 Percentages of Han and minority ethnicity individuals with potential disadvantages

	Han ethnicity	Minority ethnicity	Missing ethnicity
Rural hukou	62	72	60
Working class	47	53	60
Parents' education: compulsory education or below	65	70	60
Self-employed (father)	58	66	60
Other workplace (father)	3	3	0
Self-employed (mother)	56	68	60
Other workplace (mother)	2	2	0
Not very good at speaking Mandarin	38	50	20
Not good at understanding of Mandarin	56	66	60
Total (N)	9,176	902	11

Despite these links with disadvantage, minority ethnicity does not reveal an obvious disadvantage in HE participation. Therefore, there might be some more complicated reason for this. For instance, as was explained in Chapter 2, there are 55 different minority ethnicities in China and it is hard to imagine that they are all disadvantaged to the same extent. Instead, they are likely to face different barriers to HE entry. The first barrier could be a difference between their mother tongue and the official language. Individuals who use ethnic languages in their daily communication might be more disadvantaged, as understanding, speaking, reading or writing Mandarin could be more difficult for them.

Besides language, another barrier could be their home provinces. The provinces in China have developed very unevenly, and the wealthy eastern provinces are not only richer in economic, educational and cultural resources than the others but are also allocated higher quotas for HE admissions. Minority ethnicity students who were born and who live in these provinces might not be as disadvantaged as their counterparts who were born and who live in remote western provinces. Alternatively, they might be more advantaged than Han ethnicity students living in poor provinces.

Furthermore, the gaps between ethnic groups in SES-relevant indicators are not as significant as those between hukou status, which may indicate that ones' ethnicity is not neccessarily associated with their SES. Therefore, the sole use of minority ethnicity as an indicator could be dangerous.

In the CGSS data there is a small amount of missing data on ethnicity. Only five individuals lack ethnicity information, which is too small a number to draw any particular conclusions.

12.2 Relationships Between Potential Indicators and Others

This section explores relationships between other potential indicators of disadvantage that were linked to HE participation disparities in Chapter 11.

12.2.1 Social Class

Being working class might be a strong indicator of disadvantage. Students stating that they are working class are remarkably under-represented in HE participation and there are close relationships between being working class and other indicators of disadvantage. As Table 12.2.1 shows, in the working-class group there are more individuals with other potential indicators of disadvantage such as rural hukou (73%) and less-educated parents (74%) than in the middle-class group (54% and 58% respectively) and the upper-class group (57% and 54% respectively). Working class students are likely to be doubly disadvantaged.

Table 12.2.1 Percentages of working class, middle class and upper class individuals with potential disadvantages

	Working class	Middle class	Upper class	Missing class
Rural hukou	73	54	57	65
Minority ethnicity	10	8	14	8
Parents' education: compulsory education or below	74	58	54	65
Self-employed (father)	67	52	52	54
Other workplace (father)	3	2	1	3
Self-employed (mother)	66	49	51	54
Other workplace (mother)	3	2	2	4
Not very good at speaking Mandarin	46	32	39	50
Not good at understanding of Mandarin	61	53	44	64
Total (N)	4,798	4,934	277	80

Despite these clear links, however, it would still be problematic if being working class were used as the sole indicator in CA policies. The first reason is that it is vaguely defined and hard to measure, as was discussed in Chapter 7. In addition, as Table 12.2.1 indicates, the sole use of working class as an indicator in CA policies might incorrectly exclude about half the rural hukou holders, half the non-proficient users of Mandarin, half the students with self-employed parents and half the students with less-educated parents. It might also incorrectly include some possibly advantaged groups, such as those whose parents have a degree.

Furthermore, the upper-class group curiously seems more disadvantaged than the middle-class group. The former group has a little higher percentages of other indicators of disadvantage such as rural hukou, minority ethnicity and less fluency in speaking Mandarin (see Table 12.2.1). Besides this, the upper-class group shows a lower percentage of HE participation in Table 11.3.2. The possible reasons for the disadvantage of upper class have been discussed in Chapter 11 including vague definition, inaccurate information collection, misunderstanding for the questions or/and the term and smaller sample size. These problems might further damage the appropriateness of social class in the CGSS data being used as an indicator to identify the disadvantaged.

80 of the 10,089 CGSS cases lack social class information, and it seems that this group of cases suffer from considerable disadvantage. These cases have less-educated parents (65%) and less proficient language ability (speaking 50%; understanding 64%) and more of them are rural hukou holders (65%). Although 80 out of 10,089 is a small proportion, it would still be unjust to ignore this group or to assume the missing data are missing at random if social class is regarded as an indicator of disadvantage in HE admissions. However, it would also be unfair if the fact itself that class information is missing is regarded as an indicator of disadvantage, which might not only result in "false positives" but might also "provide an incentive" for individuals and families to conceal their social class information deliberately (Gorard, Siddiqui & Boliver, 2017, p82).

12.2.2 Parental education

Another important indicator, parental education, also seem to be linked to other disadvantages. Table 12.2.2 shows that low parental education is related to common indicators of disadvantage, in particular having self-employed parents and being a rural hukou resident. That is, individuals with parents who only completed basic education or below are more likely to come from a working-class family (54%), have self-employed parents (father 71%; mother 68%), have difficulty using Mandarin (speaking 46%; listening 63%), have rural hukou (84%) and belong to minority ethnicity groups (10%) than their counterparts with better-educated parents.

On the other hand, individuals whose father or mother has a bachelor's qualification are notably advantaged. For instance, they are less likely to have rural hukou (15%) and to have difficulty using Mandarin (speaking 14%; listening 35%). This might be because access to HE is a way to change one's hukou status. Parents who went into HE could get urban hukou, and their children would then inherit this status. As for language ability, well-educated parents might be more fluent in Mandarin themselves, which would help their children's Mandarin ability to some extent.

Furthermore, parents with bachelor's degrees are less likely to be self-employed (father 15%; mother 15%). This may be related to a policy implemented in the 1980s and 1990s. In that period the government allocated a job to graduates after they graduated from HEIs. Although in reality self-employment is not necessarily a disadvantage, in Chapter 11 this characteristic was linked with under-representation in HE participation. The last important thing is the highest percentage (11%) of minority ethnicity people in the group with well-educated parents. This high percentage shows the complexity of minority ethnicities and again emphasises that there could be bias if all minority ethnicities were regarded as disadvantaged.

Table 12.2.2 Percentages of individuals in parental education groups with potential disadvantages

	Both compulsory education or below	At least one received post-compulsory education	At least one received higher education	Parental education information missing
Rural hukou	84	58	15	
Minority ethnicity	10	7	11	6
Working class	54	37	19	48
Self-employed (father)	71	38	15	49
Other workplace (father)	3	1	2	1
Self-employed (mother)	68	38	15	44
Other workplace (mother)	3	1	3	1
Not very good at speaking Mandarin	46	26	14	42
Not good at understanding of Mandarin	63	46	35	62
Total (N)	6,596	3,062	337	94

In summary, parental education is closely linked to other acknowledged indicators of disadvantage. It is also relatively reliable because there is an official website where people's educational qualifications can be checked. This is a kind of micro-level indicator. Therefore, parents having a basic level of education or below could be used as an indicator in CA policies if there are no better options.

94 out of 10,089 cases lack parental education information, and they represent 16% of the cases with HE participation (see Table 11.3.3). According to Table 12.2.2, people missing this information are more closely linked with other potential indicators of disadvantage. Again, this group, although small in this sample, should neither be ignored nor directly labelled disadvantaged.

12.2.3 Parental workplaces

Parental workplace is another indicator related to HE participation. The previous chapter showed that individuals with self-employed parents or parents working in a "social community/committee" (father) or "other" workplace are disadvantaged in HE participation. It is clear from Table 12.2.3 that having a self-employed father and having a father working in an "other" workplace are strongly linked to rural hukou residence, being working class, having parents with little education, and having difficulty in speaking and understanding Mandarin. For instance, 54% of those whose parents are "self-employed," 42% of those whose parents work in a "social community/committee" and 46% of those whose parents work in an "other" workplace identify themselves as working class and the corresponding figures are 21%, 37% and 32% for people whose parents work in "party/government agencies," "companies" and "institutions" respectively.

Table 12.2.3 Percentages of individuals with different father's workplaces who have potential disadvantages

	Party/ government	Companies	Institutions	Social community/committee	Self-employed	Army	Other	Workplace missing

	agencies							
Rural hukou	18	31	29	86	80	35	81	49
Minority ethnicity	8	5	11	10	10	4	9	10
Working class	21	37	32	42	54	48	46	58
Parents education: compulsory education or below	17	46	23	66	79	52	83	62
Self-employed (mother)	20	18	29	38	82	39	16	26
Other workplace (mother)	1	1	1	3	0	0	66	2
Not very good at speaking Mandarin	24	25	27	42	46	43	48	34
Not good at understanding of Mandarin	42	46	44	59	62	48	65	56
Total (N)	276	1,929	719	97	5,930	23	261	854

Maybe worse, having a self-employed father is closely related to having a self-employed mother (82%) and vice versa (84%, see Table 12.2.4). These links are more pronounced for those whose mothers are self-employed or whose workplaces are unknown, as Table 12.2.4 shows. Therefore, individuals with a self-employed parent might be doubly disadvantaged.

Given the strong link between having self-employed parents, having parents with unknown workplaces and other indicators of disadvantage, it seems reasonable to view "self-employed parents" and parents working in “other” workplaces as indicators of disadvantage. However, this conclusion needs more consideration. First, it might lead to artificial exclusions of possibly disadvantaged individuals and inclusions of possibly advantaged ones. More importantly, parental workplaces may not be a valid proxy for parents' occupation, as was explained in Chapter 7.

Table 12.2.4 Percentages of individuals with different mothers' workplaces who have potential disadvantages

	Party/ government agencies	Companies	Institutions	Social community/committee	Self-employed	Army	Other	Workplace missing
Rural hukou	14	22	13	73	82	0	49	85
Minority ethnicity	9	5	9	8	11	0	8	7
Working class	23	34	26	50	55	0	54	42
Parents' education: compulsory education or below	11	40	17	46	78	0	81	61
Self-employed (father)	8	11	11	32	84	0	9	40
Other workplace (father)	0	0	0	0	1	0	72	2
Not very good at speaking Mandarin	14	21	19	41	47	0	44	33
Not good at understanding of Mandarin	30	40	41	53	63	0	62	56
Total (N)	104	1,346	531	66	5,759	1	238	2,044

854 and 2044 out of 10,089 individuals respectively lack information on their fathers' and mothers' workplaces, as this is somewhat sensitive information. Compared with the other groups, there is no explicit evidence that missing information on parents' workplace indicates more disadvantage than the previously discovered group of those with self-employed parents. However, this group does show a closer link with other indicators of disadvantage than

having parents who work in more prestigious places. Nevertheless, missing parents workplaces cannot be used as a CA indicator because the information might be deliberately concealed.

12.2.4 Language ability

The ability to understand and speak Mandarin also seems to be related to other possible indicators of disadvantage in HE admissions. Table 12.2.5 and Table 12.2.6 show that among those with difficulty in using Mandarin there are higher percentages of people with other disadvantages than among those who are more proficient. They are more likely to be rural hukou residents (84%, 84%, 76% and 66% for understanding; 92%, 83% and 75% for speaking) or have a minority ethnicity (22%, 28%, 12% and 9% for understanding; 25%, 18% and 10% for speaking). Furthermore, individuals who are not good at listening to (understanding) Mandarin are also less likely to be good at speaking Mandarin (84%, 98%, 94% and 32%), and vice versa (83%, 87% and 80%).

However, it would still not be appropriate to use ability to understand and speak Mandarin as an indicator on the basis of analysis of CGSS data. The first problem is the variables themselves. This is because in the CGSS, the interviewees are aged over 18. At many of these ages the individuals have already attended university or graduated from university and their self-reported answers on Mandarin language ability might not be the same as their ability before entering HE, as language ability is changeable. In other words, the reason why there are higher percentages of proficient Mandarin users among HE entrants in Table 11.3.6 and Table 11.3.7 could be that university entrants get more opportunities to learn or practise Mandarin after being admitted by universities rather than that they were excellent at Mandarin and then were more likely to be admitted to HEIs, as how things were before HE entry is not known from CGSS data.

Second, ability to understand spoken Mandarin and to speak Mandarin might be poor indicators of Mandarin language ability because reading and writing are also important, and even more important in education, as Mandarin is required to be used in most examinations in China. Therefore, it is not hard to imagine that students who use different oral and written languages, such as Tibetan, might be more disadvantaged than those who only speak a Chinese dialect. If the ability to understand and speak Mandarin is used to represent language ability as an indicator of disadvantage, this indicator might inadvertently doubly disadvantage people who use different written languages.

Table 12.2.5 Percentages of individuals with different levels of ability to understand (listen to) Mandarin with potential disadvantages

	Cannot understand at all	Bad at understanding	Not bad at understanding	Good at understanding	Very good at understanding	Understanding ability missing
Rural hukou	84	84	76	66	54	25
Minority ethnicity	22	28	12	9	7	0
Working class	53	68	56	48	42	50
Parents' education: compulsory education or below	84	83	79	68	56	75
Self-employed (father)	78	78	70	61	51	75
Other workplace (father)	2	5	3	3	2	0
Self-employed (mother)	82	65	70	59	49	50
Other workplace (mother)	2	4	3	3	2	0
Not very good at speaking Mandarin	84	98	94	32	17	25

Total (N)	49	193	1,884	3,594	4,365	4

Table 12.2.6 Percentages of individuals with different levels of ability to speak Mandarin with potential disadvantages

	Cannot speak Mandarin at all	Bad at speaking Mandarin	Not bad at speaking Mandarin	Good at speaking Mandarin	Very good at speaking Mandarin	Ability to speak Mandarin missing
Rural hukou	92	83	75	61	47	25
Minority ethnicity	25	18	10	8	7	0
Working class	69	67	54	45	39	50
Parents education: compulsory education or below	87	83	76	63	52	75
Self-employed (father)	78	76	68	57	46	50
Other workplace (father)	3	4	3	2	2	0
Self-employed (mother)	81	74	68	55	43	50
Other workplace (mother)	3	3	3	2	2	0
Not good at understanding of Mandarin	83	87	80	74	5	25
Total (N)	122	613	3,184	3,283	2,883	4

Third, even if the information were collected at an appropriate time and covered a broader language skill range, the variable would still be problematic as a policy indicator because it is difficult to measure. The self-reported language ability in the CGSS data certainly needs to be treated with caution, but even if standardised tests were used to evaluate language ability and the reliability were fully ensured, there would still be a risk of bias. If lower proficiency in Mandarin is used as an indicator of disadvantage in CA policies, some people might deliberately get lower scores in language ability tests to obtain contextualised university admission. After all, a measure of language ability is not as objective or accurate as that of the type of first language.

Only 4 of the 10,089 cases lack language ability data. Although in Table 12.2.5 and Table 12.2.6 there do not seem to be clear links between language ability information being missing and possible indicators of disadvantage except being working class, the group of cases is really too small for any conclusions to be reached.

12.3 Discussion and Conclusion

This chapter has examined missing data and the relationships between all possible indicators of disadvantage in the CGSS data. Because of the small numbers of cases missing data on some indicators, including hukou status, ethnicity and language ability, it is difficult to arrive at any persuasive conclusions about whether the missing data are related to disadvantages.

Only the missing data on social class, parental education and parental workplaces are sufficient for some conclusions to be drawn. Individuals missing social class data or parental education have other indicators of disadvantage, while those missing parents' workplace data are not among the most disadvantaged. However, it is still not safe to consider missing data as a proxy for disadvantage in CA policies even though there might be a strong link with disadvantages, because some parents might deliberately conceal this information to pretend to be disadvantaged and get contextualised admission for their children.

As for the relationships between the indicators, although all of them seem related to other potential indicators of disadvantage, few of these relationships could be considered indicators of disadvantage. One of these is hukou status and parental education. The combination of rural hukou and parents with only compulsory education or below are closely related to lower HE participation and other indicators of disadvantage. In addition, they are easy to measure and are collected as administrative information, which can be more accurate and has fewer missing data. They are relatively safe, but hukou status is still subject to a risk of 'ecological fallacy.'

However, parents' workplaces, language ability, social class and the currently-used indicator of ethnicity might need more investigation before making any decision on using them as indicators in CA policies because of doubtful representativeness of the variables, difficulties in measuring and other underlying complicated issues.

The next chapter employs a series of binary logistic regressions with CGSS data to explore how likely it is that indicators can predict participation in HE. Robust predictions might help answer the research questions about better CA indicators.

Chapter 13 Regression Analyses of the CGSS Data

Chapters 11 and 12 focused on the CGSS data and showed clustering of HE participation by various indicators and explored relationships between the indicators. This chapter looks at finding the extent to which these indicators predict HE participation outcomes in order to select suitable indicators.

13.1 Percentage Variation

Chapter 7 showed the order that the variables were included in the following regression models. The birth month variable was not put in the models because it was not found to indicate any particular disadvantage. Moreover, birth month data are only included in three cross-sectional datasets. Around 2,000 of the 10,089 cases would lack these data if this variable were included.

Although individual-level indicators (sex, ethnicity and hukou status) in CGSS make slightly more contributions to the prediction for HE participation than family-level indicators (parents' education, parents' workplaces and social class), with the increase of 19.1 and 18.6 for the correctly predicted percentage respectively, family-level indicators are more micro-level and could more accurately indicate ones' socio-economic status. Therefore, this chapter only presents the results for Model 2, in which family-level indicators are put into the first block. The results for Model 1 can be found in Appendix 10.

Table 13.1.1 presents the results for the second model and the correctly precited percentages averaged from ten estimates. The complete results of the ten regressions can also be found in Appendix 10. According to Table 13.1.1, the family-level indicators increase the average correctly predicted percentage from 50 to 68.6, language abilities contribute 2.2 more percentage points, and sex, ethnicity and hukou status increase the prediction by nearly 2 more percentage points.

Table 13.1.1 Summary of average correctly predicted percentages of HE entries in ten regressions (Model 2; enter method)

	Average correctly predicted percentage	Percentages increase
Base	50.0	
Block 1 (parents' education, parents' workplaces, social class)	68.6	18.6
Block 2 (ability to understand and speak Mandarin)	70.8	2.2
Block 3 (sex, ethnicity, hukou status)	72.6	1.8

N=4194

The results above indicate that once we know students' family background, their self-reported language abilities and their sex, ethnicity and hukou status become less important in predicting HE attendance.

The Appendix 10 also presents the results for both models by using forward stepwise (conditional) method. When the method was changed from enter to forward stepwise (conditional), the patterns in the two models did not change very much but the results from the forward stepwise (conditional) method showed which indicators were more important in these models. For instance, sex and father's workplace were always excluded from Model 1 with the forward stepwise (conditional) method. Only adding hukou status in Model 1 increased the correctly predicted percentage from 50 to nearly 69. In Model 2, sex and ethnicity were excluded. Although father's workplace was included in the model, this indicator contributed very little to the prediction, as did the mother's workplace and social class indicators. However, parents' education increase the accuracy of Model 2 from 50% to 67.4% (see Appendix 10). Therefore, hukou status and parental education may be very important in predicting individuals' HE admissions.

13.2 Coefficients

Table 13.2.1 list the average coefficients on each indicator in ten estimates for Model 2. The reference values are both parents with compulsory education or below, self-employed mother, self-employed father, working class, cannot understand Mandarin at all (listening) and cannot speak Mandarin at all, female, minority ethnicity and rural hukou.

To sum up, the higher parental education is, the greater the likelihood of HE participation. The gaps between students with parents who have completed HE and those with parents who only finished compulsory education are more prominent than those between hukou status groups. To a small certain extent missing parental educational information seems related to disadvantage.

Having a self-employed mother is linked with disadvantage in the model, while other types of workplace, including "other" and missing workplace information increase the, chance of HEI admission to different extents. So does having a self-employed father, but the differences in the coefficients for father's workplace are a little smaller than those for mother's workplace.

The chance of entering HE for upper class individuals is about 10% lower than the reference value for working-class peers, although in Chapter 11 the upper-class group had a higher percentage of HE admissions than the working-class group. This strange disadvantage might be attributed to the limitation of the variable itself, as was discussed in Chapters 11 and 12. It is also possible that the variable of social class has been over corrected after controlling other variables or it might be the result of smaller sample size of upper class group.

Compared to those who cannot speak Mandarin at all, individuals using Mandarin fluently are more likely to enter universities. Those who self-evaluated as "very good at speaking Mandarin" are around ten times more likely to enter HE than those who cannot speak at all. Although it is narrower, the gap in listening to/understanding Mandarin is also remarkable at nearly five times. However, individuals who self-evaluated as understanding Mandarin "badly" have a higher chance of entering HE than those who rated themselves "not bad" and "good" at understanding Mandarin. This may be because of the relatively small size of the group of people who chose “bad”. There are a very few missing data about language ability.

Table 13.2.1 Summary of the coefficients on the variables in Model 2

	Enter method
both parents have compulsory education qualifications or below	
at least one parent has a post-compulsory education qualification	2.06
at least one parent has a HE degree	4.53
parental education qualification missing	0.97
self-employed (father)	
company (father)	1.28
institution (father)	1.21
social community/committee (father)	0.95
party/government agency (father)	1.19
army (father)	1.20
other workplace (father)	1.07
workplace missing (father)	1.02
self-employed (mother)	
company (mother)	1.22
institution (mother)	1.66
social community/committee (mother)	2.16
party/government agency/army (mother)	1.58
other workplace (mother)	1.25
workplace missing (mother)	1.32
working class	
middle class	1.20
upper class	0.86
class missing	1.03
cannot understand at all	
bad at understanding	4.38
not bad at understanding	1.96
good at understanding	3.64
very good at understanding	4.78
cannot speak at all	
bad at speaking	1.76
not bad at speaking	4.61
good at speaking	7.00
very good at speaking	9.50
female*	
male*	1.10
minority ethnicity*	

Han ethnicity*	1.15
rural hukou	
urban hukou	2.43

Note: * means that the variable has been excluded using the forward stepwise method

Finally, male, Han ethnic, and urban hukou students are more advantaged in HE transition than female, minority ethnic and rural hukou students.

Although there are few differences in coefficients for each variable in regression models when the forward stepwise (conditional) method was employed, sex and ethnicity were always excluded from the regressions in forward method (see Appendix 10). This might be because these two indicators are less associated with HE participation when other indicators are known. This finding again warns that minority ethnicity is not a suitable indicator in Chinese CA policies even though it has been used.

Although hukou status and parental education qualifications both are important, there is much more risk of false-negatives and false-positives with hukou status. They are more like group-level indicators. It is hard to imagine, for example, that all rural hukou holders are disadvantaged while all urban hukou holders are not. This might not be possible even in the same province (i.e. rural hukouers in Hunan province are not necessarily more disadvantaged than urban hukouers in the same province). This does not mean there are no gaps and inequalities between rural and urban hukou holders in HE participation. In the previous chapters and in preliminary studies in Chapter 9 considerable gaps were found. And it might be true that many, if not most, rural hukou holders are more disadvantaged than their urban counterparts. However, this group-level indicator might need a better replacement or might need to be used together with other indicators to identify disadvantage more accurately. The same applies to sex and ethnicity, which are also very likely to fall into an ecological fallacy.

On the other hand, parental education qualifications are readily available on an official website called xuexin wang. Besides good accessibility they are objective measures. Furthermore, they are associated with HE participation and are likely to be associated with parents' occupations or/and family income. If only one indicator is chosen for use in CA policies, according to the analysis of CGSS data parental education might be better than the currently employed hukou status and minority ethnicity indicators.

As has been explained before, parents' workplaces are not accurate indicators of occupations. Even if they were, they would still be problematic because the information is not officially collected. Social class suffers from almost the same problems, although researchers have recognised this variable for a long time (Wang, 2011; Wu, 2017; Xu & Fang, 2020; Yang, D. P., 2010; Zhang & Chen, 2014; Zhao, 2018). Neither of these variables in the CGSS data are appropriate to be used as indicators in CA policies. However, if there are similar but more objective family-level indicators, they deserve more consideration.

Finally, speaking and listening/understanding language abilities are also problematic. They might be too inaccurate to use in regressions, but it does not make much difference to the regression result when the two predictors are removed. The percentage variation that is lost is only around 2 percent. However, the coefficients of other predictors generally remain stable. Therefore, they were left in the models.

13.3 Discussion and Conclusion

To sum up, according to the results of the descriptive analysis and the regression analyses of CGSS data, parental education might be the best choice to use as an indicator to identify disadvantage in HE admissions. This information contributes a lot to increasing the correctly predicted percentages. Although hukou status seems to contribute more when added alone (see Appendix 10), it is possible to false-negatively exclude disadvantaged urban hukouers or false-positively include privileged rural hukouers. Parental education is more micro-level than hukou status.

Parents' workplace and social class also help increase the percentage variation a little. However, it is hard to obtain reliable and accurate information about these indicators. Besides low reliability, the father's workplace indicator does not show particular importance in predicting HE entry. Language abilities also make little contribution to predicting HE participation and self-evaluated language abilities in CGSS data are not good-quality indicators to identify disadvantaged students. sex and ethnicity are less relevant in outcomes when there are other better predictors, as they are usually excluded from the regression models with very few losses in the percentage variation. On the other hand, sex and ethnicity also risk ecological fallacies.

Furthermore, according to the coefficients of each variable, compared to their advantaged counterparts, people with parents who only finished compulsory education, people with self-employed parents, people from the working and upper class, less fluent language users and rural hukou residents are much more disadvantaged in HE admissions. That is, they are much less likely to receive HE.

However, it would be incorrect to assume that the non-HE group has the same disadvantage. Students who had to drop out of primary school and those who left education because they failed the NCEE would have been disadvantaged differently. However, as it is a cross-sectional study, the CGSS does not allow researchers to follow respondents or trace them back to their earlier educational experiences, but it provides more specific education levels of interviewees. The next chapter makes a static comparison of clusterings of education levels by different indicators to discover earlier gaps before selection by the NCEE.

Chapter 14 Clustering of Educational Qualifications by Different CGSS Indicators

According to Chapter 9, nearly half of students leave education at the transition to high schools, where the inequality in education might have existed. Although the CGSS, which is a cross-sectional survey, cannot trace back earlier educational experiences of cases, it is still possible to compare the educational qualifications they obtained to investigate whether there are differences or inequalities in continuation of education which happened before the transition to HE. This chapter will make this comparison to answer the question of whether there are accumulated disadvantages in education.

14.1 Educational Qualifications in the Overall CGSS Sample

As Table 14.1.1 shows, most of the CGSS interviewees at least finished compulsory education, and then 33% of them failed to go to post-compulsory education. Only 65.2% attended and finished high school or equivalent. Then 40.7% gained qualifications from colleges. Only around 21% of these respondents finally participated in HE.

Table 14.1.1 The numbers and percentages of each education level group

	Not educated	At least finished compulsory education	At least finished high school or equivalent	At least finished 2- or 3-year vocational college	Finished higher education	Education level missing
Number	140	9,943	6,577	4,101	2,097	6
Percentage	1.4	98.6	65.2	40.7	20.8	

Differences in HE participation do not start at the point of applying for HE. About a third of the cases left education once they finished compulsory education at the age of 15. Therefore, if only NCEE candidates are viewed as the target group in research on HE equality and equity, or in CA policies, there would be a risk of missing truly disadvantaged students and only providing less disadvantaged ones with extra help.

Only 6 cases lack information on their education levels, which is a very small number considering the total number of cases is 10,089. This might be because the 10,089 cases selected from the larger merged dataset of 46,753 cases are the younger ones (born after 1980), who would have less difficulty in identifying their educational level.

The following sections examine differences in educational qualifications for each indicator of disadvantage to explore when disparities in HE participation actually start for each group.

14.2 Differences in Educational Qualifications by Currently Employed Indicators

This section explores how education levels are patterned by the indicators currently used: hukou status and ethnicity.

14.2.1 Hukou status

As Table 14.2.1 shows, there is a considerable gap between the education qualifications of rural hukou and urban hukou holders and this gap occurs much earlier than the transition to HE. Nearly half of the rural hukou residents left education after finishing compulsory education. However, 91% of their urban hukou counterparts remained. Then less than half of the remaining rural hukou holders went to post-compulsory education, while the corresponding figure for urban hukou holders is more than two-thirds.

Table 14.2.1 Percentages of different educational levels by hukou status

	Not educated	At least finished compulsory education	At least finished high school or equivalent	At least finished 2- or 3-year vocational college	Finished higher education	Total (N)
Rural hukou	1.9	98.0	50.0	24.3	10.7	6,331
Urban hukou	0.4	99.5	90.9	68.2	37.8	3,734
Hukou missing	-	-	-	-	-	24

The educational inequality between rural hukou and urban hukou holders seems to start at the transition to post-compulsory education. This inequality might be attributed to several causes. The first is an unbalanced distribution of resources. In development over decades, cities have accumulated more economic and educational resources than rural areas. Students in cities can enjoy richer resources. The other issue is the hukou policy, which restricts internal migration for education and social welfare. Rural hukou students cannot go to cities to participate in good-quality education whenever they want.

Therefore, it is more difficult for rural hukou holders to enjoy good-quality basic education. In order to improve the HE participation of rural hukou students, it is necessary to help them during their compulsory education and improve their participation in high schools. Or the government would better expedite the progress of abolishing the hukou status differences, as the disadvantages of rural hukou students could be imputed to this artificial distinction to an extent.

14.2.2 Minority Ethnicity

In Chapter 11, no large gap was found between the HE participation of minority ethnicity groups and that of the Han majority, although the former were slightly under-represented in HE. However, as Table 14.2.2 shows, minority ethnicities are notably over-represented among those who only obtained compulsory education qualifications or even lower. Less than half the minority ethnicity people obtained a high school education qualification or equivalent, but two-thirds of Han ethnicity people went to high school and finished high school education. As for the transition to post-secondary education and HE (which does not include other types of post-secondary education such as vocational colleges in this study, as mentioned in Chapter 2), minority ethnicity groups were still less likely to remain in education, but the gaps are smaller than those in the transition to high school.

The gaps in education between Han and minority ethnicity groups are more serious in the period after compulsory education. Maybe more effort needs to be made at this stage rather than after the NCEE, which could mean that it might be better to give minority ethnicity students extra scores during the transition to post-compulsory education rather than in the NCEE. Current CA policies ignore gaps between groups before the NCEE, which is problematic in terms of equity.

Table 14.2.2 Percentages of different educational levels by ethnicity

	Not educated	At least finished compulsory education	At least finished high school or equivalent	At least finished 2- or 3-year vocational college	Finished higher education	Total (N)
Minority ethnicity	4.3	95.7	49.2	30.3	15.2	902
Han ethnicity	1.1	98.8	66.8	41.7	21.3	9,176
Ethnicity missing						11

14.3 Differences in Educational Qualifications by Other Indicators

This section explores how education levels are patterned by other potential indicators of disadvantage.

14.3.1 Sex

No remarkable sex differences in HE participation were found in Chapter 11. However, more females had no education (2.1%) or just basic education (35.7%) up to the end of middle school. The uneducated rate for males was only 0.6%, and 31% of males left education after the end of compulsory education. In subsequent transitions the leaving rates of the two sexes are similar.

Sex inequality in education in China does not seem to be serious but as Table 14.3.1 shows females are a little more disadvantaged than their male counterparts. They face more barriers in the transition to post-compulsory education than their male counterparts. Although it might not be necessary to provide female HE applicants with additional support in CA policies, some attention should be paid to decreasing female uneducated rates and improving female entry to post-compulsory education. These uneducated or basically-educated females are more likely to have rural hukou or belong to minority ethnic groups, according to the structured review in Chapter 9.

Table 14.3.1 Percentages of different education levels by sex

	Not educated	At least finished compulsory education	At least finished high school or equivalent	At least finished 2- or 3-year vocational college	Finished higher education	Total (N)
Female	2.1	98	62.3	40.3	20.1	5,275
Male	0.6	99.4	68.4	41.1	21.5	4,814

14.3.2 Social Class

As Table 14.3.2 shows, working class students are remarkably disadvantaged in their educational qualifications. They are more likely to be uneducated (1.8%) than middle class (1%) and upper class (0.4%) students. Importantly, working-class students are less likely to obtain a high school qualification or equivalent (56%) than the other two groups after finishing compulsory education. This disadvantaged group continued to be disadvantaged in obtaining at least post-secondary education. The figures for the middle-class and upper-class groups are 66% (49.1%/74%) and 69% (41.5%/68.2%). The differences in obtaining a HE qualification, which receive the most attention, are the smallest.

The largest percentage of students leaving education was in the transition to high school or equivalent. This is another indication that current CA policies which only pay attention to NCEE candidates are unfair.

Table 14.3.2 Percentages of different educational levels by social class

	Not educated	At least finished compulsory education	At least finished high school or equivalent	At least finished 2- or 3-year vocational college	Finished higher education	Total (N)
Working Class	1.8	98.2	56	32	15.2	4,798
Middle Class	1	99	74	49.1	26.3	4,934
Upper Class	0.4	99.6	68.2	41.5	21.3	277
Class missing	2.5	97.5	57.5	36.3	16.3	80

14.3.3 Parents' Education

Students with only basically educated parents and those not reporting their parents' education qualification information are both over-represented in the uneducated group, at 1.9% and 2.1% respectively (see Table 14.3.3).

Gaps in education levels again start early in school education. Only 98.1% of people whose parents only finished compulsory education got compulsory education qualifications, 56% (54.6%/98.1%) of who further finished high school education or equivalent. Then around 52% (28.4%/54.6%) of those who finished high school education got a post-secondary education qualification. On the other hand, more than 99% of people whose father or mother at least completed post-compulsory education and those whose father or mother completed HE got a compulsory education qualification. 86% (85.3%/99.5%) and 94% (81.6%/93.2%) of those with compulsory education qualifications in the last step in these two groups obtained a high school or equivalent qualification. 74% (62.8%/85.3%) and 88% (81.6%/93.2%) of these high school graduates then got into post-secondary education. Finally, students with better-educated parents are more over-represented in HE, as discussed in Chapter 11.

Again, the more significant gaps in education levels are long before selection in the NCEE, particularly in continuation after compulsory education. Children's education levels are highly associated with their parents' education levels.

Table 14.3.3 Percentages of different educational levels by parental education

	Not educated	At least finished compulsory education	At least finished high school or equivalent	At least finished 2- or 3-year vocational college	Finished higher education	Total (N)
Both parents had only compulsory education or below	1.9	98.1	54.6	28.4	12.2	6,596
At least one parent had post-compulsory education	0.4	99.5	85.3	62.8	35.1	3,062
At least one parent had higher education	0.6	99.4	93.2	81.6	62	337
Parental education missing	2.1	94.7	56.4	36.2	16	94

14.3.4 Parents' Workplaces

There are similar patterns of education level according to mother's and father's workplaces. As Tables 14.3.4 and 14.3.5 show, students whose parents work in party/government agencies, companies and institutions are advantaged in their education level. On the other hand, students with a self-employed mother or father and those with a mother or father working in an "other" place are more likely to be disadvantaged. These gaps emerge after the end of compulsory education. Only a little more than a half of this disadvantaged group successfully transited to post-compulsory education, while the corresponding figures range from 87% to 96% for people whose parent(s) worked in more prestigious places. Then another half of the group of high school graduates transited to and completed post-secondary education, while the rates for the advantaged groups range from 70% to 80%.

Individuals with a father working in a social community/committee are somewhat disadvantaged, with 66% of compulsory education graduates obtaining high school education qualifications and 57% of high school graduates obtaining post-secondary education qualifications. However, individuals with a mother working in a social community/committee do not have this disadvantage, as 81% of the compulsory education graduates and 79% of the post-compulsory education graduates in this group continued and completed further education stages.

Table 14.3.4 Percentages of different education levels by father's workplace

	Not educated	At least finished compulsory education	At least finished high school or equivalent	At least finished 2- or 3-year vocational college	Finished higher education	Total (N)
party/government agency	0.4	99.6	92	72.8	45.7	276
Company	0.4	99.6	87.5	61.9	33.9	1,929
institution	0.4	99.6	86.9	67	39.4	719
social community/committee	2.1	97.9	65	37.1	14.4	97
self-employed	1.8	98.2	53.8	28.9	13.1	5,930
Army	4.4	95.7	60.9	56.5	34.8	23
Other	2.7	96.9	51.3	31	13.4	261
Father's workplace missing	1.5	98.2	71	44.9	23.9	854

Table 14.3.5 Percentages of different education levels by mother's workplace

	Not educated	At least finished compulsory education	At least finished high school or equivalent	At least finished 2- or 3-year vocational college	Finished higher education	Total (N)
party/government agency	0	100	96.2	76	54.8	104
Company	0.4	99.6	91.2	67.2	36.1	1,346
institution	0	100	96.2	78.3	50.3	531
social community/committee	3	97	78.8	62.1	28.8	66
self-employed	1.9	98.1	52.9	27.8	12.4	5,759
Army	0	100	100	100	100	1
Other	2.5	97.1	55	35.4	14.3	238
Mother's workplace missing	0.9	99	73.7	47.7	25.4	2,044

As parents' workplaces are somewhat sensitive variables, they suffer from a large amount of missing data. Students whose information on their parents' workplaces remains unclear are not necessarily disadvantaged, and they even show an advantage in obtaining education qualifications. Again, this may be a reminder that this self-reported indicator is not necessarily a good proxy for parents' occupation categories.

14.3.5 Language Ability

The educational levels of students in different language ability groups are very different. As Table 14.3.6 and Table 14.3.7 show, people who cannot understand or speak Mandarin at all and those who are bad at understanding or speaking Mandarin are much more likely to be uneducated than other groups. The biggest gap is between compulsory education and high school education. At this transition, the lowest rates are for those who are "bad" at understanding Mandarin and the group that "cannot speak Mandarin at all," at 29% (25.9%/89.1%) and 16% (12.3%/77.9%) respectively. The highest rates are for the groups that are "very good at" understanding and speaking Mandarin, at 78% and 83%, respectively.

Table 14.3.6 Percentages of different education levels by ability to understand Mandarin

	Not educated	At least finished compulsory education	At least finished high school or equivalent	At least finished 2- or 3-year vocational college	Finished higher education	Total (N)
cannot understand at all	18.4	81.6	30.6	10.2	2	49
bad	10.4	89.1	25.9	11.9	5.2	193
not bad	2.2	97.7	44.2	19.7	6.4	1,884
good	1.4	98.6	64.2	37	17.9	3,594
very good	0.4	99.5	77.2	54.4	30.3	4,365
Understanding ability missing	-	-	-	-	-	4

Table 14.3.7 Percentages of different education levels by ability to speak Mandarin

	Not educated	At least finished compulsory education	At least finished high school or equivalent	At least finished 2- or 3-year vocational college	Finished higher education	Total (N)
cannot speak at all	22.1	77.9	12.3	6.6	1.6	122
bad	6.2	93.6	28.6	10.6	3.8	613

not bad	1.3	98.6	52.2	25.7	10.1	3,184
good	0.8	99.2	71	43.1	21.8	3,283
very good	0.3	99.7	83	62.1	35.8	2,883
Missing speaking ability	-	-	-	-	-	4

There are also gaps in continuation to although they are less prominent than those above. Nevertheless, the less fluent Mandarin user groups are more likely to be less educated, and many of them miss out on the last two stages.

14.3.6 Birth Month

Table 14.3.8 compares the educational attainments of summer-born and autumn-born students. There seem to be few differences between the two groups, although summer-born students show a negligible advantage over autumn-born students in high school, college and university entry of 2, 1.6 and 0.5 percentage points. In all, the two groups are in similar situations.

Table 14.3.8 Percentages of different education levels by birth month

	Not educated	At least finished compulsory education	At least finished high school or equivalent	At least finished 2- or 3-year vocational college	Finished higher education	Total (N)
summer-born children	1.4	98.6	65.1	39.5	19.2	3,381
autumn-born children	1.1	98.9	63.1	37.9	18.7	3,435
Birth month missing	-	-	-	-	-	148

There do not seem to be remarkable differences between summer-born and autumn-born students in nearly all the stages of education. This might be attributed to the high aspirations and expectations of Chinese parents for their children’s education, which prompts parents to invest in children’s education and narrows the gaps in educational attainments due to birth months On the other hand, the classifications in the table are rough without any distinctions according to the prestige of schools/HEIs. There might be some differences between the two birth-month groups if this were taken into account.

14.4 Discussion and Conclusion

This chapter has shown inequalities in education levels by the different indicators. Almost all indicators (except birth month) discussed in the previous chapters identify some disadvantaged students. They are more likely to be less educated such as only finishing compulsory education or, even worse, not being educated at all.

Although it is true that the groups with these indicators, such as rural hukou residents, are under-represented in HE participation, which is certainly a problem that needs more attention, most of them leave education after compulsory education rather than after failing in the NCEE, as is commonly believed. The most remarkable gaps in education equality and equity are always at this point. Therefore, in order to improve HE participation for disadvantaged groups it is not enough to only give them extra credits or support in the NCEE. Those who can arrive at the NCEE stage, even if they are more disadvantaged than other candidates, are already the winners in the “semi-final”. Among the whole sample at least 35% of the competitors are excluded from the competition for HE after the end of middle school, but according to the analysis above the percentages for the disadvantaged groups are much higher, up to 88%.

Even worse, this exclusion might happen in middle school enrolments, in completion of primary school or in primary school enrolments, especially for disadvantaged students. More effort and attention therefore need to be given to disadvantaged students in the earlier stages of their education experience.

In conclusion, first, according to CGSS, rural hukou residents face substantial disadvantage, but as a group-level indicator rural hukou is vulnerable to an ecological fallacy. If there are no better individual- or family-level indicators, rural hukou could be used for a while in combination with other indicators rather than being used alone, but it still may not be suitable in the long run because it might lead to another injustice.

Second, it does not seem that minority ethnicity is a suitable indicator of disadvantage in HE admissions. Students with a minority ethnicity are only slightly disadvantaged, which might be a good result from long-implemented CA policies for minority ethnicity groups. Moreover, according to the relationships between minority ethnicity and other indicators and the differences in the educational attainments of different minority ethnicities, it would be problematic to view all minority ethnicity people as disadvantaged or disadvantaged at the same level. The extra scores for this group, therefore, might need more consideration.

Parental education qualifications might be suitable to identify disadvantaged students. This indicator is not only related to disadvantage but can be more easily measured accurately and objectively than other indicators. Moreover, as has been mentioned, this information is more available than other information in China, as it is recorded on an official website called xuexin wang. Companies, institutions and other agencies can apply to legally check this information.

Social class and parental workplaces are far from appropriate as CA indicators, at least according to the analyses of the CGSS dataset. So is language ability. Even though language ability, especially the ability to speak Mandarin, is linked with disadvantage, the risk of bias, such as difficulty in measuring this ability, would be a challenge if this indicator were utilised in CA policies.

More importantly, inequalities in HE participation do not only happen in the NCEE but far before this selection. Nearly all potential indicators of disadvantage examined in previous chapters have also shown drop out in earlier stages of education in this chapter. The most serious exclusion from education for these groups is in the transition to post-compulsory education in high schools or equivalent. If students have not completed this period of education they can never enter the NCEE, let alone apply for HE. There seems to be a cumulative disadvantage. CA policies might be less valuable and helpful if they do not take these early-excluded groups into account.

The above arguments should be critically thought through before accepting them. This is because although the CGSS was chosen for analysis due to its sample size, it has some limitations such as a lack of school-level information and problematic variables, as was explained in Chapter 7. Therefore, the CFPS dataset will be introduced and analysed in the following chapters to help answer the research questions in this study. It is an informative longitudinal dataset with five waves from 2010 to 2018 which traced a group of children from primary school to the traditional age for HE.

Chapter 15 Exploring the Clustering of Education Participation by Different CFPS Indicators

Due to the limitations of the CGSS, another large-scale social survey named the China Family Panel Studies (CFPS) was analysed. The CFPS is a longitudinal survey that can trace the early educational experiences of interviewees, in which disadvantages might start to accumulate. Therefore, the chapters based on the CFPS trace individuals' early educational trajectories.

Chapters 15, 16 and 17 present both descriptive and regression results of analyses of CFPS data. The descriptive analyses of CFPS data are partly to identify gaps in HE participation and high school education participation revealed by different indicators. Relationships between indicators and the regression analyses are discussed to help understand the extent to which these indicators are appropriate for use in CA policies.

15.1 HE Participation Among the Overall CFPS Sample

As Table 15.1.1 shows, among the 2,490 cases remaining after the data were cleaned there are 610 HE participants and 1880 non-HE participants, accounting for 24.5% and 75.5% of the total sample. Only a quarter of the CFPS interviewees successfully entered HE. The following sections explore the disparities in HE participation by different indicators based on simple descriptive analyses.

Table 15.1.1 Percentages of HE participants and non-HE participants in the CFPS data

	HE	Non-HE	Total (N)
Sample distribution	24.5	75.5	2,490

The next section discusses the HE participation disparities revealed by officially used indicators.

15.2 Clustering of HE Participation by Currently Employed Indicators

This section compares percentages of HE participation revealed by three official CA indicators: hukou status, ethnicity and hukou province.

15.2.1 Hukou Status

The hukou status recorded by the CFPS is suitable for analysis because this information was collected before the age to enter HE. There are 484 urban hukou students and 1,986 rural hukou students, respectively accounting for 19% and 80% of the sample. Non-registrants and missing information account for 1% of the total sample. Table 15.2.1 shows disparities in HE participation by hukou status. 42% of the urban hukou students entered HE, which is much higher than the average rate of 24.5%, while only 20% of the rural hukou students participated in HE. The rural hukou group is thus under-represented in HE participation. Furthermore, according to the odds ratio, urban hukou students are nearly three times (2.9 times) as likely to enter HE than their rural hukou counterparts.

Table 15.2.1 Percentages of HE participation and non-HE participation by hukou status

	Total (N)	HE (%)	Non-HE (%)	Odds ratio
Urban hukou	484	42	58	2.9
Rural hukou	1,986	20	80	
No registration	17	6	94	
Hukou missing	3	-	-	
Total	2,490	24.5	75.5	

The third hukou status category in the CFPS is no registration. No registration means that the citizen does not have an official hukou. There are various possible reasons for this such as their birth not having been registered. Non-registered people have many inconveniences in their daily lives, including being unable to have good-quality education or even just basic compulsory education, so they are likely to have less chance of accessing HE. As Table 15.2.1 shows, non-registered students are much less likely to enter HE (6%) and are possibly the most disadvantaged. However, the sample size is small, so the result is not robust. Nevertheless, the enormous gap in HE participation between non-registered students and registered ones deserves attention in future research.

The number of missing cases is so small that no meaningful interpretation of the missing data is possible.

As was discussed in Chapter 2, hukou status does not necessarily indicate the area where somebody lives. Some previous studies (Jia & Ericson, 2017; Xu, 2021) have mentioned disadvantage resulting from living in a rural area. The CFPS includes information on a binary bureaucratic categorisation of the interviewees' living areas.

Table 15.2.2 Percentages of HE participation and non-HE participation by living area

	Total (N)	HE (%)	Non-HE (%)	Odds ratio
Urban area	929	33	67	2.0
Rural area	1,561	20	81	
Total	2,490	24.5	75.5	

As Table 15.2.2 shows, 929 students live in urban areas and 33% of them successfully entered HE, which is nine percentage points lower than those with urban hukou. Students living in urban areas are 2.0 times as likely to enter HE than those living in rural areas. That again indicates that the hukou status classifications play a more crucial role in exacerbating education inequality than living areas. Therefore, again, the abolition of this artificial classification of hukou status or equal right to internal migrate and attend education for all individuals, especially rural hukou individuals, in China are required.

15.2.2 Minority Ethnicity

Table 15.2.3 presents the results of clustering HE participation by another officially employed indicator: minority ethnicity. There are 2,213 cases with Han ethnicity and 269 with minority ethnicities. Within the Han group, 26% entered HE, while the corresponding figure is only 15% for minority ethnicity cases. People with Han ethnicity are 1.9 times as likely to be admitted to HE than minority ethnicity people. This gap is nearly as large as that for living areas in Table 15.2.2. Only 8 cases lack ethnicity information, which is too small a number to draw conclusions, so there is no percentage comparison for this category.

Table 15.2.3 Percentages of HE participation and non-HE participation by ethnicity

	Total (N)	HE (%)	Non-HE (%)	Odds ratio
Han ethnicity	2,213	26	74	1.9
Minority ethnicity	269	15	85	
Ethnicity Missing	8	-	-	
Total	2,490	24.5	75.5	

In short, according to the CFPS data it would be fair to say that minority ethnicity students are disadvantaged in HE participation compared to Han majority ethnicity ones.

15.2.3 Hukou Province

The CFPS asks for information on children's hukou provinces. However, if the original hukou province information were directly used in the analysis, some provinces have small representations in the sample, and so are vulnerable to extreme values. Therefore, the original province data was re-coded into area information, as is shown in Table 15.2.4.

Table 15.2.4 Re-coding of provinces into living areas

	Province
Municipalities	Beijing, Tianjin, Shanghai, Chongqing
Northeastern China	Liaoning, Jilin, Heilongjiang
Central China	Henan, Anhui, Hubei, Jiangxi, Hunan
Eastern China	Shandong, Jiangsu, Zhejiang, Fujian
Northern China	Hebei, Shanxi, Inner Mongolia
Southern China	Guangdong, Guangxi, Hainan
Southwestern China	Sichuan, Guizhou, Yunnan, Tibet
Northwestern China	Shaanxi, Gansu, Xinjiang, Ningxia, Qinghai

As Table 15.2.5 shows, students who live in the four municipalities are the most likely to enter HE. These are followed by those living in northern China, northeastern China, central

China, eastern China and northwestern China. Students who live in southern China and southwestern China are the most disadvantaged in HE participation.

Table 15.2.5 Percentages of HE participation and non-HE participation by hukou province living area

	Total	HE (%)	Non-HE (%)	Odds ratio
Municipalities	116	34	66	
Northeastern China	266	28	72	1.2
Central China	388	29	71	1.2
Eastern China	268	25	75	1.5
Northern China	303	30	70	1.2
Southern China	322	18	82	2.3
Southwestern China	361	16	85	2.7
Northwestern China	457	24	76	1.5
Province living area missing	9	-	-	
Total	2,490	24.5	75.5	

The most advantaged group, those living in municipalities, are more than twice as likely to participate in HE as the least advantaged groups, those living in southern and southwestern China. Those living in municipalities are also 1.5 times as likely to enter HE than those who live in eastern and northwestern China, and they are 1.2 times as likely to do so than those living in northeastern, central and northern China. The CFPS data show that there are disparities in HE participation between different areas, which supports the findings in Chapter 10.

The number of cases missing hukou province data is too small to draw any conclusions.

15.3 Clustering of HE Participation by Other Potential Indicators

This section explores HE participation disparities by potential indicators of disadvantage. Three groups are examined: individual-level, family-level and school-level indicators.

Individual-level indicators

15.3.1 Sex

Clustering of HE participation by sex groups is slightly imbalanced, as Table 15.3.1 shows. A higher percentage of females (29%) enter HE, slightly more than the average percentage of 24.5%, while fewer male students do so.

Table 15.3.1 Percentages of HE participation and non-HE participation by sex

	Total (N)	HE (%)	Non-HE (%)	Odds ratio
Female	1,215	29	71	1.5
Male	1,275	21	79	
Total	2,490	24.5	75.5	

Females are around 1.5 times as likely to enter HE than their male counterparts. According to the CFPS data, males are disadvantaged in HE participation.

15.3.2 Birth Month

The participants' months of birth were examined. Summer-born and autumn-born children were categorised in the same way as in Chapter 11. Children who were born between 1 September and 28 or 29 February were labelled autumn-born, while those born between 1 March and 31 August were labelled summer-born.

Table 15.3.2 Percentages of HE participation and non-HE participation by birth month

	Total (N)	HE (%)	Non-HE (%)	Odds ratio
Autumn-born children	1,146	26	74	1.1
Summer-born children	1,333	24	76	
Birth month missing	11	-	-	
Total	2,490	24.5	75.5	

Table 15.3.2 shows no great differences between the two groups, with 26% and 24% of autumn-born and summer-born children respectively participating in HE, and an odds ratio of 1.1. It might be possible to ignore the small gap between the birth-month groups in the CFPS

data. Nevertheless, it is possible that there is bias in the CFPS samples so that better data should be employed in future research. There are too few missing cases to make any important differences.

Family-level indicators

15.3.3 Parental Education Qualifications

Not surprisingly, higher percentages of students with better-educated parents enter HE. As Table 15.3.3 shows, interviewees whose father or/and mother completed HE and obtained a bachelor's degree were much more likely to be admitted to HE, with 62% doing so. Students whose father or/and mother stayed in post-compulsory education and got a corresponding qualification are also over-represented in HE. In contrast, students whose parents only finished compulsory education are under-represented among HE entrants, with only 20% participating in HE, fewer than the average of 25%.

Table 15.3.3 Percentages of HE participation and non-HE participation by parental education qualifications

	Total (N)	HE (%)	Non-HE (%)	Odds ratio
Both parents had compulsory education qualifications	1,953	20	80	
At least one parent had a post-compulsory education qualification	477	39	61	2.8
At least one parent had a HE qualification	50	62	38	
Parental education information missing	10	-	-	
Total	2,490	24.5	75.5	

Furthermore, compared with the disadvantaged compulsory-education group, the privileged group which includes the other two groups of students with parents completing post-compulsory education or HE is nearly 2.8 times as likely to enter university. The group with basically educated parents is disadvantaged in HE participation.

There are only ten cases missing parental education information. Again, this number is not large enough to draw persuasive conclusions.

15.3.4 Parental Political Status

Political status has not been discussed in the Chinese context as much as other indicators, possibly because it is politically sensitive. Being a party member is not possible for everyone but it is a kind of meritocracy. Therefore, party members are likely to come from advantaged elite groups. This assumption seems to be supported by the results presented in Table 15.3.4. The table shows that a high percentage (around 37%) of students whose parents are party members participate in HE. Students whose parents are not party members are slightly under-represented in HE admissions, at one to three percentage points less than the average percentage of 24.5%. Students with party member parents are nearly twice as likely to enter HE as those with non-party member parents.

Table 15.3.4 Percentages of HE participation and non-HE participation by parents' political status

		Total (N)	HE (%)	Non-HE (%)	Odds ratio
Mother	Party member	254	37	63	2.0
	Non-party member	2,129	23	77	
	Political status missing	107	20	80	
Father	Party member	433	37	63	2.1
	Non-party member	1,948	22	78	
	Political status missing	109	26	74	
Total		2,490	24.5	75.5	

Students whose mother's political status is missing have five percentage point lower HE participation than the overall average but those missing their father's political status do not show this tendency.

15.3.5 Parental Occupations

Table 15.3.5 and Table 15.3.6 present clustering of HE participation by parents' occupations. Unlike the similar variables in the CGSS, the parental occupation-related variables in CFPS do not refer to workplaces but occupation categories. The CFPS employs ISCO-88 (International Standard Classification of Occupations) to categorise occupations in nine different types, as is shown in the following tables.

Table 15.3.5 shows that students with mothers doing skilled agricultural and fishery work and mothers doing elementary occupations are somewhat under-represented in HE participation, with percentages below 20%. The other groups, especially the ones with mothers working as technicians and associated professionals, and professionals and clerks, are advantaged.

Table 15.3.5 Percentages of HE participation and non-HE participation by mother's occupation

	Total (N)	HE (%)	Non-HE (%)	Odds ratio
Skilled agricultural and fishery workers	1287	19	81	0.4
Elementary occupations	54	19	82	
Legislators, senior officials and managers	41	32	68	
Professionals	68	50	50	
Technicians and associated professionals	28	68	32	
Clerks	30	50	50	
Service workers and shop and market sales workers	216	35	65	
Craft etc. trade workers	168	29	71	
Plant and machine operators and assemblers	69	26	74	
Mother's occupation missing	529	25	75	
Total	2,490	24.5	75.5	

The disadvantaged group that includes students whose mother works as skilled agricultural and fishery workers and works in elementary occupations has only 0.4 of the chance that the groups showing advantages in HE participation rates have of accessing HE.

As for students' fathers' occupations (Table 15.3.6), those whose fathers work as skilled agricultural and fishery workers or work in elementary occupations are under-represented in HE participation compared with the average percentage of 24.5%. Again, those with fathers working as professionals, clerks, legislators, senior officials and managers, and technicians and associated professionals are to some extent over-represented. The other groups – service workers, shop and market sales workers, craft trade workers, plant and machine operators and assemblers and elementary occupations – are slightly over-represented in HE participation but do not have remarkable advantages.

Table 15.3.6 Percentages of HE participation and non-HE participation by father's occupation

	Total (N)	HE (%)	Non-HE (%)	Odds ratio
Skilled agricultural and fishery workers	1056	19	81	0.4
Elementary occupations	46	20	80	
Legislators, senior officials and managers	70	44	56	
Professionals	67	51	49	
Technicians and associated professionals	35	43	57	
Clerks	15	47	53	
Service workers and shop and market sales workers	159	27	73	
Craft etc. trade workers	400	29	71	
Plant and machine operators and assemblers	279	27	73	
Father's occupation missing	363	22	78	
Total	2,490	24.5	75.5	

The disadvantaged groups of skilled agricultural and fishery workers and elementary occupations are also 0.4 times as likely to participate in HE as the other groups.

15.3.6 Family Income and Expenditure on Education

Both family income and expenditure on education are numerical variables, so their averages can be calculated and compared. It is clear from Table 15.3.7 that the average family income in the HE participant group (26,743 yuan a year) is much higher (nearly 1.5 times) than that in the non-HE participant group (17,849 yuan).

Table 15.3.7 Average income and expenditure on education of families of HE participants and non-HE participants

	HE	Non-HE
Average family income (yuan a year)	26,743.1	17,849.4
Average expenditure on education (yuan a year)	2,413.3	1,236.6

The gap in annual expenditure on education between the two groups is even more pronounced. The HE participating group spends an average of 2,413 yuan a year, while the non-HE participating group spends only 1,236 yuan. The former group spends around twice as much as the latter.

15.3.7 Communication Language at Home

Instead of self-evaluated language ability, the CFPS asks interviewees about the language used at home for daily communication with their family, which is somewhat more objective although it is still not a good indicator of language ability. Table 15.3.8 shows the clustering of HE participation by various language types. Students who communicate with their families in Mandarin or minority ethnicity languages are over-represented in HE admissions with respective participation percentages of 33% and 27%. Students who use a Chinese dialect or another language at home have lower than average participation rates. Additionally, those whose home language information is missing are highly disadvantaged, with only 3% being admitted to HE.

Table 15.3.8 Percentages of HE participation and non-HE participation by the language used to communicate with the family

	Total (N)	HE (%)	Non-HE (%)	Odds ratio
Mandarin Chinese	426	33	67	1.6
A minority ethnicity language	619	27	73	
A Chinese dialect	1,350	22	78	
Other language	62	13	87	
Home language missing	33	3	97	
Total	2,490	24.5	75.5	

Students who speak Mandarin or minority ethnic languages at home are 1.6 times as likely to enter HE than students who speak a Chinese dialect or other languages.

It is not difficult to understand why students who speak Mandarin Chinese are advantaged. Mandarin is the official language for Chinese education and examinations. However, it is a little surprising that minority ethnicity language students are advantaged because if students speak a minority ethnicity language at home they might have minority ethnicity status. This small over-representation in HE participation of students who speak a minority ethnicity language might be evidence that not all minority ethnicity students are disadvantaged, and so this indicator risks producing false positives.

School-level indicators

15.3.8 School Type

As Table 15.3.9 shows, students from key schools are much more likely to be admitted to universities, and they have an HE participation percentage of 64%. All the other school types are under-represented. Students from ordinary schools and those from private schools are slightly under-represented (24% and 23%), while those lacking school-type information are notably under-represented (9%).

Table 15.3.9 Percentages of HE participation and non-HE participation by school type

	Total (N)	HE (%)	Non-HE (%)	Odds ratio
Key school	168	64	36	5.7
Ordinary school	1,945	24	76	
Private school	60	23	77	
School for migrant workers' children	1	-	-	
School type missing	316	9	91	
Total	2,490	24.5	75.5	

Although the sample size is too small for any serious comparison, the one interviewee who attended a school for migrant workers' children did not enter HE. Furthermore, previous studies (see Chapter 9) have shown that students segregated in schools for migrant children are always more disadvantaged. This issue deserves further study.

To summarise, there are large gaps in HE participation between different school types. Only students studying in selective schools are advantaged in the competition for places in HE. They are 5.7 times as likely to enter HE than the other groups.

15.3.9 Class Type

There is also a class type variable in the CFPS. This variable can take one of four values: key class, ordinary class, the school attended does not distinguish between class types and class type is missing. Again, like key schools, key classes are classes that are selective and require extra certifications of students' academic abilities. Ordinary classes do not require this. However, some schools have started to eliminate the distinction between these two class types in response to government policy.

As Table 15.3.10 shows, students in key classes are much more likely to attend HE than other students, but those in ordinary classes are also a little advantaged. Students in schools that do not distinguish between ordinary and key classes are a little under-represented. The advantaged group - students in key classes or ordinary class- is 2 times as likely to attend HE than their counterparts whose schools do not distinguish between classes. Cases in the CFPS data missing class-type information are highly disadvantaged in HE participation.

Table 15.3.10 Percentages of HE participation and non-HE participation by class type

	Total (N)	HE (%)	Non-HE (%)	Odds ratio
Key class	262	50	50	2.0
Ordinary class	530	29	71	
The school does not distinguish between ordinary and key classes	1,376	21	79	
Class type missing	322	10	90	
Total	2,490	24.5	75.5	

The reason for the under-representation of the third group might be that schools that do not distinguish between ordinary and key classes are likely to be ordinary schools or less-elite schools. As the last subsection has shown, these schools have lower percentages of HE participation.

15.3.10 Chinese/Mathematics Scores and Cognitive Scores

Table 15.3.11 compares the average standardised scores in Chinese tests, mathematics tests and cognitive tests of HE participants and non-HE participants. Because the ages of the cases in the CFPS data are different, the raw scores in these tests were standardised by age to make them more comparable.

Table 15.3.11 Average standardised scores in Chinese/mathematics tests and cognitive tests of HE participants and non-HE participants

	HE	Non-HE
Chinese score (average)	0.54	-2.00
Mathematics score (average)	0.57	-0.22
Word cognitive test score (average)	31.38	26.42
Mathematics cognitive test score (average)	20.53	16.79

As Table 15.3.11 shows, the HE participants generally had higher average scores in all the tests examined and especially in Chinese tests and word cognitive tests. This is not surprising as lower educational achievement could be one of the main barriers to HE entry. However, it is important to recognise that gaps in prior educational achievement are already stratified by SES characteristics.

15.3.11 Communication Language with Classmates

Table 15.3.12 shows the clustering of HE participation by the language that students use for daily communication in class or school. Students who speak Mandarin with their classmates are somewhat over-represented in HE participation, and there is a reasonable proportion of students who speak a Chinese dialect. However, students who reported that they speak a minority ethnicity language in class are highly under-represented in HE, with a percentage of 5.1%. Those who did not provide this information are also under-represented, with a HE participation percentage of 6.6%.

Students who speak Mandarin or Chinese dialects at school are 3.3 times as likely to attend HE than those who speal minority ethnic languages or other languages. If the groups are compared singly, the gaps are larger. The most advantaged group of Mandarin speakers, for instance, are 1.4 times as likely to participate in HE than the dialect group and eight times as likely than the minority ethnicity language group.

Table 15.3.12 Percentages of HE participation and non-HE participation by communication language with classmates

	Total (N)	HE (%)	Non-HE (%)	Odds ratio
Mandarin Chinese	855	31	69	3.3
Chinese dialect	1,329	25	75	
Minority ethnicity language	39	5	95	
Other languages	40	15	85	
Language at school missing	227	7	93	
Total	2,490	24.5	75.5	

These results differ from those shown in Table 15.3.8, in which minority ethnicity language speakers are not under-represented but even a little over-represented. This might be because the schools where students communicate with each other in minority ethnicity languages are likely to be minority ethnicity segregated schools, which are located in places inhabited by minority ethnicity people. These places are usually less developed and lack education and economic resources.

Table 15.3.13 Percentages of HE participation and non-HE participation by communication language with classmates and communication language with family

	HE	Non-HE	Total
Other language & Chinese dialect	100	0	1
Chinese dialect & Other language	40	60	5
Mandarin & Mandarin	35	65	370
Chinese dialect & Minority language	35	65	325
Mandarin & Minority language	35	65	130
Chinese dialect & Mandarin	33	67	12
Mandarin & Chinese dialect	25	75	340
Chinese dialect & Chinese dialect	21	79	987
Mandarin & Other language	13	87	15
Other language & Other language	13	87	39
Missing either or both	7	93	227
Minority language & Minority language	6	94	35
Minority language & Other language	0	100	2
Minority language & Mandarin	0	100	2

Total	24.5	75.5	2490

Table 15.3.13 combines the language used at school and the language used at home (e.g. “Other language & Chinese dialect” means that students use other languages to communicate at school and use Chinese dialects to communicate at home). Ignoring the special values with few cases, it is clear that students speaking minority languages at home, no matter they speak Mandarin or Chinese dialects at school, are not necessarily disadvantaged in HE participation. They show over-represented HE participation percentages.

However, students who communicate with their classmates or teachers in minority ethnic languages, whatever the languages they use to communicate with their family, are greatly disadvantaged in HE. These results echo those presented in Table 15.3.8 and Table 15.3.12. All of these remind us again that minority ethnicity is not necessarily disadvantaged in HE, as those living in urban cities and enjoying good-quality education may be in the privileged group.

15.4 High School Participation

Before drawing any conclusions regarding HE admissions based on the above results, it is important to consider the previous step. As was explained in Chapter 14, the transition to HE is not the first selection in students' education careers. In China high school education is not compulsory and students need to take an examination to compete for a place in a high school. Only students who have completed high school or equivalent can apply for HE.

Table 15.4.1 Percentages of HE participation and non-HE participation by high school attendance

		HE	Non-HE	Total (N)
Did not attend high school		0	100	550
Attended	an ordinary academic high school	44	56	1,104
	an adult high school	0	100	1
	an ordinary specialised high school	3	97	160
	a specialised adult high school	0	100	3

	a vocational high school	4	96	117
	a technical high school	2	98	53
Missing	High school education information missing	23	77	502
Total		24.5	75.5	2,490

As Table 15.4.1 shows, it is almost impossible for students who did not go to high school or equivalent to enter HE. None of the 550 cases who failed to attend high school were admitted to HE.

Furthermore, only students who went to ordinary academic high schools, which teach academic knowledge, were proportionately advantaged in HE participation. Around 44% of them were admitted to university. In contrast, students who attended other types of high schools, such as adult high schools, ordinary specialised high schools, specialised adult high schools, vocational high schools and technical high schools, were very unlikely to enter HE. That is, continuation to HE is normally associated with previous continuation to high school education, or more specifically to ordinary academic high school education. Therefore, in order to know who has been enrolled in HE and who has been excluded, it is important to explore previous continuation to high school.

Table 15.4.2 shows the clustering of high school participation. 1,104 students participated in high school education, accounting for 44.3% of the sample, while the non-high-school group is a little larger, with 1,386 cases accounting for 55.7%. In order to make less confusion, the high school (HS) group in this study only includes people who attended ordinary academic high schools, instead of also including other categories.

Table 15.4.2 Percentages of HS participants and non-HS participants in the CFPS total sample

	HS	Non-HS	Total (N)
Percentage	44.3	55.7	2,490

The result presented in Table 15.4.2 indicates that nearly half of students could not attend high schools (and this half could hardly attend HE). The disparities in high school (HS) participation between different groups are compared, where the disadvantaged groups and advantaged groups are nearly the same as those when exploring HE participation. The results of detailed comparisons between indicator groups are demonstrated in Appendix 12.

15.5 Discussion and Conclusion

This chapter has compared HE participation by different indicators in the CFPS dataset. First, three officially used CA indicators, namely hukou status, minority ethnicity and hukou province, reveal disparities in HE participation. However, this does not necessarily mean that they are all appropriate to be used as CA indicators. For instance, the advantage in HE attendance of the group who speak minority ethnicity languages in their daily communication with their families means that students with minority ethnicity do not all need extra support in the transition to HE.

Second, some other indicators also reveal gaps. Females are more likely to participate in HE than males. Students with better educated parents, with parents who have better jobs and with parents who are party members, which are highly likely to be correlated with each other and with two other crucial indicators, namely family income and expenditure on education, show higher percentages of participation in HE. Students studying in key schools or key classes and students with more satisfactory test scores and cognitive scores also display advantages in HE participation. However, school type, class type and test scores may already be stratified by personal characteristics such as SES, as discussed in Chapter 9.

In addition, as few students who did not go to high school continue into HE, it is important to examine this earlier stage of education. Exploring the clustering of high school education participation has revealed disparities between groups with different characteristics, and the groups that are disadvantaged in high school education participation are nearly the same as those disadvantaged in HE participation. These results show how disadvantage can be cumulative in the education process.

However, neither the disparities in HE participation nor those in HS participation are adequate reasons to support these indicators being used in CA policies. There is a need to explore more the underlying relationships between indicators and outcomes, and between the indicators themselves. Therefore, the following chapters explore the relationships between potential indicators of disadvantage and try to predict HE participation and HS participation according to these indicators.

Chapter 16 Relationships Between Indicators in the CFPS data

As has been mentioned in the last few chapters, missing data is unlikely to be missing by chance but is more likely to indicate disadvantaged groups. Therefore, it is necessary to explore these potentially disadvantaged groups when discussing HE equity. The missing cases in the CFPS data were briefly explained in Chapter 7. Rather than missing cases, this chapter focuses on the missing values in the CFPS dataset. Furthermore, disadvantaged students may be doubly or triply disadvantaged by their characteristics, so the relationships between the key indicators are also examined in this chapter.

16.1 Relationships Between Currently Employed Indicators and Others

This section explores the relationships between officially recognised indicators and other potential indicators to see whether there are underlying doubly disadvantaged groups.

16.1.1 Hukou Status

Table 16.1.1 presents the relationships between different hukou status and other potential indicators of disadvantage according to the descriptive results in the last chapter. In general, rural hukou students tend to be more closely associated with other potential indicators of disadvantage than urban hukou students. For instance, they are much more likely to live in rural areas (75%) and to belong to a minority ethnicity (10%). As for their parents, rural hukou students have higher percentages of less-advantaged parents, such as less-educated parents, non-party-member parents and parents with lower-class occupations. Furthermore, more rural hukou students communicate with their family members in Chinese dialects or a language other than Mandarin or a minority ethnicity language. In addition, they are more likely to study in less-selective ordinary or private schools. These close links imply that rural hukou students are more disadvantaged than urban hukou students to a certain extent.

Table 16.1.1 Percentages of students in the CFPS data with potential indicators of disadvantage by hukou status

	Urban hukou	Rural hukou	No registration or hukou missing
Rural living area	10	75	90
Minority ethnicity	6	12	15
Lower-participation area	23	28	30
Male	48	52	45
Basically educated parents	39	88	85
Non-party-member parents	49	76	75
Less-advantaged mother's occupation	10	64	65
Less-advantaged father's occupation	5	52	45
Chinese dialect or other language used at home	41	62	75
Ordinary or private school	74	82	70
Non-distinguished classes	51	56	65
Minority ethnicity language or other language used at school	2	3	0
Total (N)	484	1,986	20

Besides officially recognised rural hukou status, some students without legal hukou registration can be disadvantaged and have other characteristics of disadvantage. It seems that less-favourable hukou status is closely associated with less-favourable education opportunities, occupations and political status, or maybe vice versa. Students inherit characteristics of disadvantage, such as hukou status and family environment, from their parents.

However, hukou status is not related to noteworthy differences in sex and province living area, which might mean that being male and living in a lower-participation province are not necessarily associated with hukou disadvantage. In addition, there are no remarkable gaps between rural hukou and urban hukou students by school-level indicators, including school type, class type and school language.

If rural hukou or non-registered hukou were used as the sole indicator in CA policies in China it would be problematic as there would be risks of false-positive inclusions and false-negative exclusions, as was discussed in Chapter 12.

16.1.2 Living Area

Table 16.1.2 shows the relationship between students' living areas and other potential indicators of disadvantage. Nearly all the students living in rural areas have rural hukou, and most of them come from families with parents who only completed basic education (88%), who are not party members (77%) and who work in less-privileged occupations (70% and 57%). Students living in urban areas are proportionately less likely to have rural hukou (53%), parents who only obtained compulsory education qualifications (62%), parents who work in less advantaged occupations (27% and 18%) and parents who are not party members (61%). Furthermore, students living in rural areas tend to use minority ethnicity languages at home. However, the gaps between living areas are smaller than those between hukou statuses. In addition, there are few differences between students living in rural and urban areas by lower-participation areas and sex.

Table 16.1.2 Percentages of students in the CFPS data with potential indicators of disadvantage by living area

	Urban living area	Rural living area
Rural hukou	53	96
Minority ethnicity	8	13
Lower-participation area	28	27
Male	51	51
Basically educated parents	62	88
Non-party-member parents	61	77
Less-advantaged mother's occupation	27	70
Less-advantaged father's occupation	18	57
Chinese dialect or other language used at home	52	62
Ordinary or private school	78	82
Non-distinguished classes	51	58
Minority ethnicity language or other language used at school	3	4
Total (N)	929	1,561

Although they are narrower, some gaps can be found between the two groups by school-level indicators. Slightly higher percentages of students living in rural areas attend ordinary and private schools (82%) and schools that do not distinguish between classes (58%), and use minority ethnicity languages at school (4%) than students that live in urban areas (78%, 51% and 3% respectively).

It seems clear that living area is not appropriate to be used as a contextualised indicator even though it reveals differences in school-level indicators that were not clearly revealed by hukou status. Moreover, living area might be even more vulnerable to an ecological fallacy than hukou status. However, it might be feasible to use hukou status and living area together as a combined indicator. This possibility will be discussed further below.

16.1.3 Minority Ethnicity

Table 16.1.3 shows relationships between ethnicity and other indicators of disadvantage. Not surprisingly, minority ethnicities have closer relationships with rural hukou, rural living areas, less-privileged parents and minority ethnicity languages used at school than their Han counterparts. However, the disparities between the two ethnicity groups are smaller than those between hukou statuses and between living area groups. This seems to imply that ethnicity status, which is like a biographical characteristic, is not necessarily associated with other socio-economic characteristics.

Table 16.1.3 Percentages of students in the CFPS data with potential indicators of disadvantage by ethnicity

	Minority ethnicity	Han ethnicity	Ethnicity missing
Rural hukou	89	79	75
Rural living area	75	61	63
Lower-participation area	74	22	50
Male	53	51	63
Basically educated parents	86	78	88
Non-party-member parents	76	70	100
Less-advantaged mother's occupation	77	51	63
Less-advantaged father's occupation	64	40	38
Chinese dialect or other language used at home	43	60	63
Ordinary or private school	78	81	75
Non-distinguished classes	55	55	63
Minority ethnicity language or other language used at school	13	2	0
Total (N)	269	2,213	8

Although the differences between hukou statuses and living areas are small, a much higher percentage of minority ethnicity students come from areas with lower HE participation (74%). Only 21.6% of their Han ethnicity counterparts come from these areas. A possible reason is that lower-participation areas are likely to be remote less-developed provinces densely populated by minority ethnicity people.

Few disparities can be seen in sex groups, school types and class types.

16.1.4 Province Area

Another official indicator is province, which is re-categorised area-level information. As Table 16.1.4 shows, there are few differences in the percentages of rural living areas, male students and less advantaged students between low-participation areas and high-participation areas. However, low-participation areas have a noteworthy close link with minority ethnicity groups, which echoes the result in Table 16.1.3, and use of less-advantaged languages. Furthermore, although the disparities are not great, lower HE participation areas seem slightly more disadvantaged because they have higher proportions of rural hukou students, less-educated parents, parents with poorly paid jobs and ordinary and private schools.

Table 16.1.4 Percentage of students in the CFPS data with potential indicators of disadvantage by hukou province

	Low-participation areas	High-participation areas	Hukou province missing
Rural Hukou	83	79	44
Rural living area	62	63	56
Minority ethnicity	30	4	22
Male	54	50	44
Basically educated parents	82	77	78
Non-party-member parents	76	69	78
Less-advantaged mother's occupation	60	52	56
Less-advantaged father's occupation	47	41	33
Chinese dialect or other language used at home	89	55	22
Ordinary or private school	85	79	67
Non-distinguished classes	54	56	56

Minority ethnicity language or other language used at school	10	1	0
Total (N)	1,140	1,341	9

The results above warn us of the risk of ecological fallacy resulting from the sole use of province as a CA indicator. If province is used as a criterion then there is a high possibility that advantaged students in less-advantaged provinces can get boosted chances of HE.

16.2 Relationships Among Other Potential Indicators

This section investigates relationships among potentially helpful CA indicators. These indicators are discussed in three groups, as in Chapter 15.

Individual-level indicators

16.2.1 Sex

The results of the relationships between sex and other indicators are shown in Table 16.2.1. It is fair to say that there are no noteworthy differences between sex groups for all these indicators.

Table 16.2.1 Percentages of students in the CFPS data with potential indicators of disadvantage by sex

	Female	Male
Rural hukou	79	81
Rural living area	63	63
Minority ethnicity	11	12
Lower-participation area	26	29
Basically educated parents	78	79
Non-party-member parents	71	71
Less-advantaged mother's occupation	52	55
Less-advantaged father's occupation	41	44
Chinese dialect or other language used at home	60	56
Ordinary or private school	81	80
Non-distinguished classes	55	56
Minority ethnicity language or other language used at school	3	4
Total (N)	1,215	1,275

The two sex groups have similar patterns of relationships with other indicators. Therefore, sex does not identify actual disadvantage.

Family-level Indicators

16.2.2 Parental Education Qualifications

According to the descriptive results in Chapter 15, family-level indicators seem to be the most promising ones for implementation in CA policies. As Table 16.2.2 shows, students with parents who only completed compulsory education, or even less, are nearly twice as likely to have rural hukou and live in rural areas. Maybe worse, they are 2.5 times more likely to have parents who do not work in well-paid or decent occupations. Partly because of their education qualifications, these less-educated parents are also less likely to be elite party members. Furthermore, their children are more likely to speak less-favourable languages.

Table 16.2.2 Percentage of students in the CFPS data with potential indicators of disadvantage by parental education qualifications

	Basically educated parents	At least one parent has a post-compulsory or HE qualification	Parental education missing
Rural hukou	90	44	80
Rural living area	71	33	70
Minority ethnicity	12	7	10
Lower-participation area	29	23	40
Male	52	49	70
Non-party-member parents	80	40	30
Less-advantaged mother's occupation	62	25	10
Less-advantaged father's occupation	49	19	10
Chinese dialect or other language used at home	61	47	80
Ordinary or private school	82	76	80
Non-distinguished classes	56	51	60
Minority ethnicity language or other language used at school	4	2	20
Total (N)	1,953	527	10

As for school indicators, the disparities are smaller. Students with basically educated parents are slightly over-represented in ordinary and private schools and less-advantaged class types than their counterparts with one or more parents completing post-secondary education or HE.

Due to the close association with other potential indicators of disadvantage, parental educational qualifications, and specifically parents who only completed compulsory education or less, would seem to be the most suitable CA indicator found so far.

16.2.3 Parental Political Status

Table 16.2.3 shows the relationships between students' parents' political status and other potential indicators of disadvantage. Students whose parents do not belong to the party are more likely to have rural hukou and live in rural areas. As previously mentioned, party membership is somewhat a sign of being elite, which might mean well-educated parents having professional jobs. To a certain extent the former is a necessary condition for the latter. This elite preference for party membership means that parental political status is a socio-economic indicator.

However, no differences in biographical characteristics such as ethnicity and sex, language-related characteristics, area-based distinctions and school-level characteristics are notable. Both non-party-member parents and party-member parents show similar patterns in the relationships with these characteristics.

Despite some links with indicators of disadvantage and under-representation in HE participation, non-party-member parents might not be an appropriate contextualised indicator. The first reason is that it is true that only if individuals are excellent and elite can they be able to become party members, but excellence is not a sufficient condition for membership of the party. Applying to be a party member is voluntary, so not all advantaged students have party-member parents. Therefore, if non-party-member parents were used as a contextualised indicator there would be a danger of many false positives. Second, the non-party-member group is large, which not only threatens accuracy but also makes it difficult to provide additional support.

Table 16.2.3 Percentages of students in the CFPS data with potential indicators of disadvantage by parental political status

	Non-party-member parents	Party-member parent(s)	Political status missing
Rural hukou	86	63	72
Rural living area	68	50	53
Minority ethnicity	12	9	10
Lower-participation area	29	22	28
Male	51	51	52
Basically educated parents	88	51	71
Less-advantaged mother's occupation	61	42	18
Less-advantaged father's occupation	48	32	14
A Chinese dialect or other language used at home	60	52	61
Ordinary or private school	81	78	83
Non-distinguished classes	55	55	61
Minority ethnicity language or other language used at school	4	3	2
Total (N)	1,767	587	136

16.2.4 Parental Occupation

Table 16.2.4 and Table 16.2.5 compare the relationships between fathers' and mothers' occupation types and other indicators of disadvantage. Students with a father or a mother working in a less-advantaged occupation are highly likely to have rural hukou and live in a rural area. A higher percentage of them are also of minority ethnicity, which might indicate an underlying inequality in ethnicity and employment. Furthermore, parents working in less favoured jobs are more likely to have only completed basic education. They are also more likely to not be party members.

Table 16.2.4 Percentages of students in the CFPS data with potential indicators of disadvantage by father's occupation

	Less-advantaged father's occupation	Other father's occupation	Father's occupation missing
Rural hukou	97	67	68
Rural living area	85	46	49
Minority ethnicity	17	6	9
Lower-participation area	30	24	28
Male	53	50	51
Basically educated parents	91	68	74
Non-party-member parents	81	65	60
Less-advantaged mother's occupation	90	30	22
Chinese dialect or other language used at home	59	56	61
Ordinary or private school	82	78	84
Non-distinguished classes	55	55	57
Minority ethnicity language or other language used at school	3	3	1
Total (N)	1,102	1,025	363

Table 16.2.5 Percentages of students in the CFPS data with potential indicators of disadvantage by mother's occupation

	Less-advantaged mother's occupation	Other mother's occupation	Mother's occupation missing
Rural hukou	95	57	67
Rural living area	81	37	46
Minority ethnicity	16	6	6
Lower-participation area	30	23	25
Male	53	48	53
Basically educated parents	90	60	71
Non-party-member parents	7	57	65
Less-advantaged father's occupation	70	8	12
Chinese dialect or other language used at home	59	54	60
Ordinary or private school	81	76	84
Non-distinguished classes	57	55	53
Minority ethnicity language or other language used at school	4	2	4
Total (N)	1,341	620	529

Nevertheless, neither less-advantaged fathers' occupations nor less-advantaged mothers' occupations show strong associations with school-level indicators or language-related indicators compared with the counterpart groups of other fathers' occupations and other mothers' occupations. There are very high percentages of all the groups in disadvantaged schools and classes. These similarities might be attributable to the rough binary categorisation. If the groups of other fathers'/mothers' occupations were divided into more detailed classifications there might be more visible gaps in school types and class types.

16.2.5 Family Communication Language

Table 16.2.6 shows the relationships between the daily communication language at home and other indicators. Students who speak a Chinese dialect or another language with their family members, who in the last chapter were found to be under-represented in HE participation, are not necessarily associated with other potential indicators of disadvantage. However, there are comparatively high percentages of them in the rural hukou group, the rural living area group, lower-participation areas, the group with less socio-economically advantaged parents and in less advantaged schools and classes.

Such higher percentages are sometimes also present in the other two groups. For instance, although 82% of students who use a Chinese dialect or another language at home have parents who only completed compulsory education or less, the corresponding percentages for the first two groups are also high, at 61% and 81%. In addition, 67% of students in the dialect group live in rural areas, but the percentages are 40% and 70% for the other two groups. These inexplicit differences might mean that utilising family communication languages as a CA indicator would be vague and inappropriate.

Furthermore, as it is unverified and like other self-reported indicators would be hard to verify, the communication language at home could not seriously be used as a CA indicator.

Table 16.2.6 Percentages of students in the CFPS data with potential indicators of disadvantage by communication language at home

	Mandarin used at home	Minority ethnicity language used at home	A Chinese dialect or another language used at home
Rural hukou	51	86	85
Rural living area	40	69	67
Minority ethnicity	11	18	8
Lower-participation area	3	34	32
Male	49	56	50
Basically educated parents	61	81	82
Non-party-member parents	58	75	73
Less-advantaged mother's occupation	31	67	55
Less-advantaged father's occupation	26	52	43
Ordinary or private school	77	59	91
Non-distinguished classes	54	30	67
Minority ethnicity language or another language used at school	1	6	3
Total (N)	426	619	1,445

School-level indicators

16.2.6 School Type

Despite criticisms of them in the last chapters, school-level indicators revealed some disparities in HE participation. Therefore, they deserve to be analysed more deeply. Table 16.2.7 shows that students studying in ordinary and private schools are more likely to have rural hukou, live in a rural area, belong to a minority ethnicity group, live in a lower-participation area, have less-advantaged parents and be disadvantaged language users. However, students in non-key schools do not seem to be the most disadvantaged, as their counterparts with missing school-type information have closer links with other indicators of disadvantage. Again, this reminds us that missing data should never be simply ignored when considering equity. There is no significant sex gap between the different school types.

Table 16.2.7 Percentages of students in the CFPS data with potential indicators of disadvantage by school type

	Key schools	Ordinary or private schools	Schools for migrant workers' children	School type missing
Rural hukou	52	81	-	84
Rural living area	36	64	-	68
Minority ethnicity	5	11	-	17
Lower-participation area	16	29	-	25
Male	52	51	-	54
Basically educated parents	52	80	-	85
Non-party-member parents	52	72	-	78
Less-advantaged mother's occupation	30	54	-	63
Less-advantaged father's occupation	18	43	-	51
Chinese dialect or other language used at home	43	65	-	19
Non-distinguished classes	36	65	-	1
Minority ethnicity language or another language used at school	1	4	-	1
Total (N)	168	1,945	1	316

Despite its somewhat close association with other indicators of disadvantage, school type would not be a good choice for CA policies. First, non-key schools are the majority in China. More importantly, identification of disadvantages by school type is vulnerable to the "creamy layer" (Frisancho & Krishna, 2016). Furthermore, the official classification of key/non-key schools is gradually being abolished. This might create difficulties if school type is used as a CA indicator.

16.2.7 Class type

Table 16.2.8 further examines how class type correlates with other indicators. The disadvantaged group of students studying in schools that do not distinguish between key classes and non-key classes are only slightly disadvantaged. They have similar distributions to the other two groups in almost all the other potential indicators of disadvantage.

On the other hand, more attention might be paid to the group missing class-type information. This group is not likely to claim they study in ordinary or private schools, and there is a similar result to that shown in Table 16.2.7, as students who miss school information are less likely to claim they study in a non-distinguished class. This might be because students missing school type information are highly likely to miss class type information and vice versa.

Table 16.2.8 Percentages of students in the CFPS with potential indicators of disadvantage by class type

	Non-key class	Key class	Classes not distinguished	Class type missing
Rural hukou	79	69	81	84
Rural living area	57	51	66	68
Minority ethnicity	9	8	11	17
Lower-participation area	31	26	27	24
Male	54	42	51	53
Basically educated parents	74	72	80	85
Non-party-member parents	69	68	71	78
Less-advantaged mother's occupation	48	50	55	62
Less-advantaged father's occupation	40	37	42	51
A Chinese dialect or another language used at home	55	51	70	19
Ordinary or private school	88	83	95	3
A minority ethnicity language or another language used at school	3	4	4	1
Total (N)	530	262	1,376	322

Class type, like school type, does not have the potential to be a high-quality CA indicator because of its inherent problems such as inaccuracy and less close links with other indicators of disadvantage.

16.2.8 School communication language

As for the daily communication language at school, as Table 16.2.9 shows, it would be untenable to say that the disadvantaged group that speaks a minority ethnicity language or another language at school have a considerable disadvantage. On the contrary, this group displays very similar patterns to the one that speaks a Chinese dialect at school for many socio-economic indicators such as rural hukou, parental political status and parental education qualifications.

Table 16.2.9 Percentages of students in CFPS with potential indicators of disadvantage by communication language at school

	Mandarin used at school	A Chinese dialect used at school	A minority ethnicity language or another language used at school	Language used at school missing
Rural hukou	63	88	86	91
Rural living area	45	72	68	72
Minority ethnicity	10	8	45	22
Lower-participation area	15	32	81	27
Male	50	51	58	46
Basically educated parents	68	83	87	89
Non-party-member parents	63	75	77	78
Less-advantaged mother's occupation	37	61	60	70
Less-advantaged father's occupation	29	48	46	58
Chinese dialect or another language used at home	42	75	53	25
Ordinary or private school	83	91	96	1
Non-distinguished classes	59	62	66	0
Total (N)	855	1,329	79	227

However, the group speaking a minority ethnicity language has closer links with minority ethnicity and lower HE participation areas. This possibly supports the results shown in Table 16.1.3, in which minority ethnicity students are more likely to live in distant remote areas, which often have lower HE participation rates.

16.3 The possibility of combined indicators

The analyses show that the officially used CA indicators are not general of high quality or utility. Therefore, this section considers some possible combined indicators.

First, if both the living area and hukou status are taken into consideration, some results are found. As Table 16.3.1 shows, students who have urban hukou and live in urban areas are the most advantaged in HE participation, and their rural hukou counterparts who live in rural areas are the most disadvantaged. However, students with rural hukou but live in urban areas are less likely to go into HE than those with urban hukou who live in rural areas. This finding supports revisiting the hukou categorisations of all the citizens in the country because they lead to unfairness in access and opportunities.

Table 16.3.1 Percentages of HE participation and non-HE participation by living area and Hukou status

	HE	Non-HE	Total (N)
Urban living area and urban hukou	43	57	435
Urban living area and rural hukou	24	76	492
Rural living area and rural hukou	19	81	1494
Rural living area and urban hukou	39	61	49
Data missing	10	90	20
Total (N)	24.5	75.5	2,490

The most advantaged group (urban living area and urban hukou) are nearly 2.3 times (urban living area and rural hukou), 3.2 times (rural living area and rural hukou) and 1.2 times (rural living area and urban hukou) more likely to access HE than the other groups. The HE participation gap indicated by the combined hukou and living area indicator is even more extensive than that for hukou status. This indicates that people living in rural areas and having rural hukou are the most disadvantaged. Although there is a risk of an ecological fallacy, this combined indicator might increase the accuracy of targeting CA students.

Second, minority ethnicity, an indicator that might risk mistargeting more than hukou status, could be combined with hukou province because some provinces with large minority ethnicity populations are disadvantaged in HE participation. The advantaged provinces are the municipalities and those in northeastern China, central China, eastern China and northern China, while the less advantaged provinces are in southern China, southwestern China and northwestern China.

Table 16.3.2 Percentages of HE participation and non-HE participation by hukou province and ethnicity

	HE	Non-HE	Total (N)
Advantaged province and Han ethnicity	29	71	1,285
Advantaged province and minority ethnicity	25	75	56
Less-advantaged province and minority ethnicity	13	87	219
Less-advantaged province and Han ethnicity	21	79	921
Data missing	33	67	9
Total (N)	24.5	75.5	2,490

The results for the combined minority ethnicity and hukou province indicator are presented in Table 16.3.2. The most advantaged group – advantaged province and Han ethnicity – are more than 2.7 times more likely to attend HE than the most disadvantaged group – less advantaged province and minority ethnicity. This gap is more prominent than that indicated by only using ethnicity or province.

Furthermore, as was discussed in Chapter 15, speaking a minority ethnicity language at school might indicate segregation of schools for minority ethnicities, which could be associated with disadvantage. If we accept this assumption, Han ethnicity students speaking Mandarin at school are 5.8 times more likely to enter HE than minority ethnicity students in minority-ethnicity segregated schools (Table 16.3.3).

Table 16.3.3 Percentages of HE participation and non-HE participation by language used at school and ethnicity

	HE	Non-HE	Total (N)
Mandarin language and minority ethnicity	21	79	185
Mandarin language and Han ethnicity	27	73	1,992
A minority language and minority ethnicity	6	94	36
A minority language and Han ethnicity	0	100	3
Another language and minority ethnicity	0	100	48
Another language and Han ethnicity	10	90	218
Data missing	25	75	8
Total (N)	24.5	75.5	2,490

Despite the remarkable HE participation gap revealed by the combined language at school and ethnicity indicator, we should remember that the measurement of language has problems of accuracy.

16.4 Conclusion

This chapter has investigated relationships between indicators that revealed under-representation in HE participation in Chapter 15. Some of them are closely linked with other indicators of disadvantage. For instance, rural hukou, living in a rural area, parents who only finished compulsory education, parents working in less-privileged occupations, non-party-member parents, and non-key schools are more strongly associated with other indicators of disadvantage than other indicators.

However, these close links do not necessarily mean these indicators should be used in CA policies. Rural hukou, minority ethnicity and lower-participation province areas, three officially employed indicators, are problematic. They are vulnerable to an ecological fallacy when they are used alone. Many other possible indicators that have been examined also have doubtful appropriateness.

Worse, other indicators such as parental occupations and living areas are more likely to be self-reported rather than recorded administrative information and so are unverified and lack reliability and trustworthiness. There are some possible solutions to this unreliability, such as making this kind of information officially required by the government. This solution might partly help verify job information, but it could be at the expense of cost, time, mobility, flexibility and maybe individual liberty.

It might be feasible to combine two or more indicators to identify disadvantaged groups, such as rural hukou and living in a rural area, minority ethnicity and living in a lower participation area, and minority ethnicity and speaking a minority ethnicity language at school (in minority ethnicity segregated schools). However, they are also not problem-free.

The next chapter uses regression analysis to examine how these indicators can predict individuals’ HE participation.

Chapter 17 Regression Analyses of the CFPS Data

In order to find suitable CA indicators, in this chapter regression analyses are conducted to examine the extent to which potential indicators can accurately predict participation in HE and, given that it is a crucial step towards transiting to HE, high school education.

17.1 Predicting HE Participation

This section presents regression results predicting HE participation for the indicators considered in the previous two chapters.

17.1.1 Percentage Variation

As was explained in Chapter 7, ten new data subsets each covering 1,220 cases (with 610 HE participants and 610 non-HE participants) were created to remedy the problem of unevenly distributed values of the outcome variable. The results shown in the following tables are averages of ten logistic regressions. The complete results for each estimate can be found in Appendix 13.

The Model 1 where individual-level indicators are put in the first block and Model 3 where school-level indicators are put in the first are not presented in the chapter but in Appendix 13. As explained in Chapter 13, because this study prefers more micro-level indicators, the results of the models that family-level indicators are put in at first are presented here.

As Table 17.1.1 shows, parents' education qualifications, occupations, political status, annual income and family expenditure on education improve the percentage of correct predictions to 63. Next, the individual-level indicators sex, hukou status, ethnicity and birth month only contribute one more percentage point. School-level indicators and neighbour indicators further improve the percentage to 72.

A comparison with the corresponding figures in the Model 1 (see Appendix 13) shows that family-level indicators increase the percentage of correct predictions more than individual-level indicators do. But bringing the school-level indicators forward in Model 3 improves the percentage by nearly 20 percentage points. The other individual-level, family-level and area-level indicators only contribute two more percentage points to the prediction. It seems that the school-level indicators are more important than the other kinds of indicators, or they are just a better summary of the other indicators.

Table 17.1.1 Average percentages of correct predictions of HE entry by ten regressions (Model 2; enter)

	HE entry	Percentages increase
Base	50.0	
Block 1 (parents' education qualifications, father's occupation, mother's occupation, parents' political status, communication language at home, annual family income, annual education expenditure)	62.9	12.9
Block 2 (sex, hukou, ethnicity, birth month)	63.9	1
Block 3 (school type, class type, communication language at school, word cognitive test score, maths cognitive test score)	71.3	7.4
Block 4 (living area, hukou province area)	72.0	0.7

N=1,220

However, attending a selective school or a selective class and scoring highly in cognitive tests might already be results of stratification of family or individual characteristics, but it does not mean that students from higher SES background must attend more prestigious schools or students from poor background must study in the non-selective schools. Therefore, despite impressive improvements in percentages, if school-level indicators such as non-selective schools are used in CA policies to identify disadvantaged students, new inequalities could be expected.

Adding another new variable: whether the students attended high schools can further improve 4.6 percentage points for the model in Model 4, which can be found in Appendix 13.

Table 17.1.2 puts the results of high school participation earlier in order to disclose how important this variable is. Attending high school increases the percentage of correct predictions by 23.2 percentage points on its own. The individual-level and family-level indicators then only increase the percentage a little. The school-level indicators make a contribution by increasing the percentage by 2.5 percentage points. Hukou province and living area make little differences in the model.

Table 17.1.2 Average percentages of correct predictions of HE entry by ten regressions (Model 5; enter)

	HE entry	Percentages increase
Base	50.0	
Block 1 (attended high school or not)	73.2	23.2
Block 2 (parents' education qualifications, father's occupation, mother's occupation, parents' political status, communication language at home, annual family income, annual education expenditure)	73.7	0.5
Block 3 (sex, hukou, ethnicity, birth month)	74.0	0.3
Block 4 (school type, class type, communication language at school, word cognitive test score, maths cognitive test score)	76.5	2.5
Block 5 (living area, hukou province area)	76.6	0.1

These results emphasise the importance of attending high school. When including this variable, all the other indicators become less important than they were in the first three models. Therefore, attending high school mediates the contributions of other indicators to a certain extent. However, this result might already be stratified by these indicators. This will be examined in the next section.

17.1.2 Coefficients

Table 17.1.3 lists the coefficients for each variable in the regression Model 2. Unsurprisingly, students with better-educated parents are more likely to continue into HE. Compared with students whose parents do skilled work in agriculture or the fishing industry, those with parents working in elementary occupations are disadvantaged in HE participation. Students whose mothers work in professional jobs and those who do not provide clear information on their mothers' occupations have an advantage in access to HE.

Furthermore, students with party-member parents are slightly advantaged. Compared with Mandarin users, students who speak a minority ethnicity language with their family members are advantaged in HE admission, while those who speak a Chinese dialect or another language at school do not show this advantage.

Annual family income and annual education expenditure seem irrelevant to HE participation, once other factors are controlled.

As for individual-level indicators, male students and Han ethnicity students tend to be less likely to participate in HE than their female and minority ethnicity counterparts. Minority ethnicity would seem to not be a suitable CA indicator because of the result of an advantage of minority ethnicity shown in Table 17.1.3.

Rural hukou students are less likely to go to university than urban hukou students, and so are students without hukou registration. There is no remarkable difference between the birth-month groups.

Key schools and key classes are related to HE participation, while students in private schools, those missing school-type information and those in schools that do not distinguish between key classes and non-key classes are less likely to be accepted by universities than those in ordinary schools and ordinary classes. Furthermore, higher scores in cognitive word and maths tests are associated with a higher probability of participating in HE. Again, these high scores could be stratified by family resources.

Students who speak minority ethnicity languages at schools are greatly disadvantaged. A possible reason might be the location of the school determined by the communication language, as was explained in Chapter 15.

Students living in urban areas are slightly more likely to participate in HE than their counterparts living in rural areas. In addition, compared with students whose hukou province is in the municipalities, those whose hukou provinces are in the northeastern China, central China, eastern China, northern China, southwestern China and northwestern China are advantaged in HE participation. Only students from the southern China are negligibly less likely to get a place in HE.

Table 17.1.3 Coefficients for the variables in Model 2 (enter)

	Coefficients
both parents have compulsory education qualifications	
at least one parent has post-compulsory education qualifications	1.5
at least one parent has HE degree	1.7
skilled agricultural & fishery workers (father)	
elementary occupations (father)	0.8
higher skilled/professional occupations (father)	0.9
occupation missing (father)	0.7
skilled agricultural & fishery workers (mother)	
elementary occupations (mother)	0.7
higher skilled/professional occupations (mother)	1.3
occupation missing (mother)	1.3
neither parent is a party member	
at least one parent is party member	1.1
parents' political status missing	1.0
Mandarin Chinese at home	
a Chinese dialect at home	0.8
minority ethnic language at home	1.2
another language at home	0.4
annual family income	1.0
Annual education expenditure	1.0
female	
male	0.7
minority ethnicity	
Han ethnicity	0.8
rural hukou	
urban hukou	1.2
no hukou registration	1.0
summer-born students	
autumn-born students	1.0
ordinary school	
key school	2.6
private school	0.5

missing school type information	0.2
not in a key class	
in a key class	1.9
the school does not distinguish key classes	0.9
class type information missing	2.3
Mandarin Chinese at school	
a Chinese dialect at school	1.3
minority ethnic language at school	0.4
another language at school	2.3
school communicative language information missing	1.1
cognitive math test score	1.1
Cognitive word test score	1.1
living in a rural area	
living in an urban area	1.2
municipalities	
northeastern China	1.6
central China	1.6
eastern China	1.6
northern China	1.7
southern China	0.9
southwestern China	1.3
northwestern China	1.8

Table 17.1.4 presents a very similar pattern to Table 17.1.3. The coefficients for each variable in Model 2 and Model 5 are almost the same, but students who attended high schools are 4.7 times more likely to participate in HE than those who did not. Furthermore, after controlling for high school attendance, students without hukou registration, students communicating with their family members in minority ethnicity languages and students missing school language information become slightly disadvantaged compared to their counterparts. Nevertheless, these changes in the coefficients are not massive. The indicator attendance at high school shows the most considerable difference, which indicates the importance of this indicator.

Table 17.1.4 Coefficients for the variables in Model 5 (enter)

	Coefficients
not attend high schools	
attended high schools	4.7

both parents have compulsory education qualifications	
at least one parent has post-compulsory education qualifications	1.4
at least one parent has HE degree	1.8
skilled agricultural & fishery workers (father)	
elementary occupations (father)	0.9
higher skilled/professional occupations (father)	0.9
occupation missing (father)	0.7
skilled agricultural & fishery workers (mother)	
elementary occupations (mother)	0.7
higher skilled/professional occupations (mother)	1.3
occupation missing (mother)	1.4
neither parent is a party member	
at least one parent is party member	1.1
parents' political status missing	1.0
Mandarin Chinese at home	
a Chinese dialect at home	0.8
minority ethnic language at home	0.9
another language at home	0.3
annual family income	1.0
annual education expenditure	1.0
female	
male	0.8
minority ethnicity	
Han ethnicity	0.8
rural hukou	
urban hukou	1.2
no hukou registration	0.9
summer-born students	
autumn-born students	1.0
ordinary school	
key school	2.6
private school	0.5
school type information missing	0.8
not in a key class	
in a key class	1.6
the school does not distinguish key classes	0.9
class type information missing	3.1
Mandarin Chinese at school	
a Chinese dialect at school	1.3
minority ethnic language at school	0.4
another language at school	2.9
school communicative language information missing	0.8
cognitive math test score	1.1

cognitive word test score	1.1
living in a rural area	
living in an urban area	1.1
municipalities	
northeastern China	1.5
central China	1.7
eastern China	1.4
northern China	1.5
southern China	0.9
southwestern China	1.2
northwestern China	1.5

Due to the importance of the transition to high school, the next section studies indicators that might be able to predict continuation after compulsory education.

17.2 Predicting High School Participation

As attending high school is crucial for HE participation, this section presents the results of regression analyses predicting high school participation using the cases in the CFPS data.

17.2.1 Percentage Variation

The results of predicting high school education are listed in the following table. Table 17.2.1 presents the results of Model 3 and the results of Models 1 and 2 could be found in Appendix 13.

As shown in Table 17.2.1, adding parents' education qualifications only can increase the percentage of correct predictions to 60.6, a little higher than just adding hukou does (see Table 4 in Appendix 13). Annual expenditure on education, father's job and family language also improve the model to a certain extent. The individual-level indicators in Block 2 only increase the percentage by less than one point. The school-level indicators in Block 3 and the area-level indicators in Block 4 display similar patterns as they did in other models.

Mother's occupation, parents' political status, annual family income, birth month, communication language at schools and living areas are excluded from the model. It somewhat indicates that these indicators are less relevant to the outcome variable.

Table 17.2.1 Percentages correctly predicting high school education (Model 3; forward)

		Percentage	Percentages increase
Base		55.7	
Block 1	Step 1 (parental education qualifications)	60.6	4.9
	Step 2 (education expenditure)	61.7	1.1
	Step 3 (father's occupation)	62.5	0.8
	Step 4 (communication language at home)	61.8	-0.7
Block 2	Step 1 (sex)	62.2	0.4
	Step 2 (ethnicity)	62.7	0.5
	Step 3 (hukou status)	62.7	0
Block 3	Step 1 (school type)	69.3	6.6
	Step 2 (cognitive word test score)	69.9	0.6
	Step 3 (cognitive math test score)	71.4	1.5
	Step 4 (class types)	72.0	0.6
Block 4	Step 1 (hukou province area)	72.9	0.9

N=2490

17.2.2 Coefficients

The coefficients for each variable in Model 2 are shown in Table 17.2.2. Not surprisingly, students with well-educated parents are more likely to attend high school after finishing compulsory education. However, those whose parents' education qualifications are missing in the CFPS data are the most likely to attend high school, but this result might only be the result of there being a small number of cases in this category.

Compared with students whose father work in agriculture or fishing, those with parents working in professional occupations are advantaged in high school education participation. Students with a father doing an elementary occupation are nearly half less likely to attend high school. The group with father's occupations unknown are also disadvantaged. The indicator of mother's occupations is excluded by forward method.

Table 17.2.2 Coefficients for the variables in Regression Model 2

	Forward
Annual education expenditure	1.0
both parents have compulsory education qualifications	
at least one parent has post-compulsory education qualifications	1.5
at least one parent has HE degree	1.5
parental education information missing	2.6
skilled agricultural & fishery workers (father)	
elementary occupations (father)	0.6
higher skilled/professional occupations (father)	1.2
occupation missing (father)	0.8
Mandarin at home	
Chinese dialects at home	1.0
Minority ethnic languages at home	3.5
another language at home	1.0
missing language at home	0.8
skilled agricultural & fishery workers (mother)	
elementary occupations (mother)	
higher skilled/professional occupations (mother)	
occupation missing (mother)	
neither parent is a party member	
at least one parent is party member	
parents' political status missing	
annual family income	
female	
male	0.8
minority ethnicity	
Han ethnicity	1.7
rural hukou	
urban hukou	1.4
no hukou registration	1.4
summer-born students	
autumn-born students	
ordinary school	
key schools	1.6
private schools	0.6
school types information missing	0.0
not in a key class	
in a key class	1.9
the school does not distinguish key classes	1.2
class type information missing	0.7
cognitive word test score	1.1
cognitive math test score	1.1
Mandarin Chinese at school	

a Chinese dialect at school	
minority ethnic language at school	
another language at school	
school communicative language information missing	
living in a rural area	
living in an urban area	
municipalities	
northeastern China	1.7
central China	1.2
eastern China	2.0
northern China	1.9
southern China	1.5
southwestern China	1.8
northwestern China	2.4
Missing hukou province information	3.1

Annual family income and annual educational expenditures seem less relevant to HE participation. So does parental political status. They are either excluded from the model or show no noteworthy differences in coefficients.

As for the daily language with family members, compared with students who speak Mandarin at home, their counterparts who speak a minority ethnicity language are much more likely to go to high school. There are no disparities in the other two groups that use a Chinese dialect or another language to communicate with family members.

Males, minority ethnicity and rural hukou students face more barriers in continuing to high school education. Rural hukouers are even less likely to go to high school than individuals missing hukou information. Birth month does not make a notable difference; as this is excluded by using the forward stepwise (conditional) method.

Students in key schools are 1.6 times more likely to continue to high school than those attending ordinary schools (non-selective schools). On the contrary, students in private schools are highly disadvantaged. Students in key classes are also more likely to transit to high school, and those in schools that do not distinguish between key classes and non-key classes are also a little advantaged, although they are under-represented in the descriptive analysis in Chapter 15.

The higher students cognitive math scores and word scores are, the more likely they are to go to high school. However, the differences are not very impressive.

Students in the municipalities seem to suffer fiercer competition in the transition to high school than students in any other province area. The variable of living in the rural areas or urban areas has been automatically excluded.

17.3 Conclusion

In this chapter regression analyses have been conducted to examine the extent to which potential indicators can accurately predict participation in HE and high school. Individual-level official indicators, including hukou status and minority ethnicity, contribute to the prediction to a certain extent and area-based indicators are less helpful as they improve the percentage of correct predictions by less than one point. On the other hand, family-level indicators tend to make more difference than individual-level ones. As they tend to be more micro-level, family-level indicators deserve consideration for use as contextualised indicators. When school-level indicators are added, they make the most difference to the result. However, regardless of whether they are school/class type or cognitive test score, school-level indicators are vulnerable to ecological fallacies. Therefore, they are not good choices for contextualised indicators.

More importantly, after adding the binary indicator of whether students went to high school into the model, few other predictors can make any noteworthy increases to the percentage predicted. This result again reminds us that if we want to deal with inequalities in HE participation, it is never enough to regard only NCEE candidates as targets needing support. Inequalities exist much earlier, at least at the point in which students have to compete in the High Schools Entrance Examination for an offer from an ordinary academic high school.

According to regression results predicting high school participation, hukou status is one of the most important individual-level indicators. Parents' education qualifications still perform well in increasing the percentage of correct predictions by the model, as does family annual expenditure on education. Other indicators such as ethnicity, sex, family communication language and parents' occupations only make small differences to the outcome variable. School-level indicators make notable improvements as they did in the previous regressions for HE, but they are problematic for the reasons explained above.

In both the regressions predicting HE participation and the ones predicting high school education, according to the coefficients the advantaged groups remain nearly the same. Urban hukou students, Mandarin speakers, those with better educated parents and those whose parents have better jobs remain among the privileged. Students from key schools or key classes are also more likely to continue their education, but again it is highly possible that school/class type is the result of stratification.

Compared with students living in municipalities, those living in other areas of China, such as the northeast, central and eastern China, are more likely to go to high school and then into HE. Including Chongqing, a municipality without outstanding HE admissions, in this group is a problem. Although Chongqing is a municipality according to its administrative level, it does not seem able to catch up with the other three municipalities – Beijing, Tianjin and Shanghai – in HE admissions. Including Chongqing might alleviate the potential remarkable advantages of the three other municipalities.

To sum up, one of the most important contributions of the CFPS data is that they remind us to pay attention to earlier educational selection. Disadvantages are accumulated much earlier than the NCEE portrays. Therefore, a supplementary survey was conducted in order to obtain a clearer picture of this issue. The next two chapters focus on analyses of the primary survey data.

Chapter 18 Descriptive Analysis of the Primary Survey Data

This chapter presents the results of the primary survey data collected with the purpose of providing a snapshot of disadvantage at school. The survey asked middle school students about their educational aspirations. Therefore, this chapter reveals the relationship between the educational aspirations of middle school students and their individual-level and family-level characteristics to see whether there are earlier cumulative disadvantages before the NCEE.

18.1 Results for the Overall Sample

18.1.1 The Educational Aspirations of the Overall Sample

802 middle school students were involved in the survey. Table 18.1.1 presents their aspirations for high school education and for HE. As going to vocational and skill high schools is very different to going to ordinary academic high schools, high school education in the survey results does not include vocational/skill schools. The survey only focused on academic high school education, including selective and ordinary (non-selective) high schools.

As Table 18.1.1 shows, in the survey most students (76%) reported that they wanted to continue their education after middle school and nearly half of them intended to go into HE. Therefore, generally these students had high aspirations to continue their education after the compulsory phase.

Table 18.1.1 Percentages of aspirations for high school education and for HE

	High School Education		Higher Education	
	Yes	No	Yes	No
Overall sample	76	24	49	51

As Table 18.1.2 shows, students who aspired to participate in high school education were much more likely to have aspirations for HE too (62%). On the contrary, few students who planned to go to vocational/skill high schools or who planned to leave school to work reported that they had plans for HE after the age of 18 (7%). This finding echoes similar results in Chapters 15, 16 and 17, that students who did not go to ordinary academic high schools were unlikely to go into HE.

Table 18.1.2 Percentages of aspirations for HE by aspirations for HS

	Aspiration for HE	No aspiration for HE	Odds ratio
Aspiration for high school education	62	38	21.7
No aspiration for high school education	7	93	

This was illustrated in the interviews. One student in Sangzhi stated:

"The government should expand enrolment into ordinary academic high schools, which is too low. Only half the middle school graduates have the chance to attend high school."

This statement was supported by other interviewees from Changsha.

The students were also asked to what extent they agreed that three factors – the family financial situation, school performance and health problems – can impact their plans to continue to study after completing middle school and after the age of 18, the theoretical age for completing high school. The range of scores is from 0 to 10, with 0 meaning not at all and 10 meaning very likely.

As Table 18.1.3 shows, the students tended to agree that their plans regarding continuing with post-compulsory education would be most affected by their school performance, with average scores of 8.1 and 7.9. These high scores partly mean that students at this stage, shortly before leaving middle school, knew what their prospects were of entering high school and HE. They realised that being a lower performer would mean they could not aspire to continue their academic careers.

Table 18.1.3 Mean scores for factors impacting continuation after compulsory education

		Mean score
Plan after leaving middle school	Financial barriers	5.8
	School performance	8.1
	Health issues	6.2
Plan after reaching the age of 18	Financial barriers	6.9
	School performance	7.9
	Health issues	6.9

Financial barriers are a joint second highest factor (6.9) after school performance that students thought could impact their prospects of HE, while they regarded financial issues as less important in the decision-making process for after leaving middle school (5.8). This worry about financial issues impacting post-18 plans might be because HE has a high financial cost for students and their families. Tuition fees, living costs and other expenditure, for instance, might be barriers to HE for students from less-favoured backgrounds. These students might not be able to afford the rising cost of HE in China. On the other hand, the higher score for school performance than financial barriers might indicate that students are more likely to be hindered by weaker educational achievement than by lack of financial resources. However, this is just an assumption without adequate evidence. The close association between school performance and family financial status is well-known.

In order to explore earlier educational gaps further, Table 18.1.4 shows differences in HE aspirations between students who planned to go to selective high schools and those who planned to go to ordinary high schools. High school education is mainly state-funded in China. Students graduating from middle school are required to take the High School Entrance Examination to get a place in a high school, and selective high schools usually set much higher enrolment thresholds than ordinary high schools. In other words, selective high schools prefer students who perform very well in the High School Entrance Exam.

As Table 18.1.4 shows, students who stated that they wanted to study in selective high schools had higher aspirations to participate in HE (77%). Less than half the students with aspirations to go to ordinary high schools reported that they wanted to go into HE, but the real percentage of ordinary high school students entering HE might be lower than the 48% who aspired to do so. There is an early educational gap in the transition to post-compulsory education in Table 18.1.4, even though the results are only aspirations.

Table 18.1.4 Percentages of aspirations for HE by aspirations for different types of high school

	Aspiration for HE	No aspiration for HE
Aspiration for selective high school education	77	23
Aspiration for ordinary high school education	48	52

To sum up, the students generally showed a reasonably high aspiration to continue their education, at least with high school education, after compulsory education. Huang and Gove (2015) attributed this high aspiration to Confucianism, which emphasises the importance of education. However, students who do not plan to go into HE might be disadvantaged. Students who reported that they would leave education after high school were possibly clear about their prospects of being accepted into HE, and they were likely to be disadvantaged by family poverty. Current CA policies ignore the group more at risk of leaving education. This is severely unfair.

18.1.2 Self-evaluations of School Performance

Apart from aspirations to continue education, the survey also asked students to self-evaluate their school performance with a score from 0 to 10. Table 18.1.5 shows that the average score of the overall sample was 5.5.

Table 18.1.5 Mean score for the overall sample self-evaluating their school performance

	Mean score
Self-evaluation of school performance	5.5

A possible reason for this modest self-evaluation might be related to the students' backgrounds, as many of them have rural hukou or belong to minority ethnicity groups, both of which are traditionally less advantaged. These students might have more difficulty in doing schoolwork or they might be less confident to report that they do well in schoolwork.

18.1.3 Parents' Aspirations for their Children's Education in the Overall Sample

The parents' aspirations for their children's education are very high, with an average score of 9.4 for high school education and 9.2 for HE (Table 18.1.6). High aspirations by parents are not unusual in China, as education is highly valued in the Chinese context, as it is in many other contexts (Davey et al., 2007). However, what should be noted is that the parents' aspirations in the survey are just estimations by the students of their perceptions of their parents' aspirations.

Table 18.1.6 Parents' aspirations for their children's education

	Mean
Parental aspirations after middle school	9.4
Parental aspirations post-18	9.2

The next section compares differences between three groups in education aspirations after leaving middle school, self-evaluations of school performance and parents' aspirations for their children's education. The aim is to explore whether there are some gaps between the groups in terms of these factors.

18.2 Results for Three Groups

Considering that there might be differences between students who planned to go into both high school education and HE, students who only planned to go to high school and students who did not plan to continue education after leaving middle school, the students were categorised in three groups as is shown in Table 18.2.1. Around half the students in the survey reported that they aspired to go to both high school and HE, and a quarter only aspired to go to high school. More than a fifth of them did not plan to continue after compulsory education. The following analyses are based on these three groups.

Table 18.2.1 Three categories of aspirations for high school and HE

	Percentage
Both HS and HE	47
HS but not HE	27
Neither HS nor HE	22
Aspirations missing	3
Total	100

18.2.1 Educational Aspirations of the Three Groups

It is worth exploring the differences in the characteristics of the three groups with different aspirations for post-compulsory education. As Table 18.2.2 shows, more female students intended to continue with high school education and HE than males. Few females expected to leave education after middle school or after they were 18. The female students are somewhat more advantaged than the male students.

Table 18.2.2 Percentages of students with educational aspirations by student background characteristics

		Both HS and HE	HS but not HE	Neither HS nor HE	Aspirations missing	Total (N)
sex	Male	41	31	25	3	412
	Female	54	24	19	3	390
Ethnicity	Han ethnicity	48	22	24	5	304
	Minority ethnicity	47	31	21	2	498
Hukou status	Urban hukou	57	14	22	7	127
	Rural hukou	46	30	22	2	675
Parental education qualifications	Bachelor's degree or above	66	14	14	7	73
	Post-compulsory education qualification	52	26	20	3	284
	Compulsory education or below	42	31	25	2	445
Parents' occupations	Prestigious jobs	57	17	23	2	82
	Blue-collar workers or peasants	46	29	23	2	486
	Unemployed	41	16	34	9	32
	Self-employed	51	32	12	4	136
	Other or missing	39	26	29	6	66

Tables 18.2.3 and 18.2.4 detail differences in educational aspirations between sexes. Female students are not only more likely to plan to continue into post-compulsory education but are also more likely to target selective high schools (Table 18.2.3) and prestigious HEIs (Table 18.2.4) as their educational destinations than their male counterparts.

Table 18.2.3 Percentages of aspirations for after middle school (age 15) by sex

	Aspiration for high school education		No aspiration for high school education				
	Selective high school	Ordinary high school	Vocational/skill high school	Work	Not decided yet	Other	Total (N)
Male	36	37	6	2	16	3	127
Female	40	39	3	1	16	1	675

Table 18.2.4 Percentage of aspirations for after age 18 by sex

	Aspiration for HE			No aspiration for HE				
	DFC/985/211 university	First-tier university	Any other university	Vocational college	Work	Armed forces	Not decided yet	Other
Male	21	19	2	9	15	30	2	2
Female	31	22	4	4	4	33	1	1

A possible reason for females having these higher educational aspirations could be that female students often perform better academically in school subjects (Gray et al., 2019; Reily, Neumann & Andrews, 2019). They also agreed slightly more than males that their school performance would affect their decisions after middle school (Table 18.2.5).

Table 18.2.5 Differences in how sex groups thought three factors would affect their plans after middle school and after age 18

		Male		Female		Total		
		Mean	SD	Mean	SD	Mean	SD	Effect Size
Plans for after leaving middle school	Financial barriers	5.8	3.2	5.7	3.0	5.7	3.1	0.0
	School performance	8.0	2.5	8.2	2.4	8.1	2.5	0.1
	Health issues	6.4	3.8	6.0	3.6	6.2	3.7	-0.1
Plans for after	Financial barriers	6.8	3.0	7.0	2.8	6.9	2.9	0.1

age 18	School performance	7.8	2.6	7.9	2.4	7.9	2.5	0.0
	Health issues	7.1	3.5	6.7	3.4	6.9	2.9	-0.1

As Table 18.2.6 shows, female students self-evaluated their school performance lower than male students. Considering this, the reason why females tended to have higher education aspirations may not be that suggested above. Although the effect size is almost zero and it is true that self-evaluation of school performance does not necessarily reliably reflect school achievement and is likely to be a subjective judgment, and the lower self-evaluation by females might only result from caution, this lower average self-evaluation cannot support the assumption that it results from better performance.

Table 18.2.6 Self-evaluations of school performance by sex

	Male		Female		Total		
	Mean	SD	Mean	SD	Mean	SD	Effect Size
Self-evaluation	5.5	2.7	5.4	2.5	5.5	2.6	0.0

Another reason might be that females with low education qualifications, such as middle school qualifications, find it much harder to make a living than less-educated males do. Therefore, female students feel more need to continue into post-compulsory education. In other words, higher female aspirations for education might be because of the difficulties females have in finding an excellent job with lower education qualifications than, or even the same qualifications as, their male counterparts (Guo, Tsang & Ding, 2010; Liu, 1998).

Moreover, it might also be because male students from low-income families are more likely to leave education and go to work or do agricultural work for their family after compulsory education. In contrast, female students born into a similar family situation might be allowed to continue their education due to the greater difficulty they have in finding a job. However, it is also possible that female students from such impoverished families do not finish or do not even do compulsory education at all. That is, they are not covered by this survey. Therefore, it cannot be argued that females are more advantaged than males only because they have higher educational aspirations. Again, we would need to look even earlier in students' trajectories.

As for ethnic groups, although there is no enormous difference in aspirations for both high school education and HE between Han ethnicity students and minority ethnicity students, a higher percentage of the latter replied that they might leave education after age 18 (Table 18.2.2). However, we cannot necessarily conclude that minority ethnicity students are disadvantaged in educational aspirations because more Han ethnicity students planned to leave education after middle school, as was shown in Table 18.2.2. If these results are trustworthy, they suggest that for students living in the same area it would be problematic to view minority ethnicity students as more disadvantaged. These disparities again cast doubt on the appropriateness of using minority ethnicity as an indicator in Chinese CA policies.

Table 18.2.7 Percentages of aspirations for after middle school (age 15) by hukou status

	Aspiration for high school education		No aspiration for high school education				
	Selective high school	Ordinary high school	Vocational/skill high school	Work	Not decided yet	Other	Total (N)
Urban hukou	55	20	6	1	15	3	127
Rural hukou	35	42	4	1	16	2	675

Table 18.2.8 Percentages of aspirations for after age 18 by hukou status

	Aspiration for HE			No aspiration for HE					
	DFC/985/ 211 university	First-tier university	Other universities	Vocational college	Work	Armed forces	Not decided yet	Other	Total (N)
Urban hukou	36	20	3	2	7	24	3	5	127
Rural hukou	24	21	3	8	10	33	1	1	675

Not surprisingly, urban hukou students and students with better-educated parents had higher aspirations for both high school and HE (Table 18.2.2). These students might understand the importance of academic high school education for them to get a place in HE or even to live a better life in the future. Compared with their rural hukou counterparts, urban hukouers tended to say that they planned to go to high school, especially good-quality high schools (Table 18.2.7), rather than find a job after completing middle school. Furthermore, they were more likely to state they intended to attend a prestigious HEI and less likely to choose vocational college, the armed forces or work (Table 18.2.8). It seems clear that urban hukou students not only aspire for post-compulsory education but also in prestigious schools.

Furthermore, as Table 18.2.2 shows, students whose parents work in prestigious jobs, including working in government departments/institutions and the professions, had the highest aspirations for both high school education and HE. In addition, they were more likely to state that they planned to go to a selective high school and a prestigious HEI, while their counterparts whose parents are unemployed or do agricultural work were the least likely to do so, as is shown in Table 18.2.9 and Table 18.2.10. This is not surprising because better occupations are likely to mean better family income and better parental educational attainment, both of which are associated with children's education.

Table 18.2.9 Percentages of aspirations for after middle school (age 15) by parental occupation

	Aspiration for high school education		No aspiration for high school education				
	Selective high school	Ordinary high school	Vocational/skill high school	Work	Not decided yet	Other	Total (N)
Government department/institutions	53	26	2	0	18	2	51
Professional job	45	29	10	0	16	0	31
Skilled worker such as gardener	35	41	5	2	15	2	413
Agriculture-related work	32	37	3	1	26	1	73
Unemployed	31	28	3	3	31	3	32
Self-employed as street peddler or similar	40	50	5	0	5	0	40
Self-employed running a company or similar	37	47	4	0	10	2	96
Other	47	21	3	2	19	9	50

Students with self-employed parents, including those working as street peddlers or similar and those running a company or similar, also showed high aspirations to continue their education not only after middle school but also after age 18. More specifically, as Table 18.2.9 shows, those with parents working as street peddlers, which is usually regarded as a disadvantaged occupation, had notably high aspirations to transit to academic high school education (90%). A possible reason might be that, although they are not stable, the incomes of street peddlers are not meagre so that these families are not really struggling with poverty, compared with, for instance, families in which the parents do agricultural or manual work.

Table 18.2.10 Percentages of aspirations for after age 18 by parental occupation

	Aspiration for HE			No aspiration for HE					
	DFC/985/211 university	First-tier university	Other university	Vocational college	Work	Armed forces	Not decided yet	Other	Total (N)
Government department/ institution	33	24	0	4	12	24	2	2	51
Professional work	39	19	0	3	7	26	3	3	31
Skilled worker such as gardener	25	19	3	8	10	33	2	1	413
Agricultural work	19	21	8	11	8	34	0	0	73
Unemployed	25	22	0	3	6	38	3	3	32
Self-employed as street peddler or similar	28	18	5	5	13	30	0	3	40
Self-employed running a company or similar	23	29	3	3	10	28	1	2	96
Other	24	17	0	9	7	38	5	0	58

Lower percentages of students whose parents work as skilled workers or do agricultural work and students with unemployed parents aspired to continue with post-compulsory education and HE (Tables 18.2.9 and 18.2.10). They were more likely to leave education after finishing middle school. However, the most disadvantaged group seems to be students who answered 'other' (missing data for this question). Therefore, once again, missing data should rarely be directly deleted or ignored as they are unlikely to be missing at random.

Tables 18.2.11 and 18.2.12 show findings on students' opinions about the effects of three factors on their plans. There are few differences in their opinions about financial and health issues between the group with aspirations for both high school and HE and the group with aspirations for high school but not HE. But the latter group have lower scores for school performance, with an effect size of -0.2. Furthermore, the seemingly disadvantaged group that report no aspiration to continue their education after middle school have the lowest scores for all three factors in Table 18.2.12.

Table 18.2.11 How those aspiring for both HS and HE and those aspiring for HS but not HE thought three factors would affect their plans for after middle school and after age 18

		Both HS and HE		HS but not HE		Total		Effect size
		Mean	SD	Mean	SD	Mean	SD	
Plan for after leaving middle school	Financial barriers	5.7	3.2	6.1	3.0	5.7	3.1	0.1
	School performance	8.5	2.2	8.1	2.5	8.1	2.5	-0.2
	Health issues	6.4	3.7	6.3	3.6	6.2	3.7	0.0
Plan for after age 18	Financial barriers	7.0	3.0	7.1	2.8	6.9	2.9	0.0
	School performance	8.2	2.4	7.6	2.6	7.9	2.5	-0.2
	Health issues	7.0	3.5	7.3	3.2	6.9	3.5	0.1

Table 18.2.12 How those aspiring for both HS and HE and those not aspiring for HS or HE thought three factors would affect their plans for after middle school and after age 18

		Both HS and HE		Neither HS nor HE		Total		Effect size
		Mean	SD	Mean	SD	Mean	SD	
Plan for after leaving middle school	Financial barriers	5.7	3.2	5.6	2.8	5.7	3.1	0.0
	School performance	8.5	2.2	7.4	2.7	8.1	2.5	-0.4
	Health issues	6.4	3.7	5.8	3.7	6.2	3.7	-0.2
Plan for after age 18	Financial barriers	7.0	3.0	6.5	2.8	6.9	2.9	-0.2
	School performance	8.2	2.4	7.4	2.5	7.9	2.5	-0.3
	Health issues	7.0	3.5	6.3	3.6	6.9	3.5	-0.2

These results might mean that there are other barriers than finance, educational achievement and health problems that are seen as preventing disadvantaged students from persevering with their education. However, these results might be attributable to the wording of the survey questions, as they ask "*How much do you agree with the following statements: Finance (/Performance at school/ Health) is an important issue affecting my plan for after leaving middle school*" rather than *"How much do you agree with the following statements: Finance (/Performance at school/Health) is an important issue affecting my plan to* ***not pursue high school education*** *after leaving middle school."* In other words, the questions are two-tailed. Students who plan to attend high school and HE could regard school performance (or other factors) as crucial only because they are advantaged in it.

18.2.2 Self-evaluations of School Work and Parents' Aspirations for their Children's Education of Three Groups

Tables 18.2.13 and 18.2.14 compare self-evaluations of their school performance between students in three groups with different educational aspirations. The lower their educational aspirations are, the lower their self-evaluations tend to be, with effect sizes of -0.5 between the group aspiring to continuation to HE after age 18 and the group only aspiring to go to high school. The gap between the first group and the group that does not have any aspiration for post-compulsory education is even larger with an effect size of -0.8. These gaps to a certain extent again emphasise the relationship between self-evaluated school performance and educational aspirations, which indicates that the findings in Tables 18.2.11 and 18.2.12 might be two-tailed.

Table 18.2.13 Self-evaluations of school performance by the group with aspirations for both HS and HE and the group with aspirations for HS but not HE

	Both HS and HE		HS but not HE		Total		Effect size
	Mean	SD	Mean	SD	Mean	SD	
Self-evaluation	6.3	2.4	5.0	2.3	5.5	2.6	-0.5

Table 18.2.14 Self-evaluations of school performance by the group with aspirations for both HS and HE and the group without aspirations for HS or HE

	Both HS and HE		Neither HS nor HE		Total		Effect size
	Mean	SD	Mean	SD	Mean	SD	
Self-evaluation	6.3	2.4	4.2	2.6	5.5	2.6	-0.8

Furthermore, it is not surprising that parental aspirations for their children's education are also positively related to students' own educational aspirations (Tables 18.2.15 and 18.2.16). The group with aspirations only for high school education reported lower average scores for parental aspirations (9.4 and 9.2) than the group aspiring for both high school education and HE (9.7 and 9.7).

Table 18.2.15 Parental aspirations for their children's education by the group aspiring for both HS and HE and the group aspiring for HS but not HE

	Both HS and HE		HS but not HE		Total		Effect size
	Mean	SD	Mean	SD	Mean	SD	
Parental aspirations after middle school	9.7	1.0	9.4	1.5	9.4	1.6	-0.2
Parental aspirations after age 18	9.7	1.1	9.1	1.9	9.2	1.9	-0.3

However, a larger gap exists between the group not aspiring to continue their education after middle school (8.6 and 8.3) and the group planning to go to high school and then HE

(9.7 and 9.7).

Table 18.2.16 Parental aspirations for their children's education by the group aspiring for both HS and HE and the group not aspiring for HS or HE

	Both HS and HE		Neither HS nor HE		Total		Effect size
	Mean	SD	Mean	SD	Mean	SD	
Parental aspirations after middle school	9.7	1.0	8.6	2.3	9.4	1.6	-0.7
Parental aspirations after age 18	9.7	1.1	8.3	2.6	9.2	1.9	-0.7

It seems clear that there are some differences in students' aspirations for post-compulsory education between groups with different characteristics. Students from lower SES backgrounds seem less likely to have plans to continue with their education. Students tend to agree that school performance is important when making educational plans. Furthermore, students with lower educational aspirations usually reported lower self-evaluations of their schoolwork and lower parental educational aspirations. Chapter 19 explores the relationships between students' educational aspirations and background characteristics further with regression analyses. Before that, however, the next section investigates how students perceive the fairness of currently implemented CA policies in China.

18.3 Students' Understanding of the Fairness of CA Policies

This section examines how much the middle school students sampled in the survey think the currently used indicators of rural hukou and minority ethnicity are fair. On average, these two indicators are widely acknowledged as fair, with scores out of 10 of 7.9 for rural hukou and 7.7 for minority ethnicity indicators.

The following tables show how much student groups with different characteristics approve of these two indicators. The students were asked to rate these indicators with scores out of 10. According to Table 18.3.1, males tended to approve the statement "I agree that rural hukou/minority ethnicity is a fair CA indicator," with scores of 8.1 and 7.8, more than females did, with scores of 7.7 and 7.5. The effect sizes are -0.2 and -0.1 respectively.

Table 18.3.1 Average scores for perceived fairness of the two currently used indicators by sex

	Female		Male		Total		
	Mean	SD	Mean	SD	Mean	SD	Effect Size
Rural hukou	7.7	2.7	8.1	2.8	7.9	2.8	-0.2
Minority ethnicity	7.5	3.0	7.8	3.0	7.7	3.0	-0.1

Table 18.3.2 and Table 18.3.3 show that, compared with Han ethnicity students and urban hukou students, minority ethnicity students and rural hukou students are much more likely to agree that ‘rural hukou’ and ‘minority ethnicity’ are fair and valid indicators of disadvantage in contextualised admission policies. In addition, the effect sizes between the ethnicity groups (0.4 and 0.6) and the hukou groups (0.4 and 0.1) are much larger than those between sex groups, which means that the differences in endorsements of the two official indicators are more significant between ethnicity and hukou groups than sex groups.

Table 18.3.2 Average scores for perceived fairness of the two currently used indicators by ethnicity

	Minority ethnicity		Han ethnicity		Total		
	Mean	SD	Mean	SD	Mean	SD	Effect Size
Rural hukou	8.3	2.5	7.2	3.0	7.9	2.8	0.4
Minority ethnicity	8.3	2.6	6.5	3.3	7.7	3.0	0.6

Table 18.3.3 Average scores for perceived fairness of the two currently used indicators by hukou status

	Rural hukou		Urban hukou		Total		
	Mean	SD	Mean	SD	Mean	SD	Effect Size
Rural hukou	8.1	2.7	7.0	3.2	7.9	2.8	-0.4
Minority ethnicity	7.7	3.0	7.3	3.0	7.7	3.0	-0.1

It is not surprising that minority ethnicity students and rural hukou students highly value the two currently used indicators as they are the potential beneficiaries of these indicators. In interview, a minority ethnicity student in Sangzhi said:

“Yes, I think extra scores for minority ethnicity students and rural hukou students

are fair ... because I can get those scores ...”

However, the reason why Han ethnicity students and urban hukou students gave these official indicators lower scores in their evaluations of their fairness and reasonableness might be more complicated. It could be because they think these indicators are problematic in identifying disadvantage. They might have encountered disadvantaged peers with Han ethnicity or with urban hukou. Alternatively, they might have met students with minority ethnicity or rural hukou who were more advantaged in terms of resources. Of course, another reason could just be that they are not the targeted beneficiaries of these indicators.

Furthermore, it is interesting to see the gap, although smaller, between male students and female students in Table 18.3.1. A possible reason for this gap might be that more male students involved in the survey come from less-advantaged backgrounds, because girls from families with the same poverty level are much less likely to get the chance to attend middle school, even though it is compulsory. That is, female students from impoverished families might not be covered by this survey. However, as Table 18.3.4 shows, there is no evidence to support this assumption. On the contrary, there is even a slightly higher percentage of males in the urban hukou group, which is the potentially advantaged group.

Table 18.3.4 Percentages of the distributions in the ethnicity groups and hukou groups by sex

	Han ethnicity	Minority ethnicity	Urban hukou	Rural hukou	Total
Male	37.6	62.4	17.5	82.5	412
Female	38.2	61.8	14.1	85.9	390
Total	37.9	62.1	15.8	84.2	802

Another possible explanation could be that female students are more likely to study hard in order to achieve good results in school subjects or to satisfy their parents and teachers, and it has already been shown that female students pay more attention to school performance when asked about factors that will influence their decisions in the future. Therefore, they might be less in favour of extra support provided by contextual indicators as this support is not directly associated with the effort one makes. This non-effort-oriented ‘reward’ might seem unfair to hard-working students.

According to Table 18.3.5, students with better-educated parents are less likely to approve of the rural hukou indicator. A possible reason could be that better-educated parents, such as ones who finished HE or vocational college education, are more likely to have urban hukou, as attending HE is one of the most common ways to change one's hukou status. As urban hukouers they will be less happy with the additional support in HE access for rural hukou students.

However, this tendency is not seen regarding the minority ethnicity indicator. Students with parents with first degrees tend to evaluate this indicator more positively than the other three groups. Given the results shown in Table 18.3.2, there could be quite a few students with well-educated parents or coming from more affluent families belonging to minority ethnicity groups. This again might mean that minority ethnicities are not necessarily associated with disadvantage.

Table 18.3.6 compares the perceptions of the currently employed CA indicators by students with parents with different occupations. Students with parents working for the government, as professional workers, as peasants, and staying at home unemployed are more likely to consider it fair to treat students with minority ethnicities and rural hukouers as disadvantaged and give them extra scores. In contrast, students with parents who are skilled workers or self-employed and those missing their parents' occupations give the indicators lower scores.

It seems that there are differences in students' perceptions of the extent to which the currently employed CA indicators are fair between groups with different characteristics. Less advantaged groups which might be beneficiaries of these indicators tend to agree on their fairness more than others. However, it is possible that students only have a very vague impression of the fairness of CA indicators as several respondents in the interviews made replies like this:

"Yes, I think it [giving extra scores to minority ethnicity students and rural hukou students] is fair ... because it is a government regulation."

Table 18.3.5 Average scores for the fairness of the two currently used indicators by parents' education level

	Compulsory education qualification or below		High school qualification or equivalent			Vocational college qualification			Bachelor's degree or above			Total	
	Mean	SD	Mean	SD	Effect size	Mean	SD	Effect size	Mean	SD	Effect size	Mean	SD
Rural hukou	8.1	2.6	7.7	2.9	-0.1	7.7	3.0	-0.1	7.5	3.0	-0.2	7.7	3.0
Minority ethnicity	7.7	3.1	7.6	3.0	0.0	7.3	3.2	-0.1	8.0	2.6	0.1	7.9	2.8

Table 18.3.6 Average scores for the fairness of the two currently used indicators by parents' occupation

	Skilled workers		Government department/ institution			Professional worker			Agricultural worker			Staying at home unemployed			Self-employed as street peddler or similar			Self-employed running a company or similar			Other			Total	
	Mean	SD	Mean	SD	Effect size	Mean	SD	Effect size	Mean	SD	Effect size	Mean	SD	Effect size	Mean	SD	Effect size	Mean	SD	Effect size	Mean	SD	Effect size	Mean	SD
Rural hukou	8.0	2.7	8.3	2.4	0.12	8.2	2.5	0.05	8.1	2.8	0.04	8.4	2.5	0.13	7.7	2.9	-0.12	7.8	2.9	-0.09	6.6	3.3	-0.51	7.9	2.8
Minority ethnicity	7.8	2.9	8.4	2.4	0.21	8.1	2.5	0.10	7.8	3.3	0.00	8.0	2.5	0.04	7.4	2.9	-0.14	7.5	3.2	-0.11	6.0	3.4	-0.60	7.8	3.0

18.4 Discussion and Conclusion

To sum up, in general, the students in this survey had high aspirations to continue to post-compulsory education, but some disadvantaged students did not even intend to attend high school. Again, this reminds us that selection for HE begins at least at the transition to high school education.

About half of the students in the survey planned to attend both high school and HE. In contrast, a fifth of them reported that they would leave education after finishing middle school. According to Table 18.2.2 the latter were more likely to be from disadvantaged backgrounds.

More female students than male students in the survey generally showed an aspiration for high school education and HE. However, this does not necessarily mean that females are more advantaged than males. Moreover, female students tend to disagree more than male students about the fairness and appropriateness of using minority ethnicity and rural hukou to identify the disadvantaged in CA policies.

Ethnicity, a criterion in currently employed CA policies, has been found to be an unsatisfactory indicator of disadvantage. Despite there being few differences in aspirations for both high school education and HE, a higher percentage of minority ethnicity students reported that they would leave education after age 18, but more Han ethnicity students planned to leave education after middle school. These gaps again warn us that minority ethnicity is not a synonym for disadvantage as there can be stratification within this group. Not all minority ethnicities are socioeconomically disadvantaged. There are disadvantaged minority ethnicity people who might face many barriers against their education, but there also are advantaged minority ethnicity students who can enjoy prestigious education. The same is true for Han ethnicity students. In China it is not satisfactory to identify the disadvantaged by their ethnicity, although minority ethnicity students think it is fair for them to be given extra scores in the NCEE.

Students with different hukou statuses had some disparities in their plans for their future education. Urban hukou students aimed to pursue more education, or more elite education, after compulsory education than rural hukou students. Furthermore, students with parents working in prestigious occupations and with better educated parents are more likely to have higher educational aspirations. These advantaged groups are often less likely to regard extra scores for minority ethnicity students and rural hukouers as just or fair.

On the other hand, there is a close link between self-evaluated school performance and educational aspirations, which are positively correlated. However, because of the lack of any actual record of students' achievements, higher self-evaluation scores might reflect the truth, meaning that these students indeed do better in their studies, but they might just reflect self-confidence.

The students tended to think their parents had very high aspirations for their educational attainment. If these opinions are trustworthy it might be a manifestation of embodied Confucianism in Chinese traditional culture, which not only emphasises the importance of education but also links education to the development of, or even success in, future careers.

In conclusion, based on these results, it might be unfair to argue that disadvantaged students do not work hard so that they do not go into HE, or that disadvantaged students do not go into HE because they do not want to. Aspirations for HE and even for high school education are clearly stratified by different characteristics.

The results from this survey present some possible issues which are not addressed by the large-scale survey studies in China and so have policy implications.

Chapter 19 Regression Analyses of the Primary Survey Data

The last chapter presented findings on students' education aspirations patterned by their socio-economic backgrounds, self-evaluations of their school performance and their parents' aspirations. In this chapter, logistic regression models are used to predict students' aspirations for HE and for high school education.

19.1 Predicting Aspirations for Higher Education

In this section students' aspirations for HE are predicted using information on their characteristics collected in the survey.

19.1.1 Percentage Variation

In the survey, 49% of the students reported that they aspired to continue their education at the HE level and 51% did not. This half-and-half split in their responses is suitable for conducting a binary logistic regression. Therefore, the regression model predicting aspirations for HE covers all of the cases.

Table 19.1.1 presents the result of the logistic regression Model 2 using forward stepwise method, where family-level indicators are put into Block 1. Students' parents' highest education qualification increases the percentage of correct predictions by 5.3 percentage points and students' parents occupations are excluded from the model. Sex improves for another 1.5 percentage points and self-evaluation of school performance makes the greatest contributions. On the other hand, the official indicators: hukou and ethnicity are excluded from the regression model.

If we add another binary variable in Block 4, namely aspirations for high school (HS) education or not, the prediction becomes more accurate, with the proportion of correct predictions rising to 73% (see Table 2 in Appendix 14). This new predictor contributes another 7 percentage points to the result.

Table 19.1.1 Percentages of correct predictions of aspirations for HE (Model 2; forward)

		Percentage	Percentage increase
Base		51.1	
Block 1	Step 1 (parental highest education qualification)	56.4	5.3
Block 2	Step 1 (sex)	57.9	1.5
Block 3	Step 1 (self-evaluation of school work)	66.0	8.1

N=801

Furthermore, in order to show how important the aspiration for high school education or not variable is, we modify the model a little as we did in Chapter 17 by putting this variable in Block 1, as is shown in Table 19.1.2. The contribution of this variable then seems more straightforward. Aspirations for high school education increase the percentage of correct predictions by nearly 19 percentage points and the other background information, parental educational qualifications, sex and hukou, does not increase it any more. Ethnicity and parents' occupations have been excluded from the model. Students' self-evaluations of their school performance, which in the previous models explained quite a lot of variation, increases the percentage of correct predictions by 2 percentage points in Table 19.1.2.

Table 19.1.2 Percentages of correct predictions of aspirations for HE (Model 4; forward)

		Percentage	Percentage increase
Base		51.1	
Block 1	Step1 (aspiration for HS or not)	69.8	18.7
Block 2	Step1(parental highest education qualification)	69.8	0
Block 3	Step1 (sex)	69.8	0
	Step 2 (hukou)	71.0	1.2
Block 4	Step1 (self-evaluation of school work)	73.0	2

N=801

Therefore, for middle school students at the end of compulsory education, aspirations for HE are strongly related to previous aspirations for high school education. This finding supports the result of the analysis of the CFPS data in Chapter 17.

19.1.2 Coefficients

Table 19.1.3 shows the coefficients for each indicator of Model 2. Parents' highest education qualifications show a strong association with students' aspirations for HE. Compared with students whose parents have only attended compulsory education, students with parents with at least a bachelor's degree were 1.9 times as likely to aspire to go into HE. Students whose parents had completed post-compulsory education, including high school education and vocational college education were also somewhat advantaged, as they were 1.3 times more likely to report aspiring to go into HE.

Male students were only half as likely to report an aspiration for HE as their female counterparts. This is consistent with the findings in Chapter 18.

There is an association between self-evaluations of school performance and students' aspirations for HE. When these self-evaluations increase by one point, aspirations for HE increase by 1.3.

The model excludes ethnicity, hukou status and parents' occupations. This exclusion of ethnicity and hukou status warns us that these two could be a questionable indicator of disadvantage in HE participation.

On the other hand, the reason for the exclusion of parents' occupations might be more complicated. Parents' occupations and parents' education qualifications might be highly correlated, and the model excludes predictors when there is high multi-collinearity. However, the correlation between parental occupations and parental education in the survey sample is only 0.31, which does not support the above assumption. Alternately, this predictor might be excluded because it is less important. This less importance may be a result of inaccurate information collecting, as the answers in the survey are self-reported by students. Or it might be that parental occupation does not well reflect the family's economic characteristics, such as family income, which is complex and can vary according to the area, company, contracts and many other factors.

Table 19.1.3 Coefficients for the variables in Regression Model 2

	Coefficients
both parents only have compulsory education qualifications or below	
at least one parent has a bachelor's degree	1.9
at least one parent has a post-compulsory education qualification	1.3
parents are blue-collar workers or peasants	
parents working in prestigious jobs	
parents' occupations missing	
parents unemployed	
parents self-employed	
female	
male	0.5
minority ethnicity	
Han ethnicity	
rural hukou	
urban hukou	
Self-evaluation for school work	1.3

Note: an empty cell means that the variable was excluded from the model

After adding the predictor of aspirations for high school education or not into Model 4, as Table 19.1.4 shows, an enormous gap is revealed between students wanting high school education and those not aspiring to go to high school. Those with high school aspirations are nearly 20 times more likely to aspire to go into HE.

Students with better-educated parents maintain their advantage in aspirations for HE after controlling for aspirations for high school education. Male students are still less likely to aspire for HE, while Han students and urban hukou students remain advantaged. Urban hukou students are more likely to report aspirations for HE than rural hukou students. In addition, every unit increase in self-evaluations of school performance increases aspirations for HE by

1.2.

Table 19.1.4 Coefficients for the variables in Model 4

	Coefficients
no aspirations for high school	
aspirations for high school	19.7
both parents only have compulsory education qualifications or below	
at least one parent has a bachelor's degree	1.5
at least one parent has a post-compulsory education qualification	1.2
parents are blue-collar workers or peasants	
parents working in prestigious jobs	
parents' occupations missing	
parents unemployed	
parents self-employed	
female	
male	0.5
rural hukou	
urban hukou	1.6
minority ethnicity	
Han ethnicity	
self-evaluation for school work	1.2

Note: an empty cell means that the variable was excluded from the model

To sum up, there are some disparities within each group of characteristics. Male students, rural hukou students and students with less-educated parents are less likely to have aspirations for HE. Ethnicity and parental occupations are usually excluded from the regression model with the forward stepwise (conditional) method, which might indicate that they are less important predictors of aspirations for HE.

Biographical characteristics and family background information can help predict aspirations for HE to a certain extent, but self-evaluations of school performance and especially aspirations for high school education improve the prediction to a much larger extent. Because of the importance of aspirations for high school education in the model in predicting aspirations for HE, in the next section aspirations for this earlier educational transition are predicted.

19.2 Predicting Aspirations for High School Education

In this section regression analyses are conducted to predict students' aspirations for high school education using students' background information and to examine whether and how these aspirations are stratified by this information

19.2.1 Percentage Variation

The results of the regression are presented in Table 19.2.1. It shows the average summary results for ten sub-datasets of Model 2, and the complete results are presented in Appendix 14. The results of Model 1 are also presented in Appendix 14.

Table 19.2.1 Summary of average percentages of correct predictions of aspirations for HS in ten regressions (Model 2; enter)

	Percentage	Percentages increase
Base	50.0	
Block 1 (parental education qualifications, parents' occupations)	58.1	3.1
Block 2 (sex, hukou, ethnicity)	58.6	0.5
Block 3 (self-evaluations of school work)	67.3	8.7

N=386

According to Table 19.2.1, there is a faint sign that family-level indicators increase the percentage by more percentage points than sex, hukou and ethnicity do. These latter three indicators only improve the prediction by 0.5 percentage points after parental education qualifications and parents' occupations have already been entered. The self-evaluation of school performance indicator changes little, which still makes a great contribution in increasing percentage points to 67.3.

19.2.2 Coefficients

Table 19.2.2 presents the average coefficients in ten regressions for each variable in Model 2. First, and not surprisingly, parental education qualifications show some stratifications in aspirations for high school education. Compared with students whose parents have only completed compulsory education or below, students with better-educated parents were from 1.4 to 2.4 times more likely to report that they wanted high school education.

There are some surprising results for parental occupations. Compared with their counterparts whose parents work as skilled physical labourers or peasants, students whose parents work in prestigious jobs, who are expected to have the highest aspirations for high school education, and students with missing information about their parents' occupations are less likely to aspire to aspire to go to high school. The gaps are not remarkable, however. Students with unemployed parents are still the most disadvantaged, as less than half of them reported that they were going to continue to high school after leaving middle school. However, students with self-employed parents seem the most advantaged in their aspirations for high school education. They were 2.1 times more likely to report this aspiration than students with skilled worker or peasant parents.

Table 19.2.2 Summary of the average coefficients for the variables in ten binary logistic regression models (Model 2; enter)

	Coefficients
both parents only have compulsory education qualifications or below	
at least one parent has a bachelor's degree	2.41
at least one parent has a post-compulsory education qualification	1.43
parents are blue-collar workers or peasants	
parents working in prestigious jobs	0.79
parents' occupations missing	0.95
parents unemployed	0.44
parents self-employed	2.11
female	
male	0.66
minority ethnicity	
Han ethnicity	0.60
rural hukou	

urban hukou	0.97
self-evaluation of school work	1.27

Male students are nearly half less likely to report an aspiration for high school education, which is the same as their aspirations for HE. Interestingly, Han ethnicity students in the survey sample are more disadvantaged in their aspirations for high school education. Although the cases are not representative and the sample size is small, this result again indicates that minority ethnicity is an inappropriate indicator of disadvantage in CA policies, or at least it requires us to rethink this indicator. Hukou status does not show huge differences, but surprisingly urban hukouers are a little bit disadvantaged.

Self-evaluations of school performance in Table 19.2.1 show a very similar result to those in Table 19.1.3 and Table 19.1.4. As these self-evaluations increase by one, the probability that students plan to attend high school education increases by 1.3.

To sum up, it seems clear that aspirations for high school education, which is the most important predictor of aspirations for HE, are stratified by individuals' backgrounds. The disadvantaged students suffer from the accumulated disadvantages at least from high school period.

19.3 Discussion and Conclusion

In conclusion, individual background information, such as sex, ethnicity, hukou status, parental education qualifications and parental occupations, is somewhat associated with students' aspirations for high school education and for HE. However, after adding self-evaluations of school performance and aspirations for high school education (only to the model predicting aspirations for HE), these background variables seem less important. Despite remarkable contributions by these two predictors, they are possibly already stratified and are results of selection.

As for comparisons between groups, male students and students with less-educated parents always show lower aspirations for both high school education and HE. Their disadvantage is more likely to be a result of less-abundant resources than less consciousness of the importance of education, as they all reported high parental aspirations for their education in Chapter 18, which means that at least they know how important their parents think education is.

Rural hukou, which is one of the indicators currently used in Chinese CA policies, does not seem associated with aspirations for high school education. Another official indicator, namely minority ethnicity, shows the opposite pattern to that expected. Minority ethnicity students are more advantaged and Han ethnicity students were less likely to report an aspiration for high school education when controlling for other characteristics. The unsatisfying performance of the two official indicators may be an alarm increasing our awareness that the official indicators need to be rethought.

Parental occupations show an interesting pattern. This variable does not seem important in aspirations for HE. However, in the model predicting aspirations for high school education, the traditionally privileged occupation group does not have a notable advantage in aspirations for high school education or HE. Students with parents working in traditionally privileged occupations are even less likely to report that they want to go to high school than their counterparts with parents who are blue-collar workers or peasants. Nevertheless, the gaps are very small. When other information is considered, disparities between students with parents doing prestigious jobs and those with parents doing physical jobs become unimportant.

Students with self-employed parents are more likely to show aspirations for HE and for high school education. As was explained in Chapter 18, self-employment does not necessarily mean a lack of economic resources, which might be one reason for the advantage shown by the self-employed group. Importantly, students who did not clearly report their parents' occupations are usually the most disadvantaged. Once again, this reminds us of the importance of paying attention to missing data.

The results of the survey are generally similar to those obtained from the large-scale datasets. However, what needs to be stressed again is that the results of the survey are just aspirations self-reported by students. They are not only difficult to measure but are also easily changeable. It is possible that, even though less-advantaged students involved in the survey showed high aspirations to continue into post-compulsory education, they will leave education because of a variety of barriers that they might come across later in their lives. There will be more discussion of this in the next chapter, which is the conclusion of this study.

Chapter 20 Conclusions

This chapter summarises the findings of the whole study, which shows how HE participation disparities in China look according to different indicators. The following two sections then evaluate the extent to which the currently employed contextualised indicators and potentially effective indicators are appropriate to be used in CA policies. Next, there is a reminder, or warning, that ignoring students who leave education earlier is a mistake. The limitations of this study and implications for future studies, policymaking and practice are then discussed in the final sections.

20.1 The Disparities in HE Participation by Different Indicators

20.1.1 Official Indicators

According to the results of the structured review, the gaps in HE participation between hukou provinces are remarkable. Many researchers have focused on this issue and compared provincial HE participation as measured by different indices such as Gini coefficients, admission rates and Wilson coefficients. Most of them have found that although there was an overall increase in HE enrolments nationwide after the HE expansion in the 1990s, the gaps between provinces did not decrease. The municipalities, especially Beijing, Tianjin and Shanghai, are widely agreed by researchers to be the "absolutely advantaged areas." (Zhang & Li, 2019) These areas have rich HE resources and well-developed economies. Students whose hukou is in one of these municipalities have much higher chances of getting HE access. The second advantaged areas are the provinces/autonomous regions that benefit from the preferential treatment of minority ethnicity groups, such as Qinghai and Ningxia.

In contrast, provinces in the centre of China and provinces with large populations are found to be the most disadvantaged in terms of HE admissions. In these provinces – Henan, Hebei, Hunan and Jiangxi – getting a place in HE is more competitive than in the two previously mentioned groups of provinces.

These HE admission gaps between provinces in previous studies were supported in Chapter 10 with more recent and complete data. Regardless of whether they are compared using admission rates, the Gorard Segregation Index or the Admission Opportunity Index, provincial disparities in HE enrolments are found. The most advantaged areas remain the same as previous studies have concluded. These include the well-developed municipalities, namely Beijing, Tianjin and Shanghai. These municipalities have remained in privileged positions regarding both regular HEIs and prestigious DFC universities. Even when earlier educational trajectories are taken into account, the municipalities still show higher HE student enrolment numbers.

The relatively advantaged provinces display somewhat complicated patterns when it comes to enrolments in regular HEIs and in elite universities. In admissions to regular universities, regardless of their prestige levels, Liaoning, Heilongjiang, Fujian, Jiangsu, Zhejiang, Inner Mongolia, Shandong, Shaanxi, Hubei and Chongqing show advantages, whereas in admissions to prestigious DFC universities, Jilin, Liaoning, Heilongjiang, Fujian, Hainan, Shaanxi, Qinghai, Ningxia, Hubei and Shandong are more advantaged. There are some overlaps, but what is more important are the differences between these two groups. Some provinces such as Jiangsu and Zhejiang are only privileged in regular university admissions. The reasons for these differences could be, first, that although there are considerable HE resources in the first group there are not necessarily elite universities in these provinces. Most of these universities might be local universities, which actively or passively accept more local students due to their geographical locations or investments by local government or local enterprises. These universities provide these provinces with many places in HE.

Second, many of the provinces in the first group are not covered by CA policies so that less-advantaged students in these provinces might face more barriers in the competition for places in prestigious universities. On the other hand, the provinces in the second group either have relatively rich elite universities or they benefit from preferential policies, so they get larger admission quotas in DFC universities.

The most disadvantaged provinces seem to remain the same, and largely coincide with the conclusions of previous studies. From 2016 to 2019 provinces with under-developed economies, such as Gansu, Guizhou and Guangxi, and provinces with huge populations, such as Henan, Anhui, Jiangxi, Hunan and Guangdong had clear disadvantages in admissions to not only regular HEIs but also selective DFC universities.

Therefore, according to the results of both previous studies collected in the structured review and analyses of up-to-date admission data in Chapter 10, provincial gaps in HE enrolments have lasted more than two decades since the expansion of HE. These gaps have not even narrowed in the last five or six years.

Gaps in HE participation between ethnicity groups are smaller. According to the structured review, minority ethnicity groups were in a disadvantaged position because they inhabited less developed areas or because of language barriers. Nevertheless, more and more researchers have found that minority ethnicities are not necessarily under-represented in HE admissions. On the contrary, they even tend to be over-represented in prestigious universities, which might be attributable to the long-implemented CA policies favouring minority ethnicity students. According to the analyses of individual-level datasets, although the disparities might be less remarkable, minority ethnicity groups are indeed left behind the Han majority in HE participation, with odds ratios of 1.5 in the CGSS data and 1.92 in the CFPS data. Due to data limitations, it is not clear what ethnicity disparities look like in prestigious universities, but minority ethnicities do seem to still remain in a disadvantaged position in regular HEIs.

The disparities in HE participation between different Hukou statuses are more impressive than those between ethnicity groups. There is almost a consensus on the inequality in HE admissions between rural hukou students and urban hukou students in the structured review.

Even worse, this inequality, or specifically the disadvantage of rural hukou students, has been a long-term problem for several decades. Not only the conclusions of previous studies but also the analyses of the CGSS and CFPS data show that students with rural hukou are always under-represented in HE. Much lower proportions of them than urban hukou students are always admitted. These gaps can be clearly shown by odds ratios. In the CGSS sample, urban hukou students are 5.06 times more likely to get access to universities than their rural hukou counterparts. In the CFPS sample, although the gaps are slightly narrower, urban hukouers are still 2.89 times more likely to enter higher education. It is not difficult to conclude that a disadvantage exists for rural hukou students in HE participation.

20.1.2 Other Indicators

Sex disparities in HE participation used to be severe, but they became much more moderate after the expansion of HE, as was found in the structured review. However, there are some differences between the sex groups. For example, female students with urban hukou, with better-educated parents or from more affluent families are even more advantaged in access to universities than male students from the similar background. These within-group differences might be evidence that there are few disparities in HE participation between sex groups, but disparities do exist between other groups of characteristics. The results of the analyses of the CGSS and CFPS data confirm that the gaps in HE admissions between the sex groups are negligible.

There are significant differences in HE participation according to family socio-economic status. Not only previous studies but also the analyses in this study of the two large-scale social survey datasets provide evidence of this. Higher family SES, or specifically families with well-educated parents, party-member parents, parents with good occupations and with considerable family income, is closely associated with a high level of HE participation. According to the CGSS results, students from higher social class families are more likely to get access to universities. The analysis of CFPS data revealed a positive relationship between annual family expenditure on education and HE attendance. Furthermore, the gaps between students with different family SES are more substantial than those between students with different hukou status and with different ethnicities.

Few of the previous studies that investigated HE access in the Chinese context paid attention to language barriers or birth months, but these two indicators have some associations with educational outcomes and so they were investigated in the present study. The CGSS collects information on self-evaluated language ability in understanding and speaking Mandarin, and the CFPS collects data on the languages that students use for daily communication with family members and with classmates. It is not surprising that students in the CGSS sample who evaluated their language ability as ‘good’ or ‘very good’ show a much higher proportion of HE participation than those saying that they are ‘not bad’ or ‘bad’ at understanding or speaking Mandarin, or even ‘cannot understand/speak Mandarin at all.’ In addition, students in the CFPS sample who speak Mandarin with their family and at school are more advantaged in getting a place in HE than their counterparts who speak a Chinese dialect, a minority ethnicity language or another language. There are some gaps in HE access between different language users.

As for birth month, there is no noteworthy disparity between summer-born and autumn-born students. The two groups in both the CGSS sample and the CFPS sample have similar percentages of HE participation.

School-level characteristics are very important in getting access to HE. Studies in the structured review found that students in selective high schools are more advantaged in access to universities, or even elite universities, than their counterparts who study in non-selective high schools. However, according to the CFPS results, although the gap between key-school students and non-key-school students is large, the more important gap is that between students who attend high school and those who do not. Students who do not attend high school cannot go to university even if they want to in the future.

A little surprisingly, there are few stratifications in class types. Students studying in selective classes are not necessarily more advantaged in HE admissions than those who study in non-selective classes, but students in schools that do not distinguish between selective and non-selective classes have a clear disadvantage. Finally, students who reported that they got higher scores in their last Chinese/mathematics tests and those who got higher scores in word/mathematics cognitive tests in the CFPS sample are more likely to attend a university later.

To sum up, disparities, or inequalities, exist in Chinese HE transitions. Students from under-developed or populous provinces, with rural hukou, living in rural areas, belonging to some minority ethnicities, from a family with lower SES status such as poorly educated parents and unemployed parents, who are bad at using Mandarin, who study in non-selective schools, who even fail to go to high school and who perform less well in tests are more likely to be disadvantaged, according to this study. However, this does not necessarily mean that all these disadvantageous characteristics are appropriate to be used as contextualised indicators in CA policies. In the following few sections potential suitable indicators are evaluated.

20.2 Do the Currently Employed Contextualised Indicators Fairly Identify the Disadvantaged Students?

The contextualised indicators currently employed are hukou province, hukou status and minority ethnicity. In short, these three indicators are not wholly suitable indicators to identify disadvantaged students. All of them risk false-positively or false-negatively targeting students. In fact, hukou province is not exactly an indicator in CA policies but a basic admission unit in the primary admission policy, although the province-based quota admission policy was intended to help increase HE admissions for students from impoverished provinces (Bateer, 2009). Nowadays, there are inequalities in the allocations of both regular HE and elite HE places between provinces, and there is a consensus that these allocations need some adjustments. Shang (2018) suggests allocating HE places according to the population or the number of NCEE candidates in provinces. This new allocation might be able to improve HE equity between areas, but the suggestion has been criticised as it ignores differences in the talents of students from different provinces and then has side effects for the reputations of universities and the quality of HE (Zhang & Li, 2019).

Nevertheless, these criticisms or worries need more justifications, as the effectiveness of standardised tests in selecting talented students is questionable because the results are likely to be stratified by family resources. Then, it seems to come back to Rawls's question of whether so-called talents, which are socially related or socially constructed, really deserve to be rewarded. On the other hand, as will be discussed further in the following sections, CA policies at present only work at the point of selection after NCEE. One crucial aim is, or should be, to keep providing support for disadvantaged students in their studies, psychological health, career plans etc. in order to help them complete HE and, hopefully, achieve upward social mobility.

To sum up, hukou province might not be appropriate to be used as a contextualised indicator to identify disadvantaged students, but it might need to be kept as the primary admission unit in the short run. As there are different levels of development between provinces it could result in new inequalities if a nationally unified admission standard is used. However, this does not mean that there is no improvement that can be made to the province-based quota admission policy. On the contrary, more reasonable admission quota allocations and more accurate contextualised indicators are required within each admission unit.

Minority ethnicity is also far from being a suitable indicator, or at least it cannot be used separately. First, compared with other indicators, disparities in HE participation between ethnicities are weaker. The relationships between minority ethnicity and other potential indicators of disadvantage are also obscurer than those between other indicators. More importantly, when controlling for family socio-economic indicators such as parents' education qualifications and occupations in regression models, the ethnicity predictor does not seem relevant in predicting HE participation and is excluded from the model when using the forward (conditional) stepwise method. This indicates that once we know family SES characteristics then ethnicity is no longer helpful in targeting disadvantaged students. Therefore, minority ethnicity is not necessary if more micro-level indicators can be used.

On the other hand, if micro-level indicators cannot be employed in Chinese CA policies minority ethnicity might need to be used in combination with other information in order to identify disadvantaged students more accurately. For instance, in Chapter 16 minority ethnicity had an extremely close relationship with the low-HE-participation area indicator in the CFPS data. The low-participation areas are often areas with huge populations and less developed economies. Furthermore, students who reported that they speak a minority ethnicity language at home do not have a lower HE participation percentage, but those who speak such a language with their classmates have a much lower percentage of HE participation. As was explained in Chapter 15, speaking a minority ethnicity language at school is likely to indicate that the school is a minority ethnicity segregated school. These schools are usually located in poorly developed areas with minority ethnicity populations suffering from a shortage of resources. It might be safer to combine minority ethnicity with living area or school type in order to more accurately target disadvantaged students.

However, this would still problematic. The combination of these three macro-level indicators might help reduce the possibility of mistargeting CA students, but there would be a risk of mistakenly excluding disadvantaged students who are not in the intersection set, for example Han ethnicity students from impoverished families.

Finally, hukou status is also not a good indicator, although it might be the best so far compared with the other currently used CA indicators. There is a visible gap in HE participation between rural hukouers and urban hukouers. Rural hukou is strongly associated with other potential indicators of disadvantage such as less privileged parental occupations and less educated parents. It has even been claimed that rural hukou is the primary reason for inequality of opportunity in educational outcomes in China (Golley & Kong, 2018). The rural hukou indicator seems to precisely target disadvantaged students, but it might not necessarily be so.

First, as has previously been emphasised, the rural hukou indicator is vulnerable to ecological fallacy. For example, it would be difficult to imagine that rural hukou students from Jiangsu province, which is one of the wealthiest provinces in China, would be disadvantaged, or more disadvantaged than urban hukou students from Gansu province, which is one of the poorest provinces. It is also difficult to imagine that all rural hukou students have the same level of disadvantage even if they come from the same areas. Therefore, rural hukou could be misleading if it is used alone. There might be less risk of ecological fallacy if the indicator were combined with other indicators such as living area, hukou province or even better hukou county/district and school background information. When students with rural hukou come from areas with lower GDP and study in rural-hukou-segregated schools or schools located in poor rural areas, they deserve more support from CA policies.

However, a combination of indicators could only alleviate the risk of false-positives to a certain extent. The problem of false-negatives would still be severe, as the binary categorisation of hukou status already neglects possibly disadvantaged urban hukou students, even though the reality could be that students with urban hukou living in the same city could be stratified by family resources in their access to HE (Xu, Song & Liu, 2018).

20.3 Are There Potential CA Indicators That More Accurately Identify the Disadvantaged Students?

In the previous section it was seen that not all the contextualised indicators currently used in Chinese CA policies are of good enough quality to perform their function. Strictly speaking, none of them are able to accurately identify truly disadvantaged students. This section discusses the extent that non-official but potential indicators are appropriate for use in CA policies.

The most promising indicator in this study seems to be parental education qualifications. Lower education qualifications of individuals' parents are not only associated with lower HE participation and other potential indicators of disadvantage but they can also be checked on an official website (xuexin wang), which means this information is reliable, accessible, officially verified and officially defined. Universities and government departments that are involved in HE admissions can access this information before making decisions. Nevertheless, given that HE participants are not a large proportion of the Chinese population, it might not be suitable to use binary classifications of parents' education qualifications. If the first generation in HE indicator is employed in CA policies it might cover too many people. A more recommended way, which has been used in the analyses, is to use a lower level of qualification. Students whose parents only completed compulsory education, rather than students whose parents did not go into HE, might be appropriate for use as a CA indicator. Students whose parents completed post-compulsory education (completed high school education or equivalent, or vocational college or equivalent) is not a suitable CA indicator either as this group is not necessarily disadvantaged in HE participation according to the analyses of the two datasets. It is possible that the post-compulsory group will become disadvantaged after the development of education and further expansion of HE in the future, and at that time a dichotomy (parents attended HE or not) might be reasonable.

Neither sex or birth month is appropriate for use as CA indicators. They both do not show noteworthy disparities in HE participation in the CGSS and CFPS samples and are both risky to ecological fallacy.

As for family-level indicators, social class, parents' political status or parents' occupations would not be suitable CA indicators, either. Social class is vague and contentious in its definition and measurement of the concept. Students whose parents are not party members are somewhat under-represented in HE, but they are too large a group. The information of parents' occupations is also less reliable and non-officially verified.

If parents' occupations could be more reliably and accurately collected they might be able to be used together with family income as a combined indicator, as the former somewhat indicates social resources and the latter indicates economic resources. However, more discussion is needed on what kind of occupations can be confidently regarded as socio-economically disadvantaged. A vague categorisation, such as self-employment, would not be satisfactory. Furthermore, it is also necessary to clarify to what extent occupations could or should be regarded as salary thresholds to identify disadvantage because there are differences in salary within the same occupations. For instance, the average annual salary for people working as staff in state-owned enterprises is 199,278 RMB (£23,815) in Beijing, but the figure is only 88,145 RMB (£10,540) in Gansu (China Statistical Yearbook, 2021). Annual family income also suffers this difficulty to be nationally-unified, so it is less promising to be a CA indicator.

Annual expenditure on education would not be a promising indicator even though it is associated with disadvantage. It is self-reported and hard to verify. Utilisation of annual expenditure on education as a CA indicator might result in lies, concealment and then new inequalities.

Less fluent Mandarin should not be used as a contextualised indicator either, given its inevitable shortcomings, including inaccuracy of self-reporting, difficulties in measuring language ability and deliberate cheating in language tests. In addition, the daily communication language at home or at schools cannot be used as a CA indicator due to its inaccuracy.

School-level indicators are popular in the British context, but neither school type nor class type, which were collected in the CFPS, are suitable to be used as contextualised indicators in China. Students in non-selective schools and in schools that do not distinguish between key classes and non-key classes are under-represented in HE, but this is a huge group of students, not all of whom are necessarily disadvantaged.

20.4 Who Are the Missing Cases in CA, and Why They Matter?

Importantly, a group that is highly likely to be disadvantaged seems to be ignored by current CA policies: those that leave education much earlier than the NCEE. According to the analyses of the CFPS data, nearly half the overall sample left academic education after the end of compulsory education. This group of students went to skill schools, vocational high schools, or even left education completely before they could be targeted by CA policies.

This missing group in the CFPS data seem to be more disadvantaged than their counterparts who successfully moved to ordinary academic high school education and then had the chance to be supported by CA policies. For instance, the students in this missing group are more likely to have rural hukou, live in rural areas, belong to minority ethnicities and have lower SES family backgrounds. They fail to survive for selection in the competition for a place in HE because of accumulated disadvantages, but they seem to be invisible for CA policies, as none of the policies take any action to improve their participation in academic high schools.

Similar patterns can also be found in the CGSS data. Although the CGSS is not a longitudinal study that can trace the educational trajectories of cases, the cases are categorised by their highest educational qualifications. Besides the gaps in HE participation, there is also a clear gap in completion of high school education. Rural hukouers, minority ethnicity students and individuals self-reporting as working class or from less privileged families are less likely to transit from compulsory education to high school and get the corresponding qualification. In addition, sex groups and ethnicity groups show disparities in illiteracy rates. More females and more minority ethnicity students were found to be illiterate in the CGSS sample. All these individuals leaving education early, who might be victims of accumulated disadvantage, are excluded from any support by CA policies.

The more recent data from middle school students in the supplementary survey show similar results. Despite there being high aspirations for high school education on average, less-advantaged students tended to report planning to leave education after graduating from middle school, and few of these students reported an aspiration for HE as the following step after age 18.

All three datasets indicate accumulated disadvantages and early exclusion from education for some students, but few CA policies and even few studies have paid attention to these students, let alone offered help. However, if truly disadvantaged students are not targeted in earlier steps and less disadvantaged students enjoy extra support, is it reasonable and responsible to say that CA policies are fair, just and effective?

20.5 Limitations of this Study

This study has some limitations. First, the evaluation of currently used contextualised indicators did not cover all the indicators used. One of these, namely whether students' hukou is in (and also whether they live and study in) a poor prefecture as verified by the government, has not been examined because of data deficiency. Although poor prefecture is a macro-level indicator that is vulnerable to ecological fallacy, it deserves some examination. This is because, if the province-based admission policy continues to be the fundamental HE admission policy in China, the poor prefecture indicator seems a little more promising than rural hukou and minority ethnicity as these are larger groups, not all of which are necessarily disadvantaged. This does not mean that poor prefecture is definitely appropriate to be used as a CA indicator. Instead, this is just an assumption taking into account the unsatisfactory performance of the two other currently used CA indicators. The two large-scale datasets employed in this study lack good-quality information on prefectures, so the study has failed to evaluate the poor prefecture indicator.

The second limitation is that the study used planned quotas, rather than real intakes, in admissions to DFC universities. As was mentioned in Chapter 10, the admission quotas for provinces of DFC universities that were collected are just plans published on the university websites. Although HEIs in China are highly administrative and their admission plans are set according to government regulations, there will still be some (but few) variations between planned admissions and real ones. Besides the approximately 1% float that is allowed in enrolments, one such example is that some students are admitted by universities but they reject the offer. These variations might not make huge differences to the results, but it would always be good to have more precise data.

Third, although this study has tried to make its sample size as large as possible, such as by merging several datasets from the cross-sectional survey into a new one, the representativeness of the samples is still doubtful. A more, if not the most, representative and valuable dataset would be that resulting from the census with individual-level data, which researchers are not allowed to get access to.

In addition, there is also a weakness in the survey data. Because of time limits and shortages of resources, students involved in the primary survey were not traced back. This means that the information collected was only students' plans for their future education, not their destinations after leaving compulsory education and after age 18. Their educational aspirations and real destinations could be surprisingly different. On the other hand, only 800 students in five cities joined the survey. It is a small sample size considering the population of Chinese middle school students. Therefore, by being better prepared in terms of time and costs future studies could try to enlarge the sample size and collect students' destinations after graduating from middle school and, if they attend it, high school.

20.6 Recommendations for Future Research

First, future studies could attempt to focus on the CA indicator of “impoverished prefectures” if they can find better secondary datasets such as individual-level census data, or they could even collect new high-quality primary data on this. This CA indicator deserves more evaluation to its reliability and accuracy.

Second, researchers could try to get access to the individual-level census data when they conduct research in the future. They would benefit to a large extent if the National Bureau of Statistics could provide these data.

Furthermore, future studies could make more effort to explore the accumulated disadvantage in the earlier educational period, such as conducting some longitudinal studies that focus more on middle school students or primary school students and even earlier as target groups. The disadvantages in HE participation should be traced back.

20.7 Recommendations for Policy

20.7.1 Abolish current CA indicators

First, some indicators officially used in CA policies need reconsideration. In order to reduce disparities between provinces, it might be better to abolish province-based admission quotas and hukou province policy. This artificial restriction for applying for HE is not just or reasonable in the long run. HEIs are supposed to accept students more fairly at the national level instead of at the province level.

However, the abolishment of these policies might be difficult to realise in the short term not only because of the barriers to change a policy but also the current huge gaps between provinces in educational and economic development. Students from richer provinces who enjoy more advanced educational resources might be more likely to go into HE in a nation-level competition, which would create new inequality and injustice.

Therefore, it might be more realistic to use hukou province as a basic admission unit in the current situation, but it does require a fairer allocation of quotas. These quotas nowadays are not fairly allocated even though we only take the NCEE candidates into account. More importantly, it is not adequate to only consider the NCEE candidates, as half of the students in that age cohort leave education before that stage. A fairer admission quota allocation at least could be decided according to the number of students in the relevant age cohort rather than just the number of NCEE candidates, so taking the early leavers into account.

Furthermore, it is crucial to find a nationally unified criterion to measure socio-economic status before Province-based Quota admission policy and hukou province policy are abolished and HEIs accept student at the national level. Otherwise, new inequalities could be expected. There might be some worries about the renaissance of "NCEE migrants" ("Gaokao yimin" in Chinese, see Chapter 2) due to the abolition of the restriction of hukou province. These worries could be illusions if there is a nationally based admission process or a fairer province-based admission quota allocation, as there are no additional benefits for students to move to another province for HE applications.

Minority ethnicity is not an appropriate CA indicator due to its inaccuracy. Some improvements in the use of this indicator are needed, maybe by replacing it with more precise indicators or using it in combination with other indicators that are more closely associated with socio-economic disadvantage.

The CA indicator of rural hukou is not suitable, either. Not only should hukou status not be used as a CA indicator, but the whole hukou status policy should be abolished. There are few persuasive reasons to keep this discriminatory artificial classification and it is actually being abolished gradually. This indicator would be less accurate with the progress of abolishing hukou status policy. Some worries about new inequalities, such as richer families moving to urban cities, along with the abolition of hukou policy at the initial stage. Nevertheless, this is not an excuse for continuing to use such discriminatory classification. Instead, this should be should be a motivation to look for fairer CA indicators and better admission policies to discourage people from migrating for “potential additional benefits”.

However, in the short-term, rural hukou may still need to be used as a temporary CA indicator considering the huge gaps in resources between rural hukou and urban hukou.

20.7.2 Focus on wider target groups

The second implication for future CA policies is that it will never be enough to only take NCEE candidates into account. Instead, a long-term plan to support disadvantaged students is needed. This would cover the early education period and HE after entry. This study has found accumulative inequalities in early education, especially in the transition from middle school to high school, when truly disadvantaged students might leave or, even worse, have left academic education. These students constitute nearly half of their age cohort. This means that CA policies neglect half of those who deserve to be targeted. Even worse, these disadvantages might start during compulsory education (Liu, H., 2014; Young & Hannum, 2018) or even preschool education (Su, Lau & Rao, 2020; Wang & Gong, 2018; Zhou, 2020). However, current CA policies only select cases to give help to from among those who have completed their schooling, NCEE candidates. This is hardly fair. Fairness, or equity, in education should start from equity of educational opportunities and achieve equity of educational outcomes by providing equity of educational processes (Long & Fan, 2013; Salmi & Bassett, 2014). Regardless of their background, individuals should have aspirations for education, and for HE (McCowan, 2007; Vignoles & Murray, 2016).

Therefore, in order to help the earlier missing group access high school education, future CA policies (or there is a requirement for more relevant policies) should aim to alleviate negative effects of individuals’ family socio-economic status on educational attainment at least after compulsory education. For instance, these policies can help disadvantaged students attend high schools as much as possible by requiring high schools to accept these students at a certain proportion. If disadvantaged students feel difficult to participate high school education because of cost, these policies can provide financial support such as student loan or scholarship for them. If disadvantaged students do not, or are not plan to, go to high school because of grades, future CA policies or relevant policies can lower down the entry requirements to an extent as they do for disadvantaged HE applicants.

On the other hand, CA policies should pay attention to students who have already entered HE with lower entry requirements. Some prestigious universities worry about the quality of CA students. They are concerned about lower academic competence, graduation rates and less prestigious future careers, but to a certain extent it should be the responsibility of universities to help them with these issues. Instead of criticising the possibility that CA students might damage their fame and reputation, universities should continue CA policies and provide these students with support in their daily lives, on financial issues and in their studies, socialisation and employment. For example, universities can provide induction sessions for disadvantaged students to help them with their academic abilities if they worry the mismatched quality of these students. Furthermore, universities can provide some scholarships, loans or part-time jobs for disadvantaged students if they are hindered by financial barriers.

20.8 Recommendations for Practice

20.8.1 Publish admission data

First, it would be ideal if all HEIs and just not DFC universities publish their admission data or share such data with researchers rather than only posting admission plans. If information on student intakes is reliably and accurately published, future research could explore more those who are left out in the planned admissions and real admissions. The reason for not accepting offers from universities could be that the students concerned want to go to a more prestigious university so they choose to repeat the last year of high school or go abroad for HE, but there is also the possibility that they cannot afford the university fees so they have to give up. These different stories might reflect HE inequalities.

Alternatively, there could be an institution that manages and publishes datasets of CA-related information such as CA policy introduction, CA criteria eligibility and CA implementation in a standard format (Sundorph, Vasilev & Coiffait, 2017). This would be beneficial not only for future research but also makes the HE admission process more transparent.

20.8.2 Reduce the entry requirements for CA students

Second, the entry requirements for CA students could perhaps be further reduced. Although there are some concerns about the quality of CA students, as was mentioned above (Wang, Pan & Wu, 2017), evidence suggests that CA students do not, or do not necessarily, perform worse than non-CA students academically (Boliver et al., 2017; Ma & Bu, 2019; Xie et al., 2018). Therefore, instead of being wary about the possible side effects of CA students on the reputations of universities, it might be more important to make it clear to what extent entry requirements could or should be reduced for CA students without causing the problems that universities care about. It is more equitable if all CA students are given the same amount of bonus scores when applying to HEIs.

However, before exploring the extent requirements should be reduced, another question is in what ways universities are concerned about being damaged by more CA. Is it something related to academic issues such as graduation rates, percentages of students doing postgraduate education, test scores, pass rates in College English Band 4 and Band 6 (CEB4 & CEB6) or academic publications? Or is it something related to non-academic issues such as the campus environment, campus security or student satisfaction? Or is it something involving both? Only when the answers to these questions are clear will it be meaningful to discuss the possible negative impacts of enrolling CA students. However, in fact, regardless of what the concerns of universities are, CA policies are expected to continue helping disadvantaged students in both academic and non-academic respects before and after their entry into HE. Otherwise, the effectiveness of CA policies in supporting disadvantaged students, improving HE equity and promoting social mobility may be questioned due to the insistence of prestigious HEIs on keeping their prestige and rankings (Evans, 2014).

20.8.3 Reconsideration of HE admission criteria

The reliability of standardised tests, the fairness of HE selection process and the purpose of HE need to be reconsidered. Standardised tests might be convenient for comparisons of students' educational attainment. In combination with a blind marking system they might also create procedural justice. But are not necessarily useful or accurate for selecting appropriate candidates for HE. There is much empirical evidence indicating that educational achievement is strongly associated with family socio-economic status. So, it is possible that selection by test performance may actually be selection by family SES background (Bartholo & Costa, 2014; 2016). If selection by family background is not considered appropriate or fair, then any selection criteria that are highly associated with family background should also be considered inappropriate an unfair.

On the other hand, there is still controversy over what the purposes of HE are or should be, but one of them might be meritocracy in education. Meritocracy involves hierarchies such as a hierarchy of HEIs and meritocracy views it as normal to select qualified students for this hierarchy. However, this hierarchy of HEIs, and then the hierarchy of students, violates the principle of fairness (Yang, Wang & Chen, 2019). There is some support for the view that educational meritocracy is beneficial to students from lower SES backgrounds who are intelligent or diligent and want to attend HE (Bell, 1972; Saunders, 1997), but it might be that, apart from differences in educational attainment, as was discussed in Chapter 3, aspirations to enter HE and willingness to work hard are stratified by resources that families possess. Therefore, meritocracy still favours advantaged backgrounds. Then, how far is the hierarchy of HEIs and selection for entry into HE oriented by meritocracy?

Even if disadvantaged students have the same education aspirations, willingness to work hard and educational attainment with their advantaged counterparts, they might abandon HE purely because of poverty. In this case, which might not be rare cases, meritocracy refers more to "affluent economic elites" than "intelligent academic elites". It surely requires some fundings for helping disadvantaged students have access to HE, such as student loan or scholarships as mentioned above, but it might be more important to rethink the reasonability of meritocracy of HE selection. Some more creative ways, such as alternative routes to HE, could be considered. Maybe it is a good idea to discard grades or scores as an entry criterion for HE, as Gorard et al. (2007) suggested.

Appendices

Appendix 1 HEIs That Provided Special Programmes for Colleges and Universities in 2021

HEIs affiliated with the MOE			
Peking University	Tsinghua University	Renmin University of China	Beijing Jiaotong University
University of Science and Technology Beijing	Beijing University of Chemical Technology	Beijing University of Post and Telecommunications	China Agricultural University
China University of Political Science and Law	North China Electric Power University	Beijing Forestry University	Beijing University of Chinese Medicine
Beijing Normal University	Beijing Foreign Studies Universities	Beijing Language and Culture University	Communication University of China
Central University of Finance and Economics	University of International Business and Economics	China University of Mining and Technology, Beijing	China University of Petroleum – Beijing (CUP)
China University of Geosciences	Nankai University	Tianjin University	Fudan University
Tongji University	Shanghai Jiao Tong University	East China University of Science and Technology	Donghua University
East China Normal University	Shanghai International Studies University	Shanghai University of Finance and Economics	Nanjing University
Southeast University	China University of Mining and Technology	Hohai University	Jiangnan University
Nanjing Agricultural University	China Pharmaceutical University	Zhejiang University	Hefei University of Technology
Xiamen University	Shandong University	Ocean University of China	China University of Petroleum
Dalian University of Technology	Northeastern University	Jilin University	Northeast Normal University
Northeast Forest University	Wuhan University	Huazhong University of Science and Technology	China University of Geosciences, Wuhan
Wuhan University of Technology	Huazhong Agricultural University	Central China Normal University	Zhongnan University of Economics and Law
Hunan University	Central South University	Sun Yat-sen University	South China University of Technology
Chongqing University	Southwest University	Sichuan University	Southwest Jiaotong University
University of Electronic Science and Technology of China	Southwestern University of Finance and Economics	Xi'an Jiaotong University	Xidian University
Chang'an University	Northwest Agriculture	Shaanxi Normal	Lanzhou University

	and Forestry University	University	
Other HEIs			
Beihang University	Beijing Institute of Technology	Harbin Institute of Technology	Harbin Engineering University
Nanjing University of Aeronautics and Astronautics	Nanjing University of Science and Technology	Northwestern Polytechnical University	Dalian Maritime University
University of Science and Technology of China	Beijing University of Technology	Heilongjiang University	Shanghai University
Soochow University	Nanjing Normal University	Fuzhou University	Zhengzhou University
Hunan Normal University	Guangxi University	Southwest University of Political Science and Law	Sichuan Agricultural University
Guizhou University	Yunnan University	Northwest University	

Appendix 2 The Questionnaire in Chinese and in English

Chinese Version

1）您目前初中毕业之后的计划是？（　　　）

1. 去重点高中
2. 去普通高中
3. 去职高或中专
4. 去工作
5. 我不知道/还没有决定
6. 其他（请填写）______________

请用 0-10 表达你同意的程度，0=不同意；10=非常同意

2）您多大程度上同意下列这些陈述符合您的实际情况？

1. 经济因素（如：钱）是影响我初中毕业之后的计划的一个重要因素

0 1 2 3 4 5 6 7 8 9 10

2. 学校学习成绩是影响我初中毕业之后的计划的一个重要因素

0 1 2 3 4 5 6 7 8 9 10

3. 身体健康问题，是影响我初中毕业之后的计划的一个重要因素

0 1 2 3 4 5 6 7 8 9 10

4. 我的父母希望我读高中

0 1 2 3 4 5 6 7 8 9 10

5. 我的学习成绩很不错

0 1 2 3 4 5 6 7 8 9 10

3）就目前来说，您 18 岁之后的计划是？（　　　）

1. 去双一流/985/211 大学
2. 去除 1 中提到的其他一本院校
3. 去本科院校就可以
4. 去大专
5. 去工作
6. 参军
7. 我不知道/还没决定
8. 其他（请填写）_______________

请用 0-10 表达你同意的程度，0=不同意；10=非常同意

4） 您多大程度上同意下列这些陈述符合您的实际情况？

1. 经济因素（如：钱）是影响我 18 岁之后的计划的一个重要因素

0 1 2 3 4 5 6 7 8 9 10

2. 学校学习成绩是影响我 18 岁之后的计划的一个重要因素

0 1 2 3 4 5 6 7 8 9 10

3. 身体健康问题，是影响我 18 岁之后的计划的一个重要因素

0 1 2 3 4 5 6 7 8 9 10

4. 我的父母希望我读大学

0 1 2 3 4 5 6 7 8 9 10

我国有一些针对少数民族和农村学生的高考优惠政策，请您评价您认为它们有多公平 （0=不公平，10=非常公平）

5）您认为少数民族学生高考加分有多公平？

0 10

6）您认为农村户口的优秀学生高考时能够通过申请和考试，得到加分的机会有多公平？

0 10

为了帮助我们更好地理解您的回答，请让我们多了解一些您的基本信息。

7）您的性别是？（ ）

1. 男 2. 女

8）您的民族是？（ ）

1. 汉
2. 少数民族
3. 其他（请填写）________________

9）您的户口是？（ ）

1. 城市户口
2. 农村户口
3. 其他（请填写）________________

10）您父母中，拥有的最高教育程度是？（ ）

1. 本科及以上
2. 专科
3. 高中及同等学历
4. 义务教育学历或以下

11）下列哪个最接近您父母的工作？（ ）

1. 在政府部门或事业单位工作
2. 老师、律师、医生等工作
3. 务工
4. 务农
5. 在家待业
6. 经营公司等
7. 经营小商店/小摊贩等
8. 其他（请填写）____________________

12）关于问卷里提到的内容，您还有别的想告诉我们的吗？

English Version

1）What is your current plan for after leaving middle school?（ ）

1. Go to a selective high school
2. Go to an ordinary high school
3. Go to a vocational high school/skill high school
4. Go to work
5. No idea/not decided yet
6. Other (please describe briefly) __________________

Please rate your agreement with each item on a scale from 0=do not agree at all to 10=very strongly agree

2）How much do you agree with the following statements?

1. Finance is an important issue in my decision on my plan for after leaving middle school.

0 10

2. My performance at school is an important issue in my decision on my plan for after leaving middle school.

0 10

3. Health is an important issue in my decision on my plan for after leaving middle school.

0 10

4. My parents want me to go to high school.

0 10

5. I am doing well in my school subjects.

0 10

3）What is your current plan for after age 18? （ ）

1. Go to a WDF/985/211 university/HEI
2. Go to a first-tier university/HEI (except those in option 1)
3. Go to another university/HEI
4. Go to a vocational college
5. Go to work

6. Join the armed forces
7. No idea/not decided yet
8. Other (please describe) ____________________

Please rate your agreement with each item on a scale from 0=do not agree at all to 10=very strongly agree

4) How much do you agree with the following statements?
1. Finance is an important issue in my decision on my plan for once I am 18 years old.

0 10

2. My performance at school is an important issue in my decision on my plan for once I am 18 years old.

0
10

3. Health is an important issue in my decision on my plan for once I am 18 years old.

0
10

4. My parents want me to go to HE

0
10

There are plans/policies to help disadvantaged students to gain places at university. Please rate how fair you think each of the following plans is using a scale from 0=not fair to 10=very fair.

5) How fair do you think it is that minority ethnicity students get extra credits in the NCEE?

0
10

6) How fair do you think it is that students with rural hukou get the opportunity to earn extra credits to attend university?

0
10

To help us understand your responses better we would like you to tell us a little about yourself.

7) What is your sex? ( )
1.Male

2.Female

8) What is your ethnicity? (　　　)

1.Han

2.Minority ethnicity

3.Other (please describe) ________________

9) What is your hukou identity? (　　　)

1.Urban hukou

2.Rural hukou

3.Other (please describe) ________________

10) What is the highest education qualification of your most qualified parent? (　　)

1.Bachelors' degree or above

2.Vocational college qualification

3.High school qualification or equivalent

4.Compulsory education qualification or below

11) Which one is the closest to the main occupation of your parents? （　　）

1.Working in a government department/institution

2.Working as a professional such as a teacher, lawyer or doctor

3.Skilled workers such as gardeners

4.Doing agricultural work

5.Staying at home without being employed

6.Self-employed as a street peddler or similar

7.Self-employed running a company or similar

8.Other (please describe) ____________________

12) Is there anything else you want to tell us about the issues raised in this questionnaire?

Appendix 3 Numbers of NCEE Candidates in Provinces 2016-2019 (thousand)

	NCEE Candidates			
	2016	2017	2018	2019
Beijing	61.2	60.6	63.0	59.0
Tianjin	60.0	57.0	55.0	56.0
Shanghai	51.0	50.0	50.0	50.0
Chongqing	248.9	247.5	250.0	264.0
Hebei	423.1	436.2	486.4	559.6
Shanxi	339.0	317.0	305.0	314.0
Liaoning	218.2	208.5	185.0	244.0
Jilin	148.5	143.0	150.3	162.7
Heilongjiang	197.0	188.0	190.4	204.0
Jiangsu	360.4	330.1	331.5	339.0
Zhejiang	307.4	291.3	306.0	320.0
Anhui	509.9	498.6	499.0	513.0
Fujian	175.0	188.2	200.9	207.8
Jiangxi	360.6	364.9	380.0	421.0
Shandong	710.0	583.0	592.0	559.9
Henan	820.0	865.8	983.8	1,000.0
Hubei	361.0	362.0	374.0	384.0
Hunan	401.6	410.8	451.8	499.0
Guangdong	733.0	757.0	758.0	768.0
Hainan	60.4	57.0	58.8	60.0
Sichuan	570.0	582.8	620.0	650.0
Guizhou	373.8	411.9	440.0	458.7
Yunnan	281.1	293.5	300.0	320.0
Shaanxi	328.0	319.0	319.0	325.9
Gansu	296.0	284.8	273.0	266.8
Qinghai	44.6	46.3	50.0	57.0
Inner Mongolia Autonomous Region	201.1	197.4	195.0	199.0
Guangxi Zhuang Autonomous Region	330.0	365.0	400.0	460.0
Tibet Autonomous Region	24.0	28.5	25.3	——
Ningxia Hui Autonomous Region	69.1	69.2	69.5	71.7
Xinjiang Uygur Autonomous Region	166.1	183.7	207.4	——

Appendix 4 DFC universities the admission quotas of which were collected from 2016 to 2019

	2016	2017	2018	2019
Xiamen University	√	√	√	√
Zhongnan University	√	√	√	√
Zhongshan University	√	√	√	√
Wuhan University	√	√	√	√
Southeast University	√	√	√	√
Tongji University	√	√	√	√
Shandong University	√	√	√	√
Shanghai Jiaotong University	√	√	√	√
South China University of Technology	√	√	√	√
Jilin University	√	√	√	√
East China Normal University	√	√	√	√
Hunan University	√	√	√	√
University of Science and Technology of China	√	√	√	√
Ocean University of China	√	√	√	√
Beijing Normal University	√	√	√	√
Northeastern University	√	√	√	√
Beijing Institution of Technology	√	√	√	√
Northewest A & F University	√	√	√	√
Zhengzhou University	√	√	√	√
Dalian University of Technology	√		√	√
Northwestern polytechnical University	√			√
Peking University	√			
Fudan University		√	√	√

Sichuan University		√	√	√
Nankai University		√	√	√
Chongqing University		√	√	√
Xi'an Jiaotong University		√	√	√
Harbin Institution of Technology		√	√	√
Huazhong University of Science and Technology			√	√
Nanjing University			√	
Lanzhou University			√	√
China Argiculture University			√	√
Yunnan University			√	√
Renmin University				√
Beihang University				√
Minzu University of China				√

Appendix 5 Provinces the admission percentages of which in regular HEIs were collected from 2016 to 2019

Province	2016	2017	2018	2019
Beijing	√	√	√	√
Tianjin			√	√
Shanghai				√
Chongqing	√	√	√	√
Hebei		√	√	√
Shanxi			√	
Liaoning			√	
Jilin				
Heilongjiang			√	√
Jiangsu		√	√	√
Zhejiang	√			√
Anhui	√	√	√	√
Fujian	√		√	√
Jiangxi	√	√	√	√
Shandong	√	√	√	√
Henan	√	√	√	√
Hubei	√			
Hunan	√	√	√	√
Guangdong	√	√	√	√
Hainan	√			
Sichuan			√	√
Guizhou				
Yunnan			√	
Shaanxi			√	√
Gansu	√		√	√
Qinghai	√		√	
Inner Mongolia Autonomous Region	√	√	√	√
Guangxi Zhuang Autonomous Region	√	√	√	
Tibet Autonomous Region				
Ningxia Hui Autonomous Region			√	
Xinjiang Uygur Autonomous Region	√	√	√	√

Appendix 6 De-localisation of Admission Proportions (%) in Four DFC Universities

Table 1 Admission Proportions for Provinces, China Agriculture University

	China Agriculture University (Shandong)	
	2018	2019
Beijing	5.10	5.24
Tianjin	3.50	3.54
Shanghai	0.25	0.24
Chongqing	2.43	2.40
Hebei	5.43	5.33
Shanxi	3.50	3.54
Liaoning	2.80	2.76
Jilin	2.18	2.15
Heilongjiang	2.80	2.80
Jiangsu	2.84	2.85
Zhejiang	3.54	3.46
Anhui	3.79	3.82
Fujian	2.10	2.11
Jiangxi	2.55	2.64
Shandong	22.18	21.99
Henan	4.94	4.96
Hubei	2.47	2.44
Hunan	3.25	3.25
Guangdong	2.59	2.64
Hainan	0.49	0.53
Sichuan	5.02	5.04
Guizhou	1.40	1.42
Yunnan	2.39	2.36
Shaanxi	2.92	2.89
Gansu	1.93	1.99
Qinghai	0.25	0.28
Inner Mongolia Autonomous Region	2.30	2.28
Guangxi Zhuang Autonomous Region	2.72	2.72
Tibet Autonomous Region	0.08	0.04
Ningxia Hui Autonomous Region	0.41	0.45
Xinjiang Uygur Autonomous Region	1.85	1.83

Table 2 Admission Proportions for Provinces, Shanghai Jiaotong University

	Shanghai Jiaotong University (Shanghai)			
	2016	2017	2018	2019
Beijing	1.47	1.48	1.46	1.43
Tianjin	0.94	0.89	0.88	1.43
Shanghai	11.05	9.04	8.67	8.19

Chongqing	2.78	2.62	2.68	2.62
Hebei	2.15	2.47	2.44	2.57
Shanxi	3.61	3.95	3.95	4.05
Liaoning	3.20	3.36	3.31	3.29
Jilin	1.20	1.48	1.46	1.43
Heilongjiang	1.05	1.04	1.12	1.10
Jiangsu	11.52	11.12	11.06	10.67
Zhejiang	1.41	1.33	1.22	1.29
Anhui	4.24	4.59	4.63	4.76
Fujian	3.46	4.00	3.99	4.19
Jiangxi	4.45	4.30	4.58	4.57
Shandong	8.91	8.65	8.62	8.58
Henan	6.13	6.03	6.04	6.10
Hubei	3.82	5.29	5.31	5.29
Hunan	4.98	4.84	4.92	4.81
Guangdong	5.92	6.23	6.14	6.29
Hainan	0.68	0.64	0.63	0.57
Sichuan	6.76	6.77	6.92	6.81
Guizhou	1.99	1.88	1.75	1.72
Yunnan	1.26	1.14	1.27	1.14
Shaanxi	1.05	1.09	1.22	1.48
Gansu	0.94	0.94	1.12	1.14
Qinghai	0.26	0.25	0.29	0.29
Inner Mongolia Autonomous Region	0.89	0.84	0.68	0.62
Guangxi Zhuang Autonomous Region	2.51	2.42	2.34	2.29
Tibet Autonomous Region	0.10	0.10	0.10	0.10
Ningxia Hui Autonomous Region	0.68	0.69	0.73	0.71
Xinjiang Uygur Autonomous Region	0.58	0.54	0.49	0.48

Table 3 The Admission Proportions for Provinces, Minzu University of China

	Minzu University of China (Beijing)
	2019
Beijing	2.99
Tianjin	1.71
Shanghai	0.53
Chongqing	2.83
Hebei	3.20
Shanxi	1.55
Liaoning	2.83
Jilin	5.61
Heilongjiang	4.06
Jiangsu	1.98
Zhejiang	2.14
Anhui	1.33
Fujian	2.24
Jiangxi	1.28
Shandong	2.78

Henan	2.35
Hubei	3.84
Hunan	4.91
Guangdong	0.96
Hainan	2.56
Sichuan	4.75
Guizhou	4.75
Yunnan	5.13
Shaanxi	1.87
Gansu	3.74
Qinghai	3.95
Inner Mongolia Autonomous Region	5.98
Guangxi Zhuang Autonomous Region	5.71
Tibet Autonomous Region	3.20
Ningxia Hui Autonomous Region	3.52
Xinjiang Uygur Autonomous Region	5.71

Table 4 The Admission Proportions for Provinces, Beijing Normal University

	Beijing Normal University (Beijing)			
	2016	2017	2018	2019
Beijing	5.86	5.84	5.82	5.36
Tianjin	0.80	0.79	0.79	0.61
Shanghai	0.21	0.21	0.21	0.16
Chongqing	6.08	6.05	6.19	4.75
Hebei	3.84	3.68	3.65	2.92
Shanxi	5.12	5.05	5.08	3.90
Liaoning	2.13	2.16	2.22	1.71
Jilin	4.53	4.47	4.49	3.37
Heilongjiang	3.62	3.63	3.54	2.72
Jiangsu	1.71	1.68	1.69	1.30
Zhejiang	2.13	2.16	2.17	1.66
Anhui	1.33	1.79	1.80	2.35
Fujian	5.22	5.16	5.18	3.98
Jiangxi	5.22	5.16	5.18	3.98
Shandong	7.68	7.63	7.67	7.31
Henan	7.46	7.36	7.40	6.50
Hubei	3.20	3.10	3.17	2.44
Hunan	3.36	3.26	3.33	3.49
Guangdong	1.65	2.21	2.22	12.79
Hainan	0.43	0.42	0.42	0.45
Sichuan	7.46	7.52	7.56	9.70
Guizhou	2.56	2.52	2.49	2.68
Yunnan	3.68	3.63	3.65	3.37
Shaanxi	1.87	1.84	1.85	1.46

Gansu	1.12	1.10	1.11	1.34
Qinghai	0.96	0.95	0.95	1.14
Inner Mongolia Autonomous Region	2.45	2.42	1.90	1.46
Guangxi Zhuang Autonomous Region	3.89	3.84	3.86	3.74
Tibet Autonomous Region	0.53	0.53	0.53	0.41
Ningxia Hui Autonomous Region	1.07	1.05	1.06	0.81
Xinjiang Uygur Autonomous Region	2.83	2.79	2.80	2.15

Appendix 7 Gorard Segregation indices of Admissions by Regular HEIs in Provinces 2016-2019

Figure 1 GS indices of HE admissions by regular HEIs in 2016

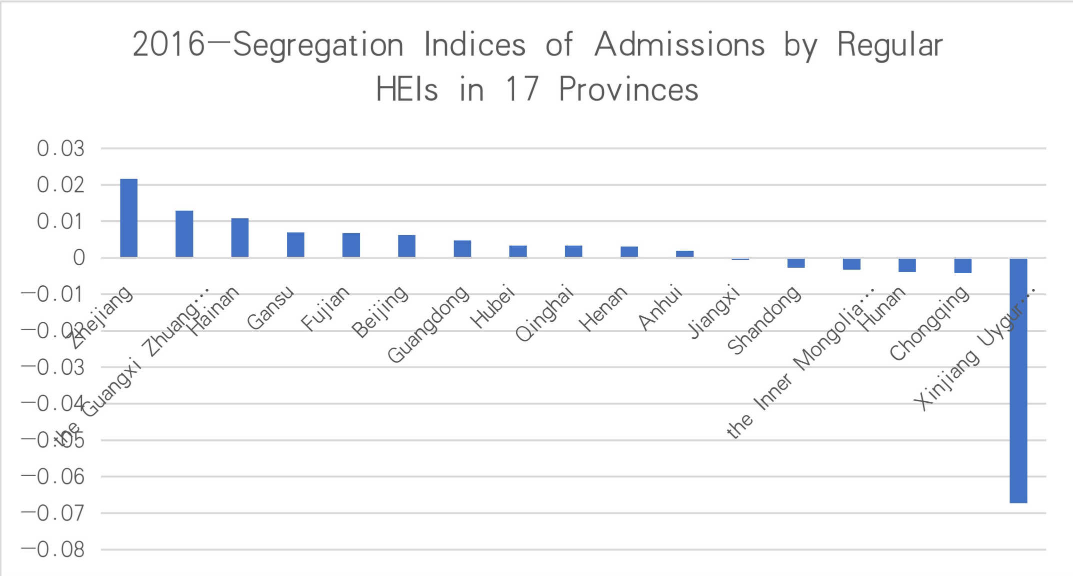


Figure 2 GS indices of HE admissions by regular HEIs in 2017

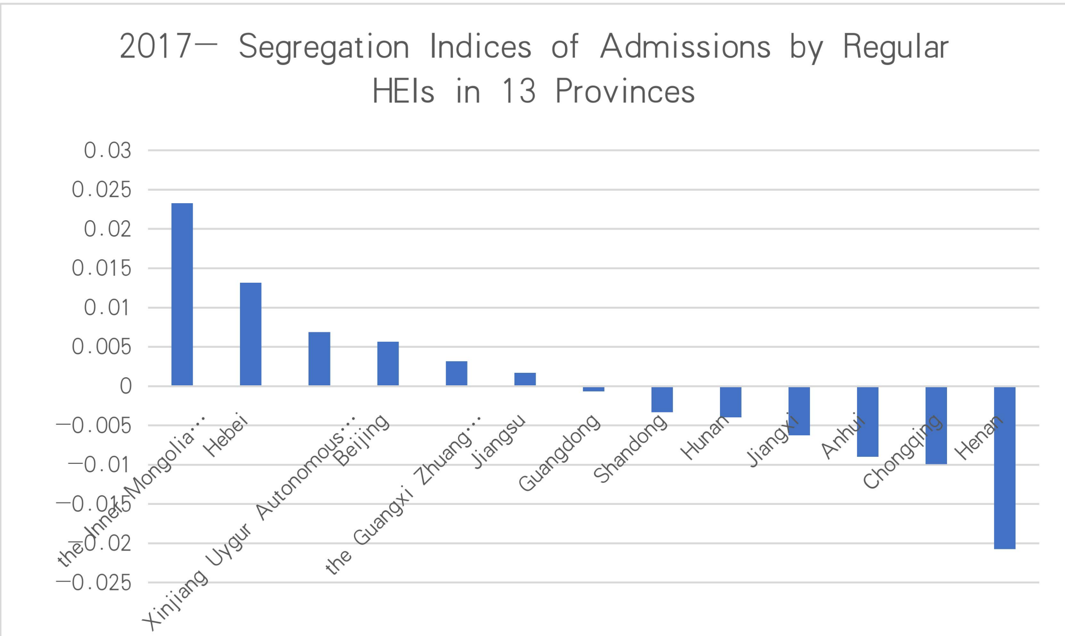


Figure 3 GS indices of HE admissions by regular HEIs in 2018

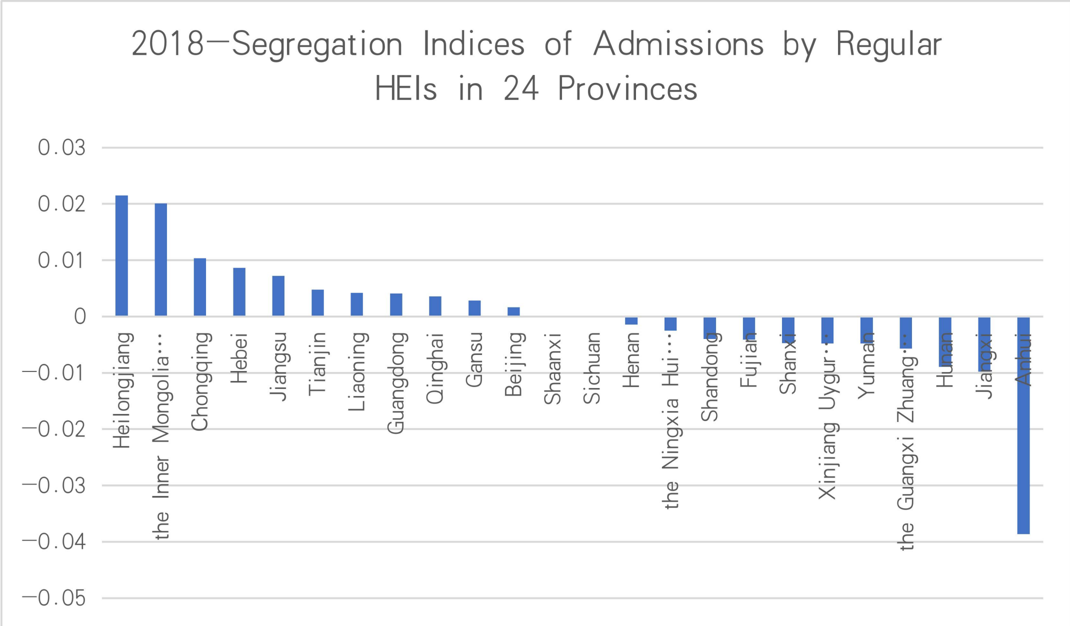
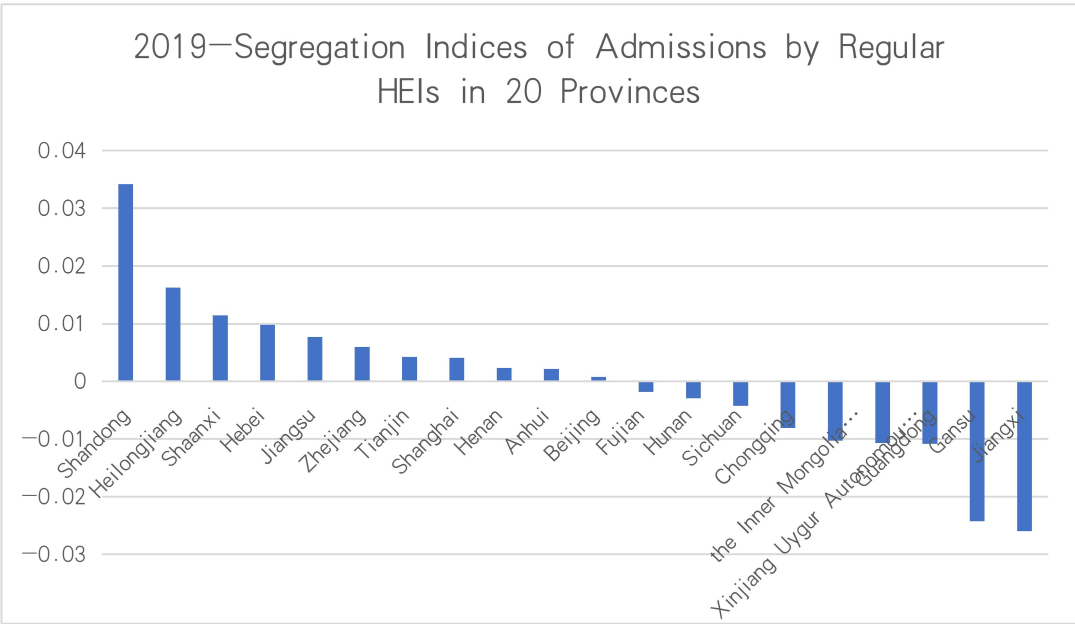
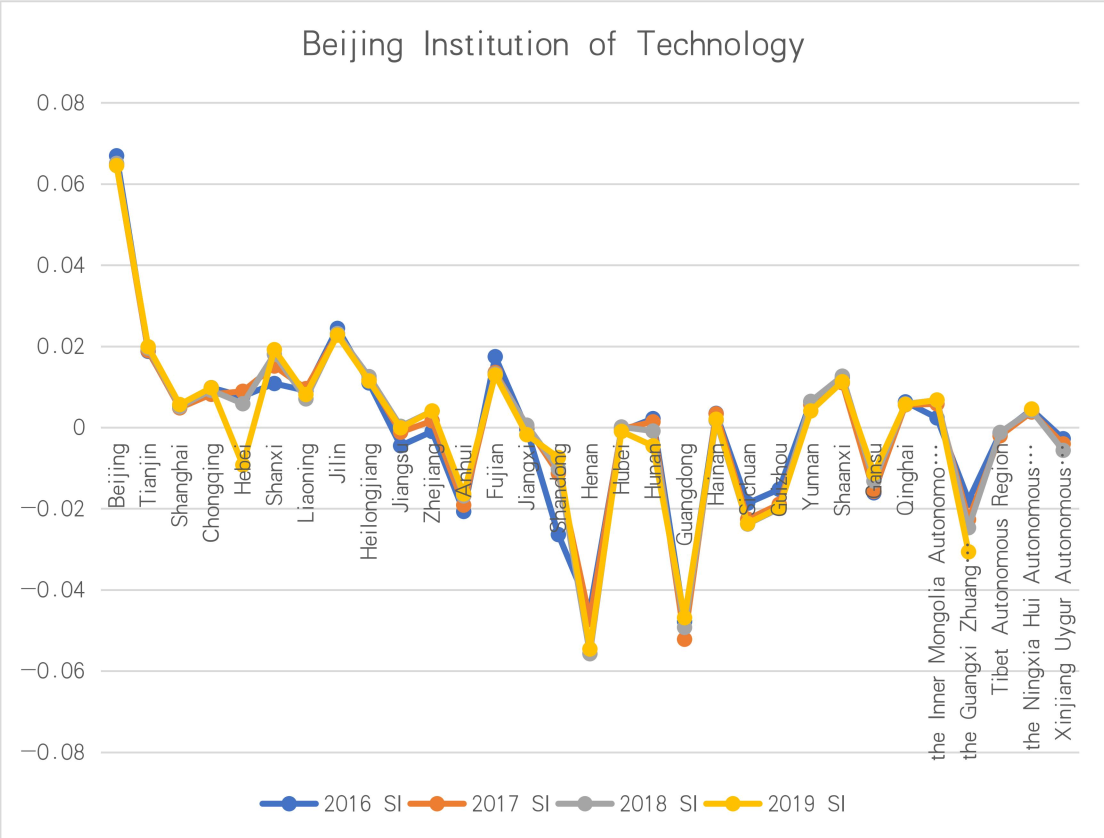
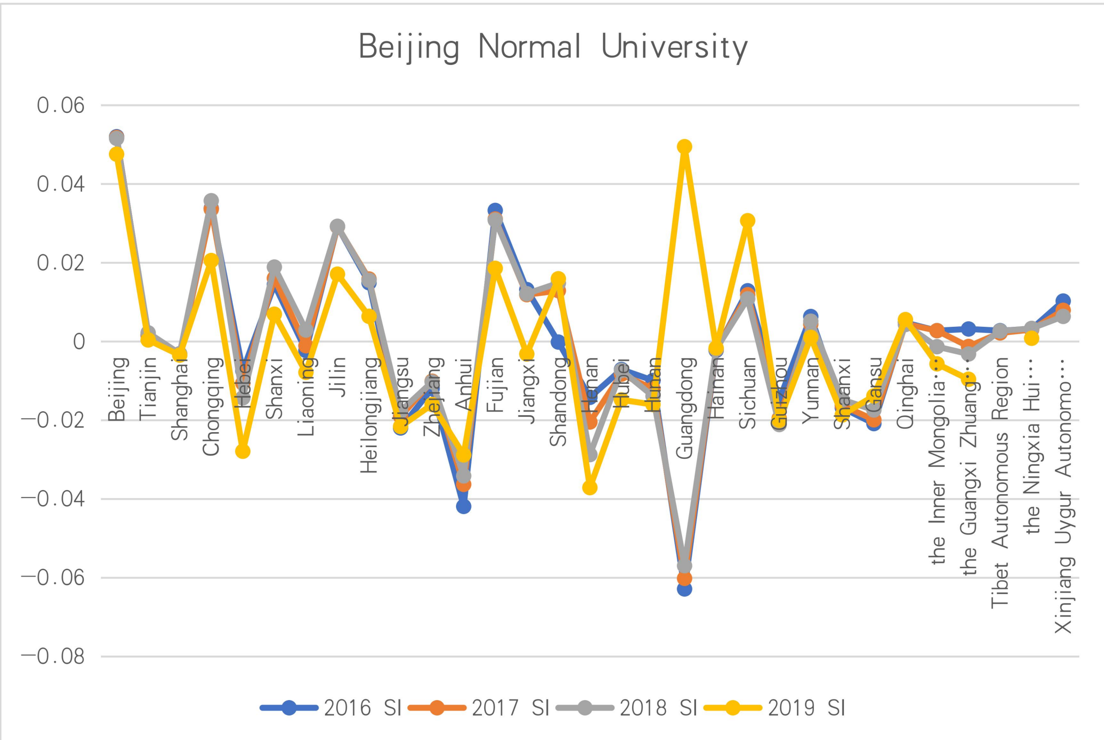


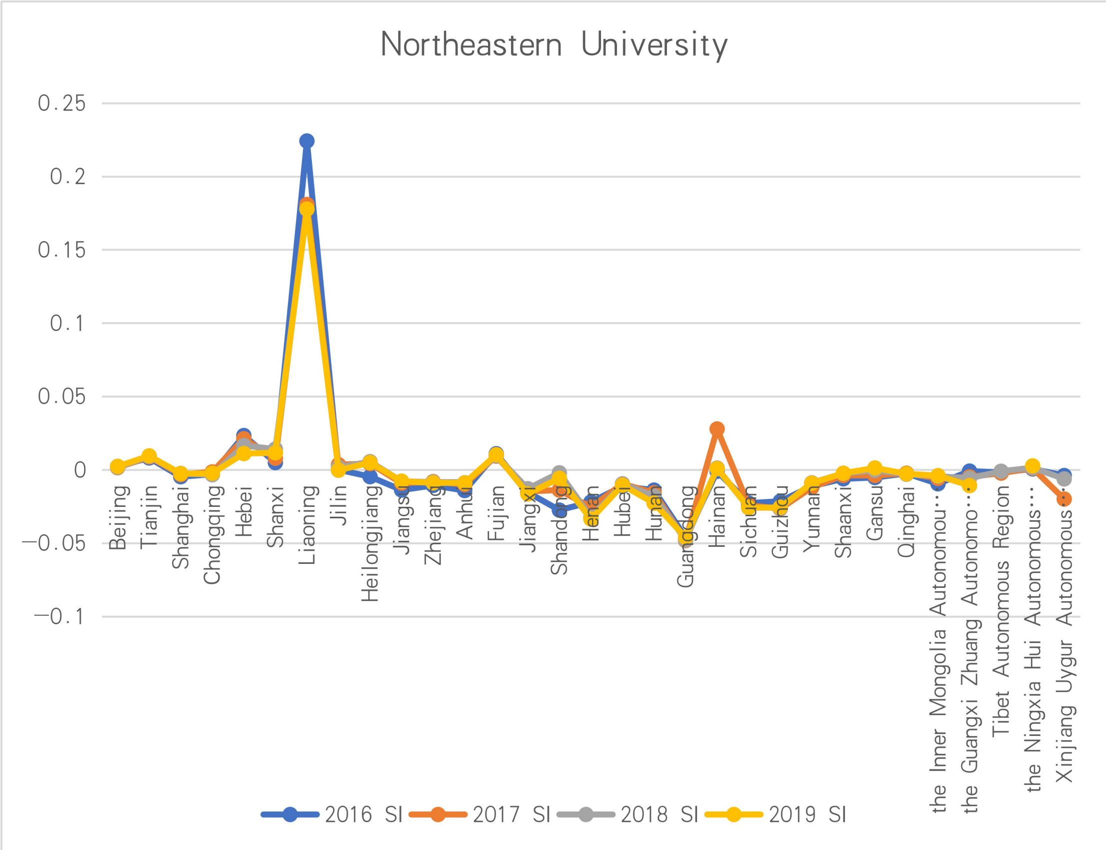
Figure 4 GS indices of HE admissions by regular HEIs in 2019

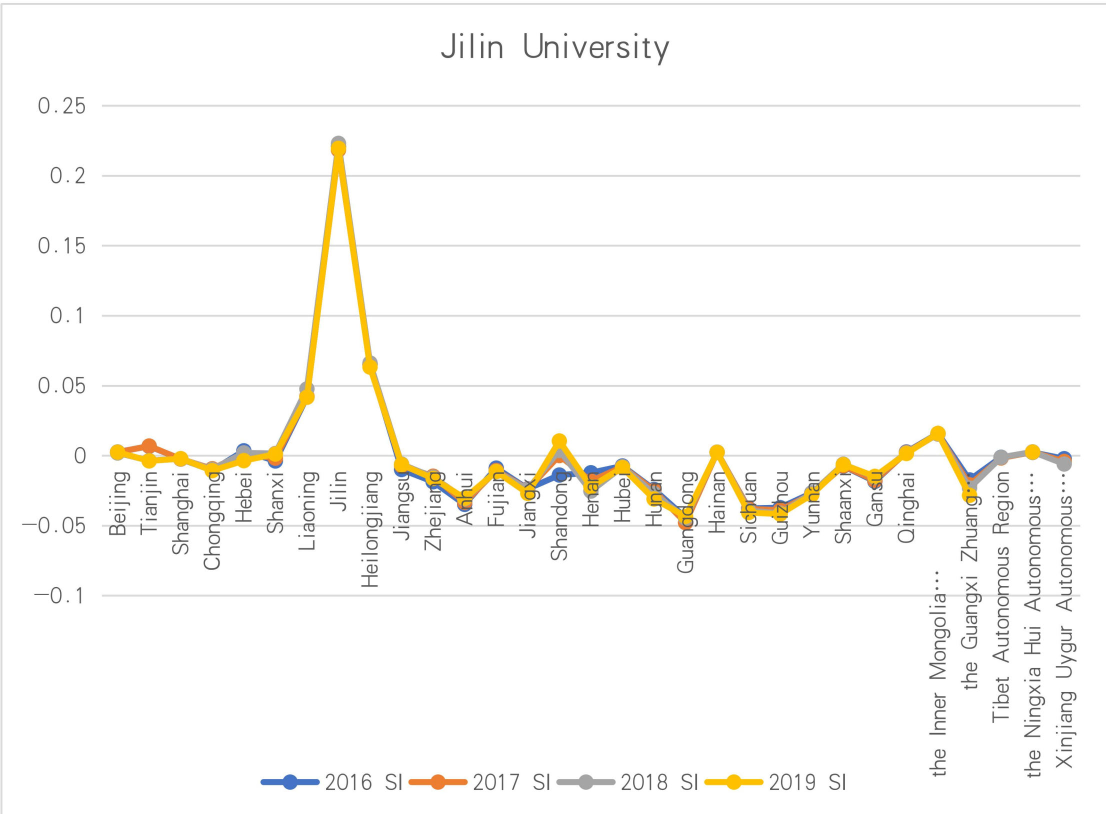


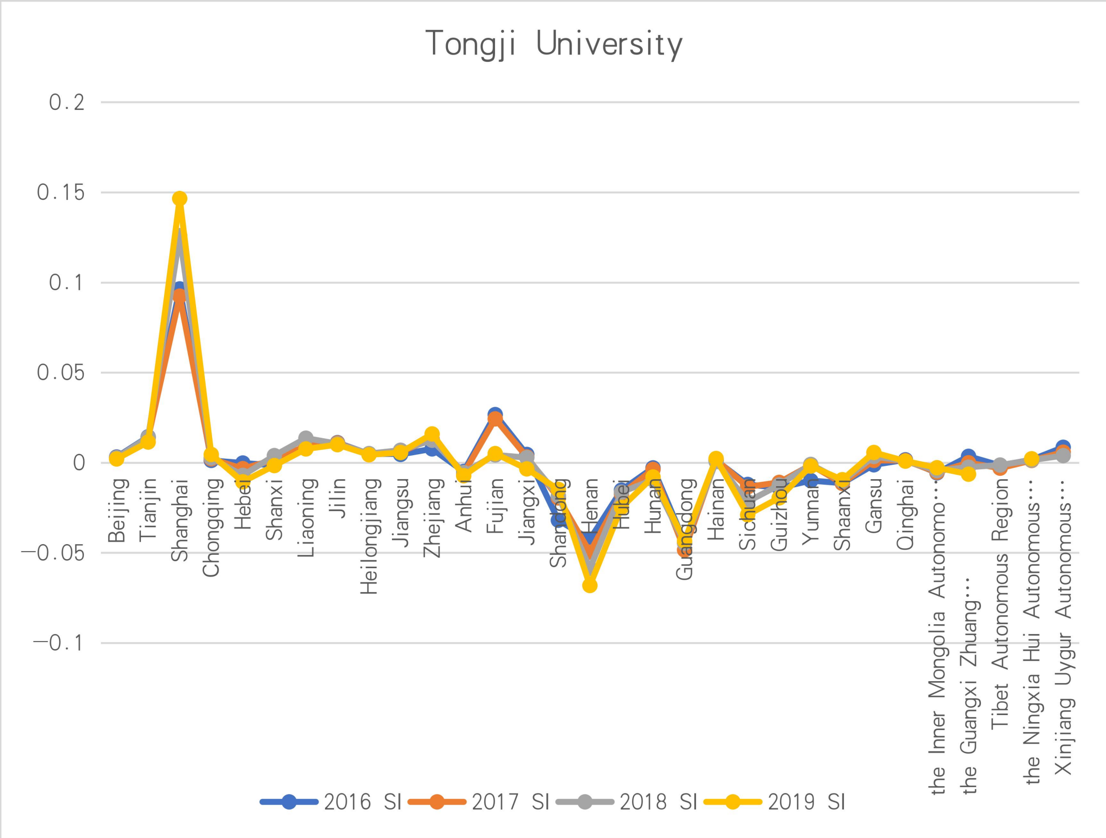
Appendix 8 Gorard segregation indices of DFC universities 2016-2019

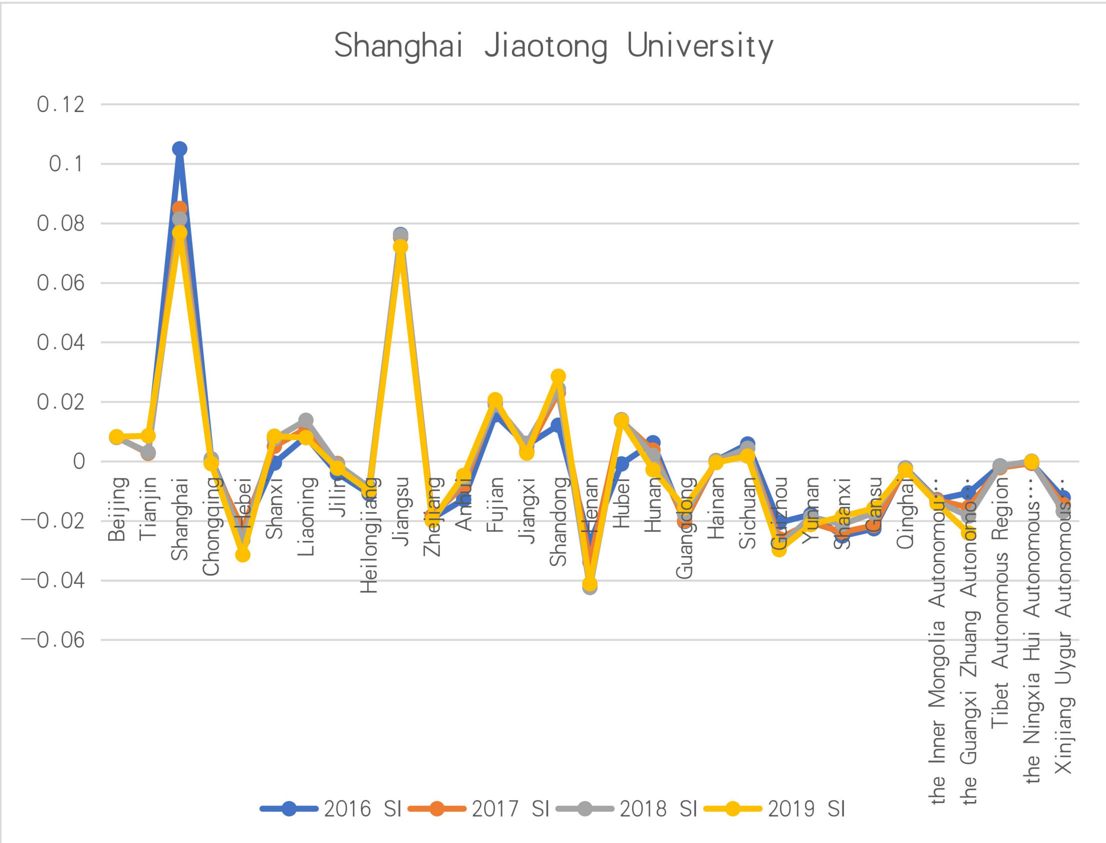


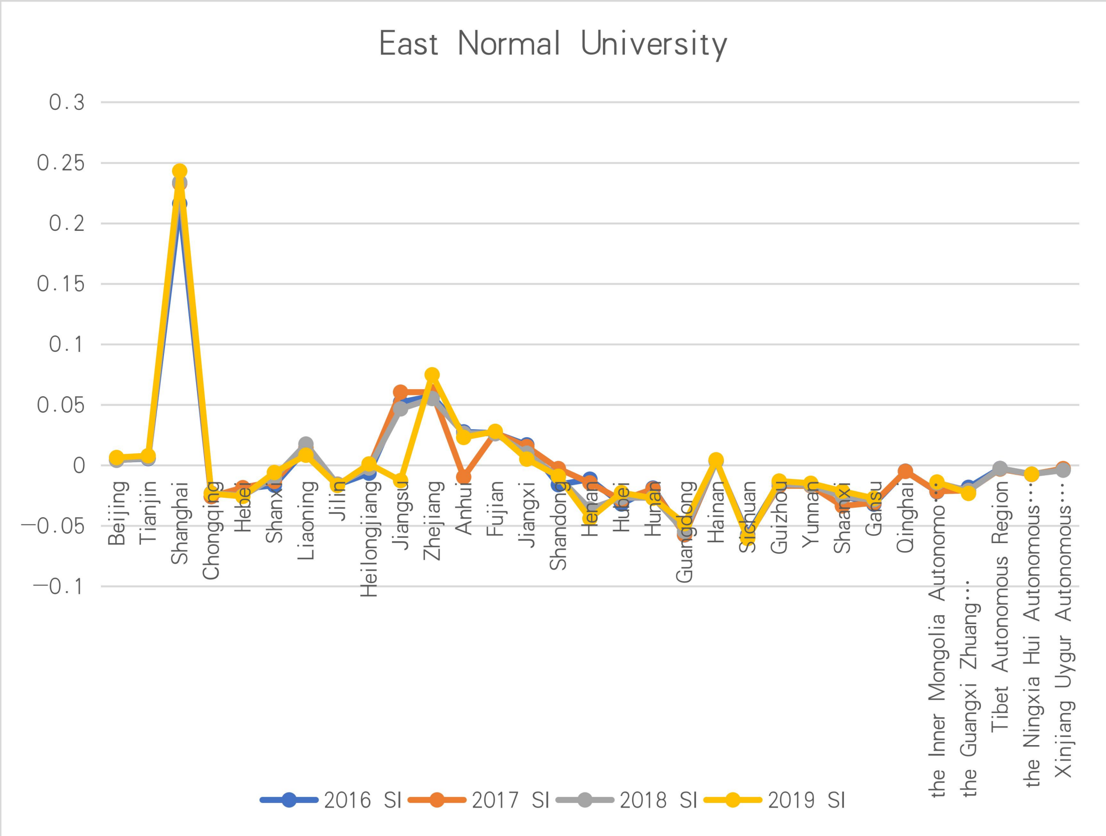


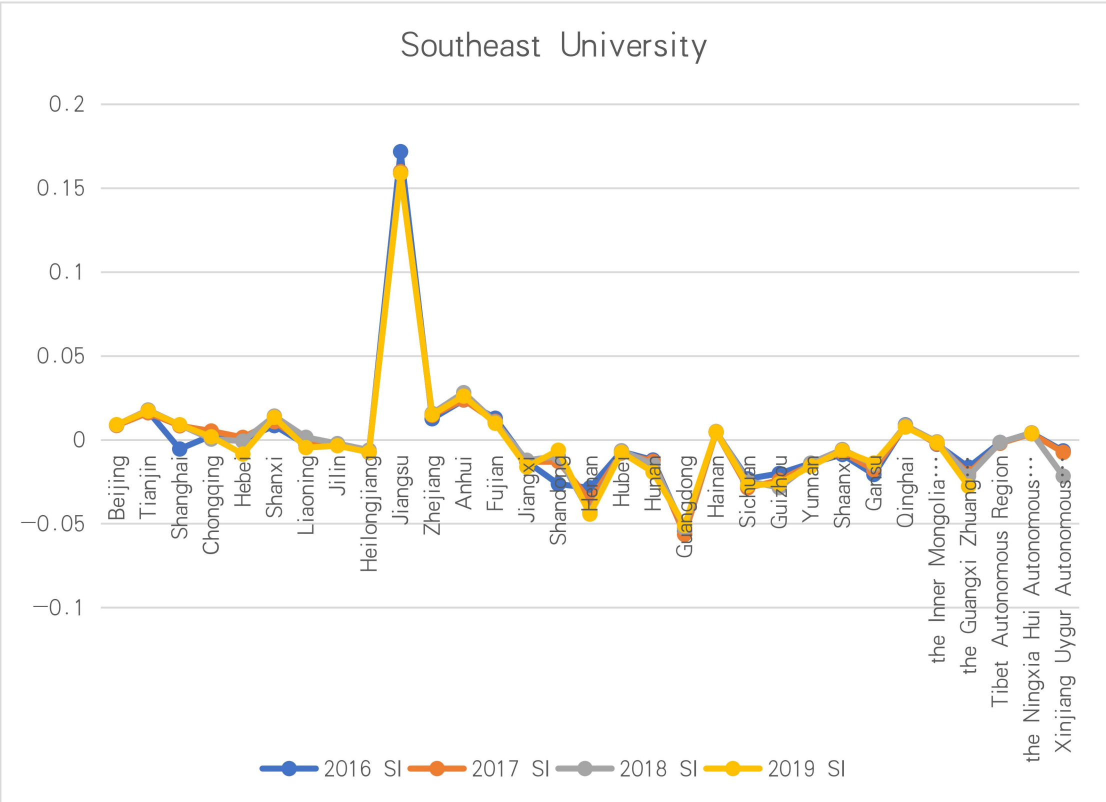

Northeastern University
0.25
0.2
0.15
0.1
0.05
0
−0.05
−0.1
Beijing
Tianjin
Shanghai
Chongqing
Hebei
Shanxi
Liaoning
Jilin
Heilongjiang
Fujian
Guangdong
Yunnan
Gansu
Qinghai
the Inner Mongolia Autonomou…
the Guangxi Zhuang Autonomo…
Tibet Autonomous Region
the Ningxia Hui Autonomous…
Xinjiang Uygur Autonomous…
2016 SI
2017 SI
2018 SI
2019 SI

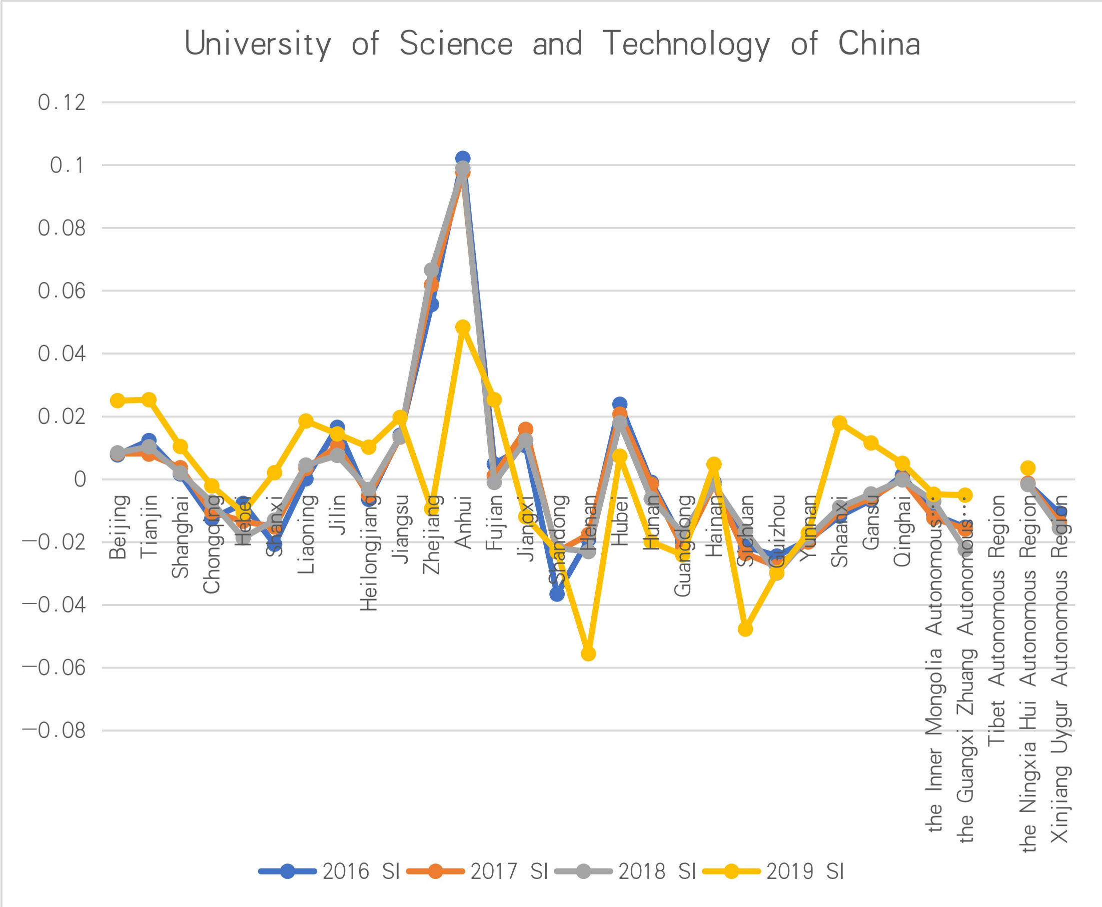

Jilin University
0.25
0.2
0.15
0.1
0.05
0
−0.05
−0.1
Beijing
Tianjin
Shanghai
Chongqing
Hebei
Shanxi
Liaoning
Jilin
Heilongjiang
Fujian
Shandong
Hubei
Guangdong
Hainan
Yunnan
Shaanxi
Qinghai
the Inner Mongolia…
the Guangxi Zhuang…
Tibet Autonomous Region
the Ningxia Hui Autonomous…
Xinjiang Uygur Autonomous…
2016 SI
2017 SI
2018 SI
2019 SI

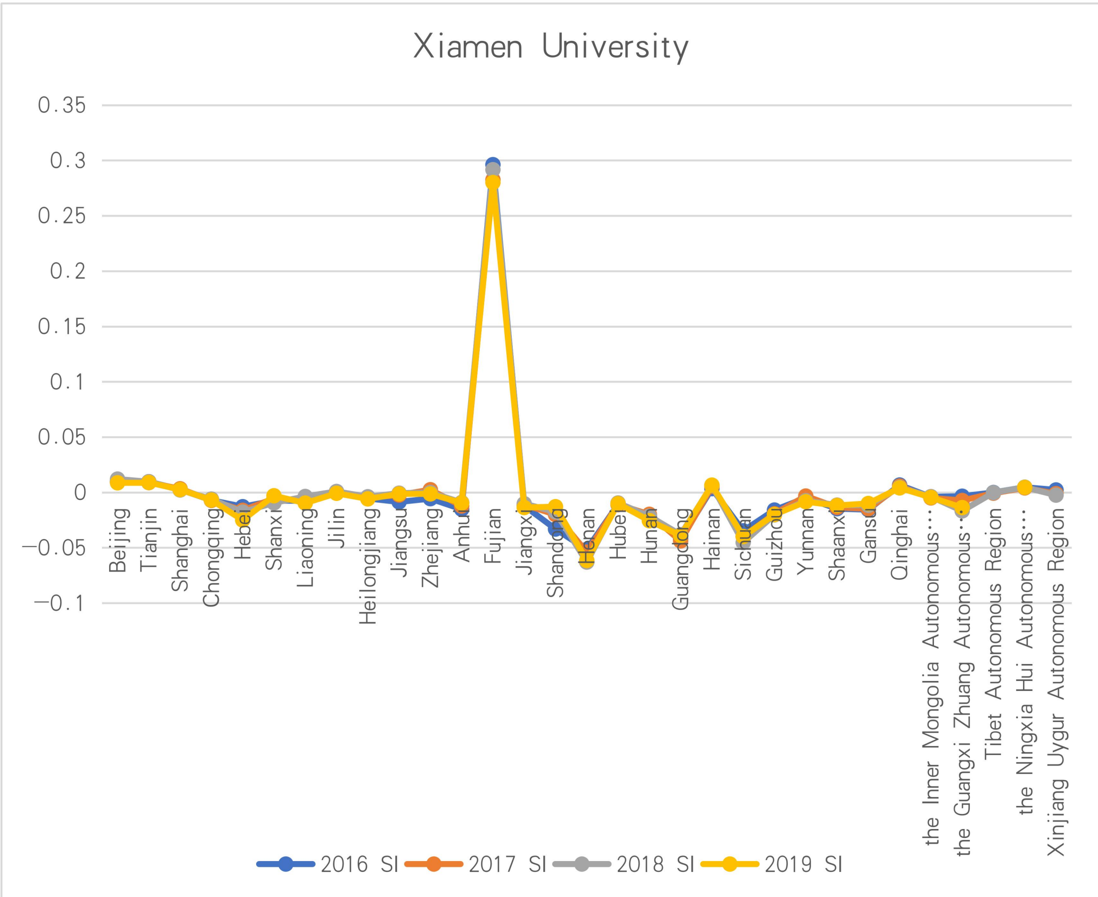

Tongji University
0.2
0.15
0.1
0.05
0
−0.05
−0.1
Beijing
Tianjin
Shanghai
Chongqing
Hebei
Shanxi
Liaoning
Jilin
Heilongjiang
Jiangsu
Zhejiang
Anhui
Fujian
Jiangxi
Shandong
Henan
Hubei
Hunan
Guangdong
Hainan
Sichuan
Guizhou
Yunnan
Shaanxi
Gansu
Qinghai
the Inner Mongolia Autonomo…
the Guangxi Zhuang…
Tibet Autonomous Region
the Ningxia Hui Autonomous…
Xinjiang Uygur Autonomous…
2016 SI
2017 SI
2018 SI
2019 SI

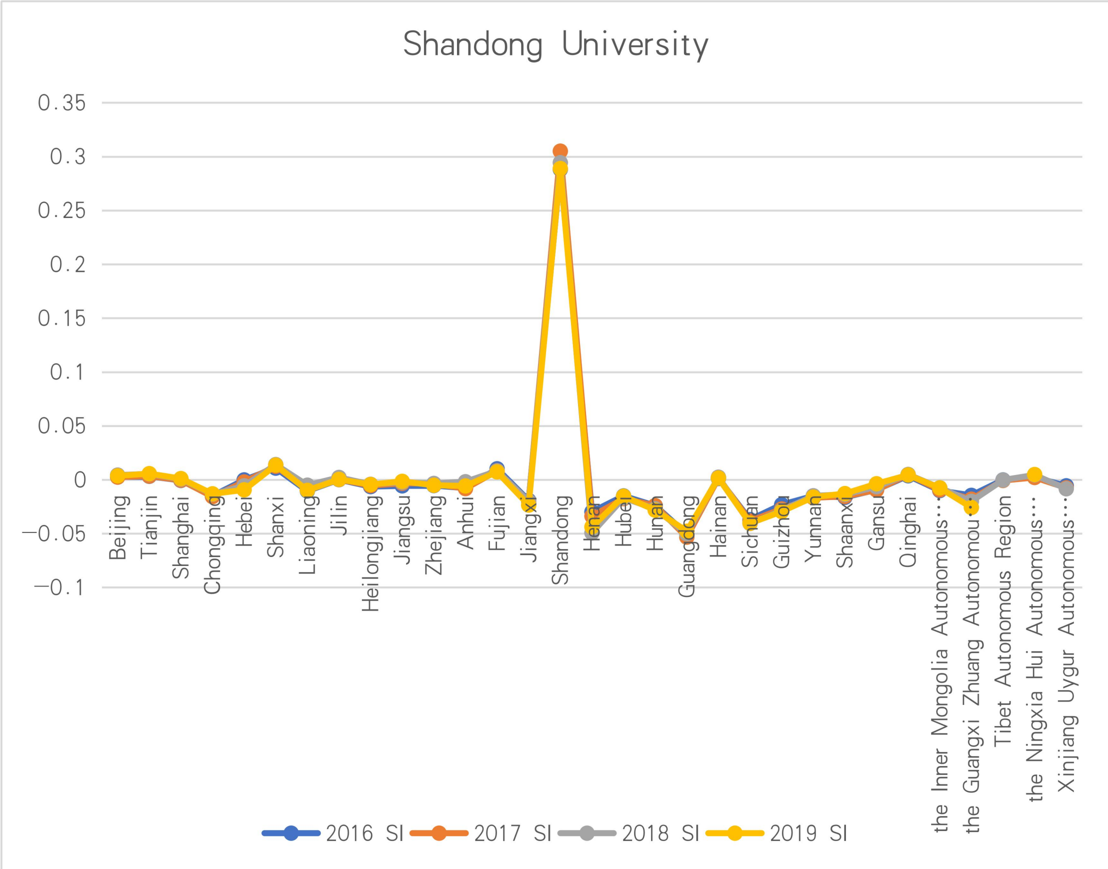

Shanghai Jiaotong University
0.12
0.1
0.08
0.06
0.04
0.02
0
−0.02
−0.04
−0.06
Beijing
Tianjin
Shanghai
Chongqing
Hebei
Shanxi
Liaoning
Jilin
Heilongjiang
Jiangsu
Zhejiang
Anhui
Fujian
Jiangxi
Shandong
Henan
Hubei
Hunan
Guangdong
Hainan
Sichuan
Guizhou
Yunnan
Shaanxi
Gansu
Qinghai
the Inner Mongolia Autonomous…
the Guangxi Zhuang Autonomous…
Tibet Autonomous Region
the Ningxia Hui Autonomous…
Xinjiang Uygur Autonomous…
2016 SI
2017 SI
2018 SI
2019 SI

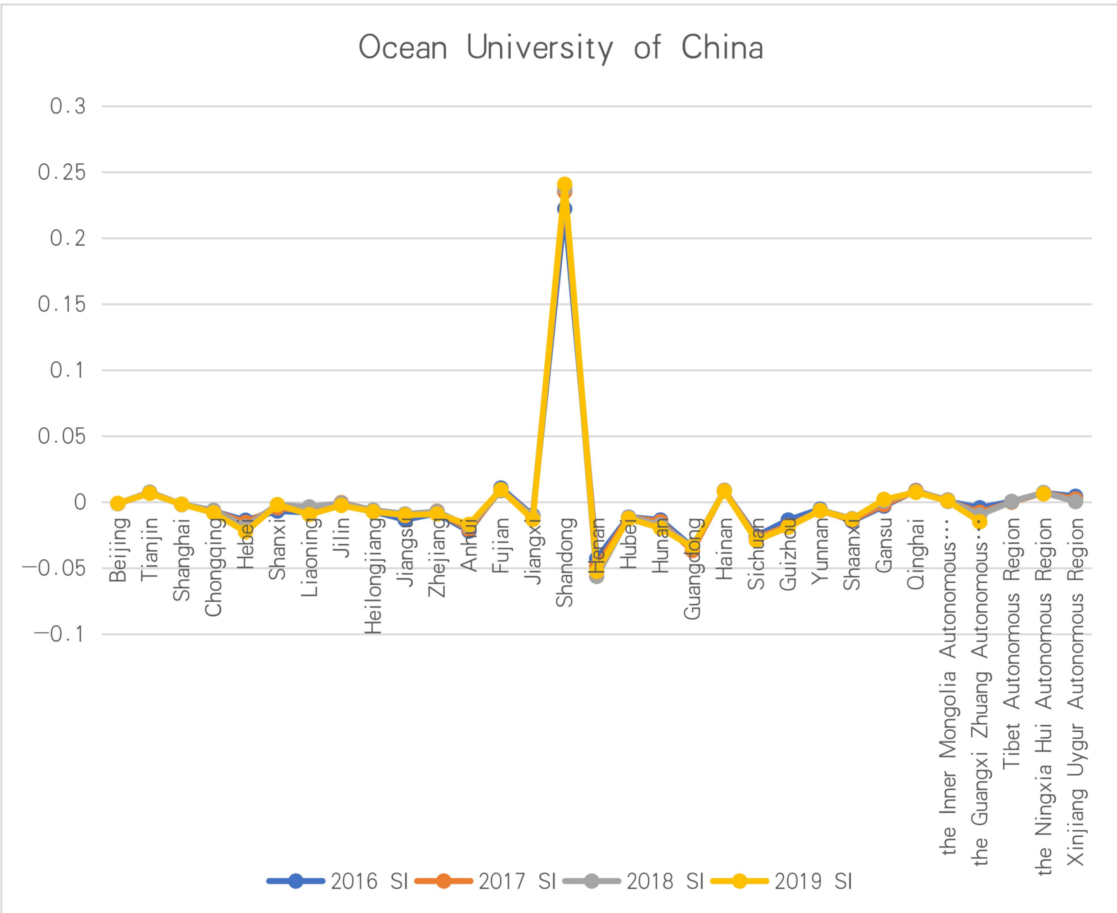

East Normal University
0.3
0.25
0.2
0.15
0.1
0.05
0
−0.05
−0.1
Beijing
Tianjin
Shanghai
Chongqing
Hebei
Shanxi
Liaoning
Jilin
Heilongjiang
Jiangsu
Zhejiang
Anhui
Fujian
Jiangxi
Shandong
Henan
Hubei
Hunan
Guangdong
Hainan
Sichuan
Guizhou
Yunnan
Shaanxi
Gansu
Qinghai
the Inner Mongolia Autonomo…
the Guangxi Zhuang…
Tibet Autonomous Region
the Ningxia Hui Autonomous…
Xinjiang Uygur Autonomous…
2016 SI
2017 SI
2018 SI
2019 SI

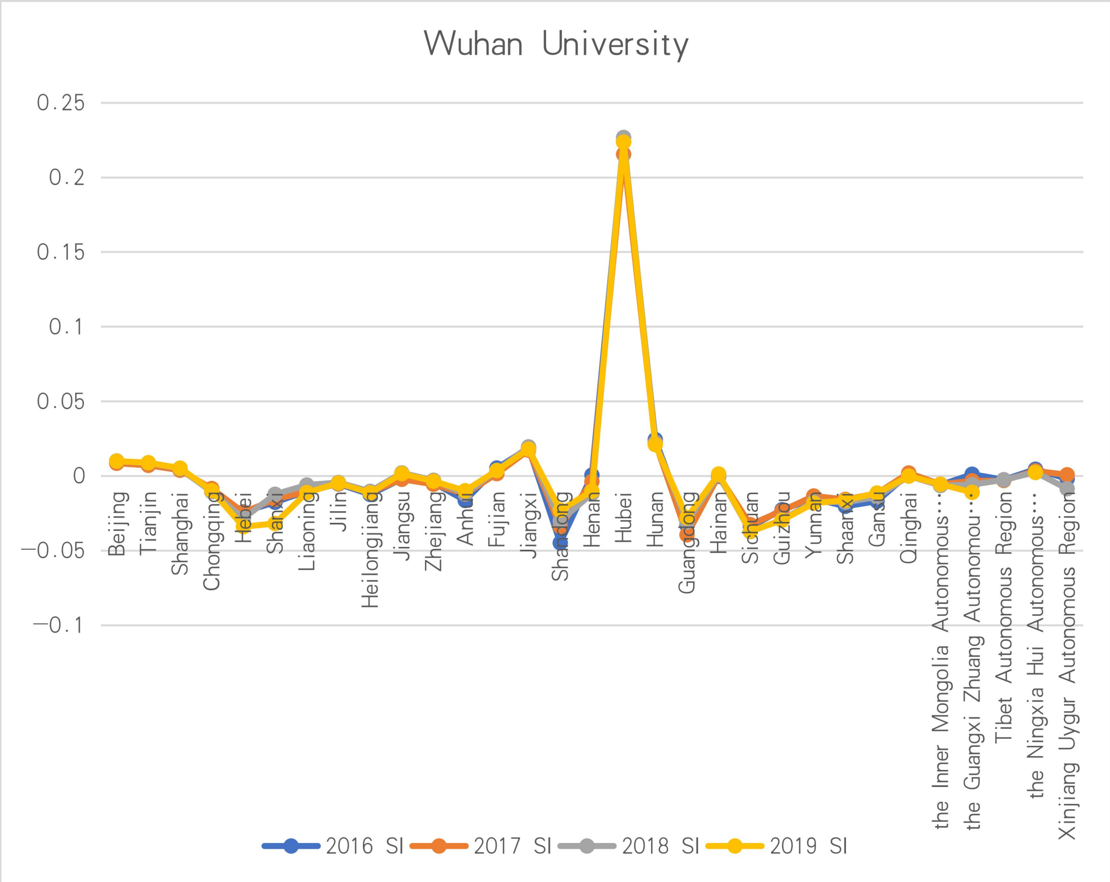

Southeast University
0.2
0.15
0.1
0.05
0
−0.05
−0.1
Beijing
Tianjin
Shanghai
Chongqing
Hebei
Shanxi
Liaoning
Jilin
Heilongjiang
Jiangsu
Zhejiang
Anhui
Fujian
Jiangxi
Shandong
Henan
Hubei
Hunan
Guangdong
Hainan
Sichuan
Guizhou
Yunnan
Shaanxi
Gansu
Qinghai
the Inner Mongolia…
the Guangxi Zhuang…
Tibet Autonomous Region
the Ningxia Hui Autonomous…
Xinjiang Uygur Autonomous…
2016 SI
2017 SI
2018 SI
2019 SI


University of Science and Technology of China
0.12
0.1
0.08
0.06
0.04
0.02
0
−0.02
−0.04
−0.06
−0.08
Beijing
Tianjin
Shanghai
Chongqing
Hebei
Shanxi
Liaoning
Jilin
Heilongjiang
Jiangsu
Zhejiang
Anhui
Fujian
Jiangxi
Shandong
Henan
Hubei
Hunan
Guangdong
Hainan
Sichuan
Guizhou
Yunnan
Shaanxi
Gansu
Qinghai
the Inner Mongolia Autonomous…
the Guangxi Zhuang Autonomous…
Tibet Autonomous Region
the Ningxia Hui Autonomous Region
Xinjiang Uygur Autonomous Region
2016 SI
2017 SI
2018 SI
2019 SI

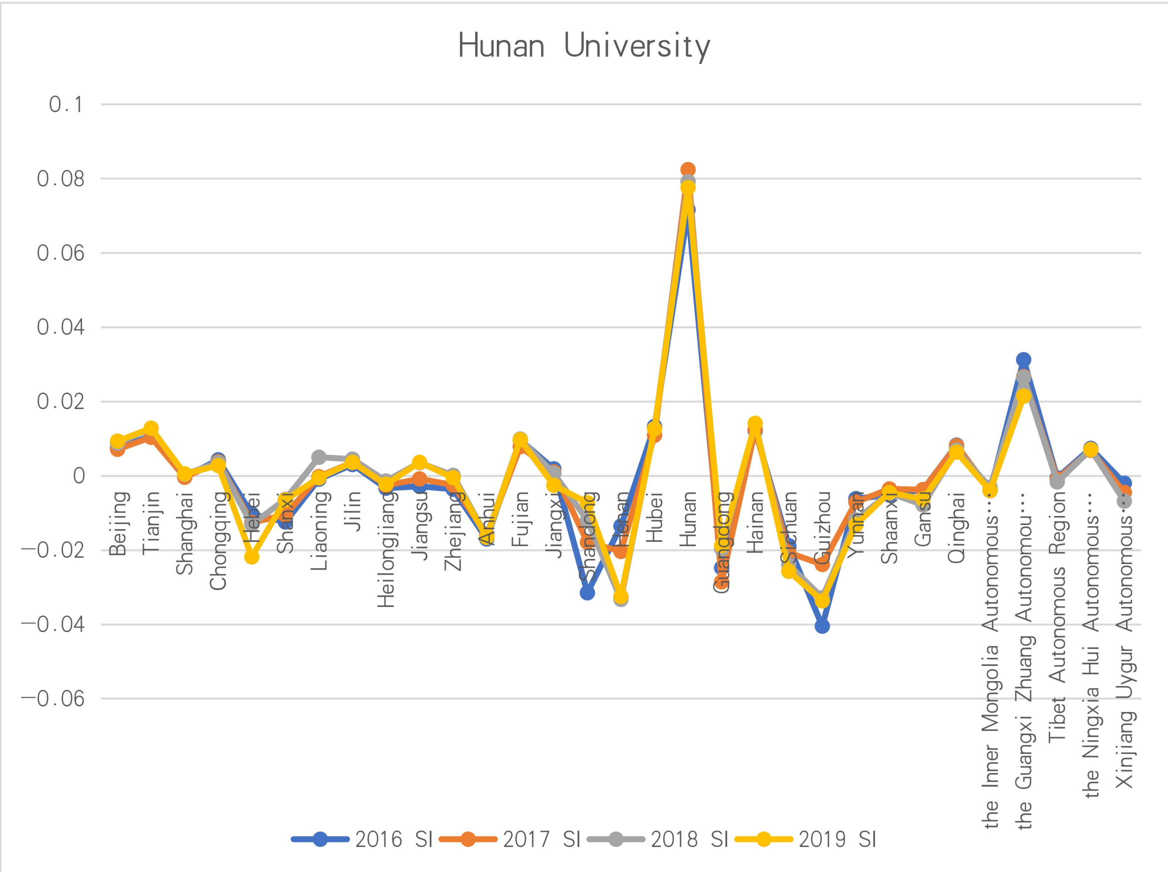

Xiamen University
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
−0.05
−0.1
Beijing
Tianjin
Shanghai
Chongqing
Hebei
Shanxi
Liaoning
Jilin
Heilongjiang
Jiangsu
Zhejiang
Anhui
Fujian
Jiangxi
Shandong
Henan
Hubei
Hunan
Guangdong
Hainan
Sichuan
Guizhou
Yunnan
Shaanxi
Gansu
Qinghai
the Inner Mongolia Autonomous…
the Guangxi Zhuang Autonomous…
Tibet Autonomous Region
the Ningxia Hui Autonomous…
Xinjiang Uygur Autonomous Region
2016 SI
2017 SI
2018 SI
2019 SI


Shandong University
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
−0.05
−0.1
Beijing
Tianjin
Shanghai
Chongqing
Hebei
Shanxi
Liaoning
Jilin
Heilongjiang
Jiangsu
Zhejiang
Anhui
Fujian
Jiangxi
Shandong
Henan
Hubei
Hunan
Guangdong
Hainan
Sichuan
Guizhou
Yunnan
Shaanxi
Gansu
Qinghai
the Inner Mongolia Autonomous…
the Guangxi Zhuang Autonomou…
Tibet Autonomous Region
the Ningxia Hui Autonomous…
Xinjiang Uygur Autonomous…
2016 SI
2017 SI
2018 SI
2019 SI

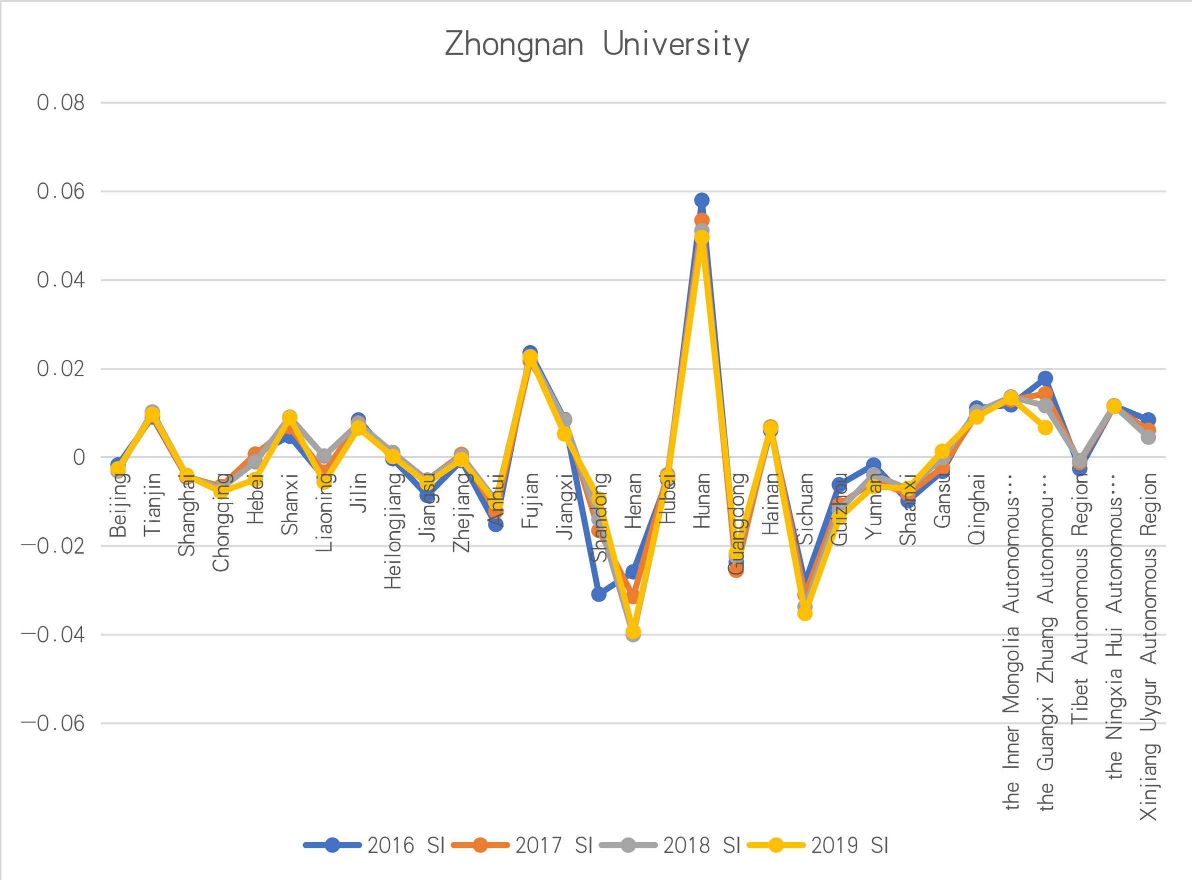

Ocean University of China
0.3
0.25
0.2
0.15
0.1
0.05
0
−0.05
−0.1
Beijing
Tianjin
Shanghai
Chongqing
Hebei
Shanxi
Liaoning
Jilin
Heilongjiang
Jiangsu
Zhejiang
Anhui
Fujian
Jiangxi
Shandong
Henan
Hubei
Hunan
Guangdong
Hainan
Sichuan
Guizhou
Yunnan
Shaanxi
Gansu
Qinghai
the Inner Mongolia Autonomous…
the Guangxi Zhuang Autonomous…
Tibet Autonomous Region
the Ningxia Hui Autonomous Region
Xinjiang Uygur Autonomous Region
2016 SI
2017 SI
2018 SI
2019 SI

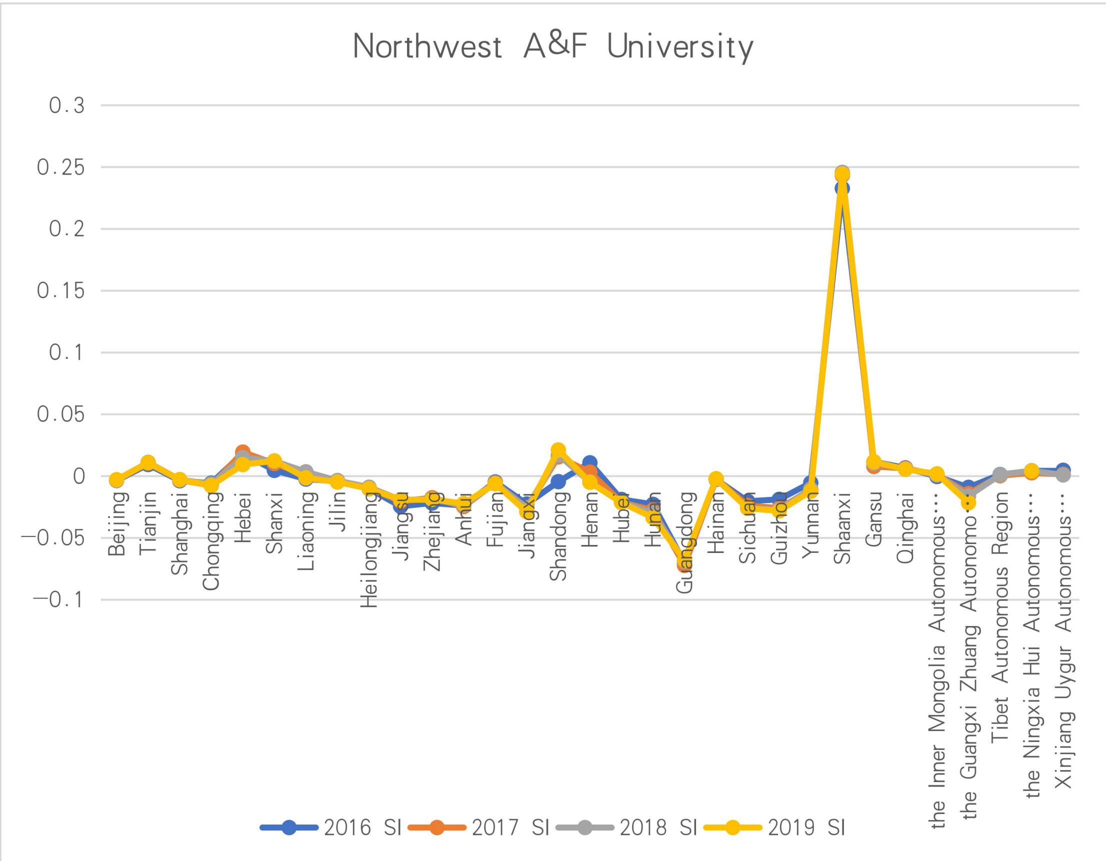


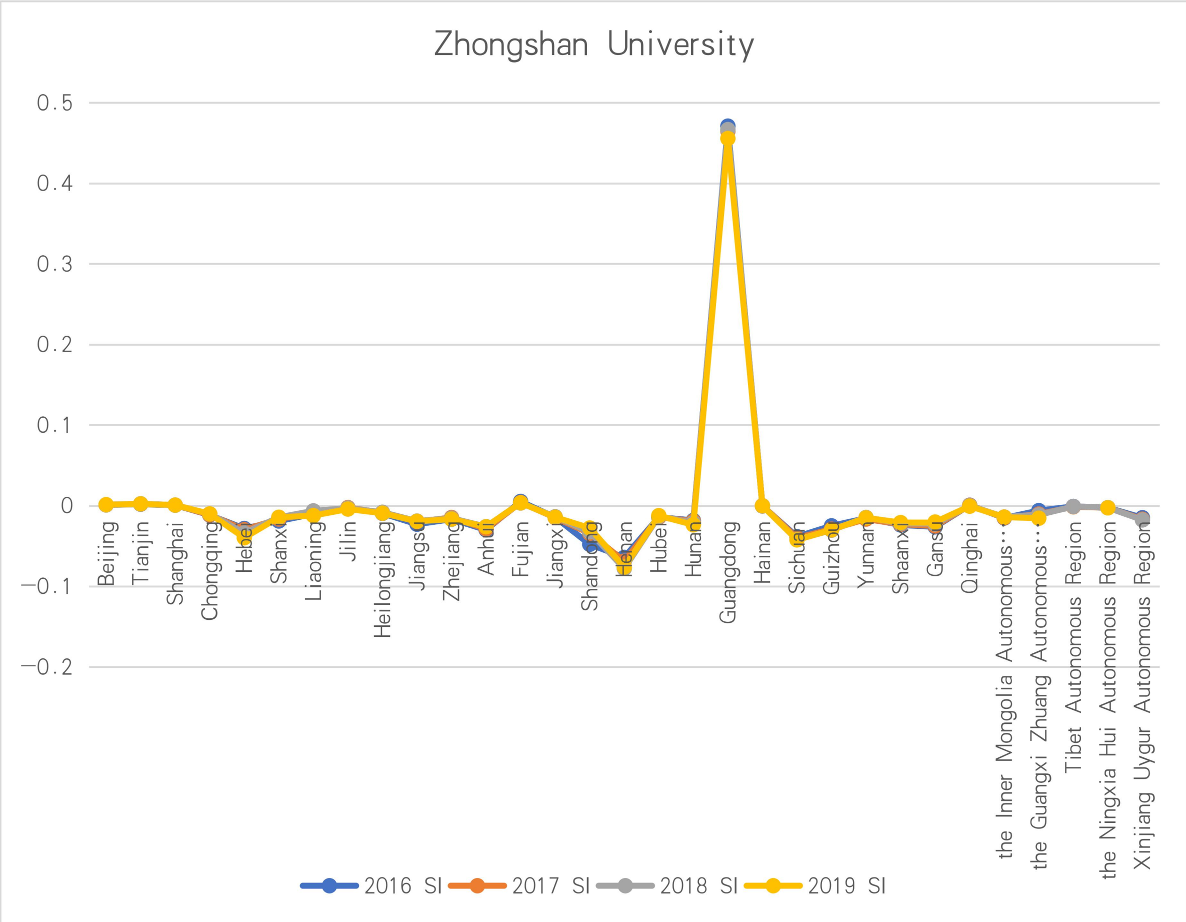
Wuhan University
0.25
0.2
0.15
0.1
0.05
0
−0.05
−0.1
Beijing
Tianjin
Shanghai
Chongqing
Hebei
Shanxi
Liaoning
Jilin
Heilongjiang
Jiangsu
Zhejiang
Anhui
Fujian
Jiangxi
Shandong
Henan
Hubei
Hunan
Guangdong
Hainan
Sichuan
Guizhou
Yunnan
Shaanxi
Gansu
Qinghai
the Inner Mongolia Autonomous…
the Guangxi Zhuang Autonomou…
Tibet Autonomous Region
the Ningxia Hui Autonomous…
Xinjiang Uygur Autonomous Region
2016 SI
2017 SI
2018 SI
2019 SI

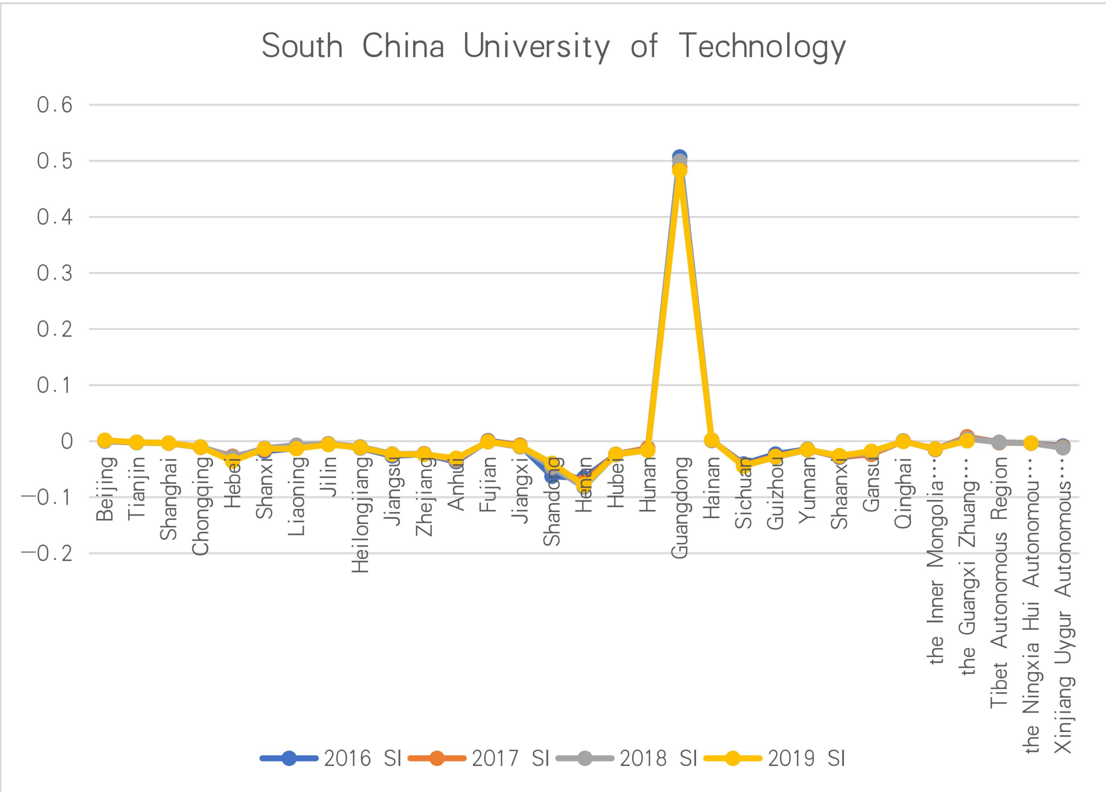


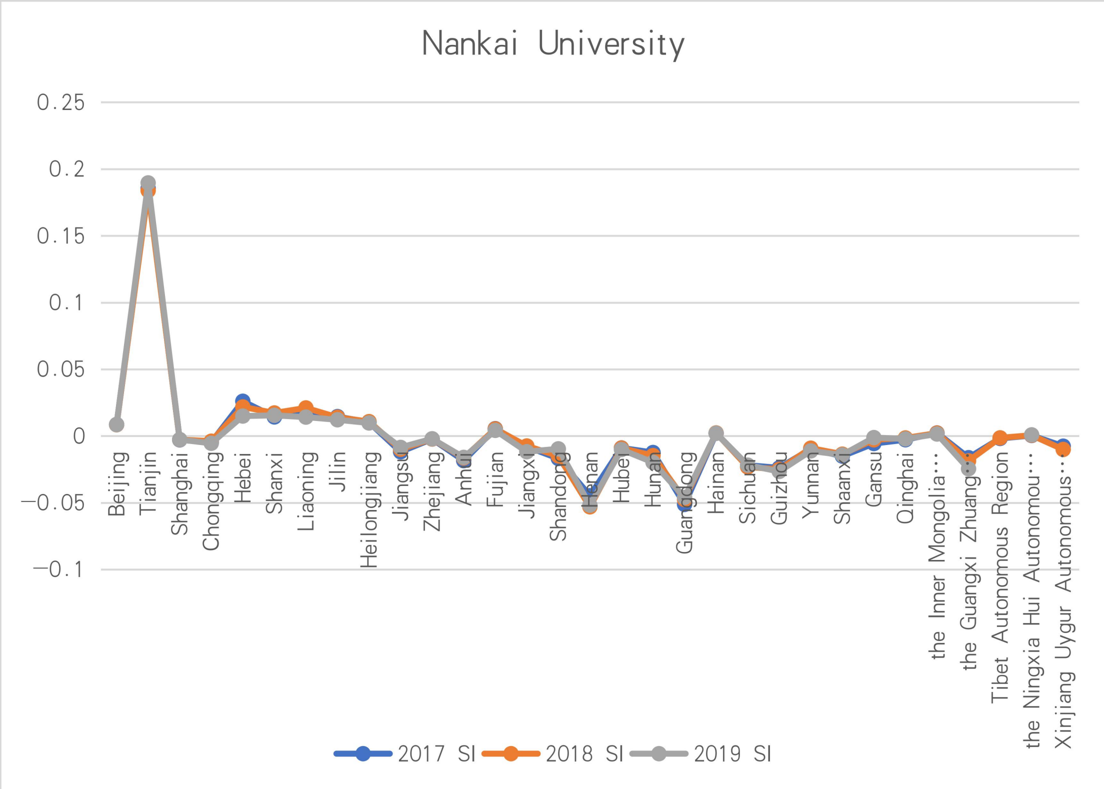
Hunan University
0.1
0.08
0.06
0.04
0.02
0
−0.02
−0.04
−0.06
Beijing
Tianjin
Shanghai
Chongqing
Hebei
Shanxi
Liaoning
Jilin
Heilongjiang
Jiangsu
Zhejiang
Anhui
Fujian
Jiangxi
Shandong
Henan
Hubei
Hunan
Guangdong
Hainan
Sichuan
Guizhou
Yunnan
Shaanxi
Gansu
Qinghai
the Inner Mongolia Autonomous…
the Guangxi Zhuang Autonomou…
Tibet Autonomous Region
the Ningxia Hui Autonomous…
Xinjiang Uygur Autonomous…
2016 SI
2017 SI
2018 SI
2019 SI

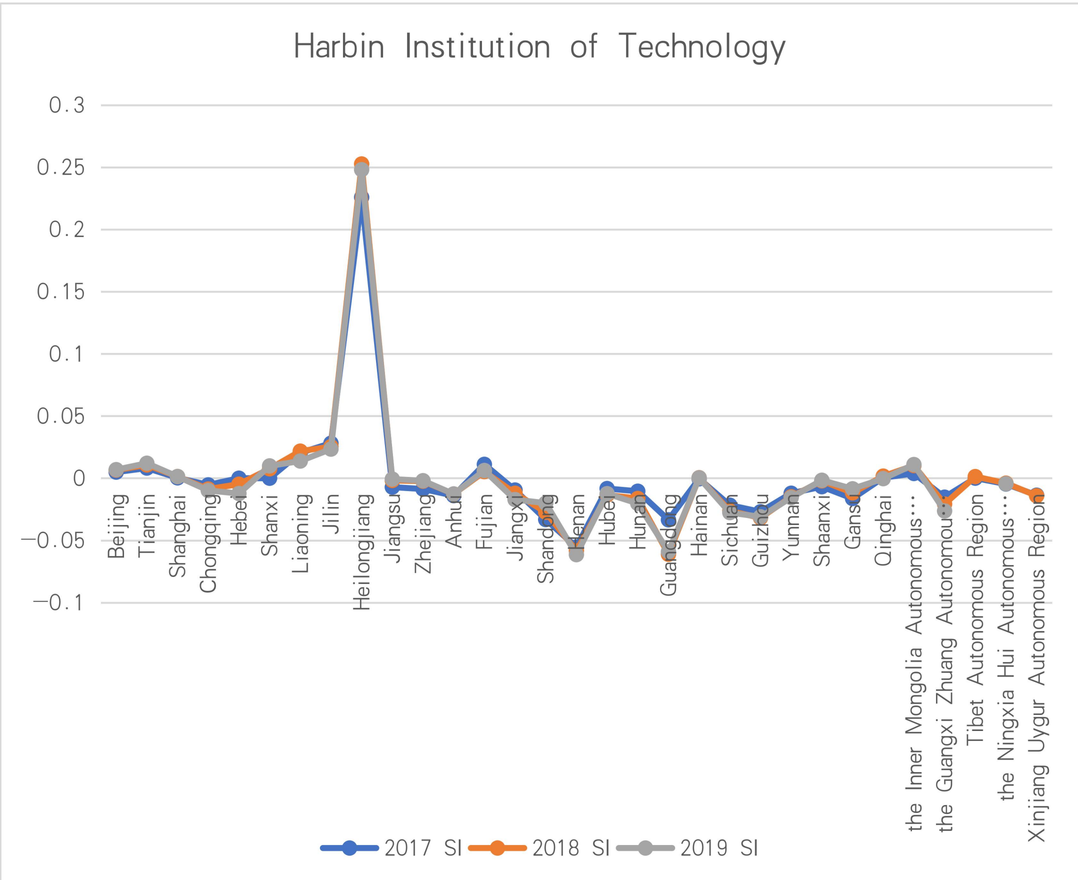

Zhongnan University
0.08
0.06
0.04
0.02
0
−0.02
−0.04
−0.06
Beijing
Tianjin
Shanghai
Chongqing
Hebei
Shanxi
Liaoning
Jilin
Heilongjiang
Jiangsu
Zhejiang
Anhui
Fujian
Jiangxi
Shandong
Henan
Hubei
Hunan
Guangdong
Hainan
Sichuan
Guizhou
Yunnan
Shaanxi
Gansu
Qinghai
the Inner Mongolia Autonomous…
the Guangxi Zhuang Autonomou…
Tibet Autonomous Region
the Ningxia Hui Autonomous…
Xinjiang Uygur Autonomous Region
2016 SI
2017 SI
2018 SI
2019 SI

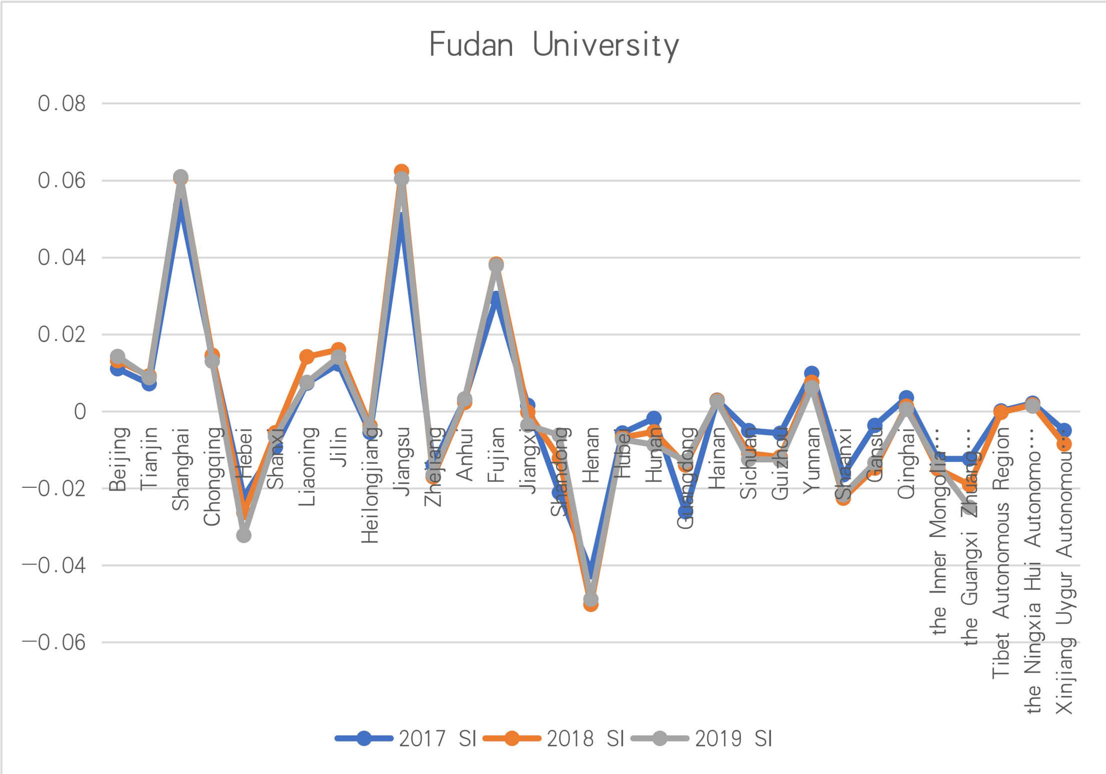

Northwest A&F University
0.3
0.25
0.2
0.15
0.1
0.05
0
−0.05
−0.1
Beijing
Tianjin
Shanghai
Chongqing
Hebei
Shanxi
Liaoning
Jilin
Heilongjiang
Jiangsu
Zhejiang
Anhui
Fujian
Jiangxi
Shandong
Henan
Hubei
Hunan
Guangdong
Hainan
Sichuan
Guizhou
Yunnan
Shaanxi
Gansu
Qinghai
the Inner Mongolia Autonomous…
the Guangxi Zhuang Autonomo…
Tibet Autonomous Region
the Ningxia Hui Autonomous…
Xinjiang Uygur Autonomous…
2016 SI
2017 SI
2018 SI
2019 SI

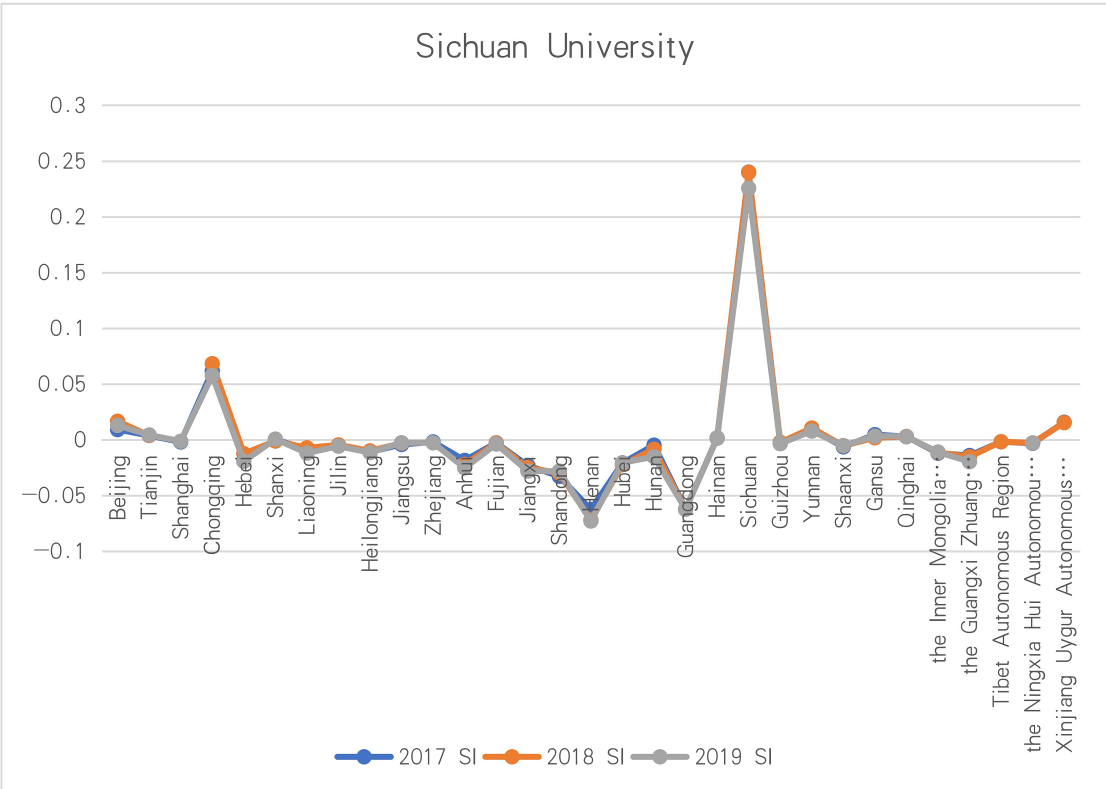

Zhongshan University
0.5
0.4
0.3
0.2
0.1
0
−0.1
−0.2
Beijing
Tianjin
Shanghai
Chongqing
Hebei
Shanxi
Liaoning
Jilin
Heilongjiang
Jiangsu
Zhejiang
Anhui
Fujian
Jiangxi
Shandong
Henan
Hubei
Hunan
Guangdong
Hainan
Sichuan
Guizhou
Yunnan
Shaanxi
Gansu
Qinghai
the Inner Mongolia Autonomous…
the Guangxi Zhuang Autonomous…
Tibet Autonomous Region
the Ningxia Hui Autonomous Region
Xinjiang Uygur Autonomous Region
2016 SI
2017 SI
2018 SI
2019 SI

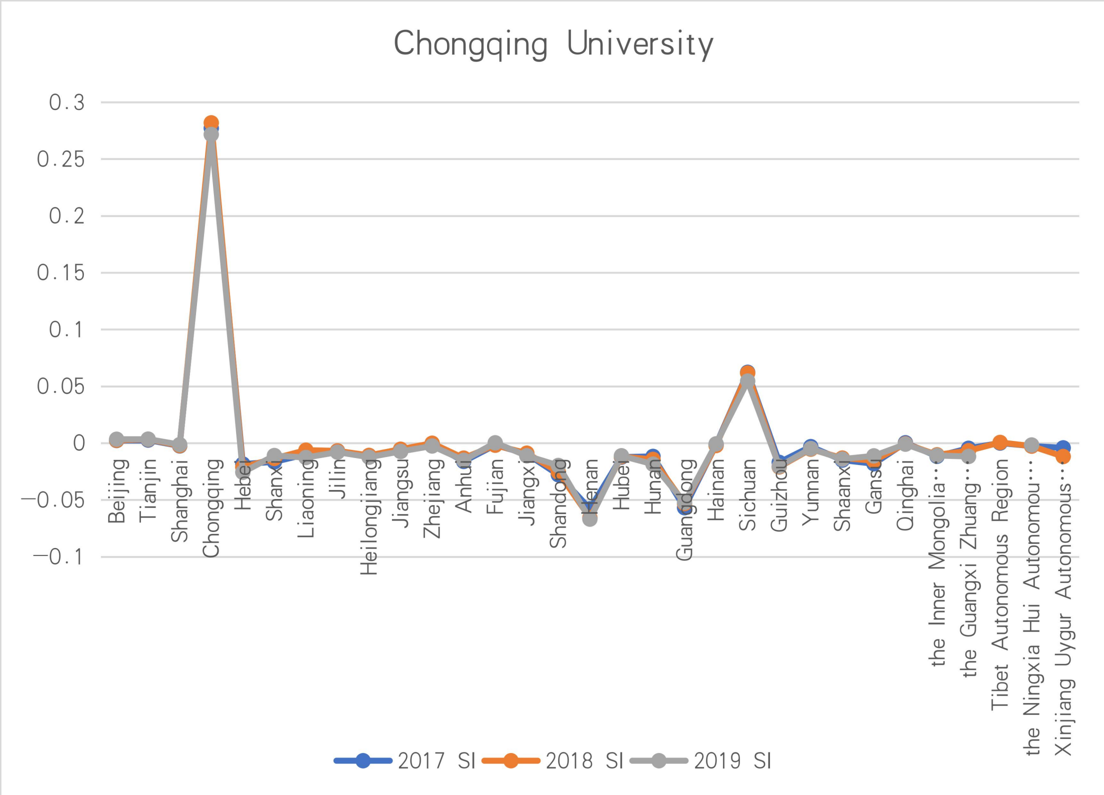

South China University of Technology
0.6
0.5
0.4
0.3
0.2
0.1
0
−0.1
−0.2
Beijing
Tianjin
Shanghai
Chongqing
Hebei
Shanxi
Liaoning
Jilin
Heilongjiang
Jiangsu
Zhejiang
Anhui
Fujian
Jiangxi
Shandong
Henan
Hubei
Hunan
Guangdong
Hainan
Sichuan
Guizhou
Yunnan
Shaanxi
Gansu
Qinghai
the Inner Mongolia…
the Guangxi Zhuang…
Tibet Autonomous Region
the Ningxia Hui Autonomou…
Xinjiang Uygur Autonomous…
2016 SI
2017 SI
2018 SI
2019 SI

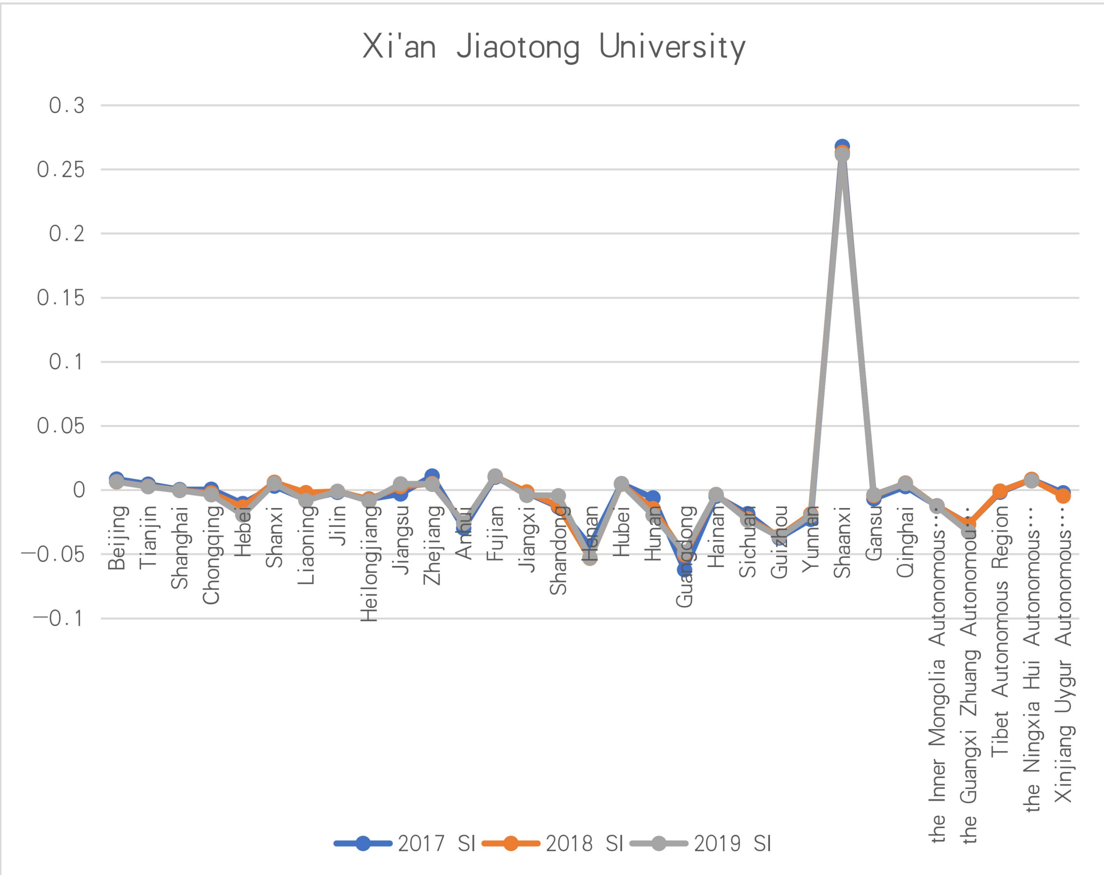

Nankai University
0.25
0.2
0.15
0.1
0.05
0
−0.05
−0.1
Beijing
Tianjin
Shanghai
Chongqing
Hebei
Shanxi
Liaoning
Jilin
Heilongjiang
Jiangsu
Zhejiang
Anhui
Fujian
Jiangxi
Shandong
Henan
Hubei
Hunan
Guangdong
Hainan
Sichuan
Guizhou
Yunnan
Shaanxi
Gansu
Qinghai
the Inner Mongolia…
the Guangxi Zhuang…
Tibet Autonomous Region
the Ningxia Hui Autonomou…
Xinjiang Uygur Autonomous…
2017 SI
2018 SI
2019 SI

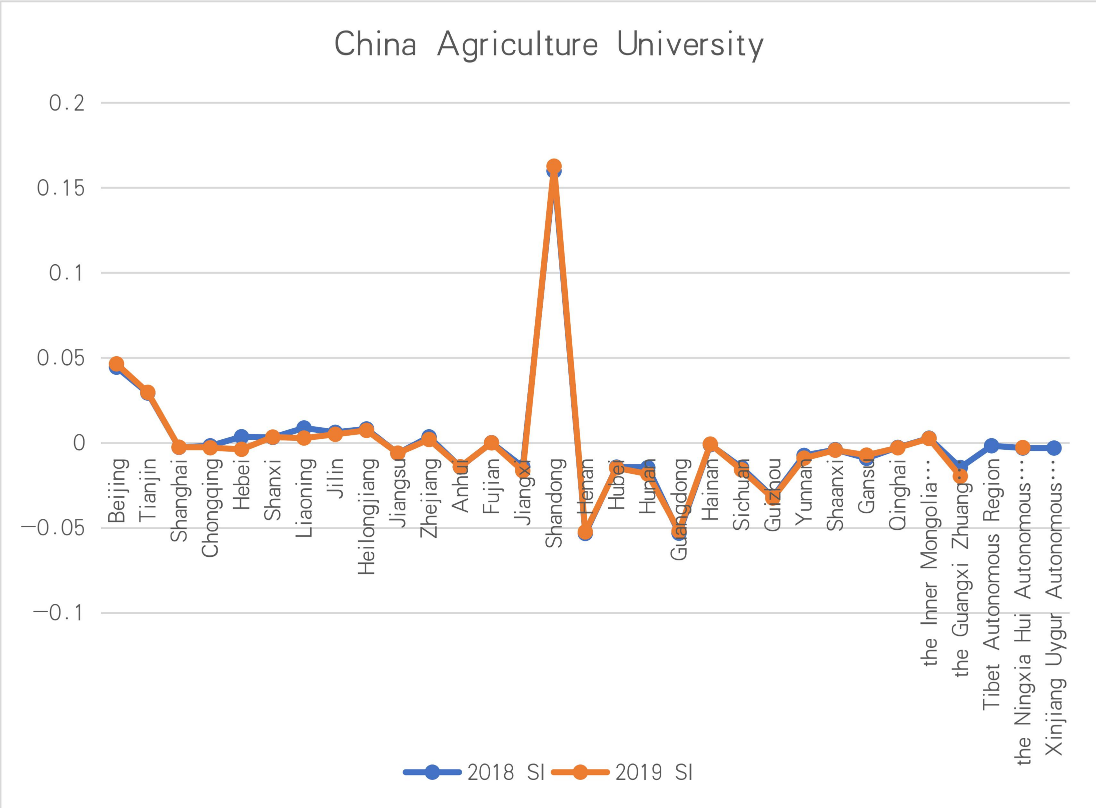

Harbin Institution of Technology
0.3
0.25
0.2
0.15
0.1
0.05
0
−0.05
−0.1
Beijing
Tianjin
Shanghai
Chongqing
Hebei
Shanxi
Liaoning
Jilin
Heilongjiang
Jiangsu
Zhejiang
Anhui
Fujian
Jiangxi
Shandong
Henan
Hubei
Hunan
Guangdong
Hainan
Sichuan
Guizhou
Yunnan
Shaanxi
Gansu
Qinghai
the Inner Mongolia Autonomous…
the Guangxi Zhuang Autonomous…
Tibet Autonomous Region
the Ningxia Hui Autonomous…
Xinjiang Uygur Autonomous Region
2017 SI
2018 SI
2019 SI

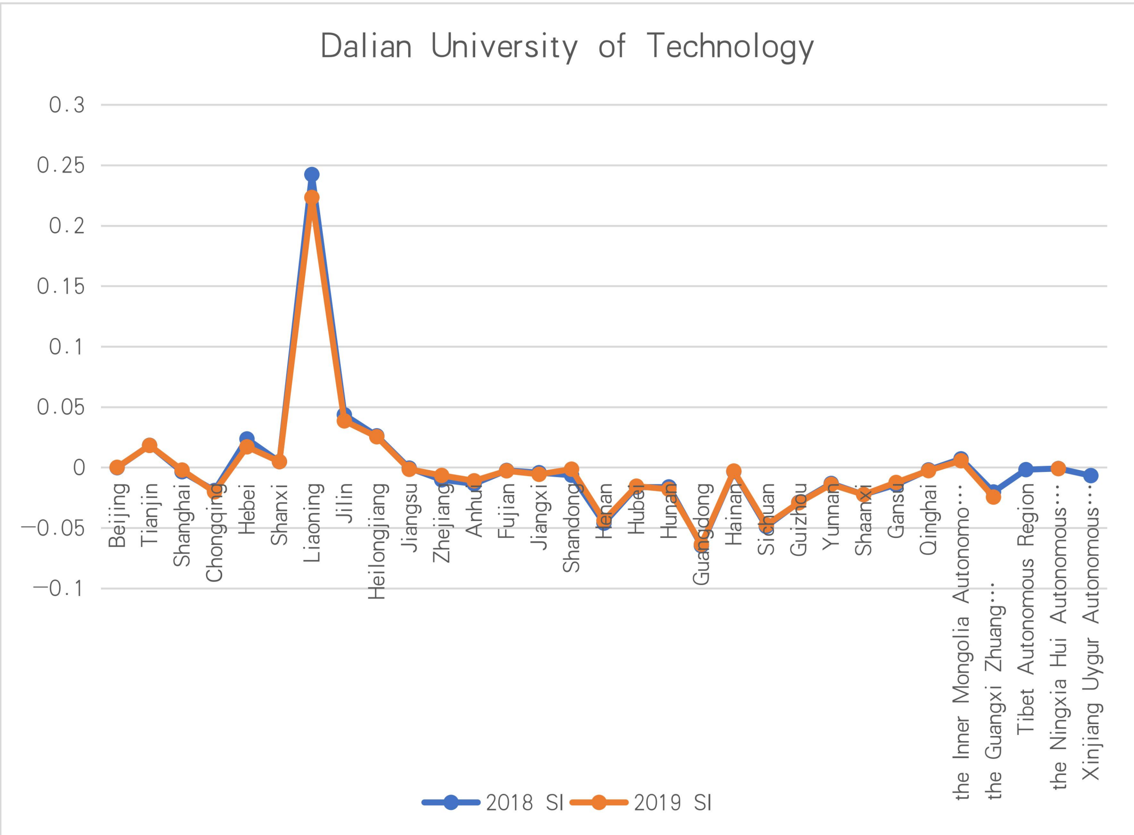

Fudan University
0.08
0.06
0.04
0.02
0
−0.02
−0.04
−0.06
Beijing
Tianjin
Shanghai
Chongqing
Hebei
Shanxi
Liaoning
Jilin
Heilongjiang
Jiangsu
Zhejiang
Anhui
Fujian
Jiangxi
Shandong
Henan
Hubei
Hunan
Guangdong
Hainan
Sichuan
Guizhou
Yunnan
Shaanxi
Gansu
Qinghai
the Inner Mongolia…
the Guangxi Zhuang…
Tibet Autonomous Region
the Ningxia Hui Autonomo…
Xinjiang Uygur Autonomou…
2017 SI
2018 SI
2019 SI

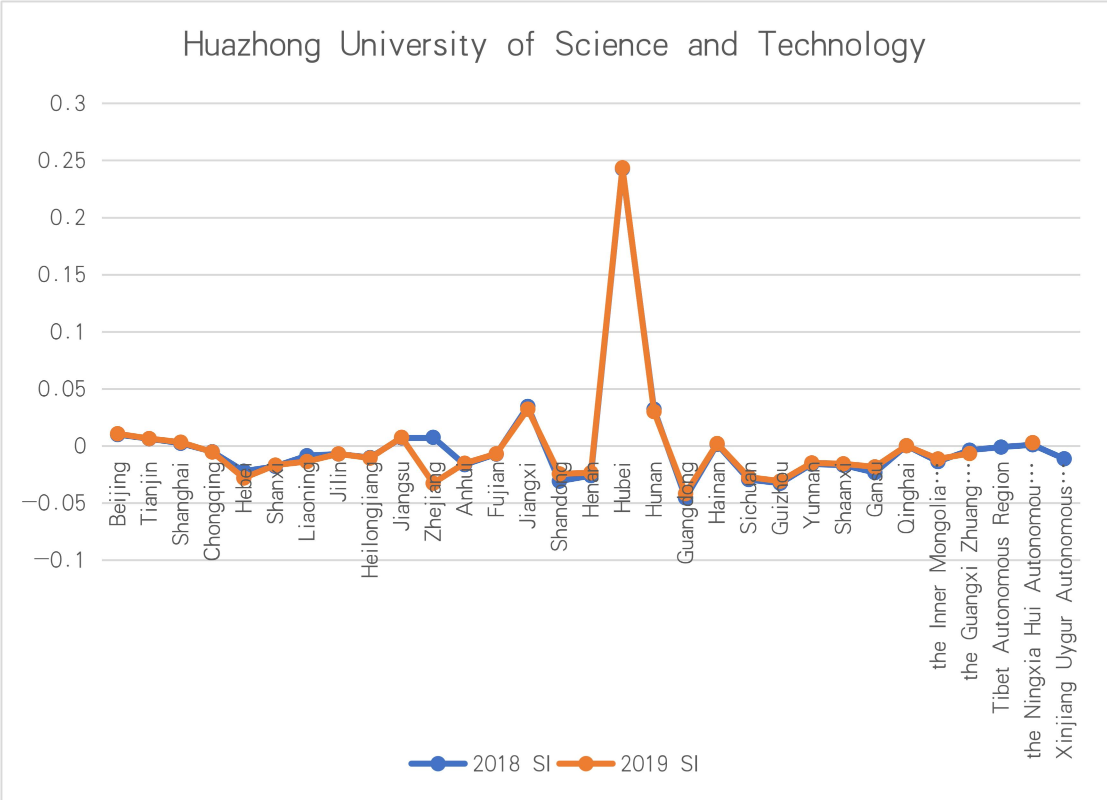

Sichuan University
0.3
0.25
0.2
0.15
0.1
0.05
0
−0.05
−0.1
Beijing
Tianjin
Shanghai
Chongqing
Hebei
Shanxi
Liaoning
Jilin
Heilongjiang
Jiangsu
Zhejiang
Anhui
Fujian
Jiangxi
Shandong
Henan
Hubei
Hunan
Guangdong
Hainan
Sichuan
Guizhou
Yunnan
Shaanxi
Gansu
Qinghai
the Inner Mongolia…
the Guangxi Zhuang…
Tibet Autonomous Region
the Ningxia Hui Autonomou…
Xinjiang Uygur Autonomous…
2017 SI
2018 SI
2019 SI

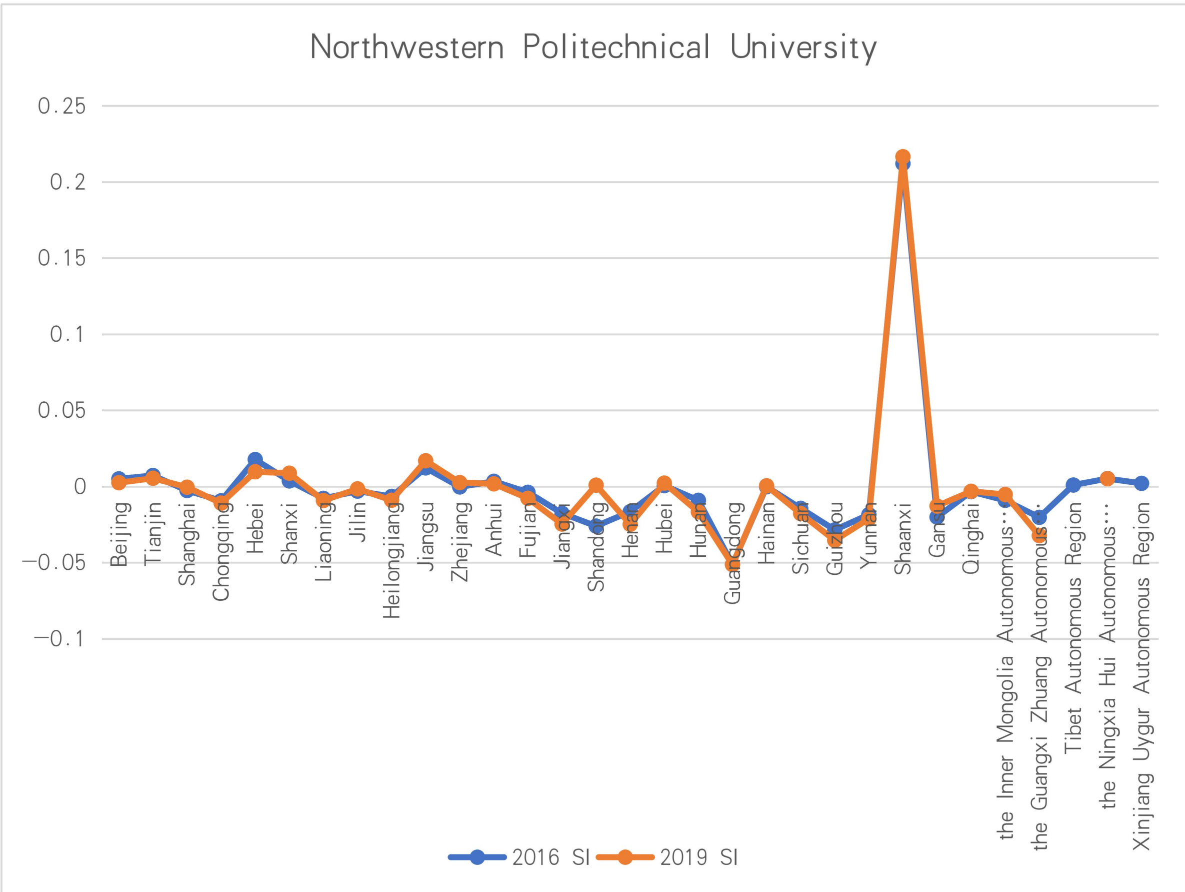

Chongqing University
0.3
0.25
0.2
0.15
0.1
0.05
0
−0.05
−0.1
Beijing
Tianjin
Shanghai
Chongqing
Hebei
Shanxi
Liaoning
Jilin
Heilongjiang
Jiangsu
Zhejiang
Anhui
Fujian
Jiangxi
Shandong
Henan
Hubei
Hunan
Guangdong
Hainan
Sichuan
Guizhou
Yunnan
Shaanxi
Gansu
Qinghai
the Inner Mongolia…
the Guangxi Zhuang…
Tibet Autonomous Region
the Ningxia Hui Autonomou…
Xinjiang Uygur Autonomous…
2017 SI
2018 SI
2019 SI

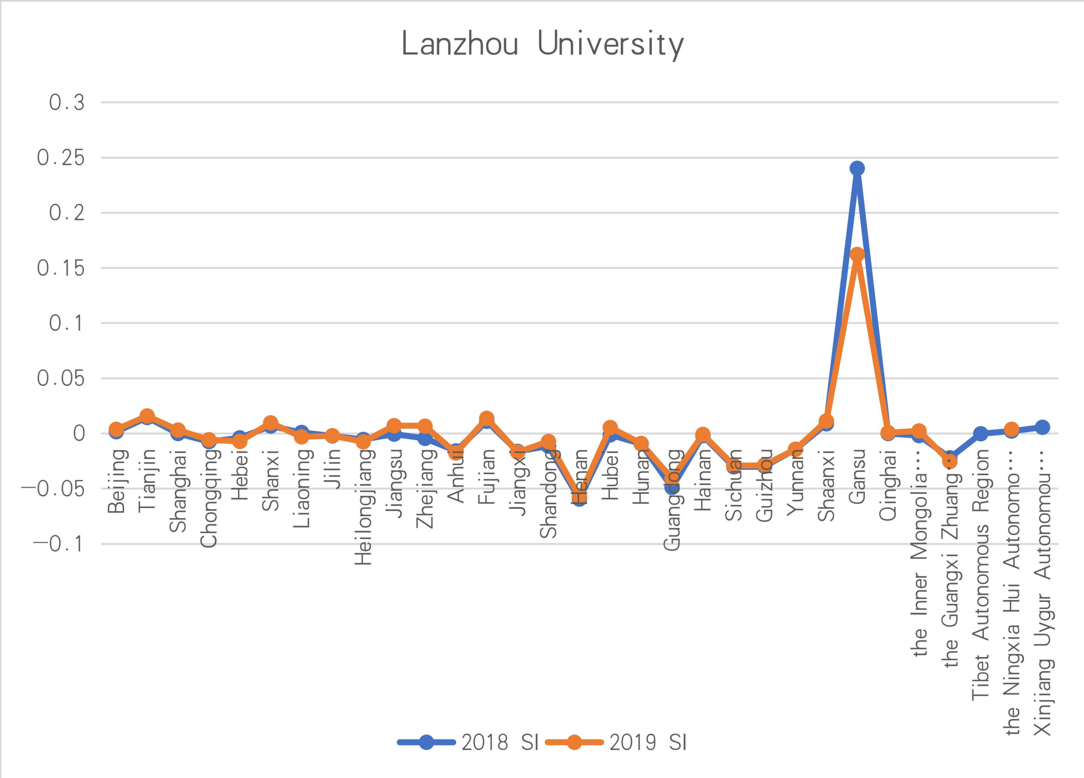

Xi'an Jiaotong University
0.3
0.25
0.2
0.15
0.1
0.05
0
−0.05
−0.1
Beijing
Tianjin
Shanghai
Chongqing
Hebei
Shanxi
Liaoning
Jilin
Heilongjiang
Jiangsu
Zhejiang
Anhui
Fujian
Jiangxi
Shandong
Henan
Hubei
Hunan
Guangdong
Hainan
Sichuan
Guizhou
Yunnan
Shaanxi
Gansu
Qinghai
the Inner Mongolia Autonomous…
the Guangxi Zhuang Autonomous…
Tibet Autonomous Region
the Ningxia Hui Autonomous…
Xinjiang Uygur Autonomous…
2017 SI
2018 SI
2019 SI

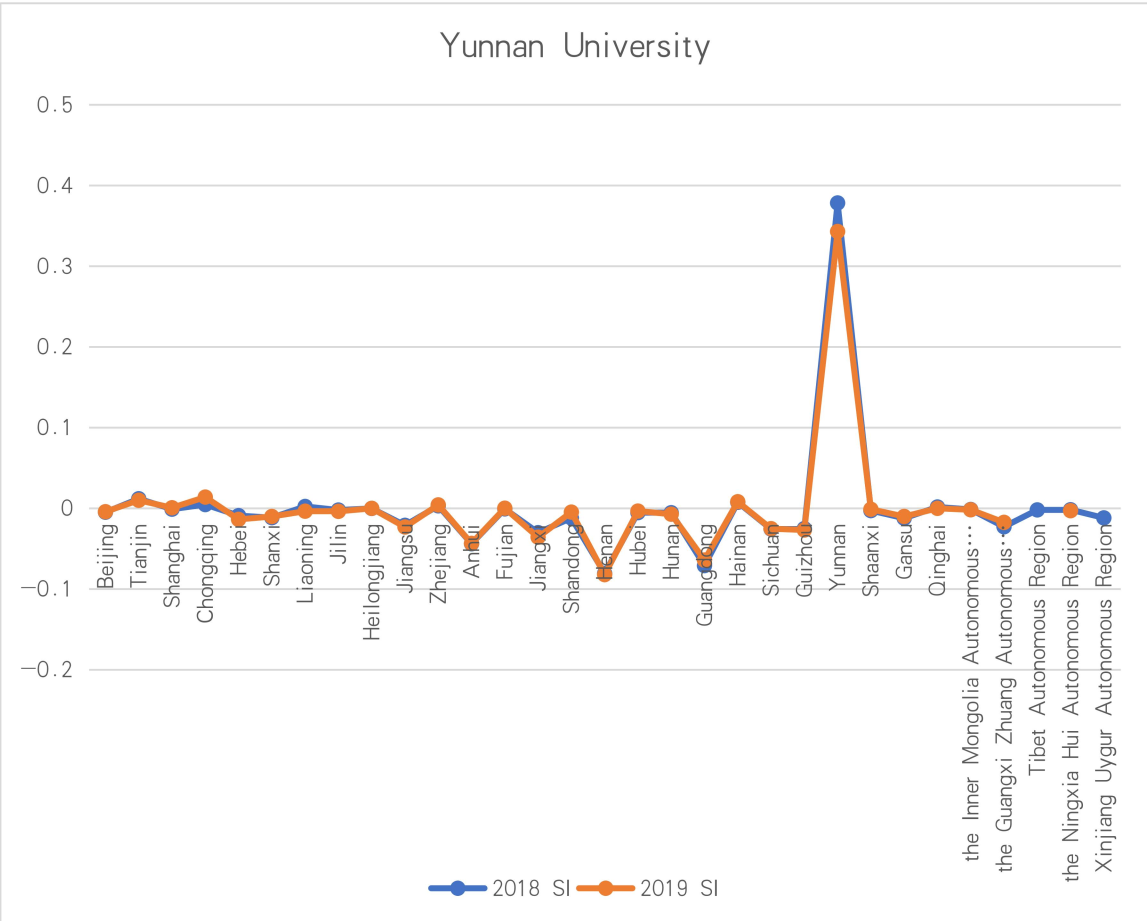

China Agriculture University
0.2
0.15
0.1
0.05
0
−0.05
−0.1
Beijing
Tianjin
Shanghai
Chongqing
Hebei
Shanxi
Liaoning
Jilin
Heilongjiang
Jiangsu
Zhejiang
Anhui
Fujian
Jiangxi
Shandong
Henan
Hubei
Hunan
Guangdong
Hainan
Sichuan
Guizhou
Yunnan
Shaanxi
Gansu
Qinghai
the Inner Mongolia…
the Guangxi Zhuang…
Tibet Autonomous Region
the Ningxia Hui Autonomous…
Xinjiang Uygur Autonomous…
2018 SI
2019 SI

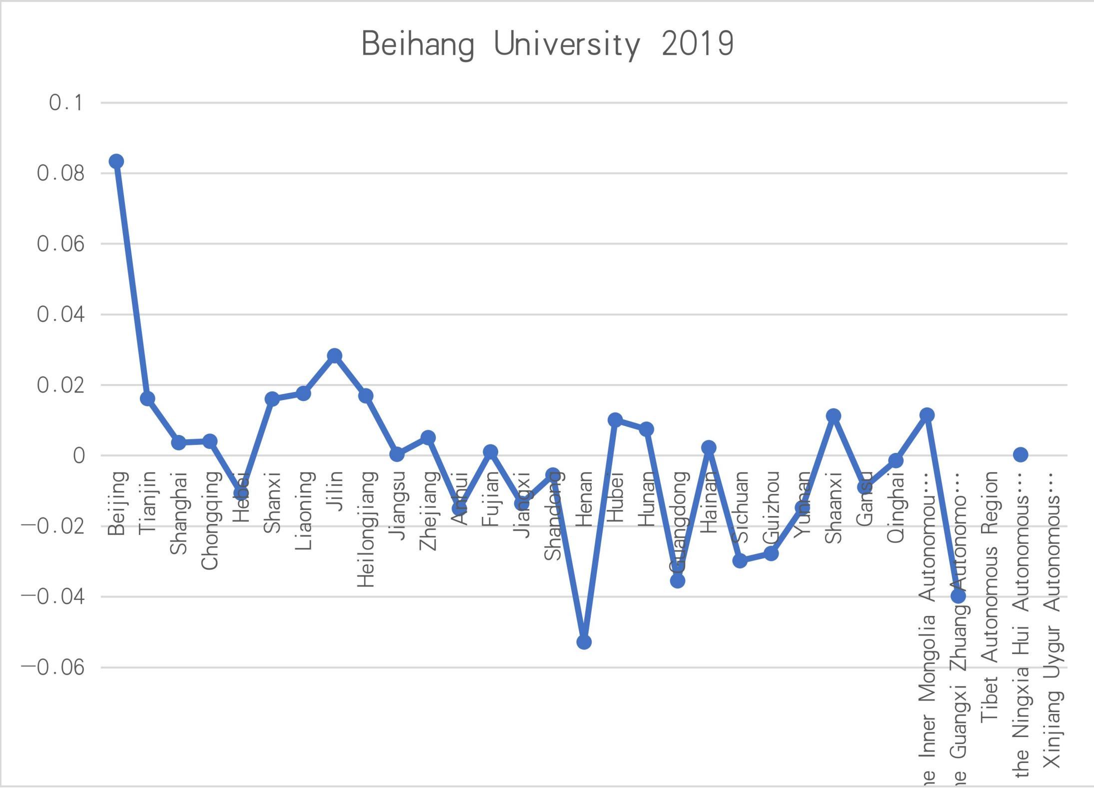

Dalian University of Technology
0.3
0.25
0.2
0.15
0.1
0.05
0
−0.05
−0.1
Beijing
Tianjin
Shanghai
Chongqing
Hebei
Shanxi
Liaoning
Jilin
Heilongjiang
Jiangsu
Zhejiang
Anhui
Fujian
Jiangxi
Shandong
Henan
Hubei
Hunan
Guangdong
Hainan
Sichuan
Guizhou
Yunnan
Shaanxi
Gansu
Qinghai
the Inner Mongolia Autonomo…
the Guangxi Zhuang…
Tibet Autonomous Region
the Ningxia Hui Autonomous…
Xinjiang Uygur Autonomous…
2018 SI
2019 SI


Huazhong University of Science and Technology
0.3
0.25
0.2
0.15
0.1
0.05
0
−0.05
−0.1
Beijing
Tianjin
Shanghai
Chongqing
Hebei
Shanxi
Liaoning
Jilin
Heilongjiang
Jiangsu
Zhejiang
Anhui
Fujian
Jiangxi
Shandong
Henan
Hubei
Hunan
Guangdong
Hainan
Sichuan
Guizhou
Yunnan
Shaanxi
Gansu
Qinghai
the Inner Mongolia…
the Guangxi Zhuang…
Tibet Autonomous Region
the Ningxia Hui Autonomou…
Xinjiang Uygur Autonomous…
2018 SI
2019 SI

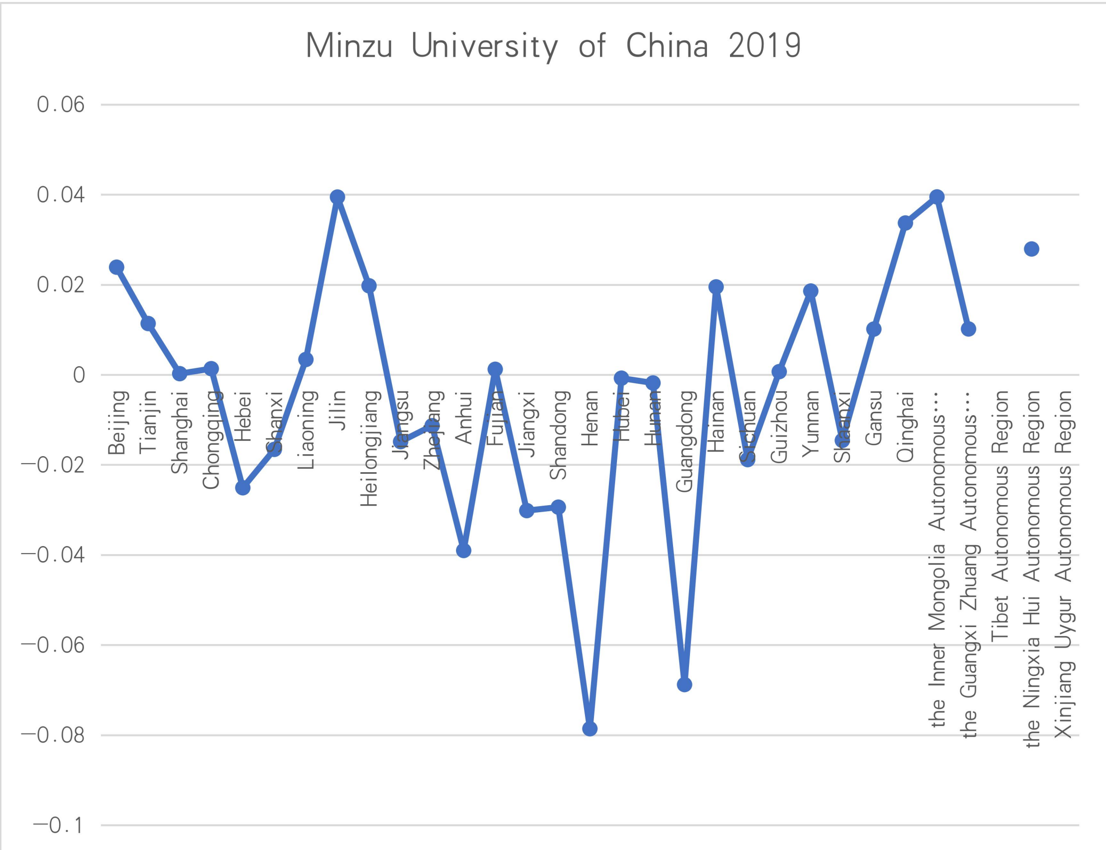

Northwestern Politechnical University
0.25
0.2
0.15
0.1
0.05
0
−0.05
−0.1
Beijing
Tianjin
Shanghai
Chongqing
Hebei
Shanxi
Liaoning
Jilin
Heilongjiang
Jiangsu
Zhejiang
Anhui
Fujian
Jiangxi
Shandong
Henan
Hubei
Hunan
Guangdong
Hainan
Sichuan
Guizhou
Yunnan
Shaanxi
Gansu
Qinghai
the Inner Mongolia Autonomous…
the Guangxi Zhuang Autonomous…
Tibet Autonomous Region
the Ningxia Hui Autonomous…
Xinjiang Uygur Autonomous Region
2016 SI
2019 SI


Lanzhou University
0.3
0.25
0.2
0.15
0.1
0.05
0
−0.05
−0.1
Beijing
Tianjin
Shanghai
Chongqing
Hebei
Shanxi
Liaoning
Jilin
Heilongjiang
Jiangsu
Zhejiang
Anhui
Fujian
Jiangxi
Shandong
Henan
Hubei
Hunan
Guangdong
Hainan
Sichuan
Guizhou
Yunnan
Shaanxi
Gansu
Qinghai
the Inner Mongolia…
the Guangxi Zhuang…
Tibet Autonomous Region
the Ningxia Hui Autonomo…
Xinjiang Uygur Autonomou…
2018 SI
2019 SI

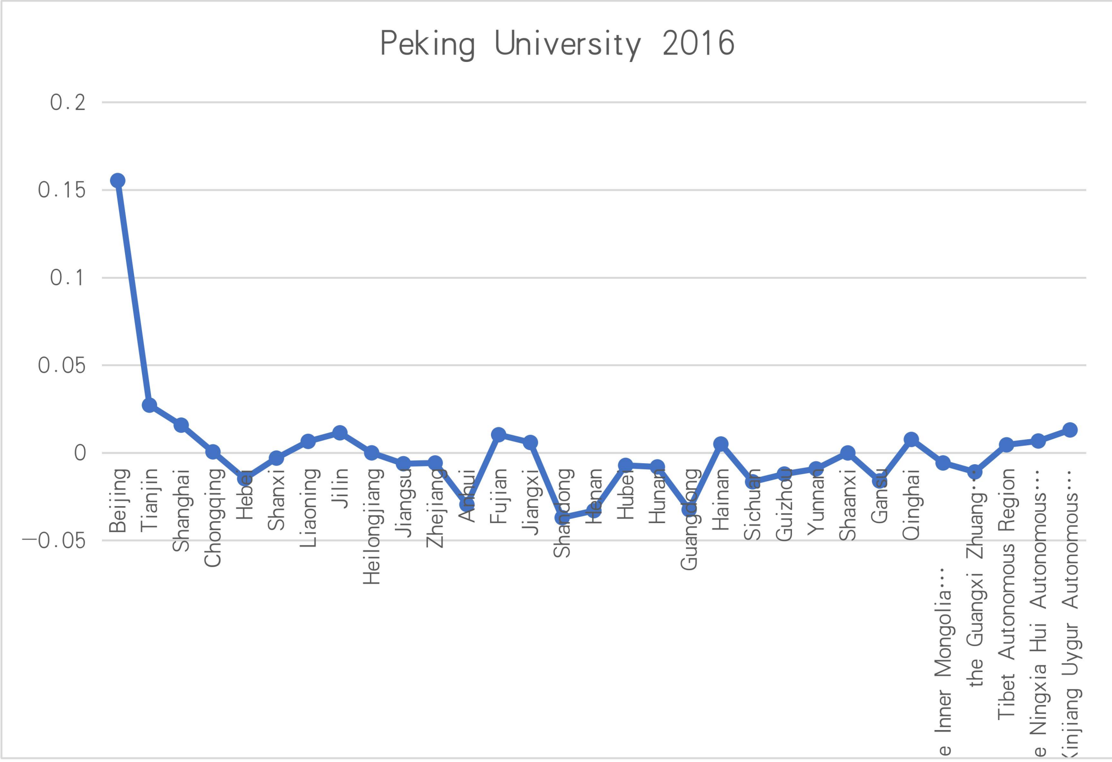

Yunnan University
0.5
0.4
0.3
0.2
0.1
0
−0.1
−0.2
Beijing
Tianjin
Shanghai
Chongqing
Hebei
Shanxi
Liaoning
Jilin
Heilongjiang
Jiangsu
Zhejiang
Anhui
Fujian
Jiangxi
Shandong
Henan
Hubei
Hunan
Guangdong
Hainan
Sichuan
Guizhou
Yunnan
Shaanxi
Gansu
Qinghai
the Inner Mongolia Autonomous…
the Guangxi Zhuang Autonomous…
Tibet Autonomous Region
the Ningxia Hui Autonomous Region
Xinjiang Uygur Autonomous Region
2018 SI
2019 SI

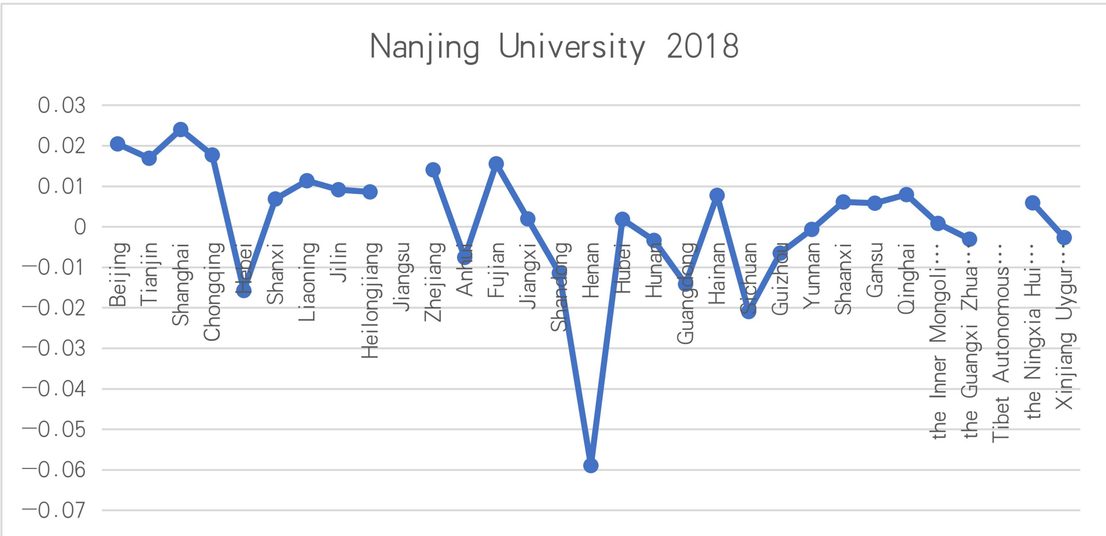


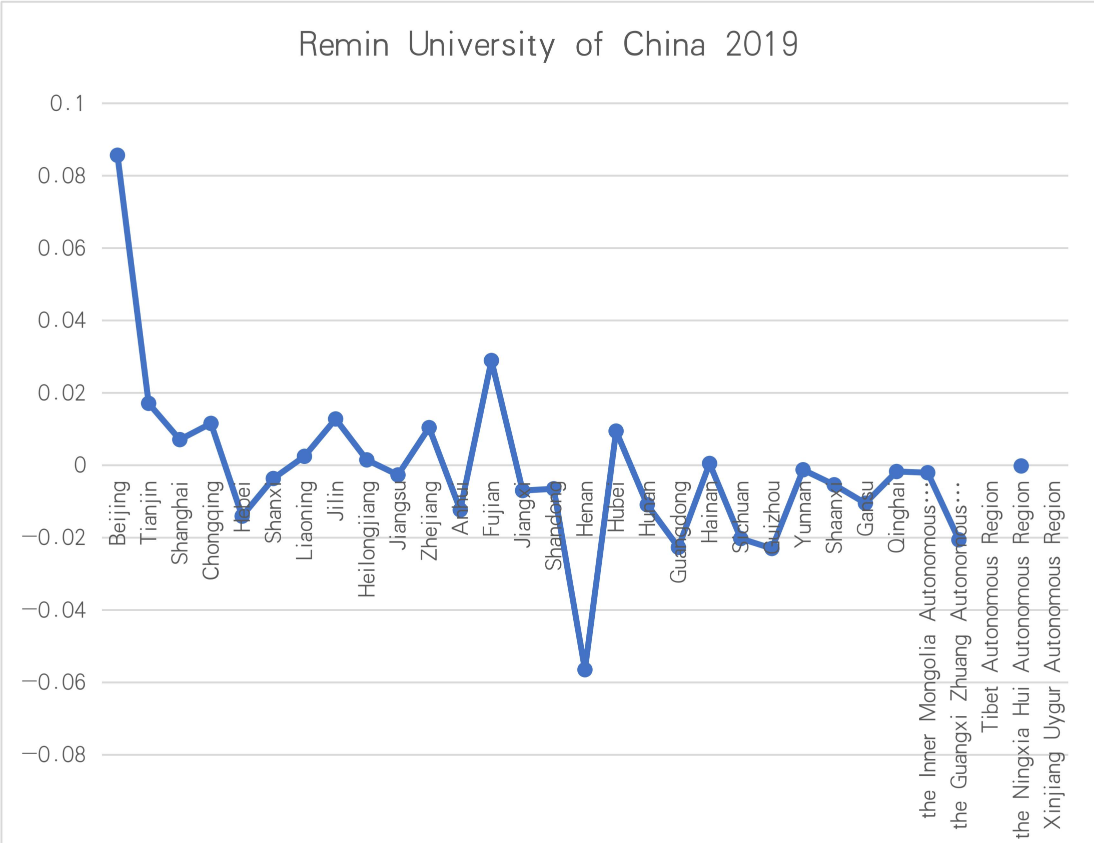
Beihang University 2019
0.1
0.08
0.06
0.04
0.02
0
−0.02
−0.04
−0.06
Beijing
Tianjin
Shanghai
Chongqing
Hebei
Shanxi
Liaoning
Jilin
Heilongjiang
Jiangsu
Zhejiang
Anhui
Fujian
Jiangxi
Shandong
Henan
Hubei
Hunan
Guangdong
Hainan
Sichuan
Guizhou
Yunnan
Shaanxi
Gansu
Qinghai
Inner Mongolia Autonomou…
Guangxi Zhuang Autonomo…
Tibet Autonomous Region
the Ningxia Hui Autonomous…
Xinjiang Uygur Autonomous…



Minzu University of China 2019
0.06
0.04
0.02
0
−0.02
−0.04
−0.06
−0.08
−0.1
Beijing
Tianjin
Shanghai
Chongqing
Hebei
Shanxi
Liaoning
Jilin
Heilongjiang
Jiangsu
Zhejiang
Anhui
Fujian
Jiangxi
Shandong
Henan
Hubei
Hunan
Guangdong
Hainan
Sichuan
Guizhou
Yunnan
Shaanxi
Gansu
Qinghai
the Inner Mongolia Autonomous…
the Guangxi Zhuang Autonomous…
Tibet Autonomous Region
the Ningxia Hui Autonomous Region
Xinjiang Uygur Autonomous Region


Peking University 2016
0.2
0.15
0.1
0.05
0
−0.05
Beijing
Tianjin
Shanghai
Chongqing
Shanxi
Liaoning
Jilin
Heilongjiang
Jiangsu
Fujian
Jiangxi
Hainan
Yunnan
Shaanxi
Qinghai
the Guangxi Zhuang…
Tibet Autonomous Region


Nanjing University 2018
0.03
0.02
0.01
0
−0.01
−0.02
−0.03
−0.04
−0.05
−0.06
−0.07
Beijing
Tianjin
Shanghai
Chongqing
Shanxi
Liaoning
Jilin
Heilongjiang
Jiangsu
Zhejiang
Fujian
Jiangxi
Henan
Hainan
Yunnan
Shaanxi
Gansu
Qinghai
the Inner Mongoli…
the Guangxi Zhua…
Tibet Autonomous…
the Ningxia Hui…
Xinjiang Uygur…


Remin University of China 2019
0.1
0.08
0.06
0.04
0.02
0
−0.02
−0.04
−0.06
−0.08
Beijing
Tianjin
Shanghai
Chongqing
Shanxi
Liaoning
Jilin
Heilongjiang
Jiangsu
Zhejiang
Fujian
Henan
Hubei
Sichuan
Yunnan
Shaanxi
Qinghai
the Inner Mongolia Autonomous…
the Guangxi Zhuang Autonomous…
Tibet Autonomous Region
the Ningxia Hui Autonomous Region
Xinjiang Uygur Autonomous Region

Appendix 9 Percentages of individuals potentially disadvantaged by hukou status

		Rural hukou	Urban hukou	Missing	Total (N)
sex	Male	60.1	39.7	0.3	4814
	Female	65.2	34.6	0.2	5275
Ethnicity	Han ethnicity	61.9	37.9	0.2	9182
	Minority ethnicity	71.6	28	0.3	902
	Missing	60	40	0	5
Social class	Working class	72.6	27.2	0.2	4798
	Middle class	53.5	46.3	0.2	4934
	Upper class	57	42.2	0.7	277
	Missing	65	33.8	1.3	80
Language ability (listening to Mandarin)	Cannot understand at all	83.7	16.3	0	49
	Bad	83.9	15.5	0.5	193
	Not bad	75.7	24	0.3	1884
	Good	65.7	34.2	0.1	3594
	Very good	53.6	46.1	0.3	4365
	Missing	25	75	0	4
Language ability (speaking Mandarin)	Cannot speak at all	91.8	8.2	0	122
	Bad	82.9	16.6	0.5	613
	Not bad	74.5	25.3	0.2	3184
	Good	60.7	39	0.2	3283
	Very good	46.7	53.1	0.3	2883
	Missing	25	75	0	4
Father's workplace	Party/government agencies	18.1	81.5	0.4	276
	Companies	31.1	68.7	0.3	1929
	Institutions	29.1	70.4	0.6	719
	Social community/committee	85.6	14.4	0	97
	Self-employed	80.1	19.7	0.2	5930
	Army	34.8	65.2	0	23
	Other	80.5	19.5	0	261
	Missing	49.3	50.4	0.4	854
Mother's workplace	Party/government agencies	13.5	85.6	1	104
	Companies	21.6	77.9	0.5	1346
	Institutions	13.4	86.4	0.2	531
	Social community/committee	72.7	27.3	0	66
	Self-employed	81.7	18.2	0.1	5759
	Army	0	100	0	1
	Other	84.9	14.3	0.8	238
	Missing	49	50.6	0.3	2044
Parental education	Both parents basic education or below	84.2	15.6	0.2	3343
	At least one parent received intermediate education	57.5	42.3	0.2	5917
	At least one parent received higher education	17	82.2	0.8	829

Appendix 10 Complete Results of Ten Regressions on CGSS data

Table 1 Summary of average correctly predicted percentages of HE entries in ten regressions (Model 1; enter method)

	Average correctly predicted percentage
Base	50.0
Block 1 (sex, ethnicity, hukou status)	69.1
Block 2 (parents' education, parents' workplaces, social class)	70.8
Block 3 (ability to understand and speak Mandarin)	72.7

N=4,194

Table 2 Percentages predicted correctly for HE entry in ten regressions (Model 1; enter)

Estimate	1	2	3	4	5	6	7	8	9	10	Average
Base	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0
Block 1 (sex, ethnicity, hukou status)	69.3	69.3	69.5	68.3	68.6	69.3	69.5	69.3	68.3	69.1	69.1
Block 2 (parents' education qualifications, mother's workplace, father's workplace, social class)	71.2	71.1	70.8	70.2	70.3	71.1	70.8	71.2	70.2	70.9	70.8
Block 3 (language ability in understanding Mandarin, language ability in speaking Mandarin)	73.0	72.7	73.0	72.2	72.2	72.7	73.0	73.0	72.2	72.9	72.7

N=4194

Table 3 Percentages predicted correctly for HE entry in ten regressions (Model 2; enter)

Estimate	1	2	3	4	5	6	7	8	9	10	Average
Base	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0
Block 1 (parents' education qualifications, mother's workplace, father's workplace, social class)	68.8	68.2	69.1	68.3	68.3	68.2	69.1	68.3	68.3	69.2	68.6
Block 2 (language ability in understanding Mandarin, language ability in speaking Mandarin)	71.2	70.4	71.3	70.5	70.4	70.4	71.3	70.4	70.5	71.8	70.8
Block 3 (sex, ethnicity, hukou status)	73.0	72.7	73.0	72.2	72.2	72.7	73.0	72.2	72.2	72.9	72.6

N=4194

Table 4 Percentages predicted correctly for HE entry in ten regressions (Model 1; forward stepwise (conditional))

Estimate		1	2	3	4	5	6	7	8	9	10	Average
Block 1	Step 1 (hukou)	69.3	69.3	69.5	68.3	68.6	69.3	69.5	68.6	68.3	69.1	69
	Step 2 (ethnicity)	-	69.3	-	68.3	68.6	69.3	-	68.6	68.3	69.1	68.8
Block 2	Step 1 (parents' education qualifications)	69.6	69.6	69.7	68.2	69.1	69.6	69.7	69.1	68.2	70.1	69.3
	Step 2 (mother's workplace)	71.0	70.7	70.7	69.7	70.1	70.7	70.7	70.1	69.7	71.1	70.5
	Step 3 (social class)	70.9	70.7	70.7	70.1	70.4	70.7	70.7	70.4	70.1	71.0	70.6
Block 3	Step 1 (language ability in speaking Mandarin)	72.6	72.2	72.2	71.7	71.9	72.2	72.2	71.9	71.7	72.0	72.1
	Step 2 (language ability in understanding Mandarin)	73.0	72.2	72.8	72.2	72.3	72.2	72.8	72.3	72.2	72.3	72.4

N=4194

Note: an empty cell means that there is no Step 2 in Block 1 in the forward entry model

Table 5 Percentages predicted correctly for HE entry in ten regressions (Model 2; forward stepwise (conditional))

Estimate		1	2	3	4	5	6	7	8	9	10	Average
Block 1	Step 1 (parents' education qualifications)	67.7	67.4	67.6	67.1	66.8	67.4	67.6	66.8	67.1	68.1	67.4
	Step 2 (mother's workplace)	68.3	67.5	67.8	67.6	67.5	67.5	67.8	67.5	67.6	68.7	67.8
	Step 3 (social class)	68.0	67.4	67.8	67.4	67.4	67.4	67.8	67.4	67.4	68.4	67.6
	Step 4 (father's working places)	68.8	68.2	69.1	68.3	68.3	68.2	69.1	68.3	68.3	69.2	68.6
Block 2	Step 1 (language ability in speaking Mandarin)	71.0	71.0	71.3	70.2	70.6	71.0	71.3	70.6	70.2	71.4	70.9
	Step 2 (language ability in understanding Mandarin)	71.2	70.4	71.3	70.5	70.4	70.4	71.3	70.4	70.5	71.8	70.8
Block 3	Step 1 (hukou)	73.1	72.7	72.9	72.2	72.2	72.7	72.9	72.2	72.2	72.7	72.6

N=4194

Table 6 Summary of the coefficients on the variables in Model 1

	Enter method	Forward stepwise method
female		
male	1.10	-
minority ethnicity		
Han ethnicity	1.12	1.25
rural hukou		
urban hukou	2.45	2.51
both parents had compulsory education or below		
at least one parent had post-compulsory education	2.07	2.11
at least one parent had HE	4.52	4.67
parental education missing	0.94	0.96
self-employed father		
company (father)	1.28	-
institution (father)	1.21	-
social community/committee (father)	1.01	-
party/government agency (father)	1.19	-
army (father)	1.25	-
other workplace (father)	1.11	-
workplace missing (father)	1.04	-
self-employed mother		
company (mother)	1.25	1.41
institution (mother)	1.66	1.84
social community/committee (mother)	2.27	2.16
party/government agency/army (mother)	1.53	1.73
other workplace (mother)	1.17	1.29
workplace missing (mother)	1.33	1.38
working class		
middle class	1.19	1.20
upper class	0.86	0.85
class missing	1.06	1.05
cannot understand at all		
bad at understanding	4.53	4.39
not bad at understanding	2.04	2.02
good at understanding	3.84	3.74
very good at understanding	5.07	4.95
cannot speak at all		
bad at speaking	1.85	1.73
not bad at speaking	4.88	4.49
good at speaking	7.26	6.79

very good at speaking	9.87	9.17

Note: an empty cell means that the variable has been excluded using the Forward stepwise method

Table 7 Summary of the coefficients on the variables in Model 2

	Enter method	Forward stepwise method
Both parents with compulsory education or below		
at least one parent had post-compulsory education	2.06	2.06
at least one parent had HE	4.53	4.52
parental education missing	0.97	0.98
self-employed (father)		
company (father)	1.28	1.28
institution (father)	1.21	1.20
social community/committee (father)	0.95	0.95
party/government agency (father)	1.19	1.18
army (father)	1.20	1.21
other workplace (father)	1.07	1.07
workplace missing (father)	1.02	1.01
self-employed (mother)		
company (mother)	1.22	1.22
institution (mother)	1.66	1.66
social community/committee (mother)	2.16	2.16
party/government agency/army (mother)	1.58	1.59
other workplace (mother)	1.25	1.26
workplace missing (mother)	1.32	1.32
working class		
middle class	1.20	1.20
upper class	0.86	0.85
class missing	1.03	1.03
cannot understand at all		
bad at understanding	4.38	4.27
not bad at understanding	1.96	1.92
good at understanding	3.64	3.58
very good at understanding	4.78	4.72
cannot speak at all		
bad at speaking	1.76	1.81
not bad at speaking	4.61	4.74
good at speaking	7.00	7.15
very good at speaking	9.50	9.70
female		
male	1.10	-

minority ethnicity		
Han ethnicity	1.15	-
rural hukou		
urban hukou	2.43	2.45

Note: an empty cell means that the variable has been excluded using the Forward stepwise method

Appendix 11 The Characteristics of Missing Cases in the CFPS data

Figure 1 The pattern of missing cases in CFPS by hukou status

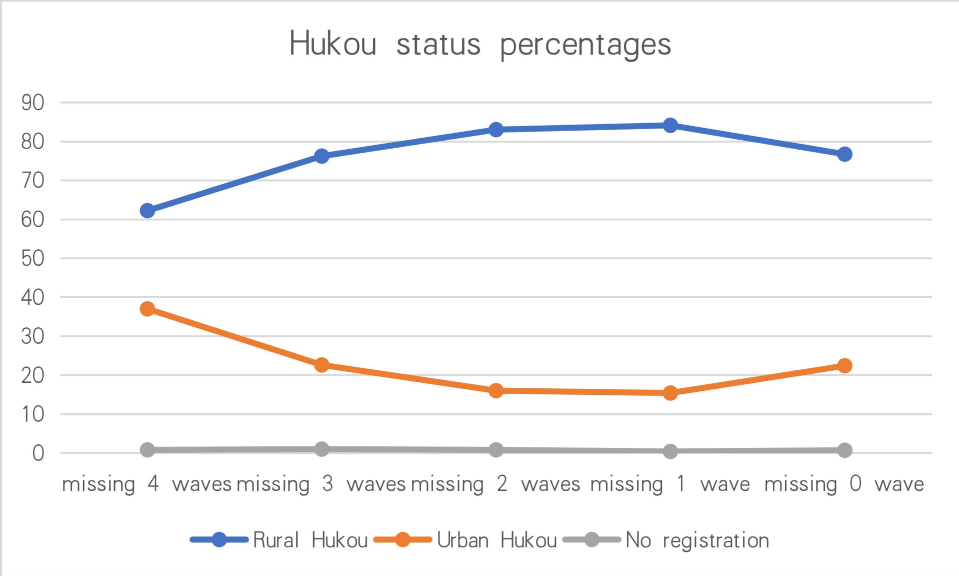


Figure 2 The pattern of missing cases in CFPS by living places

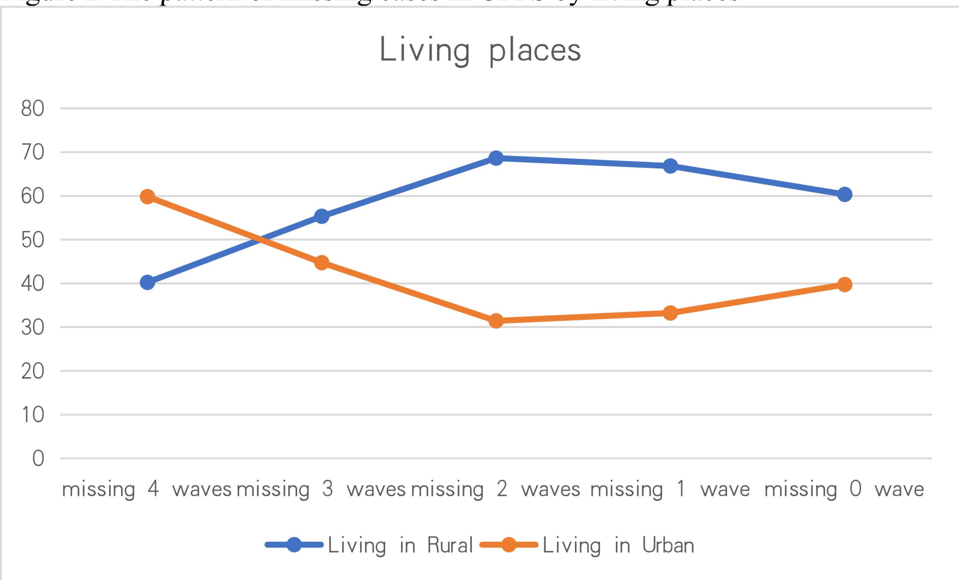


Figure 3 The pattern of missing cases in CFPS by sex

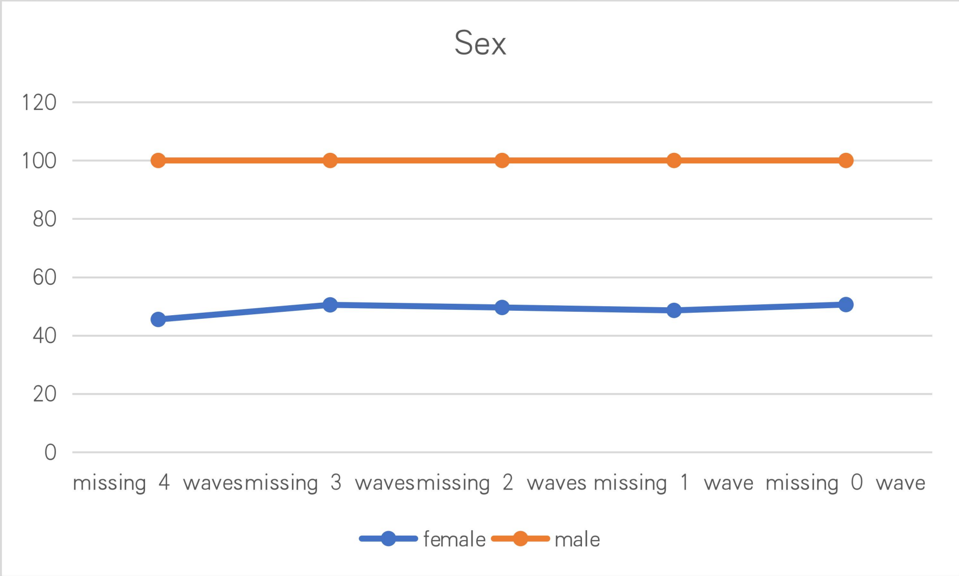


Figure 4 The pattern of missing cases in CFPS by birth month

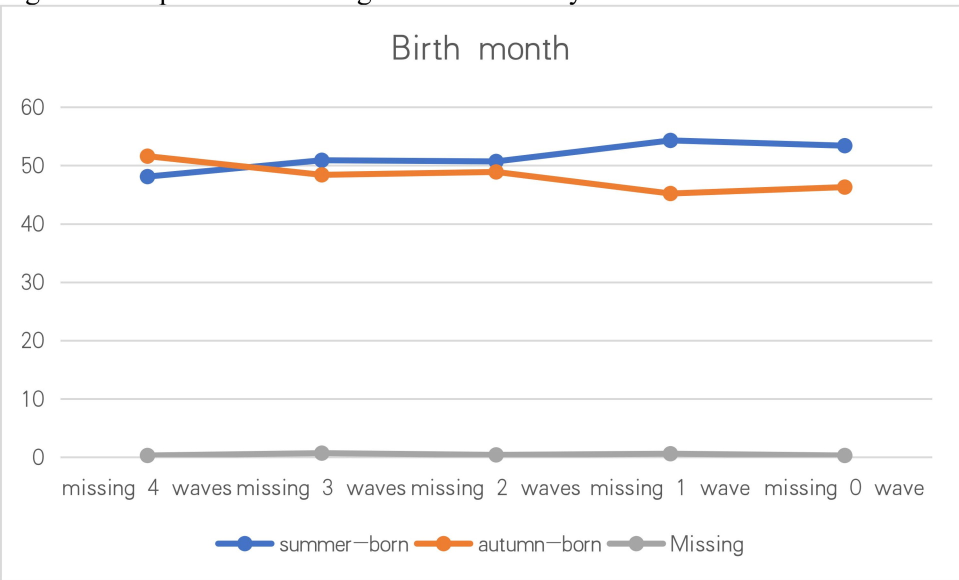


Figure 5 The pattern of missing cases in CFPS by ethnicity

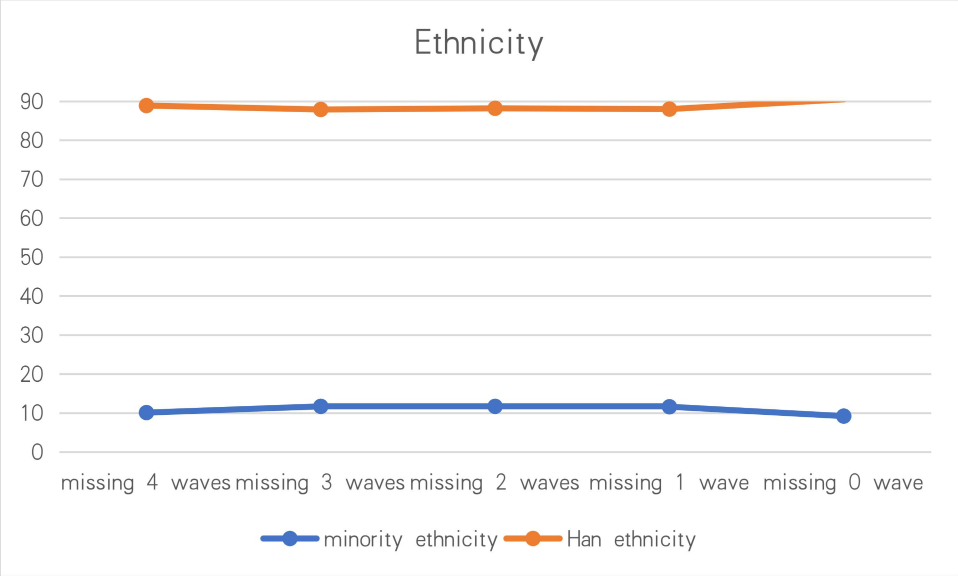


Figure 6 The pattern of missing cases in CFPS by parents' education qualifications

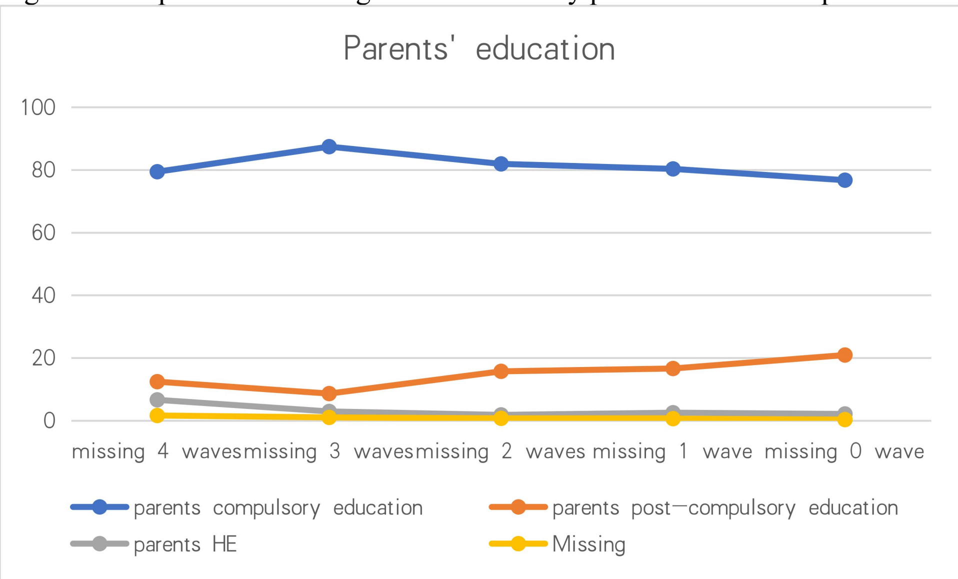


Figure 7 The pattern of missing cases in CFPS by disadvantaged parents occupations

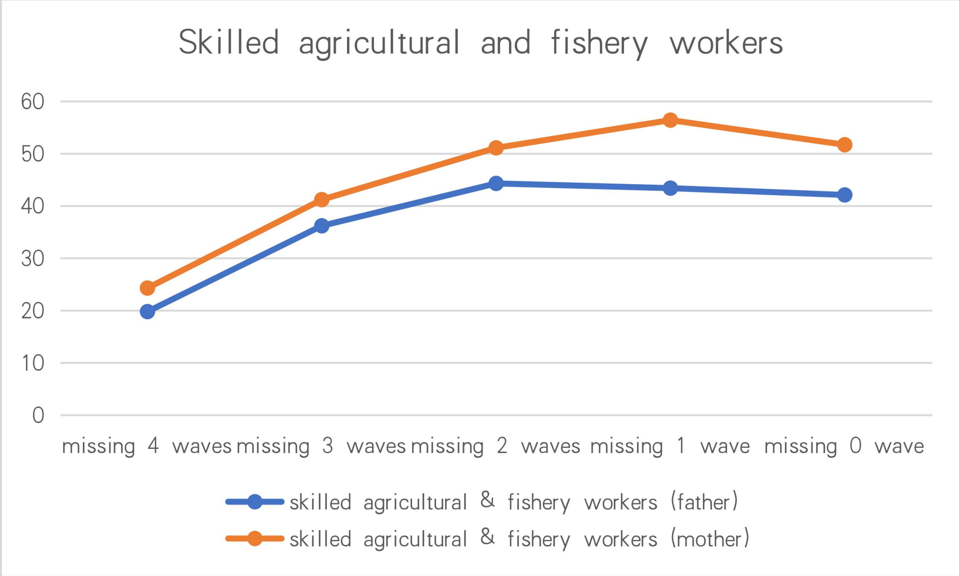


Figure 8 The pattern of missing cases in CFPS by parents' political status

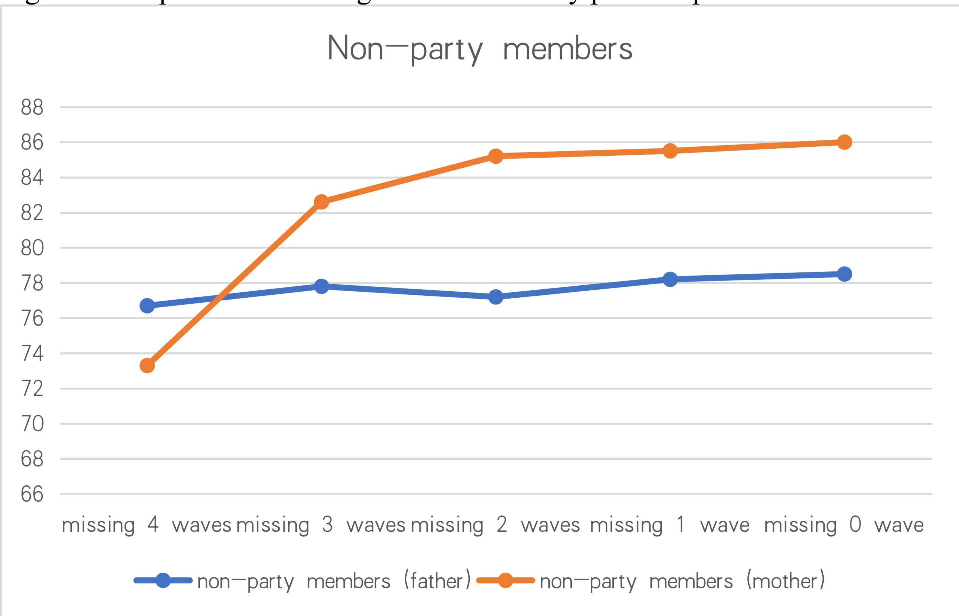


Figure 9 The pattern of missing cases in CFPS by annual family income

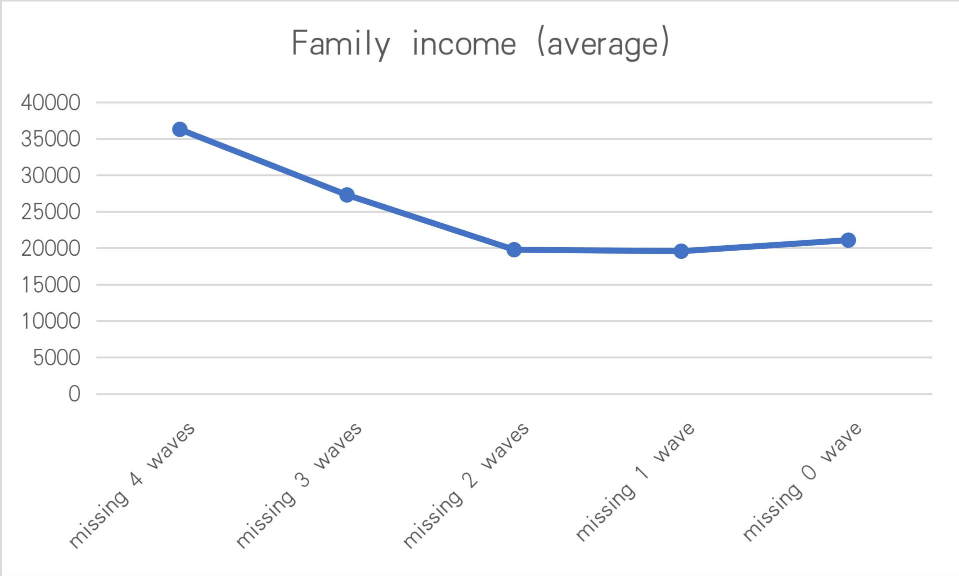


Figure 10 The pattern of missing cases in CFPS by annual expenditure on education

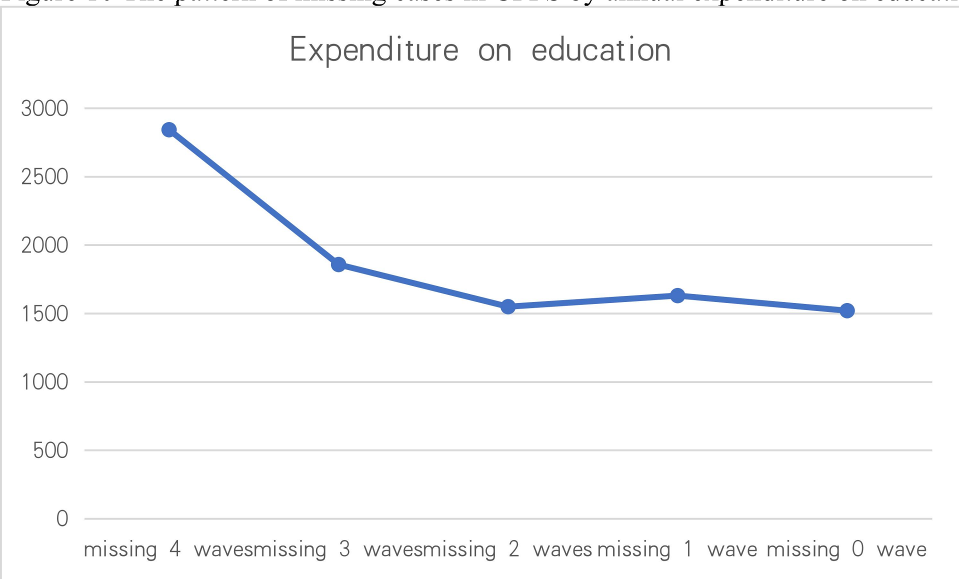


Figure 11 The pattern of missing cases in CFPS by educational aspirations

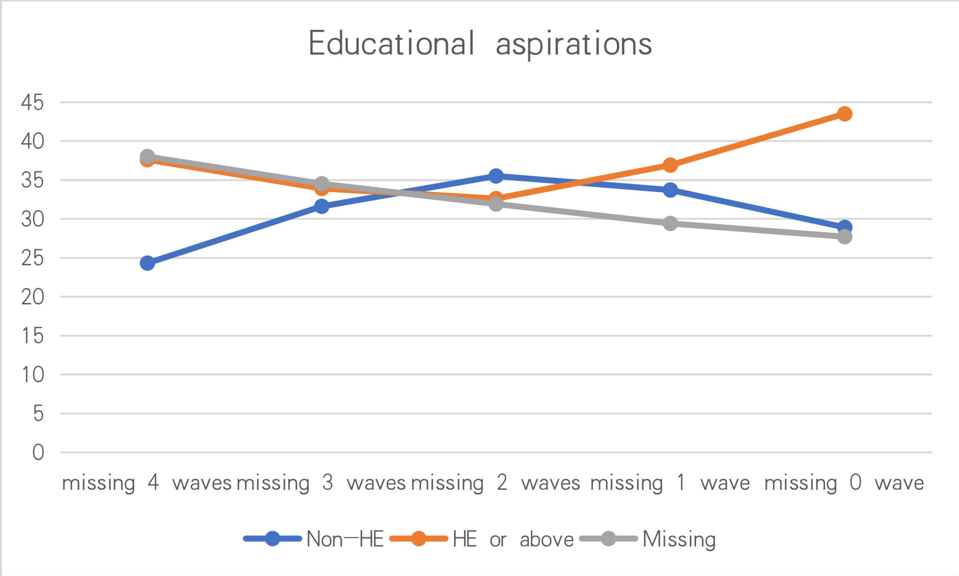


Figure 12 The pattern of missing cases in CFPS by Chinese scores

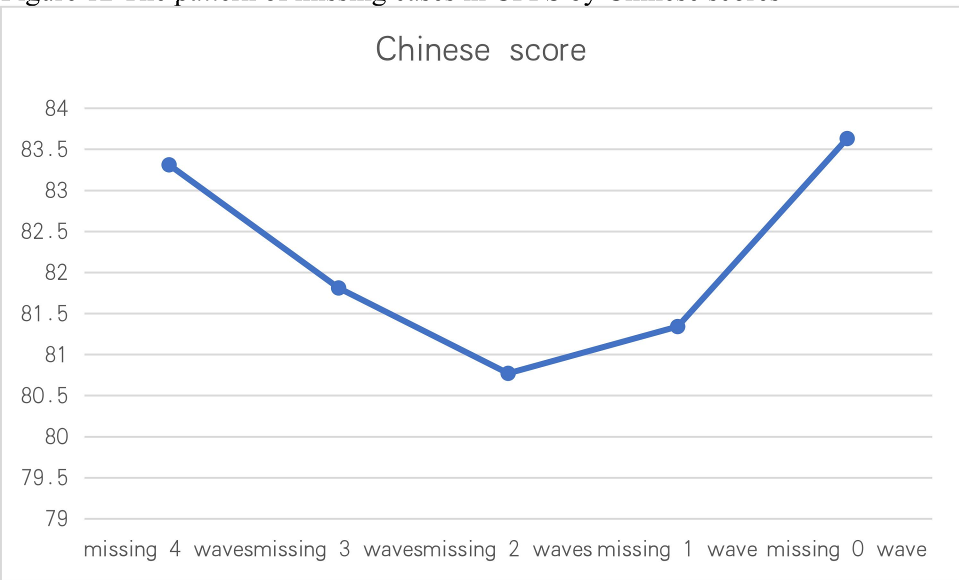


Figure 13 The pattern of missing cases in CFPS by math scores

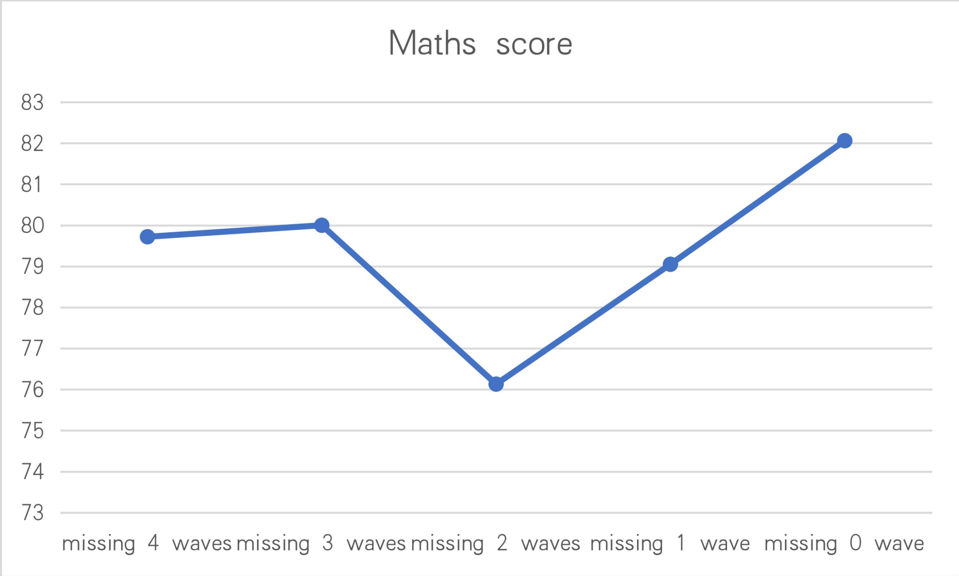


Figure 14 The pattern of missing cases in CFPS by word cognitive test scores

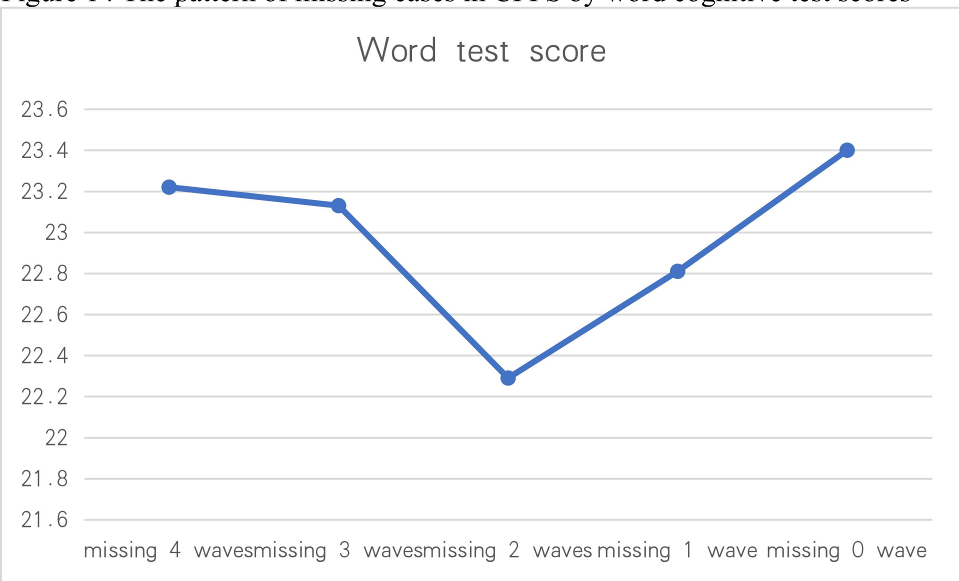
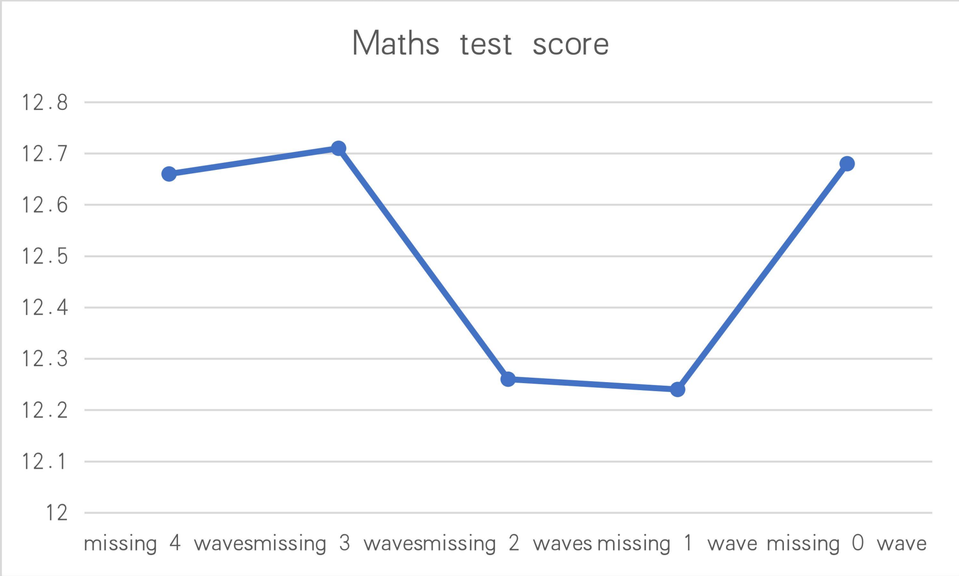


Figure 15 The pattern of missing cases in CFPS by math cognitive scores



Appendix 12 Exploring the *Clustering of High School Participation by Currently Employed Indicators*

Table 1 Percentages of HS participation and non-HS participation by hukou status

	Total (N)	HS (%)	Non-HS (%)	Odds ratio
Urban hukou	484	61	39	2.3
Rural hukou	1,986	40	60	
No registration or hukou status missing	20	35	65	
Total	2,490	44.3	55.7	

Table 2 Percentages of HS participation and non-HS participation of students living in urban and rural areas

	Total (N)	HS (%)	Non-HS (%)	Odds ratio
Living in urban	929	54	46	1.8
Living in rural	1,561	39	61	
Total	2,490	44.3	55.7	

Table 3 Percentages of HS participation and non-HS participation by ethnicity

	Total (N)	HS (%)	Non-HS (%)	Odds ratio
Han ethnicity	2,213	46	54	2.0
Minority ethnicity	269	30	70	
Ethnicity missing	8	-	-	
Total	2,490	44.3	55.7	

Table 4 Percentages of HS participation and non-HS participation by hukou area

	Total (N)	HS (%)	Non-HS (%)	Odds ratio
Municipalities	116	48	52	
Northeastern China	266	43	57	1.2
Central China	388	43	57	1.2
Eastern China	268	46	54	1.1
Northern China	303	52	48	0.9
Southern China	322	43	58	1.3
Southwestern China	361	35	65	1.7
Northwestern China	457	48	52	1.0
Hukou province missing	9	-	-	
Total	2,490	44.3	55.7	

Table 5 Percentages of HS participation and non-HS participation by sex

	Total (N)	HS (%)	Non-HS (%)	Odds ratio
Female	1,215	48	52	1.4
Male	1,275	41	59	

Total	2,490	44.3	55.7	

Table 6 Percentages of HS participation and non-HS participation by birth month

	Total (N)	HS (%)	Non-HS (%)	Odds ratio
Autumn-born	1,157	45	55	1.0
Summer-born	1,333	44	56	
Total	2,490	44.3	55.7	

Table 7 Percentages of HS participation and non-HS participation by parental education qualifications

	Total (N)	HS (%)	Non-HS (%)	Odds ratio
Both parents had compulsory education qualifications	1,953	40	60	2.4
At least one parent had a post-compulsory education qualification	477	60	40	
At least one parent had a HE qualification	50	74	26	
Parents' education information missing	10	-	-	
Total	2,490	44.3	55.7	

Table 8 Percentages of HS participation and non-HS participation by parents' political status

		Total (N)	HS (%)	Non-HS (%)	Odds ratio
Mother	Party member	254	59	41	2.0
	Non-party member	2,129	43	57	
	Mother's political status missing	107	37	63	
Father	Party member	433	56	44	1.7
	Non-party member	1,948	42	58	
	Father's political status missing	109	44	56	
Total		2,490	44.3	55.7	

Table 9 Percentages of HS participation and non-HS participation by mother's occupation

	Total (N)	HS	Non-HS	Odds ratio
Skilled agricultural and fishery workers	1,287	39	61	
Plant and machine operators and assemblers	69	41	59	
Legislators, senior officials and managers	41	54	46	1.5
Professionals	68	72	28	
Technicians and associated professionals	28	61	39	
Clerks	30	63	37	
Service workers and shop and market sales workers	216	55	45	
Craft etc. trade workers	168	50	50	
Elementary occupations	54	56	44	
Mother's occupation missing	529	43	57	

Total	2,490	44.3	55.7	

Table 10 Percentages of HS participation and non-HS participation by father's occupation

	Total (N)	HS	Non-HS	Odds ratio
Skilled agricultural and fishery workers	1,056	38	62	
Elementary occupations	46	35	65	
Legislators, senior officials and managers	70	57	43	1.6
Professionals	67	73	27	
Technicians and associated professionals	35	57	43	
Clerks	15	53	47	
Service workers and shop and market sales workers	159	59	42	
Plant and machine operators and assemblers	279	51	49	
Craft etc. trade workers	400	47	53	
Father's occupation missing	363	41	59	
Total	2,490	44.3	55.7	

Table 11 Average family income and expenditure on education of students who did and did not attend HS

	HS	Non-HS
Average family income	21,793.76	18,746.17
Average expenditure on education	1,960.81	1,205.65

Table 12 Percentages of HS participation and non-HS participation by the language to communicate with the family

	Total (N)	HS	Non-HS	Odds ratio
Mandarin Chinese	426	48	52	1.3
A minority ethnicity language	619	49	51	
A Chinese dialect	1,350	42	58	
Another language	95	25	75	
Total	2,490	44.3	55.7	

Table 13 Percentages of HS participation and non-HS participation by school type

	Total (N)	HS	Non-HS	Odds ratio
Key school	168	78	22	31.1
Ordinary school	1,945	48	52	
Private school	60	48	52	
School type missing	317	3	97	
Total	2,490	44.3	55.7	

Table 14 Percentages of HS participation and non-HS participation by class type

	Total (N)	HS	Non-HS	Odds ratio
Ordinary class	530	51	49	22.2
Key class	262	71	29	

The school does not distinguish between ordinary and key classes	1,376	46	54	
Class type missing	322	4	96	
Total	2,490	44.3	55.7	

Table 15 Average academic ability scores of students who went to HS and those who did not

	HS	Non-HS
Chinese score (average)	0.34	-0.34
Mathematics score (average)	0.33	-0.34
Word test score (average)	30.10	25.73
Mathematics test score (average)	19.43	16.38

Table 16 Percentages of HS participation and non-HS participation by language for communication with classmates

	Total (N)	HS	Non-HS	Odds ratio
Mandarin Chinese	855	52	48	2.4
A Chinese dialect	1,329	47	53	
A minority ethnicity language	39	23	77	
Another language	40	35	65	
Language at school missing	227	2	98	
Total	2,490	44.3	55.7	

Appendix 13 Complete Results of Ten Regressions on the CFPS data

Table 1 Percentages predicted correctly for HE entry (Model 1; enter)

Estimate	1	2	3	4	5	6	7	8	9	10	Average
Base	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0
Block 1 (sex, hukou, ethnicity, birth months)	60.2	60.0	59.3	60.2	60.0	59.3	60.2	59.7	58.3	59.0	59.6
Block 2 (parents' education qualifications, father's occupation, mother's occupation, parents' political status, communication language at home, annual family income, annual expenditure on education)	64.5	64.0	63.5	63.1	63.3	63.8	64.8	64.8	63.4	63.6	63.9
Block 3 (school type, class type, communication language at school, word cognitive test score, maths cognitive test score)	71.1	72.2	70.7	71.3	71.1	71.5	72.0	72.9	69.0	70.8	71.3
Block 4 (living area, hukou province area)	72.1	73.0	71.6	72.1	72.4	72.5	72.1	73.3	69.8	71.4	72.0

N=1,220

Table 2 Percentages predicted correctly for HE entry (Model 2; enter)

Estimate	1	2	3	4	5	6	7	8	9	10	Average
Base	50.	50.	50.	50.	50	50.	50.	50	50.	50.	50.0

	0	0	0	0	.0	0	0	.0	0	0	
Block 1 (parents' education qualifications, father's occupation, mother's occupation, parents' political status, communication language at home, annual family income, annual expenditure on education)	63.1	64.3	63.0	62.1	62.9	63.7	64.1	63.0	61.1	62.0	62.9
Block 2 (sex, hukou, ethnicity, birth month)	64.5	64.0	63.5	63.1	63.3	63.8	64.8	64.8	63.4	63.6	63.9
Block 3 (school type, class type, communication language at school, word cognitive test score, maths cognitive test score)	71.1	72.2	70.7	71.3	71.1	71.5	72.0	72.9	69.0	70.8	71.3
Block 4 (living area, hukou province area)	72.1	73.0	71.6	72.1	72.4	72.5	72.1	73.3	69.8	71.4	72.0

N=1,220

Table 3 Percentages predicted correctly for HE entry (Model 3; enter)

Estimate	1	2	3	4	5	6	7	8	9	10	Average
Base	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0
Block 1 (school type, class type, communication language at school, word cognitive test score, maths cognitive test score)	71.2	69.8	70.0	69.7	70.3	70.8	70.3	69.9	67.5	69.8	69.9
Block 2 (sex, hukou, ethnicity, birth months)	72.0	71.1	70.7	71.1	72.0	70.8	71.4	72.8	68.0	71.1	71.1
Block 3 (parents' education qualifications, father's occupation, mother's occupation, parents' political status, communication language at home, annual family income, annual expenditure on education)	71.1	72.2	70.7	71.3	71.1	71.5	72.0	72.9	69.0	70.8	71.3

Block 4 (living area, hukou province area)	72.1	73.0	71.6	72.1	72.4	72.5	72.1	73.3	69.8	71.4	72.0

N=1,220

Table 4 Percentages predicted correctly for HE entry (Model 4; enter)

	1	2	3	4	5	6	7	8	9	10	Average
Block 1 (parents' education qualifications, father's occupation, mother's occupation, parents' political status, communication language at home, annual family income, annual expenditure on education)	63.1	64.3	63	62.1	62.9	63.7	64.1	63	61.1	62	62.9
Block 2 (sex, hukou, ethnicity, birth months)	64.5	64	63.5	63.1	63.3	63.8	64.8	64.8	63.4	63.6	63.9
Block 3 (school type, class type, communication language at school, word cognitive test score, maths cognitive test score)	71.1	72.2	70.7	71.3	71.1	71.5	72	72.9	69	70.8	71.3
Block 4 (living area, hukou province area)	72.1	73	71.6	72.1	72.4	72.5	72.1	73.3	69.8	71.4	72.0
Block 5 (high school participation or not)	78.9	77	76.4	76.6	76.1	76.9	76.8	77	74.4	76.1	76.6

Table 5 Percentages predicted correctly for HE entry (Model 5; enter)

	1	2	3	4	5	6	7	8	9	10	Average
Block 1 (high school participation or not)	73.7	74.3	72.7	72.3	74.1	72.7	74.5	72.9	71.5	72.9	73.2
Block 2 (parents' education qualifications, father's occupation, mother's occupation, parents' political status, communication language at home, annual family income, annual expenditure on education)	74.3	74.8	73.4	72.9	74.7	73.9	74.8	73.4	71.8	73.4	73.7
Block 3 (sex, hukou, ethnicity, birth months)	74.8	75.1	73.5	73.4	74.7	74.3	75.6	73.4	72	73.4	74.0
Block 4 (school type, class type, communication	78	7	76	76	76	7	76	7	74	75	76.5

language at school, word cognitive test score, maths cognitive test score)	.3	7.5	.3	.1	.5	6.6	.8	6.5	.1	.9	
Block 5 (living area, hukou province area)	78.9	77	76.4	76.6	76.1	76.9	76.8	77	74.4	76.1	76.6

Table 6 Percentages predicted correctly for high school education entry (Model 1; forward)

		Percentage
Base		55.7
Block 1	Step 1 (hukou status)	60.0
	Step 2 (ethnicity)	59.8
	Step 3 (sex)	59.8
Block 2	Step 1 (education expenditure)	61.8
	Step 2 (parental education qualifications)	62.5
	Step 3 (communication language at home)	62.8
	Step 4 (father’s occupation)	62.7
Block 3	Step 1 (school type)	69.3
	Step 2 (cognitive word test score)	69.9
	Step 3 (cognitive maths test score)	71.4
	Step 4 (class type)	72.0
Block 4	Step 1 (hukou province area)	72.9

N=2490

Table 7 Percentages correctly predicting high school education (Model 2; forward)

		Percentage
Base		55.7
Block 1	Step 1 (school type)	59.4
	Step 2 (cognitive maths test score）	68.5
	Step 3 (cognitive word test score)	69.3
	Step 4 (class type)	69.3
Block 2	Step 1 (communication language at home)	71.6
	Step 2 (education expenditure)	71.6
	Step 3 (parental education qualifications)	71.2
	Step 4 (father's occupation)	71.8
Block 3	Step 1 (ethnicity)	71.9
	Step 2 (sex)	71.7

Block 4	Step 1 (hukou province area)	73.0
	Step 2 (living area)	72.7

N=2490

Appendix 14 Complete Results of Ten Regressions on the Primary Survey Data

Table 1 Percentages of correct predictions of aspirations for HE (Model 1; forward)

		Percentage
Base		51.1
Block 1	Step1(sex)	57.1
	Step2 (hukou)	58.6
Block 2	Step1 (parental education qualifications)	58.1
Block 3	Step1(self-evaluation of school performance)	66.0

N=801

Table 2 Percentages of correct predictions of aspirations for HE (Model 3; forward)

		Percentage
Base		51.1
Block 1	Step 1 (parental highest education qualification)	56.4
Block 2	Step 1 (sex)	57.9
Block 3	Step 1 (self-evaluation of school performance)	66.0
Block 4	Step 1 (high school aspiration or not)	73.2

N=801

Table 3 Summary of percentages predicted correctly for aspirations for HS (Model 1; enter)

Estimates	1	2	3	4	5	6	7	8	9	10	Average
Base	50	50	50	50	50	50	50	50	50	50	50
Block 1 (sex, hukou, ethnicity)	55.4	53.9	53.9	56.1	52.6	54.9	56.9	54.4	56.4	53.2	54.8
Block 2 (parental education qualifications, parents' occupations)	60.6	56.2	56.2	59.5	56.7	58.5	59	60.1	59.7	59.7	58.6
Block 3 (self-evaluation of school performance)	71.2	64.8	64.8	69.9	67.6	63.5	65.2	67.6	66.5	71.4	67.3

N=386

Table 4 Summary of percentages predicted correctly for aspirations for HS (Model 2; enter)

Estimates	1	2	3	4	5	6	7	8	9	10	Average
Base	50	50	50	50	50	50	50	50	50	50	50
Block 1 (parental education qualifications, parents' occupations)	59.8	56.5	56.5	58.2	58.3	59.1	57.9	56.7	59.2	58.7	58.1
Block 2 (sex, hukou, ethnicity)	60.6	56.2	56.2	59.5	56.7	58.5	59	60.1	59.7	59.7	58.6
Block 3 (self-evaluation of school performance)	71.2	64.8	64.8	69.9	67.6	63.5	65.2	67.6	66.5	71.4	67.3

N=386

Table 5 Summary of coefficients for variables in the last step in ten binary logistic regression models to predict the aspirations for HS (Model 1; Enter)

Estimates	1	2	3	4	5	6	7	8	9	10	Average
female											
male	0.7	0.6	0.6	0.7	0.7	0.6	0.5	0.7	0.6	0.8	0.7
minority ethnicity											
Han ethnicity	0.5	0.7	0.7	0.5	0.7	0.7	0.6	0.5	0.6	0.5	0.6
Rural hukou											
Urban hukou	0.8	1.2	1.2	0.9	0.9	1.1	0.9	1.0	0.8	1.0	1.0
both parents only have compulsory education qualifications or below											
at least one parent has a bachelor's degree	2.6	1.7	1.7	3.1	2.6	2.5	2.1	2.5	3.1	2.3	2.4
at least one parent has a post-compulsory education qualification	1.6	1.2	1.2	1.4	1.6	1.5	1.4	1.4	1.5	1.6	1.4
parents blue-collar workers or peasants											
parents working in	0.9	0.7	0.7	0.9	0.7	0.6	0.7	1.0	0.9	0.7	0.8

traditionally good jobs											
parents with missing information on occupations	1.3	0.7	0.7	1.0	0.8	0.8	1.0	1.2	1.0	1.0	0.9
parents unemployed	0.5	0.4	0.4	0.3	0.3	0.4	0.4	0.6	0.5	0.5	0.4
parents self-employed	2.2	1.9	1.9	2.1	1.6	1.8	2.5	2.4	2.7	2.2	2.1
Self-evaluation of school performance	1.3	1.2	1.2	1.3	1.3	1.2	1.3	1.2	1.3	1.4	1.3

N=386

Table 6 Summary of coefficients for variables in the last step in ten binary logistic regression models to predict the aspirations for HS (Model 2; Enter)

	1	2	3	4	5	6	7	8	9	10	Ave rage
both parents only have compulsory education qualifications or below											
at least one parent has a bachelor's degree	2.6	1.7	1.7	3.1	2.6	2.5	2.1	2.5	3.1	2.3	2.4
at least one parent has a post-compulsory education qualification	1.6	1.2	1.2	1.4	1.6	1.5	1.4	1.4	1.5	1.6	1.4
parents blue-collar workers or peasants											
parents working in traditionally good jobs	0.9	0.7	0.7	0.9	0.7	0.6	0.7	1.0	0.9	0.7	0.8
parents with missing information on occupations	1.3	0.7	0.7	1.0	0.8	0.8	1.0	1.2	1.0	1.0	0.9
parents unemployed	0.5	0.4	0.4	0.3	0.3	0.4	0.4	0.6	0.5	0.5	0.4
parents self-employed	2.2	1.9	1.9	2.1	1.6	1.8	2.5	2.4	2.7	2.2	2.1
female											
male	0.7	0.6	0.6	0.7	0.7	0.6	0.5	0.7	0.6	0.8	0.7
minority ethnicity											
Han ethnicity	0.5	0.7	0.7	0.5	0.7	0.7	0.6	0.5	0.6	0.5	0.6
Rural hukou											

Urban hukou	0.8	1.2	1.2	0.9	0.9	1.1	0.9	1.0	0.8	1.0	1.0
Self-evaluation for school work	1.3	1.2	1.2	1.3	1.3	1.2	1.3	1.2	1.3	1.4	1.3

N=386

References

Allen, W. R., Teranishi, R., Dinwiddie, G., & González, G. (2002). Knocking at freedom's door: Race, equity and affirmative action in US Higher Education. *Journal of Public Health Policy*, 23(4), 440-452.

Alton, A. & Massey, A. (1998). Date of birth and achievement in GCSE and GCE A-level. *Educational Research*, 40(1), 105-109.

Arcidiacono, P., Lovenheim, M., & Zhu, M. (2015). Affirmative Action in Undergraduate Education. *Annual Review of Economics*, 7(1), 487–518.

Baguley, T. (2009). Standardized or simple effect size: What should be reported? *The British Psychological Society*, 100, 603-617.

Bai, J. (2018). 家庭子女数量与子女高等教育入学率关系的比较分析——以 CFPS 五省数据为例 A comparative analysis of the relationship between the number of children in a family and the enrolment of children in higher education - using data from the five provinces of the CFPS. *课程教育研究 Curriculum Education Research*, 02, 230.

Banerjee. P. A. (2018). Widening participation in higher education with a view to implementing institutional change. *Perspectives : Policy and Practice in Higher Education,* 22(3), 75–81.

Bao, M. D. L (2011).大众化视域下西部民族地区高等教育入学机会研究 A Study on Opportunities of Enrollment into Higher Learning Institutions in China ' s Western Nationality Areas under the Backdrop of Popular Views. *内蒙古民族大学学报(社会科学版)Journal of Inner Mongolia University of Nationalities (Social Science Edition),* 02,96-99.

Barker, E. (1948). *The Politics of Aristotle*. Clarendon press.

Bartholo, T. L. (2013). Measuring between-school segregation in an open enrollment system: the case of Rio de Janeiro. *Journal of School Choice*, *7*(3), 353-371.

Bartholo, T. L., & Costa, M. D. (2014). Shift allocation and school segregation: discussing intra-school inequalities. *Cadernos de Pesquisa*, *44*, 670-692.

Bartholo, T. L., & Costa, M. D. (2016). Evidence of a school composition effect in Rio de Janeiro public schools. *Ensaio: Avaliação e Políticas Públicas em Educação*, *24*, 498-521.

Bathmaker, A. M., Ingram, N., & Waller, R. (2013). Higher education, social class and the mobilisation of capitals: Recognising and playing the game. *British Journal of sociology of Education*, *34*(5-6), 723-743.

Bell, D. (1972). Meritocracy and equality. *The public interest*, 29, 29.

Berben, L., Sereika, S. M. & Engberg, S. (2012). Effect size estimation: Methods and examples. *International Journal of Nursing Studies*, 49, pp. 1039-1047.

Bie, D. R. & Zhu, X. G. (2003). 我国高等教育大众化道路上的公平问题研究 On the Equity Problems in Massification of Chinese Higher Education. *Peking University Education Review*, 1(3), 54-59.

Blanden, J., & Machin, S. (2004). Educational inequality and the expansion of UK higher education. *Scottish Journal of Political Economy*, 51(2), 230-249.

Boliver, V. (2011). Expansion, differentiation, and the persistence of social class inequalities in British higher education. *Higher education*, 61(3), 229-242.

Boliver, V. (2013a). Access to Britain's top universities is far from fair. *University world* *news.*, 268.

Boliver, V. (2013b). How fair is access to more prestigious UK universities?. *The British journal of sociology*, 64(2), 344-364.

Boliver, V. (2015). Lies, damned lies, and statistics on widening access to Russell Group universities. *Radical statistics.*, 113, 29-38.

Boliver, V. (2016). Exploring ethnic inequalities in admission to Russell Group universities. *Sociology*, 50(2), 247-266.

Boliver, V., Banerjee, P., Gorard, S., & Powell, M. (2022). Reconceptualising fair access to highly academically selective universities. *Higher Education*, 84(1), 85-100.

Boliver, V., Crawford, C., Powell, M., & Craige, W. (2017). Admissions in context: the use of contextual information by leading universities.

Boliver, V., Gorard, S., Powell, M., & Moreira, T. (2017). Mapping and evaluating the use of contextual data in undergraduate admissions in Scotland.

Boliver, V., Gorard, S., & Siddiqui, N. (2015). Will the use of contextual indicators make UK higher education admissions fairer?. *Education Sciences*, *5*(4), 306-322.

Boliver, V., Gorard, S., & Siddiqui, N. (2017). How can we widen participation in higher education? The promise of contextualised admissions. In *The university as a critical institution?* (pp. 95-109). SensePublishers, Rotterdam.

Boliver, V., Gorard, S., & Siddiqui, N. (2019). Using contextual data to widen access to higher education. *Perspectives: Policy and Practice in Higher Education*, *25*(1), 1-7.

Boliver, V., Gorard, S., & Siddiqui, N. (2022). Who counts as socioeconomically disadvantaged for the purposes of widening access to higher education?. *British Journal of Sociology of Education*, *43*(3), 349-374.

Booth, A. L. & Kee, H. J. (2009). Birth order matters. *Journal of Population Economics*, 22(2), 367–397.

Bratton, D. (1979). University Admissions Policies in China, 1970-1978. *University of California Press, Asian Survey*, 19(10), 1008-1022.

Butcher, J., Corfield, R., & Rose-Adams, J. (2012). Contextualised approaches to widening participation: a comparative case study of two UK universities. *Widening Participation and Lifelong Learning*, *13*(1), 51-70.

Callinicos. (2000). *Equality*. Polity Press.

Cardak, B. A., & Ryan, C. (2006). Why are High-Ability Individuals from Poor Backgrounds Under-Represented at University?. *Available at SSRN 914025*.

Cao, Y. & Zhang, R. J. (2016). 我國一流大學的入學機會及其地區差異: 2008-2015Access to top universities in China and regional differences: 2008-2015. *華東師範大學學報 (教育科學版)Journal of East China Normal University (Education Science Edition)*, *34*(4), 52-65.

Chen, A. L., Zheng, Y. F. & Xu, J. X. (2019). 教育能促进社会阶层代际流动吗?——基于中国综合社会调查 (CGSS) 的经验证据 Does Education Promote Intergenerational Mobility of Social Class?--Empirical Evidence Based on the CGSS. *Education & Economy*, 35(6), 27-34.

Chen, J. H. & Gao, H. (2019). 我国高等教育入学机会的县域差异及其影响因素研究——基于甘肃 2007-2016 年高考录取 A study of county differences in access to higher education in China and their influencing factors--based on 2007-2016 college entrance examination admissions in Gansu. *高教探索 Higher Education Exploration,* 1, 10-14.

Chen, L. (2019). 少数民族高考加分政策研究回顾与展望 The retrospect and expectation of extra score policy for minority ethnicity groups. *收藏 Collection*, Vol. *3*.

Chen, W. S. & Peng, X. H. (2018). 高校经费自给率与高校发展的关系研究——以教育部直属高校为例 A Study on the Relationship between the Self sufficiency Rate of University Funding and the Development of Universities: Taking Universities Directly under the Ministry of Education as an Example. *会计师 Accountant,* Vol. 05, pp. 65-67.

Chen, Y.Y. & Feng, S.Z. (2013). Access to public schools and the education of migrant children in China. *China Economic Review*, *26*, 75-88.

Cheng, Y. K. (2020). 教育公平视阈下高校本科招生生源状况研究 A Study on the Status of Undergraduate Enrollment from the Perspective of Educational Equity ——A Case Study of One University in Nanjing (硕士学位论文,南京邮电大学 Master dissertation of Nanjing University of Posts and Telecommunications). https://kns.cnki.net/KCMS/detail/detail.aspx?dbname=CMFD202101&filename=1020431813.nh

China Statistical Yearbook (2021). Compiled by National Bureau of Statistics of China. 中国统计年鉴—2021 (stats.gov.cn)

CNSDA (2020, 2021). http://www.cnsda.org/index.php.

Coleman, J. S. (1966). Equal schools or equal students? *The Public Interest; Summer 1966; 4, Periodicals Archive Online*, 70-75.

Coleman, J. S. (1968). EQUALITY OF EDUCATION OPPORTUNITY. *Integrated*

Education, 6:5, 19-28.

Coleman, J. S. (1975). Equal Educational Opportunity: A Definition. *Oxford Review of Education*, 1:1, 25-29.

Coulson, S., Garforth, L., Payne, G., & Wastell, E. (2017). Admissions, adaptations, and anxieties: Social class inside and outside the elite university. *In Higher Education and Social Inequalities (1st ed., Vol. 1, pp. 2–21)*. Routledge.

Cui, Y., Liu, H., & Zhao, L.Q. (2019). Mother's education and child development: Evidence from the compulsory school reform in China. *Journal of Comparative Economics*, *47*(3), 669-692.

Curtis, S., Blundell, C., Platz, C. & Turner, L. (2014). Successfully widening access to medicine. Part 1: recruitment and admissions. *Journal of the Royal Society of Medicine*, 107 (9), 342-346.

Davey, G., Lian, C. D. & Higgins, L. (2007). The university entrance examination system in China. *Journal of Further and Higher Education*, 31(4), 385-396.

Deng, D., Seepho, S., & Lian, A. (2020). Preferential Admission Policies for Ethnic Minority Students in Yunnan: Help or Hindrance. *International Journal of Learning, Teaching and Educational Research*, *19*(4), 356-376.

Deng, F. & Fu, W. X. (2020). 新中国 70 年教育不平等的演变轨迹实证研究——基于 CGSS 同期群数据的历史考察 Trajectories of Educational Inequality in China -- An Empirical Study Based on CGSS Database. *Education & Economy*, *36*(2), 37-48.

Deng, L. (2021). 农村留守儿童家庭教育现状与对策研究 Study on the current situation of family education for rural left-behind children and countermeasures. *成才之路 The road to* *success*, 2021(28), 50-51.

Department for Education (DfE). (2012). *Download of historical free school meals eligibility:user guide for the download of historical FSM eligibility data from DfE to school and local authorities.*

Ding, Y., Wu, Y., Yang, J., & Ye, X. (2021). The elite exclusion: stratified access and production during the Chinese higher education expansion. *Higher Education*, *82*(2), 323-347.

Dodd, E., Ellis, S., & Singh, S. (2020). Making the invisible, visible: a twenty-first century approach to tertiary preparation, attainment and access for student equity. *International Journal of Inclusive Education*, 1-21.

Dong, M. Y. (2009). 教育机会均等视阈下重点高校大学生来源的历史研究 The historical research on the source of university students in key universities from the view of equality of educational opportunity --taking a key university in Shanghai for example (博士学位论文,华东师范大学 PhD Thesis of East China Normal University). https://kns.cnki.net/KCMS/detail/detail.aspx?dbname=CDFD0911&filename=2009186528.nh

Dougherty, K. J., & Callender, C. (2017). English and American higher education access and completion policy regimes: similarities, differences and possible lessons.

Dougherty, K. J. & Callender, C. (2020). Comparing and learning from English and American higher education access and completion policies. *Policy Reviews in Higher Education*, 4(2), 203–227.

Downing, R., Lubensky, M. E., Sincharoen, S., Gurin, P., Crosby, F. J., Queirolo, S., & Franco, J. (2002). Affirmative action in higher education. *Diversity Factor*, *10*(2), 15-20.

Du, J. D., Ni, L. & Huang, Y. (2018). 个体受教育程度的影响因素研究——基于CFPS2014 数据的实证分析 A study of the factors influencing individual educational attainment - an empirical analysis based on CFPS2014 data. *高教论坛 Higher Education Forum,* 10,119-124.

Duan, W.J., Guan, Y., & Bu, H. (2018). The effect of parental involvement and socioeconomic status on junior school students' academic achievement and school behavior in China. *Frontiers in psychology*, 9, 952.

Eerduntu (2008). 和谐社会视阈下高考录取制度改革的思考——基于高等教育入学机会公平的视角 Reflections on the Reform of the College Entrance Examination Admission System from the Perspective of a Harmonious Society: From the Perspective of Fair Opportunities for Higher Education Enrollment. *黑龙江高教研究 Heilongjiang Researches on Higher Education*, Vol.12, pp. 8-12.

Eerduntu & Chen, B. T. E. (2009). 和谐视域下高等教育入学机会均等研究 A study on equal access to higher education in a harmonious context. *学术论坛 Academic Forum,* 02,180-184.

Evans, C. (2014). Impact and effectiveness of widening access to HE in wales (Vol. 11). Working Paper Series WAQLCE2014.

Fan, X. G. (2013). 关于我国户口制度改革的思考 Reflections on the reform of China's hukou system. *法治与社会 Rule of Law and Society*, 06, 36-38.

Fang, C. C. & Feng, X. T. (2018). 社会出身与教育获得——基于 CGSS 70 个年龄组数据的历史考察 Social Origins and Educational Attainment: A Historical Analysis Based on 70 Birth Cohorts from CGSS Datasets. *社会学研究 Sociological research,* 02,140-163+245.

Fang, C.C., & Feng, X.T. (2020). Social origin and educational attainment in China: a historical analysis of 70 birth cohorts. *The Journal of Chinese Sociology*, *7*(1), 1-18.

Fang, S., Huang, J., Wu, S. Y., Jin, M.C., Kim, Y.M., & Henrichsen, C. (2020). Family assets, parental expectation, and child educational achievement in China: A validation of mediation analyses. *Children and Youth Services Review*, 112, 104875.

Frisancho, V. & Krishna, K. (2016). Affirmative action in higher education in India: targeting, catch up, and mismatch. *Higher Education*, 71(5), 611–649.

Gao, D. D. (2017). 优质大学招生分省定额:实施现状、区域差异与优化路径 Provincial Quota System of the High Quality of Universities: Current Situation regional Differences and Optimization Path. *重庆高教研究 Chongqing Higher Education Research,* 02, 37-47.

Gao, H. (2018). 高校招生计划统筹分配的实践与策略研究——以甘肃省为例 A Study on the Practice and Strategy of Overall Allocation of College Enrollment Quotas: Taking Gansu Province as an Example. *中国考试 China Examinations*, 08,8-14.

Garrison-Wade, D. F. & Lewis, C. W. (2004). Affirmative Action: History and Analysis. *The Journal of College Admissions*, 185, 23.

Gil-Flores, Padilla-Carmona, M. T., & Suárez-Ortega, M. (2011). Influence of gender, educational attainment and family environment on the educational aspirations of secondary school students. *Educational Review (Birmingham)*, 63(3), 345–363.

Golley, J. & Kong, S. T. (2018). Inequality of opportunity in China’s educationl outcomes. *China Economic Review*, 51, 116-128.

Gorard, S. (2007). What does an index of school segregation measure? A commentary on Allen and Vignoles. *Oxford Review of Education*, 33:5, 669-677.

Gorard, S. (2013). *Research design: creating robust approaches for the social sciences*. SAGE.

Gorard, S. (2018). *Education policy, equity and effectiveness: Evidence of equity and effectiveness (1st ed.)*. Policy Press.

Gorard, S. (2021). *How to make sense of statistics (1st ed..)*. SAGE Publishing.

Gorard, S., Adnett, N., May, H., Slack, K., Smith, E., & Thomas, L. (2007). *Overcoming the barriers to higher education*. Stoke-on-Trent: Trentham Books.

Gorard, S., Boliver, V., Siddiqui, N., Banerjee, P. & Morris, P. (2017). Which are the most suitable contextual indicators for use in widening participation to HE?. *Working Paper. School of Education and School of Applied Social Sciences,* Durham University, Durham.

Gorard, S., Boliver, V., Siddiqui, N., & Banerjee, P. (2019). Which are the most suitable contextual indicators for use in widening participation to HE?. *Research Papers in Education*, *34*(1), 99-129.

Gorard, S. & Cook, T. (2007). Where does good evidence come from?. *International Journal of Research & Method in Education*, *30*(3), 307-323.

Gorard, S., See, B. H., & Davies, P. (2011). *Do attitudes and aspirations matter in education? : a review of the research evidence.* Lambert Academic Publishing.

Gorard, S., & See, B. H. (2013). *Overcoming Disadvantage in Education*. Routledge.

Gorard, S., Siddiqui, N., & Boliver, V. (2017). An analysis of school-based contextual indicators for possible use in widening participation. *Higher education studies.*, *7*(2), 79.

Gorard, S., Smith, E., & Benadusi, L. (2010). *Equity in education: an international comparison of pupil perspectives.* Palgrave Macmillan.

Gorard, S., Smith, E., May, H., Thomas, L., Adnett, N., & Slack, K. (2006). Review of widening participation research: addressing the barriers to participation in higher education.

Gorard, S. & Taylor, C. (2002). What is Segregation? A Comparison of Measures in Terms of 'Strong' and 'Weak' Compositional Invariance. *Sociology*, 36(4), 875-895.

Gou, R. M. (2006). 高等教育领域的平等权问题研究 Research on the right to equality in higher education（博士学位论文,吉林大学 PhD Thesis of Jinlin University). https://kns.cnki.net/KCMS/detail/detail.aspx?dbname=CDFD9908&filename=2006115856.nh

GOV. UK (2021, January), https://www.ethnicity-facts-figures.service.gov.uk/style-guide/writing-about-ethnicity

Gray, H., Lyth, A., McKenna, C. *et al.* Sex differences in variability across nations in reading, mathematics and science: a meta-analytic extension of Baye and Monseur (2016). *Large-scale Assess Educ* **7**, 2 (2019).

Gu, J. F. (2011). Spatial recruiting competition in Chinese higher education system. High Education, 63, 165-185.

Gu, X.R., & Yeung, W. J. J. (2020). Hopes and hurdles: Rural migrant children's education in urban China. *Chinese Sociological Review*, *52*(2), 199-237.

Guo, C., Tsang, M. C., & Ding, X. (2010). Gender disparities in science and engineering in Chinese universities. *Economics of Education Review*, *29*(2), 225-235.

Guo, L. L. (2010). 城乡高等教育公平的制度归因 Institutional Attribution of Equity in Urban and Rural Higher Education. *高等农业教育 Higher agricultural education*, 07,15-18.

Guo, Y. M., Song, Y., & Chen, Q.M. (2019). Impacts of education policies on intergenerational education mobility in China. *China Economic Review*, *55*, 124-142.

Guo, Y. Q. (2019). 高考加分政策对农村学生的影响 The Impact of the Higher Education Entrance Examination Plus Policy on Rural Students. *Research on Transmission Competence*, 261.

Hamnett, C., Shen, H. & Liang, B. J. (2019). The reproduction of regional inequality through university access: the Gaokao in China. *Area Development and Policy*, 4(3), 252-270.

Hannum, E., An, X., & Cherng, H. Y. S. (2011). Examinations and educational opportunity in China: Mobility and bottlenecks for the rural poor. *Oxford Review of Education*, 37(2), 267-305.

Hao, L., Hu, A., & Lo, J. (2014). Two aspects of the rural-urban divide and educational stratification in China: A trajectory analysis. *Comparative Education Review*, 58(3), 509-536.

Harkonen, J. (2014). Birth order effects on educational attainment and educational transitions in West Germany. *European Sociological Review*, 30(2), 166–179.

Harrison, N. (2017). Patterns of participation in a period of change: Social trends in English higher education from 2000 to 2016. In *Higher education and social inequalities* (pp. 54-80). Routledge.

Heckman, J. J., & Yi, J. J. (2012). *Human capital, economic growth, and inequality in China* (No. w18100). National Bureau of Economic Research.

Higher Education Law of the People's Republic of China (2018). http://www.npc.gov.cn/npc/c30834/201901/9df07167324c4a34bf6c44700fafa753.shtml

Hu, A.N. & Hibel, J. (2014). Changes in college attainment and the economic returns to a college degree in urban China, 2003–2010: Implications for social equality. *Social science research*, 44, 173-186.

Hu, A.N. & Vargas, N. (2015). Economic consequences of horizontal stratification in postsecondary education: evidence from urban China. *Higher Education*, *70*(3), 337-358.

Huang, G. H. C., & Gove, M. (2015). Confucianism, Chinese families, and academic achievement: exploring how Confucianism and Asian descendant parenting practices influence children's academic achievement. In *science education in East Asia* (pp. 41-66). Springer, Cham.

Huang, G.Y., Xie, Y., & Xu, H.W. (2015). Cognitive ability: Social correlates and consequences in contemporary China. *Chinese sociological review*, *47*(4), 287-313.

Hurst, E. (2015). 'The thing that kill us': student perspectives on language support in a South African university. *Teaching in Higher Education*, 20(1), 78-91.

Jacob, W. J. (2006). Social Justice in Chinese Higher Education:Regional Issues of Equity and Access. *Review of Education*. 52, 149-169.

Jerrim, J. (2021). Measuring Disadvantage. *Research Brief.* Sutton Trust.

Ji, L . & Zhu, W. G. (2011). 我国区域高等院校入学机会公平问题探析 A Study on the Fairness of Entrance Opportunities of Regional Universities and Colleges in China. *阅江学刊 The Journal of Yuejiang, 04*,69-75.

Jia, Q., & Ericson, D. P. (2017). Equity and access to higher education in China: Lessons from Hunan province for university admissions policy. *International Journal of Educational Development*, 52, 97-110.

Jian, C. M. & Peng, S. (2019). 家庭社会资本对农村中学生自我教育期望的影响分析——基于社会闭合理论的视角. Analysis of the Influence of Family Social Capital on Rural Middle School Students' Self-education Expectation. *中南财经政法大学研究生学报 Journal of Postgraduate Studies, Zhongnan University of Economics and Law*, (1), 101-107.

Jiang, F. (2020). 五大发展理念视阈下西南地区高等教育公平问题研究 Master's Dissertation of Chongqing University of Technology Research on Education Path of Ideological Politics in Universities From The Perspective of Five Development Concepts (硕士学位论文,重庆理工大学). https://kns.cnki.net/KCMS/detail/detail.aspx?dbname=CMFD202002&filename=1020060565.nh

Jiang, H. C. (2017). 迁移降低了儿童的教育机会吗:基于 CFPS 的实证分析 Does migration reduce children's educational opportunities: an empirical analysis based on CFPS. *青年探索 Youth Quest,* 01, 59-67.

Jiao, W. Q. (2020). 非民族学校少数民族学生数学阅读能力的研究策略 Strategies for studying the mathematical reading skills of ethnic minority students in non-ethnic schools. *数学教育 Mathematics Education*.

Jin, M. H. & Li, R. (2022). 家庭资本与个体高等教育获得相关性分析——基于中国家庭追踪调查 CFPS(2018)数据 Analysis of the correlation between household capital and individual access to higher education - based on data from the China Household Tracking Survey CFPS (2018). *山西青年 Shanxi Youth,* 02, 1-6.

Kandiko Howson, C., Cohen, E., & Viola, J. K. (2022). Inertia in elite STEM widening participation: the use of contextual data in admissions. *British Journal of Sociology of Education*, 1-20.

Kenty-Drane, J. L. (2009). Early isolation: Racial and economic segregation in US public elementary schools. *Race, Gender & Class*, 45-62.

Lai, G., Song, J., Wong, O., & Feng, X. (2016). Transition to higher education in contemporary China: A study of high school graduates in urban Nanjing. *Journal of Sociology*, 52(1), 83-102.

Lasselle, L., McDougall-Bagnall, J., & Smith, I. (2014). School grades, school context and university degree performance: Evidence from an elite Scottish institution. *Oxford Review of Education*, 40(3), 293-314.

Lei, X.Y., Shen, Y., Smith, J. P., & Zhou, G.S. (2017). Sibling gender composition's effect on education: Evidence from China. *Journal of population economics*, 30(2), 569-590.

Li, A. (2019). Unfulfilled promise of educational meritocracy? Academic ability and China's urban-rural gap in access to higher education. *Chinese Sociological Review*, 51(2), 115-146.

Li, D. K. (2002). Gender Inequality in Access to Knowledge in China. *The China Review*, 2(1), 121-147.

Li, D. X. & Lu, H. X. (2015). 高等教育机会获得与家庭资本的相关性研究——基于中国家庭追踪调查 CFPS 数据的分析 The Study of the Relativity between Higher Education Entrance Chance Acquirement and Family's Capital：Based on CFPS Data Binary Logistic Regression Analysis. *全球教育展望 Global Education Outlook,* 04,50-60.

Li, D. W. (2015).1999—2010 年农村学生高等教育入学机会的状况分析——基于 CGSS2010 项目数据 Rural Students'Access to Higher Education from 1999—2010: Based on CGSS2010 Project Data. *中国人民大学教育学刊 Journal of Education, Renmin University of China,* 02, 143-155.

Li, F. C., Wang, M. & Han, F. F. (2020). 云南省义务教育质量差距的实证研究——基于 CEPS 模型分析 An empirical study of the quality gap in compulsory education in Yunnan Province - an analysis based on the CEPS model.

Li, J. X. (2016). 不同地区家庭文化资本与高等教育机会获得关系研究 The relationship between family cultural capital and higher education opportunities in different regions (硕士学位论文,广西大学 MA Thesis of Guangxi University) https://kns.cnki.net/KCMS/detail/detail.aspx?dbname=CMFD201701&filename=1016195603.nh

Li, L. H. (2013). 高考制度恢复后河南省高等教育入学机会的公平研究 Fairness Research of Henan Province University Entrance Examination after the Recovery of Higher Education (硕士学位论文，福建师范大学 MA Thesis of Fujian Normal University). https://kns.cnki.net/KCMS/detail/detail.aspx?dbname=CMFD201401&filename=1013349441.nh

Li, M. Z. & Wang, Z. Z. (2017). 跨越四十年:中国高等教育机会公平政策演进、现状测度与提升路径 Over 40 Years: Policies development, Status Quo and Improvement Path of Entrance Opportunity Equity for Higher Education. *国家教育行政学院学报 Journal of the National Academy of Educational Administration*, 10,52-57.

Li, W. L., & Min, W. F. (2001). Tuition, private demand and higher education in China. PhD Thesis of Peking University.

Li, W. L. (2006). 高等教育财政政策对入学机会和资源分配公平的促进 The contribution of higher education fiscal policy to equity in access and resource allocation. *北京大学教育评论 Peking University Education Review,* 4(2), 34-46.

Li, W. L. (2007). Family background, financial constraints and higher education attendance in China. *Economics of Education Review*, 26(6), 724-734.

Li, W. Y., & Xie, Y. (2020). The influence of family background on educational expectations: a comparative study. *Chinese Sociological Review*, 52(3), 269-294.

Li, X. F. (2011). 当代中国的户籍分层研究——基于中国综合社会调查 (CGSS2006) 数据的实证分析 Household Stratification in Contemporary China - An Empirical Analysis Based on China General Social Survey (CGSS2006) Data. *云南行政学院学报 Journal of Yunnan Administrative College*, 13(1), 94-97.

Li, X. L. (2014). Questioning China's preferential college admission policies. *Widening Participation and Lifelong Learning*, 16(1), 70-90.

Li, Y. A., Whalley, J., Zhang, S., & Zhao, X. (2011). The higher educational transformation of China and its global implications. *The World Economy*, 34(4), 516-545.

Li, Z. J. & Yang, Z. (2011). 我国高考加分政策演变的制度分析-基于历史制度主义的分析范式 The Transition of Extra Credit Policies in the China College Entrance Examination: The Analysis from the Perspective of Institutionalism Paradigm. *Tsinghua Journal of Education*, 32(1), 61-67.

Ling, M.H. (2015). "Bad students go to vocational schools!": Education, social reproduction and migrant youth in urban China. *The China Journal*, 73, 108-131.

Ling, M.H. (2017). Returning to no home: Educational remigration and displacement in rural China. *Anthropological Quarterly*, 715-742.

Liang, Y. (2010).高等教育入学机会公平研究 The research of the Enrollment Equality In the Higher Education (硕士学位论文,华南理工大学 MA Dissertation of South China University of Technology). https://kns.cnki.net/KCMS/detail/detail.aspx?dbname=CMFD2011&filename=2010229323.nh

Liang, Y. C. & Wu, X. Y. (2016). 教育中的户籍隔离与教育期望——基于 CEPS 2014 数据的分析. Education Segregation and Educational Aspiration under China's Household Registration System: A Study Based on CEPS2014, *社会发展研究 Social Development Research,* (1), 22-47.

Liu, B. Z., Zhang, Y. Y. & Li, J. X. (2014). 社会经济地位, 文化观念与家庭教育期望 Socioeconomic Status, Cultural Idea and Family Educational Expectation. *青年研究 Youth Studies*, 6, 46-55.

Liu, B. Z., Zhang, Y. Y. & Li, J. X. (2015). 家庭社会经济地位与青少年教育期望: 父母参与的中介作用 Family SES and Adolescent Educational Expectation: Mediating Role of Parental Involvement. *Peking University Education Review*, 13(3), 158-176.

Liu, C. (2015). Regional Disparity in higher education enrollment in China. BA thesis (REGIONAL DISPARITY IN HIGHER EDUCATION ENROLLMENT IN CHINA
A Thesis submitted to the Faculty of the Graduate School of Arts and Sciences of Georgetown University in partial fulfillment of the requirements for the degree of
Master of Public Policy In Public Policy By Chang Liu, B.A. Washington, DC April 15, 2015)

Liu, F. T. (2013). 关于我国高考制度的深度思考 Deep thinking on China's college entrance exam system -- in the pespective of education fairness (硕士学位论文,哈尔滨师范大学 MA

Dissertation of Harbin Normal University).
https://kns.cnki.net/KCMS/detail/detail.aspx?dbname=CMFD201401&filename=1013312938.nh

Liu, H. (2014, July). Uneven Distribution of Educational Resources and Educational Inequality in Rural China, 1978-2008. In East Asia Junior Sociologists Forum, Yokohama, Japan.

Liu, H. F., & Li, M. Z. (2014). 高考分省定额制的形成与调整 The Formation and Adjustment of Provincial Quota System of College Entrance Examination. *教育研究*, 35(6), 73-80.

Liu, L. T. (2019). 寒门如何出"贵子"——基于家庭背景和高等教育的获得关系研究 How to produce "noble children" from poor families - a study on the relationship between family background and access to higher education. *青年与社会 Youth and Society,* 17, 250-251.

Liu, Q. M. (2020). 高考移民：高校招生地区不平等下的学生移民现象 Higher education migration: student migration under regional inequalities in university admissions.(硕士学位论文 暨南大学 MA Dissertation, Jinan University)

Liu, Y. X., Jiang, Q.B., & Chen, F.N. (2020). Children's gender and parental educational strategies in rural and urban China: the moderating roles of sibship size and family resources. *Chinese sociological review*, 52(3), 239-268.

Liu, H.F., & Wu, Q. (2006). Consequences of college entrance exams in China and the reform challenges. *KEDI Journal of Educational Policy*, 3(1).

Liu, J. (2011). EXPANSION AND EQUALITY IN ACCESS TO CHINESE HIGHER EDUCATION: A CULTURAL PERSPECTIVE (A Thesis submitted in conformity with the requirements For the degree of Doctor of Philosophy Department of Theory and Policy Studies in Education Ontario Institute for Studies in Education University of Toronto)

Liu, J. M. (2006). 高等教育扩展与入学机会差异: 1978~ 2003Higher education expansion and access disparities: 1978 to 2003. *社会Society*, 26(3), 158-179.

Liu, J. M. (2007). 扩招时期高等教育机会的地区差异研究 Examination of Regional Disparities of Higher Education Opportunity during the Rapid Expansion Period. *Peking University Education Review*. 5 (4), 144-155.

Liu, J. Q. (1998). Education of Females in China: Trends and Issues. *The Journal of Educational Thought*, 32 (1), pp 43-55.

Liu, N. N. (2020). 我国优质高等教育入学机会的地区差异研究 A study of regional differences in access to quality higher education in China. *重庆高教研究 Chongqing Higher Education Research,* 01, 37-46.

Liu, X. (2011). 从我国高考录取区域差异看教育不公问题的存在 The Existence of Educational Inequity from Regional Differences in China's College Entrance Examination Admissions. *黑龙江教育学院学报 Journal of Heilongjiang College of Education,* 11, 26-28.

Liu, X. H.(2014). 改革开放以来湖南高等教育发展研究（1978-）Study on the development of higher education in Hunan since the reform and opening-up (1978-) --based on the analysis of some important problems (博士学位论文，湖南师范大学). https://kns.cnki.net/KCMS/detail/detail.aspx?dbname=CDFD1214&filename=1014244982.nh

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, 108-117.

Liu, Y. S. & Wang, Z. M. (2009). Women entering the elite group: A limited progress. *Frontiers of Education in China*, 4(1), 27-55.

Liu, Y. S., Wang, Z. M. & Yang, X. F. (2012). Selecting the Elite: Status, Geography and Capital—Admission of Rural Students into Peking University (1978-2005). *International Journal of Chinese Education*, 1(1), 19-53.

Lu, X. F., Deng, F. & Guo, J, R. (2016). 高等教育扩招对入学机会均等化的影响 The impact of higher education expansion on equalisation of access. *北京大学教育评论 Peking University Education Review*, 14(3), 131-143.

Lu, Y. X. (2018). 高等教育入学机会的公平性研究 The Research Of The Enrollment Equality In The Higher Education (硕士学位论文,华东政法大学 MA Dissertation of East China University of Political Science and Law). https://kns.cnki.net/KCMS/detail/detail.aspx?dbname=CMFD201901&filename=1019804633.nh

Llie, S., Sutherland, A., & Vignoles, A. (2017). Revisiting free school meal eligibility as a proxy for pupil socio‐economic deprivation. *British Educational Research Journal,* 43(2), 253-274.

Long, A. B. & Fan, W. (2013). 我国教育公平研究的现状及特点 The Status Quo and the Characteristics of Education Equity Research in China. *现代教育管理 Modern education management,* 1, 16-21.

Lu, Y., & Zhou, H. (2013). Academic achievement and loneliness of migrant children in China: School segregation and segmented assimilation. *Comparative education review*, 57(1), 85-116.

Lucas, S. R. (2001). Effectively maintained inequality: Education transitions, track mobility, and social background effects. *American journal of sociology*, 106(6), 1642-1690.

Luo, L. Z. (2011). 高校招生考试制度对城乡子女高等教育入学机会差异的影响 The Influence of College Admission System on the Difference of Access to Higher Education between Children in Rural and Urban Areas. *高等教育研究 Higher Education Research*, 01,32-41.

Luo, X. Y. (2020). 基于教育公平理论的高考移民问题分析 An analysis of the problem of migration to higher education based on the theory of educational equity. *教学与管理 Teaching and Management*, 2020(03), 23-26.

Luo, Y., Guo, F., & Li, R. (2021). Gender equality and expansion of higher education: Testing effectively maintained inequality theory against the case of China. *International Journal of Educational Research*, 110, 101855.

Ma, L. P. & Pu, S. C. (2019). 重点大学自主招生政策的选拔效果分析 The Effects of the Independent Admission Policy of Talents Selection in Chinese Elite Universities. *北京大学教育评论 Peking University Education Review*, 17(2), 109-126.

Ma, W. H. (2012). Why the rural poor get fewer opportunities to leading research universities?. *Asia Pacific Education Review*, *13* (2), 263-271.

Machin, S. (2006). Social disadvantage and education experiences.

McCowan, T. (2007). Expansion without equity: An analysis of current policy on access to higher education in Brazil. *Higher education*, *53*(5), 579-598.

McCulloch, A. (2017). EDUCATIONAL ASPIRATIONS TRAJECTORIES IN ENGLAND. *British Journal of Educational Studies*, 65(1), 69–85.

Milburn, A. (2012). *University challenge: how higher education can advance social mobility*. London: Cabinet Office.

Mill, J. S. (2014). *Utilitarianism*. Cambridge University Press.

MOE (2020, January)
http://www.moe.gov.cn/s78/A03/moe_560/jytjsj_2018/gd/201908/t20190812_394204.html.

MOE. (1998). *Opinions on the enrolling of colleges and universities in 1977.*

MOE. (2020a). *2021 年普通高等学校部分特殊类型招生基本要求 Basic requirements for some special types of admissions to ordinary higher education institutions in 2020*.

MOE. (2020b). *2021 年普通高等学校招生工作规定 Regulations on enrolment of regular institutions of higher education in 2020.*

MOE. (2021a). 2022 年普通高等学校部分特殊类型招生基本要求 *Basic requirements for some special types of admissions to ordinary higher education institutions in 2022.*

MOE. (2021b). *2021 年普通高等学校招生工作规定 Regulations on enrolment of regular institutions of higher education in 2021.*

MOE. (2022). http://www.moe.gov.cn

MOE, NEAC, MPS & GASCAST. (2014). *关于进一步减少和规范高考加分项目和分值的意见 Opinions on further reducing and regulating the number and value of extra points in the college entrance examination*.

Mountford-Zimdars, A., Moore, J. & Graham, J. (2016). Is contextualized admission the answer to the access challenge? *Perspectives: policy and practice in higher education*, 20 (4), 143-150.

Mountford-Zimdars, A & Moore, J. (2020). Identifying merit and potential beyond grades: opportunities and challenges in using contextual data in undergraduate admissions at nine highly selective English universities. *Oxford Review of Education*, 46(6), 752–769.

Mullen, F. (2011). Widening access to higher education: admissions (SPICe Briefing; 11/07).

Nahai, N. R. (2013). Is meritocracy fair? A qualitative case study of admissions at the University of Oxford. *Oxford Review of Education*, 39 (5), 681-701.

National Bureau of Statistics (NBS). (2021)
http://www.stats.gov.cn/tjzs/cjwtjd/201308/t20130829_74322.html.

Ngok, K. (2007). Chinese education policy in the context of decentralization and marketization: Evolution and implications. *Asia Pacific Education Review*, 8(1), 142-157.

Ngok, K. (2008). Massification, bureaucratization and questing for "world - class" status: Higher education in China since the mid - 1990s. *International Journal of Educational Management*.

Ning, W. W. (2018). 教育分层、阶层地位获得与阶层流动——基于 CGSS2011 的实证分析 Educational stratification, class status acquisition and class mobility - an empirical analysis based on CGSS2011. *武汉科技大学学报(社会科学版)Journal of Wuhan University of Science and Technology (Social Science Edition),* 06, 668-674.

Orlans, H. (1992). Affirmative action in higher education. *The Annals of the American Academy of Political and Social Science*, 523(1), 144-158.

Ou, D., & Hou, Y. (2019). Bigger pie, bigger slice? The impact of higher education expansion on educational opportunity in China. *Research in Higher Education*, 60(3), 358-391.

Owen, L. E., Anderson, S. A. & Dowell, J. (2018). Free text adversity statements as part of a contextualised admissions process: a qualitative analysis. *BMC Medical Education*, 18:58. 1-11.

Pan, K. F., Xu, S., Chen, Y., Kang, L. & Lan, Y. H. (2010). 央属高校招生名额分配的原则和方案设计 Principles and scheme design for the allocation of places in central universities. *北京大学教育评论 Peking University Education Review,* 8 (2), 43-55.

Pan, Y. H. & Wu, M. (2020).教育获得的家庭决策影响因素及其演变——基于 CFPS 数据的应用 Influencing Factors and Evolution of Family Decision Making on Educational Attainment. *江西师范大学学报(哲学社会科学版)*Journal of Jiangxi Normal University (Philosophy and Social Science Edition), 05,104-114.

Pang, X. (2019).我国教育代际传递的城乡差异研究 Study on the Urban-Rural Differences between Intergenerational Transmission of Education in China -- The Analysis of based on the Data of Chinese General Social Survey (CGSS) (硕士学位论文,云南大学 MA Dissertation of Yunnan University). https://kns.cnki.net/KCMS/detail/detail.aspx?dbname=CMFD202001&filename=1019202473.nh

Peng, Z. P. & Jin, Y. (2015). 高考加分政策的历史透视及其治理 A Historical Perspective of the Higher Education Entrance Examination Plus Policy and its Governance. *中国教育学刊 Chinese Journal of Education*, 11, 15--19+24.

Pickering, N. (2019). No such thing as a free lunch: widening participation in English higher education for those on free school meals. *Widening Participation and Lifelong Learning*, 21(3), 57-80.

Princeton University. (2022a). In an extraordinary year, Princeton offers admission to 1,498 students for the Class of 2025

Princeton University. (2022b). First-generation, lower-income alumni raise awareness during FLi is Fly Week | Princeton Alumni

Qi, D., & Wu, Y.C. (2020). Family's social economic status and child educational outcomes in China: The mediating effects of parenting practices and children's learning attitudes. *Children and Youth Services Review*, 118, 105387.

Qian, X. L. & Smyth, R. (2008). Measuring regional inequality of education in China: widening coast-inland gap or widening rural-urban gap? *Journal of International development, J. Int. Dev*. 20, 132-144.

Qiao, J. Z. (2007).优质高等教育入学机会分布的区域差异 Regional Differentiation of Entrance Opportunity f or High -quality Higher Education. *北京师范大学学报(社会科学版)Journal of Beijing Normal University (Social Science Edition)*, 01, 23-28.

Qin, X.L., & Buchanan, R. (2019). Policy and public preferences regarding the university enrollment quotas system across Chinese provinces. *Higher Education Policy*, 1-21.

Qu, X. F. (2009). 部属高校"地方化"与受教育权平等 Localisation of ministerial universities and equality of access to education, *法学 Law Science,* 11, 25-29.

Rawls. (1999). *A theory of justice (Rev. ed.)*. Oxford University Press.

Reilly, D., Neumann, D. L., & Andrews, G. (2019). Gender Differences in Reading and Writing Achievement: Evidence From the National Assessment of Educational Progress (NAEP). *The American Psychologist*, 74(4), 445–458.

Riddell, S., Raffe, D., Croxford, L., Weedon, E., & Minty, S. (2013). Widening access to higher education: Scotland in UK comparative context.

Salmi, J., & Bassett, R. M. (2014). The equity imperative in tertiary education: Promoting fairness and efficiency. *International Review of Education*, 60(3), 361-377.

Sandel. (2010). *Justice: what's the right thing to do?* Penguin Books.

Saunders, P. (1997). Social mobility in Britain: an empirical evaluation of two competing explanations. *Sociology*, 31(2), 261-288.

Schwartz, S. (2004). Fair admissions to higher education: recommendations for good practice. Department for Education and Skills.

Shang, C.M. (2018).扩招以来我国高等教育入学机会差异的研究 Research on the Differences of Entrance Opportunity for Higher Education in China since Enrollment Expansion on an Empirical Analysis of Focusing on the Allocation of High-quality Higher Education Resources (硕士学位论文,云南大学 MA Dissertation of Yunnan University). https://kns.cnki.net/KCMS/detail/detail.aspx?dbname=CMFD202001&filename=1018250499.nh

Sheng, X. (2014). Parental expectations relating to children's higher education in urban China: Cultural capital and social class. *Journal of Sociology*, 50(4), 560-576.

Siddiqui, N. (2017). Socio-economic segregation of disadvantaged children between schools in Pakistan: comparing the state and private sector. Educational Studies, 43(4), 391-409.

Siddiqui, N., Boliver, V., & Gorard, S. (2019). Reliability of longitudinal social surveys of access to higher education: the case of next steps in England. *Social inclusion*., 7(1), 80-89. 8

Sina Education (2020), http://edu.sina.com.cn/

Song, B. Liu, H. & Wang, L. (2019).高校扩招、阶层分化与农户高等教育投资决策——基于 CGSS 数据的分析 Higher education expansion, stratification and farmers' investment decisions in higher education - an analysis based on CGSS data. *教育学术月刊/Educational Academic Monthly,* 12,101-108.

Song, H. X. (2009).社会阶层与高等教育机会均等的相关性研究 A study on the correlation between social class and equality of opportunity in higher education (博士学位论文,陕西师范大学).
https://kns.cnki.net/KCMS/detail/detail.aspx?dbname=CDFD0911&filename=2010235161.nh

State Education Commission (SEC). (1987). *普通高等学校招生暂行条例 Provisional Regulations on Admissions to General Institutions of Higher Education.*

Steinberg, S. (1995). *Turning back: The retreat from racial justice in American thought and policy*. Boston: Beacon Press.

Strand, S., Malmberg, L., & Hall, J. (2015). *English as an Additional Language (EAL) and educational achievement in England: An analysis of the National Pupil Database.*

Strom, B. (2004). *Students achievement and birthday effects. Paper prepared for presentation at the CESifo-Harvard University/PEPG Conference on "Schooling and Human Capital in the Global Economy: Revisiting the Equity-Efficiency Quandary,"* CESifo Conference Center, Munich, Spetember 3-4, 2004. Present Version, July 2004.

Su, J., Xiao, P. & Hu, Z. Y. (2019). 教育, 社会资本与农户家庭多维贫困转化——来自 CFPS 微观面板数据的证据 Education, Social Capital and the Multidimensional Poverty Transformation of peasant Households—Evidence from CFPS Micro Panel Data. *教育与经*

济*Education & Economy*, 35(2), 17-27.

Su, Y., Lau, C., & Rao, N. (2020). Early education policy in China: Reducing regional and socioeconomic disparities in preschool attendance. *Early childhood research quarterly*, 53, 11-22.

Sullivan, G. M. & Feinn, R. (2012). Using Effect Size--or Why the P Value Is Not Enough. *J Grad Med Educ*, 4 (3), 279-282.

Sun, F.S.& Barrientos, A. (2009). The Equity Challenge in China's Higher Education Finance Policy, *Higher Education Policy*, 22, 191-207.

Sun, Y. Q. & Yan, Y. (2015). 我国教育代际传递的城乡差异研究——基于中国家庭追踪调查(CFPS)的实证分析 The Research of Urban-Rural Difference of Intergenerational Transimission of Education: A Case Study Based on China Family. *北京师范大学学报(社会科学版)Journal of Beijing Normal University (Social Science Edition)*, 06,59-67.

Sundorph, E., Vasilev, D., & Coiffait, L. (2017). Joining the elite: How top universities can enhance social mobility.

Tam, T., & Jiang, J. (2015). Divergent urban-rural trends in college attendance: State policy bias and structural exclusion in China. Sociology of Education, 88(2), 160-180.

Tan, G. Y. (2013). Higher education reforms in China: For better or for worse?. *International Education*, 43(1), 101.

Tang, H. L. (2011). 促进高等教育公平的政府责任与对策研究 Research of Government Responsibility and Strategies of Equity in Higher education (博士学位论文,吉林大学 PhD Thesis of Jilin University).
https://kns.cnki.net/KCMS/detail/detail.aspx?dbname=CDFD0911&filename=1011094736.nh

Tang, J. (2016). “Lost at the starting line”: a reconsideration of educational inequality in China, 1978–2008. *The Journal of Chinese Sociology*, 3(1), 1-18.

Tao, D. J., Wang, J. P. & Zhang, K. Z. (2017).家庭规模、资源约束与子女教育分布——基于 CFPS 的实证研究 Family Size and Children’s Educational Distribution under Credit Constraints: Empirical Evidence from CFPS. 华中科技大学学报(社会科学版)Journal of Huazhong University of Science and Technology (Social Science Edition), 02, 45-54.

Taylor, C., Gorard, S. & Fitz, J. (2000). A re-examination of segregation indices in terms of compositional invariance.

The University of Nottingham. (2023). https://www.nottingham.ac.uk/ugstudy/applying/contextual-offers.aspx

The university of Manchester. (2023). https://www.manchester.ac.uk/study/undergraduate/contextual-admissions/admissions/

The University of Cambridge. (2023). https://www.undergraduate.study.cam.ac.uk/applying/contextual-data

The university of Liverpool. (2023). https://www.liverpool.ac.uk/undergraduate/applying/policies-and-statements/contextual-data/

The University of Surrey. (2023). https://www.surrey.ac.uk/apply/undergraduate/contextual-offers

Tierney, W. G. (1997). The parameters of affirmative action: Equity and excellence in the academy. *Review of educational research*, 67(2), 165-196.

Treadaway, M. (2014). Pupil Premium and the invisible group. *FFT Research Paper*, 5.

Treiman, D. J. (2012). The “difference between heaven and earth”: Urban–rural disparities in well-being in China. *Research in Social Stratification and Mobility*, 30(1), 33-47.

Trent, W., Owens-Nicholson, D., Eatman, T. K., Burke, M., Daugherty, J., & Norman, K. (2003). Justice, equality of educational opportunity, and affirmative action in higher education. *Compelling interest: Examining the evidence on racial dynamics in higher education.*

Tsang, E. Y. H. (2013). The quest for higher education by the Chinese middle class: retrenching social mobility?. *Higher Education*, 66(6), 653-668.

Tu, D. & Wu, G. Z. (2012).重点高校入学机会城乡不公原因析论 The Reasons Causing Inequality of Access to Key Universities Between Urban and Rural Students. 教育与考试 Education and Examinations, 04, 50-55.

Vignoles, A., & Murray, N. (2016). Widening participation in higher education. *Education Sciences*, 6(2), 13.

Wang, A., & Guo, D. (2019). Technical and vocational education in China: enrolment and socioeconomic status. *Journal of Vocational Education & Training*, 71(4), 538-555.

Wang, F. L. (2014). 中国的户口制度及其变革 China's Hukou System and its Transformation. *中国治理评论 China Governance Review*, 2014(01), 167-190.

Wang, G. R. & Zhang, X. (2014). 优质高等教育入学机会公平问题探析——基于“985”院校为例 Analysis on Enrollment Fairness of Superior Higher Education -- Based on 985 project Universities. *教育科学 Educational Sciences,* 05, 62-66.

Wang, H. (2016). 当代中国教育不平等的家庭背景因素分析——基于 2008CGSS 研究 Analysis of Family Background Factors of Educational Inequality in Contemporary China - Based on the 2008 CGSS Study.《今日财富》杂志社.(eds.)*2016 年第二届今日财富论坛论文集,* 400-401.《今日财富》杂志社 Fortune Today. (eds.) Proceedings of the 2nd Wealth Today Forum 2016, 400-401.

Wang, H. X. (2010). Research on the Influence of College Entrance Examination Policies on the Fairness of Higher Education Admissions Opportunities in China. *Chinese Education & Society*, 43 (6), 15-35.

Wang, L. (2011). Social exclusion and inequality in higher education in China: A capability perspective. *International Journal of Educational Development*, 31(3), 277-286.

Wang, L. M. & Zhao, H. X. (2020). 社会分层、高校扩招与高等教育机会获得——基于 CGSS2015 的实证分析 Social Stratification, College Enrollment Expansion and Access to Higher Education ——Empirical analysis based on CGSS2015.(eds.)*全面建成小康社会与青少年发展——第十六届中国青少年发展论坛优秀论文集,* 210-225.

Wang, L. X.(2019). 我国优质高等教育入学机会的区域差异研究 Study on the Regional Differences of the Entrance Opportunities for High-quality Higher Education in China: Based on the Analysis of the Enrollment Data of “985 Project” University from 2013 to 2018 (硕士学位论文,华中科技大学 MA Dissertation of Huazhong University of Science and Technology). https://kns.cnki.net/KCMS/detail/detail.aspx?dbname=CMFD202001&filename=1019617646.nhA

Wang, M. Q. (2018). 我国少数民族高考优惠政策的实施现状及优化策略———基于高考加分 “瘦身” 背景 The current situation of the implementation of preferential policies for China's ethnic minorities in the college entrance examination and the optimization strategy -

based on the background of "slimming" of extra points in the college entrance examination. *黑龙江民族丛刊 Heilongjiang Ethnic Series*, 5, 183-186.

Wang, P. C. & Gong, X. (2018). 家庭收入与学前教育机会--基于 CFPS 数据的实证研究 Family Income and Preschool Attendance: An Empirical Research based on CFPS Survey Data, *教育发展研究 Education Development Study*, 38(15), 18-26.

Wang, S. Y. & Du, Y. H. (2013). 高等教育入学机会地域不公平研究 A study of geographical inequity in access to higher education. *Journal of national academy of education administration,* 5, 19-24.

Wang, Y. (2018). 家庭背景与高等教育机会获得关系的研究 A study of the relationship between family background and access to higher education opportunities. *时代金融 Times Financial,* 30, 316+319.

Wang, Y. P. (2021). Closing the gender gap in college attendance: Variation by family background in China over time. *Social Science Research*, 98, 102578.

Wang, Y. M. (2013). National Higher Education Entrance Examination, Education in China: Educational History, Models, and Initiatives. Ruth Hayhoe & Heidi Ross (author), Zha, Qiang (ed). Berkshire.

Wang, Y. P. (2021). Who Benefits More from the College Expansion Policy? Evidence from China. *Research in Social Stratification and Mobility*, 71, 100566.

Wang, X. (2001). A Policy Analysis of the Financing of Higher Education in China: Two decades reviewed, *Journal of Higher Education Policy and Management*, 23 (2), 205-217.

Wang, X., & Liu, J. (2011). China's higher education expansion and the task of economic revitalization. *Higher Education*, 62(2), 213-229.

Wang, X. B., Liu, C.F., Zhang, L.X., Luo, R.F., Glauben, T., Shi, Y.J., … (Rozelle, S.) & Sharbono, B. (2011). What is keeping the poor out of college? Enrollment rates, educational barriers and college matriculation in China. *China Agricultural Economic Review*.

Wang, X. H., Pan, K. F. & Wu, Q. X. (2017). 高水平大学农村和贫困地区专项计划学生的学业表现研究——以 A 大学为例 Study on Students' Academic Performance of the Rural and Poor Area Special Enrollment Program in High Level Universities: Evidence from A University. *国家教育行政学院学报 Journal of the National Academy of Educational Administration*, (5), 66-75.

Weedon, E. (2014). Widening participation to higher education of under-represented groups in Scotland: The challenges of using performance indicators. *Centre for Research in Education, Inclusion and Diversity*.

Wei, H. L. (2016). 教育公平与家庭背景, 性别, 地区差异——基于 CGSS2013 的调查数据 Educational equity and family background, gender, and regional differences - based on CGSS2013 survey data. *东北农业大学学报: 社会科学版 Journal of Northeast Agricultural University (Social Science Edition)*, (5), 30-34.

Wei, Y. & Ma, X. (2017). 中学生自我教育期望的影响因素研究——基于 CEPS 的实证分析 A study on the factors influencing secondary school students' self-education expectations - an empirical analysis based on CEPS. *教育学术月刊 Educational Academic Monthly*, (10), 69-78.

Wei, Y. M.(2016).全面建成小康社会目标下中国西部高等教育入学机会公平研究 Study on the equity of entrance to higher education aiming at building a well-off society in an all-round way in the western regions (博士学位论文,兰州大学 PhD Thesis of Lanzhou University).
https://kns.cnki.net/KCMS/detail/detail.aspx?dbname=CDFDLAST2016&filename=1016725

851.nhA

Wei, Y. M. & Liu, X. C. (2015). 中国西部高等教育入学机会公平发展研究 Development of Equity Enrollment in Higher Education in the Western Regions of China. *高教发展与评估 Higher Education Development and Evaluation,* 03,1-8+112.

Wei, Y. Z. (2013a). 转型期中国区域差异与城市居民教育不平等(1978-2006)——基于CGSS2006 的多层线性模型分析 Regional Differences and Educational Inequality of Urban Residents in China during the Transition Period (1978-2006) - A Multi-Level Linear Model Analysis Based on CGSS2006. *教育学术月刊 Educational Academic Monthly,* 01,12-20.

Wei, Y. Z. (2013b). 地区经济社会发展水平与城市居民教育不平等(1978-2006)——基于CGSS2006 的多层线性模型的分析 Regional Variation in Educational Inequality in Urban China During the Period of Transition (1978 －2006): Hierarchical Linear Model Based on the data from CGSS2006. *青年研究 Youth Studies,* 02,82-93+96.

Wen, J. Y. (2022). 二元户籍、家庭背景与规模扩张下的高等教育获得——基于CGSS(2015)的实证分析 Dual Household Registration, Family Background and Access to Higher Education Under the Scale Expansion——Based on Empirical Analysis of CGSS (2015). *现代教育科学 Modern educational science,* 01,14-22.

Wong, K. K. & Nicotera, A. C. (2004). Brown v. Board of Education and the Coleman Report: Social Science Research and the Debate on Educational Equality. *Peabody Journal of Education,* 79(2), 122-135.

Wu, J., Zheng, Y. F. Wu, Z. X., Chen, N. D. & Xu, J. X. (2020). 教育获得的性别差异及其变动趋势研究——基于 CGSS2008～2015 数据分析 Gender Differences in Education and Their Changing Trends： Based on CGSS 2008～2015 Data Analysis. *西北人口 Northwest population,* 03,104-115.

Wu, L.L., Yan, K., & Zhang, Y.Q. (2020). Higher education expansion and inequality in educational opportunities in China. *Higher Education*, 80(3), 549-570.

Wu, X. (2017). Higher education, elite formation and social stratification in contemporary China: Preliminary findings from the Beijing College Students Panel Survey. *Chinese Journal of Sociology*, 3(1), 3-31.

Wu, X., Zhang, H., & Hu, Y. (2021, October). The Urban-Rural Education Inequality in 21st Century China. In 2021 6th International Conference on Modern Management and Education Technology (MMET 2021) (pp. 332-340). Atlantis Press.

Wu, X., & Zheng, B. (2018). Household registration, urban status attainment, and social stratification in China. *Research in Social Stratification and Mobility*, 53, 40-49.

Wu, X. G. (2012). The Household Registration System and Rural-Urban Educational Inequality in Contemporary China. *Chinese Sociological Review*, 44 (2), Winter 2011-2012, 31-51.

Wu, X.G., & Zhang, Z.N. (2010). Changes in educational inequality in China, 1990–2005: Evidence from the population census data. *In Globalization, changing demographics, and educational challenges in East Asia* (17, 123-152). Emerald Group Publishing Limited.

Wu, W. (2019). 新中国成立以来我国高等教育城乡不均衡的最新演变——基于CGSS2015数据的分析 The Latest Evolution of Urban-Rural Inequality in Higher Education in China--Based on the Analysis of CGSS2015 Data. *中国农业大学学报(社会科学版)Journal of China Agricultural University (Social Science Edition),* 05,129-136.

Wu, Y. X. & Huang, C. (2015). 中国教育获得性别不平等的城乡差异研究——基于CGSS2008 数据 A study of urban-rural differences in gender inequality in access to education in China - based on CGSS2008 data. *国家行政学院学报 Journal of the National*

School of Administration, 02,41-47.

Xiang, L. L., Stillwell, J., Burns, L. & Heppenstall, A. (2020). Measuring and Assessing Regional Education Inequalities in China under Changing Policy Regimes. *Applied Spatial Analysis and Policy*, 13, 91-112.

Xie, A. L., Hong, Y. B., Kuang, H. & Bai, J. R. (2018). “寒门贵子”: 文化资本匮乏与精英场域适应——基于 “985” 高校农村籍大学生的追踪研究 Cultural Capital Deficiency as Challenges: Rural Students in Elite Universities. *北京大学教育评论 Peking University Education Review*, 16(4), 45-64.

Xie, Y., & Hu, J. W. (2014). An introduction to the China family panel studies (CFPS). *Chinese sociological review*, 47(1), 3-29.

Xie, Z. J. & Liu, H. J. (2019). 少数民族高考加分政策实施效果评价———基于西部四所民族院校汉族学生与少数民族学生的对比调查 Evaluation on the Effect of the Preferential Policy that Awards Bonus Points in National College Entrance Examination for Ethnic Students --Based on a Comparative Investigation of Han and Ethnic Students at Four Ethnic Colleges in Western China. *Journal of Research on Education for Ethnic Minorities*, 2, 31-41.

Xie, Z. X., Wang, W. H., & Chen, X. W. (2010). A Study of Women's Access to Higher Education in Rural and Urban China: An Analysis Across Different Types of Higher Education Institutions. *Chinese Education & Society*, 43(4), 32-40.

Xu, D., & Wu, X. (2022). Separate and unequal: hukou, school segregation, and educational inequality in urban China. *Chinese Sociological Review*, 1-25.

Xu, D. T. (2021).教育公平视角下我国居民学历教育的代际传递研究 A study on the intergenerational transmission of academic education among Chinese residents in the perspective of educational equity (硕士学位论文,南京财经大学).

https://kns.cnki.net/KCMS/detail/detail.aspx?dbname=CMFDTEMP&filename=1022401979.nh

Xu, J. L. (2007). 和谐社会视域下高考录取制度公平的构建 The Construction of Fairness in the Admission System of Higher Education Examination under the Perspective of Harmonious Society. *国家教育行政学院学报 Journal of the National Academy of Educational Administration,* 10, 31-33+22.

Xu, J. Y. (2017). 家庭背景与高等教育获得的关系研究 A study of the relationship between family background and access to higher education. *教育论丛/改革与开放 Education Series / Reform and Opening.*

Xu, W. J. (2017). 农村教育机会获得的性别差异研究——基于多年 CGSS 面板数据 A study of gender differences in access to education in rural areas - based on multi-year CGSS panel data. *教育导刊/Education Guide,* 04,25-30.

Xu, W. Q. (2021). 家庭背景、自我教育期望与高等教育获得——基于 CFPS 2010—2016 年追踪数据的实证研究 Family Background, Self－Education Expectation and Higher Education Attainment. *高等理科教育 Higher Science Education,* 02,10-17.

Xu, W. Q. & Fang, F. (2020). 谁获得了高等教育——基于户籍身份和家庭背景的双重视角. Study on the Access to the Higher Education: Dual Perspectives Based on Household Registration Status and Family Background. *Chongqing Higher Education Research,* 8(1), 14-26.

Xu, Y. (2016). 公共政策视角下我国高等教育入学机会公平问题的探析 ANALYSIS OF PUBLIC POLICY OF OUR ACCESS TO HIGHER REDUCATION EQUITY ISSUES (硕士学位论文,南华大学 MA Dissertation of University of South China). https://kns.cnki.net/KCMS/detail/detail.aspx?dbname=CMFD201701&filename=1016317744.nh

Xu, Y., Song, W.X., & Liu, C.H. (2018). Social-spatial accessibility to urban educational resources under the school district system: A case study of public primary schools in Nanjing, China. *Sustainability*, 10(7), 2305.

Xu, Y. R. & Montgomery, C. (2021). Understanding the complexity of Chinese rural parents' roles in their children's access to elite universities. *British Journal of Sociology of Education*, 42:4, 555-570.

Yan, G. F. & Wang, H. Y. (2012). 优质高等教育资源的获得及影响因素分析——从社会分层的视角出发 Analysis of access to quality higher education resources and the factors influencing it - from the perspective of social stratification. *现代大学教育 Modern university education*, 1, 6-11.

Yang, C. C. (2018). 高等教育扩张视角下教育代际流动性的区域比较——基于 2013 的实证研究 Regional Comparison of Intergenerational Mobility in Education from the Perspective of the Expansion of Higher Education: An Empirical Study Based on CGSS2013. *经济统计学 Economic Statistics*, 02, 220-232.

Yang, D. P. (2007). Focusing on Chinese education equity. *Front. Educ. China* 2007, 2 (2), 149-150.

Yang, D.P. (2010). An Empirical Study of Higher Education: Admissions Opportunities in China. *Chinese Education & Society*, 43(6), 59-85.

Yangguanggaokao. (2022). 2022 年三大专项计划帮农村生上好大学 Three special programmes to help rural students attend good universities in 2022 (chsi.com.cn) http://gaokao.chsi.com.cn/gkxx/zt/sdzxjh.shtml

Yang, J., Huang, X., & Liu, X. (2014). An analysis of education inequality in China.

International Journal of Educational Development, 37, 2-10.

Yang, J. H. (2014). 我国高等教育入学机会的区域差异及其变迁 The Regional Gap of Access to China Higher Education and Its Change. *高等教育研究 Higher Education Research,* 12, 27-34.

Yang, T. (2020). 家庭资本、教育获得与社会分层关系研究——基于 CGSS15 年的数据 A study of the relationship between family capital, educational attainment and social stratification - based on 15 years of CGSS data. *才智 wisdom,* 05,175-176.

Yang, X. (2010). Access to higher education for rural-poor students in China. *Educational Research for Policy and Practice*, 9(3), 193-209.

Yang, X. C., Yao, Y. & Zhang, S. (2016). 家庭社会地位对青少年教育期望影响研究——基于 CEPS2014 调查数据的实证分析 A study on the influence of family social status on adolescents' educational expectations - an empirical analysis based on CEPS2014 survey data. *中国青年研究 China Youth Studies*, 7, 67-73.

Yang, Y. K., Wang, X. & Chen, D. (2019). 高考录取批次改革的演变与研究 Research on the Admmissions Batches Reform of National College Entrance Examination. *Journal of Qiqihar University (Phi & Soc Sci)*.

Yao, D. X. (2021). 城市化进程中农村留守儿童隔代教育问题研究 A study on intergenerational education for children left behind in rural areas in the process of urbanisation.*新课程 New Courses*, 50(1).

Yao, H. & Ye, Z. (2018). 家庭背景, 教育质量与学生能力形成——基于 CEPS 的多层线性模型分析. Family Background, Educational Quality and Development of Student Ability: Multilayer Linear Model Analysis Based on CEPS. *当代教育与文化 Contemporary Education and Culture*, 10(4), 70-79.

Yao, X. Q. (2008).高考录取中的地区均衡问题探讨 Area Balance Problem in College Entrance Examination. *陕西教育学院学报 Journal of Shaanxi College of Education,* 03, 20-24.

Ye, H. (2015). Key-point schools and entry into tertiary education in China. *Chinese Sociological Review*, 47(2), 128-153.

Yeung, W. J. J. (2013). Higher education expansion and social stratification in China. *Chinese Sociological Review*, 45(4), 54-80.

You, Z. R. & Hu, Y. Z. (2013). Walking a Policy Tightrope: The Dilemma of Balancing Diversification and Equality in Chinese College Entrance Examination Reform. *Higher Education Policy*, 26, pp309-324.

Young, N. A., & Hannum, E. C. (2018). Childhood inequality in China: Evidence from recent survey data (2012–2014). *The China Quarterly*, 236, 1063-1087.

Zeng, W. T. (2018). 罗尔斯正义原则视角下对高考加分政策的审视 An Examination of the Higher Education Entitlement Policy in the Light of Rawlsian Principles of Justice. *浙江教育科学 Zhejiang Education Science*, 12, 3-7.

Zha, Q. (2011). China’s move to mass higher education in a comparative perspective. *Compare: A Journal of Comparative and International Education*, 41(6), 751-768.

Zhang, C. Z., Huan, Z. J., & Li, H.Z. (2007). An empirical study on education equity in China, 1978–2004. Frontiers of Education in China, 2(4), 536-544.

Zhang, D., Li, X., & Xue, J. (2015). Education inequality between rural and urban areas of the People's Republic of China, migrants' children education, and some implications. *Asian Development Review,* 32(1), 196-224.

Zhang, D. H. & Li, L. 2019. 扩招与高等教育入学机会地区差异的再分析 Re-analysis of Regional Differences between Enrollment Expansion and Higher Education Enrollment Opportunities. *Peking University Education Review,* 17 (1), 143-162.

Zhang, H. (2016). Literature review on educational attainment of migrant children in China. *Open Journal of Social Sciences*, 4(07), 190.

Zhang, H. F. (2017). Opportunity or new poverty trap: Rural-urban education disparity and internal migration in China. *China Economic Review*, 44, 112-124.

Zhang, K. (2017). 我国教育代际传递的城乡差异 Urban-rural differences in the intergenerational transmission of education in China (硕士学位论文,东北财经大学 MA Dissertation of Dongbei University of Finance and Economics). https://kns.cnki.net/KCMS/detail/detail.aspx?dbname=CMFD201802&filename=1018051063.nh

Zhang, L. X. (2009). 基于双重价值视角的我国大学入学公平提升研究 Research on promotion of college entrance equity in China based on dual values of education (博士学位论文，哈尔滨工业大学 PhD Thesis of Harbin Institution of Technology). https://kns.cnki.net/KCMS/detail/detail.aspx?dbname=CDFD0911&filename=2011015630.nh

Zhang, Q., & Liu, Z. (2019). Substitution or complementation: the influence of parents' educational level on sheepskin effects in the individual returns to education in China. *Asia Pacific Education Review*, 20(4), 595-606.

Zhang, W. W. (2014). *Educational Opportunities of Ethnic Groups in China: The Disparities* *in Attainment and Achievement* (PhD Thesis, The Chinese University of Hong Kong)

Zhang, X. L. (2015). 高等教育机会获得的影响因素研究——基于 CGSS2006 数据 A study of the factors influencing access to higher education - based on CGSS 2006 data. *郑州航空工业管理学院学报 Journal of Zhengzhou Institute of Aviation Industry Management*, 02 (33),127-132.

Zhang, X. P. & Zhang, L. (2015). 中国高质量大学入学机会和招生偏好研究——以"211"高校为例 On Enrollment Opportunities for High-quality Universities and Enrollment Preference: Evidence from "Project 211" Universities. *高等教育研究 Higher Education Research,* 07, 28-35.

Zhang, Z. S. & Chen, Q. (2013). 高校扩招与高等教育机会的性别平等化——基于中国综合社会调查(CGSS2008)数据的实证分析 College Expansion and Gender Equalization in Higher Education: An Empirical Study Based on 2008 Chinese General Social Survey. *社会学研究 Sociological research,* 02,173-196+245.

Zhang, Z.S., & Chen, Q. (2014). The expansion of higher education admissions and the gender equalization of higher education opportunity: An empirical study based on Chinese General Social Survey (CGSS2008) data. *The Journal of Chinese Sociology*, 1(1), 1-19.

Zhao, D. C. (2021). 家庭收入对大学入学机会的影响及其性别差异——基于 CFPS 数据的分析 The Impact of Family Income on College Entrance Opportunities and Gender Differences. *科教导刊 Science Education Guide,* 17,189-192.

Zhao, H. X. & Wang, L. M. (2020). 促进还是抑制: 高等教育对社会阶层流动的影响——基于 CGSS 混合截面数据的实证分析 Facilitating or inhibiting: the impact of higher education on social class mobility - an empirical analysis based on CGSS mixed cross-sectional data. *Higher Education Exploration*, 9, 5-11.

Zhao, J.S. & Guo, J.Y. (2002). The Restructuring of China's Higher Education: an experience for market economy and knowledge economy. *Educational Philosophy and Theory*, 34(2), 207-221.

Zhao, L. T. (2018). Stratification in China's Education. *China: An International Journal*, 16(3), 142-162.

Zhe, J. (2016). *Inequality within Chinese higher education with the focus on henan province.* A dissertation submitted to the School of Education Duquesne University in partial fulfillment of the requirements for the degree of Doctor of Education.

Zheng, L. & Sun, X. (2017). 高等教育收费政策对入学机会公平的影响——基于CGSS(2010-2013)数据的研究 The Impact of Higher Education Tuition Policy on Enrollment Opportunities: Evidence from CGSS（2010 － 2013). *山东高等教育 Shandong Higher Education,* 01, 30-37.

Zheng, M. Y. (2021).*我国高等教育城乡差异现状及影响因素分析(硕士学位论文,河南财经政法大学 MA Dissertation, Henan University of Economics and Law).* https://kns.cnki.net/KCMS/detail/detail.aspx?dbname=CMFD202201&filename=1021709508.nh

Zhou, B. H. (2018). Regional distribution of college enrollment in China under a multiple-principal framework. *International Review of Economics and Finance,* 57, 122-133.

Zhou, F. (2018).家庭背景、社会阶层认同对高等教育获得影响研究——基于CGSS2015的实证分析 The Influence of Family Background and Social Class Identification on higher education acqusition analysis based on CGSS2015. *宜宾学院学报 Journal of Yibin College,* 05,117-125.

Zhou, S. J., Li, Q. Y., Cui, L. Z. & Zhang, H. S. (2018). 父母学历与子女教育--基于 CGSS

微观数据的实证考察 Parents' Academic Qualifications and Children's Education Empirical Study: Based on Microscopic Data of CGSS. 教育经济 *Education & Economy*, 3, 46-53.

Zhou, W. H. (2012). 高考分省定额政策的历史演变 Historical Evolution of the Provincial Quota Policy . 教育考试 *Education & Examinations*, 04, 18-22

Zhou, Y. (2020). 学前教育机会与义务教育结果不平等——来自 CEPS 的经验证据 Preschool Education and Compulsory Education Result Inequality: Empirical Evidence from CEPS. 学前教育研究 *Preschool Education Research,* 1, 76-87.

Zhu, J. (2011). Who goes to university?: Inequality in access to higher education in China 1949-2005 (Master's thesis).

Zhu, J., Xu, L. & Wang, H. (2018). 教育代际传递的城乡差异研究——基于中国综合社会调查数据的验证 A study of urban-rural differences in intergenerational transmission of education - A validation based on data from the China General Social Survey. 教育经济 *Education & Economy*, 34(6), 45-55.