Kyklos

WILEY

KYKLOS

ORIGINAL ARTICLE OPEN ACCESS

Parental Gender Stereotypes and Student Well-Being in China

Shuai Chu[1,2] | Xiangquan Zeng[1,3] | Klaus F. Zimmermann[1,4,5,6]

[1]Global Labor Organization (GLO), Essen, Germany | [2]Occupation and Skills Research Office, Chinese Academy of Labour and Social Security, Beijing, China | [3]School of Labor and Human Resources, Renmin University of China, Beijing, China | [4]POP – Centre for Population, Development and Labour Economics, UNU-MERIT, Maastricht, the Netherlands | [5]School of Business and Economics, Maastricht University, Maastricht, the Netherlands | [6]Centre for Economic Policy Research (CEPR), London, UK

Correspondence: Klaus F. Zimmermann (klaus.f.zimmermann@gmail.com)

Received: 30 May 2023 | **Revised:** 2 September 2024 | **Accepted:** 29 September 2024

Keywords: gender identity | gender stereotypes | math stereotype | mental health | student well-being | subjective well-being

ABSTRACT

A prominent gender stereotype claims that "boys are better at learning mathematics than girls." Confronted with such a parental attitude, how does this affect the well-being of 11- to 18-year-old students in Chinese middle schools? Although well-being has often been shown to be not much gender-diverse, the intergenerational consequences of such stereotypes are not well studied. Expecting too much from boys and too little from girls might damage self-esteem among school kids. Using large survey data covering districts all over China reveals that one-quarter of the parents agree with the math stereotype. It is shown that this has strong detrimental consequences for the offspring's well-being. Students are strongly more depressed, feeling blue, unhappy, not enjoying life, and sad with no male–female differences, whereas parental education does not matter for this transfer. Various robustness tests including other than math stereotypes and an IV analysis confirm the findings. Moderating such effects, which is in line with societal objectives in many countries, not only supports gender equality but also strengthens the mental health of children.

JEL Classification: I12, I20, I31, J16

1 | Introduction

Whether subjective well-being (SWB) is affected by gender is debatable and previous findings in the literature have been inconclusive (Meisenberg and Woodley 2015; Batz and Tay 2018; Nikolova and Graham 2022). Studies have found stronger or lower effects for females or even no differences when properly controlled for other relevant factors. This may be because the evidence for genetic differences is weak and the observed associations have to be understood in complex and diverse social contexts.

This points to the relevance of identities, attitudes, norms, and stereotypes, which have been the concern of significant recent literature in economics (Akerlof and Kranton 2000; Alesina, Giuliano, and Nunn 2013; Carlana 2019; Bursztyn, González, and Yanagizawa-Drott 2020; Mishra and Parasnis 2022). Parental beliefs significantly influence the outcomes of their children. For instance, Chou (2022) demonstrated that beliefs in Chinese zodiac superstitions affect the educational attainment of offspring. According to Kuhn and Wolter (2023), adolescents' occupational aspirations conform to gender stereotypes in local communities and exhibit intergenerational persistence.

Adverse childhood experiences may have detrimental effects on later life SWB in children (Kelifa et al. 2021). Gender stereotypes may cause gender differences in SWB when the generated pressures lead men and women to feel and express their emotions differently (Nolen-Hoeksema and Rusting 2003). However, the SWBs for both genders could also react in the same (negative) way if the asymmetric impact of gender stereotypes on females and males causes similar emotional damage. For instance, the stereotype "boys are better in math than girls" may be a burden for girls and boys, for instance, if this is wrong for both, individually. Education seems to shape the way in which egalitarian gender role attitudes and behaviors are developing (Du, Xiao, and Zhao 2021).

Mathematics has long been considered to be a male-dominated subject (Forgasz, Leder, and Kloosterman 2004); therefore, there are significant gender differences in parental expectations of children's math achievement, and these differences can lead to math anxiety (Bieg, Goetz, and Wolter 2015). The math stereotype may cause both parents and teachers to have high expectations for boys and low expectations for girls (McCoy, Byrne, and O'Connor 2022). This is probably challenging for both boys and girls; girls may lower their own expectations, even though they may be doing better in math than boys, which is an important channel for girls to develop math anxiety (Pennington et al. 2021; Muntoni and Retelsdorf 2019). For boys, meeting their parents' expectations is critical, especially in Asian cultures, where boys are often considered the center of the family, and they can experience significant psychological stress if they are unable to meet their parents' expectations (Wang and Heppner 2002). A more noteworthy point is that Asian parents tend to believe that the difference in academic performance can be compensated to some extent through effort (Yamamoto and Holloway 2010); therefore, when children's performance does not meet their parents' expectations, parents may restrict their sleep and rest time (Yang and Shin 2008).

Our contribution to this debate is to focus on the intergenerational association that *parental gender stereotypes* may show for the SWB of their children and how this transfer is associated with *parental education*. These stereotypes can be associated with SWB, even in the absence of gender differences. Using the largest national education survey, the China Education Panel Survey (CEPS), we study the role of parental gender stereotypes and parental education on student SWB in China. We are able to use parental binary responses to the question: "Do you think boys are better at learning mathematics than girls?"

This is novel in various ways for research on well-being in China and beyond. Although there is substantial research on happiness and SWB in China in general (Chen and Davey 2008; Davey and Rato 2012) and their respective determinants (for instance, Tani 2017 documents the role of receiving *hukou* status; Ding, Salinas-Jiménez, and del Mar Salinas-Jiménez 2021 and Yang, Lu, and Li 2023 examine income inequality; and Nie, Sousa-Poza, and Nimrod 2017 investigate internet use), we are the first to study the impact of parental gender stereotypes on student well-being.[1] The SWB of students has been of growing interest in the international literature including psychological, cognitive, and social components (Zhang and Renshaw 2020; Tobia et al. 2019), but these papers have not studied the consequences of parental gender stereotypes. Hence, our paper can inform the international debate about an issue of global concern, the potential damage caused by gender stereotypes in practice in many countries.

We find that parental gender stereotypes are associated with a strongly lower student well-being. However, while dealing with a gender-specific issue there are no gender-specific differences: Neither does it matter who the parent interviewed is (the mother or the father), nor whether the child concerned is a boy or a girl. This is in line with previous findings that there are no gender differences in well-being. Parental stereotypes may undermine girls' self-confidence and make them more prone to anxiety and other mental health issues. For boys, stronger stereotypes may indicate higher expectations and pressures, which may also generate negative emotions. Additionally, parental human capital is not associated with offspring's well-being.

Section 2 explains the dataset and the institutional background in China, including a review of the gender equality situation of the country. It also provides descriptive evidence of the data used. Section 3 presents the econometric model and first regression results, identifying the core findings of the study. Section 4 imposes a model simplification suggested by the results in Section 3 and discusses the robust findings and evaluates their significance and relevance. Section 5 provides further intensive robustness analyses confirming the key message, and Section 6 explores causality issues. Section 7 reviews the contributions of this study and concludes.

2 | Data and Institutional Background

The dataset is the secondary use of the China Education Panel Survey (CEPS), the first and largest national representative education survey.[2] The CEPS adopted multistage probability-proportional-to-size sampling (PPS). Mainland China (excluding Hong Kong, Macao, and Taiwan) has 31 provincial-level units and 2870 county/district-level units covering the population of China. Taking the average education level of the population and the proportion of the floating population as stratified variables in the sampling design, the survey covers middle schools from 28 counties and city districts, in which four middle schools and four classrooms were selected to represent a given county or urban area.

China's administrative system is a regionally decentralized authoritarian system. The central government has control over personnel, whereas subnational governments operate the bulk of the economy and initiate, negotiate, implement, divert, and resist reforms, policies, rules, and laws. China's reform trajectories have been shaped by regional decentralization. Therefore, local governments have considerable independence and autonomy (Xu 2011). Household registration is governed by China's unique *hukou* regulation, which operates similarly to an internal passport system (Afridi, Li, and Ren 2015).

School education in China is generally divided into three stages: primary school, secondary school, and university or college. Secondary school is divided into middle school and senior middle school. Middle school is the stage of transition to senior middle school, which belongs to the category of secondary education. In China, children usually enter primary school at the age of six and middle school after 6 years of study. The middle school system covers 3 years: Grade 7 (Year 1), Grade 8 (Year 2), and Grade 9 (Year 3). The data that we used included seventh- and ninth-grade students.

Chinese political authorities are concerned about gender equality. Chairman Mao's 1955 call for "women can hold half sky" was seen as support for equal rights for men and women (Mow, Jie, and Bijun 2004). In October 2020, President Xi also confirmed that "equality between men and women is China's basic state policy" (Xi 2020). In order to protect the rights and interests of women in the labor market, China had already introduced in 1988 the "Regulations Concerning the Labor Protection of Female Staff and Workers."[3] This administrative regulation has 14 specific measures, such as, "Any Unit which is suitable for women to engage in Labor may not refuse to employ female staff and workers," "During the pregnancy, maternity leave and nursing period of female staff and workers, their basic salaries may not be reduced and their Labor contracts may not be cancelled," "The time spent by pregnant female staff and workers on antenatal examination during Labor hours shall be deemed to be Labor hours." A law of the "People's Republic of China on the Protection of Women's Rights and Interests," established in 1992 and amended in 2018, states that women are to be guaranteed the same status as men in politics, education, the labor market, property, personal security, and marriage.

In order to enhance the status of women in science, China promulgated the "Opinions on Strengthening the Construction of Female Scientific and Technological Talents" in 2011 and "Several Measures on Supporting Female Scientific and Technological Talents to Play a Greater Role in Scientific and Technological Innovation" in 2021 to support the development of women in scientific research in a comprehensive manner. China has also established the governmental organization of the All-China Women's Federation as well as the non-governmental organization of the Women's Congress (Bohong 2020) to monitor the protection of women's rights. In order to publicize the performance of women in the labor market, there is also a special award for the "Woman Pace-Setter," the winner of which will be publicized throughout the country as a role model.

As seen in the literature, China's socialist revolution has remarkably improved women's socioeconomic status in absolute terms and relative to men, which is reflected in women's educational attainment (Li 2016), life expectancy, labor force participation (Liu 2018), and political participation (All China Women's Federation [ACWF] 2019). However, some researchers have also pointed out that gender discrimination is still prevalent in China, especially in recruitment and executive positions (Zhang et al. 2021). Therefore, according to the current state of the literature, it is difficult to determine whether gender inequality in China has increased or decreased. In addition, some studies hold that the concept of gender roles in China is developing in a more open direction, which means that society's attitude towards women is more inclusive. From this perspective, gender differences should be developing in a better direction (Du, Xiao, and Zhao 2021).

This suggests a strong interest in investigating the prevalence of gender stereotypes and their impact on offspring's well-being. Our student data measure students' well-being and parental math stereotypes. The data were randomly collected for two samples, the *mother sample* (5364 students) and the *father sample* (5073 students) with a total of 10,437 students, including 5407 girls. Each individual in both samples consists of a student and a parent questionnaire. Students covered are 11–18 years old.

The student questionnaires report the following feelings in the last 7 days in the range "1 = never," "2 = seldom," "3 = sometimes," "4 = often," and "5 = always": "depressed," "feeling blue,"[4] "unhappy," "not enjoying life," and "sad," with respect to categories "unhappy," "depressed," "sad," "feeling blue" and "not enjoying life" for both mother and father samples. These are standard "well-being measures" used in the literature and reflect "negative" well-being or misery. Table 1 summarizes student well-being in the full (combined mother and father) sample. Misery ("often," "always") is low (in the range of 10%–15%) across all indicators, and "always" has consistently the lowest percentage. The other three categories ("never," "seldom," and "sometimes") are much more diverse. Focusing on the maxima in each category (rows in Table 1), it is "sometimes" (41.15%) for "depressed," "seldom" (33.46%) for "feeling blue," "sometimes" (37.64%) for "unhappy," "never" (46.08%) for "not enjoying life" (therefore, actually every second student is enjoying life), and "seldom" (36.64) for feeling "sad." The last column in Table 1 contains the mean index in all five well-being categories with the lowest value (1.96) for "not enjoying life" and the highest value (2.54) for "unhappy"; the mean for the category "depressed" (2.48) is close to that for "unhappy," whereas feeling blue (2.22) and "sad"

TABLE 1 | Student well-being ("misery") in the full sample.

	Never: 1	Seldom: 2	Sometimes: 3	Often: 4	Always: 5	Mean index
Depressed	16.41	32.67	41.15	6.45	3.32	2.48
Feeling blue	29.17	33.46	26.54	7.33	3.51	2.22
Unhappy	14.68	34.24	37.64	8.94	4.50	2.54
Not enjoying life	46.08	26.83	16.96	5.64	4.49	1.96
Sad	24.64	36.64	28.73	6.05	3.94	2.28

Note: (1) To measure students' well-being, we use student responses to questionnaire items. Specifically, five questions asked students about the frequency of the following feelings during the previous 7 days on a scale from 1 (never) to 2 (seldom), 3 (sometimes), 4 (often) to 5 (always): (a) depressed, (b) blue, (c) unhappy, (d) not enjoying life, or (e) sad. (2) Numbers are the percentage (%) for each well-being measure in the combined mother and father sample. (3) The last column contains the mean value of each misery index.

(2.28) rank in between. The marked variety between the well-being (or misery) indicators suggests analyzing the indicators separately.[5]

How different are the subsamples of mothers and fathers? A detailed analysis comparing the outcomes for both parents is provided in Table 2. For each student, only one parent (father or mother) was interviewed randomly. Student girls in both samples have mostly a smaller mean than boys, but the differences are very small. This table (last panel "sample differences") also reveals that the mother and father samples do not differ according to the provided difference *t*-tests.

The two key variables we focus on in our investigation are parental gender (math) stereotypes and parental education, controlling for a larger number of students and parental characteristics. The parent questionnaires contain responses on "*Do you think boys are better at learning mathematics than girls?*" (*1 = yes*; *0 = no*). We treat "yes" as parental gender (math) stereotype. Further, we use "*years of schooling*" to measure the education of either the father or the mother according to the following rules: "0 = no education," "6 = primary school," "9 = middle school," "12 = high school," "15 = college," "16 = undergraduate," and "19 = graduate." The father and mother samples contain educational information for both the father and mother of the student, and the parental gender (math) stereotype is only available for the parent of the respective sample.

Parental gender (math) stereotypes are somewhat but not markedly different between mothers and fathers. Details can be seen in Table 3 in the first rows of Panel A: Mother sample and Panel B: Father sample: 25.6% of mothers, but only 24.3% of fathers have the stereotype. This difference disappears if the student is a boy (27.8% for mothers against 27.5% for fathers) but is more marked if the child is a girl (23.8% for mothers against 20.9 for fathers). The stereotype is more prevalent among parents with a male child than among those with a female child; it is also more prevalent among mothers with a girl than among fathers with a girl. However, the differences between the mother and father samples are small.

The distribution of the gender (math) stereotypes across Chinese provinces is shown in Figure 1, where the horizontal axis indicates the mean value of the stereotype in a particular province, and the numbers attached reveal the respective means and standard deviations. Patterns for the father and mother samples are broadly similar, but there is some significant variation across the provinces. About a quarter of all parents agree on the gender (math) stereotypes. This is a significant size that makes this investigation of high potential importance. There appears to be significant variation in the data.

This is also clearly revealed in Figure 2, which shows the distribution of stereotypes by province on a map. Interestingly, differences on the macro level associate with regional gross domestic product (GDP): Stereotypes show a "U" curve, they are stronger in regions with both lower and higher GDP per capita (see Figure 3). This could be related to the respective industry structure and the relative position of men in the Chinese labor market during economic development. Lower GDP per capita provinces are mainly in agriculture and industry, and developed regions are mainly in production-oriented services (e.g., finance and scientific research).

The variable "years of schooling" is available in both (mother and father) samples and can be compared for consistency

TABLE 2 | Student well-being ("misery") in parent samples.

	Mother sample			Father sample			Sample differences		
	Full sample	Girl	Boy	Full sample	Girl	Boy	Full sample	Girl	Boy
Well-being	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**
Depressed	2.483	2.506	2.454	2.468	2.505	2.434	−0.014	−0.001	−0.020
	(0.937)	(0.902)	(0.979)	(0.966)	(0.913)	(1.012)	(−0.777)	(−0.04)	(−0.708)
Blue	2.215	2.165	2.278	2.236	2.158	2.309	0.021	−0.006	0.031
	(1.045)	(1.021)	(1.071)	(1.068)	(1.040)	(1.088)	(1.017)	(−0.217)	(1.003)
Unhappy	2.534	2.518	2.554	2.553	2.555	2.551	0.018	0.037	−0.004
	(0.979)	(0.939)	(1.027)	(1.012)	(0.984)	(1.038)	(0.946)	(1.407)	(−0.125)
Not enjoying life	1.940	1.871	2.026	1.974	1.880	2.061	0.033	0.008	0.035
	(1.108)	(1.058)	(1.162)	(1.136)	(1.086)	(1.175)	(1.513)	(0.289)	(1.060)
Sad	2.267	2.259	2.278	2.293	2.281	2.305	0.026	0.022	0.027
	(1.011)	(0.974)	(1.055)	(1.041)	(1.007)	(1.071)	(1.294)	(0.823)	(0.893)
Observations	5364	2965	2399	5073	2442	2631	10,437	5407	5030

Note: (1) To measure students' well-being, we use student responses to questionnaire items. Specifically, five questions asked students about the frequency of the following feelings during the previous 7 days on a scale from 1 (never) to 2 (seldom), 3 (sometimes), 4 (often) to 5 (always): (a) depressed, (b) blue, (c) unhappy, (d) not enjoying life, or (e) sad. (2) This table reports the summary statistics and the difference between the mother sample and the father sample in students' well-being. In Columns 1–6, the numbers indicate the mean of the variables, and the numbers in parentheses indicate the standard deviation of the variables. In Columns 7–9, numbers are differences of variables between both parent samples, and the numbers in parentheses are *t*-statistics.

TABLE 3 | Descriptive statistics of independent variables and control variables.

		Full sample		Girl		Boy	
		Count	**Mean/(*SD*)**	**Count**	**Mean/(*SD*)**	**Count**	**Mean/(*SD*)**
Panel A: Mother sample							
Parent mother	Stereotype	5338	0.256/(0.437)	2954	0.238/(0.426)	2384	0.278/(0.448)
	Years of schooling	5359	10.072/(3.275)	2962	9.990/(3.239)	2397	10.173/(3.318)
	Age	4574	39.511/(4.076)	2555	39.430/(4.008)	2019	39.614/(4.159)
	Hukou	5137	0.458/(0.498)	2860	0.435/(0.496)	2277	0.487/(0.500)
	Health	5210	3.814/(0.908)	2893	3.832/(0.905)	2317	3.792/(0.912)
	Occupation	5039	1.333/(0.800)	2775	1.321/(0.796)	2264	1.348/(0.805)
Parent father	Years of schooling	5359	10.414/(3.162)	2962	10.386/(3.084)	2397	10.449/(3.256)
	Occupation	5007	1.509/(0.771)	2763	1.485/(0.746)	2244	1.538/(0.801)
Individual students	Girl	5364	0.553/(0.497)	2965	-—	2399	-—
	Academic ranking in primary school	5018	15.837/(11.864)	2775	14.116/(11.007)	2243	17.966/(12.525)
	Hukou	5364	0.516/(0.500)	2965	0.530/(0.499)	2399	0.499/(0.500)
	Age	5266	13.812/(1.265)	2927	13.789/(1.284)	2339	13.840/(1.240)
	Attend kindergarten	5321	0.818/(0.386)	2953	0.826/(0.380)	2368	0.809/(0.393)
	Age when starting primary school	5308	6.512/(0.939)	2940	6.512/(0.916)	2368	6.512/(0.967)
	Family's financial situation	5188	0.914/(0.280)	2868	0.917/(0.276)	2320	0.911/(0.285)
Panel B: Father sample							
Parent father	Stereotype	5042	0.243/(0.429)	2434	0.209/(0.407)	2608	0.275/(0.446)
	Years of schooling	5070	10.015/(2.964)	2440	10.065/(2.944)	2630	9.968/(2.982)
	Age	4169	41.291/(4.723)	2093	41.312/(4.739)	2076	41.27/(4.708)
	Hukou	4807	0.376/(0.484)	2334	0.387/(0.487)	2473	0.366/(0.482)
	Health	4913	3.825/(0.938)	2404	3.849/(0.937)	2509	3.802/(0.938)
	Occupation	4749	1.446/(0.732)	2280	1.447/(0.710)	2469	1.445/(0.753)
Parent mother	Years of schooling	5070	8.729/(3.543)	2440	8.733/(3.546)	2630	8.725/(3.541)
	Occupation	4683	1.264/(0.698)	2248	1.284/(0.684)	2435	1.246/(0.710)
Individual students	Girl	5073	0.481/(0.500)	2442	-—	2631	-—

(Continues)

TABLE 3 | (Continued)

	Full sample		Girl		Boy	
	Count	**Mean/(SD)**	**Count**	**Mean/(SD)**	**Count**	**Mean/(SD)**
Academic ranking in primary school	4662	16.362/(11.936)	2221	14.457/(11.094)	2441	18.095/(12.405)
Hukou	5073	0.606/(0.489)	2442	0.598/(0.490)	2631	0.614/(0.487)
Age	4968	14.063/(1.380)	2405	14.000/(1.372)	2563	14.121/(1.385)
Attend kindergarten	5038	0.772/(0.420)	2432	0.782/(0.413)	2606	0.762/(0.426)
Age when starting primary school	5009	6.488/(0.967)	2420	6.512/(0.928)	2589	6.465/(1.002)
Family's financial situation	4871	0.886/(0.318)	2366	0.893/(0.309)	2505	0.878/(0.327)

Note: (1) "Parent's stereotype" is 1 if the answer of mother or father is "Yes" when asked: "Do you think boys are better at learning mathematics than girls?" (2) We use years of schooling to represent the education of either the father or the mother, defined according to the following rules: no education = 0; primary school = 6; middle school = 9; high school = 12; college = 15; undergraduate = 16; graduate = 19. (3) Age indicates the age of the student and the student's father or mother. (4) When the student individual or the student's father or mother has an agricultural household registration, Hukou = 1; otherwise, Hukou = 0. (5) Health is sorted from 1–5, 1 = very unhealthy, to 5 = very healthy. (6) We define Occupation = 0 if the parent has no occupation; Occupation = 1 if parents engaged in skilled workers, general workers in manufacturing or service industries and farmers; Occupation = 2 if parents engaged in teachers, engineers, doctors, lawyers, and individual business activities; Occupation = 3 if parents engaged in leadership or management positions. (7) The variable "Girl" indicates the gender of the student. (8) Academic ranking in primary school reflects the relative ranking of students' academic performance in their classes when they are in primary school. If the score is the best, the value is 1. The higher the value, the worse the students' academic performance in primary school. (9) Students who have attended kindergarten, Attend kindergarten = 1; otherwise, Attend kindergarten = 0. (10) Age when starting primary school reflects the age at which students enter primary school. (11) If the family does not receive the subsistence allowance at present, then the family's financial situation = 1; otherwise = 0.

(second and seventh rows in both panels of Table 3). In the mother sample, the mother has 10.1 years of schooling, and the father has 10.4 years of schooling; in the father sample, the mother has 8.7 years of schooling, and the father has 10.0 years of schooling. The schooling levels appear to be only marginally different for child gender within the four parent groups. Furthermore, occupation (see Table 3) is available for both parents in both (mother and father) samples. Again, the variable means between fathers (and mothers) in both samples are of similar size. Here, occupation is measured as occupational rank with values "0 = parent has no occupation," "1 = parent engaged in skilled work, general workers in manufacturing or service industries and farmers," "2 = parents engaged as teachers, engineers, doctors, lawyers and with individual business activities," and "3 = parents engaged in leadership or management positions." With respect to "years of schooling" and "occupation," the two samples (mother and father) are very similar. The exception is "years of schooling" of the mother which is 10.1 years in the mother sample and 8.7 years in the father sample.[6]

Nonoverlapping further controls for parents in both samples are "age" of the parent, his/her hukou ("1 = yes"), and his/her "health" sorted from 1 to 5 with "1 = very unhealthy" to "5 = very healthy." These data are only available for fathers in the father sample and for mothers in the mother sample but may be important for control purposes. Further controls are available and used at the student level. They include "gender" of the student ("1 = girl"; 0 otherwise), hukou ("1 = yes"; 0 otherwise), academic ranking in primary school ("rank number"), "has attended kindergarten" ("1 = yes"; 0 otherwise), "age" in years, "age when starting primary school" in years, and family's "financial situation "0 = receive subsistence allowance at present "; 1 otherwise). In general, when a family receives a subsistence allowance, the financial situation is poor. Descriptive statistics for all variables are provided in Table 3. The student controls in the two samples have very similar means; the exceptions are gender (55.3% girls in the mother sample against 48.1% in the father sample) and hukou (51.6% in the mother sample against 60.6% in the father sample).

3 | Model Specification and Regression Results

The student well-being measures Y ("depressed," "feeling blue," "unhappy," "not enjoying life," and "sad") are explained by a set of parental and student characteristics as explained in the previous section and listed in Table 3. Since the focus in this section is on parental gender stereotypes and education, the other variables are just seen as controls. They are presented and discussed in detail in Section 4. For the analysis, the two (father and mother) samples were merged, resulting in a full sample size of 6962 observations of students and their parent, where all the variables were observed.

The dummy regression specification is developed in a way that allows for direct tests for differences between the two samples and between children's gender and their interactions in one regression for each well-being measure.[7] The regression specification used is as follows:

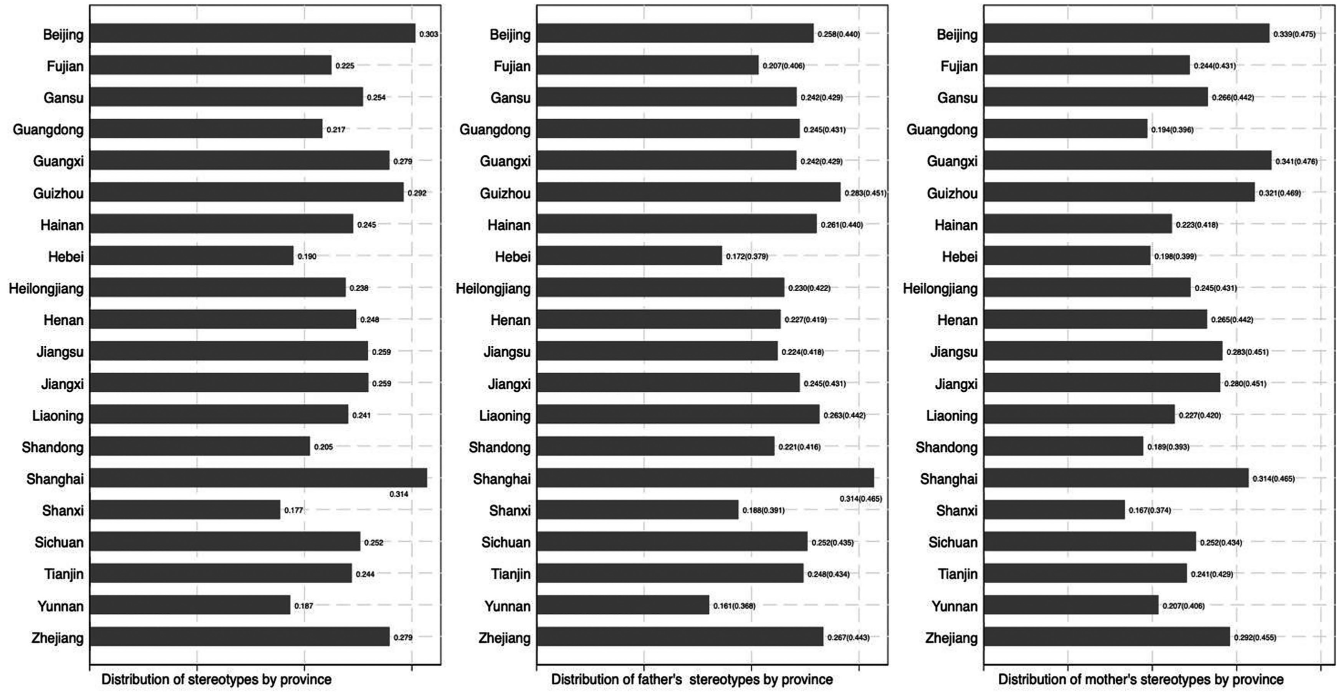

FIGURE 1 | Distribution of parental stereotypes by Chinese provinces. *Note:* The parent questionnaires ask "Do you think boys are better at learning mathematics than girls?" Responses (1 = yes; 0 = no) as a measure of parental gender stereotype. A quarter of the parents in all samples agree. Numbers are stereotype means with standard deviations in parentheses in the respective provinces. *Source:* 2014 China Education Panel Survey (CEPS), own calculations.

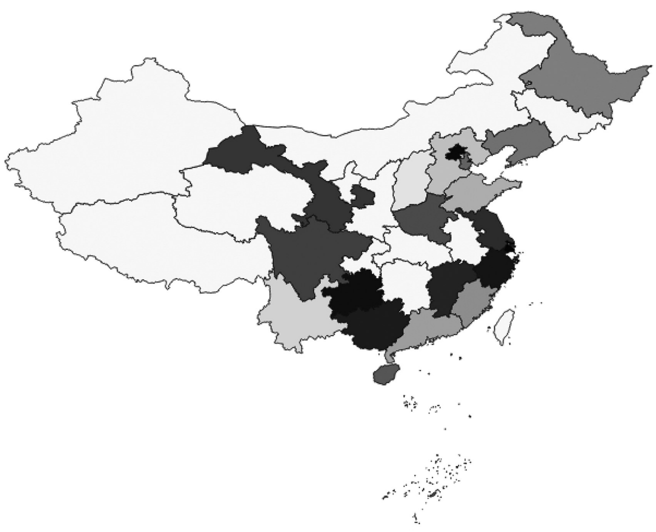

FIGURE 2 | Geographic distribution of gender stereotypes by Chinese provinces. *Note:* White parts indicate missing data, and gender stereotypes are the stronger the darker the color. *Source:* 2013–2014 China Education Panel Survey (CEPS), own calculations.

$$Y = \alpha + a_1 M + a_2 G + a_3 (M G) + b S + b_1 (S M) + b_2 (S G) + b_3 (S M G) + c_1 E_f + c_2 (E_f M) + c_3 (E_f G) + c_4 (E_f M G) + c_5 E_m + c_6 (E_m M) + c_7 (E_m G) + c_8 (E_m M G) + d X + d_1 (X M) + d_2 (X G) + d_3 (X M G) + \varepsilon \quad (1)$$

M and G are (0,1)—dummies, where M stands for mother sample and G for girl student; S stands for parental gender stereotypes (either mother or father where available); E_f and E_m are father and mother years of schooling, respectively; further, as controls: X refers to other parental or student characteristics. α is the intercept, and ε is the error term. Parameters affiliated with terms involving M and/or G measure sample differences. If they are not statistically significant, this suggests that the core effects are robust.

The results for the five well-being measures are presented in Table 4. The parameter estimates for M, G, and $M*G$ are all insignificant, with the exception of $M*G$ for "sad," implying no

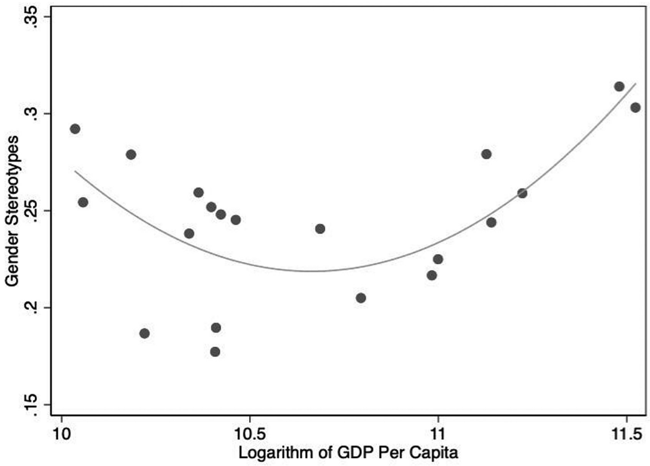

FIGURE 3 | Provincial GDP per capita in China and gender stereotypes. *Source:* 2013–2014 China Education Panel Survey (CEPS), own calculations. GDP per capita data in 2013 from the National Bureau of Statistics of China. https://data.stats.gov.cn/easyquery.htm?cn=C01.

overall average differences between the father and the mother samples, and with respect to gender differences among the students. Only girls in the mother sample feel on average statistically significantly more sad. Education of both parents (E_f and E_m) has no impact on child well-being; this is a very robust finding. Not only the direct overall effect parameters of E_f and E_m are not statistically different from 0, but there are also no significant differences across the examined subgroups. These observations and exceptions are worth mentioning: The estimated direct common parameters for E_m (mother's education) for boys and girls are all negative (besides for "feeling blue") and significantly negative at the 10% level for "sad." E_m has also a strong and statistically significant negative effect on "feeling blue" among girls. Hence, mother's education has some positive elements for student well-being.

The key issue of this study is the expected effect of parental gender stereotypes on student-kid well-being. In principle, the effects could be gender-different among children and for both parents. Table 4 provides a direct test of the potential differences. The finding for the Chinese families is surprisingly simple, sizable, statistically significant, and robust: There is a unique common parental stereotype effect that impacts all five well-being measures in a similar range. This result can be described as a *quasi-Solitary North Star* finding, indicating the strong statistical significance of the estimated effect. In general, there are no parental differences or differences with respect to student's gender. The only exception is a statistically significant negative parameter estimate for girls in the mother sample, indicating a smaller level of well-being damage for this student subgroup.

To summarize, the regression results in Table 4 show that the coefficients of S in Equation 1 are significantly positive and all of the same size. This suggests that the parental gender math stereotype significantly elevates the level of students' negative emotions, thereby reducing their well-being regardless of whether the student is a boy or girl, and whether the stereotype originates from the mother or father.

4 | Model Simplification and Detailed Findings

The analysis in the previous section has revealed that student well-being measured with five misery categories (beyond the effects of the controls) is associated with the main effect of parental gender math stereotypes only. Student gender, parental sample, and parental education do not seem to matter overall. This section explores the robustness of these findings. A first step is to explore the statistical relevance of the reduced version of Equation 1, eliminating irrelevant components to simplify further analysis:

$$Y = \alpha + b\,S + d\,X + \varepsilon. \tag{2}$$

Table 5 contains the findings of this new baseline model: The stereotype parameter estimates are broadly unchanged, and the R^2 measures are hardly smaller, supporting the simplification. Therefore, we continue the further robustness analyses with Equation 2 as the baseline or reference model.

The misery impacts of parental stereotypes are now simply captured by b. The coefficient is largest for "feeling blue" (0.500), followed by "unhappy" (0.483), "depressed" (0.440), "not enjoying life" (0.425), and "sad" (0.402). Its size evaluated at the sample mean of those students with no parental stereotypes Y_0 constructed as b/Y_0 times 100 ($b/Y_0\%$) is 18.6% for the category "depressed," 24% for "blue," 20% for "unhappy," 23% for "not enjoying life," and 18.5% for "sad." Parental stereotypes exhibit not only highly significant parameter estimates at the 1% level, but the parameters reflect also a strong contribution to student misery. According to $b/Y_0\%$, this contribution is largest for categories "blue" and "not enjoying life."

Furthermore, all regressions contain parameter vector d for parent controls (Age, Hukou, Health, and Occupation—ordered, rising with higher social rank) and student controls (Academic ranking in primary school, Hukou, Age, Attend kindergarten, Age when starting primary school, and Family's financial situation) based on the respective survey answers of parents and

TABLE 4 | Regression results for Equation 1 determining student well-being ("misery").

	Depressed (1)	Blue (2)	Unhappy (3)	Not enjoying life (4)	Sad (5)
Mother	−0.322	−0.698	−1.094	0.018	−0.609
	(0.626)	(0.857)	(0.667)	(0.683)	(0.536)
Girl	−0.381	−0.150	−0.408	0.234	−0.913
	(0.732)	(0.590)	(0.712)	(0.800)	(0.698)
Mother*Girl	1.059	0.985	1.467	−0.053	2.194**
	(1.028)	(1.002)	(1.014)	(1.081)	(0.873)
Stereotypes (math)	0.434***	0.518***	0.548***	0.466***	0.401***
	(0.061)	(0.064)	(0.055)	(0.076)	(0.044)
Stereotypes*Mother	−0.007	−0.083	−0.063	−0.116	0.013
	(0.074)	(0.090)	(0.066)	(0.082)	(0.061)
Stereotypes*Girl	0.086	0.103	−0.061	0.137	0.026
	(0.083)	(0.081)	(0.082)	(0.096)	(0.082)
Stereotypes*Mother*Girl	−0.101	−0.100	0.006	−0.191**	−0.072
	(0.098)	(0.109)	(0.097)	(0.093)	(0.097)
E_f	0.006	−0.001	−0.001	0.010	0.008
	(0.014)	(0.010)	(0.012)	(0.015)	(0.013)
E_f*Mother	0.004	−0.006	0.007	−0.019	−0.014
	(0.016)	(0.015)	(0.015)	(0.020)	(0.018)
E_f*Girl	−0.018	0.007	−0.007	−0.014	−0.010
	(0.020)	(0.015)	(0.018)	(0.020)	(0.019)
E_f*Mother*Girl	0.023	0.025	−0.008	0.021	−0.003
	(0.024)	(0.021)	(0.022)	(0.023)	(0.025)
E_m	−0.004	0.002	−0.002	−0.003	−0.016*
	(0.007)	(0.008)	(0.010)	(0.011)	(0.009)
E_m*Mother	0.000	0.002	−0.005	0.017	0.013
	(0.012)	(0.017)	(0.015)	(0.017)	(0.014)
E_m*Girl	0.008	−0.017**	−0.004	0.003	0.012
	(0.009)	(0.008)	(0.014)	(0.015)	(0.011)
E_m*Mother*Girl	−0.009	−0.005	0.009	−0.012	−0.008
	(0.014)	(0.015)	(0.020)	(0.018)	(0.017)
Constant	2.063***	1.806***	2.785***	1.360**	2.544***
	(0.452)	(0.491)	(0.443)	(0.506)	(0.417)
R^2	0.069	0.073	0.069	0.053	0.057

Note: (1) OLS estimates of Equation 1. (2) Number of observations = 6962. (3) Standard errors are robust and clustered at the city levels. (4) *p*-values. (5) All regressions contain parent controls (Parent's age, Parent's Hukou, Health; Occupation) and student controls (Academic ranking in primary school, Hukou, Age, Attend kindergarten, Age when starting primary school, and Family's financial situation) properly specified according to Equation 1. (6) Stereotypes are math stereotypes. ***$p < 0.01$, **$p < 0.05$, and *$p < 0.1$.

TABLE 5 | Regression results for Equation 2 determining student well-being ("misery").

	Depressed (1)	Blue (2)	Unhappy (3)	Not enjoying life (4)	Sad (5)
Stereotypes (math)	0.440***	0.500***	0.483***	0.425***	0.402***
	(0.030)	(0.036)	(0.037)	(0.040)	(0.030)
$b/Y_0\%$	18.551	23.998	19.950	23.048	18.491
Parent controls					
Parent Age	0.001	−0.001	−0.000	−0.005	−0.004
	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
Parent Hukou	−0.007	−0.025	0.028	0.004	0.002
	(0.059)	(0.051)	(0.053)	(0.054)	(0.066)
Parent Health	−0.075***	−0.086***	−0.082***	−0.059***	−0.070***
	(0.014)	(0.013)	(0.015)	(0.011)	(0.015)
Father's Occupation	0.004	0.041*	0.017	0.011	0.021
	(0.015)	(0.022)	(0.017)	(0.023)	(0.021)
Mother's Occupation	0.012	−0.020	−0.012	−0.000	0.003
	(0.018)	(0.020)	(0.013)	(0.019)	(0.018)
Student controls					
Rank in primary school	0.002*	0.004***	0.004***	0.006***	0.006***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Hukou	−0.109**	−0.100*	−0.032	−0.106*	−0.095*
	(0.046)	(0.051)	(0.049)	(0.059)	(0.046)
Age	0.048***	0.067***	0.038***	0.051***	0.038***
	(0.011)	(0.010)	(0.012)	(0.016)	(0.011)
Attend kindergarten	−0.043	−0.068	−0.020	−0.022	−0.080**
	(0.034)	(0.043)	(0.031)	(0.035)	(0.036)
Age when starting primary school	−0.024	−0.019	−0.037**	0.001	−0.015
	(0.014)	(0.015)	(0.016)	(0.013)	(0.015)
Family's financial situation	0.073**	0.062	0.057*	0.088	−0.045
	(0.031)	(0.044)	(0.029)	(0.056)	(0.039)
R^2	0.058	0.064	0.060	0.042	0.047

Note: (1) OLS estimates of Equation 2. (2) Number of observations = 6962. (3) Standard errors are robust and clustered at the city levels. (4) *p*-values. (5) The term $b/Y_0\%$ is the percentage increase of misery in the specific category when parents have math stereotypes with b the parameter estimate of stereotypes and Y_0 the mean of the misery category for kids with parents having no math stereotypes. (6) Stereotypes are math stereotypes. (7) The parent health index increases with better health of the parents, see footnote in Table 3 for more details.
***$p < 0.01$, **$p < 0.05$, and *$p < 0.1$.

students. Among the controls of the parents, only their reported health plays a highly statistically significant role at the 1% level for well-being categories with similar sizes of the estimated coefficients. Since the health index increases with the better health of the parents (see the footnote in Table 3 for more details), an improvement in parent health associates with a consistent reduction in student misery for all well-being categories.

Table 5 also contains parameter d for student controls. Academic ranking in primary school can be seen as a proxy for the potential performance of the students in school, and hence, is related to current student well-being. Its measurement in the survey reflects the relative ranking of students' academic performance in their classes when they were in primary school. If the score is best, the value is 1. The higher the value of "Rank" in Table 5, the worse the academic performance of students in primary school. The sizes of all the coefficients are small, although all the estimates have the expected positive sign and are statistically significant at the 1% level. Older students have larger misery indices for all categories, and the effect parameters are all statistically significant at the 1% level. Students with a Hukou have a smaller association with misery than those without a Hukou, but the significance level is only 5% or lower. This is plausible because a Hukou comes with local amenities such as healthcare and access to public schools. The other parameters of student controls, although mostly consistent in size and direction, have only a few statistically significant estimates: "Family's financial situation" exhibits a positive estimate for misery "depressed" (5% statistical significance level) and "unhappy" (1% statistical significance level). "Age when starting primary school" has a negative effect on the category "unhappy," and "attend kindergarten" has a negative effect on category "sad," both at a statistically significance level of 5%.

How relevant are the parental stereotype estimates in comparison with other relevant variables provided in Table 5? "Parent health" seems to be the best variable for a comparison since it has strongly significant and consistent estimates across misery types with the largest (absolute) coefficient for "feeling blue" (0.086), followed by "unhappy" (0.082), "depressed" (0.075), "sad" (0.070), and "not enjoying life" (0.059). Note that the sequence of these coefficients is largely the same as with the stereotype variable; only "sad" and "not enjoying life" have exchanged ranks. Further, "parent health" should be of value to children, the variable intuitively suggests relevance. And we have a similar type of measurement of the two variables: The math gender stereotype dummy variable moves from 0 to 1, and the parent health variable from 5 (very healthy) to 1 (very unhealthy). Consequently, the stereotype coefficients can be compared with 4 times the absolute value of the parent health coefficients. They are all not as large as the stereotype effects, which are $0.440/0.300 = 1.47$ as large for "depressed," 1.45 for "blue," 1.47 for "unhappy," 1.8 for "not enjoying life," and 1.44 for "sad." These are sizable effects, implying that parental stereotypes negatively affect student well-being about 50% stronger than the measured "parent health."

5 | Robustness

How robust are the findings for Equation 2 when sub-samples of student girls and boys are considered, which are either good or bad in math, leaving the middle group out? In a standard four-category scheme with A, B, C, and D and A as excellent, we can focus on samples for "good" for A + B and D for "bad." The usual practice in Chinese schools is to allocate the top 50% to "good" and the lowest 15% to "bad" (see the footnote of Table 6 for further details). Our dataset contains standardized math scores (with a maximum score of 100) for all students.[8] Applying the 50%/15% rule, we have generated four subsamples: girl–good math (1964 observations), girl–bad math (442 observations), boy–good math (1444 observations), and boy–bad math (581 observations). This implies that 54% of the girls are "good" in math, and only 12% are "bad," whereas among the boys, only 46% are good, but 18% are in the "bad" sample. Again, this is inconsistent with the gender-math stereotype story present in society. Overall, Chinese girls perform better in math than boys.

Replicating the baseline model (2) for the four subgroups, Table 6 confirms the general picture of the harmful nature of parental math stereotypes and provides more plausible results with additional insights: Girls with good math scores have worse well-being (the estimated coefficients are larger) than the total sample in the baseline. The same is true (with the exception of "blue") for boys with poor math scores. Both groups seem to suffer from not fulfilling the gender math stereotypes. Girls with bad math scores exhibit parameters for parental stereotypes, which are much lower than those for girls with good math scores (and the baseline full sample findings). They suffer less because they fulfill their expectations. Boys with good math scores still feel parental pressure; their estimates are close to those of the baseline and smaller than the estimates for boys with bad math scores. The exception is the misery category "blue," where boys with bad math scores have a smaller misery effect for "blue." We conclude that Table 6 confirms and extends the general story provided.

Table 7 reports the robustness checks when other stereotypes or teacher characteristics are added to the baseline. First, the well-being of the peers of the students is considered. Peer well-being is measured as the average well-being in the class of the student. The results show a consistent picture where the parental math stereotype parameter estimates are only slightly smaller in size, the association with peer well-being is positive and large, and all estimates are highly statistically significant at the 1% level. The same is true for students' own math stereotypes, where the estimates are all statistically significant at the 1% level, but the parameters are much smaller than those for parental math stereotypes. Again, the baseline findings are robust.

Table 7 also reports the inclusion of teacher effects in the baseline model. The additional variables are "homeroom teacher female," "homeroom teacher teaches math," and "math teacher is female." Estimates do not change the baseline findings, and only "math teacher is female" has a few statistically significant estimates with a positive size of about 0.07 (14% of a parental stereotype estimate of 0.5). Female math teachers contribute to the negative well-being (misery) of the students in the first three columns, namely, with "depressed," "blue," and "unhappy" but not with the rest ("not enjoying life" and "sad") in any relevant way (economically and statistically). Our general findings remain robust when teacher characteristics are included.

We further consider two other parental stereotypes in the context of student life (see Table 7): "Non-local students are harmful for the atmosphere of the school" (***parental migration stereotype—harmful***) and "teachers are fully responsible for the education of the students" (***parental education stereotype—teacher totally responsible***).[9] The parental harmful migration stereotype associates positively with most measures of (negative) student well-being of similar size and statistical significance, besides feeling "blue." The parental stereotype of teacher responsibility for education reduces parental pressures for their children in all well-being measures (besides "not enjoying life"), but the effect while exhibiting the right direction is only statistically significant at the 10% level for the well-being category "depressed." It is important to note that none of these estimates affect the coefficients of parental math stereotypes in comparison to the baseline in any relevant way. Again, our findings remain robust.

TABLE 6 | Student well-being ("misery"), stereotypes, and good/bad math scores subsamples.

	Depressed (1)	Blue (2)	Unhappy (3)	Not enjoying life (4)	Sad (5)
Baseline full sample (*Sample size: 6962*)					
Stereotypes	0.440***	0.500***	0.483***	0.425***	0.402***
	(0.030)	(0.036)	(0.037)	(0.040)	(0.030)
R^2	0.058	0.064	0.060	0.042	0.047
Girl–good math scores (*Sample size: 1964*)					
Stereotypes	0.540***	0.634***	0.552***	0.488***	0.425***
	(0.047)	(0.056)	(0.061)	(0.069)	(0.057)
R^2	0.077	0.082	0.069	0.049	0.062
Girl–bad math scores (*Sample size: 442*)					
Stereotypes	0.347***	0.349***	0.267**	0.234*	0.236
	(0.113)	(0.121)	(0.117)	(0.123)	(0.146)
R^2	0.081	0.060	0.050	0.067	0.036
Boy–good math scores (*Sample size: 1444*)					
Stereotypes	0.446***	0.530***	0.550***	0.474***	0.425***
	(0.057)	(0.075)	(0.069)	(0.071)	(0.065)
R^2	0.068	0.088	0.084	0.050	0.048
Boy–bad math scores (*Sample size: 581*)					
Stereotypes	0.543***	0.428***	0.668***	0.566***	0.490***
	(0.097)	(0.127)	(0.105)	(0.120)	(0.123)
R^2	0.057	0.046	0.069	0.045	0.039

Note: (1) OLS estimates of Equation 2. (2) Standard errors are robust and clustered at the city level. (3) *p*-values. (4) The sample split follows a usual practice in Chinese schools to allocate grades as A: 15%, B: 35%, C: 35%, 13% as D, and the last 2% is E (Liang et al. 2021). Using the standardized math scores for the students in our sample and ranking the top 50% of students as "good" and the lowest 15% as "bad," this implies math scores 72.85 and above for "good," and 60.4 and below for "bad." See Table A3 and Figure A1 for further descriptive information and the distribution of the standardized math scores in the sample. (5) Sample sizes are 6962 (total full sample); 3653 full girl samples with 1964 girl–good, 1247 girl–between, and 442 girl–bad; 3169 full boy samples with 1444 boy–good, 1144 boy–between, and 581 boy–bad. Due to missing data for the math score variable, the sum of the samples with this variable is 140 observations smaller. (6) Stereotypes are math stereotypes.
***$p<0.01$, **$p<0.05$, and *$p<0.1$.

Among student controls in the baseline estimates of Table 5, "rank" and "age" are the most relevant. How do additional interactions of these variables with the parental math stereotype variable perform and change these direct estimates? Such estimates are shown in Table A2, which exhibits the baseline estimates as well as the findings with both additional interactions and with both interaction variables separately. Parental math stereotype times student age has negative coefficients overall (besides "unhappy") but is only statistically significant for the category "sad" at the 10% level. Parental math stereotype times student ranking in primary school have throughout negative coefficients although only statistically significant for categories "blue" and "unhappy." The overall message is that the bad mental health impacts parental math stereotypes have on student well-being are reduced by age and rank in school. These findings are interesting and plausible, but they do not affect the overall problematic health implications of parental math stereotypes. Judging from the second panel of Table A2, by introducing both interaction terms at the same time, the direct stereotype coefficients are either of the same size as in the baseline (for the category "unhappy") or are up to twice the size.

Are these findings robust when alternative measures or treatments of well-being indicators are considered? This issue is investigated in Table 8. It repeats the baseline regression but now with different endogenous variables. The first is the self-assessed health of the student, and the second is a condensed variable of the five well-being measures using principal component analysis (PCA). PCA is commonly considered to be a statistical technique for data reduction (Kling, Liebman, and Katz 2007; Gong, Lu, and Song 2018). It helps to reduce the number of variables by describing a series of uncorrelated linear combinations of variables that contain most of the variance. We first conducted a correlation test on the five variables measuring student misery or well-being through the KMO (Kaiser–Meyer–Olkin) test. The result of the KMO test was

TABLE 7 | Student well-being ("misery"), parental math stereotypes and extra regressors: peer effects, student math stereotypes, teacher characteristics, and other stereotypes.

	Depressed	Blue	Unhappy	Not enjoying life	Sad
	(1)	(2)	(3)	(4)	(5)
Baseline full sample (*Sample size:6963*)					
Parental math stereotypes	0.440***	0.500***	0.483***	0.425***	0.402***
	(0.030)	(0.036)	(0.037)	(0.040)	(0.030)
R^2	0.058	0.064	0.060	0.042	0.047
With peer effects (*Sample size:6963*)					
Parental math stereotypes	0.378***	0.434***	0.418***	0.377***	0.345***
	(0.028)	(0.035)	(0.034)	(0.038)	(0.025)
Peer well-being (misery)	0.919***	0.922***	0.932***	0.899***	0.935***
	(0.023)	(0.027)	(0.029)	(0.028)	(0.026)
R^2	0.122	0.127	0.127	0.096	0.114
With student's own stereotype (*Sample size: 6831*)					
Parental math stereotypes	0.368***	0.450***	0.423***	0.369***	0.340***
	(0.031)	(0.035)	(0.039)	(0.038)	(0.032)
Student's own math stereotype	0.183***	0.152***	0.185***	0.147***	0.176***
	(0.034)	(0.031)	(0.033)	(0.034)	(0.026)
R^2	0.068	0.071	0.071	0.047	0.055
With teacher characteristics (*Sample size: 4961*)					
Parental math stereotypes	0.434***	0.506***	0.509***	0.443***	0.386***
	(0.033)	(0.038)	(0.039)	(0.049)	(0.038)
Homeroom teacher female	−0.014	−0.040	−0.046	0.029	−0.018
	(0.032)	(0.044)	(0.038)	(0.049)	(0.044)
Homeroom teacher teaches math	−0.012	−0.259	−0.072	−0.165	0.045
	(0.162)	(0.180)	(0.190)	(0.105)	(0.207)
Math teacher is female	0.070*	0.070*	0.077***	0.030	0.013
	(0.037)	(0.035)	(0.027)	(0.027)	(0.030)
R^2	0.061	0.067	0.068	0.047	0.045
With stereotypes of migrants (*Sample size: 6660*)					
Parental math stereotypes	0.435***	0.490***	0.477***	0.424***	0.394***
	(0.031)	(0.038)	(0.038)	(0.042)	(0.032)
Migration stereotype (harmful)	0.083**	0.052	0.086**	0.109**	0.116**
	(0.034)	(0.038)	(0.032)	(0.043)	(0.042)
R^2	0.058	0.064	0.060	0.042	0.047
With stereotype of teacher responsibility for education (*Sample size: 6817*)					
Parental math stereotypes	0.443***	0.502***	0.487***	0.427***	0.408***
	(0.031)	(0.036)	(0.037)	(0.040)	(0.030)
Teacher totally responsible	−0.058*	−0.052	−0.062	0.058	−0.048
	(0.032)	(0.036)	(0.054)	(0.048)	(0.041)
R^2	0.060	0.065	0.062	0.043	0.048

Note: (1) OLS estimates of Equation 2 with various additional regressors (see Table A1 for descriptive statistics). (2) Standard errors are robust and clustered at the city level. (3) *p*-values. (4) ***Peer well-being (misery)***: Average well-being (misery) of classmates. (5) ***Student's own stereotype***: Equal to 1, if the answer of student is "Yes" when asked: "Do you think boys are better at learning mathematics than girls?" and equal to 0 otherwise. (6) ***Migration stereotype (harmful)***: Equal to 1 if parent responded "harmful" to "What kind of effect do you think will the increase of students from nonlocal county/district have on the atmosphere of the school?" and 0 otherwise. (7) ***Teacher totally responsible***: Equal to 1 if parent answered to "Do you agree that it is totally the teachers' responsibility to educate children" by "Somewhat agree" or "Strongly agree," and 0, otherwise.
***$p<0.01$, **$p<0.05$, and *$p<0.1$.

TABLE 8 | Parental math stereotypes, self-assessed health, and misery index.

	Self-assessed health	Misery index
Parental math stereotypes	0.145***	0.983***
	(0.028)	(0.059)
R^2	0.059	0.087
Observations	6928	6963

Note: (1) OLS estimates of Equation 2 with new dependent variables (see Table A1 for descriptive statistics). (2) Standard errors are robust and clustered at the city level. (3) p-values. (4) ***Self-assessed health*** is measured by student responses to the question "How's your overall health right now from 1 (Very good), 2 (Fairly good), 3 (Fair), 4 (Not very good) or 5 (Very bad)?" (5) The ***misery index*** is calculated by the principal component analysis method condensing the 5 student well-being categories introduced in Table 1 to one single variable. See also further explanations in the text.
***$p<0.01$, **$p<0.05$, and *$p<0.1$.

0.85, which means that the commonality between the five variables is strong and satisfies the basic assumptions of PCA. In the second step, we calculated the weight of each variable for the constituent principal component indicators, targeting the linear combination of the unit length of the variables with the largest variance. Finally, according to the weight assigned to each variable, we summed up an indicator (misery index) that reflects student (negative) well-being. As Table 8 shows, both the self-assessed health of the student as well as the aggregate of her or his well-being measures strongly depend on parental math stereotypes. These results are consistent with the previous findings on the crucial role of parental math stereotypes.

6 | Causality

Thus far, we have established strong associations between parental math gender stereotypes and various misery indicators of their children in school. Are they also causal? An instrumental variable (IV) approach requires strong instruments and convincing reflections on the validity of the exclusion restriction. The two conditions for instrument validity are relevance and exogeneity, respectively. Exogeneity (or the exclusion restriction) requires that the instrument be uncorrelated with the error term in the structural equation, in our case, Equation 2, a condition that is difficult to test. Relevance implies that the instrument must be strongly correlated with parental stereotypes conditional on other covariates.

Testing for nonzero correlations or underidentification may rely on the Kleibergen–Paap rk LM statistic, a Lagrange multiplier test that uses the rank-based rk statistic (Kleibergen and Paap 2006). The null hypothesis that the model is underidentified or that the instrument is irrelevant is rejected when the smallest canonical correlation is nonzero, which is the more likely, the larger the value of the statistic is (Bazzi and Clemens 2013). A larger affiliated p-value of the statistic than the chosen significance level (e.g., 5%; 1%) indicates this. Furthermore, a strong correlation can be examined using the Kleibergen–Papp Wald F statistic (Kleibergen and Paap 2006), which is identical to the conventional F-test of variable inclusion in the case of a single endogenous regressor with a single instrument (Andrews, Stock, and J.H., and Sun, L. 2019). Lee et al. (2022) have suggested how the standard errors of the instrumented variable in the second-stage regression have to be adjusted if the F-values of the first stage are "too small."

Table 9 analyzes three potential instruments and the corresponding two-stage least squares (2SLS) analyses. The first candidate is the math gender stereotypes held by the parents' peers; the second is the average rate of parents admitting math gender stereotypes in the city where the parents reside; and the third is school-related stereotypes measured by the average rate of parents in the student's school admitting math gender stereotypes. Evaluating the validity of the exclusion restriction in these cases and assessing the likelihood that these variables directly impact student well-being is crucial.

In the case of math gender stereotypes among parents' peers, it seems plausible that students may be less aware of these stereotypes as they are less likely to be in close contact with their parents' peers. Similarly, regarding the average rate of parents admitting math gender stereotypes in the city, students might not be acutely aware of these stereotypes, especially if such attitudes are not considered politically correct in schools and society. Regarding school-related stereotypes, students were more likely to be aware of the well-being of their direct classmates (see Table 7). However, being cognizant of stereotypes across different classes in their school, particularly when these stereotypes are not socially acceptable, might be less likely. However, the possibility of direct effects on student well-being cannot be completely ruled out.

Table 9 demonstrates the robust performance of all instruments, as evidenced by their strong results at the 1% significance level for both the F-test and the Kleibergen–Paap rk LM statistic, except for city math stereotypes and the LM statistic, which exhibits a significance level of 1.75% only. Notably, the second-stage coefficients associated with the instrument "peers of parent math stereotypes" closely align with those of the baseline model, while the coefficients for the other two instruments are markedly larger.

Maintaining a 1% significance level, all parameter estimates remain highly significant, except for the instrument "city math stereotypes," which exhibits lower significance levels for miseries labeled "blue" (5%) and "sad" (10%). Given the

TABLE 9 | IV regressions using peers of parent, city, and school stereotypes as instruments.

	Depressed (1)	Blue (2)	Unhappy (3)	Not enjoying life (4)	Sad (5)
Baseline full sample					
Stereotypes	0.440***	0.500***	0.483***	0.425***	0.402***
	(0.030)	(0.036)	(0.037)	(0.040)	(0.030)
R^2	0.058	0.064	0.060	0.042	0.047
Peers of parent stereotypes					
First stage: OLS	Stereotypes				
Peers of parent math stereotypes	0.507***		*F*-test of instrument: 1391.16		
	(0.014)		$p = 0.0000$		
Second stage: 2SLS	Depressed	Blue	Unhappy	Not enjoying life	Sad
Stereotypes predicted	0.557***	0.545***	0.610***	0.488***	0.492***
	(0.038)	(0.051)	(0.057)	(0.056)	(0.049)
R^2	0.056	0.064	0.058	0.041	0.046
Kleibergen–Paap rk LM			26.178 $p = 0.0000$		
City stereotypes					
First stage: OLS	Stereotypes				
City math stereotypes	1.005***		*F*-test of instrument: 56.89		
	(0.133)		$p = 0.0000$		
Second stage: 2SLS	Depressed	Blue	Unhappy	Not enjoying life	Sad
Stereotypes predicted	1.255**	1.049**	1.376***	1.413***	1.113*
	(0.490)	(0.527)	(0.436)	(0.512)	(0.636)
0.05 tF standard error × 1.071	(0.525)^{xx}	(0.564)	(0.467)^{xx}	(0.548)^{xx}	(0.681)
0.01 tF standard error × 1.309	(0.641)	(0.690)	(0.571)	(0.670)	(0.833)
R^2	−0.077	0.015	−0.087	−0.102	−0.043
Kleibergen–Paap rk LM			5.647 $p = 0.0175$		
School stereotypes					
First stage: OLS	Stereotypes				
School math stereotypes	1.001***		*F*-test of instrument: 324.93		
	(0.056)		$p = 0.0000$		
Second stage: 2SLS	Depressed	Blue	Unhappy	Not enjoying life	Sad
Stereotypes predicted	0.912***	0.836***	0.977***	0.756***	0.928***
	(0.239)	(0.267)	(0.234)	(0.223)	(0.249)
R^2	0.013	0.046	0.015	0.026	−0.002
Kleibergen–Paap rk LM			15.36 $p = 0.0001$		

Note: (1) OLS/2SLS estimates with various instruments (see Table A1 for descriptive statistics). (2) Standard errors are robust and clustered at the city level. (3) *p*-values. (4) Sample sizes are 6963 (baseline full sample), 6948 (peers of parent stereotypes), and 6963 (city stereotypes and school stereotypes). (5) ***IV Peers of parent stereotypes***) is based on the parental response to the question "whether people around you agree that boys are better at mathematics than girls" (0 = no, 1 = yes). (6) ***IV City stereotypes*** is the average rate of parents admitting stereotypes in the city where parents are located. (7) ***IV School stereotypes*** is the average rate of parents admitting stereotypes in the student's school. (8) ***Instruments*** are set up as single alternatives. The standard *F*-test for the instrument is in this setting identical to the Kleibergen–Papp Wald *F* statistic. The Kleibergen–Paap rk LM statistic tests for underidentification. All instruments pass conventional requests for strong instruments, for example, $F > 10$ and Kleibergen–Paap rk LM with $p \leq 0.05$. (9) *t*-value inference in the second stage might suffer from weak instruments; tF adjustments suggested by Lee et al. (2022) are only of potential interest for city math stereotypes. (10) Stereotypes are math stereotypes. (11) Main findings remain robust.

***$p < 0.01$, **$p < 0.05$, and *$p < 0.1$.

robust first-stage *F*-values for school stereotypes and peers of parent stereotypes, adjustments, as recommended by Lee et al. (2022), are necessary only for city math stereotypes. Adjustments involve a factor of 1.071 at the 5% level (refer to Table 3A, p. 3271 of Lee et al. 2022) and a factor of 1.309 at the 1% level (refer to Table 3B, p. 3272). Consequently, none of the estimates associated with the instrument "city math stereotypes" attained statistical significance at the 1% level, in contrast to the two cases before instrumenting. Nonetheless, three out of the five estimates remain statistically significant at the 5% level, only one less than in the pre-instrumentation scenario, namely the case of "blue."

Importantly, none of these causality findings undermine the relevance of the negative health effects caused by parental math gender stereotypes identified in this study.

7 | Review and Conclusions

Using a large sample from the well-established China Education Panel Survey, our study investigates the intergenerational relationship between parental education and gender stereotypes for well-being measures among 11- to 18-year-old students beyond a larger number of control variables. Measures collected on a five–level intensity scale cover the well-defined well-being categories "depressed," "feeling blue," "unhappy," "not enjoying life," and "sad." Parental gender stereotypes identified on the basis of parental responses to the question "Do you think boys are better at learning mathematics than girls?" are shown to strongly decrease student well-being in China but with no relevant gender differences among parents and among students. In addition, parental human capital has no stabilizing effect on offspring well-being.

The well-being effects of gender math stereotypes we find are not gender-specific but are nevertheless relevant in terms of the size and statistical significance of the measured associations. This is consistent with the general findings in the well-being literature, which exhibits no or no robust gender differences. However, parental stereotypes may be damaging where they still exist. A quarter of all Chinese parents have gender math stereotypes with potentially detrimental effects on the well-being of their children, whereas the revealed effects on the five well-being measures are very similar.

These robust findings are supported by extensive further robustness checks, confirming the harmful nature of parental math stereotypes across different subgroups, including girls and boys with varying math performances. Additional factors, such as peer well-being, students' own math stereotypes, and teacher characteristics are considered and found to have consistent but smaller impacts compared to parental math stereotypes, without changing their basic strength and significance. The robustness checks also extend to alternative measures of well-being, including self-assessed health, and a PCA of the five well-being measures confirming the relevance of parental math stereotypes.

Furthermore, the study explores the potential causality of the observed associations using an IV approach, introducing three potential instruments related to parental math stereotypes. The instruments are found to be valid, and the IV approach supports the causal interpretation of the negative health effects of parental gender math stereotypes. The study underscores the persistence and robustness of the identified relationship, providing valuable insights into the impact of parental stereotypes on student well-being.

This suggests that parental gender "math" stereotypes are not just annoying attitudes for general societal debates, as we have in many countries around the world. They are an important topic for gender equality policies, particularly if research can confirm, as we do in this paper, that conjectures such as the math stereotype are actually wrong. This opens a new policy agenda. A better understanding of the origin of such stereotypes, and whether they are genetically or behaviorally transferred across generations through families, schools, societal norms, or public policies, provides a challenging but important agenda for future research.

Acknowledgments

We are grateful to the Editor and the anonymous referee as well as conference and seminar participants at the IEA World Congress 2023 in Medellin, EALE 2023 in Prague, ESPE 2024 in Rotterdam, EEA 2024 in Rotterdam, SITES 2024 in Naples, at Renmin University of China, Jinan University, the University of Edinburgh Business School and the Royal Society of Edinburgh, in particular Tanika Chakraborty, Shuaizhang Feng, Jun Hyung Kim, Wei Luo, Astghik Mavisakalyan, and Yun Qiu for many helpful comments and suggestions. We thank the Data Center of Renmin University of China for access to the data set used in this study. This research has received no specific funding.

Data Availability Statement

The data that support the findings of this study are openly available on the CEPS website hosted by Renmin University of China: http://ceps.ruc.edu.cn/English/Home.htm. This repository provides the entire 'core dataset' necessary to interpret, verify, and extend the research in this article. Some sensitive information, such as the geographical distribution of samples, requires additional applications.

Endnotes

[1] Chen and Davey (2008) review the large amount of subjective well-being/happiness papers published in Chinese language journals, whereas Davey and Rato (2012) evaluate the research executed in China by the International Wellbeing Group employing the then newly developed International Wellbeing Index. A recent study by Söllner et al. (2021) investigates age stereotypes and how self-regulatory behavior mediates the effects of well-being. The paper finds positive associations between positive age stereotypes and various variants of well-being.

[2] CEPS is a nationally representative large-scale tracking survey project designed and implemented by China Survey and Data Center of Renmin University of China. The project was funded by the Scientific Research Foundation of Renmin University of China, the Social Investigation Foundation of China Survey and Data Center, and the National Science Foundation (NSF) of the United States. Although only available for 2014, it is a commonly used data set for studying education issues and child development in China (Gong, Lu, and Song 2018; Hu 2018). The data that support the findings of this study are openly available on the CEPS website hosted by Renmin University of China: http://ceps.ruc.edu.cn/English/Home.htm. This repository provides

the entire "core dataset" necessary to interpret, verify, and extend the research in this article. Some sensitive information, such as the geographical distribution of samples, requires additional applications.

[3] Adapted on June 28, 1988, at the Eleventh Regular Session of the State Council and came into force on September 1, 1988. In addition, this administrative regulation was adjusted in 2012 and has been in use since then; in 2012, the name of the original administrative regulation was changed to "Special Rules on the Labor Protection of Female Employees," which, on the basis of the original administrative regulation, specifies in more detail the Scope of Prohibited Labor for Female Employees, for example, working in a mine well or discontinuously bearing a load of 25 kg or more each time.

[4] "Feeling blue" is an idiom conveying temporary sadness or melancholy, commonly caused by events like breakups, job losses, or deaths, rather than indicating a persistent mental health condition.

[5] Note that we collapse all five variables into one misery indicator in Table 8 for a robustness check, as reported in Section 4 and confirm our major finding.

[6] This could suggest a reporting bias among the responding fathers, driven by cultural stereotypes about age-differences between couples, for which we have no indication in the data. Fortunately, our empirical specification in Equation 1 explicitly tests whether the difference matters. However, the parameters for E_f * Mother in Table 4 are all small and statistically insignificant.

[7] To explain the modelling idea, note a well-known OLS equivalent applied to our multidimensional setting: If we have $Y_i = h_{0i} + h_{1i} X_i + u_i$ for subsamples $i = 1,2$, we can estimate either the two equations separately to identify parameters h_{0i} and h_{1i}; or we can estimate for the pooled sample $Y = h_0 + h_1 X + k_0 D + k_1 D X + u$ and find for $D = 1$ for $i = 1$, 0; otherwise, the identities $h_{01} = h_0 + k_0$ and $h_{02} = h_0$ as well as $h_{11} = h_1 + k_1$ and $h_{12} = h_1$. This implies that the relevant sample differences $k_0 = h_{01} - h_{02}$ and $k_1 = h_{11} - h_{12}$ are directly estimated.

[8] Standardized math scores are obtained by adjusting the children's test scores in the midterm math examination, which are centrally administered by the school. The standardized scores are calculated separately by school and grade and adjusted to a score with mean = 70 and standard deviation = 10. See Table A3 and Figure A1 for further descriptive information and the distribution of the standardized math scores in the sample. Math scores of girls show a higher mean (72) and a smaller standard deviation (9) than for the boys (mean = 70 and standard deviation = 10). The min/max values reflect this: 34/96 for the girls and 30/99 for the boys.

[9] Parental migration stereotype: "What kind of effect do you think will the increase of students from non-local county/district have on the atmosphere of the school?" 1, if response was harmful; 0, otherwise. Teacher totally responsible: Equal to 1 if parent answered to "Do you agree that it is totally the teachers' responsibility to educate children" by "somewhat agree" or "strongly agree," and 0, otherwise.

References

Afridi, F., S. X. Li, and Y. Ren. 2015. "Social Identity and Inequality: The Impact of China's Hukou System." *Journal of Public Economics* 123: 17–29.

Akerlof, G. A., and R. E. Kranton. 2000. "Economics and Identity." *Quarterly Journal of Economics* 115, no. 3: 715–753.

Alesina, A., P. Giuliano, and N. Nunn. 2013. "On the Origins of Gender Roles: Women and the Plough." *Quarterly Journal of Economics* 128, no. 2: 469–530.

All China Women's Federation. 2019. *White Book of Equal Development and Sharing: The Development and Progress of Women's Cause in the Past 70 Years in New China*. Beijing, China: Central People's Government of the PRC. Retrieved from. http://www.gov.cn/zhengce/2019-09/19/content_5431327.htm.

Andrews, I., J. H. Stock, and L. Sun. 2019. "Weak Instruments in Instrumental Variables Regression: Theory and Practice." *Annual Review of Economics* 11: 727–753. https://doi.org/10.1146/annurev-economics-080218-025643.

Batz, C., and L. Tay. 2018. "Gender Differences in Subjective Well-Being." In *Handbook of Well-Being*, edited by E. Diener, S. Oishi, and L. Tay. Salt Lake City, UT: DEF Publishers DOI: nobascholar.com.

Bazzi, S., and M. A. Clemens. 2013. "Blunt Instruments: Avoiding Common Pitfalls in Identifying the Causes of Economic Growth." *American Economic Journal: Macroeconomics* 5, no. 2: 152–186.

Bieg, M., T. Goetz, I. Wolter, and N. C. Hall. 2015. "Gender Stereotype Endorsement Differentially Predicts Girls' and Boys' Trait-State Discrepancy in Math Anxiety." *Frontiers in Psychology* 6: 1404.

Bohong, L. 2020. "The All-China Women's Federation and Women's NGOs." In *Chinese Women Organizing: Cadres, Feminists, Muslims, Queers*, edited by P.-C. Hsiung, M. Jaschok, and C. Milwertz, 141–157. Oxford and New York: Routledge.

Bursztyn, L., A. L. González, and D. Yanagizawa-Drott. 2020. "Misperceived Social Norms: Women Working Outside the Home in Saudi Arabia." *American Economic Review* 110, no. 10: 2997–3029.

Carlana, M. 2019. "Implicit Stereotypes: Evidence From Teachers' Gender Bias." *Quarterly Journal of Economics* 134, no. 3: 1163–1224.

Chen, Z., and G. Davey. 2008. "Happiness and Subjective Wellbeing in Mainland China." *Journal of Happiness Studies* 9: 589–600. https://doi.org/10.1007/s10902-008-9092-7.

Chou, A. L. 2022. "Fortunes and Misfortunes of the Dragon Sons: Direct and Cohort Effects of Superstition on Education Attainment." *Kyklos* 75, no. 4: 564–579. https://doi.org/10.1111/kykl.12312.

Davey, G., and R. Rato. 2012. "Subjective Wellbeing in China: A Review." *Journal of Happiness Studies* 13: 333–346. https://doi.org/10.1007/s10902-011-9266-6.

Ding, J., J. Salinas-Jiménez, and M. del Mar Salinas-Jiménez. 2021. "The Impact of Income Inequality on Subjective Well-Being: The Case of China." *Journal of Happiness Studies* 22: 845–866. https://doi.org/10.1007/s10902-020-00254-4.

Du, H., Y. Xiao, and L. Zhao. 2021. "Education and Gender Role Attitudes." *Journal of Population Economics* 34, no. 2: 475–513.

Forgasz, H. J., G. C. Leder, and P. Kloosterman. 2004. "New Perspectives on the Gender Stereotyping of Mathematics." *Mathematical Thinking and Learning* 6, no. 4: 389–420.

Gong, J., Y. Lu, and H. Song. 2018. "The Effect of Teacher Gender on Students' Academic and Noncognitive Outcomes." *Journal of Labor Economics* 36, no. 3: 743–778.

Hu, F. 2018. "Migrant Peers in the Classroom: Is the Academic Performance of Local Students Negatively Affected?" *Journal of Comparative Economics* 46, no. 2: 582–597.

Kelifa, M. O., Y. Yang, H. Carly, W. Bo, and P. Wang. 2021. "How Adverse Childhood Experiences Relate to Subjective Wellbeing in College Students: The Role of Resilience and Depression[J]." *Journal of Happiness Studies* 22: 2103–2123.

Kleibergen, F., and R. Paap. 2006. "Generalized Reduced Rank Tests Using the Singular Value Decomposition." *Journal of Econometrics* 133, no. 1: 97–126.

Kling, J. R., J. B. Liebman, and L. F. Katz. 2007. "Experimental Analysis of Neighborhood Effects." *Econometrica* 75, no. 1: 83–119.

Kuhn, A., and S. C. Wolter. 2023. "The Strength of Gender Norms and Gender-Stereotypical Occupational Aspirations Among Adolescents." *Kyklos* 76, no. 1: 101–124. https://doi.org/10.1111/kykl.12320.

Lee, D. S., J. McCrary, M. J. Moreira, and J. Porter. 2022. "Valid *t*-Ratio Inference for IV." *American Economic Review* 112, no. 10: 3260–3290.

Li, C. L. 2016. ""Male Disadvantage," "Left-Over Women" and Difficulty in the Job Market of Female College Students: Challenges From the Gender Reversal in Education." *Journal of Chinese Women's Studies* 2: 33–39.

Liang, Z., M. Zhang, F. Huang, D. Kang, and L. Xu. 2021. "Application Innovation of Educational Measurement Theory, Method, and Technology in China's New College Entrance Examination Reform." *Chinese/English Journal of Educational Measurement and Evaluation* 2, no. 1: 33–41.

Liu, A. Y. 2018. "Institution, Opportunity Structure and Gender Attitude: How Labor Participation of Urban Married Women Being Possible?" *Journal of Chinese Women's Studies* 6: 15–30.

McCoy, S., D. Byrne, and P. O'Connor. 2022. "Gender Stereotyping in Mothers' and Teachers' Perceptions of Boys' and Girls' Mathematics Performance in Ireland." *Oxford Review of Education* 48, no. 3: 341–363.

Meisenberg, G., and M. A. Woodley. 2015. "Gender Differences in Subjective Well-Being and Their Relationships With Gender Equality." *Journal of Happiness Studies* 16: 1539–1555. https://doi.org/10.1007/s10902-014-9577-5.

Mishra, A., and J. Parasnis. 2022. "Intentions for a Third Child: The Role of Parental Sex Composition Preferences." *Kyklos* 75, no. 2: 472–487. https://doi.org/10.1111/kykl.12298.

Mow, S., T. Jie, and Z. Bijun. 2004. *Holding up Half the Sky: Chinese Women Past, Present, and Future*. New York: The Feminist Press at the City University of New York.

Muntoni, F., and J. Retelsdorf. 2019. "At Their Children's Expense: How Parents' Gender Stereotypes Affect Their Children's Reading Outcomes." *Learning and Instruction* 60: 95–103.

Nie, P., A. Sousa-Poza, and G. Nimrod. 2017. "Internet Use and Subjective Well-Being in China." *Social Indicators Research* 132: 489–516. https://doi.org/10.1007/s11205-015-1227-8.

Nikolova, M., and C. Graham. 2022. "The Economics of Happiness." In *Handbook of Labor, Human Resources and Population Economics*, edited by K. F. Zimmermann. Cham, Switzerland: Springer Nature. https://doi.org/10.1007/978-3-319-57365-6_177-1.

Nolen-Hoeksema, S., and C. L. Rusting. 2003. "Gender Differences in Well-being." In *Well-Being: Foundations of Hedonic Psychology*, edited by D. Kahneman, E. Diener, and N. Schwarz, 330–352. New York, NY: Russell Sage Foundation.

Pennington, C. R., L. K. Kaye, A. W. Qureshi, and D. Heim. 2021. "Do Gender Differences in Academic Attainment Correspond With Scholastic Attitudes? An Exploratory Study in a UK Secondary School." *Journal of Applied Social Psychology* 51, no. 1: 3–16.

Söllner, M., M. Dürnberger, J. Keller, and A. Florack. 2021. "The Impact of Age Stereotypes on Well-Being: Strategies of Selection, Optimization, and Compensation as Mediator and Regulatory Focus as Moderator: Findings From a Cross-Sectional and a Longitudinal Study, Online." *Journal of Happiness Studies*. 23: 635–665. https://doi.org/10.1007/s10902-021-00417-x.

Tani, M. 2017. "Hukou Changes and Subjective Well-Being in China." *Social Indicators Research* 132: 47–61. https://doi.org/10.1007/s11205-016-1247-z.

Tobia, V., A. Greco, P. Steca, and G. M. Marzocchi. 2019. "Children's Wellbeing at School: A Multi-Dimensional and Multi-Informant Approach." *Journal of Happiness Studies* 20: 841–861. https://doi.org/10.1007/s10902-018-9974-2.

Wang, L. F., and P. P. Heppner. 2002. "Assessing the Impact of Parental Expectations and Psychological Distress on Taiwanese College Students[J]." *Counseling Psychologist* 30, no. 4: 582–608.

Xi, J. 2020. "Statement by H.E. Xi Jinping President of the People's Republic of China at the High-Level Meeting on the Twenty-fifth Anniversary of The Fourth World Conference on Women. October 1, 2020." https://www.fmprc.gov.cn/mfa_eng/wjdt_665385/zyjh_665391/t1821556.shtml.

Xu, C. 2011. "The Fundamental Institutions of China's Reforms and Development." *Journal of Economic Literature* 49, no. 4: 1076–1151.

Yang, L., H. Lu, and M. Li. 2023. "Multidimensional Inequality and Subjective Well-Being in China: A Generalized Ordered Logit Model Analysis." *Social Indicators Research* 165: 1021–1052. https://doi.org/10.1007/s11205-022-03060-y.

Yamamoto, Y., and S. D. Holloway. 2010. Parental Expectations and Children's Academic Performance in Sociocultural Context[J]. *Educational Psychology Review* 22: 189–214.

Yang, S., and C. S. Shin. 2008. "Parental Attitudes Towards Education: What Matters for Children's Well-Being?[J]." *Children and Youth Services Review* 30 no. 11: 1328–1335.

Zhang, D. C., and T. L. Renshaw. 2020. "Personality and College Student Subjective Wellbeing: A Domain-Specific Approach." *Journal of Happiness Studies* 21: 997–1014. https://doi.org/10.1007/s10902-019-00116-8.

Zhang, J., S. Jin, T. Li, and H. Wang. 2021. "Gender Discrimination in China: Experimental Evidence From the Job Market for College Graduates." *Journal of Comparative Economics* 49, no. 3: 819–835.

Appendix A

TABLE A1 | Descriptive statistics for variables used for robustness analysis in Sections 5 and 6.

	Define	Mean	*SD*
Peer depressed	Average student's depressed in the class where the student is located	2.476	0.277
Peer blue	Average student's blue in the class where the student is located	2.226	0.309
Peer unhappy	Average student's unhappy in the class where the student is located	2.543	0.287
Peer not enjoying life	Average student's not enjoying life in the class where the student is located	1.957	0.301
Peer sad	Average student's sad in the class where the student is located	2.28	0.294
Student's own stereotype	Which is equal to 1 if the answer of the student themself is "Yes" when asked: "Do you think boys are better at learning mathematics than girls?" and equal to 0, otherwise.	0.492	0.500
Homeroom teacher female	Which is equal to 1 if it is female; others, equal to 0.	0.629	0.483
Homeroom teacher teaches math	Which is equal to 1 if the homeroom teacher teaches math; others, equal to 0.	0.299	0.458
Math teacher is female	Which is equal to 1 if the math teacher is female; others, equal to 0.	0.627	0.484
Stereotype of migrant	Which is equal to 1 if the parent responded "harmful" to "What kind of effect do you think will the increase of students from non-local county/district have on the atmosphere of the school?" and 0, otherwise.	0.107	0.309
Stereotype of responsibility	Which is equal to 1 if the parent answered to "Do you agree that it is totally the teachers' responsibility to educate children" by "Somewhat agree" or "Strongly agree," and 0 otherwise.	0.102	0.303
Self-assessed health	Which is measured by student responses to the question "How's your overall health right now?" from 1 (Very good) to 2 (Fairly good), 3 (Fair), 4 (Not very good) to 5 (Very bad).	2.002	0.894
Misery index	The misery index is calculated by the principal component analysis method condensing the five student well-being categories introduced in Table 1.	0	1.752
Peers of parent stereotypes	Which is created by the question of "whether people around you agree that boys are better at mathematics than girls" of parents (0 = not, 1 = yes).	0.419	0.493
City stereotypes	Which is created by the average rate of parents admitting stereotypes in the city where the parents are located.	0.25	0.042
School stereotypes	Which is created by the average rate of parents admitting stereotypes in the student's school.	0.25	0.072

TABLE A2 | Parental math stereotypes and interactions with age and school performance.

	Depressed	Blue	Unhappy	Not enjoying life	Sad
	(1)	(2)	(3)	(4)	(5)
Baseline full sample					
Stereotypes	0.440***	0.500***	0.483***	0.425***	0.402***
	(0.030)	(0.036)	(0.037)	(0.040)	(0.030)
Age	0.048***	0.067***	0.038***	0.051***	0.038***
	(0.011)	(0.010)	(0.012)	(0.016)	(0.011)
Rank in primary school	0.002*	0.004***	0.004***	0.006***	0.006***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
R^2	0.058	0.064	0.060	0.042	0.047
Stereotypes * Age and Stereotypes * Academic ranking in primary school					
Stereotypes	0.724***	0.905**	0.481*	1.004***	1.078***
	(0.193)	(0.399)	(0.238)	(0.351)	(0.263)
Age	0.052***	0.072***	0.037***	0.060***	0.048***
	(0.011)	(0.012)	(0.012)	(0.015)	(0.013)
Academic ranking in primary school	0.003**	0.007***	0.005***	0.008***	0.007***
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)
Stereotypes * Age	−0.016	−0.018	0.007	−0.036	−0.044**
	(0.014)	(0.028)	(0.018)	(0.024)	(0.020)
Stereotypes * Academic ranking in primary school	−0.004	−0.010***	−0.006***	−0.005	−0.004
	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)
R^2	0.059	0.067	0.061	0.043	0.048
Stereotypes * Age					
Stereotypes	0.656***	0.718*	0.366	0.912**	1.005***
	(0.204)	(0.385)	(0.252)	(0.342)	(0.268)
Age	0.051***	0.071***	0.036***	0.059***	0.048***
	(0.011)	(0.012)	(0.012)	(0.015)	(0.013)
Academic ranking in primary school	0.002*	0.004***	0.004***	0.006***	0.006***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Stereotypes * Age	−0.015	−0.015	0.008	−0.035	−0.043**
	(0.015)	(0.028)	(0.018)	(0.025)	(0.020)
R^2	0.059	0.065	0.060	0.042	0.047
Stereotypes * Academic ranking in primary school					
Stereotypes	0.496***	0.656***	0.581***	0.500***	0.461***
	(0.047)	(0.063)	(0.052)	(0.060)	(0.047)
Age	0.048***	0.068***	0.038***	0.051***	0.038***
	(0.011)	(0.010)	(0.012)	(0.016)	(0.011)
Academic ranking in primary school	0.003*	0.007***	0.005***	0.008***	0.007***
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)
Stereotypes * Academic ranking in primary school	−0.004	−0.010***	−0.006***	−0.005	−0.004
	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)
R^2	0.059	0.067	0.061	0.042	0.047

Note: (1) OLS estimates of Equation 2. (2) *p*-values. (3) Sample sizes are 6963 (Baseline full sample) and 6953(Stereotypes * Age and Stereotypes * Academic ranking in primary school, Stereotypes * Age, Stereotypes * Academic ranking in primary school). (4) Stereotypes are math stereotypes.
***$p < 0.01$, **$p < 0.05$, and *$p < 0.1$.

TABLE A3 | Distribution of standardized math scores.

	N	Mean	*SD*	Min	Max
Standardized math scores (full sample)	6822	70.956	9.524	29.543	99.354
Standardized math scores (girl)	3653	71.826	8.963	34.093	96.009
Standardized math scores (boy)	3169	69.952	10.040	29.543	99.354

Note: Standardized math scores of students in the China Education Panel Survey (CEPS); own calculations. Standardized math scores are obtained by adjusting the children's test scores in the midterm math examination centrally administered by the school. The standardized scores are calculated separately by school and grade, and adjusted to a score with mean = 70 and standard deviation = 10.

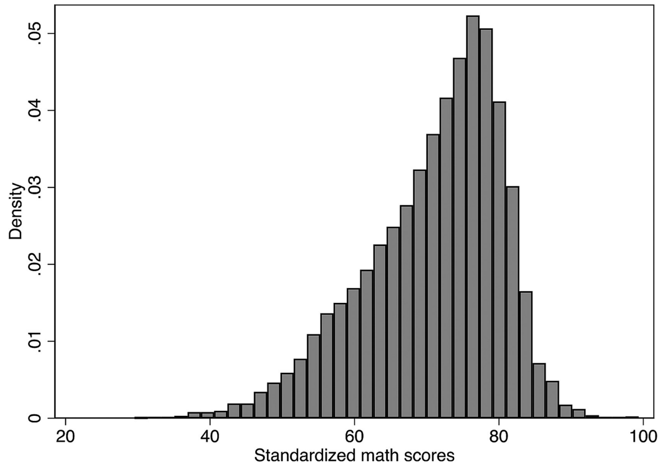

FIGURE A1 | Distribution of standardized math scores. *Note*: Standardized math scores of students in the China Education Panel Survey (CEPS); own calculations. Standardized math scores are obtained by adjusting the children's test scores in the midterm math examination centrally administered by the school. The standardized scores are calculated separately by school and grade, and adjusted to a score with mean = 70 and standard deviation = 10.