

The long shadow of overeducation: Parents' overeducation and child development

Abstract: Research has shown that overeducation has negative effects on individuals in terms of wage and job satisfaction. In this paper, we study the intergenerational implications of overeducation via child development, and its possible mechanisms. Using data from a nationwide Chinese survey and employing an instrumental variable approach, we find that the children whose parents are overeducated have lower educational attainment and cognitive test scores. And the finding appears to be attributable to the rising frustration and lower earnings caused by overeducation, which may well give rise to a worse parenting style, and thus translate the indifferent attitudes towards education onto the next generation.

Keywords: Overeducation, Intergenerational effects, Children development, Wage penalty, Parenting style

1. Introduction

As global education expands, there is an oversupply of well-educated graduates and an undersupply of adequate high-skill jobs which would lead to a situation in which not all graduates can fully utilize their education qualifications and are forced to be 'filtered down' into positions for which counterparts with worse educational background would suffice, i.e., some workers are overeducated. Convincing evidence has indicated that overeducation of today's generation causes wage penalty (Zheng et al., 2021), frustration (Mateos-Romero & Salinas-Jiménez, 2018), worse career advancement opportunities (Jiang & Guo, 2022) and even lead graduates to regret their current education levels (Kucel & Vilalta-Bufi, 2013). Yet, the impact of overeducated parents on their children's human capital development is ambiguous. If parents are stringently restricted by their education level that is not enough and to be likely to have a higher demand for the next generation's quality. But there is another possibility that overeducated parents are awaked to the unnecessary of blindly pursuing higher education and no longer attach importance to education investment, the last generation's overeducation may thus have negative consequences for their children fare. The answer to this question is essential for society as a whole, given the growing rate of overeducation and the effects of parents' overeducation are bound to extend to children in the household.

In this article, we aim at examining a potential intergenerational effect of overeducation. Using six waves of China Family Panel Studies (hereafter CFPS) 2010-2020, our paper mainly investigates the impact of parents' overeducation on the human capital outcomes of their children, in terms of children's enrollment rate, on-time grade possession and cognitive abilities. We first study the impact of both parental overeducation and maternal overeducation on their children, and compare the different effects of father and mother in children's human capital development. Then, we endeavor to identify the various possible channels through which the overeducation effect of parents transmits to the next generation, including family resources, a shift in parenting behavior and the attitudes towards education instilled by parents. And we also examine the heterogeneous effects of parents' overeducation for different types of hukou groups and different household economic condition groups.

The key empirical challenge is whether workers are overeducated is obviously subject to many unobservable variables. Such as more capable people may have a higher probability of finding a matching job, or someone's overeducation is a deliberate choice that prioritizes family over career. Our paper addresses this challenge by adopting an instrumental variable that leverages variation in the composition of highly educated labor supply that labor-market entrants face while controlling province's specific linear time trend of the graduate year. We find the more highly educated graduates workers compete with when a worker first enters the labor market, the worker is more likely to be overeducated. And the estimated effect is substantial, with an elasticity between the probability of overeducation and the employment competition status equal to 0.159. We further show when and where workers enter the labor market are hard to predict in advance, and demonstrate that the instrumental variable is exogenous. Some macroeconomic cycle changes are also kept in our specification.

Our IV results show that children whose father is overeducated may have a lower enrollment rate, are harder to get into school on time and have lower cognitive skills (including verbal test scores and math test scores) than the children whose father is non-overeducated. In contrast, maternal overeducation is only associated with a significant negative effect on children's primary school enrollment and on-time grade possession. This is may because men generally experience mental depression more severely due to a bad employment status than women (McKee-Ryan et al., 2005; Kuhn et al., 2009).

Aiming to understand the mechanisms behind the transmission of parents' overeducation to children's development, we first examine the income effect which is crucial for education investment in children. And we find that the overeducation of

parents leads to a drop in income, so that parents find it difficult to allocate extra resources to their children's education. Then considering that overeducation causes lower earnings, this may give rise to lower returns to education and make parents no longer attach importance to education. We test the channel of parenting style and find that parents who are overeducated are likely to decrease their time on educational tutoring, care less about children's schooling and take negative actions in response to children's lower school grades than those who are not overeducated. Finally, we test whether there has been a real shift in parents' perception of the importance of education, and whether similar attitudes have been instilled in their children, i.e., inter-generation transmission of attitudes toward education. Results show that overeducated parents have lower educational expectations for children and their children tend to inherit this attitude in their future life. We believe all are important channels to explain why parents' overeducation has a negative effect on children's human capital development and they may also reinforce each other's effects.

As additional evidence consistent with our channel, our study shows the heterogeneous effects of parents' overeducation on children's outcomes for different types of hukou groups and different household economic condition groups. The divergence of family hukou type and household income both reflect differences in the overall skill level of the household. Results suggest that children who live in a rural type of hukou and lower-income families perform worse when their parents are overeducated. This can be more plausibly explained in conjunction with our estimate in the mechanism test that privileged families attach greater importance to the quality of their children and it's hard to shake their attitudes toward education, so they may react less drastically to the lower return of human capital caused by increased overeducated levels.

Our study mainly contributes to three strands of literature. Firstly, our work relates to the literature that studies the impact of overeducation. Existing analyses have mainly emphasized the negative impact of overeducation on workers, such as wage penalty (Zheng et al., 2021), lower job satisfaction (Mateos-Romero & Salinas-Jiménez, 2018), cognitive decline (De Grip et al., 2008) and lower probability of promotion (Jiang & Guo, 2022). And for the firm level, there are some studies show that overeducated workers may generate a knowledge spillover effect because they have higher education qualifications than their counterparts (Battu et al., 2003). However, these positive outcomes are outweighed by the consequences of employee withdrawal, such as withholding effort at work and absenteeism, so employing overeducated workers can eventually reduce productivity (Tsang, 1987; Belfield, 2010). This paper turns the research perspective to family members of workers, and adds a new dimension to the

existing literature. Specifically, we are interested in studying whether overeducated parents transmit the negative effect of overeducation to their children, and unraveling a possible intergenerational mechanism of transmission of overeducation.

Secondly, our findings contribute to the growing empirical literatures that focus on the impact of parental career experience on the next generation's development. Previous studies have mainly focused on the effect of parents' displacement, and show that parental job loss significantly reduces children's career earnings, school enrollment, college quality (Oreopoulos et al., 2008; Hilger, 2016), and negatively affects adolescents' physical and mental health (Sleskova et al., 2006; Liu & Zhao, 2014). Overeducation and unemployment are both adverse conditions for workers in the labor market, however as an emerging but increasingly visible employment phenomenon, the former means more of a waste in educational investment (Zheng et al., 2021). Our work explores parents' responses when facing mismatching jobs in terms of income, parenting style and attitudes toward education, which advances our understanding of the status quo of today's generation and its possible future implications on children.

Thirdly, the results of this paper could inform about the unintended consequences of education reform. A vast literature has mainly focused on the benefits of this policy, for example, the positive effects on employment, entrepreneurship, health, and research capabilities (Che & Zhang, 2018; Fu et al., 2022; Piracha et al., 2022; Qin & Kong, 2021). Our finding that the depreciation of human capital suffered by parents can pass on the negative effects to the next generation by affecting the education of their children suggests there exists a loss of college enrollment expansion policy for China's economic and social developments from a long run perspective, which has not been considered before. Combined with the literature which found expansion has exacerbated a large pre-existing urban-rural gap in educational attainment and earnings (Meng et al., 2013; Huang et al., 2022), and tends to reduce relative wages, to raise the unemployment rate, and to increase the difficulties in finding matching jobs (Knight et al., 2017; Xing et al., 2018). Policymakers should therefore be more cautious with further expansions of education.

This paper proceeds as follows. Section 2 introduces the theory and channels that link parents' overeducation and children's human capital development. Section 3 describes the data, sample selection and main variables of interest. The identification strategy is specified in Section 4. Section 5 presents the results and various robustness checks. Section 6 discusses potential mechanisms. Section 7 shows the heterogeneous test. Section 8 concludes.

2. Theory and channels

At the end of the last century, to meet the needs of people to receive higher education and fuel long-term economic growth with talents, the Chinese government greatly expanded the scale of higher education since 1999, which significantly increased the supply of college graduates. Between 1998 and 1999, the increase in college enrollments reached 47.3%, and about 1.597 million students attended university. In the following 20 years, the number of higher education admissions in China is increasing substantially year by year. By 2020, the number of students enrolled in higher education up to 9.675 million.

As the proportion of graduates with high education qualifications jumped, however, the industry had not upgraded or expanded as fast as the unprecedented growth of better-educated population. Facing an excess of well-educated graduates in the labor market, the employer tends to raise the requirement to an unnecessary level to pick up a small number of employees among considerable applicants. For example, for a position that can be qualified by a bachelor's degree graduate, the employer tends to set the recruitment criteria for a master's degree to reduce the number of qualified candidates. Therefore, some highly educated graduates are forced into lower-skill and lower-income jobs due to a lack of adequate job positions, and they will find that their acquired years of schooling significantly exceed the levels at which their counterparts were once employed. Such incongruence reflects a disequilibrium in the labor market and inefficiency in human capital utilization, and that is called overeducation.

The overeducation in China may be a composite effect relying on multiple mechanisms. See Sicherman (1991) on career mobility theory or Lazear (1977) on personal preference theory, we can also interpret overeducation as a strategy to accumulate work experience in positions that require less education in order to seek future promotion opportunities, or as a voluntary behaviour for people who do not lie in returns to education and can derive utility from pursuing education. While both theories might apply to specific individuals, overeducation in China is more likely to be driven by policy-orientated shocks in both the supply and demand side at large (Zheng et al., 2021). Influenced by traditional culture, most Chinese people are often keen on investment in education, as a signal of their stronger ability, so that they can obtain a high-income job and gain prestige (Heckman et al., 2018; Chen et al., 2021). However, for overeducated workers, they have come to realize that the return to education is less than what they expected and begin to consider whether higher education is necessary or redundant, so finally no longer blindly pursue high education qualifications for their children.

The prominent and important role of family background in the process of human capital formation has been emphasized in the literature (Blau & Duncan, 1967). Parents exert their significant influence mainly in terms of parenting behaviour and family environment. Combining the above analysis, overeducated workers may attach little importance to child education, then are more likely to have neglecting parenting style - lack of ability to supervise children's study, input less time in tutoring, and lower expectations for children's education. And considering intergenerational correlation in preferences (Dohmen et al., 2012; Giménez-Nadal et al., 2019; Zumbuehl et al., 2021; Brenøe & Epper, 2022), such attitude of not valuing education may propagate from parents to the next generation. And children's expectations of their education attainment may also decline, which leads to a negative effect on human capital development. Moreover, the income effect also has an extremely direct negative impact on children's education, except for the indirect effect through parents' attitudes. Overeducation may result in wage penalty, hence tighten the monetary budget constraint, and affect the inputs of the child human capital production function.

We attribute above effects to three mechanisms, that is income effect, parenting style, and intergenerational transmission of education attitudes. And we believe these three mechanisms could reinforce each other's effects. More precisely, the decline in education investment caused by falling income may be amplified by the change in parenting style and children's negative attitudes towards learning. Meanwhile, the drop for wages also strengthens parents' belief that over-invest leads to a waste of human capital, which makes them pay less attention to their children's education, leading to a continuous chain reaction.

3. Data, Sample selection and variables

3.1 Data

In this paper, we use data from the six waves of the China Family Panel Studies (CFPS), conducted by the Institute of Social Science Survey of Peking University. The CFPS is a large-scale, nationally representative, longitudinal study on Chinese family and society, which collects data at the individual, household, and county levels and provides comprehensive information on education, migration, income, assets, psychology, family dynamics, and so forth. The national baseline survey launched in 2010 contains approximately 14,797 households in 25 provinces. Five waves of full-sample follow-up surveys were conducted in 2012, 2014, 2016, 2018 and 2020, and the survey covered 29 provinces in 2014 and 31 provinces after 2016.

In addition, we obtain annual provincial-level characteristic data from China City Statistical Yearbook 2010-2020 and the annual number of college graduates in each province from Educational Statistics Yearbook of China 1960-2010.

3.2 Sample selection

In our analysis, we focus on the sample of adolescents aged 6–15 years with at least one parent who is employed and has occupational information. We concentrate on children whose parents are under 60 years old so that their parents have not reached standard retirement age. And we delete children if their demographic information or family background is missing or if their outcome variables are all missing. The full sample contains 16,173 child-father observations and 14,133 child-mother observations. The sample sizes for the regression analysis are different for different outcomes, depending on the data availability.

3.3 Variables

3.3.1 Overeducation

The dependent variable is a binary variable indicating whether a person is overeducated, i.e. whether a worker's years of education exceed the education level required for his/her job. Regarding measures of overeducation, there are three methods commonly used in the literature: (i) the Realized Matches (RM) method, which is based on the distribution of existing workers' education levels within each occupation (Clogg & Shockey, 1984), such as using the mean together with standard deviation (Verdugo & Verdugo, 1989). (ii) the Job Analysis (JA) method, which evaluates the required level of education for a certain occupation on the basis of an evaluation by professional job analysts (Rumberger, 1981). (iii) the Self-assessment (SA) method, which asks employees whether their current job matches their education level, or asks them about education levels required for their current job (Duncan & Hoffman, 1981).

For the JA method, this approach has the advantage of being authoritative but is very costly to implement and is easy to use outdated evaluation criteria (Sicherman, 1991). The SA method has the advantage of providing measures that are precisely tailored to the respondent's job and are up to date, but it may be vulnerable to measurement errors since subjective reports can vary across respondents who do the same job (Flisi et al., 2017). And sample who answer the question of “the education level required for the job” only comprise 20.37% of the full survey sample who have a job, which is not representative for all workers in the labor market.

Combined with the analysis above, we use the RM method to measure overeducation of workers for core regression estimation. The RM method, objective as it is in assessing requirement education levels, uses the distribution of education levels of counterparts in the same occupation now. We define someone to be overeducated if his or her education level is more than one standard deviation above the required educational level within his or her occupation, otherwise, they are not¹. The occupational classification in CFPS data is based on the occupational code system of CSCO09. Its finest classification is at the 5-digit level and distinguishes more than 455 job categories². We take the first three digits as a standard for classifying occupations and get 72 occupation categories. We take the average education level of the existing labor force over each combination of occupation and province as an individual's required educational level in each survey year³. It is necessary to calculate the required education levels in each industry by province because regional economic development varies greatly in China. For example, applying for a junior high school teacher in Ningxia province (an economically backward province in China) may only require a bachelor's degree, but in Beijing, a bachelor's degree is far from enough. Admittedly, matching by occupation, province and survey year could be less than perfect, compared to considering variables like the job title which is not available. As a robustness check, we also use years of overeducation to measure parents' level of overeducation.

3.3.2 Children outcomes

We mainly use two measures for adolescent development, namely, school enrollment and cognitive ability.

School enrollment is an important measure of child quality and schooling attainment in developing countries, as has been suggested in the literature (Li et al., 2008; Glewwe et

¹ While under-education is also a part of the issue of occupational mismatch, we only focus on overeducation and serve workers who are undereducated as matched in this paper, which helps to simplify our analysis. Furthermore, the finding showed by Hartog (2000) further justifies the reason of putting them together, that undereducated workers are comparable to matched people in terms of wages and job satisfaction.

² The reason why we do not use a combination of jobs, provinces, and survey years to divide the group is that the sample size within each group divided by this criterion is so small, so that the mean of education years is not representative.

³ We calculate the requirements of years of education within each group by occupation, province and survey year. The mean of the sample size of groups is 20.43, minimum is 1 and maximum is 2033. For groups with a sample size of less than 10, we use entire samples within the geographical area (i.e., North China, Northeast China, East China, Central China, South China, Northwest China, Southwest China) in which the province is located to calculate the requirements.

al., 2001)⁴. We construct three binary variables indicating the enrollment status of a child. We look at primary school enrollment for all children aged 6 – 15 and junior high school enrollment for children aged 12 – 15. One possible scenario is that children's educational progress is postponed rather than interrupted because of parents' worsening employment status, that is the school age at a certain stage may be significantly higher than the age of normal school children. We also construct a dummy indicating whether a child is on-time grade progression, which takes a value of 1 if a child is at or above the proper grade for his/her age and 0 otherwise.

The cognitive abilities are measured by two scores from the vocabulary test and math test administered among children aged 10 or older in the CFPS. The CFPS asks respondents to take a set of Chinese language (including the application of word and word recall tests) and math (including arithmetic and number series tests) tests to determine respondents' literacy, reading, memory, and arithmetic abilities. There are 34 verbal questions and 24 standardized mathematics questions in CFPS 2010 and 2014. We obtain a verbal test score, from 0 to 34, and a math test score, from 0 to 24, that reflect an individual's vocabulary and mathematical ability, respectively. To allow for comparisons of children of different ages, we standardize the word and math test scores to have a mean of 0 and a standard deviation of 1 by survey year and age of children. And we calculate the average score as the cognitive score of children.

3.4 Descriptive statistics

Table 1 summarizes key variables used in the analysis. On average, fathers receive 8.59 years of education, and mothers receive 7.55 years of education. Out of the 16,173 fathers, 2,264 are overeducated. By contract, out of the 14,133 mothers, 1,131 are overeducated. Overall, fathers are more likely to be overeducated than mothers.

Table 1 Summary Statistics.

	(1) count	(2) mean	(3) sd	(4) min	(5) max
<i>Panel A: Child-father sample</i>					
Primary school enrollment	16173	0.84	0.36	0.00	1.00
Junior high school enrollment	6088	0.63	0.48	0.00	1.00
On-time grade progression	16173	0.71	0.46	0.00	1.00
Cognition-word	7126	0.02	0.99	-3.68	2.54
Cognition-math	7043	0.01	0.98	-3.69	2.99
Cognition	6985	0.03	1.05	-6.06	4.12

⁴ Although government has promulgated compulsory education law and set minimum age for employment to force children into school, dropout during compulsory education remains severe in China (Cui et al., 2019).

Father's overeducation(Over-educated = 1)	16173	0.14	0.34	0.00	1.00
Father's employment competition status	16173	0.07	0.08	0.00	0.86
Age	16173	10.32	2.89	0.00	15.00
Gender(male = 1)	16173	0.53	0.50	0.00	1.00
Hukou(rural hukou = 1)	16173	0.94	0.82	0.00	7.00
Father's age	16173	38.85	5.74	22.00	60.00
Father's hukou(rural hukou = 1)	16173	0.84	0.63	0.00	7.00
Father's years of education	16173	8.59	3.86	0.00	22.00
Number of siblings	16173	1.07	0.93	0.00	8.00
Ln(Family income)	16173	10.62	0.92	5.99	13.12
Ln(Family expenditure)	16173	10.28	0.90	7.09	12.77
Panel B: Child-mother sample					
Primary school enrollment	14133	0.85	0.36	0.00	1.00
Junior high school enrollment	5460	0.63	0.48	0.00	1.00
On-time grade progression	14133	0.71	0.45	0.00	1.00
Cognition-word	6323	0.01	0.99	-3.68	2.54
Cognition-math	6287	0.01	0.99	-3.69	2.99
Cognition	6232	0.02	1.07	-6.06	4.12
Mother's overeducation(Over-educated = 1)	14133	0.08	0.27	0.00	1.00
Mother's employment competition status	14133	0.07	0.09	0.00	0.86
Age	14133	10.38	2.89	6.00	15.00
Gender(male = 1)	14133	0.54	0.50	0.00	1.00
Hukou(rural hukou = 1)	14133	0.91	0.67	0.00	7.00
Mother's age	14133	37.09	5.62	22.00	60.00
Mother's hukou(rural hukou = 1)	14133	0.82	0.38	0.00	1.00
Mother's years of education	14133	7.55	4.42	0.00	22.00
Number of siblings	14133	1.07	0.93	0.00	8.00
Ln(Family income)	14133	10.63	0.93	5.99	13.12
Ln(Family expenditure)	14133	10.28	0.90	7.09	12.77

Next, we conduct rich summary data analysis at the individual level. But what should be highlighted here is that the descriptive results of the overeducation level shown below shall not be deemed as the overeducation level in China's overall labor market, it only represents the overeducation level of the samples that we impose restriction. As shown in Figure 1, the two peaks correspond to 0–1 years of overeducation and 3–4 years of overeducation, respectively. Two thirds of the population have more than zero years of overeducation, which implies that overeducation is prevalent among workers. Figure 2 illustrates the incidence of overeducation for workers entering the labor market at different times, and we find that the incidence of overeducation has been rising with the delay in entering the labor market. This is roughly in line with our analysis above that the job market for young labor-market entrants is gradually deteriorating as a result of the oversupply of labor caused by higher education expansion.

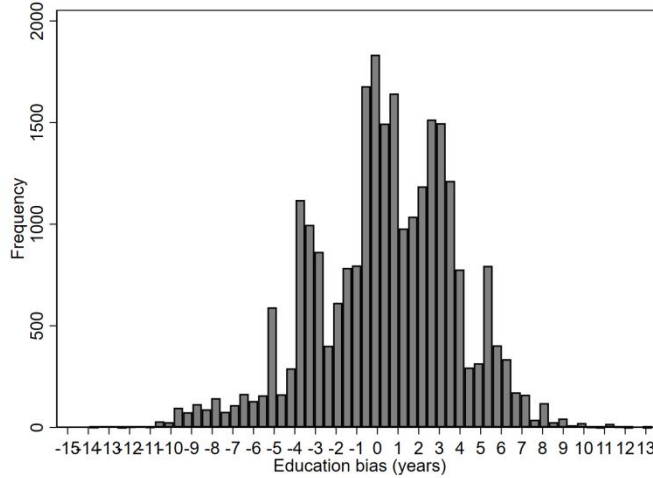


Fig. 1 Distribution of excess years of schooling.

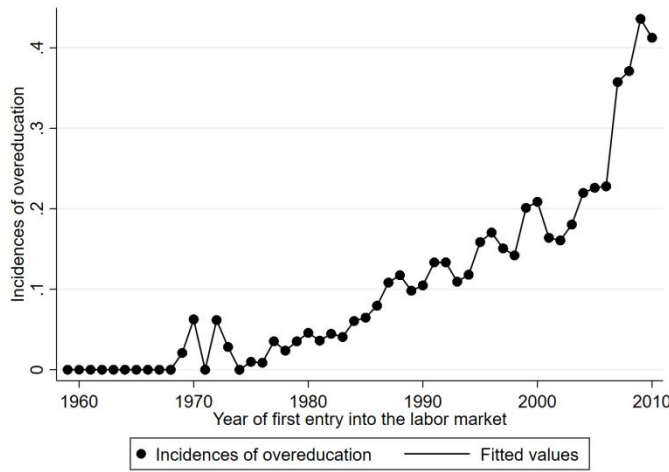


Fig. 2 Overeducation rate by the year of first entry into the labor market.

Table 2 reports the proportion of overeducation in different industries. we can see that the incidence of overeducation varies greatly by industry. Only 5.18% of workers in the education industry are overeducated, while a quarter of workers in the IT industry are overeducated. And we find some meaningful and inspiring conclusions: firstly, the misallocation of human capital factors is most likely to occur in the industries with the most active applicants competition in the labor market. The IT industry, though having a surging growing demand for highly skilled labor, can absorb only a handful of workers, and considerable young graduates fail to find matching jobs. Secondly, the incidence of overeducation in public services industries is relatively low, such as the education sector and the health sector. That's because the positions in the public service industry are characterized by stable “Bianzhi” and high welfare, which attract a large number of highly educated talents for a long time in the context of China’s special national

conditions. Hence, there is not much difference in the educational level between new entrants and existing workers in these industries. Thirdly, the proportion of overeducation is lower in labor-intensive industries, for instance, mining industry, construction industry, manufacturing industry and agriculture industry. Such phenomenon is attributed to many reasons, most is that there seems to be a natural psychological barrier for highly educated graduates voluntarily entering labor-intensive industries in China.

Table2 Incidence of overeducation by industry

Industry	Obs.	Mean	SD
IT	126	0.2857	0.4536
Electricity, gas and water production and supply	273	0.2051	0.4045
Public management and social organizations	730	0.1986	0.3992
Real Estate	234	0.1795	0.3846
Finance	242	0.157	0.3646
Agriculture, forestry, husbandry and fishery	10355	0.1354	0.3422
Leasing and business service	284	0.1268	0.3333
Mining	432	0.1065	0.3088
Culture, sports and entertainment	207	0.1063	0.3089
Transport, storage and communication	1468	0.18	0.3012
Manufacturing	5552	0.17	0.39
Construction	3116	0.0879	0.2832
Resident Services and Other Services	826	0.0872	0.2823
Water conservancy, environment and public facilities management	138	0.087	0.2828
Hotels and restaurants	1244	0.0812	0.2732
Wholesale and retail trade	3094	0.0798	0.2711
Health and medical services	449	0.069	0.2538
Education	908	0.0518	0.2217

Table 3 shows the proportion of overeducation in different regions. We find that the incidence of overeducation is highest in the least economically developed regions (Northwest China), and lowest in regions with a comparatively prosperous economy (such as South China, North China, and Central China), and the rest perform not much differently. Presumably this pattern reflects differences on the demand side of the labor market. Economically developed regions can provide more high-end jobs which can meet the needs of highly educated job seekers. Areas with a backward economy lack demand for highly skilled workers duo to their production structure, which is predominantly based on labor-intensive technologies and low innovative capabilities.

Table 3 Incidence of overeducation by region

Region	Observations	Mean	SD
North China	3695	.0966	0.2955
Northeast China	2683	.1021	0.3029
Central China	7092	.0917	0.2886
South China	4391	.0997	0.2997
Southwest China	4228	.1159	0.3201

East China	3421	.1146	0.3186
Northwest China	5064	.1536	0.3606

4. Empirical strategy

4.1 Empirical equation and IV

To identify the causal impact of parents' overeducation on child outcomes, we estimate the following equation:

$$Y_{ikpt} = \beta_0 + \beta_1 \text{Overeducation}_{ikpt}^P + \beta_2 X_{it} + \beta_3 Z_{pt} + \gamma_k + \delta_p + \omega_t + u_{ikpt} \quad (1)$$

where subscripts i , k , p and t indicate the child, father'(or mother's) year of birth, the province, and the survey year, respectively. Y_{ikpt} indicates the outcome variables of child i whose father (or mother) lived in province p in survey year t . $\text{Overeducation}_{ikpt}^P$ is a binary indicator variable indicating whether child i 's father or mother (denoted by the subscript P) who was born in year k and lived in province p in survey year t is overeducated. β_1 is the coefficient of interest, which captures the effect of father's (or mother's) overeducation on child development. X_{it} contains a vector of parents' and child's demographic characteristics, such as father's (or mother's) age at child's birth, hukou type and years of education, child's age, gender and the number of siblings, and family total income and total expenditure. Z_{pt} is a vector of provincial characteristics, including population, GDP per capita, tertiary industry growth rate, average wage of workers, selling price of commercial housing, education expenditure, number of doctors per capita and university enrollment per capita. γ_k denotes father's (or mother's) birth year fixed effects. δ_p denotes father's (or mother's) province fixed effects. ω_t denotes the survey wave fixed effects. The OLS strategy would result in biased estimates of the coefficient β_1 , because the error term u_{ikpt} may contain unobserved environmental and genetic factors plausibly correlated with both parents' overeducation and child outcomes. We cluster the standard errors at father's (or mother's) province year-of-birth level.

Overeducation is likely correlated with a complex myriad of factors. Although we have already controlled for many possible confounding ones, there may still be omitted unobservable variables that are simultaneously associated with both parents' overeducation and children's development. For instance, job participants who are overeducated may be associated with unobserved work capacity and skills for looking for jobs. If overeducated workers indeed have lower skills, that could also imply worse

quality childcare and explains lower human capital to these children⁵. In addition, individuals who value family over working may have a higher propensity to search for a relaxing job, and a simultaneously below-average job as a trade-off. And these overeducated individuals would be more likely to pay attention to parenting children⁶. Therefore, neglecting the potential unobservable variables associated with overeducation in the children's development function may result in biased and inconsistent estimators. To address these concerns, we employ the instrumental variable (IV) strategy by exploiting variation in the proportion of regional highly educated labor supply to the local population, namely regional employment competition status, as an exogenous source of changes in parents' overeducation.

The IV approach uses variation in the degree of employment competition for workers across time and space as the IV. It relies on the literatures that have confirmed that limiting the number of people receiving higher education can effectively reduce the overeducation rate both theoretically and empirically (Tarvid, 2015; Shen et al., 2023), and the increase in the number of college graduates in each province will translate into an increase in the supply of high-quality human capital in the province since most college students stay to work in the province where their university is located after graduation (Li et al., 2017). The degree of supply of highly educated graduates suggests a larger pressure of job competition in the regional labor market (Knight et al., 2017; He et al., 2020). Job seekers are more likely to be squeezed into industries that do not match their actual education when entering the labor market in a year when the supply of highly educated graduates is greater. Depending on when and where a young worker enters the labor market, he will be exposed to different propensities to looking for a matching job.

We can use a simplified example to illustrate the intuition behind the IV. Consider two bachelor's degree graduates who both hoped to find a job in Beijing of China, but graduated one year apart from each other. Specifically, the one's graduate year was 2021, and the other who coincides with a surge in the number of master's degree graduates who

⁵ Columns 1-5 of Appendix Table A.1 display the average abilities between overeducated and non-overeducated groups in terms of verbal and math test scores and the Big Five personality ((see Appendix B for details). We notice that in most dimensions the overeducated are worse off than non-overeducated ones. And as shown in columns 6-10, we find that the ability of overeducated workers is no longer significantly weaker than that of non-overeducated workers after controlling individual and family characteristics, parents' birth year fixed effects, province-specific time trend, province fixed effects, and survey year fixed effects (as shown in Eq. (3)). While the ability of agreeableness, openness, and extraversion is significantly lower when over-educated, and the overeducated individuals have significantly higher math scores than non-overeducated ones. Therefore, our results can not be exempt from the biased estimation caused by unobservable abilities.

⁶ It is difficult to make a direct test of this conjecture since there are no relevant questions in the questionnaire.

were admitted with a great number in 2020 to ease high unemployment during the COVID-19 pandemic entered the labor market in 2023. When the occupational structure was unable to absorb the sudden increase in the large number of highly educated workers, so the phenomenon of overeducation easily occurred simultaneously. It indicates that the worker graduated from 2023 will be more likely to be overeducated.

In total, when and where each worker enters the labor market determines potential competitors who happen to enter the labor market with him/her. And the education attainment distribution of “relevant competitors” will result in facing different risks of finding a job that does not match his/her education level at the beginning of a job career and a higher likelihood of being overeducated later on (Baert et al., 2013; Acosta-Ballesteros et al., 2018).⁷

We use the proportion of graduates in the regular HE (Higher Education) to the population in each province multiplied by 100 as the instrument for individuals' overeducation. The higher the value, the fiercer employment competition status in the province this year. Figure A.1 shows the variation of IV. And there is rich temporal and spatial variation in job market demand shocks.

The first-stage regression equation is:

$$\begin{aligned} \text{Overeducation}^P_{ikgpt} &= \alpha_0 + \alpha_1 \text{Competition}^P_{ikgp} + \alpha_2 X_{it} + \alpha_3 Z_{pt} + \gamma_k + \delta_p + \delta_p \cdot g + \omega_t \\ &+ u_{ikgpt} \quad (2) \end{aligned}$$

where subscripts i, k, p and t have the same meaning in Eq. (1); and subscript g represents the year in which the father (or mother) first entered the labor market. $\text{Competition}^P_{ikgp}$ is the IV, the employment competition status that child i's father (or mother) who entered labor market in year g and province p faced. Aiming at shutting down the impacts of economic cycles, based on Eq. (1), we further control the province-specific linear time trend $\delta_p \cdot g$ in Eq. (2). And other variables are defined as the same as Eq. (1)⁸.

The second-stage regression equation is:

$$Y_{ikgpt} = \beta_0 + \beta_1 \widehat{\text{Overeducation}}^P_{ikgpt} + \beta_2 X_{it} + \beta_3 Z_{pt} + \gamma_k + \delta_p + \delta_p \cdot g + \omega_t + u_{ikgpt} \quad (3)$$

⁷ Overeducation is always a trap. Mismatches at the start of a career will lead to a higher likelihood of being overeducated later on (Acosta-Ballesteros et al., 2018).

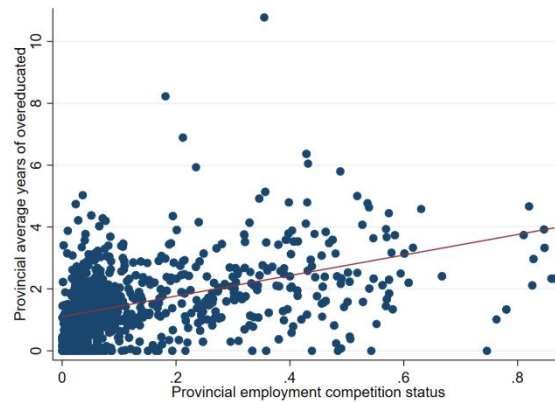
⁸ It should be noted that, with a single cross-section, when the birth year fixed effects and the province-specific linear time trend are already controlled, further including years of education as a control variable will result in multicollinearity. Therefore, years of education are not controlled in Eq. (2).

where $\widehat{Overeducation}^P_{ikpt}$ is the predicted father's (or mother's) overeducation from the first stage regression. And other variables are defined as the same as Eq. (2). The IV estimator of the coefficient β_1 identifies the average effects of father's (or mother's) overeducation on children's outcomes.

Given that the region where the worker's first job is located is not available, we only use samples who have not migrated when we use IV⁹, and see the province of the present job as the province of the first job. We follow the existing literature in using the year of birth plus years of education plus six to calculate the year of labor market entry, namely the graduate year (Schwandt & Von Wachter, 2019; Zheng et al., 2021). The implicit assumption is that people directly join the labor force after finishing schooling, and there are no gap years between different stages of education. There can be exceptions to the assumption. We discuss this in detail in the subsequent robustness checks.

4.2 IV discussion

A reasonable instrumental variable needs to satisfy correlation assumption and exclusion assumption. Figure 3 shows the unconditional relationship between employment competition status and individuals' overeducation. There is a strong unconditional positive relationship between employment competition status and individuals' overeducation levels. The correlation coefficient is equal to 0.373. Table 4 reports the results of first-stage regression for IV. Our employment competition status instrument does a good job to positively predict individuals' overeducation.



⁹ We categorize the following three groups of people as migrants: (i) people who did not live in the same counties when their children were born and when the survey was conducted as migrants. (ii) people who did not live in the same counties when their children were born and when these children were three years old. Using this measurement, the migration rate in our sample is very low as 1.785%. Therefore, biased estimate caused by migration is not a big concern.

Fig. 3 Correlation between employment competition status and overeducation, by province and the year of first entry into the labor market.

Notes: Unconditional expectation between the employment competition status and individuals' overeducation. Annual provincial average years of overeducated on the vertical axis. Annual provincial employment competition status on the horizontal axis. Sample include male workers and female workers aged below 60 years old.

Table 4 First stage regression for IV

	(1) Father	(2) Mother
IV	0.915*** (0.104)	0.429*** (0.092)
Controls	Yes	Yes
Cohort FE	Yes	Yes
Province FE	Yes	Yes
Survey Year FE	Yes	Yes
Province-specific linear time trend	Yes	Yes
Obs.	15512	13660
Adjusted R^2	0.243	0.161

Notes: Parentheses are standard errors clustered at the province-year of birth level. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

A possible concern of using employment competition status as an IV is that the timing of the graduation may be nonrandom. Some students upon graduating may have also considered expected variation in the difficulty of employment competition when they decide whether to continue studying. Then, some cohorts would be endogenously composed of more incompetent job seekers, since they do not have the strength to adjust their employment duration at short notice according to divivable employment status at the predicted graduation year. That is, the time of graduation is a potential confounder of the individuals' overeducation effect since it could impact both parents' overeducation and child outcomes through other channels.

The choice of graduation time actually represents the choice of educational levels. Therefore, we should test for endogenous education responses asking whether, conditional on specific province linear time trend, economic competition shocks do not induce endogenous education responses. However, we have no way of confirming this fact because years of education might be perfectly explained by fixed effects when doing a regression test.

We use supplementary survey data to test for the plausibility of the exclusion assumption. We study the relationship between the IV and the characteristics of workers' households when they first entering labor market. These characteristics include parents' education, hukou type and household income. Appendix Figure A.2 shows that estimates are not significantly different from zero at conventional levels and that I fail to reject the joint test = 0.

This characteristic is important in our analysis because it indicates that employment competition status cannot predict individual education and thus take any precautionary measures to change the year of entering the labor market. We find evidence that lessens this type of concern: the data do not support the hypothesis of young people changing their educational decisions in response to the variation captured in our IV.

Another potential concern about our estimation strategy is that individuals might enter the labor market outside the province of hukou registration, for the purpose of finding a suitable employment environment and obtain a matching job. This leads to some provinces gathering more incompetent job seekers who do not have the ability to migrate across provinces in a systematic manner. We contrast the ability difference between samples working in the hukou domicile and samples working outside the hukou domicile using a t-test in Table A.2. The results showed that the verbal test scores of non-migrants were significantly lower than those of migrants, and the emotional scores of non-cognitive abilities were significantly higher than those of migrants, respectively. And we do not find significant difference in math scores and the other four non-cognitive scores. To sum up, we do not find a large difference in ability between workers who work in the hukou registration province and those who work in other provinces.

The third concern is that employment competition status in the year of graduation may pick up some macroeconomic cycle changes, which could impact both the parents' overeducation and also adolescent development through other channels. In the main specifications, we included province-specific linear time trends, denoted by $\delta_p \cdot g$, to control for cross-province graduate-cohort-varying changes in the outcomes, which somewhat relax this worry. As described above, linear cohort trends at the province level are controlled. However, by focusing on the years around the year of graduation we can get a sense of when macroeconomic conditions start shaping future outcomes. Focusing on the macroeconomic conditions in the graduation year seems the most natural choice, but it is not a priori clear at which point economic conditions affect young graduates the most. If we find that employment competition status in years before graduation matters for future overeducation, we might be concerned that the effect that we find is driven by factors other than the ones we suggest.

We perform a placebo test, using employment competition status in years before or after graduation rather than the actual graduation year of each province. We construct the "placebo" instrumental variable in a similar manner as our main treatment measure, and we also control for the instrumental variable of actual graduation year in these regressions due to systematical overlaps between placebo and real instrumental variable for some cohorts. Figure A.3 illustrates that the coefficients on the "placebo" IV in years

before graduation are statistically insignificant for both father and mother. It provides supportive evidence that our effects of employment competition status in the year of graduation on individuals' overeducation are not due to other unobserved mechanisms. Moreover, we find the estimated coefficients on “placebo” IV after graduation are significant for fathers. It is probably because that our procedure to impute the year of graduation assign an incorrect year, after all there are some students who have gap years between different stages of studies (as we discuss further in Subsection 5.3 and 5.4), and consequently match an incorrect employment competition status. As we expect this measurement error to stay mostly within a close neighborhood around the true graduation year, it may not be particularly surprising that the effect does not disappear sharply when moving away from the imputed graduation year.

5. Results

5.1 Main results

We start our empirical analysis by providing OLS estimates for the effects of parents' overeducation on aspects of child development based on Eq. (1)¹⁰. In panel A of Table 5, results suggest only significant associations between father's overeducation and a lower probability of children's on-time grade progression, but no significant association for children's primary school enrollment, junior school enrollment, the verbal test scores, the math test scores and cognitive skills. And the results in panel B of Table 5 show that mother's overeducation is not significantly correlated with all dependent variables.

To avoid the possible estimation bias, we further verified the above results with the IV. In panel A of Table 6, the IV estimates document that the dummy variable father's overeducation has a negative effect on children's primary school enrollment, junior school enrollment, on-time grade progression, the verbal test scores, math test scores and average cognitive abilities. But results in panel B of Table 6 only suggest significant associations between mother's overeducation and lower probabilities of primary school enrollment or on-time grade progression. Our results hold for various specifications.¹¹ And the coefficients in the IV all shift towards the negative direction compared to the OLS results, meaning OLS may underestimate the negative effect of parents' overeducation on child outcomes. It suggests there may be negative selection into

¹⁰ We also use probit model to estimate the results, which are quantitatively similar.

¹¹ We find that the estimated coefficients on father's and mother's overeducation are nearly unchanged when we add cohort fixed effects, province fixed effects, survey year fixed effects, and province-specific time trend sequentially in the model, or when we estimate a model without any controls. These results are available upon request.

overeducation, i.e., less competent parents are more likely to be overeducated, biasing OLS toward zero.

Overall, our results indicate that fathers' overeducation affects children to have lower enrollment rate and lower cognitive skills than if the same job seeker is put in a position in which the education requirement matches his/her level. In the case of mothers, although no significant negative effect is found in most of the dependent variables, there is also no indication that children's outcomes of parents who are overeducated are better than those of parents who are not overeducated. Similar evidence that the labor market performance of parents has different effects on children between father and mother is documented by Rege et al. (2011), who show that a father's loss of employment during the year prior to college attendance does indeed have a negative effect on probability of college enrollment, but no significant effect for mother. And this is may because men generally experience more severe mental depression due to a bad employment status than women (McKee-Ryan et al., 2005; Kuhn et al., 2009)

Table 5 Impact of parents' overeducation on child development: OLS estimates

	(1)	(2)	(3)	(4)	(5)	(6)
	Primary school enrollment	Junior high school enrollment	On-time grade progression	Verbal test scores	Math test scores	Cognition
Panel A: Father						
Overeducation	-0.012 (0.8)	-0.016 (0.018)	-0.022** (0.011)	-0.050 (0.041)	-0.050 (0.040)	-0.063 (0.045)
Obs.	16170	6082	16170	7124	7041	6983
Adjusted R ²	0.337	0.448	0.073	0.203	0.268	0.249
Panel B: Mother						
Overeducation	0.003 (0.010)	0.005 (0.019)	0.005 (0.013)	0.030 (0.049)	-0.050 (0.044)	-0.006 (0.049)
Obs.	14127	5453	14127	6319	6283	6228
Adjusted R ²	0.339	0.457	0.073	0.192	0.269	0.241
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes	Yes
Survey Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Province-specific linear time trend	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Parentheses are standard errors clustered at the province-year of birth level. * p<0.1; ** p<0.05; ***p<0.01.

Table 6 Impact of parents' overeducation on child development: IV estimates

	(1)	(2)	(3)	(4)	(5)	(6)
	Primary school enrollment	Junior high school enrollment	On-time grade progression	Verbal test scores	Math test scores	Cognition
Panel A: Father						

Overeducation	-0.286*** (0.079)	-0.615* (0.368)	-0.519*** (0.106)	-1.057* (0.617)	-1.499** (0.722)	-1.778** (0.754)
KP(F-stat)	77.451	6.689	77.451	14.397	13.756	13.961
Obs.	15512	5868	15512	6872	6792	6734
Panel B: Mother						
Overeducation	-0.608*** (0.214)	-3.349 (7.039)	-1.085*** (0.314)	-0.784 (1.084)	-1.917 (1.371)	-1.674 (1.320)
KP(F-stat)	21.749	0.237	21.749	4.102	3.848	4.015
Obs.	13660	5308	13660	6161	6125	6072
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes	Yes
Survey Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Province-specific linear time trend	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Parentheses are standard errors clustered at the province-year of birth level. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

5.2 Sample self selection bias

Our results may be contaminated by a potential sample selection issue that individuals' overeducation may increase career instability (Maynard et al., 2006), and consequently affect the marriage and fertility decisions, such as delaying their marriage age and childbearing age, and even less likely to give birth to their children. If individuals' overeducation leads to a postponement of marriage or motherhood, then we observe a selected sample of children.

Figure 4 illustrates the effects of individuals' overeducation on the age of first birth at different age ranges, and Figure 5 illustrates the effects of individuals' overeducation on a binary indicator of giving birth at different age ranges. We find that individuals' overeducation significantly delays their age of first birth and decreases their fertility rates among workers under 35 years old, and there is no significant effect among workers greater than and equal to 35 years old. Therefore, we restrict our sample to children whose parents are 35 years old or older at the time of the interview. And results of Table 7 show that our main findings are robust to imposing this restriction. It implies that the sample selection due to fertility may not be a serious issue in our study.

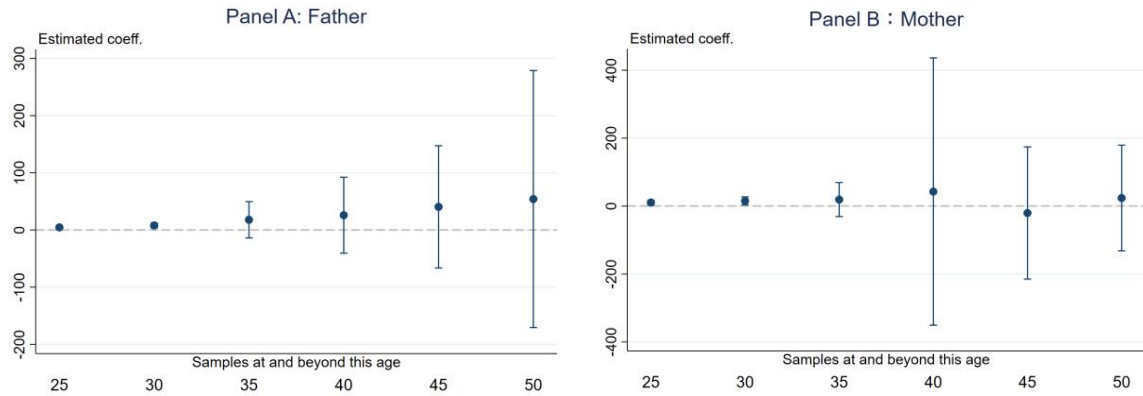


Fig. 4 Impact of workers' overeducation on the age of first birth by different age ranges.

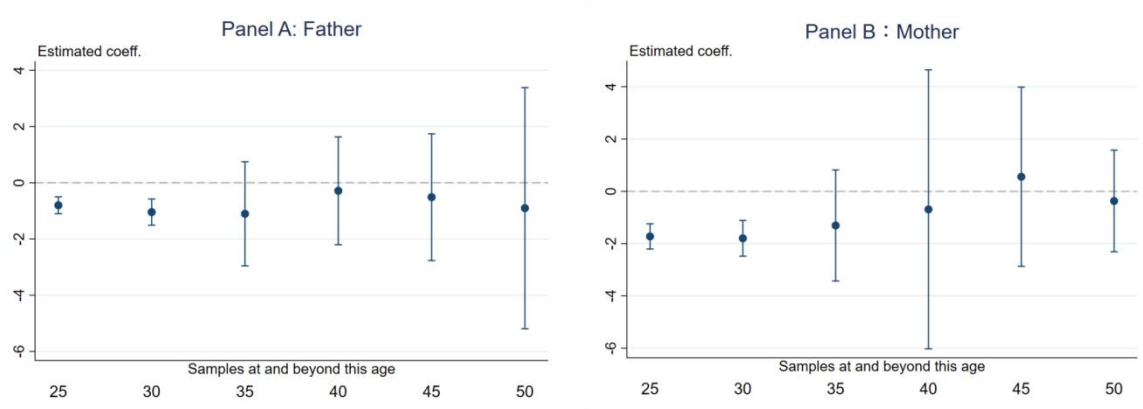


Fig. 5 Impact of workers' overeducation on fertility by different age ranges.

Table 7 Keep samples whose father parents are 35 years old or older

	(1) Primary school enrollment	(2) Junior high school enrollment	(3) On-time grade progression	(4) Verbal test scores	(5) Math test scores	(6) Cognition
Panel A: Father						
Overeducation	-0.249* (0.128)	-0.971* (0.573)	-0.508*** (0.193)	-2.398 (2.049)	-2.415* (1.447)	-3.448 (2.725)
KP(F-stat)	26.465	4.278	26.465	2.296	5.450	2.195
Obs.	11859	5674	11859	6350	6272	6220
Panel B: Mother						
Overeducation	-0.444 (0.369)	-20.891 (192.954)	-1.256** (0.603)	-2.800 (10.769)	-15.203 (56.158)	-11.602 (38.725)
KP(F-stat)	6.739	0.012	6.739	0.121	0.078	0.097
Obs.	8968	4791	8968	5176	5141	5095
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes	Yes
Survey Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Province-speci fic linear time trend	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Parentheses are standard errors clustered at the province-year of birth level. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

5.3 Placebo tests

As discussed above in 4.2, we have found that the employment competition status in years after the actual graduation year still has a significant positive effect on fathers' likelihood of overeducation, which may be due to measurement error. We further perform these tests by re-estimating Eq. (3), replacing the employment competition status in the year of graduation with the employment competition status in years after graduation, and also control for the actual employment competition status in these regressions due to systematical overlaps between placebo and real employment competition status. In Table 8 we show the estimates of our coefficient of interest for fathers when varying the employment competition status from the actual graduation year to five years after graduation¹². Reassuringly, the results show that the fitted value of fathers' overeducation from the first stage regression used by "placebo" IV has no effects on six outcomes of children. It provides supportive evidence that our main results are due to the parents' overeducation as opposed to other unobserved mechanisms.

Table 8 Placebo tests

	(1)	(2)	(3)	(4)	(5)	(6)
	0	1	2	3	4	5
Primary school enrollment	-0.286*** (0.079)	-0.908 (0.755)	-0.401 (0.275)	-0.082 (0.169)	-0.039 (0.140)	-0.006 (0.141)
KP(F-stat)	77.451	2.017	6.361	11.218	15.689	15.052
Obs.	15512	15512	15512	15512	15512	15512
Junior high school enrollment	-0.615* (0.373)	-0.566 (1.217)	-0.579 (0.542)	-0.039 (0.260)	0.107 (0.214)	0.082 (0.214)
KP(F-stat)	6.711	0.894	4.994	11.504	16.514	18.384
Obs.	5868	5868	5868	5868	5868	5868
On-time grade progression	-0.519*** (0.106)	-1.125 (0.961)	-0.659* (0.399)	-0.334 (0.243)	-0.313 (0.201)	-0.377* (0.224)
KP(F-stat)	77.451	2.017	6.361	11.218	15.689	15.052
Obs.	15512	15512	15512	15512	15512	15512
Verbal test scores	-1.057* (0.617)	-2.875 (8.317)	0.316 (1.131)	0.471 (0.635)	0.156 (0.555)	0.057 (0.541)
KP(F-stat)	14.397	0.182	4.289	10.802	13.804	14.489
Obs.	6872	6872	6872	6872	6872	6872
Math test scores	-1.499** (0.722)	-0.724 (5.889)	-0.286 (1.094)	0.173 (0.543)	-0.253 (0.467)	-0.345 (0.466)
KP(F-stat)	13.756	0.164	3.525	9.943	12.773	13.392
Obs.	6792	6792	6792	6792	6792	6792
Cognition	-1.778** (0.754)	-1.971 (6.798)	-0.189 (1.131)	0.224 (0.631)	-0.111 (0.578)	-0.168 (0.579)
KP(F-stat)	13.961	0.191	3.886	10.428	13.266	13.851

¹² The results of mother are shown in Appendix Table A3.

Obs.	6734	6734	6734	6734	6734	6734
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes	Yes
Survey Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Province-specific linear time trend	Yes	Yes	Yes	Yes	Yes	Yes

Notes: We only show the results of child-father observations. Parentheses are standard errors clustered at the province-year of birth level. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

5.4 Robustness checks

5.4.1 Considering migration across provinces

As already given earlier, making regression include migrants likely results in some measurement error. We will erroneously classify workers' location of labor market entry if they have ever changed work location before the survey year. However, our estimations may be biased when we restrict the sample to non-migrants since there are unobservable factors that are correlated with individuals' overeducation levels and migration decisions. For example, workers who pay more attention to maximizing their utilities of academic qualifications are more likely to voluntarily migrate in order to find more matching jobs when they are overeducated, and such workers also be more concerned about their children's education, then the estimated coefficients for parents' overeducation will have a downward bias. To address this concern, we use Eq. (3) to regress the individuals' migrant dummy on overeducation, and we find no evidence that individuals' overeducation affect their migration decisions¹³. In addition, we use both non-migrants and migrants to re-estimate the effects of parents' overeducation on children's outcomes, and we find similar patterns in Panel A of Table 9.

5.4.2 Considering measurement error of graduation time

Two assumptions are implied when using the predicted year of graduation calculated by the year of birth plus years of education plus six as the year of first entering the labor market, that is people directly join the labor force after finishing schooling, and there are no gap years between different stages of education. For the first assumption, we have no way to perform additional tests. For the second assumption, we can rule out samples who are influenced by the Cultural Revolution to have a more accurate identification, since some individuals may have interrupted their studies due to the influence of the Cultural Revolution. We exclude individuals who graduated from 1978 to 1988 and replicate the regressions in Panel B of Table 9. The results are consistent with our main findings.

¹³ For brevity, the results are available upon request.

We also use the year of last leaving school as the year of first entering the labor market. However, using this measurement as the year of labor market entry also needs the same assumption as the method using the year of birth plus years of education plus six. Moreover, for workers who have not been to school, their time of last leaving school are missing, which means samples with zero years of schooling are omitted from the regression. However, the results won't alter if we use the year of last leaving school.

5.4.3 Considering samples who do not have job

In the main analysis, our regression does not include the non-employed group, since only the employed sample has the job information and we can further calculate the corresponding levels of overeducation. In Panel C of Table 9, we examine the sensitivity of the estimates to including unemployed people in a control group and serve whose overeducation dummy as 0. Moreover, there may be some workers who simply quit the labor market due to severe overeducation, which results in the estimates being somewhat inconsistent. So in Panel D of Table 9, we use both the employed sample and the non-employed sample and serve the overeducation dummy of the non-employed sample as 1 to replicate the results. We both find similar patterns in these two checks.

5.4.4 Different measurements of overeducation

In the baseline specification, our overeducation variable in the baseline test is a dummy indicating whether the education level of the individual is more than one standard deviation above the required educational level. We first use the continuous variable to measure overeducation for robustness tests in Panel E of Table 9. We get the number of years overeducated by subtracting the required educational level within his or her occupation from his or her actual years of education. When the value is greater than zero, the respondents are considered to be overeducated, otherwise, they are not. And the higher the value is, the greater the level of overeducation is. We then further reconstruct a dummy variable of overeducation using another Realized Matches (RM) method, which calculates the required educational level using the mode, instead of the mean, of the workers' education levels within each occupation (Santos, 1995). The results with alternative definitions of overeducation are presented in Panel F of Table 9. In all cases, results remain the same, with some differences in significance level.

5.4.5 Keep one observation for each child

One potential concern is that there may be some father-child (or mother-child) pairs that appear multiple times in full samples since the data consists of six waves of the CFPS. If some outcomes are sustained over the course of the child's development, our results based on the pooled sample may be overestimated. To test the sensitivity of our results, we re-estimate the models using the latest observation for each child in Panel G of Table 9. Because this exercise removes some individuals from the sample, it can be more costly to the power of estimation. Again, the effects of parents' overeducation on child development are significantly negative, consistent with IV results.

Table 9 Robustness tests						
	(1)	(2)	(3)	(4)	(5)	(6)
	Primary school enrollment	Junior high school enrollment	On-time grade progression	Verbal test scores	Math test scores	Cognition
Panel A: Considering migration across provinces						
Overeducation	-0.286*** (0.079)	-0.600 (0.407)	-0.525*** (0.104)	-1.058 (0.646)	-1.669** (0.769)	-1.903** (0.802)
KP(F-stat)	80.569	5.160	80.569	12.373	11.790	11.982
Obs.	16170	6082	16170	7122	7039	6981
Panel B: Considering measurement error of graduation time						
Overeducation	-0.286*** (0.080)	-0.678* (0.385)	-0.520*** (0.107)	-1.177* (0.642)	-1.601** (0.740)	-1.928** (0.785)
KP(F-stat)	75.989	6.440	75.989	14.400	13.673	13.865
Obs.	15464	5840	15464	6839	6762	6704
Panel C: Serve the overeducation dummy of the non-employed sample as 0						
Overeducation	-0.277*** (0.072)	-0.645* (0.352)	-0.491*** (0.095)	-0.999* (0.601)	-1.323* (0.687)	-1.614** (0.695)
KP(F-stat)	89.606	7.144	89.606	14.049	13.622	13.829
Obs.	16951	6453	16951	7687	7540	7481
Panel D: Serve the overeducation dummy of the non-employed sample as 1						
Overeducation	-0.260*** (0.067)	-0.704 (0.466)	-0.463*** (0.089)	-1.065 (0.696)	-1.432* (0.828)	-1.757** (0.877)
KP(F-stat)	91.657	3.477	91.657	7.572	7.095	7.139
Obs.	16951	6453	16951	7687	7540	7481
Panel E: Use continuous overeducation variable						
Overeducation	-0.086*** (0.026)	-0.161* (0.096)	-0.157*** (0.036)	-0.278* (0.158)	-0.414** (0.191)	-0.473** (0.191)
KP(F-stat)	43.414	3.840	43.414	11.748	10.626	11.362
Obs.	15512	5868	15512	6851	6769	6713
Panel F: Measure overeducation dummy with mode RM method						
Overeducation	-0.258*** (0.074)	-0.500* (0.297)	-0.468*** (0.101)	-1.060* (0.604)	-1.461** (0.650)	-1.753** (0.705)
KP(F-stat)	70.612	5.964	70.612	11.024	11.076	10.998
Obs.	15512	5868	15512	6872	6792	6734
Panel G: Keep one observation for each child						
Overeducation	-0.283*** (0.098)	-0.881** (0.423)	-0.509*** (0.127)	-1.806 (1.114)	-2.367* (1.369)	-3.058** (1.500)
KP(F-stat)	64.804	5.990	64.804	7.975	7.751	7.778
Obs.	7575	4282	7575	3668	3650	3619

Controls	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes	Yes
Survey Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Province-specific linear time trend	Yes	Yes	Yes	Yes	Yes	Yes

Notes: We only show the results of child-father observations, and the results of child-mother observations are shown in Appendix Table A4. Parentheses are standard errors clustered at the province-year of birth level. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

6. Mechanisms

In the previous section, we find that there is a negative effect of parents' overeducation on children's outcomes, including school enrollment and cognitive abilities. In this section, we discuss the possible mechanisms through which these effects arise.

6.1 Income effects

Adverse income shocks that families experience will have an irreversible long-term impact that can adversely affect children's human capital formation (Dooley & Stewart, 2007). Therefore, the first explanation for the negative effect of parents' overeducation on child development is that the family may not allocate extra economic resources to children's education. Columns 1 and 2 in Table 10 present results with the log of monthly income in the survey year as the dependent variable. We find that the dummy variable overeducation has a negative coefficient for both father and mother after controlling for years of education, etc. It indicates that people who are overeducated are more likely to have lower wages compared to those who have the same education level and do the same job but are not overeducated, as they can not maximize their academic qualifications and advance into the job with better remuneration and benefits. Therefore, our analysis supports the claim that parents' overeducation causes a drop in household income and they are so difficult to allocate extra resources to children, then negatively influences the development of children's human capital.

6.2 Parenting style

In the Chinese traditional view, good education serves as a prerequisite for young people to secure a decent job (Heckman et al., 2018), gain prestige (Chen et al., 2021), and achieve upward mobility (Yang & Qiu, 2016), then the parents usually invest most of household resources in children's education. As detailed in 6.1, we have found that overeducation of parents is associated with a drop in income, and meanwhile, parents

may perceive that their returns to education is declining and the educational investment they ever received was wasted. The shrinking of the expected net return on education investment has made parents realize that the preferences of blindly pursuing high academic qualifications do not foster success in life, and the circumstance is that the obsession with higher education is irrational. Then parents may no longer place as much emphasis on children’s education as before and change their parenting style, which may serve alternative mechanisms.

To investigate this potential channel, we first estimate the impact of parents’ overeducation on their tutoring time. Time input of parents is found to favor children’s cognitive development (Del Boca et al., 2014). Columns 3-4 of Table 10 show the results for the logarithm of the time spent tutoring children per week, and we find that parents who are overeducated are more likely to decrease their time on educational tutoring than parents who are non-overeducated. Then we analyze the effects of parents’ overeducation on the degree of concern for children’s studies¹⁴. As shown in columns 5-6 of Table 10, overeducated parents care less about children’s schooling than non-overeducated parents. We also construct a binary variable that measures whether parents take the initiative to help their children when their children have poor grades in columns 7-8 of Table 10. And the estimates indicate that overeducated parents are less likely to take positive actions in response to children’s lower school grades than those who are not overeducated.

Looking at these results, our finding is reconciled with the hypothesis that overeducated parents are aware that education attainment is not crucial to achieving personal success, and hence change their parenting style, which makes children of overeducated parents perform worse than those whose parents are non-overeducated. Our results also echo Kucel & Vilalta-Bufi (2023), who find that over-educated mothers devote less time to childcare than they would do were they matched and one of the reasons is that overeducated mothers are frustrated with their studies.

Table 10 Mechanism analysis - income effect and parenting style

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Ln(monthly income)		Ln(time spent tutoring children per week)		Degree of concern for children’s studies		Positive response to child's low school grades	
	Father	Mother	Father	Mother	Father	Mother	Father	Mother
Overeducation	-1.265*** (0.465)	-4.178*** (1.372)	-1.369*** (0.319)	-2.815*** (0.848)	-2.449** (1.116)	-4.168* (2.131)	-0.165** (0.080)	-0.255 (0.165)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

¹⁴ For the detailed construction of outcome variables in columns 2 and 3, see Appendix A.

Survey Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province-specific linear time trend	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
KP(F-stat)	44.379	13.341	85.824	20.793	67.707	19.615	60.275	16.217
Obs.	6807	5703	15326	13506	13526	11948	13940	12345

Notes: Parentheses are standard errors clustered at the province-year of birth level. * p<0.1; ** p<0.05; *** p<0.01.

6.3 Intergenerational transmission of education attitudes

In the previous article, we elucidate the parenting style channel and indicate that parents who are overeducated tend to relax discipline on their children. Of course, the primary cause of the change in parenting style is that education is not as important for future success in overeducated workers' minds as before. Under this circumstance, parents may transmit these preferences across generations. The economic literature has extensively supported this hypothesis that parents can influence their child's attitudes and values formation (Dohmen et al., 2012; Giménez-Nadal et al., 2019; Zumbuehl et al., 2021; Brenøe & Epper, 2022). In this section, we regress parents' education attitude on whether they are overeducated or not, and then further test children's responses.

We use educational expectation as a proxy for individuals' attitudes to education (see Appendix A for more details). In columns 1 and 2 of Table 11, results shows that parents who are overeducated are more likely to have low expectations for children's education and we can find that there is a strong transmission of education attitudes from fathers to children as the correlation of column 3 is significant. This finding suggests that overeducated parents may place lower expectations on their children's educational attainment, and children's attitudes to education are also molded by their parents and have a lower value on education, with likely consequences for children whose parents are overeducated have worse educational development.

Table 11 Mechanism analysis - intergenerational transmission of education attitudes

	(1)	(2)	(3)	(4)
	Parents' educational expectations for children	Parents' educational expectations for children	Children's educational expectations for themselves	Children's educational expectations for themselves
	Father	Mother	Father	Mother
Overeducation	-2.061*** (0.706)	-4.449*** (1.419)	-3.602* (2.026)	-3.186 (3.846)
Controls	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes
Survey Year FE	Yes	Yes	Yes	Yes
Province-specific linear time trend	Yes	Yes	Yes	Yes

KP(F-stat)	77.564	23.623	20.400	5.243
Obs.	13649	12264	8316	7475

Notes: Parentheses are standard errors clustered at the province-year of birth level. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

In summary, we find the overeducated workers have lower wages than the well-matched, which have a negative effect on children's human capital development. And there are divergent parenting styles and attitudes to education between overeducated parents and non-overeducated parents. Overeducated parents tend to take negative actions in response to children's lower grades, to show a disregard for children's schooling and to have lower expectations for children's education. By contrast, non-overeducated parents are more likely than parents to be positive. Parents also transmit similar attitudes and values for education to their children. Such divergence between overeducated parents and non-overeducated parents and the difference between their corresponding children are another two potential mechanisms that underlie the negative effect of parents' overeducation on children's development. We believe the negative effect of parents' overeducation on children's development is not through a signal channel. We can think of the phenomenon as the decrease of family resources caused by overeducation has a negative impact on children, and the effect is amplified by changes in household preference for education, then leads to an ongoing chain reaction. Or we can think of it as that parents change their attitudes towards education when realizing that education attainment is not crucial for a high-paying job, then mold their children's attitudes unconsciously, and these channels may also reinforce each other's effects leading to poor development of children's human capital. In this sense, rich family economic resources and the good household education environment are important guarantees to promote the improvement of children's education level.

7. Heterogeneous test

In addition to showing the average treatment impacts of parents' overeducation on children's development, we explore heterogeneous effects that vary across family background characteristics, including parents' hukou status and household income levels. It is worth pointing out that we can serve people with rural hukou type as those who lives in a slightly poor family background because there is inequality for different types of hukou in China due to a unique dualistic urban-rural household registration system.

Table 12 presents the empirical results for the heterogeneous effects of fathers' overeducation on child development for different types of hukou. The findings show that children from rural hukou families have more detrimental effects on their development than children from urban hukou families if their parents are overeducated. We proceed to

conduct heterogeneity analysis according to household income. And we use the median household income in each province as the dividing standard, it is considered to be a high-income family if the household income is above the median, otherwise, it is considered to be a low-income family. Results in Table 10 show that the negative effect of fathers' overeducation on child outcomes is less severe in higher-income families than in lower-income families.

Presumably this pattern reveals that people with worse family backgrounds suffer more when overeducated than their counterparts with a better family background. That is may because privileged families attach greater importance to the quality of their children and it's hard to shake their attitudes towards education, so they may react less drastically to the lower return of human capital caused by increased overeducated levels. As a result, children living in rural hukou families and low-income families perform worse when their parents are overeducated than those living in urban hukou families and high-income families.

Table 12 Heterogeneous tests

	(1) Primary school enrollment	(2) Urban	(3) Junior high school enrollment	(4) Urban	(5) On-time grade progression	(6) Urban	(7) Verbal test scores	(8) Urban	(9) Math test scores	(10) Urban	(11) Cognition	(12) Urban
Panel A: By types of hukou												
Overeduca tion	Rural -0.172*	Urban -0.683*	Rural -0.201	Urban -2.035	Rural -0.349***	Urban -0.703**	Rural -1.572	Urban 2.755	Rural -1.702*	Urban -53.337	Rural -2.219*	Urban 2.072
	(0.088)	(0.272)	(0.403)	(3.649)	(0.114)	(0.304)	(1.035)	(6.330)	(0.828)	(295.766)	(1.134)	(5.185)
KP(F-stat)	56.783	12.768	4.077	0.401	56.783	12.768	8.103	0.357	8.892	0.000	7.809	0.366
Obs.	12335	3065	4643	1185	12335	3065	5413	1446	5388	1381	5301	1420
Panel B: By household income												
Overeduca tion	High -0.227***	Low -0.354*	High -0.362	Low -1.075	High -0.398***	Low -0.716**	High 0.253	Low -4.205	High -0.330	Low -3.923	High -0.345	Low -5.253
	(0.088)	(0.161)	(0.388)	(0.847)	(0.113)	(0.221)	(0.767)	(2.761)	(0.785)	(2.627)	(0.807)	(3.304)
KP(F-stat)	61.622	29.959	6.056	2.672	61.622	29.959	7.035	2.930	9.650	2.686	6.747	2.794
Obs.	8368	7139	2988	2876	8368	7139	3563	3303	3531	3255	3496	3232
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Survey Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province-s pecific linear time trend	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: We only show the results of child-father observations, and the results of child-mother observations are shown in Appendix Table A5. Parentheses are standard errors clustered at the province-year of birth level. * p<0.1; ** p<0.05; ***p<0.01.

8. Conclusion

In this paper, we study how parents' overeducation affects their children's development. Using data from a nationwide Chinese survey, the empirical analysis shows that children of overeducated parents have lower school enrollment and perform worse cognitive abilities. This negative effect is mostly driven by the loss of household income and the change in parents' attitudes towards pursuing high education qualifications. Meanwhile, children's preferences for education are also molded by their parents, which mutually reinforces the bad development. We further observe strong divergence between families with different types of hukou and different levels of household income in how parents' overeducation affects child development. The results suggest that the development of children who live in better family backgrounds, i.e., with urban hukou type and higher household income, are less sensitive to parents' overeducation than those who live in poor family backgrounds, and we assert that this result is that high-skilled families' attitudes towards the importance of education won't be easily swayed by low returns to education.

The rapid expansion of college enrollment in China has been criticized for the depressing wages and education mismatch of new college graduates. This paper now provides compelling evidence that, apart from the impact on this generation, overeducation may make parents realize high education qualification is not crucial for fostering future success and changing their parenting styles, then give rise to the worse development of the next generation. It is noted that the negative impact we find on children's human capital development may be temporary. If the overeducated worker find a matching job and improve the work situation, this negative trend can be reversed to a certain extent, but the negative impact caused by it is difficult to recover all. This is also one of the micro costs caused by the reform of China's education expansion.

Our results indicate that the overeducation caused by changes in the industrial structure and a growth in the well-educated labor supply has had a long-lasting and great negative effect. This imbalanced structure could have led to over overcrowded labor market or negative peer effects, and not only on the cohort themselves, but also on the next generation. Governments may have to rethink the further sustained expansion of higher education or, at least, consider the need for reform in the higher education system and labor market industrial structure.

There is support for education reform: (i) the Chinese government should insist on strengthening the development of vocational education and put more emphasis on educational spilt-flow. We need to form a consensus that pursuing a high academic qualification should not be the best choice for all students and it is important to choose the appropriate education to maximize future benefits. (ii) there is a need for transparency

and improved information available to young people and their parents making their choices about educational investments. Forward-looking information provision should be based on the latest evidence in a changing labor market. We propose that the state should provide regular information on the distribution of the returns to graduate education, and make it become part of the remit of a modern government.

Appendix A. Additional tables

Variables	Non-overeducated		Overeducated		Diff (2)-(4)	Non-overeducated		Overeducated		Diff (7)-(9)
	Obs.	Mean	Obs.	Mean		Obs.	Mean	Obs.	Mean	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Father										
Verbal test scores	7883	0.01	1299	-0.06	0.07***	7879	0	1297	-0.01	0.01
Math test scores	7759	0	1287	0.02	-0.03	7755	0	1285	0.04	-0.04**
Conscientiousness	13	-0.06	10	0.08	-0.150	13	0	10	0	0
Agreeableness	7769	0.02	1302	-0.06	0.08***	7706	0.02	1283	-0.06	0.08***
Extraversion	6912	0.01	1171	-0.06	0.07***	6849	0.01	1152	-0.04	0.05*
Openness	6953	0.01	1104	-0.07	0.08***	6953	0.01	1104	-0.06	0.07**
Emotional Stability	6904	0.01	1171	-0.04	0.05*	6841	0.01	1152	-0.02	0.02
Panel B: Mother										
Verbal test scores	8894	-0.01	821	-0.09	0.08***	8888	-0.01	819	-0.01	0
Math test scores	8724	0	827	-0.02	0.01	8718	-0.01	825	0.02	-0.02
Conscientiousness	17	0.06	7	-0.24	0.3	17	0	7	0	0
Agreeableness	8525	0.01	753	0	0	8370	0.01	733	0	0
Extraversion	7504	0.01	710	-0.15	0.15***	7349	0	690	-0.12	0.12***
Openness	7451	0.01	646	-0.02	0.02	7451	0.01	646	-0.02	0.03
Emotional Stability	75	0.01	710	-0.04	0.04	7345	0.01	690	-0.02	0.02

Notes: Columns 1-5 reports the average abilities between overeducated and non-overeducated groups. Columns 6-10 reports the average abilities between overeducated and non-overeducated groups controlling individual and family characteristics, parents' birth year fixed effects, province-specific time trend, province fixed effects, and survey year fixed effects. * p<0.1; ** p<0.05; *** p<0.01.

Variables	Non-migrant		Migrant		Diff (2)-(4)
	Obs.	Mean	Obs.	Mean	
	(1)	(2)	(3)	(4)	(5)
Panel A: Father					
Overeducation	12445	0.143	195	0.072	0.071***
Verbal test scores	9644	0.286	171	0.468	-0.182***
Math test scores	9463	0.254	159	0.314	-0.059
Conscientiousness	22	0.470	1	-0.631	1.101
Agreeableness	9299	-0.037	150	-0.026	-0.010
Extraversion	8441	0.048	90	-0.077	0.125
Openness	8414	0.097	185	0.113	-0.017

Emotional Stability	8433	0.121	90	-0.222	0.342***
Panel B: Mother					
Overeducation	12594	0.082	222	0.104	-0.021
Verbal test scores	11778	0.166	262	0.243	-0.077
Math test scores	11301	0.023	230	0.052	-0.030
Conscientiousness	24	0.027	1	0.114	-0.087
Agreeableness	10597	-0.052	217	-0.141	0.089
Extraversion	9775	0.010	107	-0.028	0.037
Openness	9927	0.028	282	0.077	-0.050
Emotional Stability	9776	-0.071	106	-0.304	0.233**

Notes: Columns 1-2 reports the outcomes of workers who work in the hukou registration province. Columns 3-4 reports the outcomes of workers who do not work in the hukou registration province. Column 5 tests the difference between these two samples using a t-test. Standard deviations or standard errors are in parentheses.

Table A3 Placebo tests of mother panel

	(1)	(2)	(3)	(4)	(5)	(6)
	0	1	2	3	4	5
Primary school enrollment	-0.608***	-0.604	-1.354	-3.257	6.973	0.807
	(0.214)	(0.487)	(1.650)	(7.713)	(40.500)	(0.954)
KP(F-stat)	21.749	4.289	0.880	0.185	0.030	1.087
Obs.	13660	13660	13660	13660	13660	13660
Junior high school enrollment	-3.534	3.582	5.014	-1.910	-1.543	-1.418
	(7.520)	(5.939)	(12.451)	(2.881)	(1.884)	(1.198)
KP(F-stat)	0.230	0.391	0.172	0.510	0.843	1.756
Obs.	5308	5308	5308	5308	5308	5308
On-time grade progression	-1.085***	0.158	0.642	0.935	-3.389	-1.320
	(0.314)	(0.501)	(1.223)	(3.066)	(20.341)	(1.525)
KP(F-stat)	21.749	4.289	0.880	0.185	0.030	1.087
Obs.	13660	13660	13660	13660	13660	13660
Verbal test scores	-0.784	-6.812	0.651	-0.085	-0.232	-0.232
	(1.084)	(35.445)	(2.420)	(0.946)	(0.786)	(0.651)
KP(F-stat)	4.102	0.043	1.379	5.143	6.618	8.384
Obs.	6161	6161	6161	6161	6161	6161
Math test scores	-1.917	-8.487	1.204	0.217	-0.250	-0.255
	(1.371)	(54.248)	(2.239)	(0.894)	(0.795)	(0.654)
KP(F-stat)	3.848	0.026	1.279	4.913	6.159	8.177
Obs.	6125	6125	6125	6125	6125	6125
Cognition	-1.674	-12.961	1.072	0.060	-0.415	-0.451
	(1.320)	(79.104)	(2.478)	(0.990)	(0.885)	(0.754)
KP(F-stat)	4.015	0.028	1.325	5.023	6.318	8.128
Obs.	6072	6072	6072	6072	6072	6072
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes	Yes
Survey Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Province-specific linear time trend	Yes	Yes	Yes	Yes	Yes	Yes

Notes: We only show the results of child-mother observations. Parentheses are standard errors clustered at the province-year of birth level. * p<0.1; ** p<0.05; ***p<0.01.

Table A4 Robustness tests for mother panel

	(1)	(2)	(3)	(4)	(5)	(6)
--	-----	-----	-----	-----	-----	-----

	Primary school enrollment	Junior high school enrollment	On-time grade progression	Verbal test scores	Math test scores	Cognition
Panel A: Considering migration across provinces						
Overeducation	-0.617*** (0.211)	-3.640 (8.068)	-1.077*** (0.310)	-0.845 (1.231)	-2.286 (1.592)	-2.005 (1.580)
KP(F-stat)	21.777	0.206	21.777	3.259	3.061	3.166
Obs.	14127	5452	14127	6318	6282	6227
Panel B: Considering measurement error of graduation time						
Overeducation	-0.609*** (0.215)	-4.205 (10.549)	-1.089*** (0.316)	-0.752 (1.102)	-1.890 (1.397)	-1.653 (1.345)
KP(F-stat)	21.615	0.163	21.615	3.897	3.647	3.808
Obs.	13653	5302	13653	6154	6118	6065
Panel C: Serve the overeducation dummy of the non-employed sample as 0						
Overeducation	-0.586*** (0.184)	-2.053 (2.909)	-1.021*** (0.259)	-1.238 (1.010)	-1.456 (1.059)	-1.574 (1.093)
KP(F-stat)	89.606	7.144	89.606	14.049	13.622	13.829
Obs.	16951	6453	16951	7687	7540	7481
Panel D: Serve the overeducation dummy of the non-employed sample as 1						
Overeducation	-0.991** (0.436)	2.721 (6.415)	-1.728** (0.698)	-3.355 (5.354)	-3.960 (6.012)	-4.234 (6.422)
KP(F-stat)	7.400	0.190	7.400	0.555	0.518	0.542
Obs.	16715	6385	16715	7613	7508	7450
Panel E: Use continuous overeducation variable						
Overeducation	-0.913 (1.457)	7.755 (151.963)	-1.630 (2.571)	-1.077 (2.918)	-3.036 (9.271)	-2.833 (8.949)
KP(F-stat)	0.417	0.003	0.417	0.160	0.111	0.103
Obs.	13660	5308	13660	6161	6125	6072
Panel F: Measure overeducation dummy with mode RM method						
Overeducation	-0.693** (0.292)	-3.707 (9.557)	-1.237*** (0.459)	-1.112 (1.546)	-2.841 (2.702)	-2.444 (2.324)
KP(F-stat)	10.620	0.147	10.620	1.643	1.412	1.523
Obs.	13660	5308	13660	6161	6125	6072
Panel G: Keep one observation for each child						
Overeducation	-0.415* (0.226)	-2.340 (2.547)	-0.827*** (0.308)	1.316 (1.497)	-1.078 (1.424)	-0.613 (1.351)
KP(F-stat)	21.321	1.089	21.321	3.218	3.628	3.158
Obs.	6726	3866	6726	3296	3284	3256
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes	Yes
Survey Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Province-specific linear time trend	Yes	Yes	Yes	Yes	Yes	Yes

Notes: We only show the results of child-mother observations. Parentheses are standard errors clustered at the province-year of birth level. * p<0.1; ** p<0.05; ***p<0.01.

Table A5 Heterogenous tests for mother panel

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Primary school enrollment		Junior high school enrollment		On-time grade progression		Verbal test scores		Math test scores		Cognition	
Panel A: By types of hukou												
Overeduca	Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban
	-0.415*	-0.913	1.539	-1.026	-0.812**	-1.217	-0.603	-1.650	-0.340	-3.658**	-2.069	-2.619*

tion	(0.229)	(0.629)	(1.913)	(0.854)	(0.329)	(0.798)	(2.935)	(1.131)	(3.085)	(1.812)	(4.335)	(1.489)
KP(F-stat)	13.933	4.414	0.711	2.490	13.933	4.414	0.535	4.692	0.389	4.789	0.515	4.727
Obs.	11025	2282	4275	904	11025	2282	4943	1199	4921	1185	4878	1175
Panel B: By household income												
Overeducation	High -0.723* **	Low -0.244	High -0.363	Low 1.028	High -1.118***	Low -0.770	High -0.454	Low 3.227	High -1.601 *	Low 46.536	High -1.314	Low 7.158
KP(F-stat)	(0.250)	(0.354)	(0.584)	(1.671)	(0.342)	(0.521)	(0.693)	(23.123)	(0.969)	(325.882)	(0.899)	(47.488)
Obs.	7538	6120	2767	2537	7538	6120	3281	2876	3269	2852	3236	2832
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Survey Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province-specific linear time trend	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: We only show the results of child-mother observations. Parentheses are standard errors clustered at the province-year of birth level. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

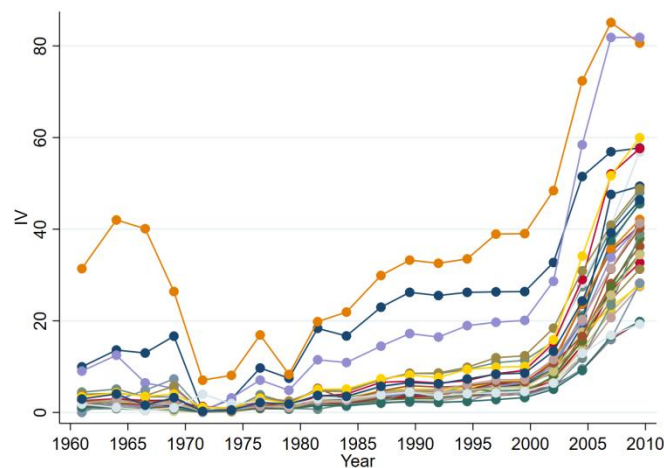


Figure A1 Instrumental variable for different provinces in 1960 - 2010

Notes: IV is calculated by the proportion of graduates in the regular HE (Higher Education) to the population in each province multiplied by 10000 as the instrument for individuals' overeducation. Each line represents a province in China. Source is Educational Statistics Yearbook of China 1960-2010. The decline of IV in 1960s and 1970s was the result of the Cultural Revolution.

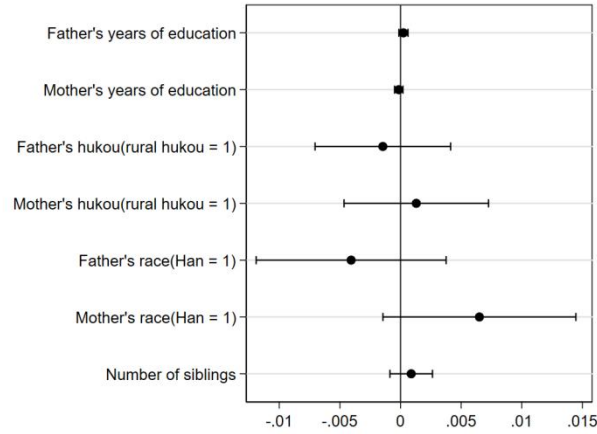


Figure A2 Instrumental variable and household characteristics

Notes: Point estimates and 95% confidence intervals of a regression of the instrumental variable on workers' household characteristics (shown in the figure), region of birth fixed effects, cohort fixed effects, survey year fixed effects and province - specific linear time trend. The Sample are individuals whose parents' information in not missing and have jobs in CFPS 2010-2020.

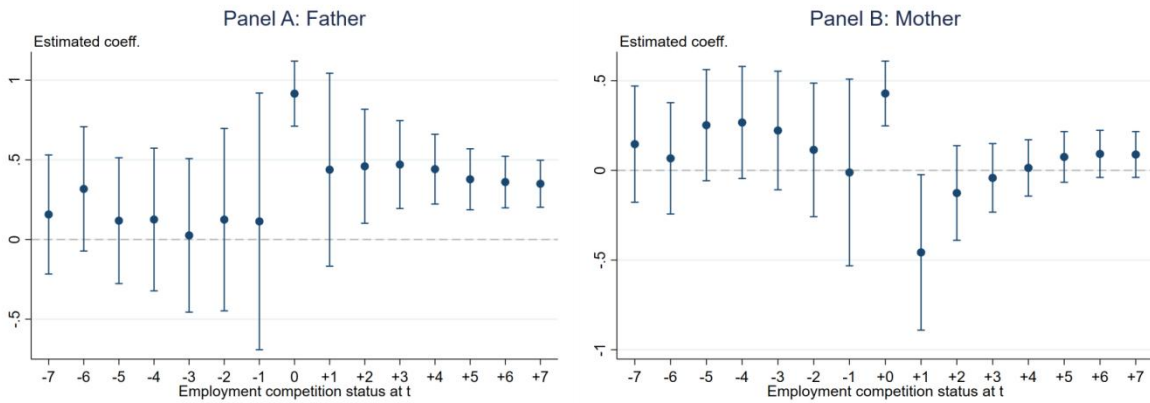


Figure A3 Effect of instrumental variable in the year of graduation and adjacent years.

Notes: The markers represent the estimated effect of the employment competition status in year t on the dummy of overeducation, where $t = 0$ is the year of graduation. The lines represent the 95% confidence intervals. The cap on each line represents the 90% confidence interval. All regressions include individual characteristics, child characteristics, family background, province characteristics, province-specific time trend, birth year fixed effects, province fixed effects, and survey year fixed effects. In addition to the regression of " $t=0$ ", we control for the employment competition state of the actual year of graduation in other regressions. Parentheses are standard errors clustered at the province-year of birth level.

Appendix B. Construction of mechanism variables

Degree of concern for children's studies

For children aged 10–15, the primary caregivers or adult family members living with them were asked to answer 6 questions about how often they gave up watching TV due to children's study, discussed with children about school life, asked children to finish

homework, checked children's homework, restricted children from watching TV, and restricted the types of TV programs children could watch. The answer is on a 5-point Likert scale from 1 (totally disagree) to 5 (totally agree). We sum up all the responses to construct an index of parenting style, where the total score ranges from 6 to 30. And higher scores indicate more effective interactions between parents and children.

Positive response to child's low school grades

For children aged 10–15, the primary caregivers are asked about the most common reaction if the children bring back a transcript with a score lower than their expectations. We construct a binary variable measuring parents’ response to child's low school grades, which equals one if the parents have a positive response, including contacting the teacher, encouraging the child to study harder, or helping the child more, and zero otherwise (i.e., physical punishment, scolding the child, restricting the child's activities, or no action).

Educational expectations

To measure educational expectation, the CFPS asked each respondents the following survey question: What is the minimum level of education you think your child should attain? The answers include “primary school,” “junior high school,” “senior high school,” “3-year college,” “4-year college/bachelor's degree,” “master's degree,” and “doctoral degree.” We construct a continuous variable that from 0 to 26 (0 represents illiteracy, 3 represents kindergarten, 9 represents primary school, 12 represents middle school, 15 represents high school, 18 represents three years of junior college, 19 represents a bachelor’s degree, 22 represents a master’s degree, and 26 represents a doctoral degree). Additionally, the CFPS asked each adolescent aged 10–15 “What is the minimum level of education you think you should attain?” The answer to the question and our variable construction is the same as above.

Non-cognitive skills

The construction method of word and math scores is the same as described early. For non-cognitive skills, we use a commonly used taxonomy which is known as the Big Five personality. The Big Five personality comprise five main dimensions, including openness, conscientiousness, extraversion, agreeableness, and emotional stability. Referring to a similar study, we use 15 questions that asks respondents the extent to which they agree to a series of statements based on the questionnaire information. And we standardize every dimension of the Big Five non-cognitive score to have a mean of 0 and a standard deviation of 1 by survey year. Details of five dimensions of Big Five are given in the Appendix Table B1. And educational expectations for oneself are constructed in the same way as educational expectations for children.

Table B1 Details for construction of non-cognitive skills

Variables	Question	Variable value	Note
Conscientiousness	Q1: I consider myself a hard worker	1 represents disagree strongly, 2 represents disagree 3 represents neutral 4 represents agree	

		5 represents agree strongly	
	Q2: I think I will concentrate on my studies	1 represents disagree strongly, 2 represents disagree 3 represents neutral 4 represents agree 5 represents agree strongly	
	Q3: I think I will check my homework	1 represents disagree strongly, 2 represents disagree 3 represents neutral 4 represents agree 5 represents agree strongly	
	Q4: I think I would obey the school rules	1 represents disagree strongly, 2 represents disagree 3 represents neutral 4 represents agree 5 represents agree strongly	
	Q5: I think I would put things in order	1 represents disagree strongly, 2 represents disagree 3 represents neutral 4 represents agree 5 represents agree strongly	
	Q6: I think I would finish my homework before I play	1 represents disagree strongly, 2 represents disagree 3 represents neutral 4 represents agree 5 represents agree strongly	
Openness	How interested are you in this investigation	1-7, 1 represents the lowest, 7 represents the highest	
Agreeableness	How good is the relationship with people (score)	0-10, 0 represents the lowest, 10 represents the highest	
Extraversion	Q1: How often you feel like your life can not go on in the past week	1 represents mostly (5-7 days), 2 represents often (3-4 days) 3 represents sometimes (1-2 days) 4 represents few (less than a day)	We reverse the scores of Q1 and Q3, so that a higher score indicates that individuals have higher emotional stability.
	Q2: How often you feel like your life is very enjoyable in the past week	1 represents mostly (5-7 days), 2 represents often (3-4 days) 3 represents sometimes (1-2 days) 4 represents few (less than a day)	
	Q3: How often you feel lonely in the past week	1 represents mostly (5-7 days), 2 represents often (3-4 days) 3 represents sometimes (1-2 days) 4 represents few (less than a	

		day)	
Emotional Stability	Q1: How often you in a blue mood in the past week	1 represents mostly (5-7 days), 2 represents often (3-4 days) 3 represents sometimes (1-2 days) 4 represents few (less than a day)	We reverse the scores of Q1, Q2 and Q4, so that a higher score indicates that individuals have higher emotional stability.
	Q2:How often you feel depressed in the past week	1 represents mostly (5-7 days), 2 represents often (3-4 days) 3 represents sometimes (1-2 days) 4 represents few (less than a day)	
	Q3: How often you feel happy in the past week	1 represents mostly (5-7 days), 2 represents often (3-4 days) 3 represents sometimes (1-2 days) 4 represents few (less than a day)	
	Q4: How often you sleep poor in the past week	1 represents mostly (5-7 days), 2 represents often (3-4 days) 3 represents sometimes (1-2 days) 4 represents few (less than a day)	

References

- Acosta-Ballesteros, J., Osorno-del Rosal, M. D. P., & Rodríguez-Rodríguez, O. M. (2018). Overeducation of young workers in Spain: how much does the first job matter? *Social indicators research*. *Social Indicators Research*, 138, 109-139.
- Battu, H., Belfield, C. R., & Sloane, P. J. (2003). Human capital spillovers within the workplace: evidence for Great Britain. *oxford Bulletin of Economics and Statistics*, 65(5), 575-594.
- Belfield, C. (2010). Over-education: What influence does the workplace have?. *Economics of education review*, 29(2), 236-245.
- Blau, P. M., & Duncan, O. D. (1967). *The American Occupational Structure*. New York: John Wiley and Sons.
- Brenøe, A. A., & Epper, T. (2022). Parenting values and the intergenerational transmission of time preferences. *European Economic Review*, 148, 104208.
- Burton, P., Phipps, S., & Curtis, L. (2002). All in the family: A simultaneous model of parenting style and child conduct. *American Economic Review*, 92(2), 368-372.
- Chan, T. W., & Koo, A. (2011). Parenting style and youth outcomes in the UK. *European sociological review*, 27(3), 385-399.
- Che, Y., & Zhang, L. (2018). Human capital, technology adoption and firm performance: Impacts of China's higher education expansion in the late 1990s. *The Economic Journal*, 128(614), 2282-2320.
- Cui, Y., Liu, H., & Zhao, L. (2019). Mother's education and child development: Evidence from the compulsory school reform in China. *Journal of Comparative Economics*, 47(3), 669-692.
- Del Boca, D., Flinn, C., & Wiswall, M. (2014). Household choices and child development. *Review of Economic Studies*, 81(1), 137-185.

- Dohmen, T., Falk, A., Huffman, D., & Sunde, U. (2012). The intergenerational transmission of risk and trust attitudes. *The Review of Economic Studies*, 79(2), 645-677.
- Dooley, M., & Stewart, J. (2007). Family income, parenting styles and child behavioural–emotional outcomes. *Health economics*, 16(2), 145-162.
- Flisi, S., Goglio, V., Meroni, E. C., Rodrigues, M., & Vera-Toscano, E. (2017). Measuring occupational mismatch: overeducation and overskill in Europe—evidence from PIAAC. *Social Indicators Research*, 131, 1211-1249.
- Fu, H., Ge, R., Huang, J., & Shi, X. (2022). The effect of education on health and health behaviors: Evidence from the college enrollment expansion in China. *China Economic Review*, 72, 101768.
- Giménez-Nadal, J. I., Mangiavacchi, L., & Piccoli, L. (2019). Keeping inequality at home: The genesis of gender roles in housework. *labor Economics*, 58, 52-68.
- Glewwe, P., Jacoby, H. G., & King, E. M. (2001). Early childhood nutrition and academic achievement: a longitudinal analysis. *Journal of public economics*, 81(3), 345-368.
- Hartog, J. (2000). Over-education and earnings: where are we, where should we go?. *Economics of education review*, 19(2), 131-147.
- He, Y., Peng, X., & Xu, H. (2020). Overeducation, market recognition, and effective labor supply. *China Economic Review*, 59, 101384.
- Heckman, J. J., Humphries, J. E., & Veramendi, G. (2018). Returns to education: The causal effects of education on earnings, health, and smoking. *Journal of Political Economy*, 126(S1), S197-S246.
- Hilger, N. G. (2016). Parental job loss and children's long-term outcomes: Evidence from 7 million fathers' layoffs. *American Economic Journal: Applied Economics*, 8(3), 247-283.
- Huang, B., Tani, M., Wei, Y., & Zhu, Y. (2022). Returns to education in China: Evidence from the great higher education expansion. *China Economic Review*, 74, 101804.
- Jiang, S., & Guo, Y. (2022). Reasons for college major-job mismatch and labor market outcomes: Evidence from China. *China Economic Review*, 74, 101822.
- Knight, J., Deng, Q., & Li, S. (2017). China's expansion of higher education: The labor market consequences of a supply shock. *China Economic Review*, 43, 127-141.
- Kucel, A., & Vilalta-Bufi, M. (2013). Why do tertiary education graduates regret their study program? A comparison between Spain and the Netherlands. *Higher Education*, 65, 565-579.
- Kucel, A., & Vilalta-Bufi, M. (2023). Overeducation and childcare time. *International Journal of Manpower*, 44(9), 108-127.
- Kuhn, A., Lalive, R., & Zweimüller, J. (2009). The public health costs of job loss. *Journal of health economics*, 28(6), 1099-1115.
- Lazear, E. (1977). Education: consumption or production?. *Journal of Political Economy*, 85(3), 569-597.
- Li, F., Morgan, W. J., & Ding, X. (2008). The expansion of higher education, employment and over-education in China. *International Journal of Educational Development*, 28(6), 687-697.
- Li, H., Ma, Y., Meng, L., Qiao, X., & Shi, X. (2017). Skill complementarities and returns to higher education: Evidence from college enrollment expansion in China. *China Economic Review*, 46, 10-26.
- Liu, H., & Zhao, Z. (2014). Parental job loss and children's health: Ten years after the massive layoff of the SOEs' workers in China. *China Economic Review*, 31, 303-319.
- Mateos-Romero, L., & Salinas-Jiménez, M. D. M. (2018). Labor mismatches: Effects on wages and on job satisfaction in 17 OECD countries. *Social Indicators Research*, 140, 369-391.
- Maynard, D. C., Joseph, T. A., & Maynard, A. M. (2006). Underemployment, job attitudes, and turnover intentions. *Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior*, 27(4), 509-536.
- McKee-Ryan, F., Song, Z., Wanberg, C. R., & Kinicki, A. J. (2005). Psychological and physical well-being during unemployment: a meta-analytic study. *Journal of applied psychology*, 90(1), 53.
- Meng, X., Shen, K., & Xue, S. (2013). Economic reform, education expansion, and earnings inequality for urban males in China, 1988–29. *Journal of Comparative Economics*, 41(1), 227-244.

- Oreopoulos, P., Page, M., & Stevens, A. H. (2008). The intergenerational effects of worker displacement. *Journal of Labor Economics*, 26(3), 455-483.
- Piracha, M., Tani, M., Zimmermann, K. F., & Zhang, Y. (2022). Higher education expansion and the rise of China in economics research. *China Economic Review*, 74, 101813.
- Qin, N., & Kong, D. (2021). Human capital and entrepreneurship. *Journal of Human Capital*, 15(4), 513-553.
- Rege, M., Telle, K., & Votruba, M. (2011). Parental job loss and children's school performance. *The Review of Economic Studies*, 78(4), 1462-1489.
- Santos, M. C. (1995). Education and earnings differentials in Portugal. PhD diss., Faculdade de Economia, Universidade do Porto.
- Schwandt, H., & Von Wachter, T. (2019). Unlucky cohorts: Estimating the long-term effects of entering the labor market in a recession in large cross-sectional data sets. *Journal of Labor Economics*, 37(S1), S161-S198.
- Shen, Y., Sun, W., & Gu, Y. (2023). College enrollment expansion, overeducation and career mobility. *Finance & Trade Economics*, 44(4), 157-172. (in Chinese)
- Sicherman, N. (1991). "Overeducation" in the labor market. *Journal of labor Economics*, 9(2), 101-122.
- Sleskova, M., Salonna, F., Geckova, A. M., Nagyova, I., Stewart, R. E., van Dijk, J. P., & Groothoff, J. W. (2006). Does parental unemployment affect adolescents' health?. *Journal of Adolescent Health*, 38(5), 527-535.
- Tarvid, A. (2015). The effectiveness of access restriction to higher education in decreasing overeducation. *Economic Analysis and Policy*, 45, 11-26.
- Tsang, M. C. (1987). The impact of underutilization of education on productivity: A case study of the US Bell companies. *Economics of education review*, 6(3), 239-254.
- Xing, C., Yang, P., & Li, Z. (2018). The medium-run effect of China's higher education expansion on the unemployment of college graduates. *China Economic Review*, 51, 181-193.
- Yang, J., & Qiu, M. (2016). The impact of education on income inequality and intergenerational mobility. *China Economic Review*, 37, 110-125.
- Zheng, Y., Zhang, X., & Zhu, Y. (2021). Overeducation, major mismatch, and return to higher education tiers: Evidence from novel data source of a major online recruitment platform in China. *China economic review*, 66, 101584.
- Zumbuehl, M., Dohmen, T., & Pfann, G. (2021). Parental involvement and the intergenerational transmission of economic preferences, attitudes and personality traits. *The Economic Journal*, 131(638), 2642-2670.