Gender Gap in University Academic Performance

A study of the China Family Panel Studies

The gender imbalance issue in Chinese higher education has received increasing attention on various aspects, including participation, attainment, and regional difference. Not enough is known about the academic performance gap between males and females mainly because of the large population and regional diversity in China. This report endeavours to in academic performance with regard to a variety of background and predictive factors. Using data from the China Family Panel Studies, approximately 3,000 university students enrolled between academic years 2010 through 2014 were studied in this report. To decompose gender differences in university rank, three year cohorts (2010, 2012, and 2014) were combined as a cross-sectional data for multiple regression analyses.

Findings show that females are outperforming their male peers for around 8%, after accounting for various socio-economic and academic background information. In fact, the female advantage exists even in those traditionally male-dominated subjects in Science, Technology, Engineering, Agriculture, and Medicine. Student activities may have similar effects on both gender groups but play an important role in widening the gender gap. Other predictive factors regarding attitude, activity, and satisfaction explain some of the widened gender gap in performance, particularly those about self-confidence, academic pressure, in-class concentration, university regulations, and student activities. Moreover, some characters of well-performing students are also identified, such as the majority ethnic group ‘Han’ and comparatively higher cognitive ability, aspirations, self-efficacy, and self-motivation.

This report helps to provide insights into the performance gender gap in Chinese higher education on a national scale. Recommendations for further research are emphasised on longitudinal data and more accurate measures of prior attainments and academic abilities.

1. Introduction

In a country with a large population and regional diversity like China, it is well known that the academic outcome varies substantially by gender, the higher education institution attended, subject, family background, etc. Since 1999, higher education admissions have witnessed an expansion providing opportunities for reshaping the gender structure (Yang, 2006). Chinese scholars found that women were the primary beneficiaries of this expansion, and in 2010, the gender ratio of enrolled higher education students reached around 50% (Zhang and Chen, 2014). However, many existing studies focus on the gender gap in higher education enrolment and completion, while there are relatively few have investigated the question of performance differences between boys and girls. As findings from research in other developed country contexts shown, females are widely and largely outperforming their male peers, and this performance disparity may be linked to a student's demographic characteristics, family backgrounds, and prior academic attainments (Crawford and Greaves, 2015; Bailey and Dynarski, 2011). In addition, some other researchers identified a variety of non-demographic and non-intellectual differences that predict academic performance, such as personalities, learning skills, aspirations, and expectations (Chamorro-Premuzic and Furnham, 2003; Conger and Long, 2010; Richardson et al., 2012). To understand the gender gap in academic performance is to contribute to gender-balanced student populations (Gibbs, 2008), and to predict future changes in the labour market driven by growing female advantage in higher education (Gonger and Long, 2010).

Based on the existing research, the following body of this quantitative data analysis report aims to study the gender disparity in academic performance among students who are currently enrolled in Chinese universities. In investigating the reasons behind the gender gap, it will endeavour to answer the following questions: 1) what factors regarding a student's socio-economic and academic backgrounds affect the academic performance differences between male and female university students? 2) What kind of predictors on a student's personality, attitudes, and satisfactions can explain the gender gap in performance? And 3) what changes in gender disparity are made by being a student representative or a leader of student organisation?

The China Family Panel Studies (CFPS) data is used to provide estimates of gender disparities in academic performances The CFPS offers not only a comparatively better opportunity for better information on a national scale (Xie and Lu, 2015) but more importantly, an insight into contemporary Chinese higher education. The university rank in the subject is the measure of academic outcome, which is expected to be explained by both background characteristics, including socio-economic status and prior attainment, and self-evaluated predictors of attitude and aspiration. Considering the fact that less than 10% of higher education students in the CFPS were followed in later sweeps, this report uses the CFPS data support a cross-sectional study.

This report will firstly start with a review of existing research on the gender gap in academic performance at university. Then, the following section will further explore Chinese university students who were enrolled between academic years 2010-2014 with multiple regression analyses. Interpretations and explanations will be presented regarding the regression results. At the end of the report, recommendations for further research on the gender imbalance issue in Chinese higher education will be provided.

2. Previous research

2.1 Gender gap in higher education

Previous research on the gender difference has been widely and thoroughly explored through investigations on the higher education participation and attainment issues in the UK and the US. When it comes to academic performance in higher education institutions (HEIs), girls are widely outnumbering and outperforming boys in various subjects. This is the phenomenon referred by some scholars as the ‘reverse gender gap’ (Nozaki et al., 2009). Graduation rates and degree levels are commonly used as criteria for assessing performance (DiPrete and Buchmann, 2006; Vincent-Lancrin, 2008), whereas exam marks and rankings are regarded as rather difficult to quantify.

As Crawford and Greaves has suggested (2015), the gender differences in attendance rates in UK higher education were found to be affected by the socio-economic circumstances in which a student was raised. Despite the rapidly rising percentage of higher education participation, white males and those with low socio-economic

backgrounds were still less likely to go to university. In fact, participation rates for girls are constantly and significantly higher than boys by around 8%. Once the prior academic attainments are accounted for, the gender differences are reduced.

Evidence from longitudinal studies carried out on young people in England showed that female students had a higher participation and graduation rate in every socio-economic status quantile than their male counterparts in recent years (Chowdry et al., 2012). This study adopted a school-fixed effects approach with a binary outcome variable and found that the gender disparity in participation is largely explained by not only personal characteristics but also prior attainments. And the chance of attending a high status institution is substantially related to socio-economic background for both boys and girls, rather than prior achievements.

The inequality in university entry has also been studied in the US, and largely explained by an inequality in the secondary education (Bailey and Dynarski, 2011). Bailey and Dynarski (2011) has indicated that gender differences exist in university course completion and graduation rates. For students with low income backgrounds, completion rates decreased to about 10% for boys but rose more than 10% for girls. As a matter of fact, the female advantage in higher education was also largely due to their higher completion rate of secondary education since the 1940s.

As for the persistence, Crawford (2014) indicated that disadvantaged socio-economic backgrounds have had a higher impact on dropout, degree completion and degree class for males than for females. The casual effects of socio-economic status on degree outcomes were investigated by regression models with the help of rich longitudinal data of children from age 11 onwards. Prior attainments at age 11, 16, and 18 were available for marking the academic achievement baseline. After accounting for demographic background and prior attainments, the gender differences become smaller, but strategies to reduced socio-economic inequalities are still desired to support students from disadvantaged families.

In fact, the differences may persist into later life events, such as employment and

earnings, as Kalleberg (2010) discovered from the US Current Population Survey from 1983 through 2008. The gender wage gap declined over the observed period, but most within-occupation inequality was found to be closely related to gender, and most gender inequality persisted across occupations. In addition, evidence from an analysis of March Current Population Survey data 1964-2002 shows that women have not only overtaken men in the completion rate, but actually gained the returns to higher education higher than those of men (DiPrete and Buchmann, 2006).

In addition to findings from previous research where prior attainment and socio-economic status play a key role in explaining the gender gap in academic outcomes, non-demographic factors have also received increasing attention. Jacob (2002) attempted to uncover the reason for the gender difference and found that greater non-cognitive skills among women were able to explain nearly 90% of the gender gap. According to a study on enrolees in Florida and Texas universities, girls generally have higher non-cognitive skills, such as organisation, dependability, and self-discipline, than boys (Conger and Long, 2010). And these differences are potentially widened in later years at university influencing not only the academic performance, but also the likelihood of retention and transfer (Allen et al., 2008).

Chamorro-Prenuzic and Furnham (2003) explored two British university samples and indicated that different personality traits might account for the variance in academic performance. In addition, personality and approach to learning have been confirmed to be related to academic outcomes among undergraduate students in other studies carried out on smaller samples (Duff et al., 2004; McKenzie and Schweitzer, 2001). Moreover, aspirations and expectations are proved to have positive effects on both university enrolment and completion. In nearly all OECD member countries, girls with high educational expectations were more likely to be expected to attend university and to achieve better outcomes than boys (Buchmann and Dalton 2002, McDaniel, 2010). It is therefore necessary to explore a student’s aspirations and expectations in addition to his or her demographic characteristics and prior attainments in order to understand the gender gap in university performance.

2.2 Chinese higher education

In a Chinese context, gender equity and access has been studied in terms of participation, attainment, and regional differences (Rong and Shi, 2001; Jacob, 2006). However, most of these studies focused on the primary and secondary education stages rather than higher education. It is largely because of the long-existing culture in China where sons are the preferred child that this has remained a key social problem for girls' educational opportunities (Honig and Hershatter, 1988; Yu et al., 1990). However, now, due in part to the one child family policy and compulsory education, girls are outpacing boys in education achievements (Wang, 2007). Thus, to what extent the gender inequality has been reduced for the younger generations becomes the main question that researchers of Chinese education are interested in.

However, research on higher education may potentially provide further evident to study the gender inequality and its relationship with academic performance. The expansion of higher education admissions enables the gender difference in enrolment to decrease to less than 1% that can be disregarded (Liu, 2006). Zhang and Chen (2014) found from a study on the China General Social Survey that it was the females from rural areas or with middle education level parents received increasingly more tertiary education opportunities, thus contributed to the gender equality. Regarding the employability after higher education, Wang (2007) has studied around 7,000 first year students at the same university, and she found that girls are more likely to achieve higher outcomes once control for ethnicity, socio-economic background, and subject. This study has also found certain personality traits, such as confidence, preciseness, adaptability, etc. are significantly related to the academic outcomes.

Interestingly, evidence from research regarding the gender differences in employment presents a completely opposite result. According to Tan (2012), male students generally receive better employment opportunities with a higher wage of approximately 9%-20%. It raises concerns on the reason why males remain overrepresented in several high-income occupations and industries in spite of their lower academic achievement. On the contrary, Xiong (2007) suggests that the gender gap in employability and annual earnings might result in the diversity of the family social network, parental education,

parental occupation, and personal ability and attitude. Further research on the gender gap in academic performance and attainments may be potentially helpful in further explaining the differences in characteristics and quality between male and female students and its effects on the future employability problem.

This report uses the China Family Panel Studies data, and extracts relevant observations for a quantitative analysis on Chinese university students. Previous research using the CFPS data has focused on several topics in different contexts: such as the wellbeing of migrant children; the relationship between cognitive development and family resources; and an investigation into the labour force (Wu, 2014; Jordan et al 2014; Chen and Qin, 2014). Information captured by the CFPS on higher education, according to the knowledge of this report, has not yet been studied. Therefore, this report expects to make a use of the national longitudinal data to contribute to a cross-sectional study on Chinese higher education.

3. Data

3.1 CFPS data

This report bores into the China Family Panel Studies (CFPS) data for anonymised Chinese domiciled households and individuals starting from 2010. The CFPS is the first national representative of longitudinal research to biannually collect the most complete, highest-quality survey data on contemporary China (Xie, 2012). It covers 95% of the total Chinese population on numerous topics including economic activities, education outcomes, family dynamics and relationships, migration, and health (Xie and Lu, 2015). Specifically, 15,000 families and almost 30,000 individuals within these families were interviewed, with a response rate of 79%. As Xie and Lu (2015) stress, an individual is the most important study unit in CFPS. All members over the age of 9 in a sampled household were regarded eligible for the research.

As a government-funded and university-led study, the CFPS is able to adopt complex sampling designs and control data representativeness. The Institute of Social Science Survey (ISSS) of Peking University was responsible for the study. Researchers specified sampling weights to account for the disproportional selection due to the regional

diversity across Chinese administrative divisions (i.e. provinces). They also paid careful attentions to adjust and control for the attrition problem by using other interview methods or the proxy response from other family members (Xie and Hu, 2015).

As the factors influencing students' academic outcomes in higher education is the topic for this report, only the adult data in each sweep will be used. The CFPS defines 'adult' as aged 16 and over, rather than the legal adult age of 18 years, for the purpose of data collection on compulsory and post-compulsory education (Xie and Hu, 2015). In addition, Chinese students normally take the National Higher Education Entrance Examination, also widely known as the *gaokao*, between the ages of 18 to 20 (Liu, 2013). The adult data, thus, is capable of reflecting all eligible higher education students. However, unfortunately, the information of *gaokao* is not included in the CFPS survey, because actual exam subjects, test formats, and admission requirements vary across the regions (Walker et al., 2011). Marking students' previous academic attainments by *gaokao* scores would be inaccurate and untrustworthy. The CFPS-adult data thus collects average scores and ranking in HEIs.

Unfortunately, after carefully exploring the data on three year cohorts, this paper found the HEI student sample is not longitudinal. When the three year cohorts were combined into one dataset, there were more than 3,500 observations available, but less than 100 of them were followed in the next sweep. A possible reason for the non-response could be off-campus final-year projects, internships, or other personal reasons that went beyond the scope of the CFPS. Besides, since the CFPS is aimed at the entire Chinese population on diverse topics, rather than the higher education sector, it is understandable that the data on current HEI students would be actually cross-sectional. Therefore, this paper will use the CFPS data to support a cross-sectional data analysis.

The CFPS collects data on students who were enrolled in higher education institutions in the academic years 2010-2014. It contains information on their academic outcomes, student activities, self-evaluations on academic performance, and satisfaction levels. Table 1 presents the individual level variables used in this report.

Table 1 Individual level variables in the CFPS database

Variables	Variable type	Detail
Basic information	Dichotomous	Gender, *hukou* status (i.e. rural/urban).
	Nominal	Ethnicity.
	Continuous	Age, cognitive ability test scores.
Academic performance	Nominal	HEI type, grade, subject.
	Continuous	Average score, ranking, total student number.
Family background	Dichotomous	Parental education (HE degree)
Activity	Dichotomous	Student leader/representative,
		Student organisations/clubs, organisation leader.
	Continuous	Leadership length, organisation number
Attitude	Ordinal (5 categories)	Self-evaluation on talent, academic performance, academic pressure, excellence, and suitability as student representatives.
		Attitudes on efforts, in-class concentration, assignment, HEI regulations, and task management.
Satisfaction	Ordinal (5 categories)	Satisfaction with HEIs and tutors.
Education expenditure (last year)	Continuous	Total education expenditures.
		Education expenditure paid by family.

In the data collection process, researchers used CAPI technology, which was provided by the Survey Research Centre (SRC) at the University of Michigan, to enable diversified questionnaire designs. To assist interviewers in completing surveys and managing data, the CAPI system provides a quicker and easier way for communication between interviewers and headquarters on problems and adjustments.

3.2 Measures

3.2.1 Population

The sample sizes included in the CFPS database and those available to this paper are given in the table below, which also shows the observations and proportions of both college and university students. Approximately 500-2,500 students are currently studying for a higher education degree in each year's cohort. For each cohort, the university is the group with the larger members. Initially, this report planned on

exploring all the sampled higher education students for a better understanding of HE students. However, because the information regarding college students only includes observations on their college types, subjects, and starting years, their academic performance becomes difficult to quantify. This paper, therefore, will focus on university students only.

Table 2 Sampled individuals by HEI type and year cohort

Year	*college*	*university*	*HE*
2010	246	295	541
	45.47%	54.53%	100%
2012	764	878	1642
	46.53%	53.47%	100%
2014	1044	1612	2656
	39.31%	60.69%	100%
Total	2054	2785	4839
	42.45%	57.55%	100%

As stressed in the previous section, information on currently enrolled HE students in the CFPS data was hardly quantified. This report, therefore, combines the three year cohorts for a cross-sectional data analysis. There were 4,839 individuals interviewed, and 2,785 of these were university students. These current university students were sampled from different types of universities or (as commonly described in the Chinese higher education context) tiers of universities. The structure and proportion of university types will be further explained in the next section.

3.2.2 Academic performance

The CFPS uses three variables to mark academic progress at the higher education levels. These are: the average final exam score; university rank among peers in the same subject; and the total student number for the previous term. Considering the average final exam score may not be comparable between universities and subjects due to the diversity of test formats and grading standards (Xie and Hu, 2015), university rank is thus used to measure the academic performance. This report also acknowledges that the university rank alone may not be as fully accurate as the key variable of academic performance. The reason for this is because the same rank may not represent the same

performance level if the student numbers for different observations are significantly different. Hence, a percentile rank is calculated to ensure that the same rank means a similar academic performance regardless of the total student number in the subject. Table 3 gives a summary of the percentile rank.

Table 3 Summary of percentile rank

Variable	Obs	Mean	Std. Dev.	Min	Max
Rank	2783	29.58508	23.03418	1	100

It is noteworthy that the interpretation of ranking is not necessarily similar to other measures of academic performance. In fact, a positive coefficient for a predictor is not linked with better performance. For example, if 'female' is the baseline group of gender, a positive coefficient of 3 would suggest that male student's rank three units lower than their female peers. All of the coefficients will be interpreted accordingly in the next chapter.

In addition, despite the fact that the *gaokao* score is not contained in the CFPS data, which makes the previous academic achievement very difficult to quantify, Chinese university types could potentially be of an alternative choice. This study follows guidance from the Chinese Ministry of Education (MOE) in defining the hierarchy of HEI tiers (MOE, 2013). Universities are normally categorised into 3 or 4 tier, and these tiers are largely used as a reference for allocating funding and recognising the quality of both the university and its graduates (Li, 2012). Tier-1 universities normally refer to those institutions directly administered by the MOE or other ministries; tier-2 universities are administered by the local government at provincial or municipal level; and most tier-3 universities are privately funded (MOE). Table 4 shows the proportion of each university tiers in the three year cohorts. Tier-2 and 3 consist of over half of universities, while tier-1 universities provide student places for the top 15% of students.

Table 4 Summary of university types

University type		Freq.	Percent
Tier 1	National key	263	9.43
	Standard	595	21.36
Tier 2		996	35.75
Tier 3		408	14.65
Other		523	18.81
Total		2785	100

Among tier-1 universities, top ranked institutions are classified into national higher education projects which are usually called 'national key universities'. The 'C9 League', also known as the 'Chinese Ivy League', together with 30 institutions in the 'Project 985', represent 39 world-class universities in China (Zhang et al., 2013). In 2007, they received an average of US$92 million for scientific research per university, and the 'Reign Supreme' Beijing and Tsinghua University's received around $132 million each (Clark, 2010). Before the announcement of the 'Project 985' in 1998, 166 HE institutions were designated as 'Project 211' universities who are nationally recognised and accounted for 96% of national key laboratories and 70% of scientific research funding (Wang and Zheng, 2013:374). With funding and prestige tied to national key universities, HE students sought out places at these institutions for a better education and future employment opportunities. However, due to regional diversity and the large population, each university may have had different admissions requirements (i.e. *gaokao* scores) for different provinces (Yu et al., 2012:30). Students enrolled at universities in the same tier tended to have similar academic achievements prior to their higher education (Yu et al., 2012:82). This report will endeavour to use university tiers to partly account for previous academic abilities, along with other factors such as the cognitive ability tests designed by the CFPS.

3.2.3 Control variables

To better account for university students' backgrounds, this report includes a variety of measures to examine socio-economic circumstances. As briefly presented in table 1, information about currently enrolled HE students are quantified with either categorical or continuous variables in the CFPS data. Categorical background variables used in this report are shown in table 5, and continuous variables are in table 6. There are three sets

of background characteristics studied:

- Personal: age, ethnic group, *hukou* status, cognitive ability test results;
- Family: parental education, education expenditure paid by family (last year);
- University: subject, university type.

Firstly, for variables on personal information such as ethnicity, this paper re-groups observations to formulate categories of the same variable easier to compare. Taking into account that the diversity between ethnic groups is reduced due to the limited university student sample, 'Han' becomes the predominant ethnic group, in comparison to other minority groups. So the ethnicity variable descriptor is changed into being a member of the Han ethnic group.

The *hukou* status is a record for a household registration system, required by law in China, to collect a resident's information based on name, gender, date of birth, place of birth, parents, and spouse (Chan and Buckingham, 2008). On the individual household's registration booklet, residents are broadly categorised as 'rural' or 'urban'. The CFPS data also makes use of the *hukou* status to collect information on individual's dwelling place and to account for the rural-urban difference (Xie and Hu, 2014). So the *hukou* status variable is also included in this report for referencing a student's family socio-economic background.

In the CFPS-adult questionnaire, all of the eligible individuals were invited to take the cognitive ability tests: which are designed to collect information on both verbal and mathematical reasoning abilities. There are eight groups of verbal questions worth 34 credits, and four groups of mathematical questions in the tests worth 24 credits. The total cognitive ability test score is calculated in percentages for each individual.

Secondly, for controls on the family background, both parental education and education expenditure are recorded in this report. According to existing research, parental education, especially parental higher education experience, has positive effects on children's high education opportunities and outcomes (Choy, 2001). Hence, the highest

degree achieved by a father or mother is used as the grouping criterion. Observations on parental education are now re-grouped as either ‘one or two parents hold a HE degree’ or ‘neither of the parents holds a HE degree’. Furthermore, family-paid education expenditure is transformed for a more complete exploration on percentile changes similar to what this paper organised for the university rank.

Thirdly, in terms of the characteristics in a higher education background, individuals in smaller universities studying less popular courses might encounter a problem of under-representation. Subjects, in addition to university types, are recorded into three main criteria to account for the variance of university rank. The three subject groups are LEM, STEAM, or others. LEM denotes subjects in Law, Economics, and Management; STEAM includes subjects in Science, Technology, Engineering, Agriculture, and Medicine; and others are mostly courses in Humanities and Arts.

Table 5 Summary of categorical control variables

Categorical variables		**Freq.**	**Percent**
Gender	Female	1,007	47.66
	Male	1,106	52.34
	Total	2,113	100
Ethnicity	Han	2,522	82
	other	535	17.51
	Total	3,057	100
Hukou status	Rural	1,091	65.37
	Urban	578	34.63
	Total	1,669	100
Subject	LEM	533	31.46
	STEAM	772	45.57
	Other	389	22.96
	Total	1,694	100
Parental education	yes	155	7.28
	neither	974	92.72
	total	1129	100

Table 6 Summary of continuous control variables

Continuous Variable	Obs	Mean	Std. Dev.	Min	Max
Age	3061	21.71121	3.239397	17	54
Cognitive ability tests	1627	79.59435	17.53919	0	100
Family-paid education expenditure	1607	87.89114	18.71299	0	100

3.2.4 Predictors

This report uses a set of measures to account for differences in student activities, self-evaluations, attitudes, and satisfaction levels. The measures included are set out in detail in the tables below. Questions regarding potential predictors of university rank are answered on a five-point Likert scale, ranging from 'strongly disagree' to 'strongly agree', or simply rated from 1 to 5. Lists of variables relating to each type of predictors are:

- Student activities: student leader/representative, student organisation, and student organisation leader;
- General self-evaluation on the previous academic year regarding: academic performance, level of academic pressure, excellence, and suitability as a student representative;
- Self-evaluation of attitudes on questions like:
    1) I study hard,
    2) I concentrate on studying while in class,
    3) I check my assignment several times before submission,
    4) I abide by the university rules and regulations,
    5) I like to put things in order,
    6) I put my assignment first;
- Satisfaction with HEI and tutor.

All of these measures will be sequentially added into regression analyses in the next section.

Table 7 Summary of predictive variables

Predictors		Freq.	Percent	Cum.
Performance	1	25	1.06	1.32
	2	99	4.21	5.52
	3	1,156	49.13	54.65
	4	918	39.01	93.67
	5	149	6.33	100
	Total	2,347	100	
Pressure	1	197	8.37	8.63
	2	481	20.44	29.07
	3	1,001	42.54	71.61
	4	573	24.35	95.96
	5	95	4.04	100
	Total	2,347	100	
Excellence	1	25	1.06	1.32
	2	163	6.93	8.24
	3	1,313	55.8	64.05
	4	763	32.43	96.47
	5	83	3.53	100
	Total	2,347	100	
Suitability	Refuse	4	0.17	0.42
	1	113	4.8	5.23
	2	386	16.4	21.63
	3	966	41.05	62.69
	4	703	29.88	92.56
	5	175	7.44	100
	Total	2,347	100	
Efforts	Strongly Disagree	144	6.12	6.37
	Disagree	906	38.5	44.88
	Neither	865	36.76	81.64
	Agree	400	17	98.64
	Strongly agree	26	1.1	99.75
	Not applicable	6	0.25	100
	Total	2,347	100	
Concentration	Strongly Disagree	181	7.69	7.95
	Disagree	1,278	54.31	62.26
	Neither	547	23.25	85.51
	Agree	300	12.75	98.26
	Strongly agree	24	1.02	99.28
	Not applicable	17	0.72	100
	Total	2,347	100	

Table 8 Summary of predictive variables (continued)

Predictors		Freq.	Percent	Cum.
Checking	Strongly Disagree	139	5.91	6.16
	Disagree	939	39.91	46.07
	Neither	592	25.16	71.23
	Agree	591	25.12	96.35
	Strongly agree	36	1.53	97.88
	Not applicable	50	2.12	100
	Total	2,347	100	
Regulations	Strongly Disagree	685	29.11	29.37
	Disagree	1,229	52.23	81.6
	Neither	126	5.35	86.95
	Agree	229	9.73	96.69
	Strongly agree	65	2.76	99.45
	Not applicable	13	0.55	100
	Total	2,347	100	
Order	Strongly Disagree	398	16.91	17.25
	Disagree	1,250	53.12	70.38
	Neither	327	13.9	84.28
	Agree	307	13.05	97.32
	Strongly agree	39	1.66	98.98
	Not applicable	24	1.02	100
	Total	2,347	100	
Assignment	Strongly Disagree	152	6.46	6.71
	Disagree	915	38.89	45.6
	Neither	743	31.58	77.18
	Agree	436	18.53	95.71
	Strongly agree	41	1.74	97.45
	Not applicable	60	2.55	100
	Total	2,353	100	
Satisfaction: HEI	1	30	1.27	1.7
	2	167	7.1	8.8
	3	714	30.34	39.14
	4	1,058	44.96	84.11
	5	374	15.89	100
	Total	2,347	100	
Satisfaction: tutor	1	91	3.87	5.01
	2	215	9.14	14.15
	3	533	22.65	36.8
	4	882	37.48	74.29
	5	493	20.95	95.24
	Total	2,235	100	

Table 9 Summary of student activity predictors

Predictors		Freq.	Percent
Student representative	No	1,424	63.23
	Yes	828	36.77
	Total	2,252	100
Student organisation	No	1,183	52.72
	Yes	1,061	47.28
	Total	2,244	100
Student organisation leadership	No	692	65.22
	Yes	369	34.78
	Total	1,061	100

Table 10 Summary of the talent variable

Variable	Obs	Mean	Std. Dev.	Min	Max
talent%	2345	44.20256	21.55267	0	100

4. Methodology

To understand to what extent a university student's academic performance is affected by his or her student activities, attitudes, and satisfaction levels, this report uses regression models to undertake the analysis and explore the effects estimated at the mean of all characteristics. These effects will be interpreted as the percentage point change in university rank of the previous term. Moreover, several hypotheses on the influencing factors of the academic performance will be tested. They are:

H_{0a}: there is no statistically significant difference of percentile rank between male and female students

H_{0b}: there is no statistically significant difference between male and female student on percentile rank in the same subject and HEI.

H_{0c}: student activities, attitudes, and satisfactions are not significantly related to the percentile rank of each gender group.

H_{0d}: being a student leader or a student representative is not significantly associated with percentile rank for both gender groups.

The first hypothesis involves only the variable of interest, university rank, and gender. It

intends to study the baseline of the relationship between gender and the outcome variable. The regression equation is:

$$Y_i = \beta_0 + \beta_1 Gender_i + \beta_2 YCohort_i + \varepsilon_i$$

where Y_i denotes the percentile rank for each individual, $Gender_i$ is the dummy variable denoting the gender group, $YCohort_i$ is the year cohort, and ε_i is the error term. Coefficients are represented by βs.

Then, to account for the background characteristics, this report adopts a multiple regression modelling approach to successively add sets of variables to the baseline model. At this stage, the sequence of adding background control variables will be based on the three sets, namely: personal, family, and university information. Aimed at the second hypothesis, the regression model with all the control variables included will be:

$$Y_i = \beta_0 + \beta_1 Gender_i + \beta_2 Age_i + \beta_3 Ethn_i + \beta_4 Hukou_i + \beta_5 PEdu_i + \beta_6 FExp_i + \beta_7 Sub_i + \beta_8 Univ_i + \beta_9 CYear_i + \varepsilon_i$$

For the third hypothesis, this report will add a detailed set of predictors concerning the unexplained variance of gender gap in academic performance on the basis of marking the background characteristics of all of the sampled university students. All sets of predictors, including activities, attitudes, and satisfaction, will be separately analysed by the regression model

$$Y_i = \beta_0 + \beta_1 Gender_i + \beta_2 X_i + \beta_3 Z_i + \varepsilon_i$$

where X_i consists of predictors on each individual, and Z_i is all the control variables. Apart from control variables studied at the previous stage, self-evaluated talent and cognitive test results will also be included to account for intelligence differences.

Finally, this report adds an interaction term to the regression model. This is to account for reasons where student activities, especially leadership activities, might have different effects on different gender groups. Now the model becomes:

$$Y_i = \beta_0 + \beta_1 Gender_i + \beta_2 Rep_i + \beta_3 Gender_i Rep_i + \beta_4 Z_i + \varepsilon_i$$

where Rep_i is the student representative variable, and $Gender_i Rep_i$ denotes the interaction between gender and leadership. The remaining academic performance differences between gender groups, after adding all of these measures, may be due to differences in other aspects that are not included in this paper. This regression model will be used to test the last hypothesis.

5. Results and discussion

5.1 Gender difference

Table 11 presents the result that gender significantly ($p = 0$) accounts for 4% of the variance in the percentile rank through observations on 1,359 individuals. As briefly introduced in Section 3, positive coefficients for explanatory variables actually suggest a negative effect on the university rank. This is because, for example, a 5% increase in rank from the 10th position would imply that a student's position in the subject has dropped to the 15th spot, and he or she may not be a top10% student anymore.

Table 11 Regression of the relationship between gender and rank

Rank	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
Gender	8.676	1.311	6.62	0	6.104	11.249
Cohort	6.479	0.356	18.18	0	5.780	7.178
_cons	25.212	0.932	27.03	0	23.382	27.041

N=1395

This paper regards the Cohort variable as a continuous variable rather than three dummies, i.e. it is coded as $2010 = 0, 2012 = 1, \text{and } 2014 = 2$. The positive coefficient 6.48 then means that the average rank has decreased each year by around 6.48 percentile points. By accounting for this inter-cohort difference, the academic performance gap between male and female students can be observed accordingly.

Since 'female' is coded in the CFPS data as the baseline category, the constant term, then, means an average rank for female students at the 25th. As mentioned in Section 3, the positive coefficient for the gender variable suggests that being male results in an

increase in the university rank. So on average, boys actually rank 8.68 percentile points lower than their female counterparts, which is in the 33rd or 34th percentile when controlling for a Cohort. Thus, the H_{0a} null hypothesis concerning no statistically significant difference of percentile rank between gender groups can be rejected. It is confirmed that there is an academic performance gap in university rank between boys and girls.

5.2 Socio-economic backgrounds

In addition to the regression on the gender difference of academic performance, other variables regarding a student's socio-economic backgrounds are then included as control variables. When information on background characteristics is added as predictors, the amount of variance accounted for in percentile rank increases to 6.6%. Among the five added predictors, only ethnicity is significant ($p = 0$). This suggests a large performance gap between the majority 'Han', and other minorities, by around 6.83 percentile points. But this difference should not be interpreted as the majority ethnic group is substantially outperforming all others because the university type and subject variables are not accounted for yet. Other variables, including age, *hukou* status, parental education, and family-paid education expenditure, are not statistically significant, thus are not associated with the percentile rank.

Table 12 Regression after adding demographic controls

Rank	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
Gender	8.829	1.295	6.81	0	6.287	11.372
Cohort	6.933	0.379	18.29	0	6.189	7.676
Age	-0.266	0.190	-1.4	0.162	-0.639	0.106
Ethnicity	-6.827	1.751	-3.9	0	-10.264	-3.390
Hukou status	-0.573	0.688	-0.83	0.405	-1.925	0.777
Parental education	0.356	0.653	0.55	0.586	-0.930	1.643
Family support	0.018	0.044	0.41	0.68	-0.069	0.106
_cons	49.648	7.779	6.38	0	34.383	64.913

N=1035

However, the gender difference only changes slightly by 0.15%, which indicates that the academic performance gap between gender groups is still the same even taking into account the background characteristics. A possible explanation for this, as the CFPS

emphasised on oversampling and weighting, is that male and female students presumably have the same average background characteristics.

5.3 Subjects, HEIs, and gender difference

As consideration is given to the idea that the gender gap in different subjects may affect a student's academic performance (Gu and Hua, 2010), variables regarding subject and university type are introduced into the regression as new controls. This report observes the sex ratio in each subject category and the results are shown in table 13. STEAM subjects are still male dominated at Chinese universities as previous research found (Gu and Hua, 2010; Zhang and Chen, 2014), and only around 36% of students in STEAM subjects are female. LEM are normally gender-balanced subjects, in which girls account for 46% of the total number of students. The Humanities and Arts subjects are made up of 62% of females, which make them mostly female dominated subjects. Thus the subject variable is coded as: male-dominated STEAM subjects as 1; normally gender-balanced LEM subjects as 2; and other subjects as 3.

Table 13 Gender proportions by subject

Gender	STEAM	LEM	Other	Total
Female	119	219	151	506
	35.94%	46.25%	61.63%	48.28%
Male	211	254	94	542
	64.06%	53.75%	38.37%	51.72%
Total	330	473	245	1,048
	100%	100%	100%	100%

By adding control variables on subject and university, the regression model now accounts for 7.73% of the variation of the university rank. Table 14 shows that despite both university background variables not being significant, the gender gap in university rank becomes narrower: by 1.54 percentile points. The finding indicates that though the preference of subject may differ between male and female students, females are still more likely to rank higher than males. Interestingly, female students rank 7.29% higher even in male-dominated subjects, i.e. those in science, technology, engineering, agriculture, and mathematics. This provides evidence to research in gender differences

where girls generally rank higher than boys, and in those traditionally male-orientated subjects, girls are still outperforming their male classmates.

Table 14 Regression after adding academic background controls

Rank	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
Gender	7.294	1.562	4.67	0	4.224	10.363
Cohort	7.374	0.440	16.76	0	6.510	8.238
Age	-0.345	0.212	-1.62	0.105	-0.762	0.072
Ethnicity	-7.679	2.188	-3.51	0	-11.978	-3.381
Hukou status	-0.363	0.823	-0.44	0.659	-1.979	1.253
Family support	0.045	0.046	0.99	0.323	-0.045	0.136
Parental education	0.106	0.716	0.15	0.882	-1.304	1.516
Subject	-1.455	1.051	-1.39	0.167	-3.520	0.608
University type	-0.087	0.092	-0.95	0.342	-0.268	0.093
_cons	42.223	6.144	6.87	0	30.159	54.286

N=828

5.4 Predictors and gender gap

As has been widely studied by previous research, a student's academic performance is typically associated with his or her intelligence (Chamorro-Prenuzic and Furnham, 2003). But since data on prior academic attainment, i.e. *gaokao* score, is not collected by the CFPS, it is rather difficult to quantify the academic baseline for each individual. Fortunately, questions for testing cognitive abilities were included in the CFPS questionnaires, which enabled a rough understanding of a student's verbal and mathematical reasoning abilities. Also, self-reported academic abilities (coded as the talent variable) are also available in the CFPS data as a supplementary measure to explain university rank. This report, thus, expects a change of percentile rank when measures of academic ability are controlled.

Table 15 Regression after adding cognitive ability controls

Rank	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
Gender	8.756	1.345	6.51	0	6.116	11.396
Cohort	7.151	0.395	18.07	0	6.374	7.927
Age	-0.349	0.199	-1.75	0.08	-0.741	0.041
Ethnicity	-5.653	1.901	-2.97	0.003	-9.385	-1.921
Hukou status	-0.218	0.716	-0.31	0.76	-1.624	1.186
Family support	0.069	0.039	1.77	0.078	-0.007	0.146
Parental education	-0.048	0.664	0.39	0.697	-1.049	1.567
Subject	-1.035	0.829	-1.25	0.212	-2.662	0.592
University type	-0.057	0.071	-0.80	0.424	-0.198	0.083
Cognitive tests	-0.134	0.051	-2.68	0.008	-0.233	-0.035
Talent (%)	0.019	0.030	0.65	0.518	-0.040	0.080
_cons	49.599	6.790	7.3	0	36.275	62.923

N=1057

The table above indeed presents a positive coefficient for the gender variable, which implies that female students are still better-performing regardless of academic ability. Since the cognitive test is significantly related to university rank ($p = 0.01$), it seems cognitive abilities have a positive effect of approximately 0.13% difference in university rank. This is a rather small effect compared to those made by other control variables, and seems quite different from findings from existing research (Furnham, Monsen and Ahmetoglu, 2009; Rohde and Thompson, 2007).

It is possible that intelligence might not be the best predictor of academic performance for university students. Or moreover, because the CFPS cognitive ability tests were designed to measure the whole population, and only about 10% of them have attended or studies in HEIs, the difficulty levels of the tests could be relatively low for most university students. If take the university type into account, tiers of university might have captured most of the information on academic ability through admissions requirements, thus left less to be further explained for the cognitive tests. However, as the cognitive test variable is highly significant, thus related to the university rank, there are very likely that other factors are influencing a student's academic performance at university. Understandably, as reviewed in Section 2, this paper accounts for attitudes and student activities at university, as these predictors represent various personalities and learning styles.

5.4.1 Self-evaluation

Regarding the general self-evaluation of the previous academic year, predictors consisted of self-ratings on academic performance, level of academic pressure, excellence, and suitability as a student representative. This report recoded the rating scale to two categories for a better understanding of self-confidence and self-efficacy of sampled university students. The top two categories, 4 and 5, are combined into a new 'positive' category, and the others are recoded as 'neutral or negative'.

Table 16 Regression of the relationship between self-evaluation and rank

Rank	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
Gender	7.079	1.556	4.55	0	4.022	10.137
Cohort	7.346	0.437	16.81	0	6.488	8.204
Age	-0.383	0.214	-1.79	0.074	-0.805	0.037
Ethnicity	-5.389	2.339	-2.3	0.022	-9.983	-0.795
Hukou status	-0.378	0.832	-0.45	0.649	-2.014	1.256
Family support	-0.038	0.702	-0.06	0.956	-1.422	1.344
Parental education	0.033	0.045	0.73	0.468	-0.056	0.123
Subject	-1.259	1.055	-1.19	0.233	-3.331	0.812
University type	-0.115	0.092	-1.25	0.21	-0.296	0.065
Cognitive tests	-0.127	0.062	-2.04	0.041	-0.249	-0.004
Talent (%)	-0.010	0.035	-0.3	0.761	-0.080	0.059
SE: Perform	2.626	3.518	0.75	0.456	-4.282	9.536
SE: Pressure	4.807	1.725	2.79	0.005	1.419	8.195
SE: Excellence	4.875	2.932	1.66	0.097	-0.883	10.633
SE: Suitability	-2.994	1.969	-1.52	0.129	-6.861	0.873
_cons	-14.737	88.001	-16.75	0	-16.465	-13.009

N=1033

As shown in table 16, the variance of percentile rank is now accounting for 11.5%. There are two predictors significantly associated with rank: self-evaluation on academic pressure ($p = 0.01$); and excellence as a university student ($p = 0.1$). When other self-evaluation predictors are controlled, studying under high academic pressure can have a negative effect on the later performance by up to 4.81 percentile points; while surprisingly, self-regarded excellence results in a lower rank of 4.88%. The results confirm findings from Wang's (2007) research that academic pressure might be overwhelming for some students and so negatively influence their academic achievements. On the other hand, believing in his or her excellence as a university

student might imply over-confidence in one's academic abilities. It is possible that those students with less self-confidence could have higher aspirations and self-efficacy, which could lead to a better academic result.

The gender gap is changed to 7.1, which implies a decrease of differences between boys and girls. It seems that if students in both gender groups have similar feelings about their academic performance, the difference in university rank might be reduced. In spite of the fact that female students still perform far better than males, similar self-evaluation on academic performance may help male students to catch up to the females. Self-efficacy probably has a positive effect on male students (Tan, 2012).

5.4.2 Attitudes

In terms of attitudes, measures provided by the CFPS data cover several aspects of efforts: in-class concentration; assignments; and university regulations. The regression results shown in table 17 explain 10.77% variance of the percentile rank. In-class concentration and attitudes towards university regulations are significant related to the student's percentile rank ($p = 0.02$) when controlling for other variables. Attitudes on university regulations are positively associated with the rank of 0.44%, but concentration during classes has a weak negative effect on rank of 0.3%. These results may partly explain why concentration during class is not necessary for an excellent academic outcome. Comparatively lower academic abilities or previous attainments might have forced some students to concentrate and keep up with the class. On the contrary, students with a negative attitude towards in-class concentration could be more self-motivated than the others. As for the incremental effects of attitudes towards university rules, this is probably due to the implications behind it. Rule abiding students are more likely to be hard working (Wang, 2007), and would therefore achieve better outcomes.

Table 17 Regression of the relationship between attitudes and rank

Rank	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
Gender	8.622	1.345	6.41	0	5.981	11.263
Cohort	6.984	0.412	16.94	0	6.175	7.793
Age	-0.375	0.203	-1.85	0.065	-0.773	0.023
Ethnicity	-5.198	1.912	-2.72	0.007	-8.950	-1.445
Hukou status	-0.366	0.716	-0.51	0.609	-1.771	1.039
Family support	0.067	0.039	1.71	0.088	-0.010	0.145
Parental education	-0.075	0.668	0.11	0.91	-1.242	1.392
Subject	-1.503	1.061	-1.42	0.157	-3.587	0.581
University type	-0.099	0.092	-1.08	0.279	-0.281	0.081
Cognitive tests	-0.119	0.050	-2.37	0.018	-0.218	-0.020
Talent (%)	0.018	0.030	0.6	0.551	-0.041	0.078
A: Effort	0.082	0.165	0.5	0.62	-0.242	0.407
A: Concentrate	0.298	0.126	2.37	0.018	0.051	0.545
A: Checking	0.070	0.094	0.74	0.458	-0.115	0.255
A: Regulations	-0.439	0.184	-2.38	0.018	-0.801	-0.076
A: Order	-0.030	0.115	-0.26	0.793	-0.256	0.196
A: Assignment	0.054	0.088	0.62	0.537	-0.118	0.227
_cons	55.582	9.113	6.1	0	37.686	73.477

N=1033

In addition, the coefficient for the gender variable is decreased by 0.13% compared to the result in table 15. It is likely that once attitudes are accounted for, the academic performance gap between boys and girls is reduced. This partly explains the H_{0c} hypothesis where attitudes towards in-class concentration and university regulations have a statistically significant relationship with the university rank.

5.4.3 Student activities

Student activities have been studied previously to determine if they are an influencing factor in academic attainment. Xiong (2007) found that student activities have a negative relationship with academic outcome, because students spend too much time on activities which led to tight schedules for assignments and even non-attendance of classes. However, Wang (2007) noted that those who participated in student organisations were more outgoing and energetic, and that being a leader of an organisation has positive effects on academic performance.

Table 18 Regression of the relationship between student activity and rank

Rank	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
Gender	9.095	1.346	6.76	0	6.454	11.737
Cohort	6.495	0.402	16.13	0	5.705	7.285
Age	-0.821	0.232	-3.53	0	-1.278	-0.364
Ethnicity	-2.802	1.953	-1.43	0.152	-6.637	1.031
Hukou status	-1.332	0.727	-1.83	0.067	-2.760	0.096
Family support	0.048	0.039	1.24	0.214	-0.0282	0.125
Parental education	-0.001	0.737	0.17	0.998	-1.454	1.450
Subject	-1.737	1.070	-1.62	0.105	-3.840	0.365
University type	-0.068	0.105	-0.65	0.517	-0.274	0.138
Cognitive tests	-0.127	0.050	-2.54	0.011	-0.225	-0.028
Talent (%)	0.023	0.030	0.76	0.446	-0.036	0.083
SA: representative	-2.534	1.376	-1.84	0.066	-5.235	0.167
SA: orgnisation	-4.490	1.387	-3.24	0.001	-7.212	-1.767
_cons	64.203	9.352	6.86	0	45.838	82.569

N=1001

Therefore, for the sampled population in the CFPS data, predictors about student activities are also considered potentially relevant to the percentile rank. Table 18 identifies significant relationships between student representatives ($p = 0.06$), student organisations ($p = 0$), and ranking. Both predictors have positive effects on the academic performance by 2.53 and 4.49 percentile points respectively. However, the gender difference is widened to 9.1 percentile points once both activity variables are controlled. These results, therefore, suggests positive relationships between activities and rank, but also points out that student activities could be a possible reason for the gender gap in academic performance. If there are more female representatives and organisation leaders, girls would also rank higher than boys on average. This possible interaction between gender and student activities will be further explored later.

5.4.4 Satisfaction

The fact that the differences in academic performance between gender groups, even after accounting for other differences in background characteristics and predictors, means there are potentially other factors to be explored. Adding variables related to satisfaction into the regression shows a 9.58% of the variance in university rank. Both variables are not statistically significant, and the gender difference is hardly changed

once one takes satisfaction levels into account. This could be because satisfaction with either the university or the tutor does not strongly affect a student's aspirations for better academic performance.

Table 19 Regression of the relationship between satisfaction and rank

Rank	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
Gender	8.719	1.351	6.45	0	6.066	11.372
Cohort	7.166	0.399	17.95	0	6.383	7.950
Age	-0.344	0.199	-1.72	0.085	-0.736	0.047
Ethnicity	-5.651	1.912	-2.96	0.003	-9.403	-1.898
Hukou status	-0.177	0.720	-0.25	0.806	-1.590	1.236
Family support	0.067	0.039	1.71	0.087	-0.009	0.144
Parental education	0.073	0.615	0.12	0.905	-1.139	1.286
Subject	-1.619	1.059	-1.53	0.127	-3.700	0.460
University type	-0.098	0.092	-1.06	0.289	-0.279	0.083
Cognitive tests	-0.135	0.050	-2.69	0.007	-0.234	-0.036
Talent (%)	0.016	0.030	0.53	0.595	-0.044	0.077
S: university	-0.581	0.657	-0.88	0.377	-1.872	0.710
S: tutor	0.009	0.046	0.2	0.844	-0.082	0.100
_cons	49.030	9.489	5.17	0	30.396	67.664

N=1033

5.5 Interaction between gender and leadership

Due to the impact of student activities on the gender gap in academic performance, as found in table 18, this report expects the activity variables, particularly the student representative and organisation leadership, to interact with gender differences. Hence, an interaction term is added to examine it. But unfortunately, both the representative and interaction variables are insignificant. It might be interpreted accordingly that although being a student representative affects his or her university rank, the influence remains the same for both boys and girls.

Table 20 Regression after adding interaction (gender*representative)

Rank	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
Gender	7.972	2.017	3.95	0	4.009	11.934
Cohort	6.839	0.449	15.21	0	5.956	7.722
Age	-0.614	0.241	-2.55	0.011	-1.089	-0.140
Ethnicity	-2.930	2.445	-1.2	0.231	-7.733	1.872
Hukou status	-1.015	0.856	-1.19	0.236	-2.697	0.666
Family support	0.029	0.047	0.62	0.537	-0.063	0.121
Parental education	0.083	0.661	0.13	0.899	-1.219	1.387
Subject	-1.878	1.071	-1.75	0.08	-3.983	0.227
University type	-0.072	0.105	-0.69	0.49	-0.279	0.134
Cognitive tests	-0.147	0.062	-2.34	0.019	-0.271	-0.023
Talent (%)	-0.007	0.036	-0.2	0.843	-0.078	0.064
Representative	-1.424	2.355	-0.6	0.546	-6.051	3.201
Gender*rep	-1.167	3.203	-0.36	0.716	-7.459	5.124
_cons	61.847	9.279	6.66	0	43.624	80.070

N=1001

According to table 21, the student organisation leadership and interaction terms are not significantly related to university rank. So, similar to the interaction between gender and student representative, the effect of being a student organiser does not differentiate between boys and girls. It can only be interpreted that student representatives and organisers are mostly higher-ranking students.

Table 21 Regression after adding interaction (gender*organisation leadership)

Rank	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
Gender	6.896	2.843	2.43	0.016	1.298	12.495
Cohort	6.315	0.681	9.26	0	4.973	7.657
Age	-0.536	0.447	-1.2	0.231	-1.418	0.344
Ethnicity	-8.856	3.760	-2.36	0.019	-16.258	-1.453
Hukou status	0.593	1.316	0.45	0.652	-1.998	3.185
Family support	0.546	1.216	0.45	0.654	-1.867	2.959
Parental education	-0.012	0.077	-0.16	0.875	-0.163	0.139
Subject	-1.574	1.651	-0.95	0.341	-4.825	1.676
University type	-0.062	0.132	-0.47	0.639	-0.322	0.198
Cognitive tests	0.018	0.1	0.18	0.857	-0.179	0.215
Talent (%)	-0.076	0.054	-1.39	0.165	-0.183	0.031
Organisation leader	-2.653	3.598	-0.74	0.462	-9.738	4.432
Gender*org lead	-1.403	5.118	-0.27	0.784	-11.479	8.673
_cons	-12662.2	1372.99	-9.22	0	-15365.33	-9959.075

N=369

6. Limits

Despite the findings demonstrated and explained in previous sections, this report has some limitations on both data availability and research design. As stressed in section 3, this report was not able to model the variation of academic performance for college students due to limitations in the CFPS data. Moreover, the cognitive ability test results and self-evaluated talents might not be accurate enough to provide previous attainment information prior to higher education. The National Higher Education Entrance Examination score could be of help if it was available in the CFPS data.

This report also acknowledges that students in the same tier of universities might not be academically identical. This is because top ranked universities, especially those in the first tier, are often obliged to take students from underdeveloped areas for the regional equality. Hence, these students might not necessarily be of the same academic standard as their peers.

Unfortunately, it is not possible for this report to explore what might be driving the unexplained differences in academic performance by gender using the CFPS data. The remaining differences could be studied with longitudinal data on Chinese university students. Further research could investigate male students, or those at risk of underperformance, further to contribute to the understanding of the gender gap.

7. Conclusion

This report conducts multiple regression analyses that shows gender differences in academic performance on around 3,000 Chinese university students sampled by the China Family Panel Studies. It also presents how gender gaps are changed by effects from demographic characteristics, student activities, attitudes, and satisfaction. It is supported by the regression results that there is a substantial gender difference in academic performance of around 8.68 percentile points among Chinese university students enrolled between the academic years 2010 to 2014.

Several background characteristics and predictors are found to be able to explain the gender disparity in academic performance. Despite the preference of subject and different gender ratio in each subject, girls still outperforming boys. Additionally, females even rank 7.29% higher in male-dominated subjects, including Science, Technology, Engineering, Agriculture, and Medicine. The cognitive ability tests result is surprisingly found to have small effect on the performance difference. This implies questions on the difficulty level of the tests or other factors regarding academic background, such as university type. In terms of predictors, findings in this report show that self-confidence, academic pressure, attitudes towards in-class concentration and university regulations present positive effects on narrowing the performance gap between male and female students. Both variables on student activities, i.e. student representative and student organisation leadership, indicates a widened gender disparity in university rank. But unfortunately, even when all the demographic characteristics and predictors are accounted for, there is still a noticeable gender gap in academic performance. This would indicate that there is much in the gender disparity that remains unexplained.

Some characters of well-performing university students are also found in this report. Those who are in the ‘Han’ majority ethnic group with comparatively higher cognitive abilities, self-confidence, and academic pressure tend to rank higher than the others. Other demographic factors regarding age, *hukou* status, parental education, and family-paid education expenditure are not linked to the outcome. In terms of attitudes, the power of negative influence from self-confidence is very similar to the one from academic pressure on a student’s academic performance, which may imply a phenomenon of the connection between higher aspirations, self-efficacy and academic outcome. On the contrary, having a positive attitude towards university regulations might emphasise characteristics of self-motivation, thus lead to a better academic performance; whereas the negative effect of in-class concentration on the university rank could be the result of comparatively lower academic abilities or previous

attainments. Satisfactions with the university attended or the tutor show no significant association with performance results. It is also tested that although the influence of student activities does not differ between boys and girls, being a student representative or organiser would potentially widen the gender gap in university rank. Overall, there is a performance difference in university rank of a 6.48% decrease between three year cohorts.

Other unexplained gender differences in academic performance could be investigated with longitudinal data on Chinese university students. Besides, the control of academic ability might be improved with more accurate measures of cognitive abilities and previous attainments, e.g. *gaokao* scores. Further research could explore the males' disadvantages in academic performance to contribute to explanations for the course completion rates in HEIs and changes in labour market structure.

References

Allen, J., Robbins, S., Casillas, A. and Oh, I. (2008). Third-year College Retention and Transfer: Effects of Academic Performance, Motivation, and Social Connectedness. *Research in Higher Education*, 49(7), pp.647-664.

Bailey, M. and Dynarski, S. (2011). 'Gains and Gaps: Changing Inequality in U.S. College Entry and Completion'. In: G. Duncan and R. Murnane, ed., *Whither Opportunity? Rising Inequality and the Uncertain Life Chances of Low-Income Children*, New York: Russell Sage Foundation, pp.117-132.

Buchmann, C. and Dalton, B. (2002). Interpersonal Influences and Educational Aspirations in 12 Countries: The Importance of Institutional Context. *Sociology of Education*, 75(2), p.99.

Chamorro-Premuzic, T. and Furnham, A. (2003). Personality predicts academic performance: Evidence from two longitudinal university samples. *Journal of Research in Personality*, 37(4), pp.319-338.

Chan, K. and Buckingham, W. (2008). Is China Abolishing the Hukou System?. *The China Quarterly*, 195.

Chen C, and Bo Qin. (2014) The Emergence of China's Middle Class: Social Mobility in a Rapidly Urbanizing Economy. *Habitat International*, (14),528-535.

Chinese Ministry of Education. (2013) *Educational Statistics in 2013 (Chinese Colleges and Universities)*. [data file] Beijing: Ministry of Education. [Accessed 29 July 2016]

Chowdry, H., Crawford, C., Dearden, L., Goodman, A. and Vignoles, A. (2012). Widening participation in higher education: analysis using linked administrative data. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 176(2), pp.431-457.

Choy, S. P. (2001). *Students whose parents did not go to college: Postsecondary access, persistence, and attainment. Findings from The Condition of Education 2001 (NCES 2001-126)*. Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Clark, N. (2010). *Chinese University Rankings - WENR*. [online] World Education News & Reviews. Available at: http://wenr.wes.org/ [Accessed 25 Jun. 2016].

Conger, D. and Long, M. (2010). Why Are Men Falling Behind? Gender Gaps in College Performance and Persistence. *The ANNALS of the American Academy of Political and Social Science*, 627(1), pp.184-214.

Crawford, C. (2014). *Socio-economic differences in university outcomes in the UK: drop-out, degree completion and degree class*. Institue for Fiscal Studies.

Crawford, C. and Greaves, E. (2015). *Socio-economic, ethnic and gender differences in HE participation*. Department for Business, Innovation and Skills.

DiPrete, T. and Buchmann, C. (2006). Gender-Specific Trends in the Value of Education and the Emerging Gender Gap in College Completion. *Demography*, 43(1), pp.1-24.

Duff, A., Boyle, E., Dunleavy, K. and Ferguson, J. (2004). The relationship between personality, approach to learning and academic performance. *Personality and Individual Differences*, 36(8), pp.1907-1920.

Furnham, A., Monsen, J. and Ahmetoglu, G. (2009). Typical intellectual engagement, Big Five personality traits, approaches to learning and cognitive ability predictors of academic performance. *British Journal of Educational Psychology*, 79(4), pp.769-782.

Gu, J. and Hua, J. (2010). *The Chinese higher education system and the impact of gender: the structure of Chinese education system and previous research and an empirical study*. Doctor of Economics and Social Science. The University of Oldenburg.

Honig, E. and Hershatter, G. (1988). *Personal voices*. Stanford, Calif.: Stanford University Press.

Jacob, B. (2002). Where the boys aren't: non-cognitive skills, returns to school and the gender gap in higher education. *Economics of Education Review*, 21(6), pp.589-598.

Jacob, W. (2006). Social Justice and Gender in Chinese Higher Education: Regional Issues of Equity and Access. *Education and Social Justice*, pp.139-159.

Jordan L.P, Ren Q, Falkingham J. (2014) Youth Education and Learning in Twenty-First Century China. *Chinese Sociological Review*,47(1):30-56.

Mouw, T. and Kalleberg, A. (2010). Occupations and the Structure of Wage Inequality in the United States, 1980s to 2000s. *American Sociological Review*, 75(3), pp.402-431.

Li, J. (2012). World-class higher education and the emerging Chinese model of the university. *Prospects*, *42*(3), 319-339.

Liu, J. 2006. Expansion of higher education in China and inequality in entrance opportunities: 1978-2003 (in Chinese). *Chinese Journal of Sociology* 26(3):158–79.

Liu, Y. (2013). Meritocracy and the Gaokao: a survey study of higher education selection and socio-economic participation in East China. *British Journal of Sociology of Education*, *34*(5-6), 868-887.

McDaniel, A. (2010). Cross‐National Gender Gaps in Educational Expectations: The Influence of National‐Level Gender Ideology and Educational Systems. *Comparative Education Review*, 54(1), pp.27-50.

McKenzie, K. and Schweitzer, R. (2001). Who Succeeds at University? Factors predicting academic performance in first year Australian university students. *Higher Education Research and Development*, 20(1), pp.21-33.

Mouw, T. and Kalleberg, A. (2010). Occupations and the Structure of Wage Inequality in the United States, 1980s to 2000s. *American Sociological Review*, 75(3), pp.402-431.

Nozaki, Y., Aranha, R., Dominguez, R. and Nakajima, Y. (2009). Gender gap and women's participation. In: D. Baker and A. Wiseman, ed., *Gender, Equality and Education from International and Comparative Perspectives*, 1st ed. London: Emerald Group Publishing, pp.217-254.

Ren, Q. and Treiman, D. (2013). The consequence of Parental Labour Mitigation in China for Children's Emotional Well-being. *Working Paper Series： WP13-001*.

Richardson, M., Abraham, C. and Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological Bulletin*, 138(2), pp.353-387.

Rohde, T. and Thompson, L. (2007). Predicting academic achievement with cognitive ability.*Intelligence*, 35(1), pp.83-92.

Rolfhus, E. and Ackerman, P. (1999). Assessing individual differences in knowledge: Knowledge, intelligence, and related traits. *Journal of Educational Psychology*, 91(3), pp.511-526.

Rong, X. and Shi, T. (2001). Inequality in Chinese Education. *Journal of Contemporary China*, 10(26), pp.107-124.

Tan, Y. (2012). Glass ceiling or sticky floor: an investigation into the gender difference in average earning after university graduation (in Chinese). *Population Journal*, 6(1), pp.24-49.

Yang, D. (2006). Access to higher education: Widening social class disparities (in Chinese). *Tsinghua Journal of Education.* 27(1):19–25.

Vincent-Lancrin, S. (2008). The Reversal of Gender Inequalities in Higher Education: An On-going Trend. In: *Higher Education to 2030 (Vol. 1): Demography*, 1st ed. Paris: OECD, pp.265-298.

Walker, A., Haiyan, Q., and Shuang, Z. (2011). Secondary school principals in curriculum reform: victims or accomplices?. *Frontiers of Education in China*,*6*(3), 388-403.

Wang, L. (2007). *A study on the relationship between university students' psychological conditions and academic achievements* (in Chinese). MSc in Management. University of Shandong.

Wang, G. and Zheng, Y. (2013). *China: Development and Governance*. Singapore: World Scientific.

Xie Y (2012) *China Family Panel Studies User's Manual* (in Chinese). Institute of Social Science Survey, Peking University, Beijing.

Xie, Y., & Hu, J. (2014). An introduction to the China family panel studies (CFPS). *Chinese Sociological Review*, *47*(1), 3-29.

Xie, Y. and Lu, P. (2015). The sampling design of the China Family Panel Studies (CFPS). *Chinese Journal of Sociology*, 1(4), pp.471-484.

Xiong, H. (2006). *Emplyability of Chinese university graduates: evidence from the*

university student career aspiration survey (in Chinese). MSc in Psychology. Jiangxi Normal University.

Yu, K., Stith, A., Liu, L. and Chen, H. (2012). *Tertiary Education at a Glance*. Rotterdam: Springer Science & Business Media.

Yu, L., Yu, Y. and Mansfield, P. (1990). Gender and changes in support of parents in china: Implications for the One-Child Policy. *Gender and Society*, 4(1), pp.83-89.

Zhang, Z. and 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. *J. Chin. Sociol.*, 1(1).

Zhang, H., Patton, D., and Kenney, M. (2013). Building global-class universities: Assessing the impact of the 985 Project. *Research Policy*, *42*(3), 765-775.