

Liang, Wenquan; Xue, Sen

Working Paper

Pandemics and Intergenerational Mobility of Education: Evidence from the 2003 Severe Acute Respiratory Syndrome (SARS) Epidemic in China

GLO Discussion Paper, No. 779

Provided in Cooperation with:

Global Labor Organization (GLO)

Suggested Citation: Liang, Wenquan; Xue, Sen (2021) : Pandemics and Intergenerational Mobility of Education: Evidence from the 2003 Severe Acute Respiratory Syndrome (SARS) Epidemic in China, GLO Discussion Paper, No. 779, Global Labor Organization (GLO), Essen

This Version is available at:

<https://hdl.handle.net/10419/229651>

Standard-Nutzungsbedingungen:

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

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

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

Terms of use:

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

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

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

Pandemics and Intergenerational Mobility of Education: Evidence from the 2003 Severe Acute Respiratory Syndrome (SARS) Epidemic in China*

Wenquan Liang**

College of Economics, Jinan University

Sen Xue***

Institute for Economic and Social Research, Jinan University

Abstract

This paper examines the impact of the SARS epidemic in 2003 on intergenerational mobility in China. Using large cross-city variation in SARS cases, our triple difference-in-differences estimates suggest that the SARS epidemic significantly increases the intergenerational transmission of education. Our results show that a one percent increase in the number of SARS cases leads to a 9.3 percent increase in the maternal intergeneration transmission coefficient. The effect of the SARS epidemic is stronger for admission to 4-year bachelor programmes and more concentrated in female students and students in large cities. This paper also investigates the potential mechanisms and finds that more highly educated mothers tend to be more engaged in children's studies during the epidemic period when teachers are absent. These results convey the warning message that pandemics may reduce intergenerational mobility of education.

Keywords: Pandemic, Epidemic, intergenerational mobility, SARS

JEL Classification: I24, I18, J62

* We are grateful to Sumit Agarwal, Tue Gørgens, Xin Meng, Chao Ma, Junsen Zhang and seminar participants at Beijing Normal University for their constructive comments. Wenquan Liang acknowledges funding by the National Natural Science Foundation of China (71803065). Sen Xue is also a GLO fellow. The authors' names are alphabetically ordered.

** Email: liangsuwenquan@163.com

***Corresponding author. Email: sen.xue@hotmail.com

1. Introduction

This paper investigates the effect of epidemics and pandemics on the intergenerational transmission and mobility of education. When negative shocks occur, people of low socioeconomic status (SES) are often disproportionately harmed. Relative to the rich, those in lower SES groups tend to have less capacity to cope with crises. This may lead to intergenerational impacts, placing the children of lower SES groups in disadvantageous positions in society. At the time of writing this paper, the COVID-19 pandemic has strongly hit many countries and is significantly affecting the global economy. This has ignited concerns about the long-run effect of COVID-19 on social inequality and mobility. This paper draws implications from the experience of the Severe Acute Respiratory Syndrome (SARS) epidemic in 2003 to answer questions about how epidemics/pandemics affect social mobility.

As a readily human-to-human transmissible disease, the SARS epidemic quickly spread in 26 countries, causing a loss of 25 billion US dollars including medical treatment costs and losses of work time and tourism to East Asian countries (World Bank, 2003). The country hit hardest by SARS was China. By 15 August 2003, there were 7,747 SARS cases and had been 829 deaths in China (Jia and Liu, 2004). China's social order was seriously jeopardised due to the SARS epidemic. To avoid human interaction, school time was substantially shortened. In many regions, schools and hospitals were even closed. In this situation, parental education became more important than usual in the human capital accumulation of children for at least two possible reasons. First, more highly educated parents tend to have more health knowledge (Currie and Moretti, 2003; Breierova and Duflo, 2004; Cutler and Lleras-Muney, 2008; see Currie (2009) for a review), which can prevent their children from being infected during an epidemic. Second, more highly educated parents tend to create better studying environment to mitigate the adverse impact of epidemics for their children. For example, when schools are closed or school time is reduced, more educated parents could play a more effective role in supervising children's self-study and/or find private tutors for children to compensate for the loss of school time. It is also possible that more highly educated parents could better relieve stress and anxiety of their children caused by the epidemic, and thereby their children could study more efficiently. Given these, the SARS epidemic may increase the intergenerational transmission of education (ITE hereafter) and thereby reduce intergenerational mobility.¹

In this paper, we empirically test whether the SARS epidemic indeed affects ITE in the mainland China in terms of children's college admission. As obtaining a college degree tends to lead to better jobs and higher income (Card, 1999) and an advantageous position in the marriage market (Charles et al. 2013; Greenwood et al., 2014; Eika et al. 2019), a college degree is an important factor for climbing up the social ladder in China. Given this, in China the college entrance examination is very competitive and usually requires students to be very well prepared and parents to be intensively involved in their children's studies, especially when students are in the last year of senior high school. In particular, we examine whether the SARS epidemic enlarges the association between parental education and having a college degree for children who took the college entrance examination during the SARS epidemic in 2003.

¹ The higher the intergenerational transmission, the lower the intergenerational mobility.

We conduct the empirical analysis mainly based on China’s 2010 population census. Merging census data with city-level SARS case numbers from reports of the Ministry of Health of China and adopting a triple difference-in-differences (DDD) strategy, we find that SARS significantly increases mother-children ITE for those children who took the college entrance examination in 2003, during the SARS epidemic.² The results suggest that if the SARS case number increases by 1 percent, the mother-children ITE increases by 9.3 percent. We also estimate the SARS effect on students who took the college entrance exam in 2002, which was before the SARS epidemic, as a falsification test, and students who took the college entrance exam in 2004 to 2006, which was after SARS, to examine the long-run effect. The results suggest a nil SARS effect on students other than the cohort taking the exam in 2003, confirming the validity of our DDD approach and suggesting the SARS effect is not long-lasting. On one hand, these results show that the SARS epidemic significantly reduces the intergenerational mobility of education for students. On the other hand, the results on the long-run effect indicate that this negative effect can be mitigated by postponing the examination time.

In analysing the heterogeneity, we find that the SARS effect is stronger for female students and students in large cities, probably due to son preference and the fact that the virus is more easily transmitted in crowded areas. In terms of education outcomes, the SARS effect is stronger for 4-year bachelor programmes relative to 3-year college programmes, perhaps because the latter require less effort from parents. In contrast to maternal ITE, we find no evidence that the SARS epidemic influences paternal ITE, regardless of whether assortative mating is controlled for. This is probably because the mother usually plays the major role in providing care to children in households (Blau and Lawrence, 2017; Kleven et al., 2019; Alon et al., 2020; Del Boca et al., 2020; Sevilla and Smith, 2020), and this role becomes particularly important during an epidemic period.

We finally discuss the potential mechanisms for SARS to affect ITE using two large datasets, i.e. China Family Panel Studies (CFPS) and Urban Household Survey (UHS). We argue that given the low incidence rate of SARS, having more health knowledge is unlikely to be the main reason for SARS to affect ITE. Instead, we provide evidence that highly educated mothers play a more critical role in the education of children. We find that, in general, mothers with more education interact more with children about their study, and in particular, they spend more on the children’s study, e.g. finding private tutors during the SARS period, in order to compensate for the loss of formal school education. We deem this finding an indication that more highly educated mothers tend to be more engaged in their children’s studies to offset the negative effect of the epidemic.

This paper enhances our understanding of the determinants of ITE. As a fundamental determinant of social mobility and inequality, economists have been long interested in ITE. A large body of literature estimates the size of ITE (see the review by Black and Devereux, 2011, and estimates for China, see e.g. Meng and Zhao, 2016, Hu et al., 2020) and investigates factors affecting ITE or intergenerational mobility, such as provision of public education (e.g. Davies et al., 2005; Bernasconi and Profeta, 2012; Uchida, 2018; Assaad and Saleh, 2018), school tracking systems

² In fact, we use the difference-in-differences strategy to identify the impact on ITE coefficients. Because a triple interaction term is included in estimation equation 2, as Section 3 shows, we call it triple difference-in-differences for simplicity.

(Bauer and Riphahn, 2006), school starting age (Bauer and Riphahn, 2009), higher education expansion (e.g. Blanden and Machin, 2004; Li and Zhang, 2017, Liu and Wan, 2019), public expenditure on education (Grawe, 2010), household assets (Huang, 2013), cigarette tax policy (Settele and Ewijk, 2018) and natural disasters (Caruso and Miller, 2015; Caruso, 2017). However, literature on the role of epidemics in ITE is rare. To the best of our knowledge, there is only one study in this area. Akbulut-Yuksel and Turan (2013) analyse the impact of the HIV/AIDS epidemic in sub-Saharan Africa and find that HIV prevalence reduces maternal ITE. As a mother-to-child transmitted disease, HIV may have a particularly strong impact on children in developing countries with limited medical resources. Whether this result can be generalised to other diseases and/or societies with different development levels remains unknown. This paper directly contributes to this strand of literature by examining the effect of a non-mother-to-child transmitted disease in a society with better medical facilities. Another difference between this paper and Akbulut-Yuksel and Turan's paper is the methodology. Akbulut-Yuksel and Turan compare the ITE parameters of mothers with different HIV statuses or mothers living in communities with different prevalence rates. However, we also include the pre-epidemic period to conduct a difference-in-differences analysis, which can remove unobserved confounding factors.

This paper also helps us to understand the human capital impact of the current COVID-19 pandemic. Given the huge negative impact of the current pandemic, a number of pioneering studies have investigated the impact of COVID-19 on student performance. These studies suggest that low socioeconomic status students are particularly vulnerable in this pandemic, implying magnified education inequality during COVID-19 (Aucejo et al., 2020; Bacher-Hicks et al., 2020; Chetty et al., 2020). Nevertheless, these papers mainly employ before-after comparison (or event study) on the usage of or search for online education programmes or directly elicit hypothetical performance in the situation without COVID-19 to identify the pandemic effect. This paper complements the literature by using a triple difference-in-differences strategy with a continuous treatment variable to identify the effect on a real measure of student performance, the outcome of college entrance examinations, with the SARS experience. If everything else is kept constant, according to the incidences of COVID-19 by 6 July 2020 (i.e. the day before the college entrance examination), our estimates suggest that in the three provinces where COVID-19 is most prevalent in the mainland China, the maternal ITE parameters would be increased by 103% (Hubei), 69% (Guangdong) and 66% (Henan), respectively. However, a caveat that should be kept in mind is that compared to COVID-19, SARS lasted for a much shorter time and was less prevalent, and also recent advancements in information technology have made online teaching popular during the current COVID-19 pandemic, which may mitigate or enlarge the negative impact. Given this, our estimated effect of SARS is just a preliminary estimate for the COVID-19 pandemic.

The remainder of the paper is structured as follows. Section 2 presents background information on the SARS epidemic. Section 3 introduces the data and econometric methodology. Section 4 details the main results. Section 5 discusses the mechanisms, and Section 6 concludes.

2. Background

As the first severe and highly infectious new disease in the 21st century, the SARS outbreak resulted

in significant losses for many countries in 2003. The first SARS incidence emerged in Foshan city of Guangdong Province in China in November 2002 (Editorial Board of Caijing Magazine, 2003). After that, the SARS virus started to spread in Guangdong, China. By February 2003, several cities, including the mega city Guangzhou, had reported the atypical pneumonia cases caused by the SARS virus to the Guangdong government. However, probably because the causes and consequences of SARS were poorly understood at that time, its importance was not recognised by the government. On 11 February, around three months after the first case, the health department of Guangdong province and the Guangzhou government held their first press conference about the atypical pneumonia. In the conference, the government stressed that the disease had started to be contained, and the prevalence was very low given the large population size in Guangzhou.³ Probably due to the lack of preparation and high transmissibility of the disease, SARS spread rapidly to other regions in China as well as to other countries. The global cumulative total case number increased quickly from 5,000 on 28 April to 7,000 on 8 May (WHO, 2003). By 25 June, 102 prefecture cities and municipalities in mainland China had found SARS cases.⁴ By 15 August, globally there were 8,422 SARS cases with 919 deaths, among which 7,747 cases and 829 deaths had occurred in China (Jia and Liu, 2004).

As the epidemic quickly spread and a large fraction of patients required intensive medical treatment, SARS caused great public panic. People were panic buying Radix Isatidis granules (*banlangen keli*) and white vinegar, which they believed could help to prevent one from catching a cold (Editorial Board of Caijing Magazine, 2003). Given the social panic, the Chinese government started to be concerned about this issue. On 13 April, SARS was included in the *Law of the People's Republic of China on Prevention and Treatment of Infectious Diseases*, which requires the government, including local governments, to make SARS information (e.g. case numbers and death numbers) public. Many government officials lost their jobs because of their unsatisfactory performance in containing SARS, including the minister of health and the mayor of the capital city, Beijing. Because there was no vaccine or effective treatment at that time, the basic control tools, isolation and quarantine, were resorted to contain the disease. On 12 April, the government required all prefecture cities to set up quarantine centres and check the health status of travellers. All suspicious and confirmed cases were required to be quarantined. In late April, the government also shut down the tourism industry in central and western China and rural areas. All public places, including food markets and restaurants, were shut down if any suspicious or confirmed case was found.⁵

School study was interrupted as well. On 23 April, the ministry of education issued a policy for students in high schools, primary schools and kindergartens.⁶ It required schools to adjust teaching progress and reduce students' workloads to guarantee that students would have enough rest to maintain good health. For final-year students in high schools, who were expected to have studied very intensively had SARS not occurred, the policy asked schools to reduce student gatherings in schools as much as possible, which shortened time at school substantially. The policy also suggested

³ For details please see Editorial Board of Caijing Magazine (2003) and <https://news.sina.com.cn/c/2003-02-12/0131904406.shtml>.

⁴ This information is from reports of the Ministry of Health of China compiled by the Jiangsu office of the CDC. There are 333 prefecture cities and 4 municipalities in China.

⁵ For details about the control measures, please see Jia and Liu (2004).

⁶ The document title is the *notice on the SARS containing job for high schools, primary schools and kindergartens* (*guanyu zuohao zhongxiaoxue he youeryuan feidian fangzhi gongzuo de tongzhi*).

that, depending on the local situation, schools could be closed to prevent SARS transmission, with approval of the local government. Given this policy, the school studies of students in different regions were affected at varying levels according to local SARS prevalence. Through a search for online news on the school interruption by SARS, we find that in the regions with many SARS cases, for example, Beijing and cities in Inner Mongolia and Shanxi, schools were shut down for different periods according to the city-level policy regardless whether there were confirmed or suspicious SARS cases found in the school, while in the regions with fewer SARS cases, for instance, cities in Jiangsu and Zhejiang, schools were shut down only when confirmed or suspicious SARS cases were found.⁷

Social panic, school interruption and other possible impacts brought about by the SARS epidemic might worsen academic performance of students. The above discussion suggests that these impacts caused by the SARS epidemic are generally proportional to the local SARS cases. Therefore, in the regions with more SARS cases, performance of students might more rely on their parents' effort. As shown in Figure 1 there was large variation in SARS cases across cities (prefectures) in 2003, in the following we first employ this large cross-city variation in local SARS intensity to identify the SARS effect on ITE and then investigate the possible mechanisms. However, a caveat we should keep in mind is that due to data constraint, in the following mechanism analysis we can only test some particular mechanisms, and our results on the mechanism are only indicative, rather than conclusive.

3. Data and Econometric Methodology

In this section we discuss the data and econometric methodology used in the main analysis.

3.1 Data Description

We use data from two sources: a random sample from the 2010 population census and the cumulative SARS case numbers from reports of the Ministry of Health of China.

The random sample of 2010 census long form data provides us with the individual characteristics used for analysis, such as gender, age, household registration type, individual educational attainment and relationship with respect to the household head.⁸ Importantly, a unique feature of the 2010 census, in comparison to previous censuses, is that it interviewed and recorded the information of migrants in both the destination place and the household registration place.⁹ This feature allows one to construct two random samples: one based on the household register address, and one based on the residence address at the census time. In this paper we conduct the analysis on the *former* sample, because the former one allows us to construct more parent-child links. Specifically, based on the former sample, we can construct parent-child links for those children who were staying with their parents in the same household, as well as those who had moved out of the household but were

⁷ Note that, however, we were only able to conduct incomplete search for the relevant online news and the search results only cover a very limited set of cities, because many relevant webpages had been removed. We include the relevant links in Online Appendix 1.

⁸ There are two types of household registration in China: agricultural and non-agricultural (i.e. rural and urban).

⁹ Theoretically speaking, a migrant should appear twice in the 2010 census. Once in the destination region, and once in the household registration address. The previous censuses only interviewed migrants on the destination place at the census time.

listed in the same household register with their parents. But the sample based on the residence address at the census time allows us to match only parents and children who resided together.

The full random sample based on household register address consists of 3,521,985 observations, accounting for 2.6‰ of the population in China. We compare the characteristics of our random sample with the information published on the website of the National Bureau of Statistics. Online Appendix Table 1 and Appendix Figure 1 show they are highly consistent. Based on this sample, we use the relationship with respect to the household head to derive parent-child links.¹⁰

We then merge the census data with the SARS case numbers. Specifically, we use the mainland China city-level cumulative SARS case numbers by 6 June in the main analysis, as the college entrance examination took place on 7 June. This information is extracted from reports of the Ministry of Health of China compiled by the Jiangsu office of CDC.¹¹ Figure 1 shows the distribution of the cumulative SARS cases.

3.2 Econometric Methodology

In general, the traditional ITE estimation model can be represented as follows:

$$y_i = r_0 + r_1 Pedu_i + \gamma_3 X_i + \epsilon_i, \quad (1)$$

where y_i is the education outcome of children, $Pedu_i$ is the measure of parental education attainment, X_i is a set of control variables, ϵ_i is the error term and γ_1 is the estimated ITE.

As our goal is to identify the impact of SARS on ITE, in the main analysis we extend Equation 1 to a triple difference-in-differences model. Specifically, in the main analysis, we define those who were born between September 1984 and August 1985 as the treatment group. Due to the facts that the compulsory education law stipulates children start schooling at six, that it usually takes 12 years to finish primary school and high school, and that the academic year starts in September, this group of individuals are supposed to take the college entrance examination in 2003. To make the control group comparable, we take only those who are supposed to take the examination between 2000 and 2002, i.e. those who were born between September 1981 and August 1984, as the control group.¹²

¹⁰ The options for the relationship with respect to household head include household head, spouse, children, parents, parents-in-law, grandparents and grandparents-in-law, children-in-law, grandchildren and grandchildren-in-law, siblings and siblings-in-law and other. We derive the parent-child links based on the combination of household head, spouse and children, the combination of parents and household head and the combination of parents-in-law and spouse.

¹¹ The data are stored in the following link:

https://web.archive.org/web/20030625152900fw_/http://www.jshealth.com/hotspots/fdx/y/w66.htm. For the relevant information, please also refer to the link:

<https://web.archive.org/web/20030808194902/http://www.jshealth.com/>.

¹² Note that while in principle the school starting age is six and the compulsory education (i.e. primary school and junior high school) and senior high school education take nine and three years, respectively, in some less developed areas the school starting age is seven and/or the compulsory education took eight years (i.e. five years for primary school and three for junior high school) due to the insufficient supply of schools and teachers. This causes three atypical cases in addition to starting schooling at six and taking twelve years to finish the compulsory and senior high school education. First, school starting age is seven and it takes eleven years to finish the compulsory education and senior high school education (five years for primary school); second, school starting age is six and it takes eleven years to finish the education; and third, school starting age is seven and it takes twelve years to finish the education (six years for primary school). In the first scenario, our estimation approach is still valid, as the treatment group still takes the college examination in 2003 (the SARS period) and the control takes the examination between 2000 and 2002. However, the other two scenarios tend to bias down our estimates. Specifically, in the second scenario both control and treatment groups take the college entrance examination before 2003. This leads the estimates to be biased to zero. In the third scenario, the treatment group takes the examination in 2004, but the control takes the examination

In the robustness check, we also test the sensitivity of our results by including more or fewer cohorts in the control group. The results remain robust.

The regression specification for this DDD model is as follows:

$$y_{ict} = \beta_0 + \beta_1 Pedu_{ict} * \mathbf{1}(t = 2003) * Intensity_c + \beta_2 Pedu_{ict} * \mathbf{1}(t = 2003) + \beta_3 Intensity_c * \mathbf{1}(t = 2003) + \beta_4 Pedu_{ict} * Intensity_c + \beta_5 Pedu_{ict} + \beta_6 \mathbf{1}(t = 2003) + \gamma X_{ict} + \epsilon_{ict}, \quad (2)$$

where y_{ict} denotes whether child i who is supposed to take the college entrance examination in year t had received college or higher education at the time of the 2010 census in city c , $Pedu_{ict}$ is the parental years of schooling, and $\mathbf{1}(t = 2003)$ is a dummy variable indicating whether child i in city c took the college entrance examination in 2003 when the SARS outbreak emerged.¹³ $Intensity_c$ is the measure for city-level SARS intensity. As shown in Section 2, more SARS cases is generally associated with larger negative impacts, such as more social panic and school interruption. In the main analysis we use the natural logarithm of *one plus* the number of city-level cumulative confirmed cases to represent the intensity of SARS, given that SARS case numbers are highly skewed as shown in Figure 1. We also test the robustness of the results by using the case number without logarithm transformation and using the natural logarithm of the case number. The results remain similar, as shown in Section 4.3. In the vector of control variables, X_{ict} , we include a gender dummy and interaction between household registration type and city to flexibly control for unobserved heterogeneity. As the interaction between household registration type and city is controlled for, we do not separately control for $Intensity_c$ to avoid the multicollinearity problem. Finally, ϵ_{ict} is the error term. This model exploits variations in cohort and SARS intensity to identify the impact on ITE. β_1 is the coefficient of interest.

One issue is that the DDD model requires the parallel trend assumption, and $Intensity$ may pick up a pre-treatment trend difference. We use two falsification tests on this assumption. First, we also assign the SARS intensity measure to those taking exam in 2002 and estimate the following model:

$$y_{ict} = \beta_0 + \beta_1 Pedu_{ict} * \mathbf{1}(t = 2003) * Intensity_c + \beta_2 Pedu_{ict} * \mathbf{1}(t = 2003) + \beta_3 Intensity_c * \mathbf{1}(t = 2003) + \beta_4 Pedu_{ict} * Intensity_c + \beta_5 Pedu_{ict} + \beta_6 \mathbf{1}(t = 2003) + \gamma X_{ict} + \phi_1 Pedu_{ict} * \mathbf{1}(t = 2002) * Intensity_c + \phi_2 Pedu_{ict} * \mathbf{1}(t = 2002) + \phi_3 Intensity_c * \mathbf{1}(t = 2002) + \phi_4 \mathbf{1}(t = 2002) + \epsilon_{ict}, \quad (3)$$

where $\mathbf{1}(t = 2002)$ is a dummy variable indicating whether child i is supposed to take the college entrance exam in 2002 before SARS appeared. If ϕ_1 is significant, then the parallel trend is violated. Similarly, we also exclude the treatment group and use the sample of the control group only to estimate the following model:

$$y_{ict} = \beta_0 + \beta_1 Pedu_{ict} * \mathbf{1}(t = 2002) * Intensity_c + \beta_2 Pedu_{ict} * \mathbf{1}(t = 2002) + \beta_3 Intensity_c * \mathbf{1}(t = 2002) + \beta_4 Pedu_{ict} * Intensity_c + \beta_5 Pedu_{ict} + \beta_6 \mathbf{1}(t = 2002) + \gamma X_{ict} + \epsilon_{ict}. \quad (4)$$

In this case, if β_1 is significant, the parallel trend is violated as well.

Theoretically speaking, SARS can also have long-term effects on college enrolment for students

between 2001 and 2003 (the SARS period), which causes downward bias to our estimates. Therefore, our estimates are the lower bounds of the true effects.

¹³ We recode the categories of educational attainment, no schooling, primary school, junior high school, senior high school, college, university and postgraduate, as the equivalent years of schooling, 0, 6, 9, 12, 15, 16, 19, respectively.

who were in the early stages of high school (e.g. first or second year in senior high school). We test the potential long-term effect by including students who were born between September 1985 and August 1988, i.e. those who are supposed to take college entrance exams between 2004 and 2006. In particular, we use two models to estimate this potential long-term effect. First, we estimate the average long-term effect over the three cohorts who took the exam between 2004 and 2006 as follows:

$$y_{ict} = \beta_0 + \phi Pedu_{ict} * \mathbf{1}(t = 2004, 2005 \text{ or } 2006) * Intensity_c + \sigma Pedu_{ict} * \mathbf{1}(t = 2004, 2005 \text{ or } 2006) + \theta Intensity_c * \mathbf{1}(t = 2004, 2005 \text{ or } 2006) + \tau \mathbf{1}(t = 2004, 2005 \text{ or } 2006) + \beta_1 Pedu_{ict} * \mathbf{1}(t = 2003) * Intensity_c + \beta_2 Pedu_{ict} * \mathbf{1}(t = 2003) + \beta_3 Intensity_c * \mathbf{1}(t = 2003) + \beta_4 Pedu_{ict} * Intensity_c + \beta_5 Pedu_{ict} + \beta_6 \mathbf{1}(t = 2003) + \gamma X_{ict} + \epsilon_{ict}, \quad (5)$$

where ϕ indicates the potential long-term effect. Second, we estimate the SARS impact for the respective cohort taking the exam between 2004 and 2006. The model can be written as:

$$y_{ict} = \beta_0 + \sum_{j=1}^3 \phi_j Pedu_{ict} * \mathbf{1}(t = 2003 + j) * Intensity_c + \sum_{j=1}^3 \sigma_j Pedu_{ict} * \mathbf{1}(t = 2003 + j) + \sum_{j=1}^3 \theta_j Intensity_c * \mathbf{1}(t = 2003 + j) + \tau \mathbf{1}(t = 2004, 2005 \text{ or } 2006) + \beta_1 Pedu_{ict} * \mathbf{1}(t = 2003) * Intensity_c + \beta_2 Pedu_{ict} * \mathbf{1}(t = 2003) + \beta_3 Intensity_c * \mathbf{1}(t = 2003) + \beta_4 Pedu_{ict} * Intensity_c + \beta_5 Pedu_{ict} + \beta_6 \mathbf{1}(t = 2003) + \gamma X_{ict} + \epsilon_{ict}, \quad (6)$$

where ϕ_j is the coefficient of interest.

It is worth noting two technical issues in the above difference-in-differences estimations. First, as we use the city-level SARS case number to construct the SARS intensity measure, in the estimation we cluster the standard errors at the city-level in case that the error term, ϵ_{ict} , is correlated with each other within the city as well. Cameron and Miller (2015) point out that cluster-robust inference requires a large number of clusters to produce consistent estimates of clustered standard errors. In the following estimation we have 336 cities in the sample. Therefore, the “large number of clusters” condition is satisfied. Second, a consistent difference-in-differences estimator requires both the treatment and control groups to be large (Conley and Taber, 2011). Given that 29% of cities (98) had non-zero SARS cases (i.e. the treatment group) as shown by Figure 1 and a large sample is used in the analysis as shown in the next section, our treatment and control groups are large enough to yield consistent estimates.

3.3 Summary Statistics

As mentioned previously, we restrict the sample to children born between September 1981 and August 1988 to construct proper control and treatment groups to analyse the SARS effect on ITE, including the relatively long-run effect. As we analyse the SARS effect on college entrance examinations, we further restrict the samples to children who reported having senior high school education or above. Removing observations with missing covariates, we construct 80,347 mother-child links as the sample for our main analysis.¹⁴

Table 1 presents the summary statistics. In the sample, 53% of students attend a university or college,

¹⁴ We also examine the effect of SARS on paternal ITE in the later analysis. As the missing values are slightly different in paternal and maternal information, the sample size for father-child links is slightly different. But the summary statistics are similar. As the later results suggest that SARS does not affect paternal ITE, we focus on the summary statistics of the sample for maternal analysis here.

and 20% attend a university.¹⁵ On average, the maternal years of schooling is 8.6 years. The average number of SARS cases is 78. 57% of students are male, and 56% have agricultural household registration. Finally, 31%, 12% and 56% of students are supposed to take the college entrance examination between 2000 and 2002, in 2003 and between 2004 and 2006, respectively.¹⁶

4 Results

In this section we discuss the results for our main analysis. We first present the main results on the cohort taking the college entrance examination in 2003 (i.e. the SARS period) and then cohorts taking the examination after 2003, and finally we move to the results on heterogeneous effects and robustness checks.

4.1 Main Results

Table 2 shows the results on the SARS effect on maternal ITE. Column 1 presents the results of the DDD model (i.e. Equation 2). The coefficient of the triple interaction term suggests that SARS intensity significantly increases maternal ITE. Specifically, for students who took the college entrance examination in 2003, if the SARS cases number increases by 1%, the magnitude of maternal ITE increases by roughly 0.0018. To put this number in perspective, we conduct the conventional ITE estimation using Equation 1 in Column 2 employing the sample of the control group (i.e. students who are supposed to take the examination between 2000 and 2002 and thereby not affected by SARS). It shows that the size of maternal ITE is 0.0194, i.e. one additional maternal year of schooling is associated with 1.94% greater likelihood of attending a university or college for children. Comparing Columns 1 and 2, we can see that a 1% increase in SARS cases leads to an increase in maternal ITE by 9.3% (0.0018/0.0194). Given that the sample average SARS case number is 78, maternal ITE is increased by 40% on average by SARS, indicating a significant decline in intergenerational mobility.

The other coefficients in Column 1 are also interesting. The interaction between maternal years of schooling and SARS cases suggests that, in the control group, maternal ITE is larger in cities which suffered from SARS more seriously in 2003. The interaction between maternal years of schooling and the indicator for taking the examination in 2003 shows that maternal ITE in 2003 declines significantly relative to the control group in general. This is probably because in 2003 it was the first time that the college entrance examination took place in June, whereas previously the examination took place in July. Given this, parents had less time to help children to prepare for the examination than before. Further, the interaction between the SARS case number and the indicator for taking the examination in 2003 indicates that students in cities which suffered from SARS more seriously are less likely to attend a university or college than before. Finally, male students are shown to have a smaller chance of attending a university or college. This is consistent with recent literature on the gender reversal in educational outcomes (Blau and Lawrence, 2019).

One concern for the DDD model is that the parallel trend assumption may be violated and the SARS intensity may pick up the pre-treatment trend difference. Given this, we conduct two falsification

¹⁵ Note that we restrict the sample to children who reported having senior high school education or above, as mentioned in the previous paragraph.

¹⁶ The proportion of the younger cohort is larger. This is probably because younger children are more likely to stay in the same household register with their parents.

tests in Columns 3 and 4 by assuming that students who took the examination in 2002 are also affected by SARS. Specifically, for the falsification test in Column 3, we use the sample of both the control and treatment groups to estimate Equation 3; for the one in Column 4, we use the sample of only the control group to estimate Equation 4. If the coefficient of the triple interaction term among maternal years of schooling, SARS case numbers and taking the examination in 2002 is significant, then the DDD estimate may be biased due to violation of the parallel trend assumption. However, Columns 3 and 4 show that the coefficients of this triple interaction term are small and statistically insignificant (i.e. -0.0008 with a standard error of 0.0009), verifying the validity of the DDD model. It is also worth noting that the coefficient of the triple interaction term for students taking the examination in 2003 in Column 3 remains similar to the one in Column 1.

In addition to the above analysis, which uses years of schooling to estimate the ITE, we also test the potential non-linearity of different educational qualifications. In particular, we replace maternal years of schooling with a set of dummy variables of maternal education qualification, junior high school, senior high school and college or above, to re-estimate Equation 2.¹⁷ Figure 2 shows the results. It is evident that relative to children of mothers who had only primary school education or no formal schooling, SARS intensity increases the chance of having college or university education for children of mothers with higher education qualifications, and the more highly educated the mother is, the greater this enlargement is.

One may wonder whether SARS affects paternal ITE and whether the SARS effect on maternal ITE is caused by assortative mating through paternal ITE. Given this, we investigate paternal ITE in Table 3. First, we directly estimate the SARS effect on both maternal and paternal ITE in Column 1. Compared to Equation 2, in Column 1 we also include the triple interaction among paternal years of schooling, the natural logarithm of *one plus* SARS case number and taking the examination in 2003 and other relevant paternal related variables in the regression, as indicated in the footnote of Table 3. The results suggest that SARS does not directly affect paternal ITE, and the SARS effect on maternal ITE is *not* caused by assortative mating. One issue with the estimation in Column 1 is that the sample size is slightly smaller than that in Table 2 due to missing information on paternal education. To check the comparability of the samples in Tables 2 and 3, we re-estimate the SARS effect on maternal ITE, without controlling for the effect on paternal ITE, using the sample in Column 1. The result in Column 2 (i.e. 0.0020) is very similar to the one in Table 2 (i.e. 0.0018), confirming the sample comparability. Finally, we estimate the SARS effect on paternal ITE without controlling for the effect on maternal ITE in Column 3. The estimate (i.e. 0.0003) is again small and statistically insignificant.

The above results suggest that SARS increases the maternal ITE for the cohort who took the college entrance examination in the year when SARS occurred (i.e. 2003). Hypothetically speaking, SARS may also affect cohorts who took the examination after 2003, because SARS also interrupted their school studies to some extent. We investigate the SARS effect on the cohorts who are supposed to take the examination during 2004 to 2006 in Table 4. We estimate Equations 5 and 6 in Columns 1 and 2, respectively. While the estimated effect on the cohort taking the examination in 2003 remains the same as that in Table 2, the results suggest that SARS does not significantly influence the cohorts

¹⁷ Primary school or no formal schooling is taken as the base group.

taking the examination after 2003.¹⁸ This indicates that the disadvantage of children of lower-educated mothers imposed by SARS fades as time goes by. The policy implication drawn from these results is that the negative effect of epidemics and pandemics can be mitigated if the government postpones important examinations.

4.2 Heterogeneity in the Effect

In this section we explore potential heterogeneity in the SARS effect on maternal ITE. First, we explore the heterogeneity according to the explanatory variables. The results are shown in Panel A of Table 5. The first two columns of Panel A show the heterogeneous effects by gender. The results suggest that SARS mainly affects girls rather than boys. This is probably because the preference for sons in China protects boys from being adversely affected by epidemic shocks. The third and fourth columns examine the heterogeneity in city size. As the epidemic was more contagious in crowded areas, SARS may affect students in big cities more. We divide the sample into two parts: one is for individuals in municipalities and provincial capital cities (labelled as ‘large city’), and the other is for observations in the remaining cities (labelled as ‘small city’). Indeed, the results confirm our conjecture, showing that the effect mainly appears in big cities.

Second, we examine the potential heterogeneous effects of SARS on 3-year college and 4-year bachelor programmes. In Panel B a multinomial logit model taking no tertiary education as the base outcome is estimated. As the admission to 4-year bachelor programmes is more competitive and requires more parental involvement in examination preparation, it is expected that the SARS effect on maternal ITE is stronger for 4-year bachelor programmes. Indeed, the results suggest that while SARS increases maternal ITE through both 3-year college and 4-year bachelor programmes, the effect is mainly concentrated in the latter.

4.3 Robustness Checks

We conduct two sets of robustness checks in Table 6. First, Panel A shows the results from a set of alternative measures of SARS intensity. In the first two columns we test whether the particular functional form of the SARS case number drives our results. In the first column we use the confirmed cases number without logarithm transformation divided by 1000 to check whether our previous results are caused by the natural logarithm transformation. In the second column we directly take the natural logarithm of the cumulative number of SARS cases as the measure, so cities with no SARS cases are excluded from the sample. As the SARS virus was novel, due to unfamiliarity three cases were incorrectly classified as SARS cases in the reports by 6 June, and these were excluded in later reports (one in Beijing and two in Hohhot). In the third column, we exclude the incorrect cases and re-estimate the effect to check the robustness. The results are all robust.

Second, we test the sensitivity of our results by constructing a set of alternative control groups. Recall that in the main analysis students who are supposed to take the college entrance examination in 2003 are treated as the treatment group, and those supposed to take the examination between 2000 and 2002 (i.e. before SARS occurred) are treated as the control group. In Panel B, we take

¹⁸ In online Appendix Table 2, we also re-estimate the SARS effect on cohorts taking the examination after 2003 by excluding the cohort taking the examination in 2003 from the sample. The results are similar.

students supposed to take the examination in 2002, between 2001 and 2002, between 1999 and 2002 as the control groups, respectively. Different constructions of the control group may lead to different estimates. However, Panel B shows that our results are robust to the construction of the control group.

5 Mechanism Analysis

The above analyses establish the total effect of SARS on ITE. In this section, we discuss the potential mechanisms through which the SARS effect works. Hypothetically, the above results can be driven by the possibility that more highly educated mothers tend to have more health knowledge and thereby are more likely to prevent their children from being infected by SARS. Alternatively, highly educated mothers may be more engaged with children's study and more likely to help their children to keep mental stability to offset the negative consequences brought about by SARS, e.g. loss of formal school education and stress and anxiety.

We first examine the direct health channel. Health is important for the examination outcome (Currie and Moretti, 2003; Breierova and Duflo, 2004; Cutler and Lleras-Muney, 2008; see Currie (2009) for a review). SARS can affect student health directly, and highly educated mothers tend to have more health-related knowledge and thereby mitigate this negative impact more than less educated mothers. However, we believe that while this channel may exist, it is unlikely to be the main reason for SARS to impact ITE, as only a limited number of students were infected by SARS. In Yu et al.'s (2006) study, it is found that in mainland China only 10.2% of SARS cases were individuals who were aged between 10 and 19 in 2003. Projected from the 2000 and 2010 censuses, there were 20 million individuals who were 18 years old in 2003. Consider the extreme case. If all the SARS cases in the group aged between 10 and 19 occurred in individuals aged 18, the incidence rate would have been around 0.0026% to 0.0027%. Given that the sample size of the cohort taking the examination in 2003 in the regression of Column 1 of Table 2 is 9,876, this incidence rate implies that only 0.27 observation would have been affected by the direct health impact of SARS. Such a small number is unlikely to drive our results.

Now we turn to the second possibility, that more highly educated mothers can help children to better prepare for the examination. First, in Figure 3 using the baseline wave (i.e. 2010) of a large household survey, China Family Panel Studies, we find that more educated mothers are more likely to save for children's education, discuss what happens in school with children, and check children's homework, and are also less likely to watch TV.¹⁹ This suggests that more highly educated mothers tend to create a better study environment for children at home, in general. As schools were closed or school time was shortened substantially during the SARS period, more highly educated mothers may substitute in the role of school teachers, to some extent.²⁰ For example, they could purchase more books (e.g. problem sets) or employ private tutors for their children. We further employ the Urban Household Survey from 2000 to 2006 to investigate the differential SARS effect on educational expenditure by maternal education.²¹ In particular, in Table 7 we estimate Equation 2

¹⁹ We choose to use the 2010 sample, because it does not have the attrition issue and thus is representative and it is also closer to the SARS period. For data details, please refer to online Appendix 2.

²⁰ It has been found that more highly educated mothers tend to devote more time to children during the COVID-19 pandemic (Del Boca et al., 2020).

²¹ For data details, please refer to online Appendix 2.

with outcome variables being educational expenditure. We restrict the sample to households with only one child who was a student and aged 18. This sample restriction is applied for two reasons. First, as the school starting age is usually six, students aged 18 are supposed to be in their final year of senior high school. Second, to avoid complications due to economic returns to scale, we exclude households with two or more such children. We take the sample from the 2003 wave as the treatment group and the rest as the control group.²² We include household size, wave dummies and city dummies as the covariates in the estimation. The first two columns present the results on the share of income spent on all kinds of educational expenditure, i.e. the total educational expenditure divided by income, and the last column shows the results on the share of income spent on private tutorials. As explained in online Appendix 2, there are differences in the survey design before and after 2002 wave. Given this, for the share of total educational expenditure we conduct two estimations. One uses all seven waves, and the other uses the latter five waves. However, since information on private tutorials is unavailable before 2002, for the share on tutorials we only employ the waves in 2002 onwards to conduct the estimation. The results show that in cities where SARS was more prevalent, more highly educated mothers tend to spend more on children's education in terms of both total expenditure and private tutorials. One may have concerns that the estimated coefficient is economically small. However, given that the mean of the dependent variable is small and that the sample average maternal years of schooling is 11, the implied effect is non-trivial. For example, even for the city with only one SARS case, the average maternal years of schooling implies that the share of income in total education expenditure increases by 2.9 percentage points, equivalent to 24% of the mean value. Given this, we are inclined to believe that the second possibility is a more important reason for SARS to affect ITE.²³

6 Concluding Remarks with Discussion on COVID-19

In this paper, we employ a triple difference-in-differences approach to investigate the impact of SARS on ITE based on the outcome of college entrance examination. Our results suggest that SARS intensity significantly increases maternal ITE. A one percent increase in SARS case numbers leads to a 9.3 percent increase in maternal ITE for the cohort taking the college entrance examination during the SARS period. Our results also suggest that the SARS effect mainly concentrates on female students and students in large cities and is stronger for admission to 4-year bachelor programmes. In the mechanism analysis, we find that the main reason for SARS to affect ITE is that more highly educated mothers may be more engaged in their children's studies, e.g. having more interaction with their children or finding private tutors for their children, so the children of more educated mothers tend to perform better in the college entrance examination during the epidemic period.

The warning message this paper conveys is that pandemics and epidemics are likely to reduce intergenerational mobility and widen the education gap between students from low and high socioeconomic status families. This paper draws an important lesson from SARS for the current

²² As the previous analysis shows, SARS does not affect the cohorts taking the college entrance examination after 2003. Thus, these cohorts can also be treated as control groups to enhance the estimation efficiency.

²³ An important caveat we should bear in mind is that we deem the above finding an indication that in general more highly educated mothers tend to be more engaged in children's studies during an epidemic. There could also be other ways for more highly educated mothers to improve their children's studies. For example, they could spend more time with their children to relieve the stress and anxiety of their children. However, due to data constraints, we are not able to test the other potential channels.

COVID-19 pandemic. Given the COVID-19 case numbers and predicting from our regression, in the COVID-19 most prevalent three provinces in China: Hubei, Guangdong and Henan, maternal ITE could increase by 103%, 69% and 66%, respectively. However, we should bear in mind that these are just rough estimates. The impacts of COVID-19 and SARS may be different. Compared to SARS, COVID-19 is more prevalent and lasting longer. Also, online teaching has been used widely during the COVID-19 pandemic, which may mitigate or magnify the adverse impacts of the pandemic to some extent.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- Akbulut-Yuksel, M. & Turan, B., 2013. Left behind: intergenerational transmission of human capital in the midst of HIV/AIDS. *Journal of Population Economics*, 26(4), 1523-1547.
- Alon, T., Doepke, M., Olmstead-Rumsey, J. & Tertilt, M., 2020. The impact of Covid-19 on gender equality, NBER Working Paper, No. 26947.
- Assaad, R. & Saleh, M., 2018. Does Improved Local Supply of Schooling Enhance Intergenerational Mobility in Education? Evidence from Jordan. *World Bank Economic Review*, 32(3), 633-655.
- Aucejo, M. E., French, J., Araya, M. P. U. & Zafar, B., 2020. The Impact of COVID-19 on Student Experiences and Expectations: Evidence from a Survey. *Journal of Public Economics*, forthcoming.
- Breierova, L. & Duflo, E., 2004., The impact of education on fertility and child mortality: Do fathers really matter less than mothers? NBER Working Paper, No.10513
- Blanden, J. & Machin, S., 2004. Educational inequality and the expansion of UK higher education. *Scottish Journal of Political Economy*, 51(2), 230-249.
- Bauer, P. & Riphahn, T. R., 2006. Timing of school tracking as a determinant of intergenerational transmission of education. *Economics Letters*, 91(1), 90-97.
- Bauer, P. C. & Riphahn, R. T., 2009. Age at school entry and intergenerational educational mobility. *Economics Letters*. 103(2), 87-90.

Black, E. S. & Devereux, J. P., 2011. Recent Developments in Intergenerational Mobility. *Handbook of Labor Economics*, Vol.4, pp. 1487-1541(Elsevier).

Bernasconi, M. & Profeta, P., 2012. Public education and redistribution when talents are mismatched. *European Economic Review*, 56(1), 84-96.

Bacher-Hicks, A., Goodman, J. & Mulhern, C., 2020. Inequality in household adaptation to schooling shocks: Covid-induced online learning engagement in real Time. NBER Working Paper, No. 27555.

Blau, F. D. & Lawrence, M. K., 2017. The Gender Wage Gap: Extent, Trends, and Explanations. *Journal of Economic Literature*, 55 (3): 789-865.

Card, David. 1999. The Causal Effect of Education on Earnings. *Handbook of Labor Economics*, Vol. 3, pp. 1801–1863 (Elsevier).

Editorial Board of Caijing Magazine (财经编辑部). 2003. SARS Investigation: A record on an unprecedented disaster (SARS 调查：一场空前灾难的全景实录), China Social Science Press.

Currie, J. & Moretti, E. 2003. Mother's education and the intergenerational transmission of human capital: evidence from college openings and longitudinal data. *Quarterly Journal of Economics*, 118(4), 1495-1532.

Cutler, David & Lleras-Muney, Adriana. 2008. Education and Health: Evaluating Theories and Evidence, in: J House, R Schoeni, G Kaplan, and H Pollack (Eds.), *Making Americans Healthier: Social and Economic Policy as Health Policy*. New York: Russell Sage Foundation, pp. 29-60.

Currie, J. 2009. Healthy, wealthy, and wise: Socioeconomic status, poor health in childhood, and human capital development. *Journal of Economic Literature*, 47 (1), 87-122.

Conley, T. G. & Taber, C. R., 2011. Inference with “difference in differences” with a small number of policy changes. *Review of Economics and Statistics*, 93 (1), 113-125.

Charles, K. K., Hurst, E. & Killewald, A. 2013. Marital sorting and parental wealth. *Demography*, 50 (1):51-70.

Cameron, A. C. & Miller, D. L., 2015. A practitioner’s guide to cluster-robust inference. *Journal of Human Resources*, 50 (2): 317-372.

Caruso, G. & Miller, S., 2015. Long run effects and intergenerational transmission of natural disasters: A case study on the 1970 Ancash Earthquake. *Journal of Development Economics*, 117:

134-150.

Caruso, G. 2017. The legacy of natural disasters: The intergenerational impact of 100 years of disasters in Latin America. *Journal of Development Economics*, 127: 209-233.

Chetty, R., Friedman, J. N., Hendren, N., Stepner, M., Team, T. O. I. & Team, I., 2020. How did COVID-19 and stabilization policies affect spending and employment? A new real-time economic tracker based on private sector data. NBER Working Paper, No. 27431.

Davies, J., Zeng, J. & Zhang, J., 2005. Intergenerational mobility under private vs. public education. *Scandinavian Journal of Economics*, 107(3), 399-417.

Del Boca, D., Oggero, N., Profeta, P. & Rossi, C. M. 2020. Women's Work, Housework and Childcare, before and during COVID-19. IZA DP No. 13409.

Eika, L., Mogstad, M. & Zafar, B., 2019. Educational Assortative Mating and Household Income Inequality, *Journal of Political Economy*, 127(6), 2795-2835.

Grawe, N. D., 2010. Primary and secondary school quality and intergenerational earnings mobility. *Journal of Human Capital*. 4(4), 331-364.

Greenwood, J., Guner, N., Kocharkov, G. & Santos, C., 2014. Marry Your Like: Assortative Mating and Income Inequality. *American Economic Review*, 104(5), 348-53.

Huang, J., 2013. Intergenerational transmission of educational attainment: The role of household assets. *Economics of Education Review*, 33, 112-123.

Hu, Y., Behrman, J.R. & Zhang, J., 2020. The causal effects of parents' schooling on children's schooling in urban China. *Journal of Comparative Economics*, forthcoming.

Jia, K. & Liu, S. 2004. Public finance and public crisis: Thoughts triggered by SARS (公共财政与公共危机——“非典”引发的思考). China Financial & Economic Publishing House.

Kleven, H., Landais, C. & Egholt Sogaard, J. 2019. Children and Gender Inequality: Evidence from Denmark. *American Economic Journal: Applied Economics*, 11 (4): 181-209.

Li, Z. & Zhong, H., 2017. The impact of higher education expansion on intergenerational mobility. *Economics of Transition*, 25(4), 575-591.

Liu, L. & Wan, Q., 2019. The effect of education expansion on intergenerational transmission of education: Evidence from China. *China Economic Review*, 57(10), 101327.

Meng, X., Shen, K. & Xue, S., 2013. Economic reform, education expansion, and earnings inequality for urban males in China, 1988–2009. *Journal of Comparative Economics*, 41(1), 227-244.

Meng, X. & Zhao, G., 2016. The long shadow of the Chinese Cultural Revolution: the intergenerational transmission of education. IZA DP No. 10460.

Settele, S. & Ewijk, R. V., 2018. Can cigarette taxes during pregnancy mitigate the intergenerational transmission of socioeconomic status. *Labour Economics*, 55(C), 130-148.

Sevilla, A. & Smith, S., 2020. Baby Steps: The Gender Division of Childcare during the COVID-19 Pandemic, IZA DP No. 13302.

Uchida, Y., 2018. Education, social mobility, and the mismatch of talents. *Economic Theory*, 65(3), 575-607.

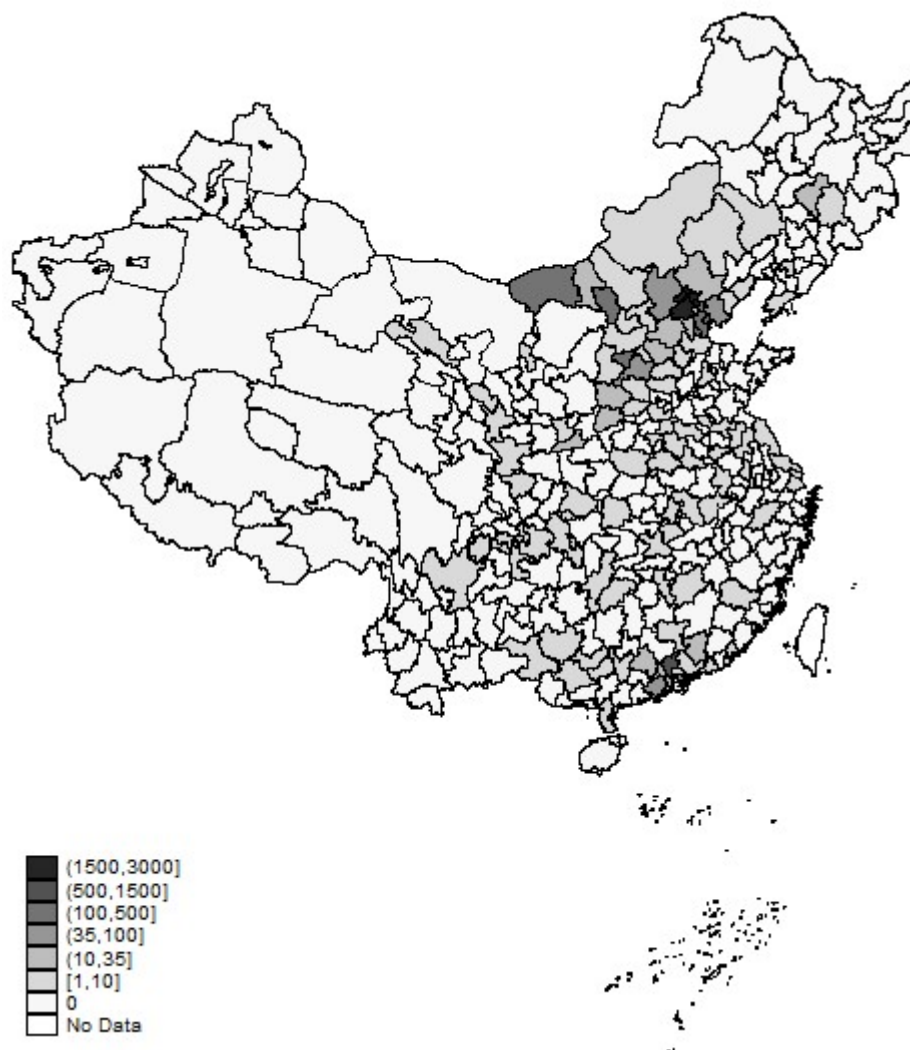
World Bank, 2003. From Cyclical Recovery to Long Run Growth: Regional Overview. <http://documents.worldbank.org/curated/en/711751468032401775/From-cyclical-recovery-to-long-run-growth-regional-overview>

World Health Organization, 2003. Severe acute respiratory syndrome (SARS): Status of the outbreak and lessons for the immediate future. https://www.who.int/csr/media/sars_wha.pdf?ua=1

Yu, P.L.H., Chan, J.S.K. & Fung, W.K., 2006. Statistical Exploration from SARS. *The American Statistician*, 60(1), 81-91.

Figures

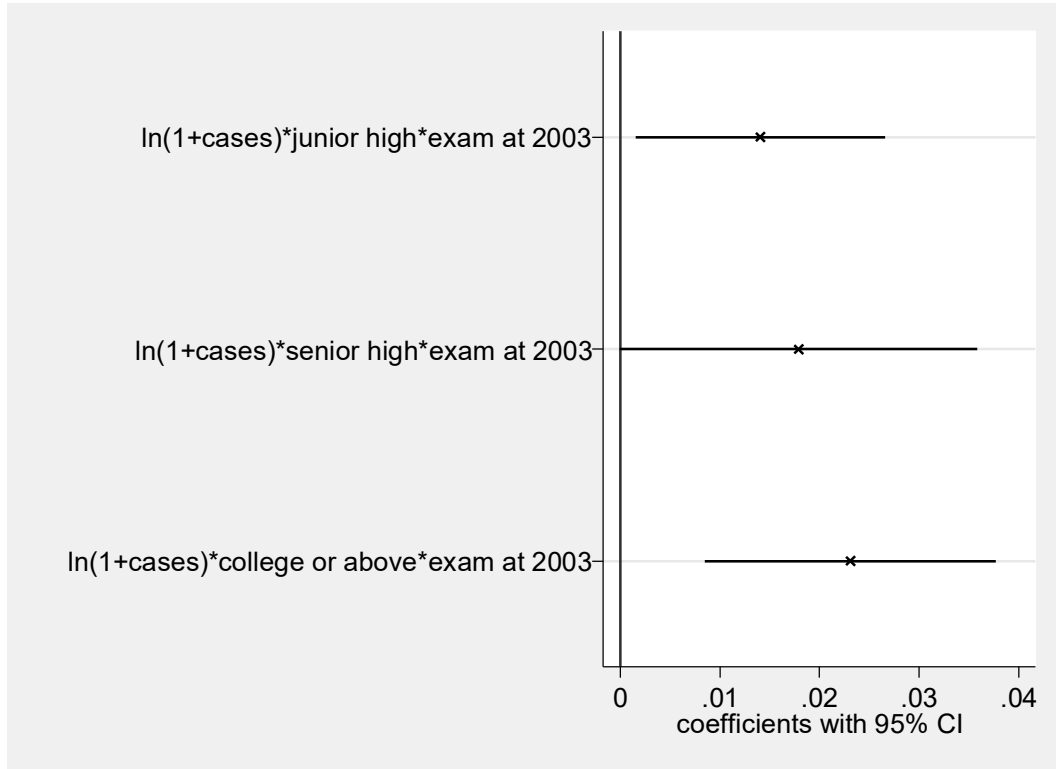
Figure 1: The Distribution of SARS Cases



Note: The cumulative case numbers by 6 June 2003 are plotted.

Source: Reports of the Ministry of Health compiled by CDC of Jiangsu Province.

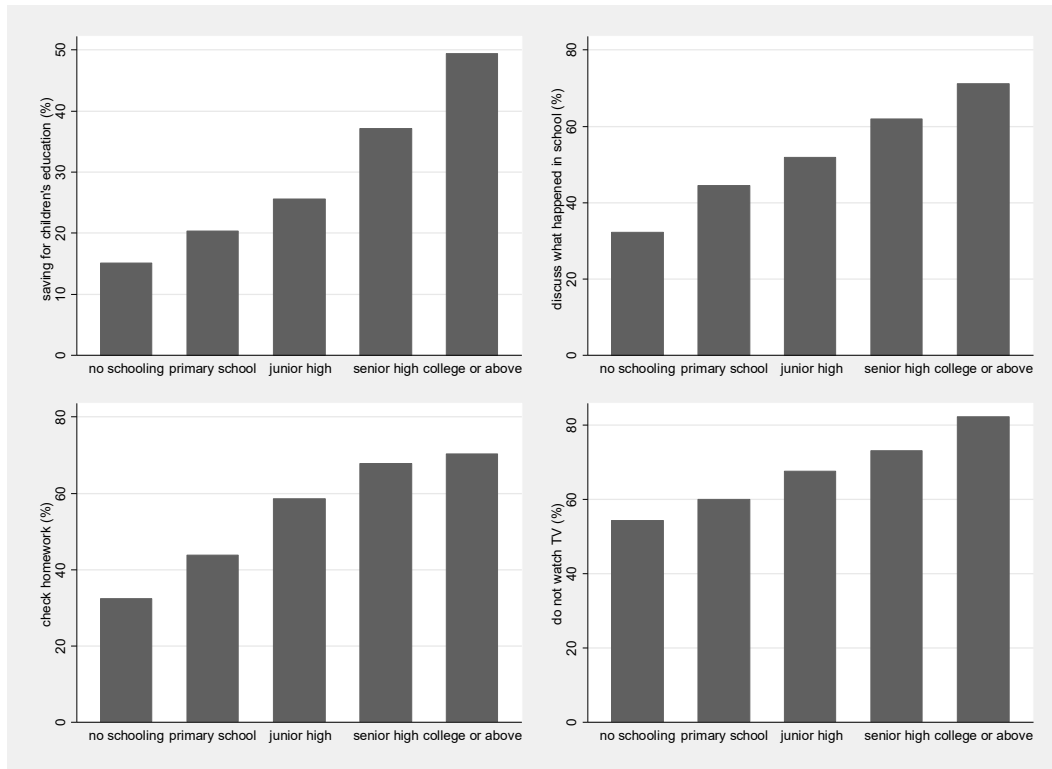
Figure 2: Nonlinear Effect of Educational Attainment



Note: In this figure, we estimate Equation 2 with the replacement of $Pedu_{it}$ with a set of dummy variables, i.e. junior high school education, senior high school education and college or above education. Note that we include different qualifications and their relevant interaction terms in one regression.

Source: The 3% random sample of the long form data of the 2010 China Census.

Figure 3: Mother's role in children's education



Source: 2010 China Family Panel Studies.

Tables

Table 1: Summary Statistics

Variable	Mean	Std. Dev.
3-year college or 4-year bachelor programme or above	0.53	0.50
4-year bachelor programme or above	0.20	0.40
mother's years of schooling	8.59	3.01
number of cases	78.19	397.39
male	0.57	0.50
rural hukou	0.56	0.50
cohort taking college entrance exam between 2000 and 2002	0.31	0.46
cohort taking college entrance exam in 2003	0.12	0.33
cohort taking college entrance exam between 2004 and 2006	0.56	0.50
Observations	80,347	

Source: The 3% random sample of the long form data of the 2010 China Census and reports of Ministry of Health of China compiled by CDC of Jiangsu Province.

Table 2: Results on the cohort taking college entrance examination in 2003

	DDD	IG estimation	falsification test	
	(1)	(2)	(3)	(4)
mother's years of schooling*ln(1+cases)*exam in 2003	0.0018** (0.0009)		0.0015* (0.0008)	
mother's years of schooling*ln(1+cases)	0.0020*** (0.0006)		0.0022*** (0.0008)	0.0021*** (0.0007)
mother's years of schooling*exam in 2003	-0.0075*** (0.0021)		-0.0071*** (0.0023)	
ln(1+cases)*exam in 2003	-0.0177* (0.0096)		-0.0155 (0.0097)	
mother's years of schooling	0.0174*** (0.0014)	0.0194*** (0.0014)	0.0168*** (0.0016)	0.0166*** (0.0016)
dummy for cohort taking exam in 2003	0.1028*** (0.0205)		0.1219*** (0.0217)	
male	-0.0569*** (0.0056)	-0.0566*** (0.0064)	-0.0560*** (0.0056)	-0.0555*** (0.0064)
mother's years of schooling*ln(1+cases)*exam in 2002			-0.0008 (0.0009)	-0.0008 (0.0009)
mother's years of schooling*exam in 2002			0.0002 (0.0021)	-0.0000 (0.0021)
ln(1+cases)*exam in 2002			0.0073 (0.0108)	0.0068 (0.0108)
dummy for cohort taking exam in 2002			0.0603*** (0.0199)	0.0617*** (0.0199)

Observations	35,037	25,161	35,037	25,161
R-squared	0.1638	0.1781	0.1662	0.1817

Note: Robust standard errors are clustered at city level. *Significant at 10% level; **significant at 5% level; ***significant at 1% level. The coefficients of the interactions of city dummies and hukou type and the constant term are omitted.

Source: The 3% random sample of the long form data of the 2010 China Census and reports of Ministry of Health of China compiled by CDC of Jiangsu Province.

Table 3: The role of assortative mating

	(1)	(2)	(3)
mother's years of schooling*ln(1+cases)*exam in 2003	0.0027* (0.0016)	0.0020** (0.0009)	
father's years of schooling*ln(1+cases)*exam in 2003	-0.0015 (0.0017)		0.0003 (0.0009)
Observations	31,281	31,281	31,281
R-squared	0.1782	0.1688	0.1762

Note: Robust standard errors are clustered at city level. *Significant at 10% level; **significant at 5% level; ***significant at 1% level. The other covariates included in the regressions in columns 2 and 3 are the same as those in Table 2. The regression in column 1 further controls for the interaction of paternal years of schooling and the dummy for the cohort taking the exam in 2003, the interaction of paternal years of schooling and ln(1+cases) and the variable of paternal years of schooling.

Source: The 3% random sample of the long form data of the 2010 China Census and reports of Ministry of Health of China compiled by CDC of Jiangsu Province.

Table 4: Selective results on the effect of SARS on the cohorts taking college entrance examinations between 2003 and 2006

	(1)	(2)
mother's years of schooling*ln(1+cases)*exam in 2003	0.0018** (0.0009)	0.0018** (0.0009)
mother's years of schooling*ln(1+cases)*exam between 2004 and 2006	0.0001 (0.0007)	
mother's years of schooling*ln(1+cases)*exam in 2004		-0.0000 (0.0006)
mother's years of schooling*ln(1+cases)*exam in 2005		0.0006 (0.0010)
mother's years of schooling*ln(1+cases)*exam in 2006		-0.0002 (0.0008)
Observations	80,347	80,347
R-squared	0.1351	0.1351

Note: Robust standard errors are clustered at city level. *Significant at 10% level; **significant at 5% level; ***significant at 1% level. Columns 1 and 2 present results from Equations 5 and 6, respectively. Coefficients of the other covariates are omitted.

Source: The 3% random sample of the long form data of the 2010 China Census and reports of Ministry of Health of China compiled by CDC of Jiangsu Province.

Table 5: Selective results the heterogeneous effect

	(1)	(2)	(3)	(4)
Panel A: Heterogeneous Effects across Covariates	OLS Regression			
	<u>gender</u>		<u>city size</u>	
	<u>female</u>	<u>male</u>	<u>small</u>	<u>large</u>
mother's years of schooling*ln(1+cases)*exam in 2003	0.0034*** (0.0013)	0.0005 (0.0011)	0.0015 (0.0023)	0.0037*** (0.0011)
Observations	14,187	20,850	26,004	9,033
R-squared	0.1767	0.168	0.1477	0.1567
Panel B: Heterogeneous Effects in the Outcome Variable	Multinomial Logit Regression			
	<u>3-year college</u>		<u>4-year bachelor or above</u>	
	<u>Coeff.</u>	<u>M.E.</u>	<u>Coeff.</u>	<u>M.E.</u>
mother's years of schooling*ln(1+cases)*exam in 2003	0.0103** (0.0051)	0.0011 (0.0008)	0.0147** (0.007)	0.0014* (0.0008)
Observations	35,037	35,037	35,037	35,037

Note: Robust standard errors are clustered at city level. *Significant at 10% level; **significant at 5% level; ***significant at 1% level. OLS regressions are estimated in Panel A, and multinomial logit model taking no tertiary education as the base outcome are estimated in Panel B. The other covariates included in both panels are the same as those in Table 2. In Panel B, Columns 1 and 3 show the estimated coefficients, and the other two columns show the estimated marginal effects.

Source: The 3% random sample of the long form data of the 2010 China Census and reports of Ministry of Health of China compiled by CDC of Jiangsu Province.

Table 6: Robustness checks

	(1)	(2)	(3)
Panel A: Different measures on the intensity of SARS	case number in a thousand w.o. any transformation	natural logarithm transformation	Excluding incorrect SARS cases
mother's years of schooling*SARS measure*exam in 2003	0.0102*** (0.0025)	0.0023** (0.0011)	0.0018*** (0.0025)
Observations	35037	15881	35037
R-squared	0.1639	0.1722	0.1639
Panel B: Different control groups	<u>exam at 02</u>	<u>control groups</u> exam at 01 & 02	exam bw. 99 & 02
mother's years of schooling*ln(1+cases)*exam in 2003	0.0024* (0.0013)	0.0017* (0.001)	0.0016* (0.0009)
Observations	18429	26643	40946
R-squared	0.1723	0.1682	0.1632

Note: Robust standard errors are clustered at city level. *Significant at 10% level; **significant at 5% level; ***significant at 1% level. The other covariates included are the same as those in Table 2.

Source: The 3% random sample of the long form data of the 2010 China Census and reports of Ministry of Health of China compiled by CDC of Jiangsu Province.

Table 7: Mother's role in education expenditure

	Total Education Expenditure		Private Tutorial Expenditure
	2000– 2006	2002– 2006	2002–2006
	(1)	(2)	(3)
mother's years of schooling*ln(1+cases)*2003 wave	0.0038** (0.0019)	0.0039** (0.0019)	0.0004* (0.0002)
Mean of the dependent variable	0.1196	0.12	0.0114
Observations	4,899	3,985	3,985
R-squared	0.0999	0.1009	0.1283

Note: Robust standard errors are clustered at city level. *Significant at 10% level; **significant at 5% level; ***significant at 1% level. The dependent variable is the corresponding education expenditure share divided by family income. The coefficients of the interaction between mother's years of schooling and 2003 wave, the interaction between ln(1+cases) and 2003 wave, the interaction between ln (1+cases) and mother's years of schooling, mother's years of schooling, household size, year dummies and constant terms are omitted. Column 1 employs the waves from 2000 to 2006, and the other two employ the waves from 2002 to 2006. The dependent variable in the first two columns is the share of income spent on education, and the one in the last column is the share of income spent on private tutorial. The sample is restricted to households with only one child who was a student and aged 18. Survey weights are applied.

Source: 2000–2006 waves of Urban Household Survey and reports of Ministry of Health of China compiled by CDC of Jiangsu Province.

Online Appendix 1: Online News for School Interruption

https://www.thepaper.cn/newsDetail_forward_6225125
<https://news.sohu.com/25/61/news208626125.shtml>
<https://news.sohu.com/78/82/news208608278.shtml?nbs>
<https://zhuanlan.zhihu.com/p/107624676>
<https://baijiahao.baidu.com/s?id=1664328984049488573&wfr=spider&for=pc>
<http://meeting.dxy.cn/576/article/471/576/582/591/5182.html>
<https://www.chinanews.com/n/2003-04-19/26/295906.html>
<http://society.eastday.com/epublish/gb/paper140/38/class014000025/hwz961004.htm>
http://tech.163.com/tm/030513/030513_93120.html
<https://idm.cctv.com/news/science/20030503/100456.shtml>
<https://baike.baidu.com/item/SARS%E4%BA%8B%E4%BB%B6>
<https://baijiahao.baidu.com/s?id=1658607212093872810&wfr=spider&for=pc>
<https://www.zhihu.com/question/374593474>
<https://news.sohu.com/63/54/news208615463.shtml>
<https://heilongjiang.dbw.cn/system/2003/04/26/012110005.shtml>
<http://edu.southcn.com/zhuanti/2003gaokao/baojian/200305210900.htm>
<https://news.sohu.com/67/88/news208948867.shtml>
<https://news.sohu.com/27/60/news208786027.shtml>
https://www.sohu.com/a/372538675_203066
<http://www.ordosedu.cn/info/1041/12853.htm>
<https://news.sina.com.cn/c/2003-05-02/15041027550.shtml>
<https://news.sina.com.cn/c/2003-05-02/15041027550.shtml>
http://blog.sina.com.cn/s/blog_49073da6010002v2.html
<https://news.sina.com.cn/c/2003-05-02/05481027104.shtml>
<https://law.lawtime.cn/d407257412351.html>
<https://baijiahao.baidu.com/s?id=1635742503805737520&wfr=spider&for=pc>

Online Appendix 2: Data Details for Section 5

The results in Figure 3 and Table 7 are derived from two large-scale household surveys, China Family Panel Studies (CFPS) and Urban Household Survey (UHS), respectively.

CFPS is a longitudinal survey conducted biannually. The baseline wave was conducted in 2010, and to date five waves have become publicly available. We use the baseline wave for the analysis in Figure 3, because it is representative of the general population in China and closer to the SARS period. The results in the four panels are from the following respective questions:

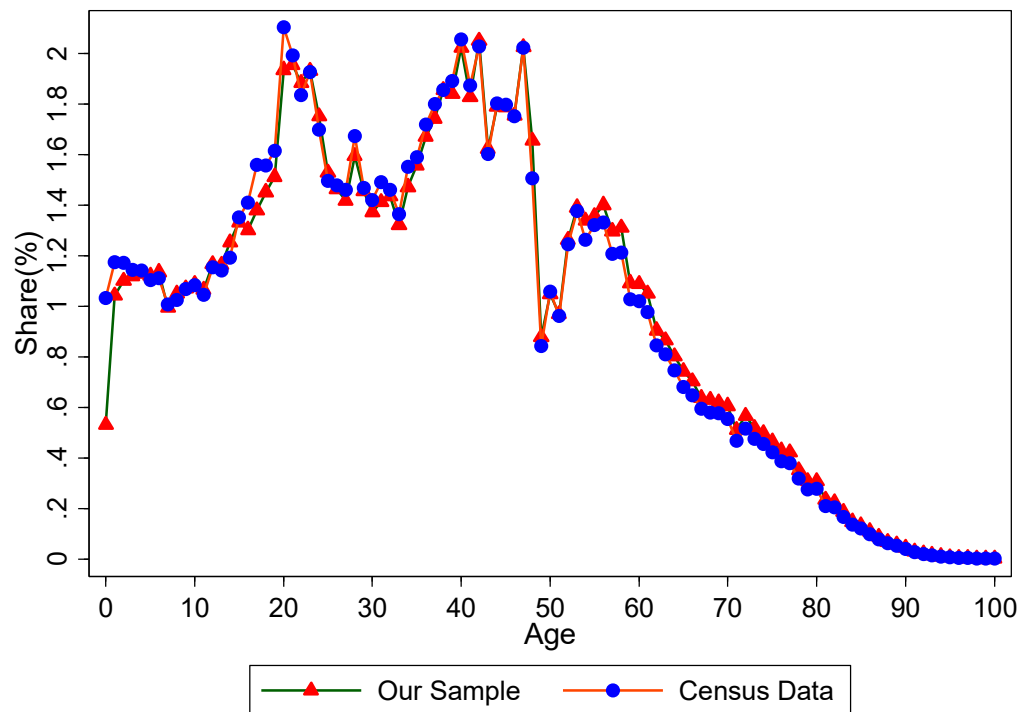
1. Have you saved money for the education of your children?
2. Since the beginning of this academic year/semester, do you often discuss what is happening in school with your children?
3. Do you often check the homework of your children?
4. When your children are studying, do you usually give up watching your favourite TV programmes to avoid interrupting the study of your children?

The answers to the first question are “Yes” (coded as 1) and “No” (coded as 0). The possible options to the other four questions are “very often (6–7 times per week)”, “often (twice or three times per week)”, “occasional (once or twice per week)”, “rare (once per month)” and “never”. The first panel of Figure 3 presents the share of “yes” for the first question, and the other three panels present the shares of “very often” and “often” for the other three questions, respectively.

UHS is a repeated cross-sectional survey aiming at documenting consumption and expenditure together with other basic information about urban households in China. We have access to the data of 16 provinces and municipalities: Beijing, Shanxi, Liaoning, Heilongjiang, Shanghai, Jiangsu, Anhui, Jiangxi, Shandong, Henan, Hubei, Guangdong, Chongqing, Sichuan, Yunnan and Gansu. To be consistent with the main results, we employ the waves from 2000 to 2006 for the analysis. Note that before 2001 (inclusive) UHS only included households with urban household registration, but since 2002 (inclusive) UHS has also included a small fraction of households with rural household registration (Meng et al., 2013). The sample size increases dramatically after 2002 (inclusive). In addition, the survey after 2002 (inclusive) includes more detailed information on educational expenditure. Given these, for the share of total educational expenditure we conduct two estimations: using all the seven waves and using 2002 to 2006 waves, respectively. For the share of expenditure on private tutorials we use only 2002 to 2006 waves, because the spending on private tutorials was not asked in the 2000 and 2001 waves. We restrict our sample to households in which the household head and spouse have children who were students and aged 18 during the survey time (i.e. supposed to be in the final year in the senior high school). This left us with 4,928 observations in all seven waves and 4,012 observations in the 2002 to 2006 waves for the analysis. We exclude households with more than one such child for simplification. The final sample sizes are 4,899 and 3,985, respectively. If we include households with more than one such child, the results are similar and available upon request.

Online Appendix Figures

Online Appendix Figure 1: The Comparison of Age Distributions between Our Sample and the Full Census Data



Source: Our sample is a 3% random sample of the long form data of the 2010 China Census. The full census data are extracted from the NBS webpage: <http://www.stats.gov.cn/tjsj/pcsj/rkpc/6rp/indexch.htm>.

Online Appendix Tables

Online Appendix Table 1: Representativeness of Our Random Sample (%)

	Our Sample	Full Census
female share	49.00	48.81
<i>share of educational attainment</i>		
no schooling	6.09	5.00
primary school	30.41	28.75
junior high	43.04	41.70
senior high	12.76	15.02
3-year college	4.60	5.52
4-year university	2.83	3.67
postgraduate	0.27	0.33

Note: The education variables are for individuals aged 6 and above.

Source: Our sample is a 3% random sample of the long form data of the 2010 China Census. The information for the full census is extracted from

<http://www.stats.gov.cn/tjsj/pcsj/rkpc/6rp/indexce.htm>, including both the long and short forms data.

Online Appendix Table 2: Results from alternative specifications on the effect of SARS on the cohorts taking college entrance examinations between 2004 and 2006

	04-06 cohorts	04 cohort	05 cohort	06 cohort
	(1)	(2)	(3)	(4)
mother's years of schooling*ln(1+cases)*exam between 2004 and 2006	0.0001 (0.0007)			
mother's years of schooling*ln(1+cases)*exam in 2004		-0.0005 (0.0007)		
mother's years of schooling*ln(1+cases)*exam in 2005			0.0010 (0.0010)	
mother's years of schooling*ln(1+cases)*exam in 2006				0.0001 (0.0008)
Observations	70,471	37,492	41,360	41,941
R-squared	0.1358	0.1613	0.1522	0.1479

Note: Robust standard errors are clustered at city level. *Significant at 10% level; **significant at 5% level; ***significant at 1% level. In all the four columns the control group are those supposed to take the exam between 2000 and 2002, and the treatment groups are those supposed to take the exam between 2004 and 2006, 2004 only, 2005 only and 2006 only, respectively. We estimate Equation 2 by replacing $1(t = 2003)$ with $1(t = 2004, 2005 \text{ or } 2006)$, $1(t = 2004)$, $1(t = 2005)$ and $1(t = 2006)$ in the four columns, respectively. Other coefficients are omitted.

Source: The 3% random sample of the long form data of the 2010 China Census and reports of Ministry of Health of China compiled by CDC of Jiangsu Province.