Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
The digital divide in online learning in China during the COVID-19 pandemic

Congbin Guo, Boshen Wan*

Graduate School of Education, Peking University, Beijing, 100871, China

ARTICLE INFO

Keywords:
Online learning
Covid-19
China
Digital divide

ABSTRACT

In recent years, online learning in the education sector has increasingly become prominent. While many believe that online learning has the potential to reduce inequity, the debate on whether it bridges the gap or widens it continues to persist. This study examined equity issues in online learning in China during the COVID-19 pandemic. The study used data from the Online Learning Survey of High School Students in China to analyze the influencing factors of the first, second, and third-level digital divide. The study found that the digital divide existed in online learning during the pandemic. It was primarily presented as differences in equipment quantity and network quality, students’ adaptability to online teaching, and their offline learning outcomes. These findings suggest that the development of online learning alone cannot eliminate achievement gaps. The promotion of education equity requires efforts from various stakeholders and interventions specifically targeting disadvantaged students.

1. Introduction

Enabled by information and communication technology (ICT), online learning has expanded dramatically in recent years. Compared to traditional offline instructions, online learning has certain advantages. First, online learning can make the learning process more student-centered by providing students with customized instructions and instant feedback [1]. Second, online learning is cost-effective, thereby promoting greater access to quality education [2]. Third, online learning is easily accessible. It offers increased learning opportunities for those who live in rural and remote areas, where local education resources are limited [3]. Given the advantages of online learning, techno-optimists believe that its diffusion has the potential to improve the quality of education and increase equity across the education sector.

However, the effects of online learning may vary from person to person. Many scholars have described these differences as the “digital divide,” which refers to the differences in the access, use, and skills of ICT among different social groups [4]. This idea argues that the distribution of ICT resources replicates the inequalities inherent in the social structure. Disadvantaged groups may not have access to online learning tools and materials. In addition, they are less prepared for remote learning without face-to-face instructions and parental guidance [5]. Therefore, as ICT is further integrated into education, the digital divide, if mishandled, could undermine the intended benefits of online learning, and even worse, reinforce existing social inequality.

Given the numerous ways in which ICT can be integrated into education, it is not surprising that opinions on equity issues around online learning vary. In reality, online learning can take place in different forms, under different contexts, and for different purposes, targeting various people. It is more advisable to examine its equity issues on a case-by-case basis. Besides, existing literature on the digital divide has primarily examined the differences in internet access and usage among various groups of people. The manifestations of the digital divide may be different depending on the context in which online learning is conducted [6].

In 2020, the sudden outbreak of COVID-19 caught the entire world off guard. During the pandemic, 1.5 billion students in 188 countries/economies were locked out of their schools [7]. Many countries were forced to adopt online learning to ensure learning continuity when social distancing was inevitable, and China was no exception [8,9]. China’s Ministry of Education required all primary and secondary schools across the country to stop in-person classes and encouraged students to continue their studies at home through online media. Since then, a large-scale online learning experiment has been conducted in China. Online learning has, to some extent, mitigated the negative impact of the pandemic on students’ learning. However, students from privileged
families may have the edge in utilizing online resources. Therefore, the gap in learning performance between different groups of students is likely to widen [10]. Considering that in the future, online learning could supplement or even replace offline education, not only in emergencies such as pandemics but also in people’s daily lives, it is necessary to examine the equity issues around online learning under such circumstances. In addition, more than 200 million Chinese students have engaged in online learning since schools were suspended. This could be the most extensive and unprecedented online education operation in the world. In this context, the analysis of the digital divide in online learning can offer valuable enlightenment for the efforts of improving digital education in China and the rest of the world.

This study aims to answer the following two questions based on data from the Online Learning Survey of High School Students During the Pandemic in China, which was conducted by the Big Data Center of the Graduate School of Education, Peking University. First, was the digital divide an issue in online teaching during the pandemic? Second, if yes, what factors contributed to the digital divide? The next section (Section 2) reviews the literature on the conception of the digital divide and research related to Chinese online learning during the COVID-19 pandemic. Section 3 introduces the research design, which includes data source and variable selection. Section 4 presents the empirical results. Section 5 concludes the paper and discusses potential mechanisms for the digital divide. The limitations of the study are also included in this section. Section 6 introduces future research perspectives.

2. Literature review

2.1. Digital divides at three levels

Discussion on the digital divide started in the 1990s. At that time, the internet and personal computers were not yet widely available [11,12]. Economically developed regions tended to have higher internet usage rates [13–15], and families with high socioeconomic status (SES) were more likely to have access to the internet [16–18]. According to literature, the digital divide is a dichotomous variable used to distinguish between those with and without internet access. This is referred to as the first-level digital divide [19].

Since the beginning of the 21st century, the popularization of ICT has made the internet readily accessible. With the gap in internet access being largely filled, academic interest in the digital divide has shifted to the differences in how the internet is used. These differences can be analyzed from two perspectives: purposes of using the internet and differences in internet skills. According to the first perspective, the internet has multiple functions that will bring different offline benefits to people [20]. Van Deursen and Van Dijk (2014) found that less-educated groups spend more time on the internet and use it primarily for gaming or social activities. In contrast, highly-educated groups use the internet less frequently and mostly use it for information retrieval and personal development [21]. These findings have been echoed by many other studies, which confirmed distinguishable patterns in internet use among different social groups [22,23]. Traditionally privileged groups prefer to convert online information into economic, human, or other types of capital to maintain their social status [24]. Moreover, people differ in their digital skills. Compared with mass media, the internet demands more advanced abilities for human-machine interaction [25]. Internet users have to extract favorable resources from a large amount of information presented on the internet. To better measure and compare the differences in internet skills, many scholars have tried to clarify the concept of internet skills and make them quantifiable [26–29]. In a study targeting the Dutch population, Van Deursen and Van Dijk (2015) divided internet skills into two broad categories, namely medium-related skills and content-related skills [30]. The study found that the general internet skills of the Dutch population and, in particular, their media-related skills, had an upward trend between 2010 and 2013. However, skill gaps by gender, age, and education remained mostly the same. Previous research found that in developed economies or districts with universal internet access, the digital divide relates more to internet usage and skills than physical access. Therefore, the difference in internet usage and skills is called the second-level digital divide.

For individuals, the offline outcomes of the internet fall into several dimensions, such as economic, political, and educational benefits [31]. However, not all groups benefit equally from the internet due to the first-level and second-level digital divide. This difference in offline benefits drawn from internet use is the third-level digital divide [32]. It has been argued that the internet is possibly an enabler and exacerbator of social inequality [33], as people with a higher social status use the internet more efficiently, which helps them expand their advantages by transforming online resources into offline benefits [34]. Regarding education, the debate remains on whether the proliferation of the internet and personal computers has bridged the education gap or, on the contrary, widened education inequality. Some scholars described the spread of information technology as a double-edged sword [35]. Information and communication technology facilitates the dissemination and sharing of educational resources, thus allowing disadvantaged groups to access high-quality education. In contrast, the presence of the internet is likely to widen the inequalities that already exist in the education system, given the differences in internet usage patterns and internet skills [36,37]. The controversy may come from the fact that there are multiple ways in which information technology and education can be integrated. The impact of ICT development on educational equity varies in different contexts. Therefore, it is necessary for one to conduct a careful examination of the digital divide in education in a specific context.

The theoretical concept of the digital divide and its citations are shown in Table 1.

2.2. Large-scale online learning in China during the COVID-19 pandemic

Since the global outbreak of COVID-19, the Chinese government has taken a series of measures to contain the spread of COVID-19 [38]. On January 27, 2020, during the peak of the pandemic, the Chinese Ministry of Education issued a policy to postpone school opening in the spring term. On February 12, 2020, the ministry advocated that all schools should offer online teaching to ensure continuity in learning. Consequently, by August 2020, as the virus had been hugely contained in China and most schools had been reopened, more than two hundred million students had been engaged in online classes as part of the official curriculum [39].

China’s nationwide online learning initiative was a response to a state of emergency following the pandemic. To some extent, it showcased China’s achievements in the promotion of integrating ICT into education. The whole country was mobilized, and all the necessary resources were fully utilized to support teachers and students [40]. The governments and schools provided free access to numerous online learning materials. Online streaming and communication platforms (for example, Zoom and Tencent Classroom) were widely used [41]. Schools
could choose from various ways of distance education depending on their actual needs. The options included live-streaming teaching, recorded online courses, online double-teacher collaborative teaching, and online mixed teaching [42]. Through online learning, schools effectively fulfilled their role to ensure the continuity of education during the pandemic. It is noteworthy that online education during the pandemic also brought about several issues. Network congestion, poor hardware quality, inadequate teacher-student interactions, and other problems compromised the quality and effectiveness of online teaching [43]. Some scholars have expressed concerns from the perspective of education equity [44]. Lei and Huang (2020) suggested that the digital divide among students with different social statuses has become a significant issue during the pandemic [45]. Guo (2021) found that during the pandemic, the first-level digital divide was nearly closed. However, rural students spent more time on online leisure activities and less focus on online learning [46].

To summarize, there are still some limitations in the existing literature on the digital divide and online learning during the COVID-19 pandemic. First, previous studies treated internet access as a dichotomous variable representing the presence or absence of internet connection while ignoring the differences in the quality of internet access. Although most students were able to study via the internet during the pandemic, differences in access quality (such as equipment fitness and network speed) might have significantly impacted students’ learning experience. Second, previous studies focused on the existence of the digital divide without considering how diverse it could be under different circumstances. Online learning during the pandemic is different from that in other periods because it is a compulsory and official means of schooling rather than a supplement to it. Therefore, the digital divide during this period is worth analyzing on a case-by-case basis. Finally, while some scholars expressed concern that online teaching during the pandemic could amplify existing inequity in the education system, empirical research has not yet verified this hypothesis.

3. Methodology

3.1. Research design

Through quantitative analysis, this study aimed to examine the existence of digital divides in online learning in China during the COVID-19 pandemic. We conducted a quantitative study for several reasons. First, we aimed to examine the equity issues for a large population. To increase external validity, we needed a large sample size [47]. It is more feasible to collect a large amount of data through quantitative questionnaires. Second, in our study, we aimed to identify the specific relationship between online learning performance and students’ family backgrounds. A quantitative study could properly address our research questions. Besides, qualitative data are more objective, helping to avoid researcher bias, which is often a challenge [48].

Considering that most students studied at home during the pandemic, their SES was the core influencing factor in our study. Moreover, given that the digital divide was formerly discussed at three circumstances, students’ adaptability to online learning is also affected by equipment and network conditions at home and, in turn, affects students’ learning performance when online education became universal during the pandemic. Therefore, as a core independent variable, the equipment and network conditions were also incorporated into the model (separately from the families’ SES) to test the path of its impact on students’ learning performance.

Adaptability to online learning: Like equipment and network conditions, students’ adaptability to online learning is also affected by their families’ SES. The adaptability index was calculated from students’ self-evaluation forms. This variable is discussed in more detail in Section 4.2 below.

3.3.3. Control variables

Demographic characteristics: The control variables used in this study included family income level, parents’ highest educational level, and parents’ highest professional status. Specifically, the measurement indicators were normalized, and then a principal component analysis was conducted to obtain the index of family SES. After the factor analysis, the final composite index of family SES explained about 90% of the variance variation. This was calculated using the following formula: Family SES index = (parents’ highest educational level × 0.57 + parents’ highest professional status × 0.21 + family income level × 0.12)/0.9. According to the SES index obtained through the principal component analysis, family SES was ranked as high, medium, or low.2

Equipment and network conditions: The families’ SES affects their equipment and network conditions at home and, in turn, affects students’ learning performance when online education became universal during the pandemic. Therefore, as a core independent variable, the equipment and network conditions were also incorporated into the model (separately from the families’ SES) to test the path of its impact on students’ learning performance.

3.3. Definitions of variables

3.3.1. Dependent variables

This study mainly focused on the overall change in high school students’ learning outcomes during the pandemic. The dependent variable of this study was a binary dummy variable; a decrease in overall test scores during the pandemic was equal to 1, whereas no decrease was equal to 0.

3.3.2. Core independent variables

Socioeconomic status (SES): Three indicators were used to measure the core independent variable of families’ SES: family income, parents’ highest educational level, and parents’ highest professional status. Specifically, the measurement indicators were normalized, and then a principal component analysis was conducted to obtain the index of family SES. After the factor analysis, the final composite index of family SES explained about 90% of the variance variation. This was calculated using the following formula: Family SES index = (parents’ highest educational level × 0.57 + parents’ highest professional status × 0.21 + family income level × 0.12)/0.9. According to the SES index obtained through the principal component analysis, family SES was ranked as high, medium, or low.

1 If you are interested in the content of questionnaire, please contact the authors for more details.

2 The classification of family socioeconomic status adopted the criteria for high and low score groups used in psychology; that is, the top 27% were part of the high SES group, whereas the bottom 27% was the low SES group.
study were mainly the demographic characteristics of individuals. They were gender, family location, and whether or not the student was a single child in their family. In existing studies, these variables were considered to have an impact on the digital divide.

**Pre-pandemic class ranking:** Students’ pre-pandemic class ranking determines the extent to which their learning outcomes could improve or worsen during the pandemic. Therefore, the pre-pandemic class ranking was also included as a control variable.

Table 2 shows the names and encoding of each variable, while Table 3 presents descriptive statistics of each variable.

### 4. Results

#### 4.1. The first-level digital divide: equipment and network conditions

This analysis begins with descriptive statistics on the equipment and network conditions of students’ online learning during the pandemic. In terms of equipment conditions, most students (95%) had at least one piece of equipment that could be used for online learning. However, owning just one piece of equipment is unlikely to suffice. Other family members utilizing the equipment during schooling hours, live-streaming software running poorly, and equipment damage can all impact students’ learning. Of all the students surveyed, 18.4% said that they had difficulty with online learning due to inadequate equipment during the pandemic. Regarding network conditions, similar to equipment access, most students reported that they had internet access to online content materials, with only 2.3% not having internet service. However, as many as 37.3% of the students reported that they had been affected by internet problems due to network congestion, insufficient mobile data, or overdue data bills.

To better perceive the differences in equipment and network conditions among students, we prepared statistics by SES, gender, ethnicity, family location, and family type. As shown in Table 4, among students with high SES, only 7.8% were affected by the lack of equipment. In contrast, the rate among students with low SES was 33.6%, 25.8% higher than the former and 15.2% higher than the sample average (18.4%). Within gender groups, the rates were 18.6% for male students and 18.3% for female students; the difference is subtle. In terms of family location, in urban areas, 10.7% of students were affected by the lack of equipment, compared with 23.6% in rural areas, with a considerable difference of 12.9% points. Among single-child families, 11.5% of students were affected by the lack of equipment, whereas the rate for children from non-single-child families was 21.4%, which is 9.9% points higher than the former. Finally, 16.5% of students with good pre-pandemic class rankings were affected by the lack of equipment, which was lower than the rate of 20.6% for students with poor performance rankings. The Chi-square test results showed that there were significant differences in equipment conditions among students with different SES, family location, family type, and pre-pandemic class ranks, except in the case of gender.

As shown in Table 4, the incidence of learning affected by network problems was also negatively correlated with the family’s SES. The lower the family’s SES, the higher the proportion of students affected by network problems. The difference between low (53.3%) and high SES (25.9%) was as high as 27.4% points. In addition, similar to the equipment problem, gender had little influence on students’ network conditions. Generally, network problems were more likely to affect students from rural areas, from non-single-child families, and with low pre-pandemic class rankings.

According to the above results, although most students had ICT equipment and internet access, they still faced difficulties in learning due to the quantity of equipment and the quality of the network. The incidence of learning difficulties was closely related to family SES, family location, and family type. The Chi-square test values indicated that compared with other factors, family SES had the greatest impact on the equipment and network conditions.

#### 4.2. The second-level digital divide: students’ adaptability to online learning during the COVID-19 pandemic

The second-level digital divide relates to how the internet is used. In the existing literature, the second-level digital divide has primarily to do with the difference in usage patterns and skills. The large-scale online education conducted in China during the pandemic had certain particularities. First, all students were using the internet for the same purpose. Students took part in online courses through network platforms to ensure learning when schools were suspended. Second, with fixed course content and teaching time organized by the schools and teachers, students were less likely to encounter difficulties in information retrieval. Online learning during the pandemic could be seen as traditional offline classes that were shifted onto online platforms. Therefore, exploring the differences in students’ daily internet use is outside the scope of this study. To capture students’ online learning performance during the pandemic, the survey required students to self-evaluate their adaptability to online learning on a scale consisting of eight questions (as shown in Table 5). The questions were measured on a five-point Likert scale, with items ranging from “very unsatisfied” to “very
Technology in Society 71 (2022) 102122

Table 6

Table: The third-level digital divide: overall changes in learning outcomes during the pandemic

As shown in Table 6, students’ family backgrounds affect their adaptability to online learning. Based on the standard deviation of all samples (0.68), the adaptability score of students with high SES was 0.46 units of a standard deviation higher than those with low SES. Additionally, males were 0.01 units of a standard deviation higher than females. Urban students were 0.31 units of a standard deviation higher than those from rural areas, and children from single-child families were 0.32 units of a standard deviation higher than those from non-single-child families. Students with good pre-pandemic class rankings were 0.44 units of a standard deviation higher than those with poor rankings. The results of the analysis of variance show that these differences were statistically significant.

4.3. The third-level digital divide: overall changes in learning outcomes during the pandemic

The third-level digital divide refers to the difference in offline benefits people gained from the internet. During the pandemic, although all students attended online courses, the learning outcomes differed among different social groups. Nearly half the high school students (46.95%) reported that their overall learning scores dropped during the pandemic. We will analyze the factors that negatively impacted students’ learning outcomes. Specifically, we constructed a binomial logistic regression model with sequential predictor entry. The dependent variable was a dummy variable of whether learning scores deteriorated. It had the following specification:

\[
\text{Logit}(P_i) = b_0 + b_1\text{SES}_i + b_2X_i + b_3\text{Equipment}_i + b_4\text{Internet}_i + b_5\text{Adaptability}_i
\]

In the above model, SES is the index of family socioeconomic status. X_i represents a series of individual-level control variables, including gender, family location, family type, and pre-pandemic class rankings, all of which are dummy coded. Equipment and the internet are dummy variables that are proxies for the first-level digital divide, whereby the value of 1 indicates that students’ learning had been negatively affected by equipment or network problems. Adaptability is the self-rated score of students’ adaptability to online learning during the pandemic, which is the proxy for the second-level digital divide. In addition, the interaction terms of each control variable with the network, equipment, and students’ adaptability were also incorporated.

A stepwise regression was then adopted for the analysis. Model 1 included the core independent variable of family SES and the control variables. Based on Model 1, Model 2 added three variables (namely equipment conditions, network conditions, and adaptability) to represent the first and second-level digital divide. Models 3, 4, 5, 6, and 7 incorporated the interaction terms. The results of the analysis are shown in Table 7.

As shown in the table above, Model 1 could distinguish those students whose learning performance deteriorated from those whose performance did not (\(\chi^2 = 163.17, p < 0.001\), Nagelkerge Pseudo- \(R^2 = 0.0042\)). In Model 2, the fitting degree of the model significantly improved after the variables representing digital divides were incorporated (\(\chi^2 = 2269.75, p < 0.001\), Nagelkerge Pseudo- \(R^2 = 0.0591\)).

Table 5

Table: Description of students’ self-evaluation scale.

Table 6

Table: Individual learning adaptability during online teaching.
According to the results of Model 1, students’ SES could significantly predict whether their learning outcomes would deteriorate. All other factors being equal, a one-unit increase in SES was associated with a 9.5% reduction in the likelihood of outcome deterioration. Specifically, in terms of outcome deterioration, the likelihood was higher for female students, rural students, students from non-single-child families, and students with high pre-pandemic class rankings (by 4.9, 8.3, 11.6, and 7.3% points, respectively).

According to the results of Model 2, after adding three variables representing the first and second-level digital divides, the coefficients of family SES, family location, and family type decreased significantly, and turned insignificant. This implies that during the pandemic, family background influenced students’ online learning mainly through the first and second-level digital divide. The regression results also showed that network conditions and adaptability could significantly predict changes in learning performance. Holding other predictors constant, students were 13.9% more likely to suffer performance deterioration if they were affected by network problems. Each unit increase in students’ adaptability to online learning was associated with a 58.1% decrease in the probability of outcome deterioration. The effect of equipment conditions on students’ learning outcomes was not significant, possibly because the overall incidence of equipment failure (18.4%) was lower than that of network problems (37.3%). It should be noted that in Model 2, the coefficient of the gender variable was still significant, suggesting that female students are more likely to suffer from outcome deterioration during the pandemic, and this may relate to factors other than the first and second-level digital divide.

The coefficients of interaction items showed that adaptability with gender, family location, family type, and pre-pandemic class rankings were significant. Specifically, the interaction coefficient of adaptability with gender was significantly positive, indicating that if the adaptability of all students improved by one unit at the same time, holding other conditions unchanged, the possibility of female students suffering performance deterioration would be lower than that of male students. Similarly, when other conditions were kept constant if the adaptability of all students improved by one unit, rural students were lower than urban students in the probability of outcome deterioration; this was also true in the case of children from non-single-child families and students with poor pre-pandemic class rankings. In other words, if we could help the traditionally disadvantaged group to better adapt to online learning, they would benefit more. It should also be noted that the interaction term between equipment conditions and pre-pandemic class rankings was significantly negative. This result indicates that when equipment conditions are improved, those with high rankings are less likely to suffer

---

3 The coefficient presented in the regression results can be converted as exp(coefficient) to show the influence of each unit change of independent variable on the incidence of the dependent variable, and the calculation method is \( \text{exp(coefficient) - 1} \times 100\%. \)
performance deterioration, whereas those with low rankings benefit less. In other words, improving equipment conditions while ignoring the process of online learning will probably fail to help students who are left behind.

5. Discussion and conclusions

The outbreak of COVID-19 has accelerated the adoption of online learning in the education arena. Although it ensured learning continuity while school-based education was disrupted, online learning did add new elements to education inequity. Considering that the diffusion of online learning is inevitable in the foreseeable future, it is important to look at the equity issue of online learning carefully.

This study examined the digital divide at three levels in online learning in China during the COVID-19 pandemic. According to the analysis, this study draws the following four conclusions. First, during the pandemic, the first-level digital divide was primarily presented as equipment quantity and network quality problems. The results show that higher SES is associated with better ICT equipment and network. Previous literature mainly focused on whether or not students have internet access [16–18]. However, the access problem has now switched to the quality of internet tools, especially when online learning is conducted in a live-stream manner. Second, students from socially advantaged groups are more adaptable to online teaching during the pandemic. This finding supports existing literature, which found that people differ in their digital skills [28–33]. Therefore, when presented with equal online resources, students’ learning outcomes may vary based on their unequal digital skills. Third, while large-scale online learning has alleviated the negative impact of the pandemic on students’ learning, the offline benefits vary for students from different social backgrounds. In particular, children from advantaged social groups were more likely to maintain their original performance during the pandemic, whereas their disadvantaged peers were more likely to experience outcome deterioration. Family backgrounds affected students’ equipment and network conditions, as well as their adaptability to online learning, which, in turn, made a difference in the offline learning outcome. This finding verifies the concern that technologies are likely to widen the existing inequities [39,40]. Fourth, other factors being equal, the more adaptable the students from traditionally disadvantaged groups were to online teaching, the less likely they were to suffer from outcome deterioration. This finding is quite interesting considering that no studies have examined such a heterogeneous effect.

5.1. Possible mechanisms

The above findings could be explained from several perspectives. First, the survey showed that most high school students (81%) participated in live-streaming online classes during the pandemic. Although live-streaming courses were not confined to a physical location, students could not alter course content and time of learning. This type of online learning is less flexible compared to online courses that are recorded in advance and can be viewed by students at any time. In live-streaming courses, students need to deal with network delays, equipment failure, and other unexpected situations from time to time. Such a situation required more equipment choices and better network quality. Students with higher family SES are less likely to be disturbed by emergencies when taking live-streaming courses.

Second, online learning places a greater emphasis on students’ ability to learn independently. Students from advantaged families probably had more experience with online learning before the pandemic and thus were more attuned to the online teaching methods. In addition, since most students stayed at home during the pandemic, companionship and supervision from parents could also affect students’ learning. Parents with higher social status could provide more guidance regarding equipment operation and self-learning to help their children get through this special period. To some extent, the family environment could be more influential when online learning and self-learning become universal.

Finally, learning gaps during the pandemic may be a reflection and extension of pre-existing inequity. Before the outbreak of COVID-19, there were gaps in learning performance among different groups of students already. Education during the pandemic merely shifted offline schooling into cyberspace. The disadvantaged state of traditionally vulnerable groups remains the same. It means that establishing a nationwide online education system may not solve the problems of education inequity. Traditionally vulnerable groups lack sufficient resources and skills to benefit from online learning. Our research results suggest that if we could help traditionally disadvantaged groups to adapt to online learning, they could gain more compared to their advantaged peers. The diminishing marginal utility of resources may explain this difference. Given that advantaged students already enjoy access to better educational resources, the marginal benefit of additional investment in them is smaller than the marginal benefit of additional investment in disadvantaged students.

5.2. Policy recommendations

Based on the above, this study puts forward the following recommendations.

First, for better online teaching in the future, educators should bridge the first-level digital divide. Before carrying out school-wide online teaching, schools must ensure that all students have good equipment and network conditions through equipment donation, financial aid, and encouraging mutual aid among students so that all students can begin their learning at the same level.

Second, efforts should be made to improve students’ digital literacy, especially those from disadvantaged groups. Schools can organize regular training for students who are not accustomed to online learning. This measure will not only improve their internet skills in the short term but will also help them master the necessary skills for lifelong learning, laying a good foundation for their long-term development.

Third, once the schools are reopened, and offline schooling resumes, schools need to provide customized support targeting disadvantaged students. For students with significant outcome deterioration, schools must build communication and cooperation with them and their families to understand their needs and introduce assistance programs. Students could be paired in such a way that those with better learning performance can assist others who are experiencing difficulties. Additionally, teachers of all subjects should pay close attention to students from disadvantaged families and provide them with extracurricular help and guidance.

In summary, while ICT technology has the potential to narrow educational gaps, its development alone cannot eliminate education inequity. If educators use the internet to share educational resources universally while ignoring pre-existing inequalities, the digital divide will persist, and the offline benefits will benefit people from the upper social class disproportionately. The promotion of education equity requires efforts from various stakeholders. The government could establish an alliance between advantaged and disadvantaged schools. These schools could share their lectures and teaching resources through shared platforms. Online resources could be used by individual schools to provide tutoring for disadvantaged students during after-school hours or holidays. Non-profit organizations can actively promote educational innovation by integrating new technology into educational practices and channeling high-quality resources to vulnerable groups at a lower cost and more efficient way.

5.3. Limitations

The following limitations are still to be addressed. As mentioned above, in addition to hardware and network conditions, many other factors could affect students’ online learning, such as parental support...
and supervision, students’ self-management capability, and their former experiences with online learning. Except for students and their families, schools and communities could also play an important role in online learning during the pandemic [49,50]. Because of the data limitation, this study could not conduct a more detailed examination of these factors and their effects.

In addition, this study collected data through questionnaires. The answers from students may not be accurate enough, thereby compromising the internal validity of our research. We distributed questionnaires on a large scale. This method of collecting data ignored the effects of these interventions should be examined thoroughly. Future research can focus on how to design online learning programs and conduct these programs effectively.

6. Future research perspectives

In the future, both qualitative and quantitative methods can be used to fully explore the mechanism that influences students’ online learning in different situations. In addition to students, the influencing factors may also involve schools, teachers, and communities. In different online learning scenarios, the effect of these factors may be different. Only by fully understanding these problems can educators improve the effect of online learning and increase education equity.

Besides, future research can explore how to improve the learning outcomes of low-income students through online learning programs. Our research shows that when policymakers equally distribute education resources, education inequity will not be closed. Therefore, interventions targeting disadvantaged students should be designed and the effects of these interventions should be examined thoroughly. Future research can focus on how to design online learning programs and conduct these programs effectively.

Author statement

Congbin Guo: Writing- Reviewing and Editing, Methodology, Project Administration. Boshen Wan: Data curation, Writing-Original draft preparation, Formal Analysis. Funding

This work was funded by the Center for Data Science, Peking University, under grant number 20202DB02.

Declaration of competing interest

We declare that we do not have any competing interest in connection with this article.

Data availability

The data that has been used is confidential.

Acknowledgements

We thank the Center for Data Science, Peking University for their financial support.

References

[1] N. Balacheff, J.J. Kaput, Computer-based learning environments in mathematics [M], in: International Handbook of Mathematics Education, Springer, Dordrecht, 1996, pp. 469–501.
[2] K. Lynch, J.S. Kim, Effects of a summer mathematics intervention for low-income children: a randomized experiment[J], Educ. Eval. Pol. Anal. 39 (1) (2017) 31–53.
[3] S. Baum, M. McPherson, The human factor: the promise & limits of online education[J], Daedalus 148 (4) (2019) 235–254.
[4] P. DiMaggio, E. Hargittai, From the ‘digital divide’ to ‘digital inequality’: studying internet use as penetration increases[J], Princeton, Center for Arts and Cultural Policy Studies, Woodrow Wilson School, Princeton University 4 (1) (2001), 4–2.
[5] C. Azorin, Beyond COVID-19 Supernova. Is Another Education coming?[J], Journal of Professional Capital and Community, 2020.
[6] M. Warschauer, M. Knobel, L. Stone, Technology and equity in schooling: deconstructing the digital divide[J], Educ. Pol. 18 (4) (2004) 562–588.
[7] Organisation for Economic Co-operation and Development, The State of Global Education: 18 Months into the pandemic[M], OECD Publishing, 2021.
[8] C. McBurnie, T. Adam, T. Kaye, Is there learning continuity during the COVID-19 pandemic? A synthesis of the emerging evidence[J], Journal of Learning for Development 7 (3) (2020) 485–493.
[9] U. Reimers, A. Schleicher, A framework to guide an education response to the COVID-19 pandemic of 2020[M], OECD 14 (2020), 2020: 2004-04.
[10] S. Dhawan, Online learning: a panacea in the time of COVID-19 crisis[J], J. Educ. Technol. Syst. 49 (1) (2020) 5–22.
[11] P. DiMaggio, E. Hargittai, C. Celeste, et al., Digital Inequality: from Unequal Access to Differentiated use[M], Social Inequality, Russell Sage Foundation, 2004, pp. 355–400.
[12] J.A.G.M. Van Dijk, The Deepening Divide: Inequality in the Information society [M], Sage Publications, 2005.
[13] R. Belock, D.V. Dimitrova, An exploratory model of inter-country internet diffusion[J], Telecommun. Pol. 27 (3–4) (2003) 237–252.
[14] M.F. Guilen, S.L. Suárez, Explaining the global digital divide: economic, political and sociological drivers of cross-national internet use[J], Soc. Forces 64 (2) (2005) 681–708.
[15] G.D.M. Wijers, Determinants of the digital divide: a study on IT development in Cambodia[J], Technol. Soc. 32 (4) (2010) 336–341.
[16] D.J. Hoffman, T.P. Novak, Bridging the Racial Divide on the internet[J], 1998, 1998.
[17] P. Attewell, Comment: the first and second digital divides[J], Soc. Educ. 74 (3) (2001) 252–259.
[18] K.R. Wilson, J.S. Wallin, C. Roiter, Social stratification and the digital divide[J], Soc. Sci. Comput. Rev. 21 (2) (2003) 133–143.
[19] F. Cruz-Jesus, M.R. Vicente, F. Baco, et al., The education-related digital divide: an analysis for the EU-28[J], Comput. Hum. Behav. 56 (2016) 72–82.
[20] N. Zillien, E. Hargittai, Digital distinction: status-specific types of internet usage[J], Soc. Sci. Q. 90 (2) (2009) 274–291.
[21] A.J.A.M. Van Deursen, J.A.G.M. Van Dijk, The digital divide shifts to differences in usage[J], New Media Soc. 16 (3) (2014) 507–526.
[22] G. Blask, D. Groselj, Dimensions of internet use: amount, variety, and types[J], J. Information, Communication & Society 17 (4) (2014) 417–433.
[23] M. Micheli, Social networking sites and low-income teenagers: between opportunity and inequality, J. Information, Communication & Society 19 (5) (2016) 565–581.
[24] J.K.H. Ma, T.E. Vachon, S. Cheng, National income, political freedom, and investments in R&D and education: a comparative analysis of the second digital divide among 15-year-old students[J], Soc. Indicat. Res. 144 (1) (2019) 133–166.
[25] J.D. Lockesby, H. Li, From the editors why we need the journal of interactive advertising[J], J. Interact. Advert. 1 (1) (2000) 1–5.
[26] A. Deursen, J.A.G.M. Van Dijk, Measuring internet skills[J], Int. J. Hum. Comput. Interact. 26 (10) (2010) 891–916.
[27] A. Ferrari, Digital Competence in Practice: an Analysis of frameworks[J], JRC IPTS, Sevilla, 2012, https://doi.org/10.2791/82116.
[28] E.J. Helprer, R. Eynon, Distinct skill pathways to digital engagement[J], Eur. J. Commun. 28 (6) (2013) 696–713.
[29] A.J.A.M. Van Deursen, E.J. Helprer, R. Eynon, Development and validation of the internet skills scale (ISS)[J], Information, Commun. Soc. 19 (6) (2016) 804–823.
[30] A.J.A.M. Van Deursen, J.A.G.M. van Dijk, Internet skill levels increase, but gaps widen: a longitudinal cross-sectional analysis (2010–2013) among the Dutch population, J. Information, Communication & Society 18 (7) (2015) 782–797.
[31] A.J.A.M. Van Deursen, E.J. Helprer, Collateral benefits of internet use: exploring the diverse outcomes of engaging with the internet[J], New Media Soc. 20 (7) (2018) 2333–2351.
[32] A.J.A.M. Van Deursen, E.J. Helprer, The Third-level Digital Divide: Who Benefits Most from Being online?[M], Communication and Information Technologies Annual, Emerald Group Publishing Limited, 2015.
[33] J.C. Witte, S.E. Mannos, The Internet and Social Inequalities [M], Routledge, 2010.
[34] D. Calderon Gomez, The third digital divide and Bourdieu: bidirectional conversion of economic, cultural, and social capital to (and from) digital capital among young people in Madrid[J], New Media Soc. 23 (9) (2021) 2534–2553.
[35] M. Wang, G.T. Xu, Y.Q. Ren, Promoting equity in education through ICT: a double-edged sword?[J], J. Int. Education 43 (2014) 38–49 (in Chinese).
[36] E.J. Helprer, A corresponding fields model for the links between social and digital exclusion[J], Commun. Theor. 22 (4) (2012) 403–426.
[37] A. Scheerder, A. van Deursen, J. van Dijk, Determinants of internet uses, skills and outcomes: a systematic review of the second-and third-level digital divide[J], Telematics Inf. 34 (8) (2017) 1607–1624.
[38] R.M. Anderson, H. Heeterbeek, D. Klinkenberg, T.D. Hollingsworth, How will country-based mitigation measures influence the course of the COVID-19 epidemic?[J], Lancet 395 (10230) (2020) Mar 21 931–934.
[39] H.R. Zhang, J.Y. Wu, Y. Liu, N.Q. Song, The digital divide between urban and rural basic education in China: characterization, causes and countermeasures—an empirical study on online teaching[J], 04, Education & Economy 37 (2021) 20–28 (in Chinese).
[40] X.H. Yang, H j Zhou, X. Zheng, Characteristics and enlightenment of online teaching in colleges and universities during the epidemic: an analysis based on online teaching quality reports and online teaching cases[J], E-Education Research 42 (10) (2021) 48–53 & 62 (in Chinese).
[41] S. Chen, J.W. Liang, Y. Yu, Y.Q. Sun, R.L. Wang, X. Qu, S. Li, Research on online teaching platforms, tools, and application effect during epidemic period[J], 05, China Educational Technology (2020) 44–52 (in Chinese).

[42] Y.R. Xie, Y. Qiu, Y.L. Huang, Q.L. Wang, Characteristics, problems and innovations of online teaching of “No Suspension of Classes” during the period of epidemic prevention and control [J], 03, E-education Research 41 (2020) 20–28 (in Chinese).

[43] J.X. Wang, Y.T. Wei, M. Zong, The current situation, problems and reflection of online teaching for primary and secondary school teachers during a large-scale epidemic—based on the investigation and analysis of “classes suspended but learning continue” in Hubei province[J], 05, China Educational Technology (2020) 15–21 (in Chinese).

[44] K. Wan, X.D. Zheng, Y.Q. Ren, Does large-scale online learning prepare: the thinking of online intelligent learning in the post-epidemic period[J], 03, J. Distance Educ. 38 (2020) 105–112 (in Chinese).

[45] W.P. Lei, X.Z. Huang, Problems in the development of online education under the COVID-19 epidemic[J], 02, Educ. Res. Exp. 38 (2020) 13–16 (in Chinese).

[46] J. Guo, The evolution of digital divide: from access to engagement—based on the undergraduate online learning survey during COVID-19[J], 07, J. East China Normal Univ. Educ. Sci. 39 (2021) 16–26 (in Chinese).

[47] E.L. Abramson, C.R. Paul, J. Peterslack, et al., Conducting quantitative medical education research: from design to dissemination[J], Academic Pediatrics 18 (2) (2018) 129–139,

[48] D.C. Watkins, Qualitative research: the importance of conducting research that doesn’t “count” [J], Health Promot. Pract. 13 (2) (2012) 153–158,

[49] E. Gandolfi, R.E. Ferdig, A. Kratcoski, A new educational normal an intersectionality-led exploration of education, learning technologies, and diversity during COVID-19[J], Technol. Soc. 66 (2021), 101637,

[50] M.A. Camilleri, A.C. Camilleri, Remote learning via video conferencing technologies: implications for research and practice, J. Technology in society 68 (2022), 101881,