Educational Competition, Mental Health, and the Allocation of Human Capital: A Macroeconomic Analysis*

Shijun Gu† Lichen Zhang‡

January 15, 2026

Abstract

Educational competition can distort not only the level of parental investment in children, but also the composition of human capital it produces. In rank-based admission systems with scarce elite education slots, parents are incentivized to invest in exam-oriented inputs even when they harm children's mental health. We study the aggregate and intergenerational consequences of this trade-off in China. Using household survey data and original parent-child paired belief data, we document that tutoring-oriented investment is associated with higher cognitive skills but worse mental health, and parents substantially underperceive children's mental health problems, especially in lower-SES families. We develop and estimate a heterogeneous-agent life-cycle model in which parents invest in children's cognitive skills and mental health while competing for limited college slots. Tutoring raises cognitive skills but worsens mental health, and both dimensions affect adult outcomes. The model implies an education-competition externality that leads to excessive private investment in exam-oriented skills, while parental unawareness further strengthens the case for intervention. Quantitatively, a tutoring ban raises short-run welfare but lowers long-run welfare because of intergenerational costs. By contrast, an optimal tutoring tax combined with higher public investment raises welfare by 0.8% in the benchmark economy and by 1.1 % when parents are unaware of children's mental health.

Keywords: Mental Health, Educational Competition, Misallocation of Human Capital, Child Development, Inequality, Intergenerational Mobility

JEL Classification: E21, E24, I24, J24, O15

*For useful comments, we thank Mariacristina De Nardi, Gabriella Conti, Matthias Doepke, Jonathan Heathcote, Kyle Herkenhoff, Jo Mullins, Todd Schoellman, Michele Tertilt, Fabrizio Perri, Weilong Zhang, Kai Zhao, and participants at the Peking University GSM seminar, HKU Brownbag, EF&G Annual Conference 2025, K21-Jiam International Conferenc, and *Female Voices in Empirical Economics* at NYU Abu Dhabi, and Demographics, Health and Inequality (DHI) Workshop in Okinawa. We thank Huan Liu for excellent research assistance. All errors are our own.

†Shijun Gu (email: shijungu@cufe.edu.cn): Central University of Finance and Economics, China.

‡Lichen Zhang (email: lichenz@hku.hk): University of Hong Kong, Hong Kong SAR.

1 Introduction

The prevalence and importance of child mental health problems have been increasingly recognized in recent years. According to a recent World Health Organization report, globally, one in seven 10-19-year-olds experiences a mental disorder, accounting for 15% of the global burden of disease in this age group.[1] Child mental health is closely related to non-cognitive skills and is an important aspect of human capital development. Mental health formed in childhood and adolescence often persists into adulthood (Cuddy and Currie, 2026),[2] with significant impacts on labor market outcomes. From a macro/growth perspective, children's and adolescents' mental health is closely tied to the accumulation of "educational human capital" and "health human capital" (Grossman, 1972; Becker, 2007; Currie, 2020), which in turn determines adult human capital and labor productivity. Because human capital is a key production input, and labor efficiency depends on both cognitive ability and mental health, these factors matter for aggregate output and productivity. From the perspective of inequality and social mobility, differences in family resources and parenting styles can largely impact adolescents' physical and mental health. These disparities may reinforce existing inequalities, which are further amplified by intergenerational dynamics.

A growing empirical literature documents links between child mental health and economically important future outcomes such as education, disability, and earnings (see Currie and Stabile, 2007; Currie, 2020, 2025 for an overview). Yet very little is known about its macro implications, its effects on inequality and intergenerational mobility, and how incorporating mental health alters our understanding of child development policies. While a few papers use structural models to examine the role of mental health for macroeconomic aggregates, inequality, and policy (Jolivet and Postel-Vinay, 2022; Abramson, Boerma, and Tsyvinski, 2025), they focus primarily on adulthood. As a result, little is known about the formation of mental health at the transition to adulthood and, more importantly, how lifetime inequality is shaped by mental health formed in childhood and by its interaction with the formation of cognitive skills.

This paper fills that gap by introducing endogenous mental health formation into a heterogeneous-agent life-cycle model of child development à la Lee and Seshadri (2019), Yum (2023), and Daruich (2022). We focus on environments characterized by intense

[1] For more details, see https://www.who.int/news-room/fact-sheets/detail

[2] The psychiatric literature (e.g., Green, McLaughlin, and Berglund (2010); Kosik et al. (2018); Johnson et al. (2018)) further indicates that depression in childhood and adolescence increases the risk of poor mental health outcomes in adulthood. In a meta-analysis, Johnson et al. (2018) found that adolescents with depression had 2.78 times higher odds of depression in adulthood (95% CI: 1.97–3.93).

educational competition, such as in countries with score-based college admissions and limited seats at elite universities. This context is prevalent across many institutional settings, including Asian countries like China, India, and South Korea—where university admission almost purely depends on exam scores and parents place a high value on education (a phenomenon often termed "education fever," as highlighted by Seth (2002)). It is also relevant for the United States, where a "rug rat race" for top college admissions persists (Bound, Hershbein, and Long, 2009; Ramey and Ramey, 2010). Such competition may incentivize parents to overinvest in cultivating cognitive skills for exam performance (e.g., through cram schools and private tutoring), potentially at the expense of their children's mental health. These investment distortions can misallocate different types of human capital, leading to substantial and persistent losses in income and welfare, with important implications for inequality and intergenerational mobility.

We begin by documenting three stylized facts that motivate our framework. First, adolescent mental health deteriorates substantially over school ages, with especially large declines among children receiving intensive tutoring. Second, tutoring-oriented parental investment is associated with higher cognitive skills but worse mental health, indicating a trade-off between academic achievement and mental well-being. Third, parents substantially underperceive their children's mental health problems, particularly in lower-SES families. Taken together, these facts suggest that educational competition may distort not only the level of parental investment, but also the composition of child human capital, and that parental misperceptions may further amplify this distortion.

We then develop a heterogeneous agent life cycle model in which parents invest in their children's human capital at multiple stages to compete for limited college slots. Admission probability depends solely on a child's relative cognitive skill rank. There are two types of human capital accumulated over childhood and adolescence, which shapes adult labor-market productivity: (i) cognitive skills, which is central for exam performance and college admission, and (ii) mental health capital, a subset of non-cognitive skills that captures psychological well-being. Parental monetary investment affects these two types of human capital differently.

In the seminal work by Cunha, Heckman, and Schennach (2010), the authors cannot empirically distinguish investments in cognitive versus non-cognitive skills and therefore model a single investment input that enters both skill-production technologies (with potentially different elasticities). Their investment concept is broad, encompassing the home environment, emotional support, cognitive stimulation, planned activities, and family surroundings; accordingly, advantaged family backgrounds improve both cognitive and non-cognitive skills.

In our model, we focus on a specific monetary input that is highly salient in test-oriented education systems: private tutoring expenditures. Investments that raise cognitive skills need not enhance all non-cognitive dimensions; in particular, tutoring can continue to boost cognitive achievement (with diminishing returns) while harming mental health when pushed beyond a threshold. While we abstract from endogenous channels through which family background positively impacts child mental health, we include an exogenous, family-background–dependent component to capture how the broad investments discussed in Cunha, Heckman, and Schennach (2010) positively affect child mental health. Formally, cognitive skills evolve via a multi-period production function in which next-period cognitive skills depends on current cognitive skills, parental cognitive skills, parental monetary investment, and public spending. Mental health capital in the current period depends on lagged mental health, parental monetary investment, parental mental health status, and the exogenous family-background term. When parental monetary investment exceeds a threshold—interpreted as intensive tutoring or cram-school participation—it has a negative marginal effect on current mental health. The magnitude of this adverse effect is disciplined by reduced-form evidence on the impact of cram-school participation on child mental health.

We discipline our model, focusing on the life cycle of one adult generation and one child generation under fixed prices, to target a cohort of children aged 6–18 from 2010 to 2018, along with their parents, in the China Family Panel Studies (CFPS). Our identification strategy proceeds in two steps. First, we estimate a set of parameters outside the model including the score-based college-admission function and the effect of mental health on labor productivity. Second, we calibrate the remaining parameters using the method of simulated moments (MSM) and indirect inference. These include parameters governing the child cognitive skill production function and the child mental health formation technology—most importantly, the effect of cram-school participation on mental health. For this step, we rely on the China Education Panel Study (CEPS), a nationally representative, school-based survey that initially samples roughly 10,000 seventh-grade students. We estimate the causal effect of attending cram schools on mental health using a district–school–level instrumental variable. We find that students receiving extracurricular tutoring have mental health capital scores 0.33 standard deviations lower than nonparticipants, indicating a substantial negative effect of cram-school participation on children's mental health. We use this reduced-form estimate to discipline the parameter governing the effect of private tutoring in the model.

Armed with a well-fitted model, we first examine how competition for college admissions shapes the allocation of two types of human capital crucial for child development

and adult labor market outcomes: cognitive skills and mental health capital. To isolate the effect of the educational competition, we eliminate it by making college admission a purely random lottery, independent of individual heterogeneity (including cognitive skill), while holding college capacity fixed at the benchmark level. Shutting off competition leads to a 61% reduction in average parental investment. This reduction, in turn, lowers children's cognitive skills by 10% but improves their mental well-being by approximately 50% of a standard deviation. Overall, expected lifetime income declines by 2.2% on average. However, by reallocating the private investment previously used for admissions competition into public investment—thus maintaining the same aggregate effective investment—we can induce a more efficient allocation of human capital. This approach preserves the improvement in child mental health capital while keeping cognitive skills nearly unchanged from the benchmark level. Consequently, both consumption-equivalent welfare and expected lifetime income increase.

We then quantitatively evaluate the aggregate, welfare, and distributional effects of regulating parental investment competition in college admissions from both positive and normative perspectives. We begin by analyzing the impacts of China's 2021 private tutoring ban. In the short run, for an average household living in current generation, parental investment declines by approximately 45% on average. This reduction leads to a 5% decrease in children's cognitive skills at college entry, but simultaneously results in an increase in mental health capital—equivalent to 38% of the standard deviation of the child mental health distribution. These two opposing effects on labor efficiency units ultimately result in a 2.3% reduction in children's expected lifetime income. Welfare, measured by the consumption equivalent variation (CEV) in ex ante lifetime utility of parents, increases by about 0.8% on average because the ban curbs over-investment in admissions-driven tutoring, improving the efficiency of cognitive investment and removing the mental-health penalty associated with parental monetary investment. The ban therefore improves the within-generation allocation between pre-college and mental health capital.

In the long run, although a ban on private tutoring raises current parental welfare, its effects on children—lower cognitive skills, higher mental health capital, and a reallocation between these two forms of capital—can propagate to future generations through two channels. First, lower pre-college human capital depresses adult income and, in turn, reduces monetary investment in the next generation's cognitive skills, generating persistent losses in skills and income. Second, better mental health raises adult income and directly boosts the next generation's mental health through parental mental health inputs and income-dependent components in child mental health formation, yielding

persistent gains. The net effect on income and investment is a priori ambiguous because these forces move in opposite directions and feed back into each other. Quantitatively, the persistent mental-health gains are not strong enough to offset the persistent losses in pre-college human capital: the ban ultimately worsens the allocation between the two forms of human capital and produces a long-run welfare loss of 2.4%.

Guided by the finding that reallocating spending from private tutoring to public investment can deliver a better balance between cognitive skills and mental health capital, we propose a revenue-neutral policy: a linear tax on private tutoring expenditures paired with a linear subsidy to public pre-college spending, financed solely by the tutoring tax. We identify a 28% tax rate that maximizes long-run ex ante lifetime utility in partial equilibrium. Relative to a ban, this policy raises welfare for all future generations by more effectively balancing two objectives: it dampens the competition incentive that creates over-investment and misallocation of cognitive skills and mental health capital while minimizing losses in cognitive skills, delivering a net welfare gain of 0.8%. When parents are unaware of children's mental health, the optimal tax rate is higher (34% vs. 28%) and the welfare gain larger (1.1% vs. 0.8%). If we abstract from the mental-health margin altogether, so the tax confers no mental-health benefits, the welfare gain falls to 0.5%. Interpreting the optimal tax-plus-subsidy allocation as a proxy for a social planner's outcome, the associated gains quantify the welfare cost of educational competition in the status quo. That cost is about 0.50 percent of aggregate consumption when mental health is excluded. Incorporating mental health raises the estimated cost to 0.79–1.06 percent—an increase of 58% to 200%. The message is clear: endogenizing child mental health is essential. Models that omit it substantially understate both the mechanisms and the welfare stakes of educational competition and the corresponding policy impacts.

Related Literature This paper contributes to several strands of literature. The first pertains to quantitative life-cycle models of child human capital formation with parental investments. This strand of work highlights various aspects, including credit constraints and uninsured labor market risk (Lee and Seshadri, 2019; Caucutt and Lochner, 2020; Abbott, 2022), household composition (especially maternal labor supply) (Del Boca, Flinn, and Wiswall, 2014; Mullins, 2022), neighborhood effects (Fogli and Guerrieri, 2019; Chyn and Daruich, 2024), parenting styles and peer effects (Agostinelli, Doepke, Sorrenti, and Zilibotti, 2022, Forthcoming), the long-run impacts of early childhood development policies (Restuccia and Urrutia, 2004; Daruich, 2022), and college admission policies (Krueger, Ludwig, and Popova, 2024). We develop our model based in particular on Lee and Seshadri (2019), Yum (2023), and Daruich (2022), which embed the skill formation

technology developed in Cunha, Heckman, and Schennach (2010) into quantitative macro models. In Cunha, Heckman, and Schennach (2010), the investment input is broad—capturing home environment, emotional support, and stimulation—and affects cognitive and non-cognitive skills, potentially with different elasticities. In contrast, we focus on a specific, policy-relevant monetary input salient in test-oriented systems: private tutoring expenditures. In our framework, tutoring continues to raise cognitive achievement (with diminishing returns) but can harm mental health when pushed beyond a threshold. Whereas much of the literature tackles underinvestment due to liquidity constraints and market incompleteness, we show that admissions competition can induce overinvestment that yields modest cognitive gains at the cost of substantial deterioration in child mental health, leading to a misallocation across the two forms of human capital.

The second strand of literature relates to macroeconomic models of health. While health economics is dominated by an empirical, reduced-form micro approach, a burgeoning literature starting with Grossman (1972) studies quantitative macroeconomic models of health, focusing on the dynamic, intertemporal trade-offs in health investment (Hubbard, Skinner, and Zeldes, 1995; French, 2005; Hall and Jones, 2007; Low, Meghir, and Pistaferri, 2010; De Nardi, French, and Jones, 2010; French and Jones, 2011; Kopecky and Koreshkova, 2014; Low and Pistaferri, 2015; De Nardi, French, and Jones, 2016; Braun, Kopecky, and Koreshkova, 2017, 2019; Cole, Kim, and Krueger, 2018; Ameriks et al., 2020; Fang and Krueger, 2022; Greenwood, Guner, and Kopecky, 2022; De Nardi, Pashchenko, and Porapakkarm, 2025; Hosseini, Kopecky, and Zhao, 2025; Mahler and Yum, 2024; Ozkan, 2025). Very few structural papers study mental health; two notable exceptions are Jolivet and Postel-Vinay (2022), who model the two-way interaction between labor market outcomes and mental health in adulthood using a search framework, and Abramson, Boerma, and Tsyvinski (2025), who highlight key margins that differentiate mental from physical health and matter for economic outcomes—particularly negative thinking and rumination—which in turn affect job choices, labor supply, asset portfolio choice, wealth accumulation, and treatment decisions. Using their quantitative macro framework, they estimate an aggregate cost of mental illness roughly 1.2 percent of consumption annually. Both papers focus on adulthood. We instead model endogenous mental health formation during childhood and adolescence—when physical health is relatively stable but mental health evolves sharply—and embed it in a heterogeneous-agent life-cycle model of child development. This allows us to quantify how mental health formed in youth interacts with cognitive skill formation, shapes lifetime inequality and intergenerational mobility, and alters the evaluation and design of education policies.

The third strand of literature examines educational competition and its implications using structural models, with different focuses. These include the varying childcare time spent by parents with different education levels in the U.S. (Ramey and Ramey, 2010), low fertility and childlessness in South Korea (Kim, Tertilt, and Yum, 2024), policy implications of regulating educational competition in China (Gu and Zhang, 2025), the role of child efforts (Kang, 2022), and innovation and growth (Celik, 2023). Our final contribution is to provide, to our knowledge, the first quantitative macroeconomic assessment of educational competition's impact on child mental health. By structurally modeling the endogenous formation of mental health and its interaction with cognitive skill development, we show that competition-driven overinvestment in tutoring worsens mental well-being, a effect amplified by intergenerational dynamics. We quantify a profound result: incorporating the mental health channel increases the estimated welfare cost of educational competition from 0.50% to between 0.79% and 1.06% of aggregate consumption—an increase of 58% to 112%. This highlights the critical importance of including mental health in any complete analysis of education policy.

2 Empirical Motivation

In this section, we present three set of empirical facts to motivate our model setup in Section 3. First, adolescent mental health declines sharply over the school years. Second, greater parental investment in tutoring is associated with higher cognitive achievement but lower child mental health, with the trade-off strongest among lower-income families. Third, parents systematically underperceive children's mental health problems, particularly in lower-SES families.

2.1 Data

2.1.1 Data Sources

The two main data sources used in this paper are the China Family Panel Studies (CFPS) and our original survey data on parent-child gaps in mental health perceptions.

CFPS The CFPS is a nationally representative, biennial longitudinal household and individual survey initiated in 2010 by the Institute of Social Science Survey at Peking University. Comparable to the Panel Study of Income Dynamics (PSID), the CFPS tracks individuals across waves. As of December 2024, seven waves of data had been released.

The baseline CFPS sample comprises approximately 15,000 households and 30,000 individuals, with a response rate of 79%. Five provinces were initially oversampled, with 1,600 families selected from each province to enable regional comparisons. The remainder of the sample is drawn from other provinces to achieve national representativeness through weighting. Overall, the CFPS represents 95% of China's population. The CFPS is key to our study for two main reasons. First, it measures cognitive ability—akin to the Armed Forces Qualification Test (AFQT) in the United States—and mental health for both adult and child respondents, enabling the construction of parent-child pairs. This allows us to examine how parental background—in terms of income, education, cognitive skills, and mental health—translates into child outcomes. We limit our sample to urban households with only one child in CFPS.

Self-Collected Parent–Child Survey Data We collected an original survey to study parental under-awareness of children's mental health difficulties. The central component of this dataset is a paired sample of 500 parent–child pairs, in which both parents and children answer comparable questions about the child's mental health and family interactions. These paired responses allow us to measure the extent to which parents' beliefs and observations diverge from children's own reports.

We supplement the paired sample with two nationally representative surveys. The first includes 1,500 parents whose child attended middle school or high school within the past five years. The second includes 1,500 young adults aged 18–25 who retrospectively report on their own middle- and high-school experiences. The young-adult sample is matched to the parent sample on parental income, parental education, and rural–urban composition. These broader samples help assess whether the reporting patterns observed in the paired data are consistent with patterns in the population.

2.1.2 Main Variables

Test Scores and Cognitive Skills The CFPS includes detailed cognitive ability assessments comparable to the AFQT in the US. It comprises four types of cognitive tests: literacy, math, word recall, and numerical series. The CFPS alternates between two sets of these assessments across waves: Set A (literacy and math tests) was used in 2010, 2014, and 2018, while Set B (word recall and numerical series tests) was used in 2012, 2016, and 2020. We focus on the 2010, 2014, and 2018 waves featuring Set A to ensure consistent cognitive ability measurement. We measure cognitive skills as the sum of math and literacy test scores. In the child survey, we can only observe children's cognitive ability scores starting at age 10, so we infer their average cognitive score before age 10

through extrapolation.[3] We normalize the average cognitive ability scores for the parent generation to be one. We report the age profile of child ability scores in Table 3 and also in Figure A8 in Appendix B.11.

Tutoring Expenditures and Parental Investment We focus on parents' monetary investments in children aged 6–18, primarily aimed at improving exam performance. Total parental monetary investment comprises two components: (i) in-school expenses, including tuition and fees, textbooks, uniforms, in-school events, transportation, insurance, meals, and boarding; and (ii) tutoring expenses, including spending on cram schools or private tutoring in core subjects such as math, Chinese, and English, which are emphasized in exams, especially the college entrance examination. Figure A4 in Appendix B.6 presents empirical evidence that parental time investment in China is not a primary concern. We adopt this definition in both CEPS and CFPS.

Mental Health Our main focus in measuring mental health is depressive mood. More specifically, we measure children's mental health capital, i.e., mental well-being, using six items from the negative emotion scale in the CEPS.[4] We measure adults (parents)' mental well-being and their children's mental well being for constructing intergenerational transmissions using four items from the Center for Epidemiologic Studies Depression Scale (CES-D) in the CFPS.[5] Note that, compared with other widely used scales for assessing the prevalence and severity of mental illness in epidemiology and psychiatry—such as the Kessler Psychological Distress Scale (K6)—the CES-D specifically screens for depressive symptomatology. It emphasizes depressed mood, anhedonia, somatic symptoms, and interpersonal difficulties, and it has a child version (CES-DC) adapted for children and adolescents. Also using CFPS, Liu et al. (2023) show that the shorter version of the CES-D (the CES-D8) is a robust instrument for measuring depressive symptoms in adolescents, especially in large-scale and longitudinal surveys.

To analyze panel responses to Likert-type items, we use the graded response model

[3]We use the average cognitive ability scores by age, ranging from age 10 to age 21, to extrapolate the average score for ages 6-9 using the Piecewise Cubic Hermite Interpolating Polynomial (PCHIP) method.

[4]The six questions are: In the past seven days, have you experienced the following feelings? (1) Depressed, (2) So down that you couldn't focus on tasks, (3) Unhappy, (4) Feeling that life is meaningless, (5) Lacking motivation to do things, and (6) Sad and upset. For each question, the response options were five ordered categories: (1) Never, (2) Rarely, (3) Sometimes, (4) Often, and (5) Always.

[5]The four anchor questions are: In the past seven days, have you experienced the following feelings? (1) I felt depressed, (2) I felt that everything I did was an effort, (3) I did not enjoy life, and (4) I felt sad. For each question, the response options are four ordered categories: (0) almost never (less than 1 day); (1) sometimes(1-2 days); (2) often(3-4 days); (3) most of the time(5-7 days). We provide more information about our measurement of mental health capital in Appendix A.2.

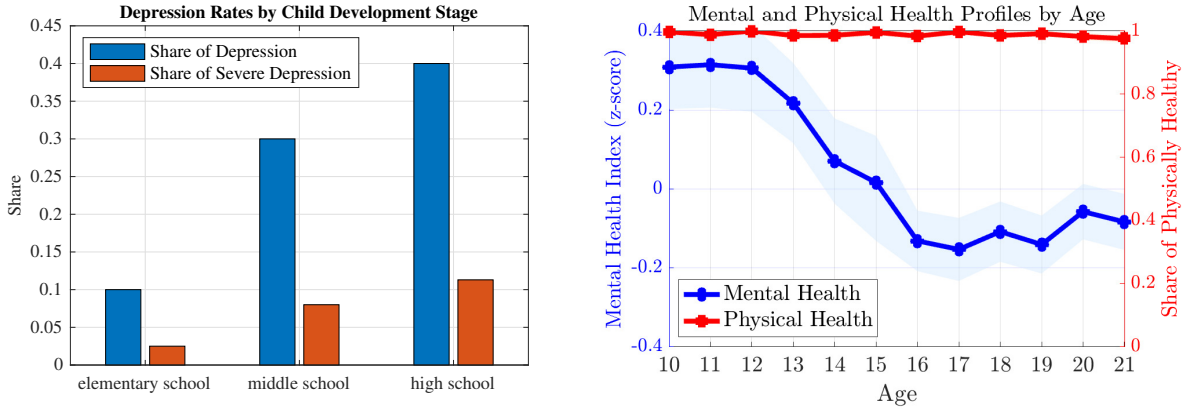
(GRM) from item response theory (IRT). The GRM assumes that the probability of endorsing each response category is a logistic function of an underlying latent trait (e.g., negative emotion states or depressive symptom severity). We estimate the model via maximum likelihood and then score respondents on the latent trait, yielding IRT-based continuous measures from the negative emotion scale or the CES-D. Unlike raw summed scores, which implicitly weight all items equally, IRT-based scores incorporate estimated item discriminations and category thresholds, so responses on more informative items contribute more to the trait estimate than responses on weaker items.[6] We recode the negative emotion scale and CES-D so that higher scores indicate fewer negative emotions and a lower likelihood of mental illness, reflecting greater mental health capital.

Measuring Perception Gap To measure parent–child perception gaps, we collected paired subjective reports from parents and children using parallel questionnaires. The central idea is to compare what parents believe or observe about their child's mental health and family interactions with what children report about their own experiences. The parent questionnaire elicits parents' assessments of how well they understand the child's psychological state, whether the child proactively discusses stress or emotions with them, whether the child has recently exhibited symptoms such as depressed mood, anxiety, sleep problems, fatigue, or school avoidance, how parents view the trade-off between academic performance and mental health, and how they respond when the child shows signs of mental health problems. The child questionnaire asks the corresponding questions from the child's perspective, including whether parents understand the child's true stress and emotional condition, whether the child can talk to parents about stress, the child's own recent emotional state, whether parents prioritize grades over mental well-being, and whether parental responses to mental health problems are helpful. This paired design allows us to quantify discrepancies between parental perceptions and children's self-reports.

We then use these matched reports to construct a measure of parental underawareness of child mental health problems. Specifically, we focus on five parent–child matched questions on recent child mental health symptoms: depression or loss of interest, anxiety, sleep problems, fatigue, and school aversion. For each symptom, the parent is asked how

[6] IRT models estimate a latent trait and item parameters (discriminations and thresholds), improving measurement precision and enabling person-level standard errors. Because the latent trait (e.g., depression severity) is not directly observable, each item serves as a noisy indicator, with response probabilities depending on the trait level. In the GRM for ordered categories, each item's discrimination reflects how strongly it relates to the trait, and its thresholds locate category boundaries along the trait continuum. Individual latent trait scores are inferred from the full response pattern and can be transformed to user-friendly scales (e.g., T-scores).

often the child has appeared to experience the symptom during the past month, while the child is asked how often they themselves experienced that symptom over the same period. Response options are identical across the two questionnaires—never, sometimes, often, and almost always—so that answers are directly comparable. Our main symptom-level underawareness measure is defined as the share of parent–child pairs in which the child reports experiencing a symptom often or almost always, while the parent does not report that symptom at the same level. This measure captures cases in which children report substantial distress that is not fully recognized by their parents.



(a) National-level Data (b) Household Survey Data

Figure 1: Child Mental Health by Age

Data Source: Panel (a): Report on National Mental Health Development in China (2019-2020). Panel (b): CFPS 2010-2022. *Notes:* Parental mental health capital is defined in Section 2.1.2 and is measured as a continuous variable.

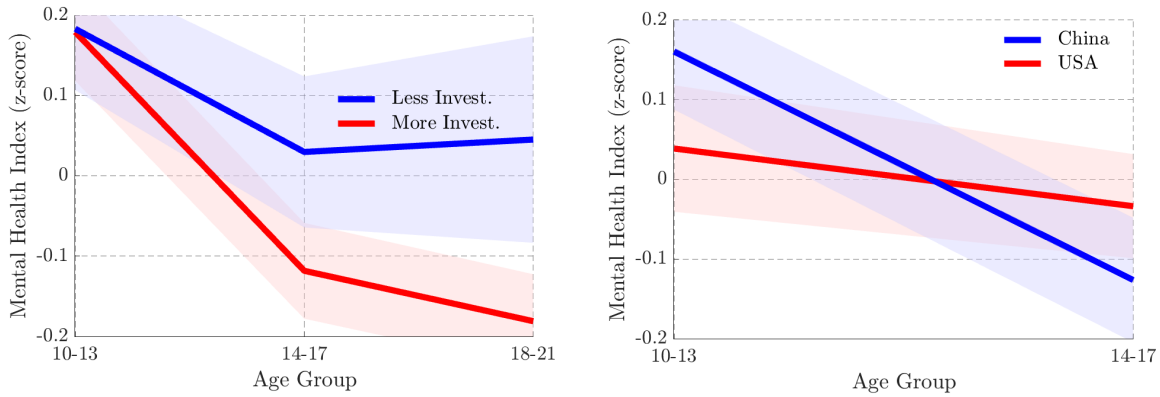
2.2 Stylized Facts

Fact 1: Mental health Deteriorates over Adolescence.

In Figure 1(a), we plot depression rates by school level using data from the Report on National Mental Health Development in China (2019–2020) published by the Chinese Academy of Sciences. Both depression and severe depression become more prevalent as children progress from elementary to middle to high school. This pattern coincides with rising academic pressure as students approach the college entrance examination.

Next, we check whether we could observe similar patterns in our household-level data. In Figure 1(b), we plot age profiles of the mental and physical health indices over childhood and adolescence using the CFPS. The mental health index is constructed as described in Section 2.1.2. Higher values indicate better health. The physical health index

(share of physically healthy individuals) is constructed based on self-reported health status.[7] As shown in the figure, the physical health index is relatively stable during adolescence, reflecting generally good physical health among young people. By contrast, the mental health index declines sharply between ages 10 and 17—when pressure from high-stakes examinations intensifies, cram-school attendance peaks, and parental-monetary-investment-to-income ratio surges (see Figure 11(B.2)).[8] This underscores why we want to focus on mental health during child and adolescent stage.



(a) Monetary Investment Intensity (b) China vs USA

Figure 2: Mental Health Worsened Dramatically During Adolescence

Data Source: CFPS and PSID. *Notes:* Parental mental health capital is defined in Section 2.1.2 and is measured as a continuous variable.

One might argue that the sharp decline in mental health during adolescence is not primarily driven by academic pressure from intensive tutoring, but is instead a natural consequence of developmental changes such as puberty. While we agree that biological and developmental factors might be part of the reasons, they are unlikely to be the whole story. In Figure 2(a), we divide children into two groups based on the intensity of parental monetary investment in education, using investment/household income = 0.2 as the cutoff. Mental health worsens during adolescence in both groups, but the decline is substantially larger in the high-investment group. In Figure 2(b), we compare the same mental health measure for adolescents in China and the United States using CFPS 2018 and PSID 2019, respectively. China, where education is more competitive due to score-based college admissions and strong cultural emphasis on schooling (including the

[7] More specifically, we rely on self-reported health status in the CFPS, with "healthy" coded as 1, "fair" as 0.5, and "unhealthy" as 0.

[8] In Figure A2 of Appendix B.5, we extend the age profiles to age 62 and find that the mental health index stabilizes in middle age and slightly increases from around age 50 onward, exhibiting an overall hump-shaped pattern with respect to age, which aligns with findings in the psychiatry literature (Gallo, Anthony, and Muthen, 1994; Husain et al., 2005).

widespread use of cram schools and tutoring), exhibits a much larger decline in mental health during adolescence than the U.S., even though both countries show a downward trend. Our findings also align with Chen, Currie, Ding, and Lo (2024) who causally identify that school-related stress periods—especially school-entry milestones and high-stakes educational transitions—raise children's mental health treatment, implying that academic stress and school pressure negatively affect child mental health.

Fact 2: Exam-oriented Investment Creates a Trade-off between Child Cognitive Skills and Mental Health

As shown in Figure 3, when we divide parental monetary investment in children (with tutoring expenditures comprising a large share) into quintiles, children's cognitive skills rise with higher investment, while their mental health capital declines. This pattern may indicate an investment trade-off between child cognitive skills and mental health capital.

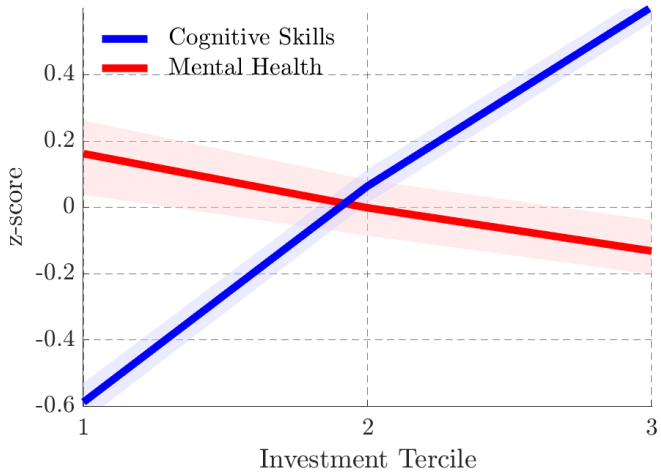


Figure 3: Investment Trade-offs between Child Cognitive Skills and Mental Health Capital

We then plot children's cognitive (blue) and mental (red) health z-scores against parental tutoring investment terciles, separately by income quartile, in Figure 4. In the lowest-income quartile (Q1), higher tutoring investment is strongly associated with better cognitive outcomes but clearly worse mental health, revealing a sharp trade-off; this trade-off progressively weakens as income rises, and almost disappears by Q4. This pattern suggests that higher-income families can cushion the mental health costs of tutoring—likely through other forms of "investment" such as better home environments and enrichment activities—so low-income children face a double whammy: they need high tutoring to keep up academically, but each extra unit of tutoring carries larger mental health costs.

Accounting for parental mental health, Figure 5 shows that cognitive skills continue to increase with investment and mental health capital continues to decline. However, parental mental health appears to have a protective effect on children's mental health: when tutoring-intensive investment is associated with lower child mental health, higher parental mental health can partly offset this deficit. For example, among children with intensive cram-school attendance (parental monetary investment in the top tercile), those with parents in the top 25% of mental health capital have mental health z-scores that are slightly higher than those of children who attend very little cram school (investment in the bottom tercile) but whose parents are in the second quartile of mental health capital. In addition, after controlling for investment tercile, children's cognitive skills do not meaningfully vary with parental mental health.

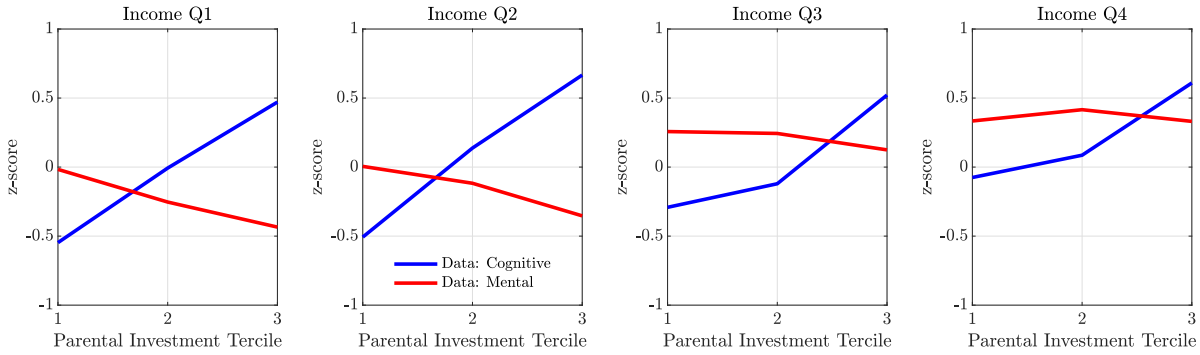


Figure 4: Investment Trade-offs for Varied Income Quartiles

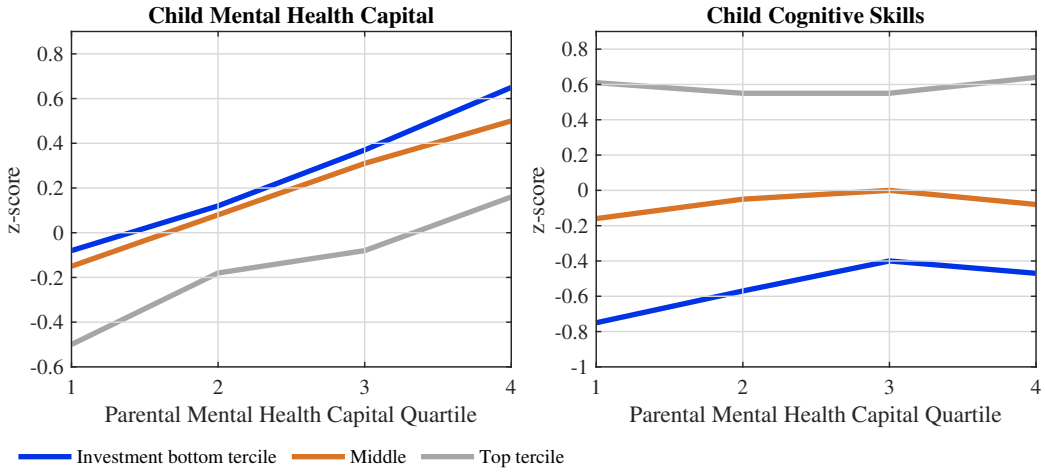
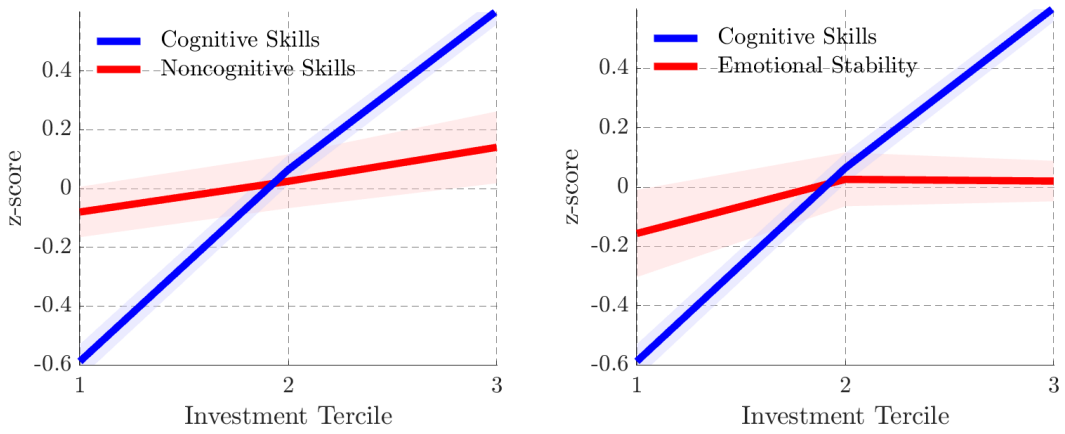


Figure 5: Child Mental Health by Parental Mental Health and Monetary Investment

Finally, we examine whether similar trade-off patterns arise for other psychologically related measures. In Figure 6, we show that the investment trade-off documented for



(a) Noncognitive Skills (b) Emotional Stability

Figure 6: Investment Patterns for Other Psychologically Related Measures

mental health does not appear for these other outcomes. Along the x-axis, investment terciles increase, and in both panels the blue line (cognitive skills) rises steeply with investment. The red lines—noncognitive skills (measured using the Rosenberg Self-Esteem Scale) in panel (a) and emotional stability (the inverse of neuroticism in the "Big 5") in panel (b)—also slope slightly upward, with confidence bands that overlap zero but never show a clear decline. Thus, higher parental investment improves cognitive outcomes without worsening these other psychological traits, suggesting that the earlier trade-off is specific to particular mental-health outcomes (in particular, depression) rather than a general feature of all noncognitive or personality dimensions.

Fact 3: Parents systematically underperceive children's mental health problems, especially in lower-SES families.

Figure 7 provides direct evidence that parents systematically underperceive their children's mental health problems. Comparing matched parent and child reports, we find large and consistent discrepancies in perceived distress, parent–child communication, and parental responses to mental health difficulties. Children are far more likely than parents to report that parents prioritize academic performance over mental well-being, view mental health costs as acceptable in exchange for exam success, and respond to distress by pressuring the child to continue working. Conversely, parents are much more likely than children to report that children can easily communicate stress and emotions and that parents understand the child's true mental state. Even for children's own symptoms, parents underreport distress: the share reporting that the child often or always felt depressed or anxious is 21.7 percentage points lower in parent reports

than in children's self-reports. Overall, the figure points to a broad pattern of parental underrecognition of adolescent psychological strain.

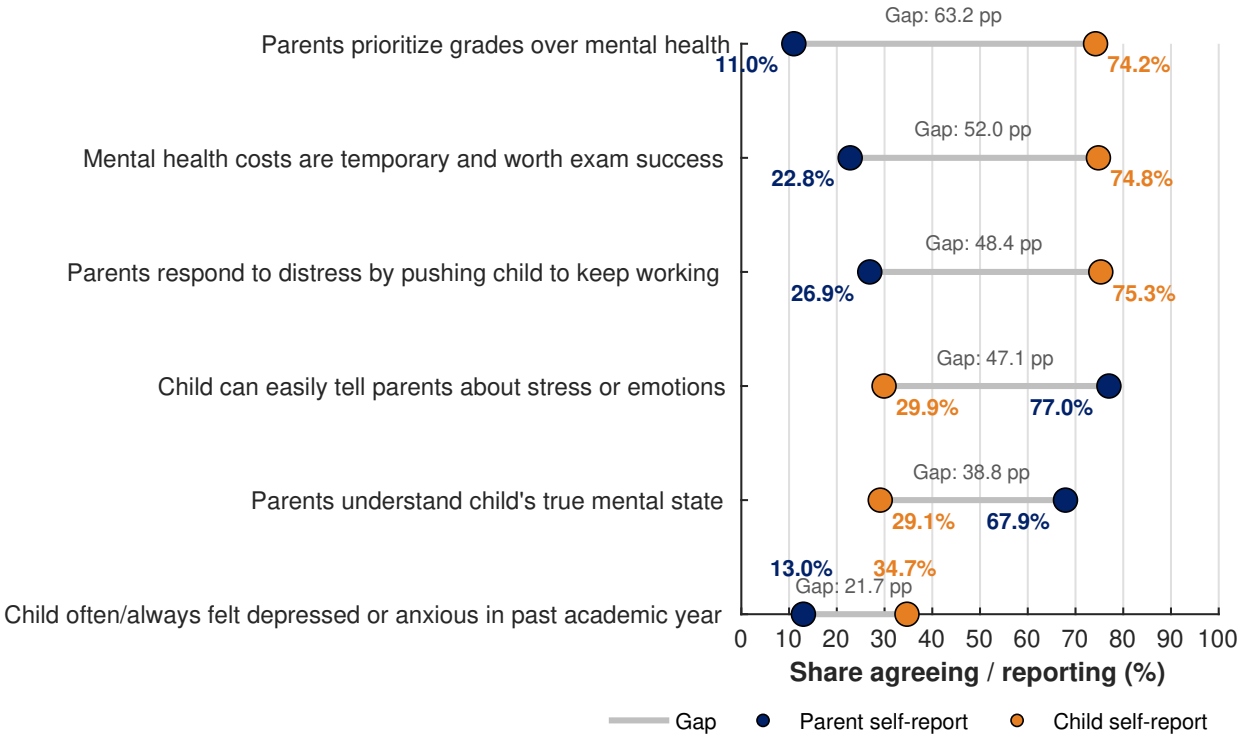


Figure 7: Parent–child gaps in reported symptoms, communication, and responses to distress

Figure 8 further shows that this underrecognition is concentrated among socioeconomically disadvantaged households. Using symptom-specific parent–child mismatches as a measure of underawareness, we find that children with non-college-educated parents are more likely to report frequent sadness, anxiety, sleep problems, fatigue, and school aversion that their parents do not recognize. The differences are especially large for fatigue and school aversion, and strikingly so for school aversion, where the underawareness rate among non-college-educated parents is more than five times that among college-educated parents. This SES gradient suggests that parental awareness of children's mental health is itself unequally distributed, implying that disadvantaged families may be less able to detect or respond to psychological distress during adolescence.

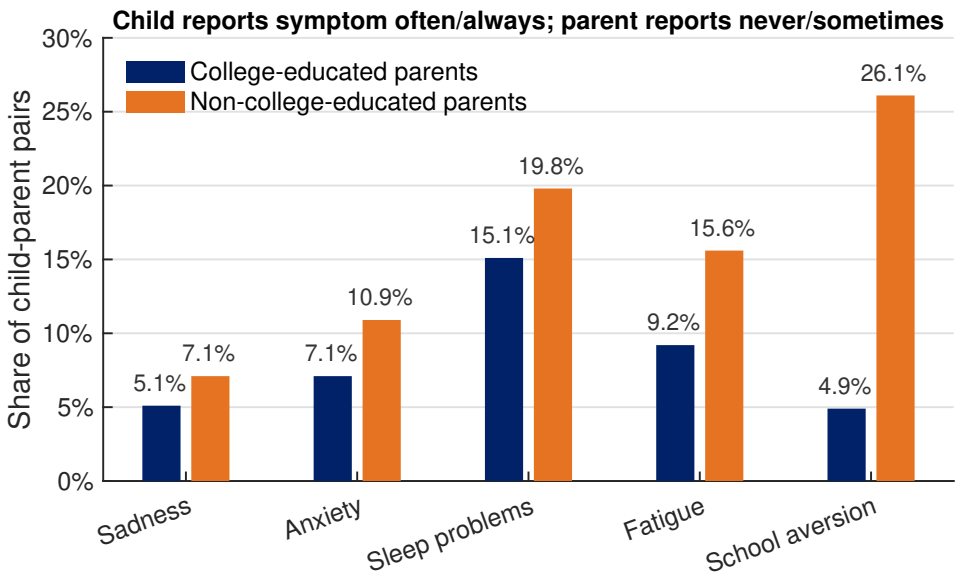


Figure 8: Parental Underawareness by Parent Education

3 Model

In this section, we introduce a general equilibrium (GE) heterogeneous-agent life-cycle model with uninsurable idiosyncratic shocks and incomplete markets *à la* Huggett (1996) and altruistic parents caring about their descendants' utility *à la* Barro and Becker (1989). This model comprises three agent types: heterogeneous households, a representative firm, and the government, including a continuum of homogeneous public colleges.

At the heart of our framework is the joint formation of two types of human capital during childhood and adolescence and how parental monetary investment shapes their accumulation. Cognitive skills evolve via a multi-period production function in which next-period cognitive skills depend on current cognitive skills, parental cognitive skills, parental monetary investment, and public education spending. Next-period mental health depends on current mental health, parental mental health, parental monetary investment, and an exogenous family-background component. When parental monetary investment exceeds a threshold—interpreted as intensive tutoring or cram-school participation—it has a negative effect on current mental health.

3.1 Demographics and Environment

Time in the model is discrete, with one period corresponding to four years. Before a child becomes independent, she does not make any choices. The model follows a dynastic framework with four stages: childhood, college, work/parenthood, and retirement. Figure 9 illustrates the life-cycle events for a sample parent (from independence to death)

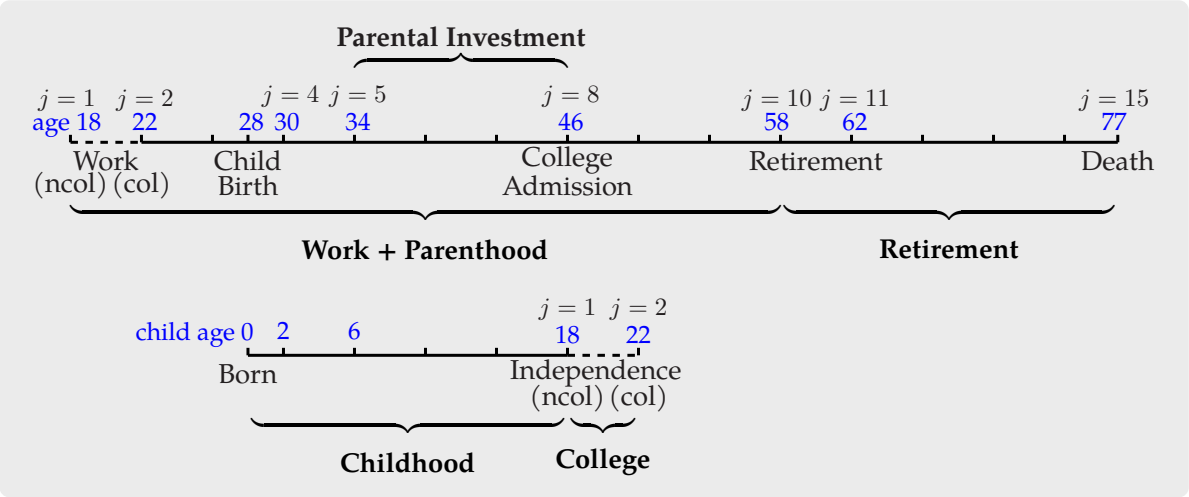


Figure 9: Timeline of the Model

and child (from birth to independence) in the economy. Let j denote age in periods ($j = 1$ refers to ages 18-21, $j = 2$ to ages 22-25, etc.). The adult agent becomes independent and enters the labor market at either period $j = 1$ or $j = 2$ (age 18 or 22), working until retirement at periods $j = 10$ or $j = 11$ (age 58 or 62), depending on whether or not she attends college. They then live for four periods after retirement and die at the end of period $j = 15$ (age 77).

In all periods, the adult agent makes consumption-savings choices, subject to uninsurable idiosyncratic labor productivity shocks. They save through a non-state-contingent asset. The child is born when the adult agent reaches the age of 28 (child age 0), corresponding to the middle between periods $j = 3$ and $j = 4$. The fertility is deterministic such that each individual parent has only one child. The model abstracts from the early childhood development stage. Parents start to invest in their children from the beginning of the period $j = 5$ (adult age 34-37, child age 6-9) until the end of the period $j = 7$ (just before a child reaches the age 18) to improve children's cognitive skills h.However, this investment also influences the evolution of the child's mental health state m. At the beginning of $j = 8$ (child's age 18), parents make college decisions for their children based on the child's cognitive skills h as well as household income and wealth a. If the value of attending college surpasses that of not attending, they will send their children to the college entrance examination, where the admission probability depends solely on the child's cognitive skills h relative to others in the same cohort. There is no population growth and no survival risk. We summarize the individual state variables in Table 1.

Table 1: Individual State Variables

State variable	Description
j	Model age $j \in \{1, \ldots, 15\}$
δ_p	Parental types $\delta_p \in \{I, U\}$
s	Education of adults $s \in \{col, ncol\}$
a	Assets
h_p	Cognitive skills of adults
m_p	Initial mental health of adults
η	Mental health shocks of adults
z	Idiosyncratic labor productivity shocks
h	Cognitive skills of children
m	Mental health of children
s_c	Education of children $s_c \in \{col, ncol\}$

3.2 Wage Income Process

An individual enters the labor market with ex ante fixed heterogeneity formed during the pre–labor market stage (childhood). This heterogeneity consists of education $s \in \{col, ncol\}$ (where col denotes college and $ncol$ denotes non-college), cognitive skills h_p, and initial mental health m_p. Cognitive skills h_p are fixed after labor market entry, whereas mental health is subject to a persistent shock η_j, which follows an AR(1) process:

$$\log(\eta_j) = \rho_{m,p} \log(\eta_{j-1}) + \epsilon_{m,p}, \qquad \epsilon_{m,p} \sim N(0, \sigma_{m,p}), \tag{1}$$

An individual's income depends on education $s \in \{col, ncol\}$ and labor efficiency units y_j. Specifically, y_j consists of the education-specific deterministic age profile A_j^s, cognitive skills h_p, mental health $m_p \cdot \eta_j$, and idiosyncratic labor productivity shocks z_j. That is,

$$\log y(h_p, m_p, \eta_j, A_j^s, z_j) = \lambda \log h_p + \psi_{m,p} (\log m_p + \log \eta_j) + A_j^s + \log z_j, \tag{2}$$

where λ denotes the return to cognitive skills, $\psi_{m,p}$ captures the effect of mental health capital on labor efficiency units, and z_j follows a Markov chain with state space $z = \{z_1, z_2, \ldots, z_M\}$, stationary distribution μ_z, and transition probabilities $\pi(z' \mid z) > 0$. Specifically,

$$\log z_j = \rho_z \log z_{j-1} + \epsilon_z, \qquad \epsilon_z \sim N(0, \sigma_z).$$

We abstract from two other types of human capital accumulation: (1) human capital

acquired in college;[9] and (2) human capital accumulated on the job.[10]

3.3 Preferences

Households are risk-averse, supply labor inelastically, and discount the future at rate β. They have identical preferences over consumption c and mental health $m_{p,\eta} = m_p \cdot \eta$, where mental health is the product of an ex ante component m_p and an ex post shock η. Preferences are represented by

$$u(c, m_{p,\eta}) = \log((1 + \kappa m_{p,\eta})c).$$

We model altruism à la Barro and Becker (1989), whereby agents care about the utility of their descendants.

3.4 Parental Types

Parents differ in the extent to which they are informed about their children's mental health. We assume two types of parents: informed (I) and uninformed (U). Informed parents observe their child's mental health state, understand how mental health evolves during childhood, and internalize the negative effect of tutoring on child mental health, as well as its implications for later adult mental health, labor market outcomes, and utility. By contrast, uninformed parents do not observe their child's true mental health state and instead base their investment decisions on the population mean level of child mental health. They do not understand the evolution of mental health during childhood and therefore fail to internalize the detrimental effect of tutoring on child mental health and its impacts on later-life outcomes. At childbirth, each parent draws their information type, δ_p, from an education-specific distribution. That is, $\delta_p \in \{I, U\}$, where $\Pr(\delta_p = U) = \Pi_{U,s}$. In other words, $\Pi_{U,s}$ captures the share of uninformed parents in the economy and varies with parental education. The share of uninformed parents is lower among college-educated parents than among non-college-educated parents. That is, $\Pi_{U,col} < \Pi_{U,ncol}$.

[9] One potential concern is that college education may improve cognitive ability. However, in Chinese data, we find, as shown in Figure A8, that an individual's cognitive test scores peak at around age 18 (the typical age at which students take the college entrance examination) and decline afterward. This suggests that college education may not improve test-measured ability, although it is likely to improve ability in other dimensions.

[10] The existing macro literature, both without endogenous initial conditions (e.g., Keane and Wolpin (1997); Huggett, Ventura, and Yaron (2011)) and with endogenous initial conditions (e.g., Lee and Seshadri (2019); Daruich (2022); Daruich and Kozlowski (2020)), has shown that most of the variation in lifetime income is attributable to conditions present before labor market entry. Guvenen et al. (2022), using a 57-year panel from the U.S. Social Security Administration, also supports this conclusion.

3.5 Firm Sector

We assume there is a representative firm that solves a static profit maximization problem as follows:

$$\max_{H_{col},H_{ncol},K} \left\{ Y - w^{col} H_{col} - w^{ncol} H_{ncol} - (r + \delta_k + \iota_o) K \right\} \tag{3}$$

Production technology is $Y = AK^{\Omega} H^{1-\Omega}$, where A denotes aggregate productivity, K is aggregate capital input, Ω denotes capital share, and $H = \left[\phi H_{col}^{\psi} + (1-\phi) H_{ncol}^{\psi} \right]^{\frac{1}{\psi}}$ denotes aggregate labor input, which is a CES aggregator of college labor H_{col} and non-college labor H_{ncol}. Specifically, $H_{col} = \sum_{j=2}^{11} \int_{X_j} y_j^{col}(x_j)\, d\Psi_j$, $H_{ncol} = \sum_{j=1}^{10} \int_{X_j} y_j^{ncol}(x_j)\, d\Psi_j$, where $x_j \in X_j$ denotes the age-specific state vector of an individual of age j, as defined in Table 1. The term ϕ captures skill (college labor) intensity, $\frac{1}{1-\psi}$ captures the elasticity of substitution between college and non-college labor where $\psi \in (-\infty, 1]$, and $(r + \delta_k + \iota_o)$ is the capital rental cost, where ι_o captures financial intermediation cost.

3.6 Public Sector

The public sector comprises two components: the government and the college sector.[11]

Colleges In our benchmark model, the college sector consists of homogeneous colleges, each with a fixed number of slots that determine the total capacity, denoted by Z (the fraction of a cohort that can attend college). We exclude low-quality and less competitive 4-year private colleges and 3-year tertiary vocational colleges from our consideration. We begin with this no-quality-differentiation setup because our primary goal is to quantify the effects of college expansion in China. In contrast to countries like the U.S., where overall college capacity is high and nearly everyone who wants to attend college can do so, the primary consideration is quality rather than availability when modeling college attendance. However, in China, college capacity was extremely limited in the 1990s. Even today, after more than two decades of expansion, admission to second-tier public universities remains highly competitive, with demand far exceeding supply (let alone elite institutions).

Government The centralized government admits students based on the ranks of cognitive skills at the beginning of period $j = 8$, given fixed college capacity Z. In our model, this admission policy is manifested as follows. The government takes Z as given

[11] In China, the higher education landscape is primarily characterized by public colleges and universities with central-planned nature, which is why our analysis focuses on the public college sector.

and adjusts the shifter ζ in the admission function $\chi(\cdot, \zeta)$ to ensure the college capacity constraint is satisfied (Equation (**??**) in Appendix **??** holds), where $\chi(h, \zeta)$ is increasing in an individual's cognitive skills h, specifying the admission probability of each level of h relative to the cohorts. Moreover, the government collects tax revenues in terms of a proportional consumption tax at rate τ_c and provides subsidies for college students $\varsigma_{college}$ and public investment in children's cognitive skills by age, $g_j, j \in \{5, 6, 7\}$. Note that τ_c is pinned down when the government's budget is balanced (Equation (**??**) holds), and $\varsigma_{college}$ and g_j are exogenously given.

3.7 Recursive Problems

3.7.1 Working Stage without Children

During this stage (starting from $j = 1$ for individuals who do not go to college or $j = 2$ for those who go to college), individual households face a standard life-cycle problem. Given the linear consumption tax τ_c, households make consumption-saving decisions based on education s, assets a, cognitive skills h_p, initial mental health m_p, mental health shocks η, and idiosyncratic shocks z:

$$v_j(s, a, h_p, m_p, \eta, z) = max_{c,a'} \left\{ u(c, m_{p,\eta}) + \beta \mathbb{E}_{z'} \left[v^s_{j+1}(s, a', h_p, m_p, \eta', z') \right] \right\} \tag{4}$$

$$\text{subject to} \quad (1 + \tau_c)\, c + a' = (1 + r)a + w^s y(h_p, m_p, \eta, A^s_j, z), \quad c, a' \geq 0. \tag{5}$$

3.7.2 Working Stage with Children and Parental Investment

At the beginning of $j = 5$, a child is born with some cognitive skill endowment h_o, and some mental health capital endowment m_o. Our model abstracts from early childhood stages and focus exclusively on children aged 6-17, a critical stage when they begin attending school, prepare for college entrance exams, and receive significant parental monetary investments for improving exam performance. Child cognitive skills at age 18 (beginning of $j = 8$) is the outcome of parental monetary inputs i, government inputs g_j, $j = 5, 6, 7$, and cognitive skills at an early stage. Parental monetary investment, however, simultaneously affects the evolution of the child's mental health. We assume that the child shares household consumption c and does not make any independent decisions relevant to the household's economic status during childhood.

Intergenerational Transmission of Skills A child's cognitive skill endowment h_o is stochastic but correlated with the parent's cognitive skills h_p by parental education level,

which is specified as follows:

$$h_o = (1-\rho_h)\mu_h^s + \rho_h h_p + \epsilon_h^s, \;\; \epsilon_h^s \sim N(0, \sigma_h^s) \;\; s \in \{col, ncol\} \tag{6}$$

where ρ_h captures intergenerational persistence of cognitive skills, μ^s is the average child cognitive skill endowment with parental education s, and σ_h^s is the standard deviation of intergenerational genetic shocks for parental education s. The heterogeneity in h_o stems from both genetic influences and early childhood development outcomes, which are treated as exogenous in our model.

Likewise, a child's mental health capital endowment is stochastic but correlated with the parent's mental health capital m_p by parental education level, which is specified as follows:

$$log(m_o) = (1-\rho_m)\mu_m^s + \rho_m log(m_p) + \epsilon_m^s, \;\; \epsilon_m^s \sim N(0, \sigma_m^s) \;\; s \in \{col, ncol\} \tag{7}$$

Child Cognitive Skill Production Function The next-period cognitive skills depends on the current-period cognitive skills h, parental cognitive skills h_p, parental investment i, and age-specific public investment g_j that is homogeneous across children:

$$h' = \theta_j \left[\alpha_j h^\gamma + \xi h_p^\gamma + (1-\alpha_j-\xi) I_j^\gamma\right]^{\frac{1}{\gamma}}, I_j = \left[\eta\, i^\mu + (1-\eta)\, g_j^\mu\right]^{\frac{1}{\mu}} \tag{8}$$

where θ_j is the age-specific efficiency of child skill production technology, α_j captures the age-specific self-productivity, ϖ denotes the share of parental human capital in child skill production, γ is the elasticity of substitution between h and the *effective investment* I_j, a CES aggregation of private investment i and public investment g, where η denotes share of i relative to g, and μ denotes the elasticity of substitution between i and g.

Child Mental Health Production Function The next-period mental health of children m depends on the current-period mental health m, current-period parental monetary investment i, current-period parental mental health $m_{p,\eta}$ and income, and idiosyncratic shocks. In particular, parental background affects the development of children's mental health through two main channels. First, parental monetary investment influences children's mental health via cram-school enrollment. Parents enroll their children in a cram school if their monetary investment i exceeds a threshold $\bar{i}$. We choose $\bar{i}$ so that the share of households with $i > \bar{i}$ matches the cram-school participation rate observed in the data, defined as the share of parents who purchase extracurricular tutoring in addition to other educational expenditures. This modeling choice reflects the idea that parents prioritize

in-school necessities—such as stationery, books, boarding, and transportation—before allocating funds to extracurricular tutoring. Second, children from better family backgrounds on average have better mental health. As shown by Cunha, Heckman, and Schennach (2010), advantaged family backgrounds can positively affect non-cognitive skills through better "investments" captured by a better home environment, greater emotional support, more planned activities, and richer family surroundings. Thus, family background can affect a child's mental health through channels other than tutoring expenditures. For example, better-educated parents with higher income may have more effective parenting skills or a greater capacity to provide enriching experiences—such as travel, museum visits, or concerts—which can all contribute positively to children's mental health. We model this channel exogenously; details are provided below.

The evolution process of child mental health is specified as follows:

$$\begin{aligned}\log(m') = \varsigma_{m,c}\log(m_p) + \rho_{m,c}\log(m) + \varpi_{m,c,j} + \vartheta_{m,c}\log(w^s y(h_p, m_p, z)) \\ +\phi_{m,c}\log(1+\max\{i-\bar{i},0\}) + \epsilon'_{m,c},\ \epsilon'_{m,c} \sim N(0,\sigma_{m,c})\end{aligned} \tag{9}$$

where $\varsigma_{m,c}$ measures how child mental health depends on parental mental health in each period, $\rho_{m,c}$ denotes the persistence of mental health capital across periods, $\phi_{m,c}$ captures the effect of cram-school spending on mental health. Note that the strength of the effect is determined not only by the extensive margin (the decision to participate in tutoring) but also by the intensive margin (the level of tutoring expenditure, i.e., the portion where $i > \bar{i}$).[12] The component $(\varpi_{m,c,j} + \vartheta_{m,c}\log(w^s y(h_p, m_p, z)))$ captures how family background affects child mental health through channels other than tutoring expenditures. More specifically, our model abstracts from investments that directly improve child mental health. If it included only the negative impact of private tutoring on child mental health, without other investment channels that positively affect the evolution of child mental health, then the age profile of child mental health would be much lower at every period j than in the data. Hence, we introduce this component exogenously. Note that, in modeling both the production of children's cognitive skills and the evolution of children's mental health, we do not include cross-effects between cognitive skills and mental health (e.g., we do not allow next-period cognitive skills to depend on current-period mental health). We make this choice because we find that,

[12]One might argue that child mental health is more likely to be adversely affected by the time spent on homework and tutoring rather than by monetary expenditures. While we acknowledge this point, for the sake of model tractability, we capture the detrimental effect on mental health through tutoring expenditures. This is justified by the strong positive correlation between tutoring time and monetary outlays, as shown in Figure A7 of Appendix B.7. Specifically, when parental expenditures on tutoring are low, the time children spend in tutoring is also negligible.

for children aged 10–18, cognitive skills depend very little on non-cognitive skills, and vice versa. This is consistent with Cunha, Heckman, and Schennach (2010)'s estimates of cross-productivity between cognitive and non-cognitive skills.

Recursive Problem for Informed Parents Informed parents ($\delta_p = I$) have perfect information, so they make decisions based on their own characteristics and their child's cognitive skills h and mental health m for $j = 5, 6, 7$:

$$v_j^I(s, a, h_p, m_p, \eta, z, h, m) = \max_{c,a',i} \left\{ u\left(\frac{c}{1+\iota}, m_{p,\eta}\right) + \beta \mathbb{E}_{z'} \left[v_{j+1}(s, a', h_p, m_p, \eta', z', h') \right] \right\} \tag{10}$$

$$\text{subject to} \qquad (1+\tau_c)c + a' + i = (1+r)a + w^s y(h_p, m_p, \eta, A_j^s, z), \qquad c, a', i \geq 0,$$

together with Equations (6), (7), (8), and (9).

Recursive Problem for Uninformed Parents Uninformed parents ($\delta_p = U$) do not observe their child's true mental health state and instead base their investment decisions on the population mean level of child mental health. They do not understand the evolution of mental health during childhood and therefore fail to internalize the effect of tutoring on child mental health and its impact on later-life outcomes. Thus, for $j = 5, 6, 7$, they make decisions based only on their own characteristics and their child's cognitive skills h:

$$v_j^U(s, a, h_p, m_p, \eta, z, h) = \max_{c,a',i} \left\{ u\left(\frac{c}{1+\iota}, m_{p,\eta}\right) + \beta \mathbb{E}_{z'} \left[v_{j+1}(s, a', h_p, m_p, \eta', z', h') \right] \right\} \tag{11}$$

$$\text{subject to} \qquad (1+\tau_c)c + a' + i = (1+r)a + w^s y(h_p, m_p, \eta, A_j^s, z), \qquad c, a', i \geq 0,$$

together with Equations (6) and (8). Note that although uninformed parents do not know Equations (7) and (9) when making decisions, their child's mental health still evolves according to Equation (9).

3.7.3 College Attendance Stage

At the beginning of period $j = 8$ (child age 18), a parent compares the value of attending college $v(\cdot, s_c = col)$ with that of not attending $v(\cdot, s_c = ncol)$ for her child where $s_c \in \{col, ncol\}$ denotes education outcomes of children. If $v(\cdot, col)$ surpasses $v(\cdot, ncol)$, parents send their children to take the college entrance exam. For a child with $v(\cdot, col) > v(\cdot, ncol)$, their admission outcome relies solely on their cognitive skills at the start of $j = 8$, h_8, relative to that of other children, which is governed by the admission function $\chi(\cdot, \zeta)$.

We abstract our model from inter vivos transfers,[13] and from the option to go to college overseas, as this alternative is mainly accessible to privileged families and represents a very small portion of the total population.[14]

Let $x_j \in X_j$ be the age-specific state vector of an individual of age j. Then x_8 for informed parents ($\delta_p = I$) is $x_{8,\delta_p=I} = (s, a, h_p, m_p, \eta, z, h, m)$, and for uninformed parents ($\delta_p = U$) is $x_{8,\delta_p=U} = (s, a, h_p, m_p, \eta, z, h)$. Given $\chi(\cdot, \zeta)$, individual parents solve the following problem:

$$\begin{aligned} v_8(x_8) = {} & v_8(x_8, ncol)\mathbb{1}_{v(\cdot,col)\leq v(\cdot,ncol)} \\ & + \chi(h,\zeta)v_8(x_8, col)\mathbb{1}_{v(\cdot,col)>v(\cdot,ncol)} + (1-\chi(h,\zeta))v_8(x_8, ncol)\mathbb{1}_{v(\cdot,col)>v(\cdot,ncol)}. \end{aligned} \tag{12}$$

Informed Parents The value of informed parents whose children do not go to college $v^I(\cdot, ncol)$ is:

$$v_8^I(x_8, ncol) = max_{c,a'} \left\{ u(c, m_{p,\eta}) + \beta \underbrace{\mathbb{E}_{z'}\left[v_9(s, a', h_p, m_p, \eta', z')\right]}_{\text{Parent value}} + \nu \underbrace{\int v_1(ncol, 0, h, m, z) d\mu_z(z)}_{\text{Child value}} \right\} \tag{13}$$

subject to (5) and the value of informed parents whose children go to college $v^I(\cdot, col)$ is specified as:

$$v_8^I(x_8, col) = max_{c,a'} \left\{ u(\frac{c}{1+\iota}, m_{p,\eta}) + \beta \underbrace{\mathbb{E}_{z'}\left[v_9(s, a', h_p, m_p, \eta', z')\right]}_{\text{Parent value}} + \nu\beta \underbrace{\int v_2(col, 0, h, m, z) d\mu_z(z)}_{\text{Child value}} \right\} \tag{14}$$

$$\text{subject to} \qquad (1+\tau_c)\, c + a' + \kappa = (1+r)a + w^s y(h_p, m_p, \eta, A_j^s, z),\ c, a' \geq 0. \tag{15}$$

where ν captures parents' altruistic motives and κ denotes college tuition paid by parents. Note that children become independent at age of 18 if they do not go to college, and at age of 22 if they go to college, so they have separate value function from their parents where they start with zero assets $a = 0$ and draw labor productivity shocks z from μ_z, and the state variable h no longer appears in parents' value function. For children who

[13] Numerous studies highlight the important role of inter vivos transfers in college access and mobility in the U.S. (Abbott et al., 2020; Daruich and Kozlowski, 2020). These transfers primarily serve to finance children's costly college tuition. In China, however, inter vivos transfers are scarce during the college decision-making stage, as parents directly cover their children's tuition. See Appendix B.10 for more details.

[14] For example, in 2019, around 1,194,900 Chinese students studied abroad (including all education levels), which is about 0.43% of the total urban population aged 7-25 in China (estimated at 280.07 million).

go to college, they spend four years in college, so they enter the labor market at model period $j = 2$; accordingly, parents must finance their consumption for an additional four years (one model period) and pay the college tuition.

Uninformed Parents $v^U(\cdot, ncol)$ and $v^U(\cdot, col)$ are defined analogously to (13) and (14), except that parents no longer evaluate child conditioned on child individual mental state m. That is,

$$v_8^U(x_8, ncol) = \max_{c,a'} \Bigg\{ u(c, m_{p,\eta}) + \beta \underbrace{\mathbb{E}_{z'}\left[v_9(s, a', h_p, m_p, \eta', z')\right]}_{\text{Parent value}} + \nu \underbrace{\iint v_1(ncol, 0, h, m, z)\, d\mu_m(m)\, d\mu_z(z)}_{\text{Child value}} \Bigg\} \tag{16}$$

subject to (5), and

$$v_8^U(x_8, col) = \max_{c,a'} \Bigg\{ u\left(\frac{c}{1+\iota}, m_{p,\eta}\right) + \beta \underbrace{\mathbb{E}_{z'}\left[v_9(s, a', h_p, m_p, \eta', z')\right]}_{\text{Parent value}} + \nu\beta \underbrace{\iint v_2(col, 0, h, m, z)\, d\mu_m(m)\, d\mu_z(z)}_{\text{Child value}} \Bigg\} \tag{17}$$

subject to (15).

After the agent's child becomes independent, her individual problem is equal to (4). We leave the recursive problem at the retirement stage into Appendix E.1.

3.8 Stationary Equilibrium

A stationary recursive competitive equilibrium for this economy is a collection of: (i) private households' decision rules for consumption, asset holdings, and parental investment $\{c_j^{\delta_p}, a_j^{\delta_p\prime}, i_j^{\delta_p}\}$, and the associated value functions $\{v_j^{\delta_p}\}$ for both informed and uninformed parents $\delta_p \in \{I, U\}$; (ii) an admission shifter ζ set by the government in college admission policy function $\chi(\cdot, \zeta)$, and consumption tax τ_c; (iii) aggregate capital and labor inputs of the representative firm; (iv) prices $\{w^{col}, w^{ncol}, r\}$; and (v) a vector of measures of distributions such that:

1. Given prices, decision rules solve individual's problems (4), (10), (11), (12), and (A13);

2. Given prices, aggregate capital and labor inputs solve the firm's problem (3);
3. Capital market and labor markets are cleared;
4. The college capacity constraint is satisfied;
5. The government budget is balanced;
6. The distribution of households is stationary.

The equilibrium concept is standard and fully detailed in Appendix D.

4 The Sources of Inefficiency

Our quantitative model provides a rich environment in which to examine how educational competition and parental unawareness distort both the level of human capital investment and the composition of human capital, as well as how these two forces interact. To illustrate the main sources of inefficiency, clarify the scope for government intervention, and identify the deep parameters that govern the quantitative importance of these inefficiencies, we analyze a simple two-generation model in this section.

A standard source of inefficiency in our full quantitative model arises from imperfect capital and insurance markets, as in an Aiyagari-style environment, together with parents' inability to write enforceable contracts with their children. This friction helps rationalize why an increase in pre-college public investment, g, may improve welfare, and its implications have already been studied extensively in the existing literature (e.g., Lee and Seshadri, 2019; Daruich, 2022).

In addition to this conventional friction, our quantitative model features two novel sources of inefficiency. The first is an education-competition externality: parents do not internalize how their own educational investments affect the outcomes and incentives of other families. As a result, there is a wedge between private and social incentives to invest in children's education. The second is parental unawareness: uninformed parents fail to internalize the negative long-run effects of excessive tutoring on children's adulthood outcomes. This generates a wedge between the investment choices of informed and uninformed parents.

These two distortions also interact with each other. In particular, the overinvestment of uninformed parents intensifies the educational competition faced by informed parents, thereby amplifying the competition-driven externality. In the simple two-generation model considered here, we abstract from the first, conventional source of inefficiency and focus exclusively on these two new distortions.

4.1 Simple Model Setup

Consider a two-generation model in which a parent chooses educational investment for a child. The household is endowed with income y_p, the child's initial cognitive skill $h_0 > 0$, the child's initial mental health $m_0 > 0$, and the parent's mental health m_p. A share Π_U of parents are uninformed: they are unaware of their child's mental health state, how it evolves, and how mental health affects future labor income. The remaining share, $1 - \Pi_U$, are fully informed about their children.

The college admission probability is governed by $\chi(h, \zeta)$, which is a function of h and a shifter parameter ζ (capturing admission threshold) set by the government in equilibrium. We assume $\chi(\cdot)$ is increasing in h and decreasing in ζ. That is, $\chi_h(h, \zeta) > 0,\ \chi_\zeta(h, \zeta) < 0$.

4.1.1 Private Household's Problem

The household solves

$$\max_{c_1, i} \left\{ \log((1 + \kappa m_p)c_1) + \chi(h, \zeta) \log\big((1 + \kappa m)c_2^H\big) + \big(1 - \chi(h, \zeta)\big) \log\big((1 + \kappa m)c_2^L\big) \right\}$$

$$\text{subject to} \qquad c_1 + i = y_p,$$

$$\log h = \alpha \log h_0 + (1 - \alpha) \log i, \qquad \alpha \in (0, 1),$$

$$\log m = \rho_m \log m_0 + \phi_{m,c} \log\big(1 + \max\{i - \bar{i}, 0\}\big), \qquad \phi_{m,c} < 0,$$

$$c_2^H = w_{col} h^\lambda m^{\psi_{m,p}}, \qquad c_2^L = w_{ncol} h^\lambda m^{\psi_{m,p}}.$$

Informed parents. An informed parent understands both the effect of investment on cognitive skill and the adverse effect of excessive investment on child mental health. Using the budget constraint $c_1 = y - i$, the parent's problem can be written as

$$\max_{i \in [0, y_p]} \left\{ \log(y_p - i) + \lambda \log h + \psi_{m,p} \log m + \log(1 + \kappa m) + \chi(h, \zeta) \log\left(\frac{w_{col}}{w_{ncol}}\right) \right\}.$$

Taking the first order condition (FOC) with respect to i, we obtain:

If $i \leq \bar{i}$, since mental health is unaffected by investment below the threshold, the first-order condition is

$$\frac{1}{(1 - \alpha)(y_p - i)} = \frac{\lambda}{i} + \left(\frac{h_o}{i}\right)^\alpha \log\left(\frac{w_{col}}{w_{ncol}}\right) \frac{\partial \chi(h, \zeta)}{\partial h}. \tag{18}$$

If $i > \bar{i}$, note $\phi_{m,c} < 0$, we have

$$\underbrace{\frac{1}{(1-\alpha)(y_p - i)}}_{\text{MC of investing } i} + \underbrace{\frac{(-\phi_{m,c})}{(1-\alpha)(1+i-\bar{i})}\left(\psi_{m,p} + \frac{\kappa m}{1+\kappa m}\right)}_{\text{mental health cost of investing } (i-\bar{i})} = \underbrace{\frac{\lambda}{i}}_{\text{standard incentive}} + \underbrace{\log\left(\frac{w_{col}}{w_{ncol}}\right)\left(\frac{h_0}{i}\right)^{\alpha}\frac{\partial \chi(h,\zeta)}{\partial h}}_{\text{competition incentive}} \tag{19}$$

Uninformed parents. An uninformed parent does not internalize the effect of investment on the child's mental health. Taking m as exogenous to the investment decision, the parent's problem is

$$\max_{i \in [0, y_p]} \left\{ \log(y_p - i) + \lambda \log h + \chi(h, \zeta) \log\left(\frac{w_{col}}{w_{ncol}}\right) \right\}.$$

Regardless of $i > \bar{i}$ or not, the FOC is the same as (18). Thus, for $i > \bar{i}$, the wedge between informed and uninformed parents is exactly the *mental health cost of investing* $(i - \bar{i})$ term in Equation (19).

4.1.2 The Planner's Problem

Let households be indexed by type $x = (y_p, h_o, m_o, m_p)$, with distribution G, and let college capacity be Z. The planner chooses investment allocations $\{i(x)\}$ and the admission threshold ζ to maximize utilitarian welfare:

$$\max_{\{i(x)\},\zeta} \int \Big\{ \log\big[(1+\kappa m_p)(y_p - i(x))\big] + \chi(h(x),\zeta)\log\big((1+\kappa m(x))c_2^H(x)\big) + \big(1-\chi(h(x),\zeta)\big)\log\big((1+\kappa m(x))c_2^L(x)\big) \Big\} dG(x)$$

subject to the laws of motion for $h(x)$ and $m(x)$, and the college-capacity constraint $\int \chi(h(x),\zeta)\, dG(x) = Z$.

The planner's Lagrangian is

$$\mathcal{L} = \int \left[\log\big((1+\kappa m_p(x))(y(x) - i(x))\big) + \lambda \log h(x) + \psi_{m,p} \log m(x) + \log(1+\kappa m(x)) + \chi(h(x),\zeta)\log\left(\frac{w_{col}}{w_{ncol}}\right) \right]$$

The first-order condition with respect to $i(x)$ is

$$\frac{1}{y(x) - i(x)} = \lambda \frac{1-\alpha}{i(x)} + \psi_{m,p}\frac{\partial \log m(x)}{\partial i(x)} + \frac{\kappa}{1+\kappa m(x)}\frac{\partial m(x)}{\partial i(x)} + \left[\log\left(\frac{w_{col}}{w_{ncol}}\right) - \nu_l\right]\chi_h(h(x),\zeta)(1-\alpha)\frac{h(x)}{i(x)}.$$

For $i(x) > \bar{i}$, this becomes

$$\frac{1}{y(x)-i(x)}+\frac{-\phi_{m,c}}{1+i(x)-\bar{i}}\left(\psi_{m,p}+\frac{\kappa m(x)}{1+\kappa m(x)}\right)=\lambda\frac{1-\alpha}{i(x)}+\left[\log\left(\frac{w_{col}}{w_{ncol}}\right)-\nu_l\right]\chi_h(h(x),\zeta)(1-\alpha)\frac{h(x)}{i(x)}.$$

The first-order condition with respect to ζ is

$$\int \chi_\zeta(h(x),\zeta)\left[\log\left(\frac{w_{col}}{w_{ncol}}\right)-\nu_l\right]\,dG(x)=0.$$

This gives: for $i(x) > \bar{i}$,

$$\frac{1}{(1-\alpha)(y(x)-i(x))}+\frac{-\phi_{m,c}}{(1-\alpha)(1+i(x)-\bar{i})}\left(\psi_{m,p}+\frac{\kappa m(x)}{1+\kappa m(x)}\right)=\frac{\lambda}{i(x)}. \tag{20}$$

4.2 Insights from the Simple Model

Insight 1 The simple model highlights two distinct wedges. First, private parents internalize the effect of investment on their own child's admission probability, but not the induced tightening of the admission threshold faced by others; this generates a competition externality. Second, uninformed parents fail to internalize the mental-health cost of tutoring investment. As a result, the private investment choice of uninformed parents exceeds that of informed parents, which in turn exceeds the socially optimal level:

$$i^U > i^I > i^{SP},$$

at least generically for interior solutions with $i > \bar{i}$.

Insight 2 Parental misinformation does not only distort the choices of uninformed households themselves. Because college admission is rank-based and aggregate capacity is fixed, uninformed parents also affect the incentives faced by informed parents through equilibrium spillovers.

Decomposing the effect of parental misinformation on equilibrium investment First, there is a *direct effect*: uninformed parents choose too much investment because they fail to internalize the mental-health cost of tutoring. Second, there is an *indirect equilibrium effect*: because uninformed parents increase the level of aggregate educational competition, informed parents also increase investment in response to the tighter admission environment. Let $\bar{i}^I(\Pi_U)$ and $\bar{i}^U(\Pi_U)$ denote the average investment of informed and uninformed households in equilibrium. Then the total effect of a change in Π_U on average

investment can be decomposed as

$$\frac{d}{d\Pi_U}\left[(1-\Pi_U)\bar{i}^I(\Pi_U)+\Pi_U\bar{i}^U(\Pi_U)\right]=\underbrace{\left(\bar{i}^U(\Pi_U)-\bar{i}^I(\Pi_U)\right)}_{\text{composition/direct effect}}+\underbrace{(1-\Pi_U)\frac{d\bar{i}^I(\Pi_U)}{d\Pi_U}+\Pi_U\frac{d\bar{i}^U(\Pi_U)}{d\Pi_U}}_{\text{equilibrium spillover effect}}.$$

The first term captures the fact that replacing an informed household with an uninformed one raises investment mechanically. The second term captures the endogenous change in behavior by all households as the competitive environment becomes more intense. A higher share of uninformed parents, Π_U, affects average investment through both a composition effect and an equilibrium spillover effect. The composition effect is positive whenever uninformed parents invest more than informed parents. The spillover effect captures how a larger mass of uninformed parents changes the equilibrium intensity of educational competition. Under strategic complementarity in parental investments, an increase in Π_U raises the competitive pressure faced by both informed and uninformed parents, implying

$$\frac{d\bar{i}^I(\Pi_U)}{d\Pi_U}\geq 0 \quad \text{and} \quad \frac{d\bar{i}^U(\Pi_U)}{d\Pi_U}\geq 0.$$

In this case, a larger share of uninformed parents increases average investment not only because more parents are of the high-investment type, but also because equilibrium responses induce both types to invest more.

5 Mapping the Model to the Data

In this section, we outline the process of mapping the model to the data. Our baseline calibration focuses on the life cycle of one adult generation and one child generation.[15] We estimate our life-cycle model using a two-step identification strategy. First, we directly estimate or calibrate parameters that are identifiable outside the model. Second, we estimate the remaining parameters using the method of simulated moments (MSM), matching model outputs to their empirical counterparts constructed from our primary sample of children aged 6-21 and their parents observed between 2010 and 2018.

[15] While the standard approach in macro is to calibrate the model to a pre-college-expansion-reform steady state, we employ our current approach because: (1) the main data source we use for calibration spans from 2010 to 2018, capturing the post-reform scenario; (2) In this data, we observe a significant disparity in the college labor share between the parent and child generations. Thus, assuming a stationary distribution is not feasible, and calibrating the model to pre-reform China is not possible.

5.1 Externally Calibrated Parameters

In the first step, we directly estimate four sets of parameters from the data, including: (1) pre-college human-capital-based college admission probability function; (2) the effects of mental health capital on labor efficiency;(3) age profile of earnings by education levels; (4) parent cognitive skill distributions by education levels; and (5) public investment by the stage of pre-college education. We report the results on the externally estimated parameters in Table 2 and Figure 10.

Admission Function We estimate the function using the adult survey of the CFPS, where we can observe both the measure of pre-college human capital (i.e., cognitive ability scores) in their early career and education outcomes. We restrict the age range to 24-35, which is close enough to the college entrance examination so that cognitive ability scores primarily reflect pre-college human capital.[16] Individuals in our data sample are sorted into ten bins based on their decile rank of cognitive ability scores. For each bin, we observe the corresponding college admission probability, measured by the share of people with a college degree.[17]

To map the empirical patterns of the ability-based admission probability into the model, we first compute the logarithmic distance between the mean cognitive ability score of each bin and the overall mean ($\bar{h}$) in our data sample. Together with the admission probability in each bin, we build a correspondence between the relative standing of child ability and college likelihood. These empirical patterns are represented by red square dots in Figure 10. Next, in our model, we can easily calculate the log distance between human capital grids and the mean human capital (normalized to one in parent generation) and assign each grid an admission probability through linear interpolation. This yields a pre-college human capital-based admission function, represented by the solid blue curve in Figure 10.[18] Finally, to ensure the model-generated college capacity matches the targeted empirical counterpart Z in our benchmark model, we adjust the

[16] Ideally, with a long enough panel dataset with sufficiently large number of observations, we could observe individuals' childhood development stage, adult education, and labor market outcomes, and we would then use their cognitive ability scores at age 18 and track their education outcomes to estimate the admission function. However, for the cohort of children aged 6-21 in 2010-2018 in our CFPS sample, there are not sufficient number of observations for us to observe their college education outcome to enable us to estimate the admission function. Therefore, we also need to rely on older cohorts in the CFPS, where we can observe cognitive ability in early adulthood and their education outcomes. To mitigate potential issues, we select adults aged 24-35, who are not far from college graduation, thus minimizing changes in pre-college human capital (measured by cognitive ability scores) after entering the labor market.

[17] In other words, we do not differentiate between college admission and college completion, as the college completion rate in China is very high, often above 95% for each cohort.

[18] Note that $\chi(h)$ is a map between human capital grid h and college admission probability.

shifter parameter ζ in $\chi(\cdot, \zeta)$. This adjustment process is illustrated by the blue dashed curve in Figure 10.[19]

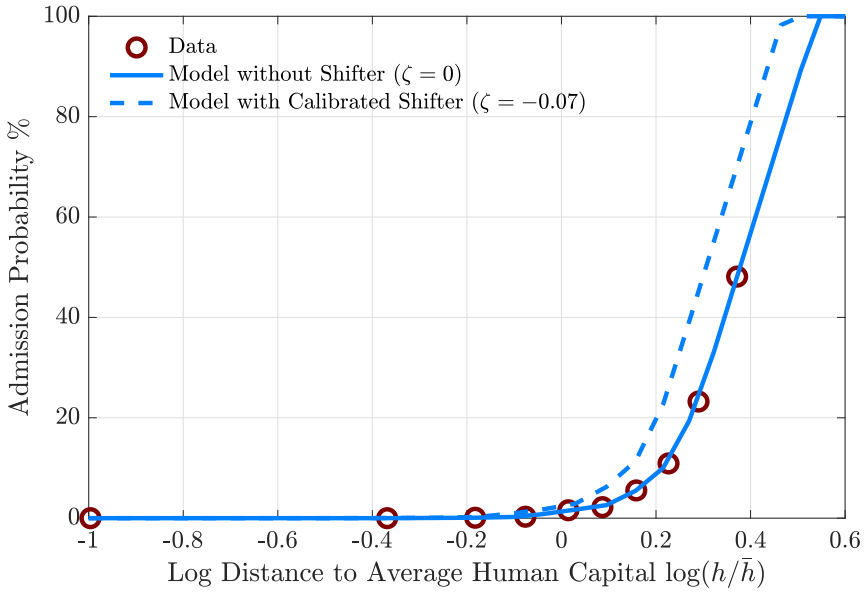


Figure 10: Pre-college Human-capital-based College Admission Policy

As shown in Figure 10(A), the admission probability is approximately an increasing convex function of a child's pre-college human capital (normalized by the average) at the time of the college entrance exam within the domain $[-1, 0.5]$.[20] A sensitive region emerges when a child's pre-college human capital is roughly 20% higher than the average, where a small increase in human capital results in a substantial rise in admission probability. We only care about the shape of $\chi(\cdot, \zeta)$, which dictates admission probabilities as a function of pre-college human capital rank, since we can always adjust the shifter ζ to match the observed college capacity. Furthermore, the data (see Figure A12 in Appendix B.13 for more details) suggests that the convexity of $\chi(\cdot, \zeta)$ remains remarkably stable (nearly parallel) as college capacity expands.

Consumption Equivalence Since we assume that the child shares household consumption c, the implied consumption share for each parent is $\frac{c}{1+\iota}$. We set $\iota = 0.3$ to be consistent with the OECD equivalence scale, which assigns a weight of 0.3 to a child.

[19] In other words, we adjust ζ to ensure the college capacity constraint (**??**) holds. More details on the whole procedure are provided in Appendix B.13.

[20] Note that there are very few individuals in the data whose h is larger than $exp(0.5)\bar{h}$.

Table 2: Externally Estimated Parameters

Parameter		Value	Source/Target
(A) Preference			
Consumption equivalence	ι	0.30	OECD equivalence scale for kids
(B) Production			
Capital share	Ω	0.48	Bai, Hsieh, and Qian (2006)
Depreciation rate	δ	0.11	
Iceberg cost	ι_o	0.05	Chen et al. (2019)
Elasticity of substitution (col. & non-col.)	$\frac{1}{1-\psi}$	2.08	Estimation using UHS 92-09
(C) Government policies			
Public education investment $j = 5$	g_5	0.17	
Public education investment $j = 6$	g_6	0.24	CEFS Yearbook
Public education investment $j = 7$	g_7	0.31	
College tuition and fees	κ	0.16	
College subsidy	$\zeta_{college}$	0.45	
Fixed college capacity	ξ	0.35	College labor share = 35% in Census 2020
(D) Parent human capital and age profile			
Mean: Parent human capital dist. (col.)	M^{col}	59.11	
Mean: Parent human capital dist. (non-col.)	M^{ncol}	5.90	CFPS 2010-18
SD: Parent human capital dist. (col.)	Q^{col}	0.02	
SD: Parent human capital dist. (non-col.)	Q^{ncol}	0.19	
Age profile of earnings by education	A_j^s	Figure A9	CHNS

Return to Mental Health Capital The goal is to back out $\phi_{m,p}$ in Equation (2). Mental health capital is approximated by the CES-D scores in the CFPS. We use labor income data from the CFPS to estimate age polynomials for different education groups. We then use these age profiles to filter out age effects from the log labor income and log mental health capital observations. We estimate an OLS regression of log individual labor income on time dummies, log cognitive ability test scores, and the mental health capital measure for adults to recover $\phi_{m,p}$. We find $\phi_{m,p} = 0.068$, with a significance level of less than 1%. In a log–log specification, this implies that a 1 percent increase in mental health capital is associated with a 0.068 percent increase in adult labor income. We also find that worse mental health is associated with a lower likelihood of employment and, conditional on employment, reduced labor supply. For hourly wages, the effects are no longer significant once we include individual and time fixed effects. Note that the dependent variable in our main specification—log individual labor income—combines both the extensive and intensive margins of labor supply. Therefore, significant effects of mental health on labor supply can generate impacts on total earnings even in the absence of significant effects

on hourly productivity. Using models with individual fixed effects, we confirm that the employment relationship holds even when identification relies on within-individual variation—i.e., adverse mental health "shocks" have immediate consequences for labor supply. These results are consistent with findings in Cronin, Forsstrom, and Papageorge (2005), who use the U.S. Medical Expenditure Panel Survey. They are also consistent with Abramson, Boerma, and Tsyvinski (2025), who argue that a key channel through which mental illness affects the macroeconomy is via employment and labor supply, including mechanisms such as rumination. Finally, Abramson, Boerma, and Tsyvinski (2025) report that, in the PSID, after controlling for individual fixed effects, the effect of mental health on hourly wages (labor productivity) is very small and statistically insignificant—consistent with our findings.

For comparison, the elasticity with respect to cognitive skills is about 0.50: a 1 percent increase in cognitive skills is associated with a 0.5 percent increase in adult labor income.

Firm Production Technology We set the capital share parameter $\Omega = 0.48$ and depreciation rate $\delta_k = 0.11$ following Bai, Hsieh, and Qian (2006). The financial intermediation cost $\iota = 0.05$ is set so that the gross rate of return to capital is 11% in the early 2010s (in line with the estimates of Chen et al. (2019)). As detailed in Gu and Zhang (2025), ψ is estimated to be 0.52, suggesting the elasticity of substitution between college-educated and non-college-educated labor $\frac{1}{1-\psi}$ to be 2.08.

Public Investments Pre-college public investment is roughly twice the private investment. The public education expenditures are 17% to 31% of the average household labor income (normalized to be one in the model) by school stage, informing the value of g_j parameter. College tuition and fees roughly equal 16% of the average household labor income, pinning down the tuition parameter κ. The government subsidy per college student $\varsigma_{college}$ is nearly three times higher than tuition paid by households.[21]

5.2 Internally Calibrated Parameters

Table 3 summarizes the internally calibrated parameters alongside the empirical moments used to discipline them. These moments include estimated parameters from misspecified translog child skill production functions, age profiles of child cognitive ability and parental monetary investments, and the income elasticity of investment. The 22 model parameters are estimated by minimizing the distance between empirical moments and

[21] See Table A12 in Appendix A for more details.

their model counterparts. While all parameters jointly influence each targeted moment, we discuss them in relation to the parameters that provide the most identification power.

Indirect Inference. In an ideal empirical setting with both private and public investments and ability at an annual frequency available, the skill production function can be directly estimated. However, since the CFPS provides skill and private investment measures only every four years without any information on public investments, Following Abbott (2022), we use an indirect approach. The main identification comes from matching pooled OLS regression estimates of two misspecified translog child cognitive skill production function parameters with those estimated in the data. The first auxiliary model is

$$\log(i_j) = \omega_0 + \omega_1 \log(h_j) + \omega_2 \log(y_j), \tag{21}$$

which describes how parental monetary investment depends on the current period's child cognitive skills and parental income. The second model is

$$\log(h_{j+1}) = \beta_0 + \beta_1 \log(i_j) + \beta_2 \log(h_j) + \beta_3 \log(i_j)\log(h_j), \tag{22}$$

which captures how parental monetary investment, the current stock of child cognitive skills, and their interaction influence future child skills. For both regressions specified in Equation (21) and (22) in the data, we control for year and age effects. In the data, we observe parental monetary investment i_j for children aged 6-9 (model period $j = 5$), 10-13 (model period $j = 6$), and 14-17 (model period $j = 7$), and the child cognitive ability scores h_{j+1} for $j = 5, 6, 7$. We perceive child cognitive skills at ages 18-21 (model period $j = 8$) as the final outcome of pre-college human capital.

Preferences We set the discount factor $\beta = 0.92$ to target a 4% annual real interest rate in equilibrium, given the iceberg cost ι_o. The altruism parameter ν is jointly calibrated with other parameters in the child skill production function, as discussed below.

Child Cognitive Skill Production Function The child skill production technology in Equation (8) draws upon particularly Cunha, Heckman, and Schennach (2010). While existing studies primarily focus on the differential skill productivity and complementarity between early (ages 0-5) and later childhood (ages 6-14+), our model exclusively focuses on children aged 6-18—the critical period when they start to attend schools and receive intensive pre-college investments aimed at improving exam performance. Unlike prior work emphasizing early childhood development, we use this multi-stage framework to analyze how families with varying resources allocate investments to enhance their

children's relative ability position within their cohort during this educationally pivotal phase when resource disparities strongly manifest through unequal access to enrichment opportunities (e.g., cram schools), contributing to intergenerational persistence.

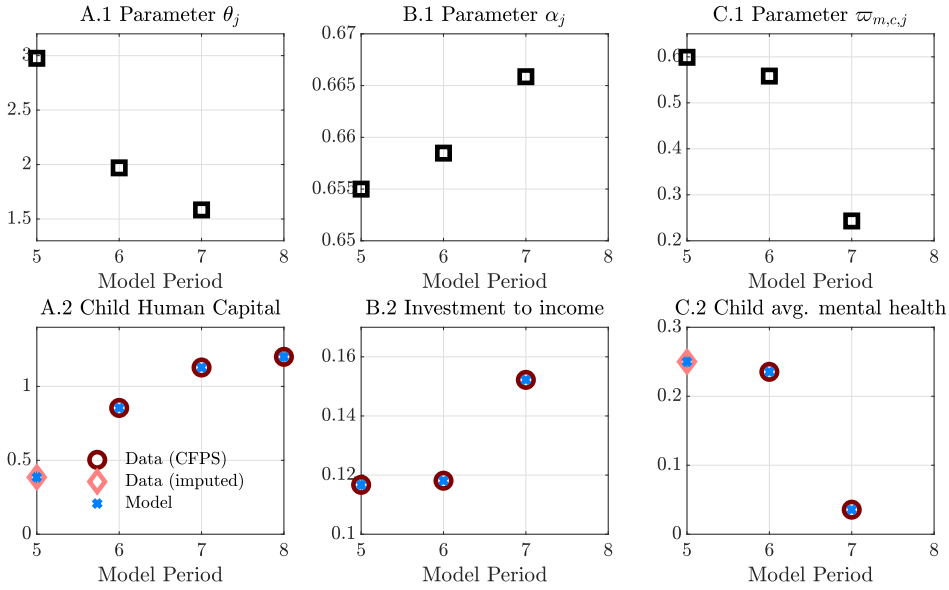


Figure 11: Age-dependent Parameters and Targeted Moments

Note: Panel (A.1) displays the estimates of θ by age j in the model. Panel (A.2) demonstrates how well the model moments (blue cross) on children's pre-college human capital by age j in the model, primarily used to discipline θ_j, align with their empirical counterparts (red circle). Panel (B.1) presents the estimates of α by age j. Panel (B.2) shows how well the ratio of parental investment to income by age j in the model (blue cross), primarily used to discipline α_j, match their empirical counterparts (red circle). Panel (C.1) presents the estimates of $\varpi_{m,c}$ by age j. Panel (C.2) shows how well the standardized mental health scores by age j in the model (blue cross), used to discipline $\varpi_{m,c}$, match their empirical counterparts (red circle).

We choose the age-dependent efficiency of child skill formation technology θ_j to match the age profile of child cognitive skills. Since the CFPS data only provide cognitive ability test scores for children aged 10 and older, we extrapolate the age profile to infer average cognitive skills for ages 6 to 9, as discussed in Section 2.1.2, corresponding to the beginning of the model period 5.[22] We report the estimated results for θ_j in Table 3, where we show that total factor productivity decreases from age 6 to age 17. This result aligns with the estimation in Agostinelli and Wiswall (2024). We jointly calibrate the altruism parameter ν, age-dependent self-productivity α_j for $j = 5, 6, 7$, and the weight on parental investment η to target three sets of empirical moments: (1) The age profile of the private monetary investment-to-household income ratio; (2) The standard deviation

[22]We have selected the initial stage to be ages 6-9, despite lack of cognitive ability score data in this range, because (1) we can still observe parental investments in children aged 6-9; and (2) The education competition intensity typically begins to increase in elementary school in China, starting at age 6.

of logarithmic child cognitive skills at period 8 (age 18-21); and (3) β_1 in Equation (22), which measures the effect of parental monetary investment on next-period child skill. Further identification details are provided in Gu and Zhang (2025). We particularly target the effect of parental investment on child human capital growth because it determines (i) how college expansion quantitatively impacts *average* child human capital by shaping parental investment incentives, and (ii) how college expansion quantitatively affects inequality and intergenerational mobility through *heterogeneous* parental investment responses. We report the estimated α_j in Table 3, which indicates that self-productivity rises from age 6 to 17, though the rate of increase slows with age. This aligns with the estimates in Krueger, Ludwig, and Popova (2024).

The dynamic complementarity parameter γ is set to match the regression coefficient β_3 in Equation (22) following Abbott (2022). In cases where parental investment exhibits a stronger complementarity with the current-period child skill, the interaction term would exert a more negative impact. The elasticity of substitution between private and public investment, μ, is set to match the regression coefficient ω_2 in Equation (21). Since public education investment, g_j, does not depend on household characteristics, a higher degree of complementarity leads to a smaller impact of household income, y, on private education expenditures, i. We obtain $\mu = 0.24$, indicating that governmental and parental monetary investments are substitutes, albeit far from perfect ones.[23]

Return to Cognitive Skills We identify λ following Gu and Zhang (2025)'s strategy, which exploits a unique empirical pattern in Chinese data: parental monetary investments spike sharply as children's score ranks approach college admission thresholds.[24] The economic intuition is revealing: when $\lambda < 1$ and takes on a small value—indicating a low return to skills and diminishing returns as pre-college human capital increases—the observed spike in investment near admission thresholds suggests that parents are primarily motivated by securing college admission rather than enhancing their children's future labor productivity. Thus, λ can be identified by examining how parental monetary investments vary with children's test score ranks after controlling for parental income. Our implementation uses the regression coefficient ω_1 in Equation (21) to discipline λ. Simultaneously, we adjust the shifter parameter ζ in the admission function to align with the 35% share of college-educated labor in the child generation, as well as the college labor intensity parameter ϕ in the representative firm's production technology (Equation

[23]This value suggests governmental and parental monetary investments are more than Cobb-Douglas. This finding is qualitatively consistent with the 0.59 estimation from Kotera and Seshadri (2017), which uses U.S. data and variations in public investment across school districts.

[24]See more details in Section 3.3.2 and Figure 4 in Gu and Zhang (2025).

(3)) to match the college labor income premium of 1.72 found in our CFPS sample.[25]

Intergenerational Transmission of Skills Since child cognitive skills are only observed from age 10 onward, we use the intergenerational correlation of cognitive skills at ages 10-13 (model period $j = 6$) to identify the persistence parameter ρ_h in Equation (6). This persistence captures the influence of parental human capital on child human capital endowment. The average and standard deviation of human capital endowments are both dependent on parental education. We discipline the four parameters $\{\mu_h^{col}, \mu_h^{ncol}, \sigma_h^{col}, \sigma_h^{ncol}\}$ to target their counterparts in the CFPS data.

For intergenerational transmission of mental health capital, we use a similar method as above to discipline the parameters. We use the intergenerational correlation of mental health at ages 10-13 (model period $j = 6$) to identify the persistence parameter ρ_m in Equation (7). This persistence captures the influence of parental mental health on child mental health capital endowment. The average and standard deviation of mental health capital endowments are both dependent on parental education. We discipline the four parameters $\{\mu_m^{col}, \mu_m^{ncol}, \sigma_m^{col}, \sigma_m^{ncol}\}$ to target their counterparts in the CFPS data.

Shocks to Mental Health Capital We assume that shocks to mental health capital—for both adults and children—follow a normal distribution. Specifically, $\epsilon'_{m,p} \sim N(0, \sigma_{m,p})$ and $\epsilon'_{m,c} \sim N(0, \sigma_{m,c})$. This assumption is based on data showing that the change in mental health scores over a four-year period is approximately normally distributed, as shown in Figure A1 in Appendix B.3. We discipline $\sigma_{m,p}$ and $\sigma_{m,c}$ to target the standard deviations of the mental health scores for both adults and children respectively.

The Evolution Process of Child Mental Health Capital In Equation (9), $\varsigma_{m,c}$ measures how child mental health capital depends on parental mental health capital in each period, which is disciplined to match the intergenerational correlation between mental health of children in age $j = 6$, i.e., age 14-17, and that of their parents. $\rho_{m,c}$ denotes the persistence of mental health capital across periods and is disciplined to match the four-year persistence of 0.420 observed in the CFPS (i.e., the dependence of current-period child mental health on its previous-period level for the targeted cohort).

Note that the component $(\varpi_{m,c,j} + \vartheta_{m,c}\log(w^s y(h_p, m_p, z)))$ in Equation (9) captures

[25]The CFPS does not provide measures on the intensive margin of labor supply, such as hours worked per year, and more importantly, our model abstracts from the intensive margin of labor supply, so we use college labor income premium, defined as the ratio of average labor income ($w^s y$, where y is specified in (2)) of college workers to that of non-college workers, instead of college wage premium ($w^{\text{col}}/w^{\text{ncol}}$) as our empirical target.

how family background affects child mental health through channels other than tutoring expenditures. Correspondingly, ($\varpi_{m,c,j}$ is disciplined to match the age profile of child mental health capital from $j = 5$ to $j = 7$ (ages 10–13 to 18–21), and $\vartheta_{m,c}$ is disciplined to match the correlation between parental income $w^s y(h_p, m_p, z)$ and child mental health m in the data.

Identification of $\phi_{m,c}$ In the model, $\phi_{m,c}$ governs the tutoring channel in the child mental-health formation technology (Equation (9)). The term $\log(1 + \max\{i - \bar{i}, 0\})$ captures cram-school (tutoring) spending beyond basic in-school necessities, and $\phi_{m,c}$ measures its marginal effect on child mental health capital. The model implies $\phi_{m,c} < 0$ if tutoring harms mental health. Because we do not have a sufficiently strong instrument to causally identify the impact of cram-school participation on mental health, our identification strategy combines internal calibration with external validation.

For internal identification, we rely on the empirical moment "Difference in child mental health between college- and non-college-educated parents for children aged 18–21" to discipline $\phi_{m,c}$. The key logic is that $\phi_{m,c}$ is pinned down by the life-cycle pattern of child mental health by parental education/SES. Before schooling, tutoring is minimal, so higher-SES children exhibit better mental health due to supportive resources (e.g., home environment and enrichment). During the school years, tutoring intensity rises and is disproportionately higher among high-SES families; the resulting deterioration in mental health among heavily tutored children causes the mental-health gap between high- and low-SES groups to narrow and potentially flip. The timing and magnitude of this gap shrinkage in the mental-health age profile discipline the magnitude of $\phi_{m,c}$, as shown in Figure 12.

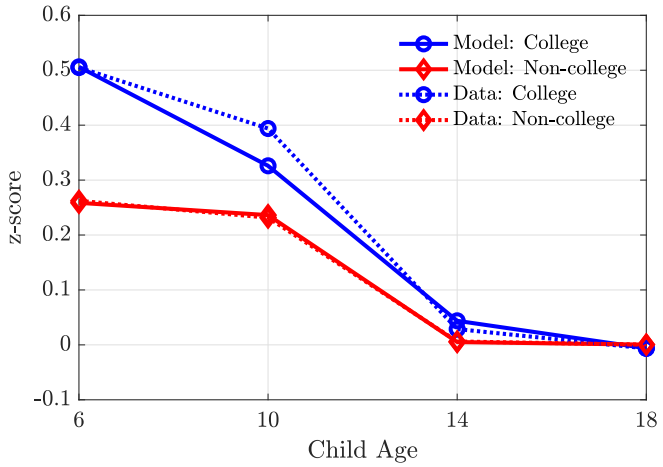


Figure 12: Identify $\phi_{m,c}$: using Child Mental Health Age Profile by Parent Education

Table 3: Internally Estimated Parameters

Description	Parameter	Value	Targeted Moment	Data	Model
(A) Preference					
Discount factor	β	0.91	Real interest rate	0.04	0.04
Altruism	ν	0.46	Regression coefficient β_1	0.09	0.09
(B) Child cognitive skill formation					
Age-dependent efficiency of skill production, $j = 5, 6, 7$	θ_j	Figure 11(A.1)	Child cognitive skill age profile	Figure 11(A.2)	
Age-dependent weight on current skill, $j = 5, 6, 7$	α_j	Figure 11(B.1)	Parental investment to income age profile	Figure 11(B.2)	
Weight on parental investment	η	0.32	SD log child cog. skill (age 18-21)	0.26	0.26
Dynamic complementarity	γ	0.52	Regression coefficient β_3	-0.02	-0.02
Elasticity of substitution between i and g	μ	0.34	Regression coefficient ω_2	0.58	0.58
(C) Child mental health formation					
Weight on parental mental health	$\varsigma_{m,c}$	0.24	IG correlation of mental health (age 10-13)	0.18	0.18
Persistence of child mental health	$\rho_{m,c}$	0.50	4-year persistence of child mental health	0.42	0.42
Age-dependent mental health shifter, $j = 5, 6, 7$	$\varpi_{m,c,j}$	Figure 11(C.1)	Age profile of child mental health profile	Figure 11(C.2)	
Weight on parental income	$\vartheta_{m,c}$	0.16	Corr. child mental health and parental income (age 10-13)	0.013	0.013
Impact of tutoring on mental health	$\phi_{m,c}$	-3.50	Cram school participants have mental health capital 0.33 SD lower than non-participants		
SD of Child mental health shocks	$\sigma_{m,c}$	0.49	SD: child mental health IRT-based scores	0.80	0.80
(D) Intergenerational skill transmission					
Persistence of IG skill	ρ_h	0.08	IG correlation of cog. skill (age 10-13)	0.29	0.29
Avg. child skill endowment (college parent)	μ^{col}	0.32	Avg. rel. child cog. skill by parent edu (age 10-13)	1.13	1.13
Avg. child skill endowment (non-college parent)	μ^{ncol}	0.33	Avg. child cog. skill (age 6-9)	0.39	0.39
SD IG genetic shocks (college parent)	σ_h^{col}	0.10	SD log child cog. skill (college parent, age 10-13)	0.27	0.27
SD IG genetic shocks (non-college parent)	σ_h^{ncol}	0.12	SD log child cog. skill (non-col. parent, age 10-13)	0.36	0.36
(E) Labor productivity Shocks					
Persistence of labor productivity shocks	ρ_z	0.72	Estimated persistence of wage income	0.79	0.79
SD of labor productivity shocks	σ_z	0.23	Household wage income Gini coefficient	0.38	0.38
Aggregate productivity (firm)	A	0.93	Average household wage income	Normalization	1.00
(F) Education and returns					
Return to cognitive skills	λ	0.50	Regression coefficient ω_1	1.14	1.14
College labor intensity (firm production)	ϕ	0.28	College labor income premium	1.72	1.72
College admission function shifter	ζ	-0.01	College-educated labor share of child generation	0.27	0.27

In the second step, we validate the calibrated $\phi_{m,c}$ using reduced-form evidence that cram-school participation lowers adolescent mental health by approximately 0.05–0.37 standard deviations (Section B.8); the model implies an effect of about 0.15 standard deviations, consistent with the empirical range.

Transition into adulthood mental health dynamics. To discipline the adult mental-health process, we need to quantify how much of adults' mental health is driven by heterogeneity formed earlier in life (i.e., prior to labor-market entry) versus shocks realized during adulthood. Using CFPS panel data, we estimate through the following regressions:

$$\log(m_{p,i,j}) = X'_{i,j}\beta + R_{i,j}, \qquad R_{i,j} = \alpha_i + z_{i,j} + u_{i,j},$$

where $X_{i,j}$ includes age polynomials and year and gender dummies. The individual effect $\alpha_i \sim N(0, \sigma_\alpha^2)$ captures permanent, ex ante heterogeneity in initial adult mental health (interpreted as mental health accumulated in earlier stages, including childhood). The component $z_{i,j}$ follows a persistent AR(1) process, $z_{i,j} = \rho_z z_{i,j-1} + \varepsilon_{i,j}$ with $z_{i,24} = 0$ and $\varepsilon_{i,j} \sim N(0, \sigma_\varepsilon^2)$, while $u_{i,j} \sim N(0, \sigma_u^2)$ is an i.i.d. term capturing both measurement error and acute mental-health events.

The estimates imply a highly persistent shock component ($\rho_z = 0.965$) and substantial dispersion in permanent heterogeneity ($\sigma_\alpha^2 = 0.236$), with smaller innovations and transitory noise ($\sigma_\varepsilon^2 = 0.059$, $\sigma_u^2 = 0.036$). An important implication is that, conditional on observables in $X_{i,j}$, ex ante heterogeneity accounts for about 71% of the cross-sectional variation in mental health at age 25, with shocks accounting for the remaining 29%. We use this decomposition as a key moment to calibrate the adult mental-health process in the quantitative model, ensuring that the model matches the relative importance of (i) mental health differences formed before adulthood versus (ii) adult shocks in generating observed mental-health dispersion.

Labor Productivity Shocks We jointly calibrate ρ_z and σ_z to match the persistence of the labor income process and the Gini index of urban household labor income in China.[26] Note that we normalize the average household income to be one by adjusting the aggregate productivity parameter A in firm technology (Equation (3)).

[26]The estimates on the persistence of the labor income process are drawn from Yu and Zhu (2013), which uses longitudinal data from the CHNS for urban China. The CHNS holds an advantage over the CFPS because of its longer panel feature. The 0.38 Gini index of urban household earnings is documented by Ding and He (2018) using the UHS data of the late 2000s.

5.3 Model Performance and Validation

We validate our model estimation through three main exercises.

6 How Educational Competition Shapes Allocation of Human Capital

Benchmark Economy vs. "Parental Unawareness" Economy Motivated by our empirical findings in Section 2.2, we consider an "parental unawareness" economy where parents are unaware of their child mental state as well as its evolution over childhood and its impact on adult outcomes.

We compare key outcomes in the benchmark and "parental unawareness" economies Table 4. Unsurprisingly, when parents disregard their child's mental health, they invest more aggressively to enhance their child's academic performance. Total educational expenditures rise by approximately 14%, driven by a 30% increase in tutoring expenditures and a rise in the cram school participation rate from 51.2% to 54.9%. This reallocation of resources leads to a 0.5% decline in average parental consumption. As a direct result of this intensified investment, children's pre-college human capital improve by 1.4% by age 18. However, this comes at a significant cost to their well-being, as their mental health deteriorates by an amount equivalent to 15% of a standard deviation in the mental health distribution. The net effect on a child's expected lifetime income is a balance between these two opposing forces. Given that our estimated return to cognitive skills (0.50) is substantially larger than the return to mental health (0.078), the net impact is a slight increase in expected lifetime income of 0.6%.

Shutting Off Competition We eliminate competition by making college admission a purely random draw, independent of any individual heterogeneity, including pre-college human capital, while fixing the college capacity at 27% as in our benchmark model. In this way, parental investment in child pre-college human capital is no longer driven by competition for college admissions, as a parent's efforts have no influence on her child's chance of being admitted; every child has an equal 27% chance of admission. A comparison between columns (i) and (ii) in Table 5 reveals the consequences of shutting off competition on parental welfare and child development outcomes. As shown in Panels (A) and (B), with reduced incentives for competition, average parental investment falls by 61.0%, which is largely driven by drastic decline in cram school participation and tutoring expenditures, resulting in a 0.9% consumption-equivalent welfare gain for parents. At the

Table 4: Benchmark Economy vs. "Ignore Mental Health" Economy

	Benchmark	Parental Unawareness	Change
(i) Aggregates			
Parent Investment	0.119	0.135	+14.0%
– Tutoring Expenditures	0.057	0.074	+30.0%
Cram School Participation Rate	51.2%	54.9%	+3.7 p.p.
Parent Consumption	0.945	0.939	-0.5%
Child pre-college human capital	1.200	1.217	+1.4%
Child mental health capital	0.036	-0.086	-0.15 sd
Child expected lifetime income	1.201	1.208	+0.6%
(ii) Inequality			
Var log lifetime earnings	0.132	0.133	+0.8%
(iii) Intergenerational Persistence			
Pre-college human capital	0.381	0.396	+4.0%
Education	0.163	0.152	+1.2%
Mental health	0.168	0.165	-9.5%

same time, reduced parental investment also lowers children's age-18 pre-college human capital by 10.3%, but improves child mental well-being by around 50% of a standard deviation of the child mental health distribution. Their expected lifetime income overall decline by 2.2% on average.

Human Capital Reallocation Next, we consider how to allocate investment such that pre-college human capital and mental health capital are allocated more efficiently. We show that admitting students based on college lottery draw scheme while reallocating the private investment previously used for admissions competition to public investment to maintain the same aggregate effective investment can induce better allocation of human capital. Column (iv) of Table 5 present the results of this reallocation. Here, effective investment I_j received by children at age j—a CES aggregation of private and public investments as specified in Equation (8)—is preserved by offsetting the large reduction in private investment (61.0%) with increased public investment g_j for all periods j.[27] This reallocation induces improved outcomes: it maintains the improvement of child mental health capital, while child pre-college human capital remains nearly unchanged from the benchmark level. Consequently, consumption-equivalent welfare and child expected lifetime income increase.

[27]The government budget constraint remains balanced through increased consumption tax τ_c.

Table 5: Shut off Competition + Human Capital Reallocation

	Benchmark	Shut off Competition Δ	+ Reallocation Δ
Parent Investment	0.119	-60.5%	-62.2%
– Tutoring Expenditures	0.057	-98.2%	-98.24%
Cram School Participation Rate	51.2%	-39.9 p.p.	-41.4 p.p.
Public Investment	0.242	+0.0%	+27.7%
Parent Consumption	0.945	+2.1%	+0.7%
Parent Welfare (CEV)	-	+1.0%	+0.9%
Child cognitive skills h	1.000	-9.0%	-0.3%
Child mental health m	1.000	+35.1%	+35.2%
Child human capital $h^{\lambda}m^{\psi_{m,p}}$	1.000	-2.1%	+2.5%
Child expected lifetime income	1.201	-5.1%	-0.6%

7 Welfare Cost of Educational Competition and Policy Implications

This section quantitatively evaluates the aggregate, welfare, and distributional consequences of regulating parental investment competition for college admissions, with a particular focus on the allocation of human capital. We use fiscal policy instruments to induce a more efficient long-run allocation than the decentralized equilibrium, not only to derive an optimal policy, but also to estimate the welfare costs of educational competition. We show that endogenous mental health formation and the joint allocation of pre-college human and mental health capital are key to the welfare losses from competition and to the resulting policy implications.

To demonstrate the importance of modeling endogenous child mental health formation, we consider an alternative model that abstracts from it, in which adult labor efficiency units y (Equation 2) no longer depend on mental health capital. In this model, investing i above $\bar{i}$ entails no mental health cost. We label this economy "abstracting from mental health" and recalibrate it to match the same moments as the benchmark, particularly for parameters governing the relative strength of competitive incentives in parental investment, ensuring policy impacts are comparable.

7.1 Short-Run Impacts of the Private Tutoring Ban

In 2021, China banned for-profit tutoring for core subjects in grades K-9, aimed at improving exam performances, to alleviate financial pressures on families from intense

college entrance exam competition. We also adjust the admission shifter, ζ, and linear consumption tax, τ_c, to match the 27% college capacity and balance the government budget. To evaluate the short-run effects, we impose the private investment cap on investments made by *the current parent generation*.

To quantify the welfare consequences for the parent generation, we define welfare based on ex ante lifetime utility of economically active individuals (i.e., after their education level is determined and they have just entered the labor market). That is,

$$W_0 = (1-\Omega_{col,0}) \int\int v_1(ncol, a, h_p, m_p, z) d\mu_z d\Gamma_{h,ncol} \Gamma_{m,ncol} + \Omega_{col,0} \int\int v_2(col, a, h_p, m_p, z) d\mu_z d\Gamma_{h,col} \Gamma_{m,ncol} \tag{23}$$

where a is the initial asset holdings (equal to zero), $\Gamma_h^s, s \in \{col, ncol\}$ are initial distributions of parental pre-college human capital h_p by education level s, $\Gamma_m^s, s \in \{col, ncol\}$ are initial distributions of parental mental health capital m_p by education level s, μ_z is distribution of z, $\Omega_{col,0}$ is the population weight of college labor in the benchmark economy, and v_1 and v_2 are defined in Problem (4). The integrals are taken over the state space of h_p and possible types z.

Human Capital Allocation and Welfare With the investment cap $\bar{i}$ in place, as shown in Table 6, parental investment declines by approximately 44.7% on average. This reduction leads to a 5% decrease in children's pre-college human capital at college entry, but simultaneously results in an increase in mental health capital—equivalent to 38% of the standard deviation of the child mental health distribution. These two opposing effects on labor efficiency units (as shown in Equation (2)) ultimately result in a 2.3% reduction in children's expected lifetime income.

Reduced investment causes a 1.7% increase in parents' consumption. Welfare, measured by the consumption equivalent variation (CEV) in the ex ante lifetime utility of parents (as defined in (23)), improves due to the private tutoring ban, which leads to a better consumption-investment allocation. Specifically, as discussed in Section 4, parents do not internalize the impact of their investments on other parents' investments. Parents with high-ability children tend to prioritize investing in college admissions at the expense of their own consumption utility. Since these children typically already possess high pre-college human capital, the decreasing returns to scale in converting pre-college human capital h_p into labor efficiency units y (determined by Equation (**??**) with $\lambda = 0.50$) mean that the return on investment in terms of y is relatively small for them. The tutoring

Table 6: Short-Run Effects of Private Tutoring Ban

	(i)		(ii)		(iii)	
	Benchmark	Ban	Ignore Mental Health	Ban	Abstract from Mental Health	Ban
(A) Aggregates						
Parent Investment	0.119	-44.7%	0.135	-51.5%	0.116	-43.4%
– Tutoring Expenditures	0.057	-100.0%	0.074	-100.0%	0.053	-100.0%
Cram School Participation Rate	51.2%	0%	54.9%	0%	51.2%	0%
Parent Consumption	0.945	+1.7%	0.939	+2.2%	0.944	+1.6%
Parent Welfare (CEV)	-	+0.76%	-	+0.76%	-	+0.64%
Child pre-college human capital	1.200	-5.1%	1.217	-6.5%	1.200	-4.4%
Child mental health capital	0.036	+0.38 sd	-0.086	+0.54 sd	-	-
Child expected lifetime income	1.201	-2.3%	1.208	-2.8%	1.197	-16.5%
(B) Inequality						
Var log lifetime earnings	0.132	-6.7%	0.133	-7.4%	0.134	-5.7%
(C) Intergen. Persistence						
Pre-college human capital	0.381	-31.9%	0.396	-34.5%	0.366	-25.4%
Education	0.163	-84.0%	0.152	-84.1%	0.137	-72.8%
Mental health	0.168	+20.7%	0.165	+33.4%	-	-
Lifetime income					0.360	-26.3%

Note: Welfare means a CEV of welfare measure for parents defined in (23). Inequality refers to the variance of log-expected lifetime earnings for the child generation. Intergenerational persistence is measured in terms of rank-rank correlation coefficients.

ban reduces over-investment in these high-ability children, which not only improves the efficiency of investment in pre-college human capital but also eliminates the negative impact of parental monetary investment on child mental health capital. This induces a better allocation of these two types of human capital. Consequently, parental welfare increases by an average of 0.76%.

As shown in Column (ii) of Table 6, when parents do not value their children's mental well-being—leading to higher tutoring expenditures and greater cram school participation than in the benchmark—the tutoring ban has a more pronounced effect on reducing parental investment. Consequently, it also leads to a larger reduction in pre-college human capital and a greater improvement in mental health capital relative to the benchmark. The impacts of the ban on average parental welfare and child expected lifetime income are quantitatively similar to those in the benchmark case.

When the model abstracts from child mental health formation, as shown in Column (iii), the tutoring ban results in a much larger decline in child expected lifetime income

due to the reduction in pre-college human capital, alongside smaller welfare gains for parents. This occurs because a model without mental health omits the key benefits of the ban: it eliminates the negative effect of tutoring investment on mental health and improves the allocation between the two types of human capital. Consequently, such a model underestimates the welfare gains and overestimates the losses in child expected income.

Inequality and Mobility As shown in **??**, the tutoring ban disproportionately reduces private investment by families with high-income parents and high-ability children. Consequently, as shown in Panel (C) of of Table 6, intergenerational persistence in pre-college human capital and educational attainment declines, regardless of whether the model incorporates child mental health formation. However, in the benchmark that incorporates child mental health formation, the tutoring ban improves children's mental health capital, which increases its intergenerational persistence. As a result, abstracting from mental health in the model implies a larger reduction in intergenerational income persistence resulting from the tutoring ban.

7.2 Long-Run Policy Impacts and the Welfare Cost of Educational Competition

Our benchmark model, calibrated to the Chinese economy of the 2010s, is a non-stationary two-generation life cycle model. This structure is necessary due to the significant gap in the college labor share between the parent and child generations. However, this model omits an important margin: intergenerational dynamic effects. These effects capture how parental investment decisions impact not only their children but also their grandchildren and subsequent generations.

Although a ban on private tutoring raises current parental welfare, its effects on children—lower pre-college human capital, higher mental health capital, and a reallocation between these two forms of capital—can propagate to future generations through two channels. First, lower pre-college human capital reduces adult income, which in turn lowers monetary investment in the next generation's pre-college human capital. This produces persistent losses in pre-college human capital and depresses income across generations. Second, better mental health raises adult income and directly improves next-generation mental health through the components $\varsigma_{m,c}\log(m_p)$ and $\vartheta_{m,c}\log(w^s y(h_p, m_p, z))$ in the child mental-health capital formation equation (Equation (9)). This yields persistent gains in mental health capital and a positive income effect across generations. The

net intergenerational effect on income is ambiguous. The positive effect of improved mental health and the negative effect of reduced pre-college human capital work in opposite directions. This ambiguity also affects the parental income component (i.e., $\vartheta_{m,c}\log(w^s y(h_p, m_p, z))$) of child mental health formation endogenously across generations, making their combined effect on long-run investment levels ambiguous as well.

To assess the quantitative importance of these intergenerational dynamics, this section first examines the long-run effects of the private tutoring ban. We then explore whether an alternative policy could mitigate the distortions caused by competitive incentives and improve the allocation of resources between pre-college human capital and mental health capital.

To evaluate the long-run policy impacts, we start from the two-generation model with an exogenous distribution of parents and then examine the policy effects from the current parent generation onwards. This means all offspring generations are shaped by various policies and are endogenously determined. This scenario is compared to a status quo where the model remains calibrated to the same initial conditions, but no policy intervention is implemented. Over sufficiently many generations, the initial distribution of households will converge to a stationary distribution. By comparing the stationary equilibrium under a certain policy scenario to the one without any policy intervention, we can determine the long-run impacts of the policy. To ensure comparability between the short-run effects discussed in Section 7.1 and the long-run effects, we fix factor prices at our benchmark level and solve for a partial equilibrium, while simultaneously keeping the college capacity constraint and the government budget satisfied. For each policy experiment, we consider three cases: (1) the baseline, i.e., the stationary-distribution version of the benchmark model in Section 3; (2) an "ignore mental health" specification; and (3) a specification without a mental-health margin.

7.2.1 Intergenerational Impacts of Banning Private Tutoring

Figure 14 presents the intergenerational impacts of the private tutoring ban from the benchmark generation to subsequent generations. Each generation in the figure encompasses a complete life cycle of parents (from independence to death) and their children (from birth to independence). Generation 1 represents the parent generation, along with their children, in our benchmark economy, while generation 2 represents the children of generation 1 who become parents and have their own children, and so on.

In the short run, tutoring ban improves the consumption-investment allocation for parents by mitigating the negative externalities of competition, resulting in an approximate 0.76% welfare gain in the benchmark model. The welfare gain is smaller when the model

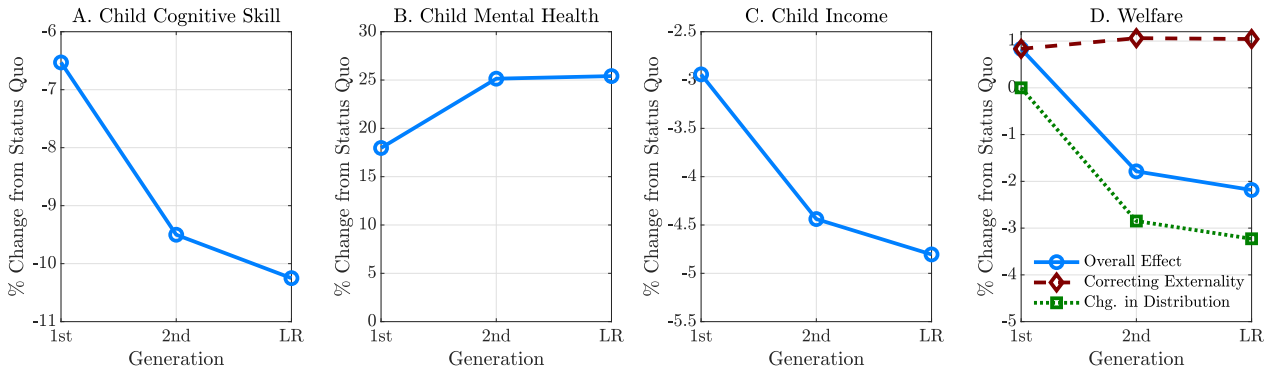


Figure 13: Intergenerational Impacts of Tutoring Ban: Decomposition

abstracts from the mental health margin. This effect is shown by the percent change in variables of interest for generation 0 relative to the status quo (an economy without policy intervention) in Figure 14, which also corresponds to the short-run impacts reported in Table 6.

In the long run, abstracting from mental health, reducing investment in a child today lowers that child's pre-college human capital and income and also produces a less productive parent in the next generation. As shown in Column (iii) of Table 7, these intergenerational dynamics generate persistent human capital losses that eventually outweigh the ban's short-run gains, culminating in a welfare loss of 2.7% (Column (iii)(b)). When we incorporate the mental health margin (Column (i) of Table 7)—so that better mental health raises adult income and directly improves the next generation's mental health through healthier parents and richer family resources—the sustained mental health gains in income and welfare are still insufficient to offset the persistent losses in pre-college human capital. Thus, the tutoring ban worsens the allocation between the two forms of human capital, yielding a net welfare loss of 2.4%.

The Role of Intergenerational Transmission of Mental Health Although persistent mental-health gains are not large enough to offset persistent cognitive-skill losses and the resulting deterioration in the intergenerational allocation of human capital across skills, the intergenerational mental-health channel still materially improves welfare. As shown by the orange lines in Figure 14, shutting down this channel—by removing the effect of better mental health on adult income and, in turn, on children's mental health through parental mental health and family resources—increases the long-run welfare losses from the tutoring ban by roughly one-third.

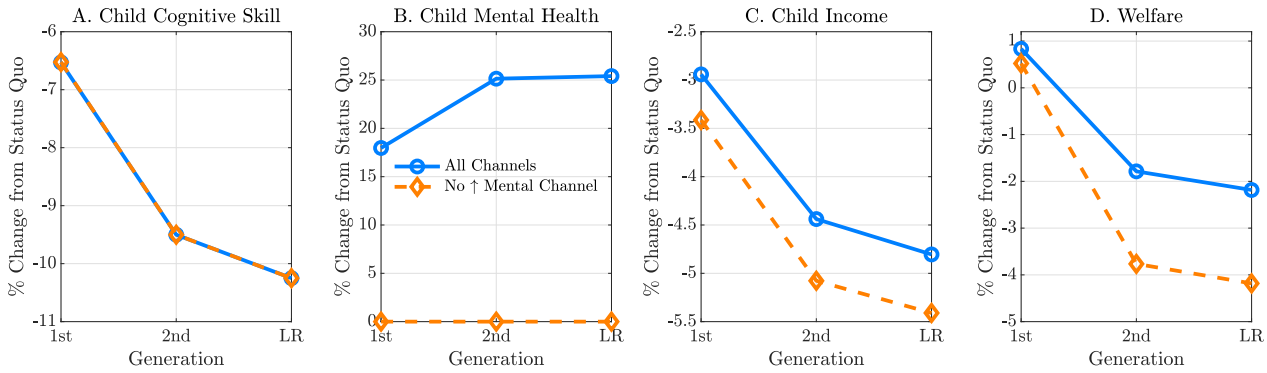


Figure 14: Intergenerational Transmission of Mental Health

7.2.2 Long-Run Impacts of Private Education Investment Tax

Building on our analysis in Section 6, we propose a policy intervention to improve the allocation of educational resources. Our results indicate that reallocating investment from private tutoring to public funding—while controlling for aggregate investment levels—fosters a more efficient balance between pre-college human capital and mental health capital, thereby improving lifetime income and welfare. To operationalize this in reality, we propose a linear tax on tutoring expenditures that exceed a threshold (i.e., the portion $i > \bar{i}$), with the generated revenue used to increase public investment g.[28] This approach is particularly advantageous in the long run: the tax reduces the mental health costs associated with educational competition, while the public subsidy compensates for the pre-college human capital losses induced by the tax. Critically, both of these positive effects are reinforced through intergenerational dynamics.

We choose a tax rate to maximize the long-run ex ante lifetime utility of newly economically active individuals under the veil of ignorance, as defined in Equation (A5), while satisfying the financing constraint, the government budget constraint, and the college capacity constraint. More details on this problem are provided in Appendix C.

We report the by-generation and long-run impacts for the baseline case in Figure 14 and Column (i)(c) of Table 7, respectively. Compared to an outright ban on private tutoring, the tax policy improves welfare for all future generations. It achieves this by more effectively balancing two objectives: it dampens the competition incentive that creates over-investment and misallocation of pre-college human capital and mental health capital while minimizing losses in pre-college human capital. This is evidenced by a much smaller reduction in private investment (15.5% under the tax versus 56.1% under

[28]This policy has a real-world counterpart in the U.K.'s recent introduction of a 20% value-added tax on private school tuition fees, the proceeds of which are directed toward public education funding.

the ban). Consequently, while the ban results in a significant long-run welfare reduction of approximately 2.4%, the tax-and-subsidy policy delivers a net welfare gain of 0.8%.

When parents do not value their children's mental health (Column (ii)(c) of Table 7), the optimal tax rate is higher (34% vs. 28% in the baseline) and the welfare gain is larger (1.1% vs. 0.8%). If we abstract from the mental-health margin altogether (Column (iii)(c) of Table 7)—so the tax generates no additional mental-health benefits—the welfare gain falls to 0.5%.

Welfare Cost of Educational Competition By interpreting the allocation under the optimal policy as a proxy for a social planner's solution, we can frame these welfare gains as the estimated cost of educational competition in the status quo economy. From this perspective, the welfare cost of the educational competition is 0.50 percent of aggregate consumption when mental health is excluded. When mental health is incorporated, this cost rises dramatically to between 0.79 and 1.06 percent—an increase of 58% to 200%. This stark contrast underscores the critical importance of endogenizing child mental health formation in models of educational competition and policy evaluation.

Table 7: Long-Run Policy Impacts

	(i)			(ii)			(iii)		
	(a)	(b)	(c)	(a)	(b)	(c)	(a)	(b)	(c)
	Baseline	Ban	Tax	Ignore Mental Health	Ban	Tax	Abstract from Mental Health	Ban	Tax
	Status Quo			Status Quo			Status Quo		
(A) Aggregates									
Optimal Tax Rate	-	-	28%	-	-	34%	-	-	23%
Private Investment	0.160	-56.1%	-15.5%	0.184	-61.9%	-17.8%	0.152	-53.6%	-11.1%
– Tutoring expenditures	0.094	-100.0%	-3.5%	0.118	-100%	-47.1%	0.085	-100%	-35.9%
Cram school participation rate	59.8%	0.0%	56.3%	63.4%	0.0%	58.4%	60.0%	0.0%	57.7%
Public investment									
Pre-college human capital	1.284	-8.7%	-0.9%	1.312	-10.7%	-1.1%	1.277	-7.4%	-0.2%
Mental health capital	-0.039	+0.55 sd	+0.19 sd	-0.195	+0.75 sd	+0.27 sd	-	-	-
Lifetime income	1.241	-4.0%	-0.3%	1.251	-4.9%	-0.3%	1.232	-3.6%	-0.1%
Consumption	1.174	-1.6%	+0.5%	1.177	-1.9%	+0.7%	1.167	-1.6%	+0.3%
(B) Welfare									
CEV	-	-2.4%	0.8%	-	-2.8%	+1.1%	-	-2.7%	+0.5%
(C) Inequality									
Var log lifetime income	0.131	-9.5%	-4.4%	0.132	-9.7%	-4.4%	0.131	-7.1%	-2.3%
(D) Intergenerational Persistence									
Pre-college human capital	0.419	-58.1%	-16.1%	0.439	-59.9%	-16.3%	0.370	-50.7%	-10.4%
Education	0.299	-90.2%	-23.5%	0.300	-90.2%	-22.9%	0221	-85.1%	-14.2%
Mental health	0.153	+34.1%	+12.5%	0.135	+52.3%	+18.8%	-	-	-
Lifetime income							0.371	-53.6%	-10.8%

Note: "Status Quo" refers to the scenario where we start from the two-generation model with an exogenous distribution of parents and iterate until the distribution converges to a stationary one. "Ban" and "Tax" refer to the scenarios where we impose the ban or tax on the parent generation in the benchmark and iterate until the distribution converges to a stationary one, and numbers reported are percent change relative to the status quo. In each steady state, factor prices are fixed at the benchmark level, while the college capacity constraint and the government budget are satisfied. Welfare is calculated as CEV based on ex ante newly economically active individuals as in (A5).

8 Conclusion

Educational competition can incentivize parents to overinvest in skills that boost exam performance—often through cram schools and private tutoring—at the expense of their children's mental health. These distorted investments can misallocate human capital across skill types, leading to persistent losses in income and welfare, with significant implications for inequality and intergenerational mobility.

In this paper, we develop a heterogeneous-agent life-cycle model in which parents invest in two types of human capital for their children at multiple stages to compete for limited college slots: pre-college human capital (skills crucial for exam performance and college admission) and mental health capital. Both forms of capital accumulate throughout childhood and adolescence and together shape adult labor-market productivity. The core element is that cram school participation improves exam performance but harms mental health. We discipline the magnitude of this trade-off using reduced-form IV evidence on the impact of cram-school participation on child mental health. Estimating the model with Chinese data, we find that educational competition imposes an aggregate welfare cost equivalent to roughly 0.8–1.1% of total consumption. Omitting the mental health channel reduces this cost by nearly half, underscoring the importance of accounting for mental health in educational outcomes. Finally, we evaluate two policy responses. While a strict tutoring ban would curb overinvestment, it also generates persistent losses in pre-college human capital. In contrast, a combined policy of a private investment tax and a public investment subsidy strikes a better balance: it dampens the incentive to overinvest (reducing the misallocation of pre-college and mental health capital) while minimizing the loss of pre-college human capital, ultimately improving overall welfare.

References

Abbott, Brant. 2022. "Incomplete markets and parental investments in children." *Review of Economic Dynamics* 44.

Abbott, Brant, Giovanni Gallipoli, Costas Meghir, and Giovanni L. Violante. 2020. "Education Policy and Intergenerational Transfers in Equilibrium." *Journal of Political Economy* 25.

Abramson, Boaz, Job Boerma, and Aleh Tsyvinski. 2025. "Macroeconomics of Mental Health." Working Paper.

Agostinelli, Francesco, Matthias Doepke, Giuseppe Sorrenti, and Fabrizio Zilibotti. 2022. "When the Great Equalizer Shuts Down: Schools, Peers, and Parents in Pandemic Times." *Journal of Public Economy* .

———. Forthcoming. "It Takes a Village: The Economics of Parenting with Neighborhood and Peer Effects." *Journal of Political Economy* .

Agostinelli, Francesco and Matthew Wiswall. 2024. "Estimating the Technology of Children's Skill Formation." Working Paper.

Ameriks, John, Joseph Briggs, Andrew Caplin, Matthew D. Shapiro, and Christopher Tonetti. 2020. "Long-Term-Care Utility and Late-in-Life Saving." *Journal of Political Economy* 128 (6).

Attanasio, Orazio, Richard Blundell, Gabriella Conti, and Giacomo Mason. 2020. "Inequality in socio-emotional skills: A cross-cohort comparison." *Journal of Public Economics* 191.

Bai, Chong-En, Chang-Tai Hsieh, and Yingyi Qian. 2006. "The Return to Capital in China." Working Paper.

Barro, R. J. and G. S. Becker. 1989. "Fertility Choice in a Model of Economic Growth." *Econometrica* 57.

Becker, Gary S. 2007. "Health as Human Capital:Synthesis and Extensions." *Oxford Economic Papers* 59 (3).

Bound, John, Brad Hershbein, and Bridget Terry Long. 2009. "Playing the Admissions Game: Student Reactions to Increasing College Competition." *Journal of Economic Perspectives* 4 (23).

Braun, R. Anton, Karen A. Kopecky, and Tatyana Koreshkova. 2017. "Old, Sick, Alone, and Poor: A Welfare Analysis of Old-Age Social Insurance Programmes." *The Review of Economic Studies* 84 (2).

———. 2019. "Old, Frail, and Uninsured: Accounting for Features of the U.S. Long-Term Care Insurance Market." *Econometrica* 87 (3).

Caucutt, Elizabeth M. and Lance J. Lochner. 2020. "Early and Late Human Capital Investments, Borrowing Constraints, and the Family." *Journal of Political Economy* 128.

Celik, Murat Alp. 2023. "Does the Cream Always Rise to the Top? The Misallocation of Talent in Innovation." *Journal of Monetary Economics* 133.

Chen, Kuan-Ming, Janet Currie, Hui Ding, and Wei-Lun Lo. 2024. "School Milestones Impact Child Mental Health in Taiwan." NBER Working Paper No. 32842.

Chen, Wei, Xilu Chen, Chang-Tai Hsieh, and Zheng Song. 2019. "A Forensic Examination of China's National Accounts." *Brookings Papers on Economic Activity* .

Chyn, Eric and Diego Daruich. 2024. "An Equilibrium Analysis of the Effects of Neighborhood-based Interventions on Children." Working Paper.

Cole, Harold L, Soojin Kim, and Dirk Krueger. 2018. "Analysing the Effects of Insuring Health Risks: On the Trade-off between Short-Run Insurance Benefits versus Long-Run Incentive Costs." *The Review of Economic Studies* 86 (3).

Cronin, Christopher J, Matthew P Forsstrom, and Nicholas W Papageorge. 2005. "What Good Are Treatment Effects Without Treatment? Mental Health and the Reluctance to Use Talk Therapy." *The Review of Economic Studies* 92 (3).

Cuddy, Emily and Janet Currie. 2026. "Rules versus Discretion: Treatment of Mental Illness in US Adolescents." *Journal of Political Economy* 134 (1).

Cunha, F., J. J. Heckman, and S.M. Schennach. 2010. "Estimating the technology of cognitive and noncognitive skill formation." *Econometrica* 78 (3).

Currie, Janet. 2020. "Child Mental Health as Human Capital."

———. 2025. "Presidential Address: Investing in Children to Address the Child Mental Health Crisis." *The American Economic Review* 115 (5).

Currie, Janet and Mark Stabile. 2007. "Mental Health in Childhood and Human Capital." NBER Working Paper No. 13217.

Daruich, Diego. 2022. "The Macroeconomic Consequences of Early Childhood Development Policies." Working Paper.

Daruich, Diego and Julian Kozlowski. 2020. "Explaining intergenerational mobility: The role of fertility and family transfers." *Review of Economic Dynamics* 36.

Del Boca, D., C. Flinn, and M. Wiswall. 2014. "Household choices and child development." *The Review of Economic Studies* 81 (1).

De Nardi, Mariacristina, Eric French, and John B. Jones. 2010. "Why Do the Elderly Save? The Role of Medical Expenses." *Journal of Political Economy* 118 (1).

De Nardi, Mariacristina, Eric French, and John Bailey Jones. 2016. "Medicaid Insurance in Old Age." *The American Economic Review* 106 (11).

De Nardi, Mariacristina, Svetlana Pashchenko, and Ponpoje Porapakkarm. 2025. "The Lifetime Costs of Bad Health." *The Review of Economic Studies* 92 (3).

Ding, Haiyan and Hui He. 2018. "A Tale of transition: an empirical analysis of economic inequality in urban China, 1986-2009." *Review of Economic Dynamics* 29.

Dixon-Roman, Ezekiel J., Howard T. Everson, and John J. Mcardle. 2013. "Race, Poverty and SAT Scores: Modeling the Influences of Family Income on Black and White High School Students' SAT Performance." *Teachers College Record* 115 (4).

Doepke, Matthias and Fabrizio Zilibotti. 2017. "Parenting With Style: Altruism and Paternalism in Intergenerational Preference Transmission." *Econometrica* 10 (2).

Evans, W.N. and C.L. Garthwaite. 2014. "Giving mom a break: The impact of higher EITC payments on maternal health." *American Economic Journal: Economic Policy* 6 (2).

Fang, Hanming and Dirk Krueger. 2022. "The Affordable Care Act After a Decade: Its Impact on the Labor Market and the Macro Economy." *Annual Review of Economics* 14.

Fogli, Alessandra and Veronica Guerrieri. 2019. "The End of the American Dream? Inequality and Segregation in US Cities." NBER Working Paper.

French, Eric. 2005. "The Effects of Health, Wealth, and Wages on Labour Supply and Retirement Behaviour." *The Review of Economic Studies* 72 (2).

French, Eric and John Bailey Jones. 2011. "THE EFFECTS OF HEALTH INSURANCE AND SELF-INSURANCE ON RETIREMENT BEHAVIOR." *Econometrica* 79 (3).

Gallo, Joseph J., James C. Anthony, and Bengt O. Muthen. 1994. "Age Differences in the Symptoms of Depression: A Latent Trait Analysis." *Journal of Gerontology* 49 (6).

Green, Jennifer Greif, Katie A. McLaughlin, and Patricia A. Berglund. 2010. "Childhood Adversities and Adult Psychiatric Disorders in the National Comorbidity Survey Replication I: Associations With First Onset of DSM-IV Disorders." *Archives of General Psychiatry* 67 (2).

Greenwood, Jeremy, Nezih Guner, and Karen A Kopecky. 2022. "The Downward Spiral." Working Paper 29764, National Bureau of Economic Research.

Grossman, Michael. 1972. "On the concept of health capital and the demand for health." *Journal of Political Economy* 80 (2).

Gu, Shijun and Lichen Zhang. 2025. "The Macroeconomic Consequences of Competition for College Admissions." Working Paper.

Guvenen, Fatih, Greg Kaplan, Jae Song, and Justin Weidner. 2022. "Lifetime Incomes in the United States over Six Decades." *American Economic Journal: Applied Economics* 14.

Hall, Robert E. and Charles I. Jones. 2007. "The Value of Life and the Rise in Health Spending*." *The Quarterly Journal of Economics* 122 (1).

Heckman, J., J. Stixrud, and S. Urzua. 2006. "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior." *Journal of Labor Economics* 24.

Heckman, James J. and Y. Rubinstein. 2001. "The importance of noncognitive skills: Lessons from the GED testing program." *American Economic Review* 91 (2).

Hosseini, Roozbeh, Karen Kopecky, and Kai Zhao. 2025. "How Important Is Health Inequality for Lifetime Earnings Inequality?" *The Review of Economic Studies* .

Hosseini, Roozbeh, Karen A. Kopecky, and Kai Zhao. 2022. "The Evolution of Health over the Life Cycle." *Review of Economic Dynamics* 45.

Hubbard, R. Glenn, Jonathan Skinner, and Stephen P. Zeldes. 1995. "Precautionary Saving and Social Insurance." *Journal of Political Economy* 103 (2).

Huggett, Mark. 1996. "Wealth distribution in life-cycle economies." *Journal of Monetary Economics* 38.

Huggett, Mark, Gustavo Ventura, and Amir Yaron. 2011. "Sources of Lifetime Inequality." *American Economic Review* 101.

Husain, Mustafa M., A. John Rush M.D., Harold A. Sackeim, Stephen R. Wisniewski, Shawn M. McClintock, Nicoletta Craven, Joel Holiner, Jeff R. Mitchell, G.K. Balasubramani, and Robert Hauger. 2005. "Age-Related Characteristics of Depression: A Preliminary STAR*D Report." *The American Journal of Geriatric Psychiatry* 13 (10).

———. 2023. "Child mental health and income gradient from early childhood to adolescence: Evidence from the UK." *SSM - Population Health* 24.

Johnson, Dylan, Gabrielle Dupuis, Justin Piche, Zahra Clayborne, and Ian Colman. 2018. "Adult mental health outcomes of adolescent depression: A systematic review." *Depression and Anxiety* .

Jolivet, Gregory and Fabien Postel-Vinay. 2022. "A Structural Analysis of Mental Health and Labour Market Trajectories." *The Review of Economic Studies* 89 (1).

Kang, Hyunjae. 2022. "Parental Investment, Child's Efforts, and Intergenerational Mobility." Working Paper.

Keane, Michael P. and Kenneth I. Wolpin. 1997. "The Career Decisions of Young Men." *Journal of Political Economy* 105 (3):473–522.

Kim, Seongeun, Michele Tertilt, and Minchul Yum. 2024. "Status Externalities in Education and Low Birth Rates in Korea." *American Economic Review* .

Kling, J. R., J.B. Liebman, and L.F. Katz. 2007. "Experimental analysis of neigh- borhood effects." *Econometrica* 75 (1).

Kopecky, Karen A. and Tatyana Koreshkova. 2014. "The Impact of Medical and Nursing Home Expenses on Savings." *American Economic Journal: Macroeconomics* 6 (3).

Kosik, R., G. Mandell, A.P. Fan, T. Nguyen, J. Chen, and W. Eaton. 2018. "The association between childhood educational attainment and adult mental health and status: A thirty-year longitudinal follow up study." *The Euroepan Journal of Psychiatry* 32 (2).

Kotera, Tomoaki and Ananth Seshadri. 2017. "Educational policy and intergenerational mobility." *Review of Economic Dynamics* 25.

Krueger, Dirk, Alexander Ludwig, and Irina Popova. 2024. "Shaping Inequality and Intergenerational Persistence of Poverty: Free College or Better Schools." NBER Working Paper.

Lee, Sang Yoon (Tim) and Ananth Seshadri. 2019. "On the Intergenerational Transmission of Economic Status." *Journal of Political Economy* .

Liu, Shuxia, Yuan Fang, Zhongyan Su, Jimin Cai, and Zhiyan Chen. 2023. "Factor structure and measurement invariance of the 8-item CES-D: a national longitudinal sample of Chinese adolescents." *BMC Psychiatry* 868 (23).

Low, Hamish, Costas Meghir, and Luigi Pistaferri. 2010. "Wage Risk and Employment Risk over the Life Cycle." *The American Economic Review* 100 (4).

Low, Hamish and Luigi Pistaferri. 2015. "Disability Insurance and the Dynamics of the Incentive Insurance Trade-Off." *The American Economic Review* 105 (10).

Maccoby, E. E. and J. A. Martin. 1983. *Socialization in the context of the family: Parent-child interaction. In P. H. Mussen (Series Ed.) E. M. Hetheringtono(Vol. Ed.), Handbook of Child Psychology: Vol. IV. Socialization, Personality and Social Development*. New York: Wilepy.

Mahler, Lukas and Minchul Yum. 2024. "Lifestyle Behaviors and Wealth-Health Gaps in Germany." *Econometrica* 92 (5).

Mullins, Joseph. 2022. "Designing Cash Transfers in the Presence of Children's Human Capital Formation." Working Paper.

Ozkan, Serdar. 2025. "Income differences and health disparities: Roles of preventive vs. curative medicine." *Journal of Monetary Economics* 150.

Papageorge, Nicholas W., Victor Ronda, and Yu Zheng. 2022. "The Economic Value of Breaking Bad: Misbehavior, Schooling and the Labor Market." Working Paper.

Ramey, G. and V. A. Ramey. 2010. "The Rug Rat Race." *Brookings Papers on Economic Activity* .

Restuccia, Diego and C. Urrutia. 2004. "Intergenerational persistence of earnings: The role of early and college education." *American Economic Review* 94.

Seth, Michael J. 2002. *Education Fever: Society, Politics, and the Pursuit of Schooling in South Korea*. University of Hawaii Press.

Wagner, Stefanie, Daniel Wollschläger, Nadine Dreimüller, Jan Engelmann, David P. Herzog, Sibylle C. Roll, André Tadić, and Klaus Lieb. 2020. "Effects of age on depressive symptomatology

and response to antidepressant treatment in patients with major depressive disorder aged 18 to 65 years." *Journal of Gerontology* 99.

Wei, Shang-Jin and Xiaobo Zhang. 2011. "The Competitive Saving Motive: Evidence from Rising Sex Ratios and Savings Rates in China." *Journal of Political Economy* 119 (3).

Wilson, Naomi, Helen Minnis, Shari McDaid, and Anna Pearce. 2025. "Household income and children's mental health outcomes: the mediating role of maternal wellbeing and parent–child relationship quality." *European Child and Adolescent Psychiatry* .

Yu, J. and G. Zhu. 2013. "How uncertain is household income in China." *Economics Letters* 120.

Yum, Minchul. 2023. "Parental Time Investment and Intergenerational Mobility." *International Economic Review* .

Zhao, Ruonan, Jian Wang, Jiaxu Lou, Mei Liu, Jiahui Deng, Derong Huang, and Huiling Fang. 2024. "The effect of education level on depressive symptoms in Chinese older adults–parallel mediating effects of economic security level and subjective memory ability." *BMC Geriatrics* 24 (635).

Zhou, Kai. 2016. "Non-cognitive skills: Definitions, measurement and malleability. Paper commissioned as background for the UNESCO 2016 Global Education Mon- itoring Report, Education for People and Planet: Creating Sustainable Futures for All." https://unesdoc.unesco.org/ark:/48223/pf0000245576. Accessed: Sept. 6, 2025.

Online Appendix

Educational Competition, Mental Health, and the Allocation of Human Capital: A Macroeconomic Analysis

Shijun Gu
CUFE

Lichen Zhang
HKU

Table of Contents

A Background Information **1**
A.1 Institutional Background on the Higher Education Systems of China 1
A.2 Mental Health and Its Relationship to Noncognitive Skills in the Economics Literature 2

B Data and Measurement **4**
B.1 Definitions for Variables 4
B.2 Estimated Impact of Cram School Participation on Mental Health for Full CEPS Sample 13
B.3 Distribution of Mental Health Shocks 14
B.4 Child Mental Health by SES 14
B.5 Mental Health over Life Cycle 14
B.6 Investment in Children 16
B.7 Private Education Expenditure 18
B.8 Reduced-form Evidence on the Impact of Tutoring on Mental Health 21
B.9 Public Education Expenditures 24
B.10 Inter-vivos Transfers 25
B.11 Age Profile of Child Cognitive Ability 25
B.12 Age Profile of Earnings and Initial Distribution of Parents 26
B.13 Admission Probability Function 26

C Welfare Measure **30**

D Full Definition of the Stationary Equilibrium **31**

E Additional Model Setup and Quantitative Analysis Results **33**
E.1 Recursive Problem at Retirement Stage 33

A Background Information

A.1 Institutional Background on the Higher Education Systems of China

The Chinese higher education system features the central-planned nature. Our model resembles China's public college system in a parsimonious manner. The system centers on the annual *Gaokao*, a mandatory standardized test for higher education admission.

The Gaokao system in China is one of the world's most rigorous college entrance exams. It is held annually, usually in early June. Unlike the ACT or SAT, which require a fee to take and are more likely to be retaken by high-income students to boost scores further (Dixon-Roman, Everson, and Mcardle, 2013), *Gaokao* is held only once a year and is free for all students.

The exact format and content of the exam can vary by region, but it generally covers Chinese literature, mathematics, and a foreign language (usually English) as compulsory subjects. In addition, students have to choose between two tracks for their elective subjects: science (which includes physics, chemistry, and biology) or social science (which includes politics, history, and geography).

After the Gaokao exam, universities set their admission cutoff scores based on the number of students they plan to admit and the overall performance of the students who have applied to each university that year. Consequently, the cutoff score can vary significantly from year to year and across different universities and provinces.

However, this does not mean that students and parents have no information regarding their chances of being admitted before the exam. Parents are generally aware of their children's probability of college admission based on their academic ability rank. In practice, universities usually announce the number of students they plan to admit before the college entrance exam. Additionally, most high schools in China administer monthly, and sometimes even weekly, practice tests that simulate the real college entrance exams. These frequent assessments provide students with a clear understanding of their rank.

As a result, students and their parents can infer the college admission probability based on the students' rank in practice tests and the planned number of college slots. However, this does not guarantee admission based solely on academic rank, as some uncertainty remains when taking the actual college entrance exam. For example, a student might underperform on the exam day due to unforeseen circumstances, preventing them from achieving results as good as in the practice tests.

Our model setup is consistent with reality. When parents decide how much to invest in their children, they consider the ability-rank-based college admission function (as shown in Figure 10) and the college capacity (determined by the shifter parameter ζ in the admission function $\chi(\cdot, \zeta)$). This endogenously determines their children's college admission probability based on their ability rank.

Once the scores are released, students submit an ordered list of their preferred universities. Universities then admit students in descending order of their Gaokao scores, starting with the highest score, until all their spots are filled.

The competition is intense due to the large number of students taking the Gaokao and the limited number of seats in universities, especially top universities. Consequently, the Gaokao score heavily influences the opportunities for higher education and future career prospects of Chinese students.

This system has been criticized for placing excessive pressure on students and for its inability to fully assess a student's capabilities beyond academic knowledge. However, it is also seen as a relatively fair way of deciding who gets admitted to university, as it is based purely on exam performance and is less susceptible to corruption or favoritism.

A.2 Mental Health and Its Relationship to Noncognitive Skills in the Economics Literature

As pointed out in Currie (2020), *"Until recently, few economists explicitly studied childhood mental health; concepts related to mental health have frequently been subsumed in the broader category of noncognitive skills."* Heckman and Rubinstein (2001) emphasize that individuals with similar levels of cognitive skill, as measured by standardized test scores, can have vastly different outcomes, implying that something beyond cognition matters. This idea was further developed in Heckman, Stixrud, and Urzua (2006), who show that measures of noncognitive skills strongly predict education, earnings, and risky behaviors. Over time, a large literature has developed on the importance of noncognitive skills, a term that embraces concepts as diverse as the Big Five personality traits (extraversion, agreeableness, openness, conscientiousness, and neuroticism), grit or perseverance, and teachable competencies such as self-control and social skills (Zhou, 2016).

A.2.1 Measurement of Noncognitive Skills in the Literature

In Cunha, Heckman, and Schennach (2010), the Behavior Problems Index (BPI) is used. They use the following BPI subscores: (i) antisocial, (ii) anxious/depressed, (iii) headstrong, (iv) hyperactive, and (v) peer problems. A child with a high score on the anxious/depressed subscale is one who experiences sudden changes in mood, feels unloved, is fearful, or feels worthless or inferior. In Heckman, Stixrud, and Urzua (2006), the noncognitive measures are the Rotter Locus of Control Scale (Rotter, 1966), administered in 1979, and the Rosenberg Self-Esteem Scale (Rosenberg, 1965), administered in 1980. The Rotter scale measures the degree of control individuals feel they possess over their lives and has been used in studies of the role of noncognitive skills in labor-market outcomes (Osborne-Groves, 2006). The Rosenberg scale measures perceptions of self-worth.

A.2.2 Relationship to Internalizing and Externalizing Behavior

Attanasio et al. (2020) divide noncognitive skills in young children into internalizing and externalizing behaviors: anxiety, depression, shyness, unassertiveness, and fearfulness fall into the former category, while conduct disorders, hyperactivity disorder, and antisocial behavior fall into the latter. Attanasio et al. (2020) show that these skills at age five are important predictors of adult outcomes. Papageorge, Ronda, and Zheng (2022) note that internalizing behavior is commonly associated with depressive disorders, which motivates our choice of depression as the dedicated measure (Regier, Kuhl, and Kupfer, 2013; Kendler and Myers, 2014). They find a positive correlation between externalizing and internalizing behavior, along with a negative correlation between these two socioemotional skills and cognition. Aligned with prior work, they find that externalizing behavior lowers educational attainment for males; however, they also provide novel evidence that it increases earnings for both males and females.

A.2.3 Measurement of Mental Health in the Literature

Mental health is a broad clinical construct referring to the overall state of an individual's psychological and emotional well-being. Measures of mental health assess the presence or absence of psychopathology (mental disorders) as well as positive states of mental wellness. In practice, mental health is measured either using standardized diagnostic criteria (e.g., DSM-5 or ICD-11) or along a continuum using screening questionnaires that often capture symptoms of specific disorders (e.g., anxiety, depression, ADHD, conduct disorder).

The Panel Study of Income Dynamics (PSID) measures respondents' mental health using the Kessler Psychological Distress Scale (K6), which is widely used in the epidemiological and psychiatric literatures to assess the prevalence and severity of mental illness and is the primary mental health measure used in U.S. government-administered health surveys as well as the WHO World Mental Health Surveys. The K6 is calculated from responses to six questions (Kessler et al., 2002, 2003). Respondents are asked: "In the past 30 days, about how often did you feel (1) sad, (2) nervous, (3) restless or fidgety, (4) hopeless, (5) that everything was an effort, and (6) worthless?" For each item, individuals respond (0) none of the time, (1) a little of the time, (2) some of the time, (3) most of the time, or (4) all of the time. The K6 score is the sum of responses to the six questions.

The K6 has been extensively validated against clinical diagnoses and has been shown to consistently predict mood and anxiety disorders (Kessler et al., 2002, 2003; Furukawa, Kessler, Slade, and Andrews, 2003; Cairney et al., 2007). Following Kessler et al. (2008), individuals can be classified into three mental health states based on the K6: scores of 13–24 indicate serious mental illness, 8–12 indicate mild mental illness, and 0–7 indicate no mental illness (healthy).

B Data and Measurement

B.1 Definitions for Variables

B.1.1 Mental Health

In our CFPS panel, the questions used to measure mental health are not consistent across waves. For example, the 2010 and 2014 waves use the K6 scale, the 2012 and 2016 waves use the CES-D 20-item scale, while the 2012, 2016, 2018, 2020, and 2022 waves use the CES-D 8-item scale (see Table A1 for more details). To maximize the use of information from all waves, we place all responses (from both the 8-item and 20-item waves) onto a single latent scale using an Item Response Theory (IRT) graded response model with linking.

We apply IRT with linking to a common metric. More specifically, we fit a graded response model (GRM) because the CES-D and K6 items are multiple-category ordered responses. There are four items common to the CES-D20, CES-D8, and K6 scales.

The four common questions are: *"In the past seven days, have you experienced the following feelings?"*

(1) I felt depressed.

(2) I felt that everything I did was an effort.

(3) I did not enjoy life.

(4) I felt sad.

For each question, the response options are four ordered categories:

- (0) almost never (less than 1 day);
- (1) sometimes (1–2 days);
- (2) often (3–4 days);
- (3) most of the time (5–7 days).

These four items act as anchor items to link the scales across waves.

We implement the following procedure step-by-step to construct a solid continuous measure for mental health for both adults and children in our CFPS sample.

1. **Response Harmonization:** Since the K6 and CES-D scales have different original response ranges, we first adjust the K6 response ranges (0–4) to make them consistent with those of the CES-D (0–3). The mapping is as follows: '0: never'; '1: sometime'; '2: half of the time'; '3: often, almost everyday'. Additionally, for the CES-D, we reverse the scores for items that are phrased positively using the transformation $r = 3 - x$.

2. **Data Stacking:** We stack all person-by-item responses from all waves into a single dataset. This includes all K6 and CES-D items. The full item set is the CES-D20 (20 items). The two subsets are the CES-D8 (8 items) and the K6 (5 items, excluding the item *"Frequency of feeling stressed last month"*). Items not administered in a particular wave are coded as missing (.).

3. **Model Fitting:** We fit a single unidimensional GRM to the entire stacked dataset. Modern IRT software handles this planned missingness by design; the four shared items anchor the scale to a common metric.

4. **Scoring:** Given the item responses, the IRT model estimates a latent trait score θ for each person and provides a standard error of measurement that varies with θ. We output person scores (EAP or MAP estimates for θ) for every observation (wave $\times$ person), along with their standard errors. These θ scores are on the same metric across all waves.

5. **Validity Checks:**

    - **Internal Consistency:** In the 2012 and 2016 waves, where both the CES-D8 and CES-D20 were administered, we estimate identical models with θ_{20} and θ_8 (the scores based on the 20-item and 8-item sets, respectively) separately as predictors in an external model (e.g., predicting a health outcome). We compare the standardized coefficients and metrics like AUC/R^2. They should be very similar if the metric is consistent.
    - **External Validity:** We examine the relationship between the estimated θ scores and external outcomes measured contemporaneously (e.g., self-rated health).

Table A1: Survey Questions on Mental Health in CFPS

Mental Health Measure	Items	Available Years		Range	Cutoff
		Adult Survey	Child Survey		
K6 Scale	(1) Frequency of feeling depressed last month (2) Frequency of feeling stressed last month (3) Frequency of feeling upset last month (4) Frequency of feeling hopeless last month (5) Frequency of feeling everything is difficult last month (6) Frequency of feeling life is meaningless last month	2010, 2014	2010, 2014	Each item ranges from 0 to 4. 0: never; 1: sometime; 2: half of the time; 3: often; 4: almost everyday	0-4: low or no distress; 5-12: moderate distress; 13+: serious mental illness
CES-D20 Scale	(1) I was bothered by things that usually don't bother me (2) I did not feel like eating; my appetite was poor (3) I felt I could not shake off the blues even with help from my family or friends (4) I felt that I was just as good as other people (5) I had trouble keeping my mind on what I was doing (6) I felt depressed (7) I felt that everything I did was an effort (8) I felt hopeful about the future (9) I thought my life had been a failure (10) I felt fearful (11) My sleep was restless (12) I was happy (13) I talked less than usual (14) I felt lonely (15) People were unfriendly (16) I enjoyed life (17) I had crying spells (18) I felt sad (19) I felt that people dislike me (20) I could not get going	2012, 2016, (2020, 2022)	2012, 2016, (2020, 2022)	Each item ranges from 0 to 3. 0: almost never; 1: sometimes; 2: often; 3: almost every day	0-15: healthy; 16+: depression (some use 20 instead of 16)
CES-D8 Scale	(1) I felt depressed (2) I felt that everything I did was an effort (3)My sleep was restless (4) I was happy (5) I felt lonely (6) I enjoyed life (7) I felt sad (8) I could not get going	2012, 2016, 2018, 2020, 2022	2012, 2016, 2018, 2020, 2022	Each item ranges from 0 to 3. 0: almost never; 1: sometimes; 2: often; 3: almost every day	0-6: healthy; 7+: depression (some use 8 instead of 7)

Table A2: Survey Questions on Non-Cognitive Skills in CFPS

Non-cognitive Ability Measure	**Items**	**Available Years**		**Range**
		Adult Survey	**Child Survey**	
Rosenberg Self-Esteem Scale	(1) I feel that I am a person of worth, at least on an equal plane with others (2) I feel that I have a number of good qualities (3) All in all, I am inclined to feel that I am a failure (4) I am able to do things as well as most other people (5) I feel I do not have much to be proud of (6) I take a positive attitude toward myself (7) On the whole, I am satisfied with myself (8) I wish I could have more respect for myself (9) At times I think I am no good at all	2014, 2016, 2018	2010, 2012, 2014, 2016, 2018	Each item ranges from 1 to 5. 1: totally disagree; 2: disagree; 3: neither agree nor disagree; 4: agree; 5: totally agree

Table A3: Survey Questions on Internalizing/Externalizing Behavior in CFPS

Internalizing/ Externalizing Behavior	**Items**	**Available Years**	**Range**
		Adult (Individual) Survey	
Internalizing	I feel angry when I have trouble learning. I worry about taking tests. I often feel lonely. I feel sad a lot of the time. I worry about doing well in school. I worry about finishing my homework I worry about having someone to play with at school. I feel ashamed when I make mistakes at school.	2018 (age 9-15) 2022 (age 9-20)	Each item ranges from 1 to 5. 1: Strongly disagree; 2: Disagree; 3: Neither agree nor disagree (do not read); 4: Agree; 5: Strongly agree
Externalizing	I often argue with other kids. It is hard for me to pay attention. I get distracted easily. It is hard for me to finish my school work. I get in trouble for talking and disturbing others. I get in trouble for fighting with other kids.	2018 (age 9-15) 2022 (age 9-20)	Each item ranges from 1 to 5. 1: Strongly disagree; 2: Disagree; 3: Neither agree nor disagree (do not read); 4: Agree; 5: Strongly agree

Table A4: Survey Questions on Big Five Personal Traits in CFPS

Big Five Personality Scale	**Items**	**Available Years**	**Range**
		Adult (Individual) Survey	
Conscientiousness	I do a thorough job. I tend to be lazy. (reverse coded) I do things efficiently.	2018 (aged 15 and above)	Each item ranges from 1 to 5. 1: Totally inapplicable; 2: Somewhat inapplicable; 3: Generally applicable; 4: Somewhat applicable; 5: Strongly agree
Extraversion	I am talkative. I am outgoing and sociable. I am reserved. (reverse coded)		
Agreeableness	I am sometimes rude to others. (reverse coded) I have a forgiving nature. I am considerate and kind to almost everyone.		
Openness	I am original and come up with new ideas. I value artistic experiences. I have an active imagination.		
Neuroticism	I worry a lot. I get nervous easily. I am relaxed and handle stress well. (reverse coded) *Note: the measure of emotional stability is the reverse of neuroticism*		

Table A5: Survey Questions on Mental Health in CEPS

Mental Health Measure	**Items**	**Range**
Wave 1	Do you have the feelings below in the last seven days? 1. feeling blue; 2. depressed; 3. unhappy; 4. not enjoying life; 5. sad	Each item ranges from 1 to 5. 1: never; 2: seldom; 3: sometimes; 4: often; 5: always
Wave 2	Do you have the feelings below in the last seven days? 1. feeling blue; 2. too depressed to focus on anything; 3. unhappy; 4. not enjoying life; 5. having no passion to do anything; 6. sad, sorrowful	Each item ranges from 1 to 5. 1: never; 2: seldom; 3: sometimes; 4: often; 5: always

B.1.2 Other Key Variables

College We determine an individual's college education status using information on years of education and school level. Following the World Bank's definition, we consider a four-year tertiary degree or its equivalent (e.g., a bachelor's degree in the United States) or higher as college education, and exclude less competitive tertiary vocational colleges as well as four-year private colleges.

Income Our baseline income measure is labor income, i.e. wages and salaries, earned in the previous survey year, deflated using the 2010 Consumer Price Index (CPI).

Test Scores and Pre-college Human Capital The CFPS includes detailed cognitive ability assessments comparable to the AFQT in the US. It comprises four types of cognitive tests: literacy, math, word recall, and numerical series. The CFPS alternates between two sets of these assessments across waves: Set A (literacy and math tests) was used in 2010, 2014, and 2018, while Set B (word recall and numerical series tests) was used in 2012, 2016, and 2020. We focus on the 2010, 2014, and 2018 waves featuring Set A to ensure consistent cognitive ability measurement. We measure pre-college human capital as the sum of math and literacy test scores. In the child survey, we can only observe children's cognitive ability scores starting at age 10, so we infer their average cognitive score before age 10 through extrapolation.[29] We normalize the average cognitive ability scores for the parent generation to be one. We report the age profile of child ability scores in Table 3 and also in Figure A8 in Appendix B.11.

Parental Investment We focus on parental monetary investments for children aged 6 to 18, primarily aimed at enhancing exam performance, especially through cram school attendance and private tutoring. Figure A4 in Appendix B.6 presents empirical evidence indicating that parental time investment in China is not a primary concern. The CFPS data cover private education expenditures including school fees, extracurricular activities and private tutoring, books, boarding, and transportation fees.[30]

Parenting Style Parenting style is a potentially important determinant of child mental health, as parents can shape their child's choices by molding their preferences (persuasion) or by constraining their choice set (Doepke and Zilibotti, 2017). Following Maccoby and Martin (1983), we categorize parenting styles into four types: permissive, neglecting, authoritative, and authoritarian. This classification is based on two dimensions: parental supervision and parental involvement. These dimensions are indices constructed from survey questions in the CEPS, as described in Section B.1.3.

Specifically:

- A **neglectful** style is defined as both supervision and involvement indices being below their respective averages.

[29] We use the average cognitive ability scores by age, ranging from age 10 to age 21, to extrapolate the average score for ages 6-9 using the Piecewise Cubic Hermite Interpolating Polynomial (PCHIP) method.

[30] We provide a detailed breakdown of private education expenditure by school stage in Appendix B.7, showing that for children aged 6-18, most expenditure, whether incurred in-school or out-of-school, is directed towards improving exam performance.

- An **authoritarian** style occurs when the supervision index is above average and the involvement index is below average.
- A **permissive** style is identified when the supervision index is below average and the involvement index is above average.
- An **authoritative** style is defined as both indices being above average.

As shown in Table A6, the authoritative parenting style is associated with the highest level of child mental health capital.

Table A6: Mental Health Capital by Parenting Style

Parenting Style	Cram Sch. Participation	School Fee	Extra. Tutoring Fee	Time (School)	Time (Tutoring)	Share of High SES	Mental Health
Neglecting	23.8%	938	865	17.8	5.5	5.1%	-0.11
Authoritarian	30.3%	829	970	19.5	7.4	4.8%	-0.14
Permissive	37.1%	874	1515	19.0	7.5	7.8%	0.04
Authoritative	40.0%	888	1214	19.4	8.6	6.3%	0.14

B.1.3 Variables in CEPS

We provide information on all the variables we use for running regressions regarding the CEPS in Section 2 as follows.

Variable Name	**Description**
Age	Students' age from wave 1 survey
Gender	0: male; 1: female
Siblings	Number of siblings
Hukou status	0: urban; 1: rural
Academic performance (Chinese)	Original standardized Chinese test score at grade 7
Academic performance (Math)	Original standardized Math test score at grade 7
Academic performance (English)	Original standardized English test score at grade 7
Physical health status	Self-rated health status. 1: very poor; 2: not very good; 3: moderate; 4: good; 5: very good
Boarding	1 if live in school at night from Monday to Thursday; 0 otherwise
Living with parents	1 if live with both biological father and mother; 0 otherwise
Intimacy with parents	1: not close; 2: not too close nor too far; 3: very close

Family SES	Use parental survey. 1: very poor; 2: somewhat poor; 3: moderate; 4: somewhat rich; 5: very rich
Socioeconomic conditions in the neighborhood	Use parental survey. 1: very low; 2: somewhat low; 3: average; 4: somewhat high; 5: very high
Parental years of education	The higher one between father and mother, range from 0 to 18
Parental expectation about children education years	Range from 8 to 22. The lowest is 8 (grade 8). "Do not care" treated as missing.
Physical cultural capital	An IRT GRM measure based on the following questions: (1) books at home: range 1–5 ("very few" to "a great number"); (2) whether the student has own writing desk: 0=no, 1=yes; (3) computer and access to internet: 0=neither, 1=computer only, 2=computer+internet.
Parental occupation (father)	1: government official/cadre; 2: cadre/official/administrator of public institutions, enterprises/corporations; 3: scientist, engineer, university professor or other professionals; 4: doctor, lawyer, teacher; 5: accountant, nurse, programmer, technical staff; 6: ordinary staff (secretary, bank clerk, librarian); 7: staff in service/business (salesperson, cook, barber, cosmetologist); 8: technical worker (driver, electrician, plumber, mechanist); 9: ordinary worker (porter, production line worker); 10: farmer, herdsman, fisherman; 11: elementary worker (cleaner, guard, housekeeper, sanitation worker); 12: self-employed; 13: unemployed/laid-off; 14: other.
Parental supervision	Total score of 6 questions: care/strictness about (1) homework/exams, (2) behavior, (3) friends, (4) dress style, (5) time online, (6) TV. 1=don't care; 2=care, not strict; 3=very strict.
Parental involvement	Total score of 4 questions: How often does your father discuss the following with you? (1) Things happened at school (2) The relationship between you and your friends (3) The relationship between you and your teachers (4) Your worries and troubles. Responses: 1=never; 2=sometimes; 3=often.
Semester	0: 2015 spring; 1: 2014 fall

Class characteristics	IRT measure of 9 questions: (1) teachers frequently ask me questions; (2) teachers frequently praise me (3) teachers frequently criticize me (reverse-score) (4) Most of my classmates are nice to me (5) My class is in good atmosphere. (6) I often take part in school/class activities. (7) I feel close to people in this school. (8) I feel bored in this school. (reverse-score) (9) I hope that I could transfer to another school. (reverse-score).. Responses 1–4 (strongly disagree → strongly agree).
Ethnicity	1: Han; 0: others

B.2 Estimated Impact of Cram School Participation on Mental Health for Full CEPS Sample

In the main text, we only focus on urban households. In this section, we examine the estimated impact of cram school participation on mental health capital for the full CEPS sample.

Table A8: Estimated IV Coefficients of Cram School Participation on Mental Health (Full Sample)

	(i) OLS	(ii) OLS	(iii) IV	(iv) IV
Cram School Participation	-0.023	-0.026	-0.300**	-0.249**
Lagged Variables	No	Yes	No	Yes
School FE	Yes	Yes	No	No
District FE	No	No	Yes	Yes
First-stage F stats			233.6	237.4
No. Observations	7378	7378	7378	7378

Note: $***p < 0.01, **p < 0.05, *p < 0.1$. Lagged variables include academic performance and mental health state in previous period.

Table A8 presents the estimated effects of cram school participation on children's mental health for sample which includes both rural and urban households. Column (i) reports the simplest OLS specification from Equation (A1). Although the estimated coefficient β_m is negative, it is not statistically significant. Column (ii) adds lagged variables—including prior academic performance and mental health—which may influence both cram school participation and current mental well-being. While the magnitude of β_m increases slightly, it remains statistically insignificant.

Columns (iii) and (iv) implement the IV specification. With IV estimation, the effect of cram school participation on mental health becomes substantially larger and statistically significant at the 5% level. Omitting lagged variables, as in Column (iii), yields an even larger estimate of β_m, but this may overstate the true effect.

For interpretation, the fully specified model in Column (iv) indicates that students who participate in extracurricular tutoring have mental health scores that are 0.27 standard deviations higher than those who do not participate. Specifically, the estimated coefficient β_m is 0.25, and the standard deviation of the IRT-based depressive mental health score is 0.94. This result suggests that participating in cram schools has a substantial negative effect on children's mental health, which informs the identification of $\phi_{m,c}$ in Equation (9) when we bring the model to the data.

B.3 Distribution of Mental Health Shocks

We assume that shocks to mental health capital—for both adults and children—follow a normal distribution. This assumption is based on data showing that the change in mental health scores over a four-year period is approximately normally distributed, as shown in Figure A1.

B.4 Child Mental Health by SES

We find that children from more advantaged family backgrounds (higher SES) have, on average, better mental health, as shown in **??** in the main text and Table A9 here. This aligns with findings in the health and psychiatry literature (Husain et al., 2023; Wilson et al., 2025) and with evidence from Cunha, Heckman, and Schennach (2010) that children from more advantaged family backgrounds have better non-cognitive skills.

Table A9: Child Mental Health and Tutoring Participation by SES

SES	Tutoring Participation	Expenses (School)	Expenses (Tutoring)	Time (School)	Time (Tutoring)	Cognitive Skills	Mental Health
Low	30.8%	1402	258	19.9	4.9	67.0	-0.17
Middle	49.1%	1450	1145	18.8	7.9	70.7	0.04
High	67.5%	2046	2304	20.5	10.0	72.6	0.18

Data Source: CEPS. *Notes:* Both mental health scores and cognitive test scores are standardized, with mean close to zero, and standard deviations equal to 0.852 and 0.882 respectively.

B.5 Mental Health over Life Cycle

In the main text, we plot age profiles of the mental health index and the physical health index for childhood and adolescence. The mental health index is constructed as described in Section 2.1.2. The physical health index is based on self-reported health status in the CFPS, with "healthy" coded as 1, "fair" as 0.5, and "unhealthy" as 0. For both indices, higher values indicate better health. Here, we extend the age profiles through age 62 (see Figure A2). As shown in the

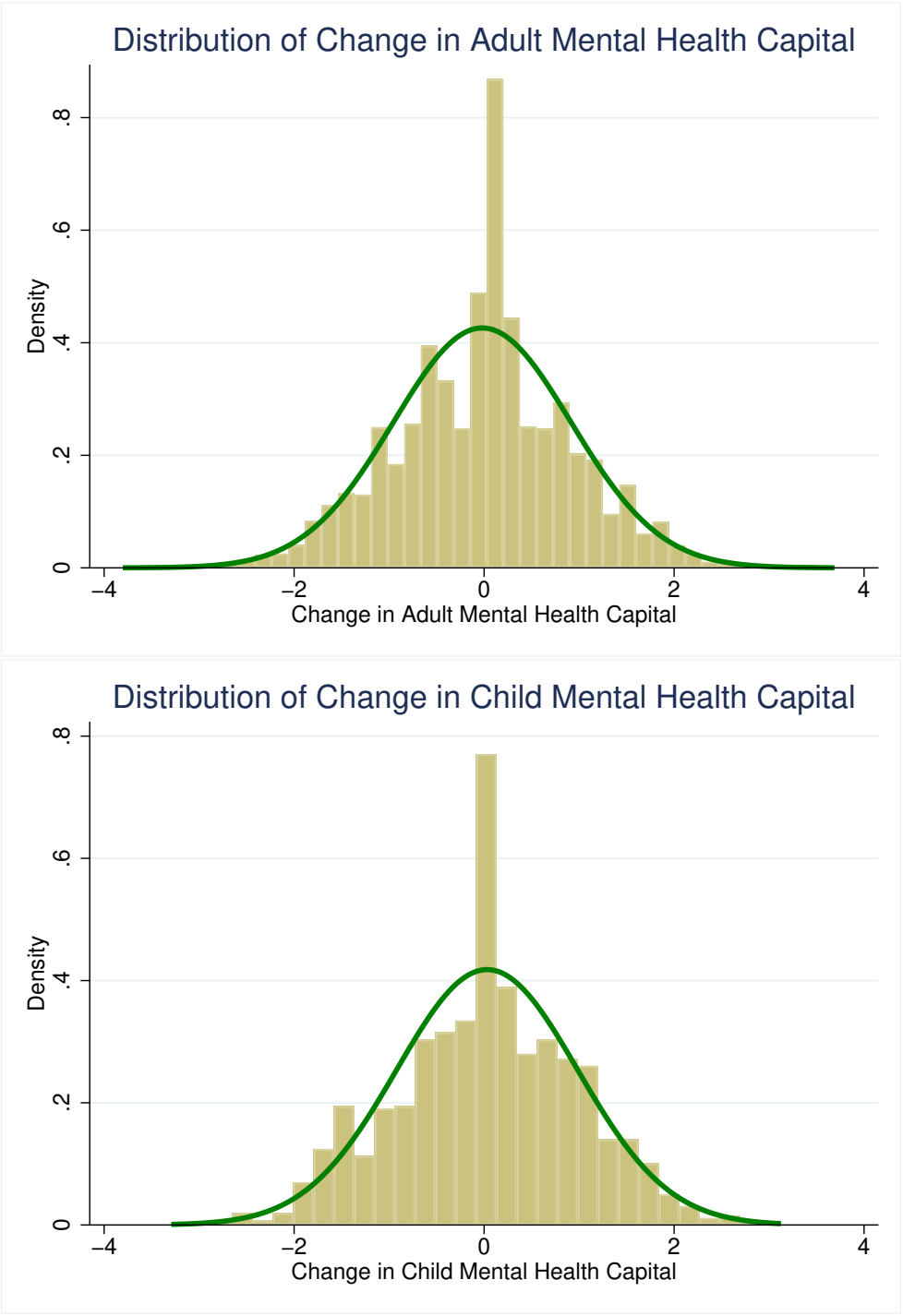


Figure A1: Distribution of Mental Health Shocks

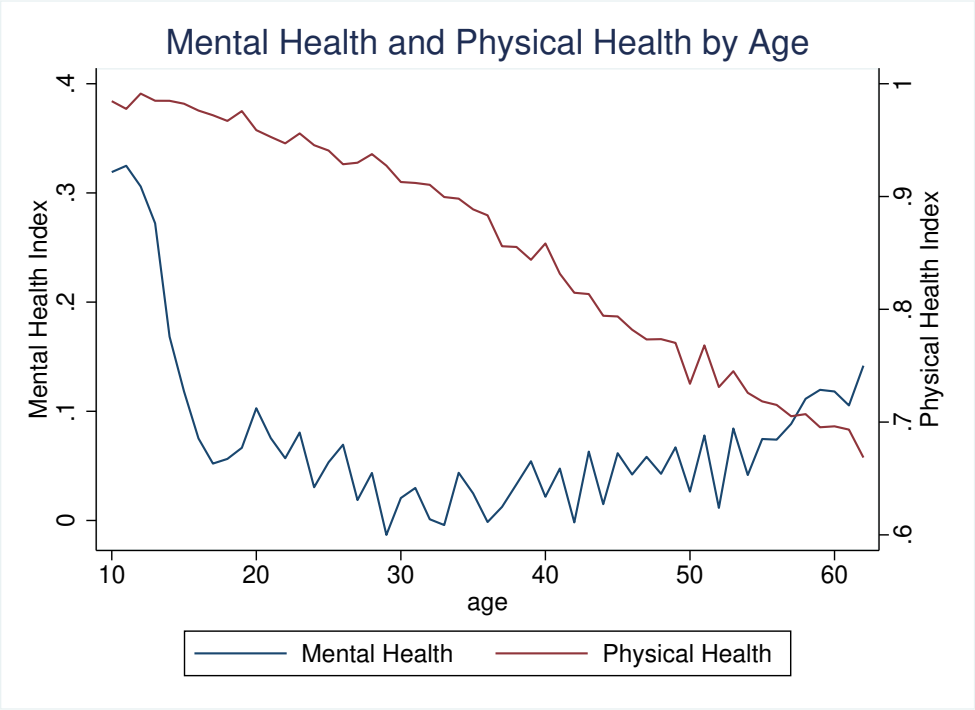


Figure A2: Mental Health over Life Cycle

figure, the physical health index declines with age, consistent with stylized facts in the health economics literature (e.g., Hosseini, Kopecky, and Zhao (2022)). By contrast, the mental health index declines sharply during adolescence—when students start to face academic pressure and high-stakes examinations—stabilizes in midlife, and increases slightly after age 50. This U-shaped age profile is consistent with findings in the psychiatry literature (e.g., Gallo, Anthony, and Muthen (1994); Husain et al. (2005); Wagner et al. (2020)).

By Education Next, we plot age profiles of the mental and physical health indices by educational attainment (non-college vs. college-educated). Consistent with existing literature (e.g., De Nardi, Pashchenko, and Porapakkarm (2025); Hosseini, Kopecky, and Zhao (2022)), college-educated individuals exhibit better physical health than their non-college-educated counterparts. This health gap widens with age. For mental health, the educational disparity is less pronounced across most of the life course. However, among older adults (e.g., age 56-62), the college-educated also show a mental health advantage, a finding consistent with the broader health literature (e.g., Zhao et al. (2024)).

B.6 Investment in Children

Time Input In this paper, we abstract from early childhood stages and focus exclusively on children aged 6-18, a critical period when they begin attending school, preparing for college entrance examinations, and receiving significant parental investments, especially attending " schools" to improve exam performance. Hence, parental investments in this paper refer only to monetary investments. To show that time investments from parents in China are indeed not the

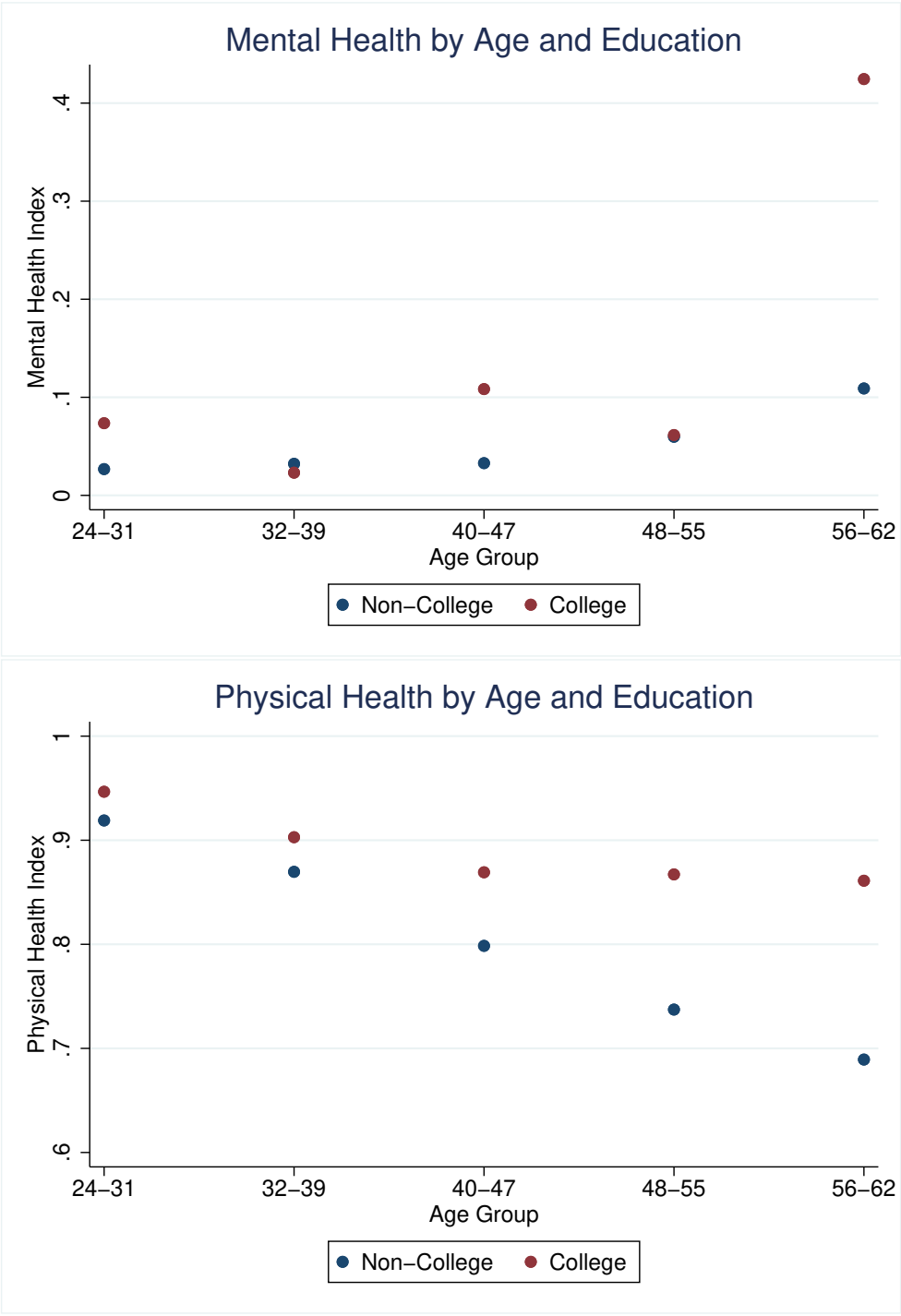


Figure A3: Mental Health over Life Cycle by Education

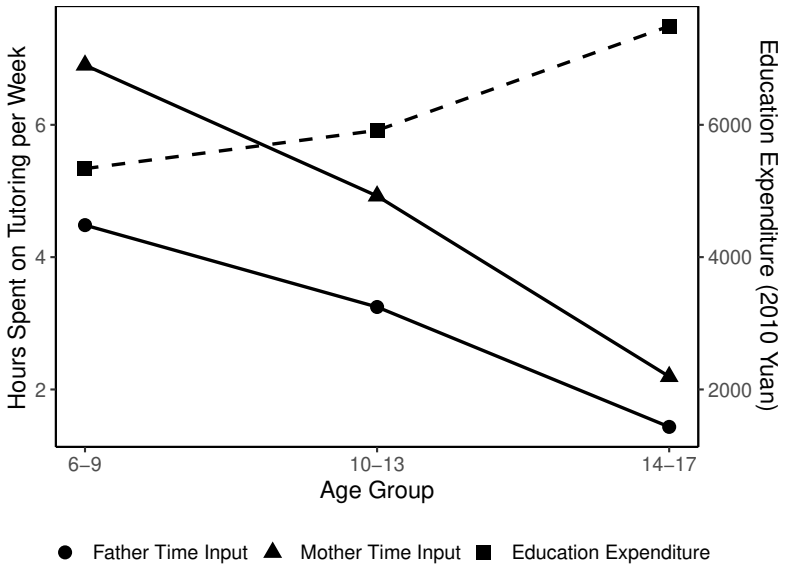


Figure A4: Parental Time Input and Monetary Input by Child Age

primary issue, we also examine parental time input patterns in the CFPS.

Unlike the CDS of the PSID, where each child submits a detailed 24-hour time diary for one weekday and one weekend day, the CFPS does not have such detailed time diaries. The only detailed data regarding time input available is "weekly hours father/mother spend on tutoring," and this data is available only from the 2014 wave onward. We merge this data into our main sample and report the results on both time input and money input with respect to child age groups: age 6-9, 10-13, and 14-17, consistent with our model setup, in Figure A4. As we can see from the figure, the time input in terms of hours per week that parents spend on tutoring declines as children age, while education expenditures simultaneously surge. Even at ages 6-9, the peak age, the tutoring hours are only 4.5 hours for fathers and 6.9 hours for mothers on average, which are much lower than the active time reported by Lee and Seshadri (2019) using PSID-CDS for American parents.

Parental Monetary Investment by Cognitive Skill Quintile and Mental Health In the main text, we examine how parental monetary investment jointly depends on children's cognitive skills and mental health in urban China. Here, we present several versions of the pattern exhibited in **??**, including adding rural areas to the sample (Figure A5) and controlling for SES (Figure A6) and parenting style, and we find the patterns are robust.

B.7 Private Education Expenditure

We categorize private education expenditure into three areas: in-school expenses, extracurricular activities and private tutoring, and other education-related costs. In-school expenses mainly include after-school programs, tuition, textbooks, and various costs incurred by the school, such as food, boarding, school-selection fees, uniforms, physical exams, and school medical insurance.[31]

[31]Note that not all these fees apply to every school and every child.

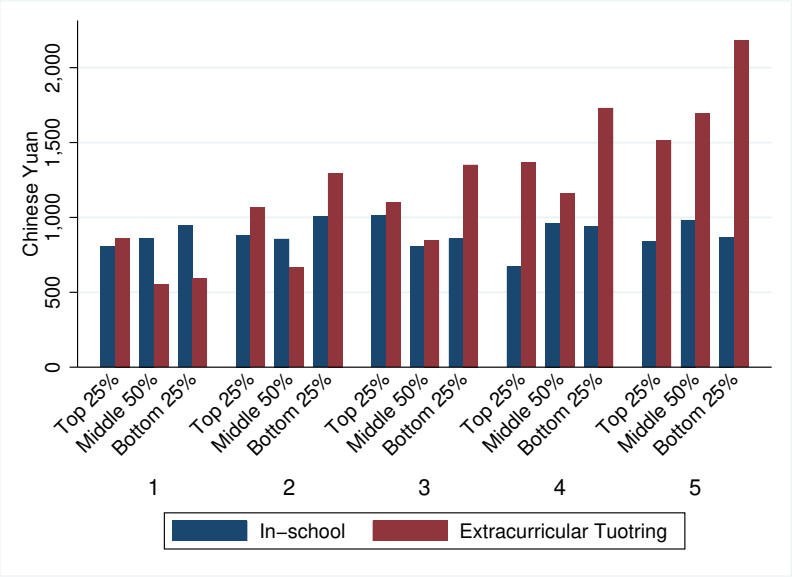


Figure A5: Parental Monetary Investment by Cognitive Skill Quintile and Mental Health for both Rural and Urban Areas

Note: The data are from a nationally representative sample of seventh-grade children in China (including both urban and rural) from the 2014 wave of CEPS. "Top 25%, Middle 50%, Bottom 25%" denote segments of the distribution of the CEPS measure of children's mental health (top quartile, interquartile range, and bottom quartile, respectively). Higher scores indicate better mental health. The values 1-5 denote quintiles of cognitive skills for the same group of children (1 = lowest, 5 = highest).

China has implemented a 9-year compulsory education system, providing free and universal education to children aged 6 to 15. School fees at the elementary and middle school levels do not include tuition fees but primarily cover after-school programs provided by the schools. These programs are designed to help students complete their homework and offer additional tutoring to cater to the needs of students at different ability levels.

Expenditure on extracurricular activities and tutoring includes costs associated with for-profit after-school programs ("cram schools") aimed at improving performance in core subjects tested in the College Entrance Examination (e.g., Chinese, Math, English, Physics, History), as well as arts and sports classes.

Other education-related costs encompass transportation, software, stationery, and other miscellaneous items directly related to education.

Notably, in-school expenses account for a significant portion of total education expenditure (around 40-44% for elementary and middle school), with most of this going towards after-school programs that enhance academic performance. The high private cost of education is likely due to the demand and limited availability of prestigious schools, which can charge much more for additional after-school services. At the high school stage, the share of private education expenditure on in-school expenses increases, as high school tuition and textbooks are no longer free. After high school, the need for tutoring decreases significantly as students no longer prepare for the college entrance examination.

To sum up, for children aged 6-18, most private education expenditure, whether incurred in-school or out-of-school, is directed towards improving academic performance.

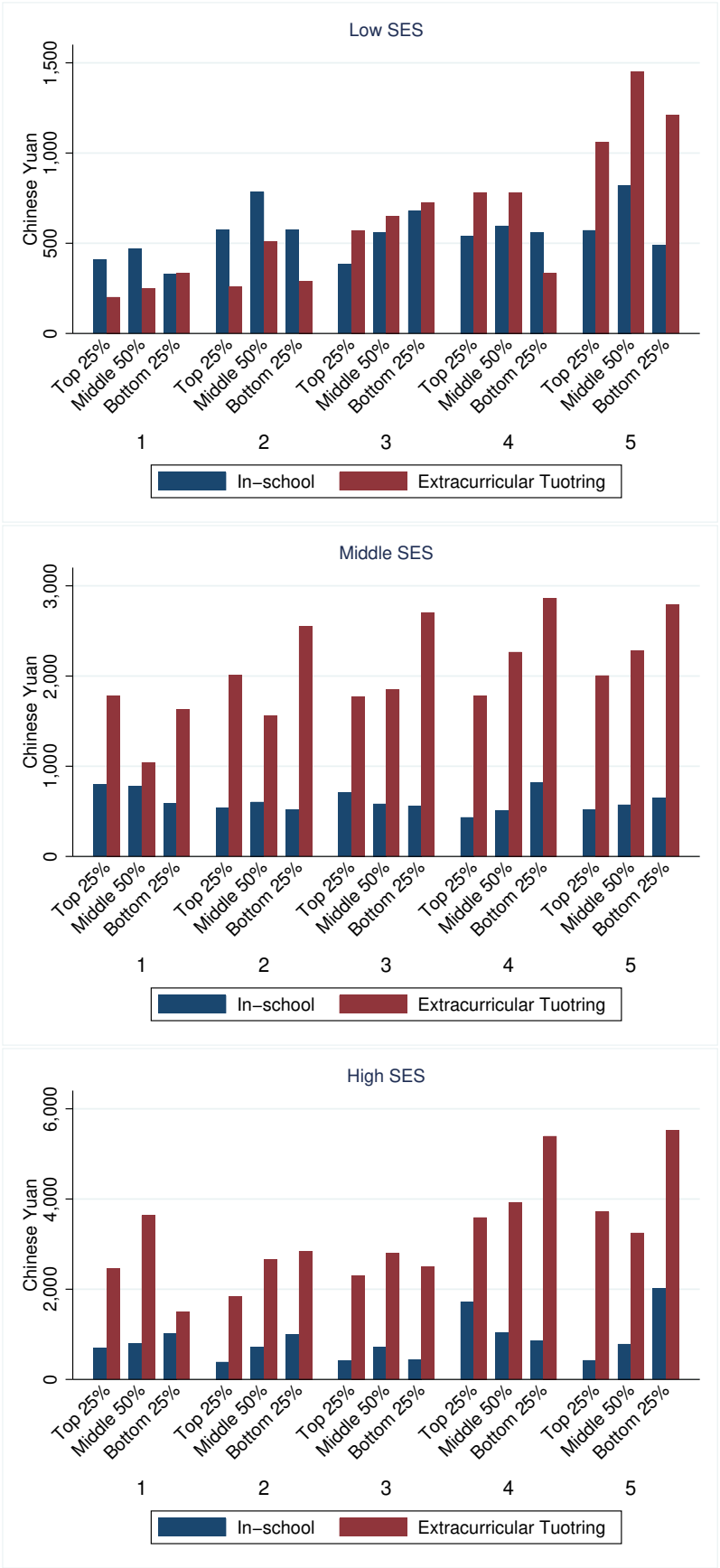


Figure A6: Parental Monetary Investment by Cognitive Skill Quintile and Mental Health, Controlling for SES

Table A10: Private Education Expenditures

By School Stage	Elementary School	Middle School	High School
Value (2010 Yuan)			
In-school expenses incl. after-school tutoring	1,987	2,379	5,179
Extra-curricular tutoring	1,958	2,422	1,892
Other expenses	974	634	859
Total	4,919	5,435	7,930
% of total private education expenditure			
In-school expenses incl. after-school tutoring	40%	44%	65%
Extra-curricular tutoring	40%	45%	24%
Other expenses	20%	12%	11%

Notes: The statistics are derived from samples taken from CFPS 2010, 2012, 2014, 2016, 2018, 2020, 2022. The samples are limited to urban households. Each household includes one child and at least one parent. A child is defined as an individual aged between 6 and 18 who is currently attending school. All results are computed at the individual child level and deflated using the 2010 CPI.

Tutoring Time vs. Tutoring Money As shown in Figure A7, we find a strong positive correlation between tutoring time and monetary expenditures. Specifically, low levels of parental spending on tutoring correspond to negligible amounts of time spent by children on these activities. The variables are defined as follows:

- **Tutoring time** refers to the average weekly hours a child spends attending private tutoring or cram school to improve their academic exam performance.
- **Extracurricular assignment time** is defined as the average weekly hours a child spends on assignments not issued by their in-school teachers.

B.8 Reduced-form Evidence on the Impact of Tutoring on Mental Health

B.8.1 Instrumental Variable Design

To estimate the effect of cram school attendance on a child's mental health well-being, we employ the following empirical design:

$$Y_{icsd,t} = \beta_m \cdot cramsch_{icsd,t} + X'_{icsd,t}\vartheta_0 + Z'_{csd,t}\vartheta_1 + D'_{sd,t}\vartheta_2 + \epsilon_{icsd,t} \tag{A1}$$

where Y_{isct} is an outcome measure of individual i in class c of school s in district d in year t for the IRT-based mental health measure. The variable $cramsch_{icsd,t}$ is a dummy variable, which denotes whether individual i attends cram school in year t. $X_{icsd,t}$ denotes a vector of controls on student characteristics including gender, age, cognitive ability, physical health, ethnicity, *hukou* status,

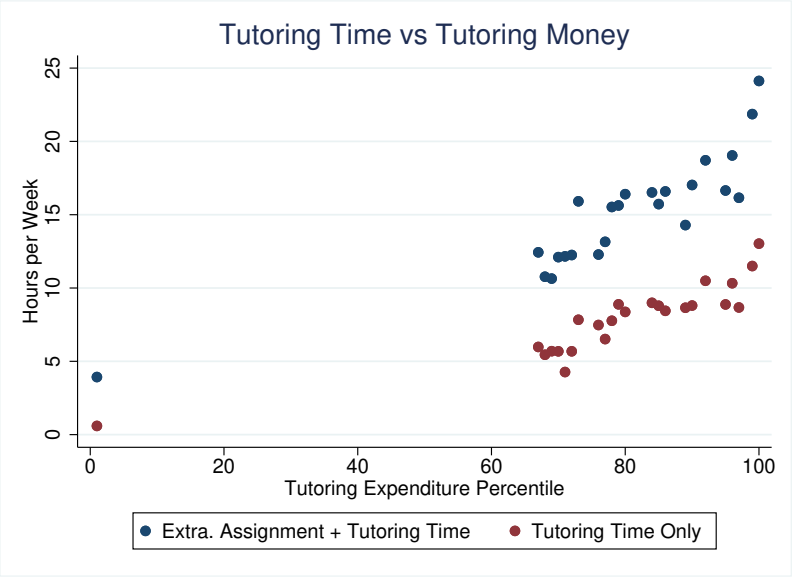


Figure A7: Tutoring Time vs. Tutoring Money

Note: Tutoring time refers to the average weekly hours a child spends attending private tutoring or cram school to improve their academic exam performance. Extracurricular assignment time is defined as the average weekly hours a child spends on assignments not issued by their in-school teachers.

parenting style, socioeconomic status (SES), parental education, parental occupation, parental expectation, and home environment. $Z_{csd,t}$ denotes a vector of class-level controls including peer effects, teacher quality, and class culture / atmosphere. $D_{sd,t}$ denotes a vector of school-level controls including average teacher quality of school s, and school culture / atmosphere. We provide more details on the description of all variables we use in Appendix B.1.

The goal is to obtain an unbiased and consistent estimate of β_m, which measures the effect of attending cram schools on child mental well-being. There are potential challenges to estimating Equation (A1) using OLS. First, cram school attendance may be endogenous due to reverse causality: a student's mental well-being in the previous period may influence her current decision to attend cram schools. Additionally, her academic performance in the previous period may affect her current decision to attend cram schools, which could, in turn, impact her current mental health. Second, there could be a correlation between a child's mental well-being and unobservable individual qualities or abilities. This bias can operate in either direction. Individuals with lower ability may exhibit poorer academic performance, which could lead to increased depressive symptoms from falling behind, as well as a higher likelihood of attending cram schools to catch up, leading to an upward bias in β_m. Conversely, it is also possible that high-ability students are more likely to come from high-income families, whose children may enjoy better mental well-being due to the potential positive effect of income on mental health (Kling, Liebman, and Katz, 2007; Evans and Garthwaite, 2014), and who are also more able to afford cram schools. This would result in a downward bias in β_m.

To address these identification challenges, we first introduce a lagged measure of child mental health $Y_{isc,t-1}$, as well as lagged variables of the student's academic performance in terms of standardized test scores in three core subjects—math, Chinese, and English—to address reverse

causality, as follows:

$$Y_{icsd,t} = \beta_m \cdot cramsch_{icsd,t} + \beta_l \cdot Y_{icsd,t-1} + X'_{icsd,t}\vartheta_0 + Z'_{csd,t}\vartheta_1 + D'_{sd,t}\vartheta_2 + \epsilon_{icsd,t} \tag{A2}$$

Then, we estimate the causal effect of attending cram schools on child mental health using a "leave-out local prevalence" district-school-level instrumental variable. The basic idea is to use the treatment rate among "nearby" units (local prevalence) as an instrument for an individual's treatment, but compute it leaving out the individual i (and often i's closest peers) to avoid mechanical correlation and peer-reflection bias. Our instrument is defined as the proportion of students participating in extracurricular tutoring among those who live in the same district (county) as student but attend different schools, denoted by $IV_{icsd,t}$. We label this case as "IV$_{school}$" in Table A11.

A valid instrument must satisfy two conditions: (1) relevance—correlation with cram school participation, and (2) exogeneity—no direct correlation with mental health. For the relevance assumption, this IV captures the district-level tutoring participation rate, reflecting the local equilibrium of the after-school tutoring market. On the demand side, the IV captures aggregate demand for tutoring among other middle school students in the district, potentially propagating through local social networks and facilitating information diffusion. On the supply side, it reflects the availability of cram schools, influencing individual access to tutoring. Thus, the IV is strongly associated with $cramsch_{icsd,t}$. However, cram school participation may be influenced by peers, which could directly affect individual mental health. For instance, if peers pursue competitive tutoring, non-participating students may experience increased peer pressure and negative emotions.

A key exogeneity concern is that peers' tutoring choices may directly affect an individual's mental health (e.g., competitive pressure causing stress among non-participants). To mitigate direct peer-effects and the Manski reflection problem, we construct the instrument by excluding same-school peers when computing the district participation rate. This leave-out design reduces mechanical correlation between the instrument and the individual's own peer environment and aims to capture broader, plausibly exogenous variation in district-level tutoring participation. While our "leave-out local prevalence" instrument exhibits a strong first stage, exogeneity remains the primary concern because district-level shocks (e.g., policy changes, economic conditions, stress climate) could influence both tutoring prevalence and mental health. To address this, we implement additional robustness checks using an alternative instrument: a finer leave-out at the class (school-by-grade-year) level within the district. We label this case as "IV$_{class}$" in Table A11.

Specifically, our empirical specification incorporating the IV is as follows:

$$\begin{aligned} \widehat{cramsch}_{icsd,t} &= \pi_m \cdot IV_{icsd,t} + \beta_{l0} \cdot Y_{icsd,t-1} + X'_{icsd,t}\varsigma_0 + Z'_{csd,t}\varsigma_1 + D'_{sd,t}\varsigma_2 + \nu_{icsd,t} \\ Y_{icsd,t} &= \beta_{m,2sls} \cdot \widehat{cramsch}_{icsd,t} + \beta_l \cdot Y_{icsd,t-1} + X'_{icsd,t}\vartheta_0 + Z'_{csd,t}\vartheta_1 + D'_{sd,t}\vartheta_2 + \epsilon_{icsd,t} \end{aligned} \tag{A3}$$

B.8.2 Estimated Impact of Cram School Participation on Mental Health

Table A11 presents the estimated effects of cram school participation on children's mental health. Column (i) reports the simplest OLS specification from Equation (A1). Although the estimated coefficient β_m is negative, it is not statistically significant. Column (ii) adds lagged variables—

Table A11: Estimated IV Coefficients of Cram School Participation on Mental Health

	(i) OLS	(ii) OLS	(iii) IV_{school}	(iv) IV_{school}	(v) IV_{class}
Cram School Participation	-0.053	-0.051	-0.391**	-0.316**	-0.313**
Lagged Variables	No	Yes	No	Yes	Yes
School FE	Yes	Yes	No	No	No
District FE	No	No	Yes	Yes	Yes
First-stage F stats			126.8	131.6	248.3
No. Observations	3392	3392	3392	3392	3392

Note: $***p < 0.01, **p < 0.05, *p < 0.1$. Lagged variables include academic performance and mental health state in previous period.

including prior academic performance and mental health—which may influence both cram school participation and current mental well-being. There is very little change in the magnitude of β_m, which still remains statistically insignificant.

Columns (iii) and (iv) implement the IV specification. With IV estimation, the effect of cram school participation on mental health becomes substantially larger and statistically significant at the 5% level. Omitting lagged variables, as in Column (iii), yields an even larger estimate of β_m, but this may overstate the true effect.

For interpretation, the fully specified model in Column (iv) indicates that students who participate in extracurricular tutoring have mental health capital scores that are 0.33 standard deviations lower than those who do not participate. Specifically, the estimated coefficient β_m is -0.32, and the standard deviation of the IRT-based depressive mental health score is 0.97. This result suggests that participating in cram schools has a substantial negative effect on children's mental health, which informs the identification of $\phi_{m,c}$ in Equation (9) when we bring the model to the data.

In Column (v), we verify robustness to tighter leave-out definitions of the instrument. Using the district-level leave-out share excluding the student's school yields a 2SLS estimate of 0.316, while excluding the student's class gives 0.313; the coefficients are statistically indistinguishable. First-stage relevance remains strong in both cases. This similarity suggests that our results are not driven by very local peer mechanical correlation and that the identified effect is stable across reasonable instrument constructions.

B.9 Public Education Expenditures

In China, the majority of children attend public schools. In 2022, public schools accounted for 89.4% of primary and junior high school students and 82.7% of senior high school students.

We utilize the China Education Finance Statistical Yearbook (CEFSY) to gather information on public investment in both pre-college and college education, and present the results in Table A12.

The currency unit used in the table is Chinese *yuan*.

Table A12: Public Education Expenditures

Panel (A)	Child age	Model period	Public expenses	$g/\text{avg}(i)$
Primary school	6-11	5, 6	8,463	1.68
Middle school	12-14	6, 7	11,724	2.49
High school	15-17	7	11,482	1.98
Four-year college	18-21	8	17,959	2.99
Panel (B)	Child age	Model period	Average earnings	$g/\text{avg}(y)$
Stage 1	6-9	5	47,442	0.17
Stage 2	10-13	6	42,459	0.24
Stage 3	14-17	7	40,153	0.31

B.10 Inter-vivos Transfers

While a number of papers argue that inter-vivos transfers play an important role in shaping college education outcomes and intergenerational mobility for U.S. households (e.g., Abbott, Gallipoli, Meghir, and Violante (2020), Daruich and Kozlowski (2020)), our model abstracts from inter-vivos transfers. Unlike the U.S. data, where a significant and common amount of monetary transfer from parents to children around age 16 is observed, only a very small portion of parents transfer money to their children during that age range in our CFPS sample for Chinese households. Even at the peak age (around 15), less than 5% of parents transfer money to their children. Moreover, when money is transferred, the average amount is quite small, around 900 Yuan (or 120 USD) at the peak age. This discrepancy arises because the large amount of inter-vivos transfers that U.S. parents provide to their children is primarily for paying college tuition and fees. In contrast, tuition and fees in China are much lower, and most Chinese parents pay these costs directly for their children. Chinese parents do transfer significant amounts of money to their children, but this occurs at a much later age and is primarily intended to enhance their son's attractiveness for marriage (Wei and Zhang, 2011). The money is usually used for paying the bride price or purchasing real estate property for marriage.

B.11 Age Profile of Child Cognitive Ability

For the child survey, we can only observe children's cognitive ability scores starting at age 10 and must infer their cognitive ability score distribution before age 10 through extrapolation. We use the data points of the average cognitive ability test scores by age, ranging from age 10 to age 21, to carry out the extrapolation, employing the Piecewise Cubic Hermite Interpolating Polynomial (PCHIP) method. In practice, we implement this method by directly applying the "pchip" command in MATLAB. We report the results in Figure A8.

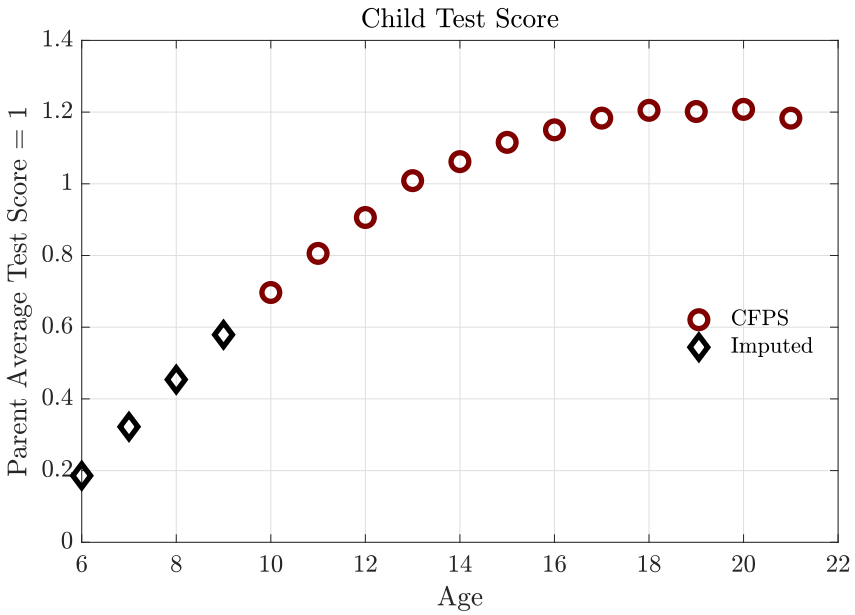


Figure A8: Children's Average Cognitive Ability Test Scores by Age

Note: The figure exhibits the data patterns on children's cognitive ability test scores by age, with the portion before age 10 imputed through extrapolation. Note that since one model period corresponds to 4 years in the data, we calculate the average cognitive ability scores of children aged 6-9, 10-13, 14-17, and 18-21, respectively, to compute the empirical moments. The average cognitive ability scores for the parent generation is normalized to be one.

B.12 Age Profile of Earnings and Initial Distribution of Parents

We use the the China Health and Nutrition Survey (CHNS) to estimate the age profile of earnings by education $A_j^s, s \in \{col, ncol\}$, and report the results in Figure A9.

We assume the initial distribution of parental pre-college human capital (h_p^s) follows an education-specific Gamma distribution:

$$h_p^s \sim \Gamma_s(M^s, Q^s), \quad s \in \{col, ncol\}\,.$$

We calculate the average and standard deviation of parental cognitive skills using the CFPS data, establishing the parameters that shape the Gamma distribution. Note that we estimate the initial distribution of parental skills outside the model, without treating it as an endogenous outcome of the stationary equilibrium. We plot the distributions of parents by education level in Figure A10.

B.13 Admission Probability Function

We estimate the pre-college human-capital-based college admission function using the adult survey of the CFPS, where we can observe the measure of pre-college human capital (i.e., cognitive ability scores) in both their early career and education outcomes. We restrict the age range to 24-35, as this allows us to observe both cognitive ability scores and education outcomes, with ages close enough to the college entrance examination so that cognitive ability scores primarily reflect pre-college human capital. Here are more details on the procedure we follow to estimate

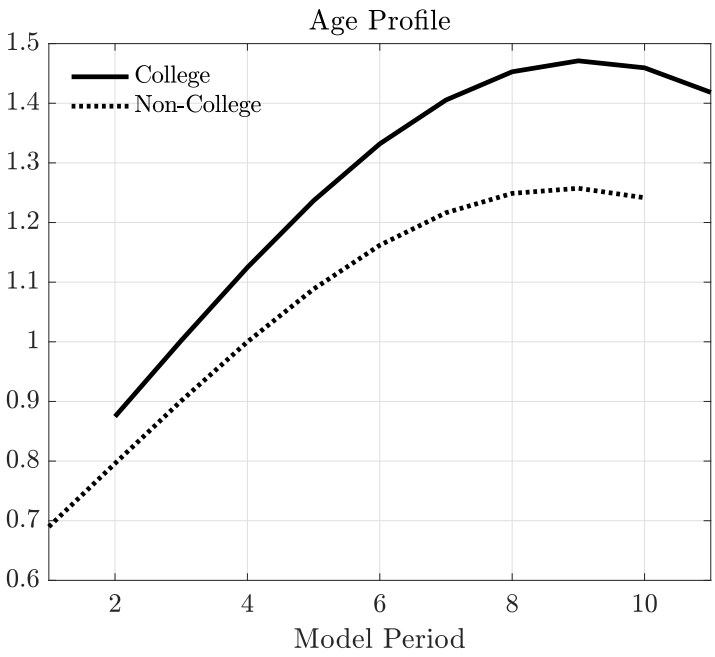


Figure A9: Age Profile of Earnings

Note: The x-axis is for model periods j. The value reported on y-axis is $exp(A_j^s), s \in \{col, ncol\}$. Note that the average household income is normalized to be one.

and feed the college admission function into the model:

(I) Computing empirical moments needed from data:

(i) Compute the mean of the cognitive ability scores for all individuals in our sample and take the logarithm, denoted by $\bar{h}^D$.

(ii) Sort individuals in the sample into ten bins according to their decile rank of cognitive ability scores.

(iii) Compute the college admission probability in terms of the share of people with a college degree for each bin.[32]

(iv) Through steps (i)-(iii), we have obtained patterns on how college admission probability depends on cognitive ability decile in data, as shown in Figure A11. This is the empirical counterpart of our admission probability function.

(II) Preparing for taking the data into the model:

(i) In order to map the empirical patterns of ability-based admission probability into the model, we need to find a way to establish a connection between the data and the model.

[32]In other words, we do not differentiate between college admission and college completion, as the college completion rate in China is very high, often above 95% for each cohort.

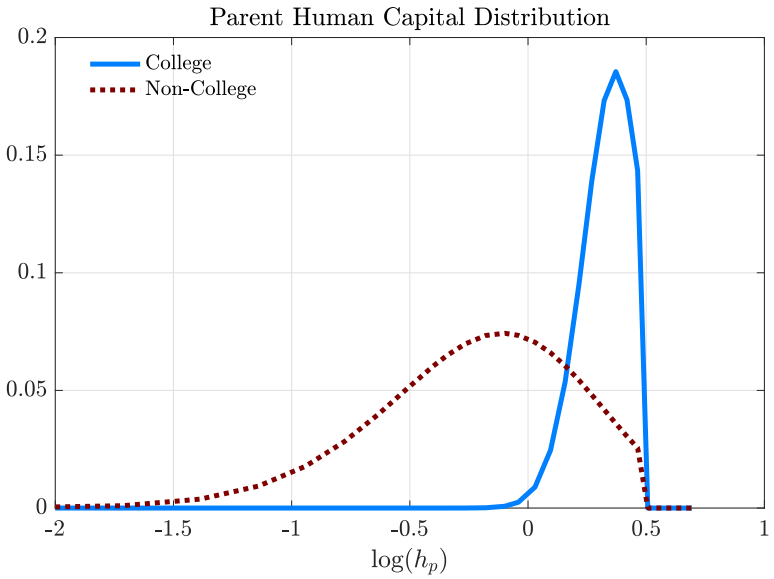


Figure A10: Initial Distribution of Parents

(ii) As a preparatory step, we first compute the average cognitive ability scores for each bin in the data and take the logarithm, denoted by $\hat{h}_i^D$ for $i = 1, 2, ..., 10$. Then we compute the distance between the mean of each bin and the mean of the whole population in logarithms in the data: $h_i^D = \hat{h}_i^D - \bar{h}^D$.

(iii) The college admission probability $\mathrm{X}(h_i^D)$ for each h_i^D is already computed in step (iii) in Part (I).

(iv) By normalizing the cognitive ability decile in the data based on the steps above, we can easily compute the counterpart in our model.

(III) Mapping the admission probability function estimated from data into the model:

(i) Now, we can assign each child human capital grid in the model an admission probability from data through piecewise linear interpolation:

$$\mathrm{X}(\widetilde{h}) = \mathrm{X}(h_i^D) + \frac{\widetilde{h} - h_i^D}{h_{i+1}^D - h_i^D}\left(\mathrm{X}(h_{i+1}^D) - \mathrm{X}(h_i^D)\right) \quad \text{for } h_i^D < \widetilde{h} < h_{i+1}^D.$$

where $\widetilde{h}$ denotes the log of a child human capital grid h in computing the model, and $\chi(h) = \mathrm{X}(\widetilde{h})$.

(ii) Finally, we choose the college admission function shifter ζ to ensure the model-generated college capacity matches its targeted empirical counterpart Z. In other words, Equation (**??**) must hold.[33] The same interpolation method as used above applies.

[33]Note that college capacity decreases as ζ increases. ζ is determined within the model because it endogenously affects the distribution of child human capital.

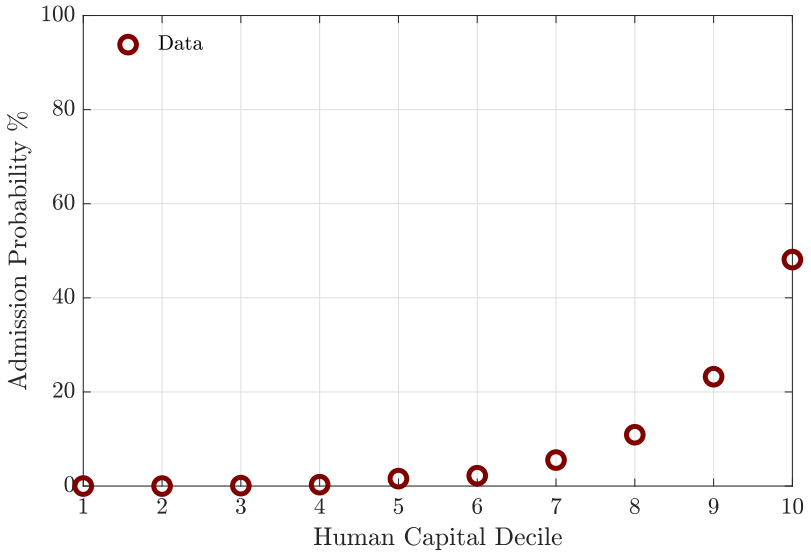


Figure A11: How College Admission Probability Depends on Cognitive Ability Decile

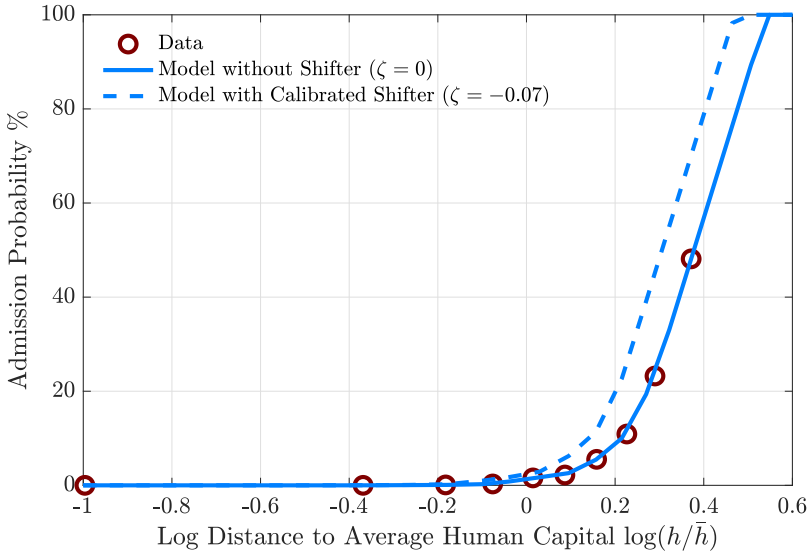


Figure A12: How Shifter ζ in Admission Function Works

As a cautious remark, ideally, with a sufficiently long panel dataset and a large number of observations, we could track individuals' childhood development stages, adult education, and labor market outcomes. We would then use their cognitive ability scores at age 18 and follow their education outcomes to estimate the admission function. However, for the cohort of children aged 6-21 in 2010-2018 in our CFPS sample, there are not enough observations to observe their college education outcomes and thus estimate the admission function. Therefore, we also rely on older cohorts in the CFPS, where we can observe cognitive ability in early adulthood and their education outcomes. However, this is not an issue for our analysis because we only care about the shape of the admission function $\chi(\cdot, \zeta)$; we can always adjust the shifter parameter ζ to match the college capacity observed in the data.

Furthermore, as shown in **??**(B), regardless of the decade when college entrance exams were taken (pre-1990s, 1990s, or 2000s), college admission probability consistently depends on pre-college human capital ranks. Individuals in the top quintiles have a much higher chance of attending college. In other words, the shape of the admission function is always convex with respect to child ability rank, regardless of the college capacity.

C Welfare Measure

Welfare by generation We define welfare for individuals in generation q based on ex ante lifetime utility of economically active individuals (i.e., after the education level is determined and just enter the labor market). That is,

$$W_q = (1 - \Omega_{col,q}) \int \int v_1(ncol, a, h, z) d\mu_z d\Gamma_{ncol,q} + \Omega_{col,q} \int \int v_2(col, a, h, z) d\mu_z d\Gamma_{col,q} \quad \text{(A4)}$$

where a is the initial asset holdings (equal to zero), $\Gamma_{s,q}, s \in \{col, ncol\}$ are (endogenous) distributions of parental human capital by education level s in generation q, which are obtained through $\Gamma_{s,q-1}, s \in \{col, ncol\}$, together with a transition function generated by decision rules on parental investment in generation $(q-1)$ to form child pre-college human capital h. The child become a parent in generation q with human capital $h_{p,q} = h$. μ_z is distribution of z, $\Omega_{col,q}$ is the population weight of college labor in the benchmark economy, and v_1 and v_2 are defined in Problem (4). The integrals are taken over the state space of human capital h and possible types z.

Social Welfare Function We finally define welfare in the long run based on ex ante lifetime utility of economically newborn individuals (i.e., after the education level is determined and just enter the labor market) under the veil of ignorance. That is,

$$W_\infty = (1 - \Omega_{col,\infty}) \int \int v_1(ncol, a, h, z) d\mu_z d\Gamma_{ncol,\infty} + \Omega_{col,\infty} \int \int v_2(col, a, h, z) d\mu_z d\Gamma_{col,\infty} \quad \text{(A5)}$$

where a is the initial asset holdings (equal to zero), $\Gamma_{s,\infty}, s \in \{col, ncol\}$ are stationary distributions of human capital by education level s, μ_z is distribution of z, $\Omega_{col,\infty}$ is the population weight of college labor in the benchmark economy, and v_1 and v_2 are defined in Problem (4). The integrals are taken over the state space of human capital h and possible types z.

Optimal Tax Problem We now investigate the effects of a private investment tax in our model. We impose a linear tax on private tutoring expenditures, denoted by τ_i, paired with a linear subsidy for pre-college public expenditures, denoted by τ_g, funded solely by the private tutoring tax revenue. That is, we add one term to the household budget constraints:

$$(1+\tau_c)\,c + a' + i + \tau_i \mathbb{1}_{i>\bar{i}}(i-\bar{i}) = (1+r)a + w^s y(h_p, A_j^s, z), \tag{A6}$$

The next period's pre-college human capital of children becomes

$$h' = \theta_j \left[\alpha_j h^\gamma + (1-\alpha_j)\, I_j^\gamma\right]^{\frac{1}{\gamma}}, I_j = \left[\eta i^\mu + (1-\eta)\left((1+\tau_g)g_j\right)^\mu\right]^{\frac{1}{\mu}}, \tag{A7}$$

where the linear subsidy imposed on pre-college public investment τ_g is determined by:

$$\sum_{j=5}^{7} \int_{X_j} \tau_g g_j d\Psi_j = \sum_{j=5}^{7} \int_{X_j} \tau_i \mathbb{1}_{i>\bar{i}}(i-\bar{i}) d\Psi_j. \tag{A8}$$

We choose a tax rate to maximize the long-run ex ante lifetime utility of newly economically active individuals under the veil of ignorance, as defined in Equation (A5), while satisfying the financing constraint (Equation (A8)), the government budget constraint (Equation (**??**)), and the college capacity constraint (Equation (**??**)).

D Full Definition of the Stationary Equilibrium

Let $x_j \in X_j$ be the age-specific state vector of an individual of age j, as defined by the recursive representation of the individual household's problems in Section 3. Let the Borel sigma-algebras defined over those state spaces be $\Psi = \{\Psi_j, \Psi_h\}$. A stationary recursive competitive equilibrium for this economy is a collection of: (i) private households' decision rules for consumption, asset holdings, and parental investment $\{c_j^{\delta_p}(x_j), a_j^{\delta_p\prime}(x_j), i_j^{\delta_p}(x_j)\}$, and the associated value functions $\{v_j^{\delta_p}(x_j)\}$ for both informed and uninformed parents $\delta_p \in \{I, U\}$; (ii) a college admission shifter set by the government ζ and consumption tax τ_c; (iii) aggregate capital and labor inputs of the representative firm $\{H_{col}, H_{ncol}, K\}$; (iv) prices $\{w^{col}, w^{ncol}, r\}$; and (v) a vector of measures of Ψ such that:

1. Given prices, decision rules solve individual's problems (4), (10), (11), (12), and (A13);
2. Given prices, aggregate capital and labor inputs solve the representative firm's problem (3);
3. The asset market clears:

$$K = \sum_{j=1}^{15} \int_{X_j} a_j'(x_j)\, d\Psi_j. \tag{A9}$$

4. Labor markets clear for both education levels:

$$H_{col} = \sum_{j=2}^{11} \int_{X_j} y_j^{col}(x_j)\, d\Psi_j, H_{ncol} = \sum_{j=1}^{10} \int_{X_j} y_j^{ncol}(x_j)\, d\Psi_j \tag{A10}$$

5. The college capacity constraint is satisfied:

$$\frac{\int_{X_8} \int_{H_h} \chi(h,\zeta)\, \mathbb{1}_{v_8(x_8,col)>v_8(x_8,ncol)} d\Psi_h d\Psi_8}{\int_{X_8} \Psi_8} = \xi \tag{A11}$$

where H_h denotes the state space for pre-college human capital, Ψ_h denotes the distribution of the pre-college human capital of children at age 18 (corresponding to the beginning of parent period $j = 8$), and ξ denotes the fixed college capacity.

6. The government budget is balanced every period as follows:

$$\varsigma_{college} \int_{X_8} \int_{H_h} \chi(h,\zeta)\, \mathbb{1}_{v_8(x_8,col)>v_8(x_8,ncol)} d\Psi_h d\Psi_8 + \sum_{j=5}^{7} \int_{X_j} g_j d\Psi_j = \sum_{j=1}^{15} \int_{X_j} \tau_c c_j(x_j)\, d\Psi_j \tag{A12}$$

7. Individual and aggregate behaviors are consistent: measure Ψ is a fixed point of $\Psi(X) = Q(X, \Psi)$ where $Q(X, \cdot)$ is a transition function generated by decision rules and exogenous laws of motion, and X is the generic subset of the Borel-sigma algebra defined over the state space.

E Additional Model Setup and Quantitative Analysis Results

E.1 Recursive Problem at Retirement Stage

At the beginning of periods $j = 11$ or $j = 12$ (age 58 for adults without college degree or age 62 for college-educated adults), the agent retires with the source of income coming from savings, and $j = 15$ being the end of the life cycle. For simplicity, we abstract our set-up from social security tax and hence social security benefits provided by the government. The recursive problem is:

$$v_j(a) = \max_{c,a'} \left\{ \log(c) + \beta v_{j+1}(a') \right\} \tag{A13}$$

subject to

$$(1 + \tau_c)\, c + a' = (1 + r)a,\ c, a' \geq 0$$