

Li, Zhengyang; Zhao, Guochang

Working Paper

The quiet revolution in women's human capital and the gender earnings gap in the People's Republic of China

ADB Working Paper Series, No. 1112

Provided in Cooperation with:

Asian Development Bank Institute (ADBI), Tokyo

Suggested Citation: Li, Zhengyang; Zhao, Guochang (2020) : The quiet revolution in women's human capital and the gender earnings gap in the People's Republic of China, ADB Working Paper Series, No. 1112, Asian Development Bank Institute (ADBI), Tokyo

This Version is available at:

<https://hdl.handle.net/10419/238469>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by-nc-nd/3.0/igo/>



ADB Working Paper Series

**THE QUIET REVOLUTION IN
WOMEN'S HUMAN CAPITAL AND
THE GENDER EARNINGS GAP IN
THE PEOPLE'S REPUBLIC OF CHINA**

Zhengyang Li and
Guochang Zhao

No. 1112
April 2020

Asian Development Bank Institute

Zhengyang Li is a PhD candidate in economics, and Guochang Zhao is an assistant professor of economics, at the Research Institute of Economics and Management of Southwestern University of Finance and Economics in Chengdu, People's Republic of China.

The views expressed in this paper are the views of the author and do not necessarily reflect the views or policies of ADBI, ADB, its Board of Directors, or the governments they represent. ADBI does not guarantee the accuracy of the data included in this paper and accepts no responsibility for any consequences of their use. Terminology used may not necessarily be consistent with ADB official terms.

Working papers are subject to formal revision and correction before they are finalized and considered published.

The Working Paper series is a continuation of the formerly named Discussion Paper series; the numbering of the papers continued without interruption or change. ADBI's working papers reflect initial ideas on a topic and are posted online for discussion. Some working papers may develop into other forms of publication.

The Asian Development Bank refers to "China" as the People's Republic of China.

Suggested citation:

Li, Z. and G. Zhao. 2020. The Quiet Revolution in Women's Human Capital and the Gender Earnings Gap in the People's Republic of China. ADBI Working Paper 1112. Tokyo: Asian Development Bank Institute. Available: <https://www.adb.org/publications/quiet-revolution-women-human-capital-gender-earnings-gap-prc>

Please contact the authors for information about this paper.

Email: zhengyang_li@foxmail.com, guochangzhao@swufe.edu.cn

Asian Development Bank Institute
Kasumigaseki Building, 8th Floor
3-2-5 Kasumigaseki, Chiyoda-ku
Tokyo 100-6008, Japan

Tel: +81-3-3593-5500
Fax: +81-3-3593-5571
URL: www.adbi.org
E-mail: info@adbi.org

© 2020 Asian Development Bank Institute

Abstract

Using the 2010 and 2014 data from the China Family Panel Studies (CFPS), this paper analyzes the effect of human capital on the gender earnings gap, both within cohorts and across cohorts using regression, Oaxaca-Blinder, and Juhn-Murphy-Pierce decomposition analyses. On the one hand, over the past few decades, Chinese women have staged a ‘quiet revolution’ in education and cognitive abilities, and especially after the large-scale higher education expansion since 1999, women have outperformed men in terms of entering college. On the other hand, the gender earnings gap is reducing among the 1980s birth cohort relative to older cohorts (the 1960s and 1970s birth cohorts). Despite the advantages in terms of education and cognitive abilities of younger women, as well as a higher rate of returns on education for women, the gender earnings gap is still large: 20%–30%. When exploring the effect across cohorts, we find that the gap effect (discrimination and unobserved skills) accounts for more than 50% of the change in the gender earnings gap across cohorts. These results imply that labor market discrimination is the main barrier to narrowing the gender earnings gap in the People’s Republic of China.

Keywords: education, cognitive ability, gender earnings gap

JEL Classification: J31, J71, I24, I26

Contents

1.	INTRODUCTION	1
2.	CHANGES IN EDUCATION, COGNITIVE ABILITY, AND EARNINGS OVER TIME..	3
3.	DATA AND SUMMARY STATISTICS	4
3.1	Data	4
3.2	Summary Statistics	6
4.	ECONOMETRIC MODEL	8
4.1	OLS	9
4.2	Oaxaca-Blinder Decomposition	9
4.3	JMP Decomposition	10
5.	EMPIRICAL RESULTS	11
5.1	OLS Results	11
5.2	Oaxaca-Blinder Decomposition Results	14
5.3	JMP Decomposition Results	15
6.	SUMMARY AND CONCLUSION	17
	REFERENCES	19
	APPENDIX	21

1. INTRODUCTION

Gender inequality has always been an essential issue in the labor market. During the past few decades, both the labor market input factors (e.g., education) and labor market output results (e.g., earnings and the labor participation rate) have exhibited a convergence in gender differences. In the United States, the gender gap in earnings¹ measured by the earnings ratio of females to males was approximately 60% from 1955 to 1980; subsequently, the earnings ratio increased by 20% in the 1980s before the reduction progress slowed, and the earnings ratio rose to approximately 80% in 2014 (Blau and Kahn 2017). Other economically advanced nations also experienced a substantial reduction in the gender earnings gap (Blau and Kahn 2008). In the People's Republic of China (PRC), the gender earnings gap was relatively small in the planned economy period because workers in state-owned enterprises were paid relatively equally. Since the 1980s, firms obtained increasing amounts of freedom to select workers and set earnings, and the market became more competitive under reform. Women's labor market conditions worsened because of their lower education and experience levels compared to men. Moreover, there also emerged more serious discrimination in the labor market (Li et al. 2014). As a result, the gender earnings gap gradually widened (Gustafsson and Li 2000; Zhang et al. 2008).

Women gained ground in education; labor force participation; work experience; labor hours; occupation choices; and industry options, among others. All these factors hold some influence on the gender earnings gap (Goldin 2014; Blau and Kahn 2017). Among these variables, education is often the most focused upon, because it is the best indicator of work skills, and highly skilled individuals have a higher return on education and are more likely to have a high-paying position (Blau and Kahn 2017). Liu (2011) finds that increasing one's education level to obtain a job or higher pay is more effective for women than men. Employed women not only have more schooling, but their education level also increases faster than that of the full sample of women, and the reduction of the gender earnings gap from 1993-2004 largely relied on the improved years of schooling of employed women (Liu 2011).

In addition to education, cognitive and noncognitive abilities also have substantial impacts on labor market outcomes (Heckman et al. 2006). Cognitive and noncognitive abilities act as a new way to understand the unexplained part of previous models, which mainly contain education, experience, occupation, and industry. It has been found that women outperform men in word tests, but men do better in math tests and are more likely to work in mathematically intensive positions (Pope and Sydnor 2010; Ceci et al. 2014). Fortin (2008) finds that earnings increase with high school math scores, and the returns to the math scores for the young cohort is twice that for the old cohort. Using the same data set, Blau and Kahn (2017) decompose the earnings gap, showing that math scores can explain 10%–14% of the change in the gender earnings gap across cohorts. Additionally, many studies give evidence on the contributions of noncognitive abilities to labor market outcomes, such as leadership skills (Kuhn and Weinberger 2005), self-esteem (Fortin 2008), and interpersonal skills (Borghans et al. 2014), among others.

¹ Blau and Kahn (2017) use the usual weekly earnings of full-time workers and the annual earnings of full-time, year-round workers from the data of Current Population Survey. The components of earnings are different in papers, and many papers use labor income of employees.

We explore the gender earnings gap from a different perspective than the extant literature. The existing studies usually focus on the gender earnings gap trend over periods of years. The results do not explicitly illustrate the different influences that the factors have on earnings in different cohorts (Blau and Kahn 1997; Gustafsson and Li 2000; Liu, 2011). Although some studies focus on specific cohorts and study the change within cohorts and across cohorts, the cohorts are too unique to represent the entire population (Fortin 2008). We divide the whole sample into three birth cohorts: the 1960–1969 cohort, the 1970–1979 cohort, and the 1980–1989 birth cohort. It is believed that these three birth cohorts have very distinct characteristics among them, and they have had very different experiences. For example, the 1960–1969 birth cohort spent most of their childhood during the Cultural Revolution, whereas the 1980–1989 birth cohort was confronted with the expansion of higher education and the introduction of the One-Child Policy. Due to the dual-track system in the PRC, looking into the gender earnings gap across birth cohorts at the same time can, to some extent, reveal more changes in the labor market than looking into the gender earnings gap over time.

Goldin (2006) has argued that there was a ‘quiet revolution’ from the late 1970s to the present in the United States: because of the usage of pills, women increased their investments in education before marriage; made greater preparation for the labor market; and developed a better career and better pay than their previous generations. In the PRC, after the implementation of the One-Child Policy, parents changed the previous gender preference and increased education investment in girls who were the only child in the families, and the gender education gap therefore gradually decreased (Tsui and Rich 2002; Lee 2012). Later, the expansion of higher education provided more opportunities to enter universities. Girls are more likely to obtain this opportunity because they have higher average scores in high school and behave more appropriately than boys in high schools. Therefore, the reverse condition appeared in the gender education gap (Goldin et al. 2006; Buchmann and DiPrete 2006; Becker et al. 2010). As the gender education gap decreased and even reversed, it is interesting to investigate how the gender earnings gap changed along with the increasing female human capital.

We use data from the China Family Panel Studies (CFPS). To enlarge the sample size, we pool the 2010 and 2014 data together. The data have detailed demographic and earnings information, as well as questions about cognitive and noncognitive abilities, which precisely suits the aim of this study. Given the reverse condition in gender education gap and the requirement to analyze the gender earnings gap by cohorts, this paper aims to investigate the effect of education and cognitive/noncognitive abilities on the gender earnings gap within and across cohorts.

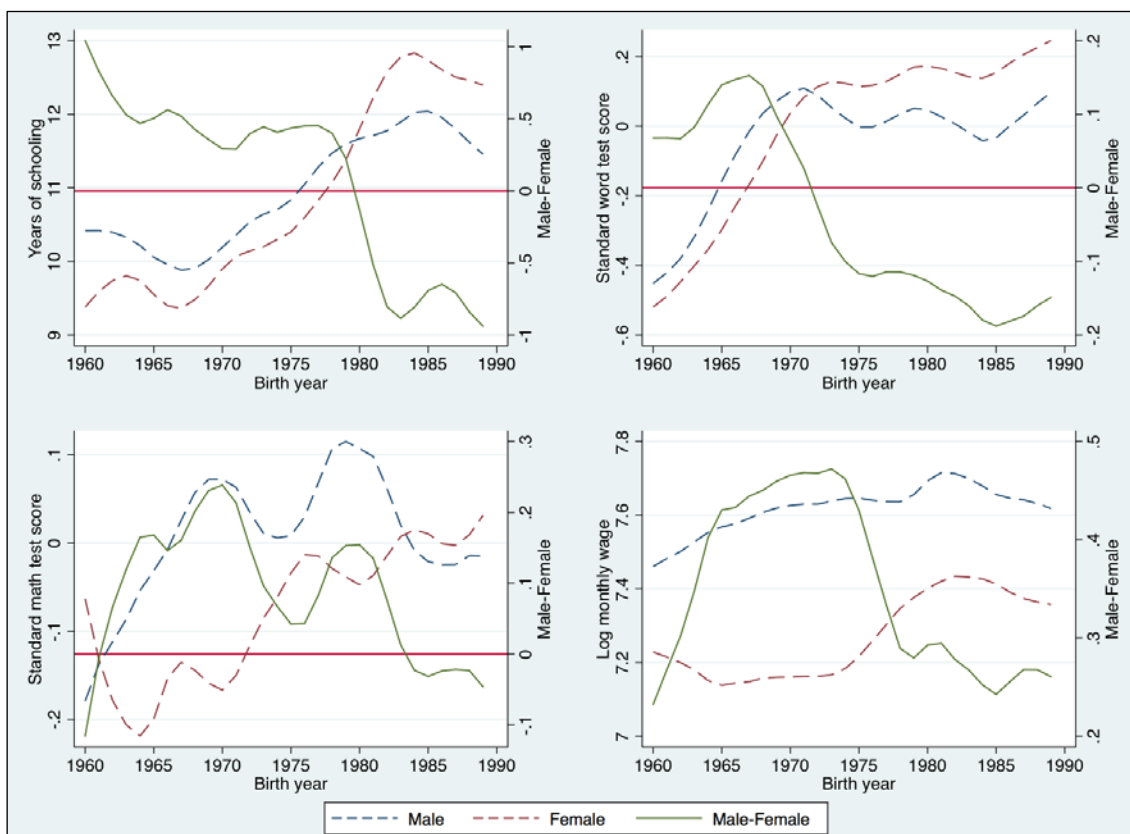
There are five main findings. First, just like the US, Chinese women also have been experiencing a ‘quiet revolution’ in terms of their human capital, especially in education and cognitive abilities. Second, the gender earnings gap remains the same in the 1960–1969 cohort and the 1970–1979 cohort, that is, 0.40 log points. However, the gap substantially decreases in the 1980–1989 cohort. Third, the returns on human capital substantially change across cohorts. Returns on education decrease across cohorts, largely because of the increased education level, which may lead to a lower education premium. Earnings increase with math test scores in the 1960–1969 birth cohort, while earnings increase with word test scores in the 1970–1979 birth cohort. For the youngest cohort, both test scores are important for earnings. Fourth, within cohorts, college degrees and cognitive abilities account for little of the total gender difference in earnings in the 1960–1969 and 1970–1979 cohorts. However, college degrees and word tests help to narrow the gender gap in the 1980–1989 cohort. Last, we find that the gap effect (unobserved skills) contributes to over 50% of the gap change across cohorts.

The rest of the paper is structured as follows. Section 2 presents the trends of key variables across cohorts. Section 3 describes the data and presents the summary statistics. Section 4 illustrates the econometric models, and Section 5 reports and discusses the results. The final section offers our conclusions.

2. CHANGES IN EDUCATION, COGNITIVE ABILITY, AND EARNINGS OVER TIME

To explain the impetus behind our research, we present the trends of key variables – human capital and earnings – across cohorts in Figure 1, which is based on the CFPS data. In all four graphs, the blue dashed curve represents men, the red dashed curve represents women, the green solid curve is the gender gap (men–women), and the red solid horizontal line represents the zero-gender gap. The level value refers to the left vertical axis, and the gap value refers to the right vertical axis.

Figure 1: The Changes in Key Variables among Birth Cohorts



Data Source: The China Family Panel Studies (2010 and 2014).

The graph in the upper left corner presents the relationship between education and cohorts. Women born before 1980 have fewer years of schooling than men. The education gap is 1 year for the 1960 cohort. Both genders receive equal years of schooling in the 1980 cohort, and women surpass men thereafter. The overall gap in years of schooling decreases from the 1960 cohort to the 1970 cohort. After the 1970 cohort, the gender gap in education first rises and then drops to zero in the 1980 cohort.

People born after 1980 experience a large reduction in the education gap. For the 1989 cohort, the gender gap in years of schooling is -1 year.

In addition to the usual educational attainment, the CFPS also conducted a word test and math test to evaluate basic cognitive abilities. The trend of standard word test scores, the upper right corner of Figure 1, is also very clear. The gender gap is decreasing. The graph in the lower left corner depicts the trend of standard math test scores across cohorts. The number of observations who were born in the early 1960s is relatively small in this sample, so we focus on the trend after this period. Both male and female math scores are rising with fluctuation. Male scores decrease a little after the 1980 cohort, and females increase their math scores at the same time; thus, women catch up with men and surpass them in math test scores.

The last graph is the trend of log monthly earnings. The gap is above 0.2 log points in all cohorts. The trend can be divided into three parts. For the 1960–1969 birth cohort, the gender gap in earnings widens greatly and then undergoes a substantial reduction for the 1970–1979 birth cohort. For the 1980–1989 birth cohort, the gender earnings gap decreases slightly.

To summarize, the gender gap in human capital is decreasing and even appears to have a reverse condition. The gender earnings gap also narrows, but the gap still exists. Specifically, as female education levels rise, the gender earnings gap decreases; the gender gaps in earnings and math and word test scores have all decreased from their levels in the 1970 cohort. Thus, there should be some relationship between education, cognitive/noncognitive abilities, and earnings, and these factors may have different contributions in different cohorts. We will discuss the relationship in detail in the following sections.

3. DATA AND SUMMARY STATISTICS

3.1 Data

The paper uses the data from the CFPS, conducted by the Institution of Social Science Survey, Peking University. It is longitudinal data, surveyed in 2010, 2012, 2014, and 2016, and the respondents are from 25 provinces in the PRC. The questionnaires consist of three parts: community, household and individual. We use the individual part from 2010 and 2014.² The individual questionnaire asks the respondents their basic demographic information, education history, marriage status, work, income, social relationship, health condition, and a set of questions testing cognitive/noncognitive abilities. These pieces of information are directly suited to this research.

To enlarge the sample, we pool the 2010 and 2014 data together. Then, we restrict the sample to individuals who were born from 1960 to 1989, were engaged in nonfarm jobs, participated in the labor market, and earned at least 50 yuan on average during the past 12 months. These sample restrictions result in 11,117 observations. Furthermore, we remove the observations that have missing values for education, cognitive/noncognitive abilities, and so on. Ultimately, we obtain a sample consisting of 9,475 observations: 4,253 from the 2010 survey and 5,222 from the 2014 survey.

The earnings are the sum of monthly earnings; monthly bonuses; and subsidies, and are deflated to 2010 prices using the consumption price index (CPI). As we know, there is a

² The 2012 CFPS has different cognitive/noncognitive ability test questions, and the 2016 CFPS was not available when the paper was being writing.

nonlinear relationship between age and earnings, and this is more complicated in the PRC because of the large-scale reforms and rapidly changing labor market. Therefore, we control for the cohort dummies to make the relationship more flexible, but this is a little tricky. If the cohort is defined too narrowly, there are too many dummies in the model, which might make estimation infeasible sometimes. If the cohort is too wide, there are too few dummies, and things may not work as we expect. Finally, given our sample size, we define the cohort group with a 5-year window – that is, we classify all individuals in our sample into 6 cohort groups: 1960–1964, 1965–1969, 1970–1974, 1975–1979, 1980–1984, and 1985–1989.

Education is one of the most important explanatory variables in this study. Here, we classify education into two groups – senior high school or below and college or above – and use a dummy for college or above to measure educational attainment. We use this classification scheme because senior high school and below aim to help individuals establish basic reading, calculation, and logic abilities, while college and above mainly helps people obtain professional abilities that are more related to work. Appendix Table A1 provides the evidence: in the full sample, the returns on education for those with a senior high school education or below are small and rarely significant, but the returns on education for college and above are much larger and are all statistically significant.

Regarding cognitive abilities, the CFPS has test questions about words and math. Word tests are designed to examine literacy, and math tests contain certain mathematical problems. The test questions are ordered from easy to hard, and the starting question depends on the respondent's highest level of education. For example, individuals with senior high school degrees start the word test from the 21st word and the math test from the 19th problem; individuals with junior high school educations start the word test from the 9th word and the math test from the 13th problem. If the respondent continuously makes mistakes in three words/problems, or finishes all the words/problems, the test is over and the respondent gets a score equal to the number of questions answered correctly. Because the starting point of the test is related to education level, we standardize the word and math test scores with zero mean and unity at each educational level.

As for noncognitive abilities, the questions are not consistent between the 2010 and 2014 questionnaires; therefore, we use factor analysis to extract noncognitive abilities and get three comprehensive variables: emotional stability, self-esteem and popularity, among which emotional stability is one of the main elements of the Five-Factor Model. These factors are found to have substantial effects on earnings (Judge et al. 1999; Nyhus and Pons 2005; Drago 2011; Borghans et al. 2014). In the CFPS data, emotional stability consists of 6 questions that ask interviewees about the frequency with which they have experienced the following feelings in the past month: “depressed, and unable to get excited about anything”; “nervous”; “restless and unable to calm down”; “no hope for the future”; “have a difficult time doing anything”; and “life is meaningless.” The interviewee will choose one answer from “almost every day”, “often”, “half the time”, “sometimes” to “never” for each question, and the answer is coded 1, 2, 3, 4, or 5. Thus, a larger value of emotional stability means more positive emotions. Self-esteem is comprised of three survey questions: “Are you happy?”, “Are you satisfied with your life?”, and “Are you confident about your future?”. The answers are also coded from 1 (lowest) to 5 (highest). Popularity is based on the following two questions: “Do you think you are popular?” and “Is it easy for you get along well with others?”. The answers are coded as variables of self-esteem.

In the PRC, *hukou* (the household registration system) has an important effect on earnings (Afridi et al. 2015). Here, we define a binary variable: 0 for an agricultural *hukou* and 1 for a nonagricultural *hukou*. Fertility status is defined as 1 for at least one child and

0 for no children. Marital status is defined as 1 for married, including windowed and divorced, and 0 for unmarried, including both single and cohabitation. Since we restrict the sample to nonfarm workers, we eliminate individuals who are engaged in farm occupations. In the paper, we classify individuals into six occupations: managers, professionals, clerical workers, business and service workers, manufacturing and transportation workers, and other occupations. There are 21 types of industries in the original data set, including the farming industry. After removing the farming industry, we classify the rest of the industries into six groups according to the industries' mean earnings and denote Industry 1 to Industry 6 according to the highest mean earnings to the lowest. Appendix A2 presents the details of these industry classifications. Employers were classified into five groups: government departments, public institutions, state-owned enterprises, private enterprises, including joint ventures and foreign enterprises, and others.

3.2 Summary Statistics

Table 1 presents the summary statistics. The mean gender earnings gap is approximately 0.40 log points³ for the 1960s birth cohort, and it remains the same for the 1970s birth cohort, although both males' and females' earnings increase relative to the previous birth cohort. However, for the 1980s birth cohort, women's earnings increased faster than men's, leading to a 0.13 log point reduction in the gender earnings gap, accounting for 32.5% of the earnings gap in the previous two cohorts.

As expected, the younger cohorts are more likely to receive higher education due to the continuous expansion of higher education, but women benefit more. Among the 1980s birth cohorts, approximately 44% of women have a college degree or above, while this is only the case for 35% of men. Overall, Chinese women experienced a 'quiet revolution' in terms of higher education and now outperform men in this respect, as was the case in the US during the past few decades.

As for cognitive abilities, women have made great progress. In terms of word tests, men did better than women in the 1960s birth cohort, but not to a statistically significant extent. In the 1970s birth cohort, women caught up and even obtained higher test scores, but not to a statistically significant extent. In regard to the 1980s birth cohort, women gain ground, just as they did when it came to obtaining college degrees. In terms of math tests, the situation is similar: women progress faster than men, but from a position of disadvantage to one of no difference.

Compared to education and cognitive abilities, the change of the gender differential in noncognitive abilities across cohorts has no clear pattern. Emotion becomes less stable in younger cohorts, but it corrodes more quickly for men than for women. As a result, in the 1960s birth cohort, men's emotion stability was still 0.13 unit higher than women's, but since the 1970s birth cohort, the advantage has disappeared. In contrast, both genders' self-esteem increases across cohorts, indicating that people feel more and more satisfaction and happiness in their lives and gradually developed more confidence in their futures, but from the 1970s birth cohort onward, women's self-esteem increased at a higher speed than men's, leading to a reversal of the gender gap in self-esteem. Popularity does not change significantly, and women are more likely to feel popular than men among all cohorts.

Table 1: Summary Statistics

³ It means male earnings is about 40% higher than that of the female.

Variables	1960s Cohort			1970s Cohort			1980s Cohort		
	Male	Female	M-F	Male	Female	M-F	Male	Female	M-F
Log monthly earnings	7.552	7.148	0.40***	7.631	7.226	0.40***	7.680	7.409	0.27***
<i>Cohort groups</i>									
1960–1964	0.439	0.373	0.07***	–	–	–	–	–	–
1965–1969	0.561	0.627	–0.07***	–	–	–	–	–	–
1970–1974	–	–	–	0.527	0.481	0.05**	–	–	–
1975–1979	–	–	–	0.473	0.519	–0.05**	–	–	–
1980–1984	–	–	–	–	–	–	0.468	0.463	0.01
1985–1989	–	–	–	–	–	–	0.532	0.537	–0.01
Age	47.124	46.757	0.37**	37.959	37.920	0.04	27.792	27.806	–0.01
Hukou	0.577	0.599	–0.02	0.562	0.571	–0.01	0.456	0.489	–0.03
College and advanced degree	0.165	0.172	–0.01	0.281	0.259	0.02	0.353	0.441	–0.09***
Standardized word test score	–0.175	–0.249	0.07	0.061	0.113	–0.05	0.015	0.182	–0.17***
Standardized math test score	–0.025	–0.184	0.16***	0.052	–0.060	0.11***	0.005	–0.013	0.02
Emotional stability	0.073	–0.059	0.13**	0.014	–0.036	0.05	–0.002	–0.045	0.04
Self-esteem	–0.031	–0.048	0.02	–0.037	0.057	–0.09**	0.062	0.143	–0.08*
Popularity	–0.042	0.047	–0.09*	–0.025	0.046	–0.07	–0.054	0.027	–0.08*
Have children	0.933	0.938	–0.01	0.910	0.954	–0.04***	0.500	0.550	–0.05**
Married	0.995	0.995	0.00	0.954	0.985	–0.03***	0.641	0.711	–0.07***
<i>Occupations</i>									
Managers	0.098	0.039	0.06***	0.080	0.047	0.03***	0.064	0.044	0.02*
Professionals	0.094	0.161	–0.07***	0.116	0.177	–0.06***	0.144	0.267	–0.12***
Clerical workers	0.158	0.104	0.05***	0.113	0.114	–0.00	0.115	0.176	–0.06***
Business and service workers	0.101	0.370	–0.27***	0.119	0.327	–0.21***	0.156	0.279	–0.12***
Prod., transp. workers	0.534	0.308	0.23***	0.555	0.323	0.23***	0.497	0.214	0.28***
Others	0.016	0.018	–0.00	0.018	0.012	0.01	0.024	0.020	0.00
<i>Industries</i>									
Industry 1	0.172	0.077	0.10***	0.143	0.066	0.08***	0.134	0.060	0.07***
Industry 2	0.108	0.053	0.06***	0.128	0.054	0.07***	0.136	0.088	0.05***
Industry 3	0.175	0.116	0.06***	0.120	0.139	–0.02	0.112	0.176	–0.06***
Industry 4	0.306	0.370	–0.06***	0.336	0.392	–0.06**	0.369	0.369	–0.00
Industry 5	0.072	0.105	–0.03**	0.112	0.176	–0.06***	0.109	0.182	–0.07***
Industry 6	0.167	0.279	–0.11***	0.161	0.174	–0.01	0.141	0.125	0.02
<i>Employers</i>									
Government departments	0.103	0.099	0.00	0.080	0.061	0.02*	0.059	0.044	0.02
Public institutions	0.114	0.135	–0.02	0.119	0.155	–0.04**	0.086	0.157	–0.07***
State-owned enterprises	0.240	0.149	0.09***	0.191	0.119	0.07***	0.159	0.112	0.05***
Other enterprises	0.485	0.560	–0.08***	0.557	0.615	–0.06***	0.661	0.649	0.01
Others	0.058	0.057	0.00	0.054	0.051	0.00	0.035	0.039	–0.00
2014 data	0.540	0.567	–0.03	0.562	0.606	–0.04**	0.587	0.623	–0.04*
<i>Regions</i>									
Northeast	0.211	0.201	0.01	0.161	0.139	0.02	0.130	0.109	0.02
North coast	0.167	0.119	0.05***	0.179	0.157	0.02	0.223	0.208	0.01
East coast	0.121	0.179	–0.06***	0.107	0.155	–0.05***	0.128	0.179	–0.05***
South coast	0.078	0.095	–0.02	0.110	0.121	–0.01	0.149	0.130	0.02
The mid-Yellow River	0.148	0.120	0.03*	0.156	0.131	0.02*	0.143	0.131	0.01
The mid-Yangtze River	0.134	0.156	–0.02	0.151	0.133	0.02	0.088	0.130	–0.04***
Southwest	0.119	0.120	–0.00	0.122	0.141	–0.02	0.126	0.099	0.03*
Northwest	0.022	0.012	0.01***	0.015	0.022	–0.01*	0.012	0.014	–0.00
N	1,749	1,181	–	1,837	1,425	–	1,823	1,460	–

Note. *** p<0.01, ** p<0.05, * p<0.10.

More than 90% of the 1960s and 1970s cohorts had at least one child, while the proportion is only approximately 50% for the 1980s cohort, which is as expected. Women are 4 and 5 percentage points more likely to have a child than men in the 1970s and 1980s birth cohorts, respectively. Younger cohorts are more likely to be single – this finding is also consistent with our expectations, but men rather than women are more likely to be unmarried in the 1970s and 1980s birth cohorts.

As mentioned previously, we classify people into 6 occupations. Older men are more likely to be managers, but this advantage of men decreases substantially for younger cohorts: the gender gap declines from 0.06 in the 1960s cohort to 0.03 in the 1970s cohort and ultimately to 0.02 in the 1980s cohort. Young cohorts are more likely to be professionals, in contrast to the occupation of managers, but the increase is much higher for women than for men since the 1980s cohort. As a result, more women are engaged as professionals than men in the 1980s cohorts.

We can see from the table that in the first two industry classes, the proportion of males is larger than that of females in all three cohorts. Both the male and female proportions of the first industry class decrease across cohorts. In the most high-paying industry class, the proportion of men is 7–10 percentage points higher than that of women. For the second industry class, male and female proportions increase along with cohorts, and the gender gap is 5–7 percentage points. In the third industry class, the gender gap changes from a significant 6 percentage points to a significantly negative 6 percentage points because the male proportion decreases and the female proportion increases. Women and men mostly choose to work in the fourth industry class. The male proportion increases gradually in the fourth industry class; however, the female proportion decreases a little over all, and the gender gap is not significant for the last cohort. In the last two industry classes, there is a larger proportion of women, except in the sixth industry class for the last cohorts, but the gender gap is not significant. In short, men are more likely to have a job in the high-paying industries, and women are more likely to work in the low-paying industries. However, this condition is changing in a convergent trend.

The number of people working in government departments and state-owned enterprises shows a decreasing trend across cohorts, and the gender gap in those types of employers only changes a little. It is worth noting that both the male and female proportions of non-state-owned enterprises are increasing across cohorts, and the gender gap indicates that men have a higher growth rate. In short, the level of marketization is on the rise, which changes people's choice of employer and, thus, their income.

4. ECONOMETRIC MODEL

To study the gender earnings gap by cohorts, we set three goals in the empirical section of this paper. First, we want to know how the unexplained gender differential varies between models and how independent variables influence the earnings. To this end, we use the ordinary least squares (OLS) model. Second, Oaxaca-Blinder decomposition can help find the magnitude of contribution that each factor contributes to the gender earnings gap. This is a static analysis within each cohort. Finally, we also want to see the gender earnings gap in a dynamic perspective. Juhn, Murphy, and Pierce (1991) developed a decomposition method (JMP decomposition) that can help understand how the changes in different factors across cohorts contribute to the changes in the gender earnings gap.

4.1 OLS

First, we establish an OLS model:

$$y_i = \beta \cdot gender_i + X_i' \gamma + \varepsilon_i \quad (1)$$

where y_i is the log monthly earnings; $gender_i$ is the gender dummy, 0 for female and 1 for male; X_i is a vector of control variables, including education, cognitive/noncognitive abilities, *hukou*, fertility and marital status, occupation, industry and the region dummies. We cluster the standard deviation at the county level. Here, we are interested in the coefficient of gender, β , which measures the gender gap between the males and females, holding the covariates equal in both groups.

4.2 Oaxaca-Blinder Decomposition

According to the method proposed by Oaxaca (1973) and Blinder (1973), the gender-specific OLS model can be written as:

$$y_g = X_g b_g + \varepsilon_g, \quad g = m, f \quad (2)$$

where g is the subscript of gender, and m and f represent male and female, respectively. The log earnings differential between the two genders is:

$$\Delta y = \bar{X}_m b_m - \bar{X}_f b_f \quad (3)$$

If we use the male earnings structure as the nondiscrimination coefficient, the gender earnings gap can be written as:

$$\Delta y = (\bar{X}_m - \bar{X}_f) b_m + \bar{X}_f (b_m - b_f) \quad (4)$$

In contrast, if women are treated without advantages and discrimination, the coefficient of their earnings structure is nondiscriminating:

$$\Delta y = (\bar{X}_m - \bar{X}_f) b_f + \bar{X}_m (b_m - b_f) \quad (5)$$

The first term in the right side of Equations (4) and (5) is the “endowment” effect, also called the explained part, which denotes the earnings differential induced by differences of characteristics. The second part is the “discrimination” effect or the unexplained part, which is the earnings gap induced by the return on characteristics.

However, using the male coefficient or female coefficient as the nondiscrimination structure will lead to an “index number problem” (Oaxaca 1973), because we will never know which gender is preferred or discriminated against in the market. To solve this problem, Neumark (1988) proposed using pooled regression coefficients as the nondiscriminating earnings structure. The gender earnings differential is:

$$\Delta y = (\bar{X}_m - \bar{X}_f) b^* + \bar{X}_m (b_m - b^*) + \bar{X}_f (b^* - b_f) \quad (6)$$

where b^* is the regression coefficient from the pooled sample. The first term of the right side of Equation (6) is the explained part. The second and third parts comprise the unexplained part, which is the sum of the advantage of men and discrimination against women. The decomposition based on the pooled estimates may overstate the explained

part of the gender earnings while underestimating the unexplained part of the gender earnings gap (Jann 2008). Fortin (2008) solves this problem by adding a gender dummy to the pooled earnings regression. We follow Fortin's approach in this paper.

4.3 JMP Decomposition

To investigate gender gap changes across two cohorts, we follow the method developed by Juhn, Murphy, and Pierce (1991), who treat the residual of the earnings equation as the product of the residual standard deviation of male earnings and a standardized residual (i.e., with mean zero and variance), decomposing the gender earnings gap into the gender-specific factors and price factors. Furthermore, we use the decomposition form proposed by Blau and Kahn (1997), based on the idea of Juhn, Murphy, and Pierce (1991), and set the latter period to be the base period. We follow this method because the gender earnings gap in our sample is smaller in the young cohort, which suggests that the young cohort has a reduced discrimination effect. Specifically, the gender earnings gap, using the male earnings equation's coefficients as benchmarks, is:

$$\Delta y_1 = \Delta X_1 b_1 + \sigma_1 \Delta \theta_1 \quad (7)$$

where ΔX_1 is the average difference of observed characteristics in period 1; b_1 is the male equation coefficients in period 1; σ_1 is the standard deviation in the male earnings equation, representing the return on unobserved skills in period 1; and $\Delta \theta_1$ is the gender difference of average standardized residual percentile ranking in period 1. As for the time difference in gender difference, it can be obtained by:

$$\begin{aligned} \Delta y_1 - \Delta y_0 &= (\Delta X_1 - \Delta X_0) b_1 + \Delta X_0 (b_1 - b_0) \\ &\quad + (\Delta \theta_1 - \Delta \theta_0) \sigma_1 + (\sigma_1 - \sigma_0) \Delta \theta_0 \end{aligned} \quad (8)$$

Here, the first term on the right side of the equation is the "observed X 's effect". This partly measures the change in the gender earnings gap, which equals the average difference of gender characteristics changes multiplied by period 1's male coefficient. The second term is the "observed price effect", which denotes the contribution to the gender earnings gap change of changes in the earnings structure. The former two terms amount to the "explained" part of the Oaxaca-Blinder decomposition. The third term reflects the standardized residual dispersion change weighted by period 1's standard deviation, called the "gap effect". This term measures the unobserved characteristics, as well as discrimination. The last term is the "unobserved price effect", which is the product of the unobserved price change multiplied by the dispersion difference in period 0. The last two terms together equal the "unexplained" part of the Oaxaca-Blinder decomposition.

The "gap effect" can be written as:

$$\begin{aligned} (\Delta \theta_1 - \Delta \theta_0) \sigma_1 &= \Delta \theta_1 \sigma_1 - \Delta \theta_0 \sigma_1 \\ &= (\theta_{1m} - \theta_{1f}) \sigma_1 - (\theta_{0m} - \theta_{0f}) \sigma_1 \\ &= -\theta_{1f} \sigma_1 + \theta_{0f} \sigma_1 \end{aligned} \quad (9)$$

The “unobserved price effect” can be extended in the same way. To calculate the “gap effect” and “unobserved price effect”, the most difficult thing is knowing the value of $\theta_{0f}\sigma_1$. Because θ_m is the standardized residual, its mean is zero. $\theta_{1f}\sigma_1$ and $\theta_{0f}\sigma_0$ can be calculated by using the female characteristics and the male coefficients of the corresponding period. To impute $\theta_{0f}\sigma_1$, first we need to know the female residual at period 0 on period 0’s male residual distribution percentile, then get the male residual of the corresponding percentile on period 1’s male residual distribution, that is, $\theta_{0f}\sigma_1$.

5. EMPIRICAL RESULTS

5.1 OLS Results

The coefficient of the gender dummy in the OLS represents the unexplained part of the gender earnings gap. Here, we add explanatory variables into the model step by step and focus on the change of the gender dummy coefficient, and Table 2 presents these results. Panel A of Table 2 reports the regression results with basic controls: cohort groups, *hukou* status, survey year dummy, and regional dummies. The coefficient of the gender dummy across birth cohorts is similar to the raw gender gap reported in Table 1. In panel B, we add college degree to the model. The unexplained gender earnings gap hardly changes across the 1960s and 1970s birth cohorts; however, it increases the coefficient of the gender dummy from 0.303 to 0.318, to a 0.015 log point increase in the 1980s birth cohort. This outcome implies that education is one driver to narrow the gender earnings gap, but it seems that the net effect is not so substantial compared to the women’s ‘revolution’ in education itself. In panels C and D, we add cognitive and noncognitive abilities consecutively. In the 1960s birth cohorts, controlling for these abilities reduces the coefficient of the gender dummy, but in the 1970s and 1980s birth cohorts, these extra controls increase the coefficient of the gender dummy. This is consistent with the progress in women’s cognitive and noncognitive abilities, but the contribution to narrowing the gender earnings gap is still not large. Panel E further controls for fertility and marital status. This does not change the gender gap in the 1960s birth cohorts, but reduces and increases the gender gap in the 1970s and 1980s birth cohorts, respectively, although the changes are trivial in both groups. The greatest change happens in Panel F, where we further control for the occupation, industry, and employer dummies. In the three birth cohorts, these additional controls reduce the coefficient of the gender dummy by 0.07, 0.05, and 0.03 log points. This suggests that there still are certain barriers for women entering high-earning jobs.

Next, we turn to the full results of the OLS regressions with all controls, which are presented in Table 3. *Hukou* has a premium in the earnings of females, but not males. Women with a nonagricultural *hukou* have more opportunities to choose different kinds of work than women with an agricultural *hukou*. For people with an agricultural *hukou*, the job opportunities they face are mostly heavy labor work which favors men. This outcome means that work opportunities for women holding an agricultural *hukou* are limited. In addition, the *hukou* premium becomes more important for the youngest cohort, which is a little surprising.

Table 2: OLS Results of Gender Dummy

Variables	1960s Cohort	1970s Cohort	1980s Cohort
Panel A			
Basic controls	0.432*** (0.032)	0.429*** (0.030)	0.303*** (0.032)
Panel B			
Further control for college degree	0.430*** (0.029)	0.421*** (0.030)	0.318*** (0.032)
Panel C			
Further control for cognitive abilities	0.422*** (0.030)	0.424*** (0.030)	0.321*** (0.032)
Panel D			
Further control for noncognitive abilities	0.420*** (0.031)	0.430*** (0.030)	0.324*** (0.032)
Panel E			
Further control for children and marriage	0.420*** (0.031)	0.422*** (0.029)	0.326*** (0.032)
Panel F			
Further control for occ., ind. and employer	0.349*** (0.030)	0.372*** (0.028)	0.296*** (0.030)
Cohort groups	Yes	Yes	Yes
2014 data	Yes	Yes	Yes
Regions	Yes	Yes	Yes
Observations	2,930	3,262	3,283

Notes: Robust standard errors in parentheses, which are clustered at the county level. *** p<0.01, ** p<0.05, * p<0.10.

As expected, education has a high association with earnings, especially in the 1960s birth cohort. Compared to those without a college degree, those with a college degree increase their earnings by about 50% on average in the 1960–1969 cohort. For the 1970s and 1980s birth cohorts, the return on a college degree is 35% and 25%, respectively. This decreasing return on college degrees may be due to the expansion of higher education and the resulting degree inflation. An interesting finding is that the return on college degrees is basically the same between genders in the 1960s birth cohort, but in the 1970s birth cohort, it is 10% higher for females than for males, and in the 1980s birth cohort, the return on college degrees becomes close between the genders again. Compared to men, more women have a college degree in the youngest cohort, which should be the main cause of the lowering premium of a college degree.

Compared to education, the effects of cognitive and noncognitive abilities, except for self-esteem, have no consistent pattern across cohorts. A few interesting finds include the fact that math test scores become more important for males in the 1980s birth cohort, but not for females. The return on self-esteem is always higher for men than for women, especially for the 1960s birth cohort, and popularity sometimes has a negative effect on earnings, which may be because of a measurement or sampling error.

Table 3: OLS Results of Full Specification

Variables	1960s Cohort			1970s Cohort			1980s Cohort		
	all	male	female	all	male	female	all	male	female
Gender	0.349*** (0.030)			0.372*** (0.028)			0.296*** (0.030)		
Hukou	0.044 (0.044)	−0.020 (0.056)	0.128*** (0.047)	0.053 (0.035)	−0.003 (0.049)	0.126*** (0.042)	0.042 (0.036)	−0.045 (0.041)	0.153*** (0.049)
College degree	0.486*** (0.065)	0.498*** (0.073)	0.509*** (0.094)	0.350*** (0.052)	0.308*** (0.060)	0.407*** (0.064)	0.251*** (0.034)	0.244*** (0.047)	0.235*** (0.047)
Standardized word score	0.012 (0.015)	−0.007 (0.017)	0.041 (0.027)	0.037** (0.016)	0.048*** (0.018)	0.017 (0.022)	0.025** (0.012)	0.025 (0.016)	0.030 (0.019)
Standardized math score	0.033* (0.018)	0.032* (0.018)	0.038 (0.032)	−0.012 (0.011)	−0.012 (0.015)	−0.005 (0.018)	0.031** (0.012)	0.044*** (0.016)	0.011 (0.018)
Emotional stability	0.001 (0.013)	−0.002 (0.016)	0.001 (0.018)	−0.008 (0.015)	0.006 (0.024)	−0.020 (0.020)	0.005 (0.013)	0.005 (0.020)	0.013 (0.018)
Self-esteem	0.050** (0.021)	0.085*** (0.019)	0.004 (0.036)	0.042*** (0.015)	0.049** (0.020)	0.030* (0.017)	0.046*** (0.014)	0.075*** (0.019)	0.020 (0.019)
Popularity	−0.009 (0.019)	−0.027 (0.026)	0.003 (0.019)	−0.002 (0.010)	0.032** (0.016)	−0.043** (0.017)	−0.005 (0.015)	−0.033* (0.020)	0.026 (0.022)
Have children	−0.049 (0.049)	−0.072 (0.057)	0.002 (0.081)	− 0.191*** (0.071)	−0.186** (0.084)	−0.203** (0.097)	−0.078* (0.043)	−0.050 (0.050)	− 0.116** (0.050)
Married	0.177 (0.133)	0.328** (0.153)	0.015 (0.240)	0.012 (0.091)	0.023 (0.117)	−0.140 (0.210)	0.091** (0.041)	0.147*** (0.054)	0.003 (0.059)
Constant	7.130*** (0.185)	7.374*** (0.222)	7.049*** (0.306)	7.364*** (0.173)	7.894*** (0.211)	7.290*** (0.313)	7.397*** (0.143)	7.715*** (0.182)	7.404*** (0.215)
Observations	2,930	1,749	1,181	3,262	1,837	1,425	3,283	1,823	1,460
R-squared	0.289	0.217	0.283	0.282	0.213	0.265	0.231	0.242	0.196

Notes: Robust standard errors in parentheses, which are clustered at county level. All regressions controls for cohort groups, occupation, industry, employer, region, and survey year dummies. *** p<0.01, ** p<0.05, * p<0.1.

Except for the female sample in the 1960s birth cohort, having at least one child leads to lower earnings, although this is often statistically insignificant. More importantly, the magnitude of the coefficient is larger for women than men in the 1970s and 1980s birth cohorts. Although this might not be a desirable result, we have to admit that fertility has a negative effect on women's careers. After the birth of a child, the mother will have a short break in employment, her job-oriented training decreases, and her productivity may decrease accordingly, leading to lower earnings (Blau and Kahn 2017). People with young children need to spend more time taking care of them, but older children can take care of themselves. In our sample, people born in the last two cohorts are more likely to have young children, while the first cohort tends to have older children, so the negative effect of fertility is larger in the last two cohorts. Marital status has a significant positive association with men's earnings in the 1960s and 1980s birth cohorts, but for women, it is insignificant in all three birth cohorts. After marriage, men pay more attention to their market outcome, acting as "breadwinners", and women spend more time on housework (Becker 1985). Given the imbalanced marriage ratio in the younger cohorts, these results may be because of a selection effect: more able men are more likely to be married, but it is hard to say why married men earn more than unmarried men on average in the 1960s birth cohorts.

5.2 Oaxaca-Blinder Decomposition Results

Table 4 presents the results of the Oaxaca-Blinder decomposition. As mentioned in the last section, we use pooled coefficients as the reference in the decomposition. Overall, the total gender earnings difference and the explained difference decrease across the cohorts, while the constant effect on earnings structure shows a rising trend to some extent. The latter accounts for 31.06%, 65.63%, and 62.12% of the total gender earnings differential across birth cohorts. This outcome is totally unexplained, and we can consider it as discrimination.

Table 4: Results of Oaxaca-Blinder Decomposition

Variables	1960s cohort		1970s cohort		1980s cohort	
<i>Gender Earnings Gap</i>	0.425***		0.416***		0.264***	
	<i>Explained</i>	<i>Unexplained</i>	<i>Explained</i>	<i>Unexplained</i>	<i>Explained</i>	<i>Unexplained</i>
<i>Total</i>	0.056**	0.349***	0.033**	0.372***	−0.025	0.296***
Cohort groups	−0.000	0.011*	−0.002	−0.056**	0.000	0.031
Hukou	−0.001	−0.088**	−0.001	−0.073**	−0.001	−0.094***
College degree or above	−0.003	−0.002	0.008	−0.026	−0.022***	0.005
Standardized word test score	0.001	0.011	−0.002	0.003	−0.004*	−0.001
Standardized math test score	0.005	0.001	−0.001	0.000	0.001	−0.000
Emotional stability	0.000	−0.000	−0.000	−0.000	0.000	0.000
Self-esteem	0.001	−0.003	−0.004*	0.000	−0.004*	0.006
Popularity	0.001	0.000	0.000	0.001	0.000	0.001
Have children	0.000	−0.069	0.008**	0.016	0.004	0.035
Married	0.000	0.312	−0.000	0.160	−0.006*	0.098*
Occupations	0.021*	0.046	0.023**	0.068**	−0.001	−0.011
Industries	0.048***	0.004	0.026***	−0.026*	0.020***	−0.020
Employers	0.004	0.045	0.000	−0.024	0.010*	0.028
2014 data	−0.003	−0.036	−0.007***	0.043*	−0.009*	0.059*
Regions	−0.018**	−0.016	−0.016*	0.011	−0.012*	−0.005
Constant	—	0.132	—	0.273	—	0.164**

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

For the 1960s birth cohort, the explained part is 0.056 log points, accounting for 13.18% of the gender earnings gap. The results show that *hukou*, education, cognitive abilities, noncognitive abilities, fertility, and marital status have no significant explanatory power for the explained gender earnings gap. The main contributors among the observed characteristics are occupations and industries. They explain 4.94% and 11.29% of the total gender earnings gap, and 37.5% and 85.7% of the explained gender earnings gap, respectively. Obviously, men tend to enter high-paying occupations and industries, which enlarges the gender earnings gap.

The unexplained part in the 1960s birth cohort is 0.349, which accounts for 82.12% of the total gender earnings gap. The return to *hukou* can significantly reduce the gender earnings gap because of its high premium for women, but in terms of social integration, the discrimination in the market based on *hukou* is not desirable, and the goal of the future should be to abolish *hukou*. Though not significant, the largest contribution to the unexplained gender gap is from marital status. It accounts for 73.41% of the total gap. The return on marriage especially favors men. Table 3 shows that a married man earns 32.8% more than an unmarried man on average in the 1960s birth cohort, which is equal to two or three years of college education. Given that we have controlled for age and experience, the premium of marriage is not because the married men are usually older and earn more. As we discussed for the OLS results, after marriage, there might be a

division of labor: men tend to focus on work, while women tend to concentrate on family. This specialization within a family makes for a totally different marriage premium between genders, and it also makes the gender earnings gap larger.

The 1970s birth cohort is somewhat different from the 1960s birth cohort. In the explained part, the main change happens to the occupations and industries. Compared to the 1960s birth cohort, the contribution of occupation to the explained gender gap does not change, but the unexplained part presents that within the same occupation, men are more likely to earn more than women. As Equation (6) shows, the unexplained part consists of $\bar{X}_m(b_m - b^*) + \bar{X}_f(b^* - b_f)$. Given the same occupation, the two parts are both bigger than 0 because of $b_m - b^* > 0$ and $b^* - b_f > 0$. For an industry, the contribution to the explained gender gap reduces from 0.048 to 0.026 log points, and it even contributes to reducing the unexplained gender gap by 0.026 log points. This outcome shows that not only were the barriers to the high-paying industries weakened for the 1970s birth cohort, but the earnings of women also increased relative to men within the same industry.

In the 1980s birth cohort, the explained part changes to -0.025 log points, but is not significant. Basically, this outcome means that in terms of observed characteristics, there is no difference between women and men. Compared to the previous cohorts, this lack of difference is largely based on better educational attainment, as well as better word test scores and self-esteem. It is worth noting that occupation no longer contributes to the gender earnings gap. This neither insignificant nor substantial value is caused by the convergence of differences in the share of managers and enhanced advantages for women in the share of professionals and clerical workers. For the unexplained part, the change of earnings structure not only counterbalances the explained part, it also covers the entire gender earnings gap. Cohort groups, fertility and marital status, and data source contribute the most. In the full specification OLS model, having at least one child will significantly decrease female earnings, while being married will significantly increase male earnings. The return on fertility and marital status are all disadvantageous for women, thus amplifying the gender earnings gap. The earnings structure effect of whether to have a child accounts for 13.26% of the total gender earnings gap. The parental leave policy aims to protect women who give birth to a child, but it may make women lose advantages in the job market. During this period, the human capital accumulation stops, job-oriented training decreases, and women will devote less time to work for the sake of raising a child after birth. These factors all contribute to a larger difference on the return of having a child, going against the reduction of the gender earnings gap.

5.3 JMP Decomposition Results

In Table 5, we estimate the contribution of each factor to the change in the gender earnings gap across birth cohorts. The gender earnings gap increases by 0.114 log points from the 1960s cohort to the 1970s cohort. The observed characteristics' effect is 0.003 log points, only accounting for 2.61% of the change in the total gender earnings gap. College degree explains the majority of the observed characteristics. It plays a role in pushing up the gender earnings gap. It contributes to 9.93% of the total differential and is much larger than the contributions of all other observed characteristics. The main reason for this is that the gender gap in the share of college degree changes from -1 percentage points in the 1960s birth cohort to 2 percentage points in the 1970s birth cohort, meaning that women lose their advantage in terms of college degrees. Occupation is the largest factor working to narrow the gender earnings gap and contributes to 8.45% of the gap change. This negative occupational

contribution to the gender earnings gap is caused by the share of women in low-paying business and service occupations declining, and the share of women in high-paying managerial occupations increasing (see Table 1).

Table 5: Results of JMP Decomposition

Variables	1970s–1960s		1980s–1970s		1980s–1960s	
	Value	Proportion	Value	Proportion	Value	Proportion
Changes in the total earnings gap	0.114		–0.194		–0.080	
Observed X's	0.003	2.61	–0.017	8.68	–0.017	21.35
Hukou	0.002	1.45	–0.004	1.92	–0.002	2.11
College and advanced degree	0.011	9.93	–0.026	13.59	–0.020	24.57
Standard word test score	–0.002	–1.80	–0.003	1.79	–0.007	9.05
Standard math test score	0.000	0.25	–0.001	0.54	–0.002	1.97
Emotional stability	0.002	1.46	0.000	0.05	–0.001	1.42
Self-esteem	–0.003	–2.86	0.000	–0.13	–0.002	2.37
Popularity	–0.001	–0.69	0.000	0.15	0.000	–0.26
Have children	0.008	6.77	0.001	–0.34	0.005	–5.90
Married	0.004	3.85	0.000	0.25	–0.001	1.12
Occupations	–0.010	–8.45	0.001	–0.76	0.000	–0.55
Industries	–0.008	–6.72	–0.001	0.68	–0.004	5.11
Employers	–0.004	–3.43	0.008	–3.99	0.010	–12.74
2014 data	–0.002	–1.77	0.002	–0.85	–0.002	2.32
Regions	0.005	4.62	0.008	–4.23	0.007	–9.25
Observed prices	–0.004	–3.69	–0.025	13.14	–0.027	33.16
Hukou	0.000	0.05	0.000	0.14	–0.001	0.74
College and advanced degree	0.001	0.56	–0.004	1.81	0.002	–2.15
Standard word test score	–0.002	–1.43	–0.001	0.36	–0.001	0.78
Standard math test score	–0.007	–6.07	0.002	–1.01	–0.004	5.21
Emotional stability	–0.003	–2.47	0.002	–0.86	0.002	–1.96
Self-esteem	0.000	0.40	0.001	–0.49	0.000	–0.35
Popularity	0.004	3.70	–0.005	2.57	–0.002	2.57
Have children	0.001	1.00	–0.004	2.18	0.001	–0.72
Married	0.000	–0.13	–0.005	2.40	0.000	0.02
Occupations	0.032	28.44	–0.016	8.45	0.007	–9.34
Industries	–0.028	–24.31	0.014	–7.41	–0.018	22.86
Employers	–0.012	–10.77	–0.002	1.02	–0.021	25.79
2014 data	0.001	0.93	–0.004	1.93	–0.001	1.51
Regions	0.007	6.41	–0.004	2.04	0.009	–11.80
Gap effect	0.126	110.78	–0.194	99.82	–0.042	52.51
Unobserved price	–0.011	–9.70	0.042	–21.64	0.006	–7.02

As for the effect of observed price, although the total effect is to reduce the gender earnings gap, the magnitude is very small. The observed price of occupation is the largest factor contributing to the increase of the gender earnings gap, and it is even larger than the effect of the total observed price effect. As shown in Equation (8), the observed price effect is $\Delta X_0(b_1 - b_0)$. If $\Delta X_0 > 0$, and the occupation is male-dominated in 1960s. If $\Delta X_0 < 0$, it is female-dominated. The positive value of occupation means either the returns on male-dominated occupations increase or the returns on female-dominated occupations decrease, thus widening the gender earnings difference. Industry and employer factors can offset the entire effect of the return on occupation. Similarly, the negative effects of industry and employer factors mean that given the distribution of the two characteristics in 1960s, either the returns decrease in the

male-dominated characteristics or the returns increase in the female-dominated characteristics.

The gap effect is 0.126 log points, accounting for 110.78% of the increase of the gender earnings differential in the first two cohorts. The gap effect represents discrimination and unobserved skills. The discrimination may result from the different treatment of the genders in the same position, such as men being more valued than women, men being given more opportunities to accumulate working skills and promotions, or other factors, which widens the gender earnings gap. The cognitive abilities in this paper may not be good measurements, resulting in the high amount of unobserved skills in the gap effect. For example, given the development of technology, computer skills may help individuals get jobs or higher pay. Meanwhile, computer skills are cognitive instead of manual and may improve women's performance in the labor market (Beaudry and Lewis 2014). The unobserved price effect is another part that can narrow the gender earnings gap. However, the value is small.

From the 1970s cohort to the 1980s cohort, the gender earnings differential decreases to 0.194 log points. The observed X 's effect works to reduce the gender earnings gap by 0.017 log points. A college degree contributes to almost all of the observed X 's effects, accounting for 13.59% of the total gender earnings gap. The PRC started higher education enrolment expansion in 1999. As a result, the 1980s cohort had more opportunities to enter college, and men and women both increased their share of getting a college degree, but women increased at a higher speed because they obtained a higher return on their college degree. In addition to college degrees, other variables in the observed X 's effect contribute little to the gender earnings gap. The observed price effect also decreases the gender earnings difference, but it is less than twice of the observed X 's effect, accounting for 13.14% of the total gap. The value of industries can counterbalance more than half of the observed price effect.

Overall, from the 1960s cohort to 1980s cohort, the gender earnings gap narrows by 0.080 log points. All effects, except the unobserved price effect, contribute to the reduction of the earnings gap. The observed characteristics' effects account for 21.35% of the total decrease. The education variable plays the most important role in this part, explaining 24.57% of the reduction of the gender earnings gap. Women gained ground in college education across the cohorts, and the high level of human capital they accumulated helped them to get a pay rise. As women show a reverse in their share of college degrees, and the return on education is higher for women, encouraging women to increase human capital investment is an effective way to narrow the gender earnings gap. Nevertheless, the effect is limited, as the return on education decreases with the rise in college degree share.

6. SUMMARY AND CONCLUSION

We use the CFPS to investigate the effect of education and cognitive/noncognitive abilities on the gender earnings gap. We examine the relation within and across birth cohorts using OLS, Oaxaca-Blinder, and JMP decompositions.

Overall, although women's human capital has experienced a "quiet revolution," the gender earnings gap is still quite large compared to the substantial progress of women in terms of human capital. The gender earnings gap is decreasing with the birth cohorts, and the return on education decreases as the younger cohort obtains higher quantities of education. As women exceed men in education and have higher returns on education relative to men, education will continue to narrow the gender earnings gap, despite the restricted effect. However, we also need to note that given that women's education has

been – on average – beyond men's in the PRC, we should expect that women's progress in education relative to men should slow, and in this case, improving women's human capital might not be a primary or powerful driver for narrowing the gender earnings gap. Particularly, the JMP decomposition results show that the gap effect makes the greatest contribution to the change in the gender earnings gap, which is attributed to the unobserved skills related to the observed skills level. The explanation of the gap effect is somewhat monotonous, and more research should be done to investigate the content of the gap effect.

These results have the following policy implications. First, governments and schools can continue to reduce the discrimination against women in education, for example, by combating the stereotype that women are bad at math. Second, it is important to develop cognitive abilities, as well as noncognitive abilities, as the return on noncognitive abilities, such as self-esteem, is positive. Third, more research on regulations and labor market discrimination against women should be encouraged in order to understand more about the gender earnings gap and its evolution.

REFERENCES

- Afridi, F., Li, S. X., and Ren, Y. (2015). Social identity and inequality: The impact of China's hukou system. *Journal of Public Economics*, 123, 17–29.
- Beaudry, P., and Lewis, E. (2014). Do male-female earnings differentials reflect differences in the return to skill? Cross-city evidence from 1980–2000. *American Economic Journal: Applied Economics*, 6(2), 178–94.
- Becker, G. S. (1985). Human capital, effort, and the sexual division of labor. *Journal of Labor Economics*, 3(1, Part 2), S33–S58.
- Becker, G. S., Hubbard, W. H., and Murphy, K. M. (2010). Explaining the worldwide boom in higher education of women. *Journal of Human Capital*, 4(3), 203–241.
- Blau, F. D., and Kahn, L. M. (1997). Swimming upstream: trends in the gender earnings differential in the 1980s. *Journal of Labor Economics*, 15(1), 1–42.
- . (2008). Women's work and earnings. In: S. N. Durlauf and L. E. Blume, *The New Palgrave Dictionary of Economics*, 2nd ed. (pp. 762–772). London: Palgrave Macmillan.
- . (2017). The gender earnings gap: extent, trends, and explanations. *Journal of Economic Literature*, 55(3): 789–865.
- Blinder, A. S. (1973). Earnings discrimination: reduced form and structural estimates. *Journal of Human Resources*, 8(4), 436–455.
- Borghans, L., Ter Weel, B., and Weinberg, B. A. (2014). People skills and the labor-market outcomes of underrepresented groups. *ILR Review*, 67(2), 287–334.
- Buchmann, C., and DiPrete, T. A. (2006). The growing female advantage in college completion: The role of family background and academic achievement. *American Sociological Review*, 71(4), 515–541.
- Ceci, S. J., Ginther, D. K., Kahn, S., and Williams, W. M. (2014). Women in academic science: A changing landscape. *Psychological Science in the Public Interest*, 15(3), 75–141.
- Drago, F. (2011). Self-esteem and earnings. *Journal of Economic Psychology*, 32(3), 480–488.
- Fortin, N. M. (2008). The gender earnings gap among young adults in the united states: The importance of money versus people. *Journal of Human Resources*, 43(4), 884–918.
- Goldin, C. (2006). The quiet revolution that transformed women's employment, education, and family. *American Economic Review*, 96(2), 1–21.
- Goldin, C., Katz, L. F., and Kuziemko, I. (2006). The homecoming of American college women: The reversal of the college gender gap. *Journal of Economic Perspectives*, 20(4), 133–156.
- Goldin, C. (2014). A grand gender convergence: Its last chapter. *American Economic Review*, 104(4), 1091–1119.
- Gustafsson, B., and Li, S. (2000). Economic transformation and the gender earnings gap in urban China. *Journal of Population Economics*, 13(2), 305–329.

- Heckman, J. J., Stixrud, J., and Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics*, 24(3), 411–482.
- Jann, B. (2008). The Blinder-Oaxaca decomposition for linear regression models. *The Stata Journal*, 8(4), 453–479.
- Juhn, C., Murphy, K. M., and Pierce, B. (1991). Accounting for the slowdown in black-white earnings convergence. In: Marvin H. Koster, *Workers and Their Earnings*, (pp 107–43). Washington, DC: AEI Press.
- Judge, T. A., Higgins, C. A., Thoresen, C. J., and Barrick, M. R. (1999). The big five personality traits, general mental ability, and career success across the life span. *Personnel Psychology*, 52(3), 621–652.
- Kuhn, P., and Weinberger, C. (2005). Leadership skills and earnings. *Journal of Labor Economics*, 23(3), 395–436.
- Lee, M. H. (2012). The one-child policy and gender equality in education in China: Evidence from household data. *Journal of Family and Economic Issues*, 33(1), 41–52.
- Li, S., Song, J., and Liu, X. (2014). The evolution of the gender earnings gap of the staff of China's cities and towns. *Management World*, 03, 53–66.
- Liu, H. (2011). Economic reforms and gender inequality in urban China. *Economic Development and Cultural Change*, 59(4), 839–876.
- Neumark, D. (1988). Employers' discriminatory behavior and the estimation of earnings discrimination. *Journal of Human Resources*, 23(3), 279–295.
- Nyhus, E. K., and Pons, E. (2005). The effects of personality on earnings. *Journal of Economic Psychology*, 26(3), 363–384.
- Oaxaca, R. (1973). Male-female earnings differentials in urban labor markets. *International Economic Review*, 14(3), 693–709.
- Pope, D. G., and Sydnor, J. R. (2010). Geographic variation in the gender differences in test scores. *Journal of Economic Perspectives*, 24(2), 95–108.
- Tsui, M., and Rich, L. (2002). The only child and educational opportunity for girls in urban China. *Gender and Society*, 16(1), 74–92.
- Zhang, J., Han, J., Liu, P. W., and Zhao, Y. (2008). Trends in the gender earnings differential in urban China, 1988–2004. *Industrial and Labor Relations Review*, 61(2), 224–243.

APPENDIX

Table A1: The Relation of Education and Earnings

Variables	1960–1969 Cohort			1970–1979 Cohort			1980–1989 Cohort		
	All	Male	Female	All	Male	Female	All	Male	Female
Primary school	0.036 (0.071)	−0.199* (0.116)	0.242** (0.089) *	−0.011 (0.056)	0.078 (0.087)	— (0.067) 0.135**	0.079 (0.091)	0.040 (0.137)	0.115 (0.118)
Secondary school	0.006 (0.066)	−0.193* (0.105)	0.154** (0.066)	−0.070 (0.057)	−0.010 (0.086)	— (0.068) 0.156**	0.091 (0.091)	0.050 (0.139)	0.127 (0.110)
Senior high school	0.103 (0.069)	−0.154 (0.113)	0.369** (0.078) *	−0.030 (0.063)	−0.006 (0.094)	−0.063 (0.082)	0.077 (0.090)	−0.016 (0.135)	0.191* (0.109)
College and above	0.576** (0.074) *	0.278** (0.114)	0.895** (0.085) *	0.297** (0.059) *	0.243** (0.090) *	0.371*** (0.073)	0.338** (0.093) *	0.237* (0.140)	0.445** (0.112) *
Gender	0.426** (0.030) *			0.422** (0.030) *			0.318** (0.032) *		
2014 data	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Regions	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	6.870** (0.092) *	7.523** (0.113) *	6.630** (0.101) *	7.016** (0.084) *	7.402** (0.112) *	7.059*** (0.090)	6.973** (0.112) *	7.338** (0.158) *	6.918** (0.124) *
Observations	2,930	1,749	1,181	3,262	1,837	1,425	3,283	1,823	1,460
R-squared	0.214	0.118	0.210	0.217	0.119	0.190	0.167	0.144	0.132

Notes: Robust standard errors in parentheses, which are clustered at the county level. *** p<0.01, ** p<0.05, * p<0.1.

Table A2: Industry Classification

Industry	Mean	Std. Dev.	Obs
Industry 1			
Scientific research, technical services, and geological prospecting	8.031	0.552	29
Finance	7.752	0.691	228
Construction	7.689	0.633	881
Industry 2			
Transport, storage, and post	7.664	0.746	656
Information, computer service, and software	7.661	0.781	88
Leasing and business services	7.605	0.729	192
Industry 3			
Culture, sports, and entertainment	7.538	0.632	98
Real estate	7.524	0.546	138
Education	7.505	0.586	652
Mining	7.501	0.586	310
Industry 4			
Manufacturing	7.465	0.573	3,161
Health, social security, and welfare	7.444	0.646	298
Industry 5			
Wholesale and retail trade	7.436	0.631	833
Other industries	7.375	0.543	67
Supply of electric power, gas, and water	7.372	0.537	218
Industry 6			
Hotels and catering services	7.300	0.619	415

Resident services and other	7.291	0.701	369
Public management and social organization	7.267	0.692	765
Water, environmental protection, and public facility management	7.236	0.555	77