Muslim Education Review
Vol. 1 No. 1, 2022, 59-84
https://journal.uiii.ac.id/index.php/mer

The Nexus between Education and Internet Use of Students: Evidence from Underdeveloped Regions in Indonesia

Taly Purwa[1], Diyang Gita Cendekia[2]
[1]Badan Pusat Statistik (BPS)-Statistics Indonesia of Bali Province
[2]Badan Pusat Statistik (BPS)-Statistics Indonesia of Kalimantan Selatan
Corresponding E-mail: taly@bps.go.id

Abstract

During the COVID-19 pandemic and the Industry 4.0 era, the role of the Internet became extremely important for connecting the society. Unfortunately, heterogeneous geographical, socioeconomic and demographic characteristics may create different roles in using the Internet, leading to a digital divide. Utilizing National Socioeconomic Survey (Susenas) data collected early in the COVID-19 pandemic, this study employs binary logistic regression to investigate the effect of education through school participation on internet use in underdeveloped regions in Indonesia. The findings show that only one-fifth of students in underdeveloped regions are using the Internet. Looking deeper, school participation plays a prominent role for students online. The more educated the students, the more likely they are to use the Internet. Moreover, the possibility of a student using the Internet is increasing for students are getting the aid of the Program Indonesia Pintar (PIP), who live in households where the head of the household has particular characteristics, which are being female, of non-productive age, having higher education, working in the non-agricultural sector, having higher socioeconomic status and where fewer students live in the household. However, this study also finds that student gender has no significant impact on internet use. Promoting and providing proportional support by the government in terms of internet use based on school participation is principal due to the existence of the digital divide. It will also be very interesting when further research may account for other potential variables from the supply side that could explain the internet use of students in underdeveloped regions of Indonesia.

Keywords: *internet activity, logistic regression, online, pandemic, school participation*

Introduction

Contribution to education is still a basic requirement for Indonesia in achieving the demographic dividend (Sulisworo, 2016). In addition, education gives the strength to meet the sustainable development goals (SDGs) which are agreed upon by almost all countries in the world. Nowadays, population quality has been increasingly disrupted since COVID-19 was declared as a global pandemic (Cucinotta & Vanelli, 2020), not only in regard to population health but also in regard to education. Currently, the global education system is facing its greatest challenge during the pandemic (United Nations, 2020) in that social restrictions have led to school closures and switching from offline learning to online distance learning. This means the Internet plays a recent important role in the implementation of educational activities. Unfortunately, this new learning method highly relies on internet access and network coverage. The high variation in these two aspects could potentially widen the education gap in Indonesia (Alifia, 2020).

Since Indonesia entered the Industry 4.0 era which was marked by "the internet of things" (Lampropoulos et al., 2019), the Internet has become the most valuable resource in the world (Dahiya et al., 2021). This is also true when many others in the world live in a digital society. Actually, before the pandemic of COVID-19, the Internet existed and was popular in many segments of the Indonesian population. Moreover, many aspects of life are highly reliant on the Internet. However, heterogeneous geographical, socioeconomic and demographic characteristics may create different possibilities in using the Internet and then lead to a digital divide (Smith & Graham, 2012), especially in underdeveloped regions. Consequently, this can bring about inequality in each region. Even so, the facilities and the capability of using the Internet could prevent education in underdeveloped regions from getting worse (Arkiang, 2021).

The Internet has created many opportunities and grown rapidly (Barua et al., 2000); the Internet makes life easier. Various studies have been conducted to observe internet use in Indonesia (Eschachasthi et al., 2022; Purwa & Cendekia, 2021; Puspitasari & Ishii, 2016; Wahid, 2007). A higher education level is correlated to internet use. More educated people have the potential to adopt the technology faster than people with lower education attainment (Vodoz et al., 2007) and they also tend to use the Internet to achieve prosperity. Therefore, education is the most important factor that influences internet use (Al-Hammadany & Heshmati, 2011). Many studies have investigated the relationship between education and internet use. Mostly, previous studies have stated that education has a positive correlation with internet use (Singh, 2004; Cooke & Greenwood, 2008; Noce & McKeown, 2008; Al-Hammadany & Heshmati, 2011; Lera-López et al., 2011; Pénard et al., 2012; Pick & Nishida, 2015; Mubarak et al., 2020). Qomariyah (2009) found that students in urban areas begin to use the Internet when entering Junior High School. On the contrary, Martin and Robinson (2007), and Middleton and Chambers (2010) found that education has no clear correlation with internet use. A study which examines internet use among students in underdeveloped regions is also very limited. Recently, as far as author's knowledge, only Arkiang (2021) who has investigated this topic.

This study fills the aforementioned research gap by observing the internet use of students from underdeveloped regions in Indonesia. As education escalates the quality of human life, which is stated in SDG 4, this study investigates the impact of education, through school participation from elementary school to high school, on internet use among students from underdeveloped regions in Indonesia in 2020. Several previous studies that discussed internet use mostly focused on factors relating to signals, facilities and infrastructure of Information Communications Technology (ICT) (Usluel et al., 2008; Chiao & Chiu, 2018; Dahiya et al., 2021). Taking this further, this study also investigates the impact of socioeconomic and demographic variables including student and household characteristics, on internet use among students, since these characteristics determine the existence of a digital divide (Hoffman & Novak, 1998; Hargittai, 1999; Dabla, 2004; Yu, 2011; Smith & Graham, 2012; Van Deursen et al., 2015). The findings obtained from this study are expected to add to information about the level of internet use of students and provide empirical results relating to the determinant of internet use of students, especially from underdeveloped regions in Indonesia at the early stage of the COVID-19 pandemic. Therefore, this study contributes to identifying important individual and household characteristics that can be used to narrow the digital divide for students during the pandemic, especially from underdeveloped regions in Indonesia.

Method

This study utilizes raw data from the National Socioeconomic Survey (Susenas) that was held in March 2020 by Statistics Indonesia (Badan Pusat Statistik). The unit of analysis is students who attend elementary, junior and senior high school in regencies or cities described as underdeveloped regions in Indonesia. Those regions accord to the Indonesian Republic Presidential Regulation No. 63, 2020 concerning the determination of underdeveloped regions in 2020–2024. There are 62 regencies or municipalities that are stated as underdeveloped regions as follows:

Table 1. *List of Underdeveloped Regions, 2020*

No	Province	Regency/ Municipality	No	Province	Regency/ Municipality	No	Province	Regency/ Municipality
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
1	Sumatera Utara	Nias	22	Sulawesi Tengah	Donggala	43	Maluku	Paniai
2	Sumatera Utara	Nias Selatan	23	Sulawesi Tengah	Tojo Una-una	44	Maluku	Puncak Jaya
3	Sumatera Utara	Nias Utara	24	Sulawesi Tengah	Sigi	45	Maluku	Boven Digoel
4	Sumatera Utara	Nias Barat	25	Maluku	Maluku Tenggara Barat	46	Maluku	Mappi
5	Sumatera Barat	Kepulauan Mentawai	26	Maluku	Kepulauan Aru	47	Maluku	Asmat
6	Sumatera Selatan	Musi Rawas Utara	27	Maluku	Seram Bagian Barat	48	Maluku	Yahukimo
7	Lampung	Pesisir Barat	28	Maluku	Seram Bagian Timur	49	Maluku	Pegunungan Bintang
8	Nusa Tenggara Barat	Lombok Utara	29	Nusa Tenggara Timur	Maluku Barat Daya	50	Papua	Tolikara
9	Nusa Tenggara Timur	Sumba Barat	30	Nusa Tenggara Timur	Buru Selatan	51	Papua	Keerom
10	Nusa Tenggara Timur	Sumba Timur	31	Maluku Utara	Kepulauan Sula	52	Papua	Waropen
11	Nusa Tenggara Timur	Kupang	32	Maluku Utara	Pulau Taliabu	53	Papua	Supiori
12	Nusa Tenggara Timur	Timor Tengah Selatan	33	Papua Barat	Teluk Wondama	54	Papua	Mamberamo Raya
13	Nusa Tenggara Timur	Belu	34	Papua Barat	Teluk Bintuni	55	Papua	Nduga
14	Nusa Tenggara Timur	Alor	35	Papua Barat	Sorong Selatan	56	Papua	Lanny Jaya
15	Nusa Tenggara Timur	Lembata	36	Papua Barat	Sorong	57	Papua	Mamberamo Tengah

No	Province	Regency/ Municipality	No	Province	Regency/ Municipality	No	Province	Regency/ Municipality
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
16	Nusa Tenggara Timur	Rote Ndao	37	Papua Barat	Tambrauw	58	Papua	Yalimo
17	Nusa Tenggara Timur	Sumba Tengah	38	Papua Barat	Maybrat	59	Papua	Puncak
18	Nusa Tenggara Timur	Sumba Barat Daya	39	Papua Barat	Manokwari Selatan	60	Papua	Dogiyai
19	Nusa Tenggara Timur	Manggarai Timur	40	Papua Barat	Pegunungan Arfak	61	Papua	Intan Jaya
20	Nusa Tenggara Timur	Sabu Raijua	41	Papua	Jayawijaya	62	Papua	Deiyai
21	Nusa Tenggara Timur	Malaka	42	Papua	Nabire			

There are 34.621 students from 18.322 households in the sample of Susenas March 2020 who were involved in the analysis. According to Badan Pusat Statistik (2020), a household is a person or group of people who inhabit part or all of a physical or census building and, usually, live together and eat from one kitchen or consume certain types of goods and services collectively. A household head is the one who is responsible for the daily needs of the household. The variables used in this study, including the variables that related to the student characteristics, with student school participation as a main predictor variable, and also household characteristics, are as follows:

Table 2. *Operational Definition and Type of Variables*

Variable	Notation	Description
(1)	(2)	(3)
Response Variable		
Internet use	*Inet*	Status of internet use in the past 3 months Categories: 0 – not using the Internet 1 – using the Internet
Main Predictor Variable – Student Characteristic		
School participation	*school*	School participation refers to the presence of students in formal and non-formal educational activities

Variable	Notation	Description
(1)	(2)	(3)
Response Variable		
		Categories: 0 – elementary school 1 – primary school 2 – high school
Control Variable – Student Characteristic		
Gender	*gend*	Categories: 0 – Female 1 – Male
Program Indonesia Pintar (PIP)	*PIP*	Getting aid from Program Indonesia Pintar (PIP) a year ago. Categories: 0 – no 1 – yes
Control Variable – Household Characteristic		
Household head gender	*gendHH*	Categories: 0 – Female 1 – Male
Household head age	*ageHH*	Categories: 0 – >64 years old 1 – 15-64 years old
Household head education attainment	*educHH*	Categories: 0 – not completed elementary school or completed elementary school 1 – completed junior high school or higher
Household head occupation	*workHH*	Occupation status in a week. Categories: 0 – not working or working in the agricultural sector 1 – working in a non-agricultural sector
Number of students	*nstudHH*	Number of students in the household in which a student lives (numeric).
Socioeconomic status	*sesHH*	Expenditure quantile of the household in which a student lives. Categories: 0 – first 40% 1 – second 40% 2 – top 20%

Note: Dummy 0 is a reference category.

First, the descriptive analysis for each variable is performed by using tables and/or graphs. Second, the inferential analysis using logistic regression is utilized to find the significant predictor variables that affect students in using the Internet by calculating the odds ratio for each variable. The logistic regression model with the logit link function is as follows (Agresti, 2007; Hosmer & Lemeshow, 2000):

$$ln\left(\frac{\pi(x)}{1-\pi(x)}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p \quad (1)$$

where $\pi(x)$ and $1-\pi(x)$ are the probability of a student using the Internet or not using the Internet in the past three months, respectively. Predictor variables are x_1, x_2,…, and x_p and β_0, β_1, β_2,…, and β_p are parameters that need to be estimated and statistically significance-tested. Since the number of samples for each class of response variable are considered as imbalance, the number of students in one class (majority class) is more than in the other class (minority class), this study utilizes the robustness check of parameter estimation by using the under-sampling scheme (King & Zeng, 2003; Lunardon et al., 2014) that is repeated for n times, with n =100, 1000,… and so on. In the under-sampling scheme, the number of students in the majority class is reduced by sampling randomly as many as the number of students in the minority class so that both classes have the same number of samples. According to King and Zeng (2003), the imbalance case in logistic regression analysis could underestimate probability of minority class or rare event. That condition also could lead to bias and higher variability of estimated parameters (Salas-Eljatib et al., 2018). To produce the best model, variable selection using backward elimination and the likelihood ratio test (LRT) are performed as described by Zhang (2016). Last, the odds ratio (OR) for each predictor, $e^{\beta_i x_i}$, $i = 1,2, \ldots, p$, is interpreted. The OR indicates the comparison between odds of outcome of interest given specific treatment or category of factor and odds of outcome of interest given the absence of specific treatment or category of factor (Szumilas, 2010).

Results and Discussion

Education faced major challenges during the COVID-19 pandemic, for example, when the offline learning method was switched to online distance learning. Although technology through the Internet cannot replace the role of the teacher, the Internet can facilitate the learning process during the pandemic. Access to the Internet certainly differs based on the development levels of the region, that is, developed, developing and underdeveloped. There is a disparity in internet connectivity between the western, central and eastern regions of Indonesia (Alifia et al., 2020; The Jakarta Post, 2020). Note that most of the underdeveloped regions in Indonesia are located in the eastern part of the country where, according to

Statistics Indonesia (2021b), most of these regions have the lowest percentage of internet users. Hence it is interesting to explore internet use in underdeveloped regions through this study.

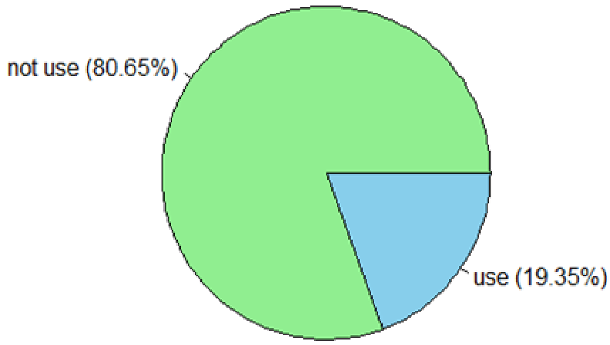


Figure 1. *Proportion of students in internet activity.*

Generally, before the COVID-19 pandemic students could access the Internet from many places, including home, school and other public places. However, the previous study conducted by Arkiang (2021) highlighted that many students in underdeveloped regions face difficulties online, especially for the learning process. In fact, this study finds that most students in Indonesia's underdeveloped regions are not online or do not use the Internet. There is only about 19 percent of students who use the Internet (Figure 1). Even if there are many student not online, they probably still need internet access (Dahiya et al., 2021). Moreover, education programs could fail if the government doesn't fulfill the requirement for availability of computers and internet access (Dahiya et al., 2021).

More detail, based on the Susenas sample in March, 2020, of the distribution of internet usage of students by student and household characteristics are presented in Table 3. For the school participation variable, the percentage of internet users in underdeveloped regions increases for higher school levels with the maximum percentage being 53.68 percent for students in high school. This finding indicates that the higher the education participation the more likelihood of using the Internet. Regarding gender and the acceptance of Program Indonesia Pintar (PIP) assistance of students, the percentage of internet users for each category is quite similar, that is, about 18 to 20 percent. The percentage of internet users for female students is slightly higher than for male students. While In regard to the acceptance of PIP assistance, the percentage of internet users for students who received PIP assistance was slightly higher than for students who did not receive PIP assistance.

Table 3. *Characteristics of Students in Underdeveloped Regions (%)*

Variable		Not Use Internet	Use Internet	Total
(1)		(2)	(3)	(4)
Student Characteristics				
School participation	High school	46.32	53.68	100.00
	Junior high school	74.02	25.98	100.00
	Elementary school	92.64	7.36	100.00
Gender	Male	81.12	18.88	100.00
	Female	80.12	19.88	100.00
Program Indonesia Pintar (PIP)	Yes	79.89	20.11	100.00
	No	80.88	19.12	100.00
Household Characteristics				
Household head gender	Male	80.99	19.01	100.00
	Female	77.41	22.59	100.00
Household head age	15–64 years old	80.70	19.30	100.00
	64 years old	79.77	20.23	100.00
Household head education attainment	Completed junior high school or higher	76.76	23.24	100.00
	Not completed elementary school or completed elementary school	86.16	13.84	100.00
Household head occupation	Non-agricultural sector	68.04	31.96	100.00
	Not working or working in agricultural sector	86.59	13.41	100.00
Socioeconomic status	Top 20%	77.26	22.74	100.00
	Second 40%	80.23	19.77	100.00
	First 40%	82.60	17.40	100.00

Source: Calculated by authors from Susenas March 2020.

In the head of household characteristics, gender and age also have a similar percentage of internet users for each category, that ranges between 19 to 22 percent. In more detail, the percentage of internet users for students who live with female household heads is slightly higher compared to the percentage of internet users for students who live with male household heads. The percentage of internet users for students who live with household heads of non-productive age (65 years and over) is also slightly higher compared to students who live with productive age (15-64 years) household heads. In contrast, significant differences in the percentage of internet users for students in each category are depicted in variables of household head education attainment, household head occupation, and socioeconomic status. A higher percentage of internet users for students is found in households with heads who have higher education attainment (completed junior high school or higher) and work in the non-agricultural sector. A higher percentage of internet users for students was also found in households with higher per capita expenditure categories that reflect household socioeconomic status.

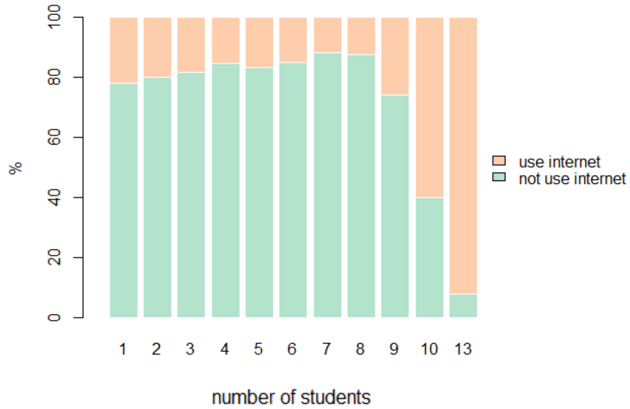


Figure 2. *Percentage of internet activity of students by the number of students in the household where the student lives.*

The only numeric variable in this study is the number of students in the household where the student lives. As presented in Figure 2, two different patterns emerge in the percentage of students that use the Internet as the number of students in the household increases. First, the decreasing pattern in the percentage of students that use the Internet appears from 1 to 8 student/s in the household. While from 9 to 13 students, the percentage of students that use the Internet increases significantly; starting from about 20 percent for 9 students in the household and becoming about 90 percent for 13 students in the household. Please noted that the household that contains 9 and more students is sparse which is only 50 households from total of 34.621 households in this study and it might be classified as outliers since have different pattern. Unfortunately, it couldn't be neglected in this study.

Estimation of Logistic Regression Models

The aim of the study is to investigate the effect of school participation on the internet use of students. Student school participation is the main response variable which is categorized into three categories, that is, elementary school, junior high school and high school, with the former category as a reference category. In order to analyze, this study uses three models of estimation (see Table 4). First, the estimation in Model 1 incorporated student school participation without any control variables. The result shows that school participation affects the internet use of students statistically significantly.

Next, the control variables are incorporated in the model. The estimation result in Model 2 shows that all these variables are statistically significant for α =0.05, except the variable of student gender. Hence, the backward elimination process is performed by excluding the insignificant variable and re-estimating the model. The result shows that the remaining variables in the model are already statistically significant, as presented in Model 3. Note that the parameter estimates in Model 1, Model 2, and Model 3 are very similar as indications of robustness, even after including and excluding other variables. The results of the likelihood ratio test in Appendix 1 show that in terms of fit to the data, Model 2 and Model 3 are significantly different compared to Model 1. While Model 2 and Model 3 are not significantly different indicates that both models are similar, hence the variable of student gender is negligible. Therefore Model 3 is chosen as the best model showing more parsimony or having fewer variables with the formula as follows:

$$\begin{aligned} ln\left(\frac{\pi(x)}{1-\pi(x)}\right) = & -2.6605 + 1.6096school_1 + 2.8422school_2 + 0.1230PIP_1 \\ & - 0.1672gendHH_1 - 0.2197ageHH_1 + 0.4046educHH_1 \\ & + 1.2166workHH_1 - 0.1895nstudHH + 0.1113sesHH_1 \\ & + 0.2769sesHH_2 \end{aligned} \quad (2)$$

As described in the previous subsection, only about one fifth of the total student population uses the Internet. This indicates there is a case of imbalance class and the estimation results are strongly influenced by this imbalance condition. To check the robustness of the estimation result, the parameter estimation in Model 3 is compared with the distribution of parameter estimation from the under-sampling scheme that is repeated for 1000 times. The result in Appendix 2 shows that all parameter estimates (red vertical lines) lay in the distribution range of parameter estimates of the repeated under-sampling scheme and that the range of parameter estimates does not contain both positive and negative values that could lead to two different interpretations. So the estimation result of Model 3 is considered to be relatively robust and could be interpreted further.

Effect of School Participation

Based on Table 4, student school participation has a significant impact on the internet use of students (with and without control variables). This means that each level of school participation has different patterns of using the Internet. As presented in Table 4, the odds ratio of student school participation in junior high school and high school are 5.0010 and 17.1540, respectively. In other words, the odds that a student uses the Internet are about 5 times higher for students in junior high school and 17 times higher for students in high school compared to students with school participation only in elementary school. In summary, a student with a higher level of education is more likely to use the Internet. This finding is in line with results from Cooke and Greenwood (2008), Smith and Graham (2012), Pick and Nishida (2015), Mubarak et al. (2020) and Shavkun et al. (2021). Moreover, Vodoz et al. (2007) also highlight that Internet usage possibilities increase for people with higher education.

Interestingly, this study finds that school participation has the highest odds ratio and Wald value. It indicates that education through school participation has the greatest effect on the internet use of students in underdeveloped regions. In contrast, several previous studies have stated that the important variable that affects internet use is the socioeconomic variable (Chiao & Chiu, 2018; Martínez-Domínguez & Fierros-González, 2022). Note that the coverage of those previous studies are not students in underdeveloped regions.

Generally, education brings many advantages in the future for many aspects. Consequently, education plays a prominent role in population quality. Through education, individual capabilities will increase which leads to increased productivity in the labor market. Further, an increase in productivity will increase output in the labor market and the receipt of higher income. Moreover, in an economic view, education highlights human capital investment even though it usually places a heavy financial burden on the household. More-educated people tend to create many ways to reach prosperity, especially by utilizing technology.

Van Deursen et al. (2015) agreed education plays an important role in the type of online activities. Alderete (2019) reveals that higher education levels deliver high digital competencies for ICT use. With a closer look, when students participate in higher education, they have more information on how to use, and on the greater benefits of using, the Internet. A higher education level is required by the higher personal skill of using the Internet (Rice & Katz, 2003; Al-Hammadany & Heshmati, 2011; Lera-López et al., 2011; Martínez-Domínguez & Fierros-González, 2022). For example, websites usually appear in English and this could be uncomfortable for lower educated people (Elena-Bucea et al., 2021).

Effect of Other Student Characteristics

Previous studies agreed that internet usage patterns differ between males and females (Wahid, 2007; Willoughby, 2008; Smith & Graham, 2012; Chiao & Chiu, 2018). Females tend to use the Internet for chatting and study-related matters, while males tend to use the Internet for searching for online information, reading online news, testing and downloading software, shopping, entertainment, seeking job vacancies and so on. On the other hand a study by Smith and Graham (2012) found that females tend to have a higher possibility of using the Internet compared to males. Conversely, this study finds that the gender of students has no significant effect on internet use in underdeveloped regions. This means that males and females tend to have no different patterns in using the Internet. Furthermore, it indicates that there is no significant issue of the gender gap in terms of the ability of accessing the Internet.

Table 4. *Estimation Results and Odds Ratio from Logistic Regression Models*

Variable		Model 1			Model 2			Model 3		
		Estimate	Wald	Odds Ratio	Estimate	Wald	Odds Ratio	Estimate	Wald	Odds Ratio
Intercept		-2.5326***	9156.841	0.0794	-2.6406***	857.773	0.0713	-2.6605***	900.591	0.0699
School participation	High school	2.6801***	5148.531	4.4181	2.8417***	5030.959	5.0008	2.8422***	5033.204	5.0010
	Junior high school	1.4857***	1624.016	14.5858	1.6096***	1730.377	17.1443	1.6096***	1730.652	17.1540
	Elementary school	reference category								
Gender	*Male*				-0.0380	1.453	0.9627			
	Female	reference category								
Program Indonesia Pintar (PIP)	Yes				0.1230***	11.186	1.1309	0.1230***	11.196	1.1309
	No	reference category								
Household head gender	*Male*				-0.1663**	10.327	0.8468	-0.1672**	10.449	0.8460
	Female	reference category								
Household head age	Productive age				-0.2196***	9.950	0.8028	-0.2197***	9.961	0.8028
	Non-productive age	reference category								
Household head education attainment	Completed junior high school or higher				0.4042**	137.249	1.4981	0.4046**	137.566	1.4987
	Not completed elementary school or completed elementary school	reference category								

Variable		Model 1			Model 2			Model 3		
		Estimate	Wald	Odds Ratio	Estimate	Wald	Odds Ratio	Estimate	Wald	Odds Ratio
Household head occupation	Non-agricultural sectors				1.2165**	1295.830	3.3755	1.2166**	1296.169	3.3759
	Not working or working in the agricultural sector	reference category								
Number of Students					-0.1898***	220.904	0.8271	-0.1895***	220.342	0.8273
Socioeconomic status	Top 20%				0.1113***	9.849	1.1177	0.1113***	9.850	1.1177
	Second 40%				0.2773***	41.988	1.3196	0.2769***	41.877	1.3191
	First 40%	reference category								

Note: *,**, and *** indicate statistically significant with p-value<0.05, p-value <0.01, and p-value<0.001, respectively.

In particular, PIP is an assistance program of the Indonesian government for children aged 6–21 years from poor/vulnerable/priority families so that they can continue to participate in school and prevent dropouts (Ministry of Education, Culture, Research, 2017). During the COVID-19 pandemic, families need to meet additional costs for the implementation of online learning in the form of buying gadgets and internet packages for students. Hence this could lead to a financial problem, especially for poor families. As reported by BPS-Statistics Indonesia, (2020; 2021a), the poverty rate in underdeveloped regions in 2020 is 26.43 percent, higher than the overall poverty rate of only 9.78 percent. The existence of PIP has the potential to ease that financial burden, especially in supporting online learning costs for students (The Government of Kapuas Hulu Regency, 2020). This study supports that statement by empirically showing the odds that a student uses the Internet is 1.13 times higher for a student that obtains PIP compared to a student that does not.

Effect of Household Characteristics

According to Nakagawa et al. (2022), there are two generations in the economy: adults and children. Consider that over a lifetime, a child consumes parental income to improve their quality of life through personal skills. Accordingly, child behavior depends on household characteristics. Moreover, household characteristics relate to the internet use of students. Household characteristics involve its particular characteristics and head of household characteristics. As mentioned before, household head characteristics used in this study are gender, age, education attainment, and occupation.

The household head represents the parent who plays an important role in supporting the student's access the Internet, particularly during the COVID-19 pandemic. Unfortunately, different capacities of the heads of households imply different ways to accompany and teach the student. In developed countries like the United States, the gender of household heads has no significant effect on the internet use of students (DeBell & Chapman, 2006). Interestingly, findings are different between developed and underdeveloped regions. This study finds that the gender of household heads has a significant effect on the internet use of students in underdeveloped regions. This means that there are different patterns of using the Internet for students who live with male or female household heads. A student who lives with a male household head has an odds ratio less than unity, that is, 0.8460. This means the odds that a student uses the Internet is 0.8460 times lower for students with a male household head compared to students with a female household head. In other words, students who live with male household heads are less likely to go online than students who live with female household heads.

Regarding age, this study finds that the age of household heads has a significant effect on internet use. This means that there are different patterns of using the Internet for students who live with productive age (15-64 years) and non-productive age (65 years and over) household heads. Interestingly, a student who lives with a productive age household head has an odds ratio less than unity, that is, 0.8028, of using the Internet. This reveals that students who live with productive age household heads are less likely to go online than students who live with non-productive age household heads.

Looking at the results, the level of educational attainment of household heads has a significant effect on internet use. A student who has a household head with a minimum education of junior high school has an odds ratio of 1.4987 of using the Internet. That means the odds that a student uses the Internet is 1.4987 times higher for a student who has a household head with a minimum education of junior high school compared to a student who has a household head with a maximum education only of elementary school. This finding confirms the previous study of DeBell and Chapman (2006) which stated that students with educated parents are more likely to use the Internet. Educated household heads have good skills in time allocation (Alifia et al., 2020) therefore these household heads have time to accompany and guide the student in their internet use. Among other reasons, a high level of parental education can guarantee a good social status and good material conditions for students.

In turn, many people believe that the occupation of household heads is interrelated with the education of household heads. Similarly, with household head education, the occupation of the household head has a significant effect on internet use. It means there is a different pattern of internet use between a student who lives with a household head who works in the non-agricultural sector and a student who has a household head who does not work or works in the agricultural sector. The student who lives with a household head who works in the non-agricultural sector tends to use the Internet 3.3759 times higher than a student who has a household head who does not work or who is working in the agricultural sector. Simply, students who live with a household head who works in the non-agriculture sector are more likely to use the Internet than students who live with a household head who does not work or works in the agricultural sector. This finding is in line with Lindblom and Räsänen (2017) and highlights people with positions in higher occupations are more likely to be frequent internet users. More generally, the head of household who is not working or working in the agriculture sector in developing countries probably has limited income or none. Therefore, their household prefers to spend the income for daily needs rather than for access to the Internet.

Meanwhile, household characteristics in this study are socioeconomic status and the number of students who live in the household. Previous studies in Spain (Rice & Katz, 2003) and in Mexico (Martínez-Domínguez & Fierros-González, 2022) found that socioeconomic status is another variable that affects internet use besides education level through school participation. Higher socioeconomic status means a higher probability of internet use (Chiao & Chiu, 2018). In addition, DeBell and Chapman (2006) also stated that students who live in poverty-stricken households are less likely to use the Internet in the United States.

The level of income of the household determines how a household allocates resources for internet use (Smith & Graham, 2012). Many studies approach the socioeconomic status of a household by income level (DeBell & Chapman, 2006; Smith & Graham, 2012; Van Deursen et al., 2015). In the economic view, budget constraints influence consumers' behavior because of the income they earn. Therefore, the household should consider the needs (goods and services) that can be met according to their income. In addition, budget constraints also shift if the socioeconomic status of the household changes. The combination of goods and services that are preferred and purchased are controlled by budget constraints. For those of lower socioeconomic status, daily basic needs are the main focus. Whereas, for those of higher socioeconomic status, they begin to spend their money for tertiary needs. In addition, they can afford to buy various advanced technologies and use them frequently (Lindblom & Räsänen, 2017).

Chiao and Chiu (2018) also composed an index of economic, social and cultural status that is calculated based on the educational attainment of the parents, the occupations of the parents, and household ownership. According to Lindblom and Räsänen (2017), socioeconomic status is related to occupation class and economic resources variables related to the ability to pay bills. However, the socioeconomic status of households in this study was approached by household expenditure. It is important to note that in developing countries, household expenditure is a great proxy of the welfare of a population since it shows long-term economic status compared to underestimated household income (Srivastava & Mohanty, 2010). Socioeconomic status is categorized into the first 40 percent, the second 40 percent, and the top 20 percent. This study confirms that the socioeconomic status of households has a significant effect on internet use. Households of lower socioeconomic status are less likely to use the Internet. The odds ratios from the second 40 percent category and the top 20 percent category are more than unity. Moreover, these odds ratios show an increasing pattern, that is, 1.1177 and 1.3191, which indicates that the odds that a student uses the Internet are 1.1177 and 1.3191 times higher for a student who lives in a household with a per capita expenditure in the second 40 percent and top 20 percent categories compared to a student who lives in a household with a per capita expenditure in the lowest 40 percent category. The increasing pattern of odds ratio was also found in Smith and Graham (2012).

Lera-López et al. (2011) describe that the presence of school-age children does not affect internet use in Spain. However, this study depicts the negative effect of the number of students in households on internet use. The number of students in the household in which a student lives has an odds ratio less than unity, that is, 0.8273 for internet use. This means that for the addition of a student in a household, the probability of a student using the Internet decreases by approximately 0.8273 times. In other words, the probability of a student using the Internet increases when fewer students live in a household. The possible explanation is the number of students in a household affects the needs that must be met. Increasing the number of students in a household will increase household expenses, including education expenses.

Conclusion

Nowadays, information and services are increasingly offered online. Hence, internet use is important to investigate. In conclusion, only one-fifth of students in underdeveloped regions are online or using the Internet. The main finding in this study reveals that education through school participation is the most important variable for internet use of students in underdeveloped regions. When students participate in higher education, they are more likely to use the Internet since they have more information on how to use it, and are aware of the greater benefits of using the Internet.

According to these results, student and household characteristics also affect internet use. Looking deeper, students who get the aid of PIP are more likely to use the Internet than students who don't get the aid of PIP. In addition, this study also finds that students who live in a household with household head characteristics such as being female, of non-productive age, having higher education, and working in non-agricultural sectors tend to use the Internet. Moreover, students who live in a household of higher socioeconomic status and with fewer students in the household, also tend to use the Internet.

While government policy is intended to increase the number of students in underdeveloped regions using the Internet, regardless of race, the policy should map priorities on both short and long-term scales. According to the main findings, the short-term policy is to promote Information and Communication Technology (ICT) use-based educational resources. In particular, digital literacy is an important way to convey the opportunities and the benefits from internet use for students and teachers, especially in the learning process during the COVID-19 pandemic. Moreover, since there is a digital divide in terms of internet use according to levels of school participation, for long-term policy proportional support for students needs to be given based on the level of school participation.

In addition, continuing PIP is also necessary to support households with lower socioeconomic status. Furthermore, the long-term policy to increase internet use of students in underdeveloped regions would be to develop internet skills for household heads whose characteristics are male, of productive age, poorly educated, of lower socioeconomic status, and have many children of school age living in the household. As a result, the household head would have a better chance of assisting the student with internet use.

Although this study is concerned with a very interesting topic, it has certain limitations, especially in the lack of data. First, cross-section data were chosen out of convenience and completeness of variables to be used since longitudinal data with complete variables are not available. Second, data do not contain the signal strength but only cover whether the internet is used or not. Third, this study only uses variables on the demand side whereas variables on the supply side, for example, availability of ICT facilities and infrastructure, are also important.

The contradictory results between the variable of age and education attainment of the household head would be interesting to investigate further. Future studies may account for the existence of unobserved variables that may have the power to explain the internet use of students in underdeveloped regions of Indonesia. It would be interesting to examine how socioeconomic variables on the demand and supply sides may influence the perceived benefits. In addition, since education is regarded as a future investment, patterns of individual changes are crucial. More would be demonstrated when using longitudinal data. Future studies are also needed on macro-level research to compare the potential of each region so that policies can be adapted to the problems of each region. Finally, these findings can serve as a foundation for further research on how education participation affects internet use among the students in Indonesian underdeveloped regions and other developing countries.

References

Agresti, A. (2007). *An introduction to categorical data analysis* (2nd ed.). John Wiley & Sons, Inc.

Al-Hammadany, F.H., & Heshmati, A. (2011). Determinants of internet use in Iraq. *International Journal of Communication*, *5*(23), 1967-1989. https://ijoc.org/index.php/ijoc/article/view/746/670

Alderete, M.V. (2019). Examining the drivers of internet use among the poor: The case of Bahía Blanca city in Argentina. *Technology in Society*, *59*(101179), 1-8. https://doi.org/DOI: 10.1016/j.techsoc.2019.101179

Alifia, U. (2020). *COVID-19 is widening Indonesia's education gap*. Research on Improving Systems of Education (RISE) Essay. https://rise.smeru.or.id/en/blog/covid-19-widening-indonesia’s-education-gap

Alifia, U.E.A., Barasa, A.R., Bima, L., Pramana, R.P., Revina, S., & Tresnatri, F.A. (2020). Belajar dari rumah: Potret ketimpangan pembelajaran pada masa pandemi COVID-19. *Catatan Penelitian Smeru*, *1*. https://rise.smeru.or.id/

id/publikasi/belajar-dari-rumah-potret-ketimpangan-pembelajaran-pada-masa-pandemi-covid-19

Arkiang, F. (2021). Analisis pembelajaran daring selama pandemi COVID-19 di daerah 3T (Nusa Tenggara Timur). *Jurnal Pendidikan*, *12*(1), 57-64. https://doi.org/http://dx.doi.org/10.31258/jp.12.1.57-64

Badan Pusat Statistik (BPS). (2020). *Konsep dan definisi Susenas Maret 2020 (Buku 4)*. Badan Pusat Statistik (BPS).

Barua, A., Whinston, A.B., & Yin, F. (2000). Value and productivity in the Internet economy. *Computer*, *33*(5), 102-105.

BPS-Statistics Indonesia. (2020). *Persentase penduduk miskin Maret 2020 naik menjadi 9,78 persen*. https://www.bps.go.id/pressrelease/2020/07/15/1744/persentase-penduduk-miskin-maret-2020-naik-menjadi-9-78-persen.html

BPS-Statistics Indonesia. (2021a). *Persentase penduduk miskin di daerah tertinggal (Persen), 2018-2020*. https://www.bps.go.id/indicator/153/1238/1/persentase-penduduk-miskin-di-daerah-tertinggal.html

BPS-Statistics Indonesia. (2021b). *Telecommunication statistics in Indonesia, 2020*. BPS-Statistics Indonesia. https://www.bps.go.id/publication/2021/10/11/e03aca1e6ae93396ee660328/statistik-telekomunikasi-indonesia-2020.html

Chiao, C., & Chiu, C.H. (2018). The mediating effect of ICT usage on the relationship between students' socioeconomic status and achievement. *The Asia-Pacific Education Researcher*, *27*(2), 109-121. https://doi.org/https://doi.org/10.1007/s40299-018-0370-9

Cooke, L., & Greenwood, H. (2008). Cleaners don't need computers: Bridging the digital divide in the workplace. *Aslib Proceedings, 60*(2), 143-157.

Cucinotta, D., & Vanelli, M. (2020). WHO declares COVID-19 a pandemic. *Acta Biomedica*, *91*(1), 157-160.

Dabla, A. (2004). The role of information technology policies in promoting social and economic development: The case of the state of Andhra Pradesh, India. *The Electronic Journal on Information Systems in Developing Countries*, *19*(5), 1-21. https://doi.org/DOI: 10.1002/j.1681-4835.2004.tb00126.x

Dahiya, S., Rokanas, L.N., Singh, S., Yang, M., & Peha, J.M. (2021). Lessons from internet use and performance during Covid-19. *Journal of Information Policy, 11*, 202-221. https://doi.org/https://doi.org/10.5325/jinfopoli.11.2021.0202

DeBell, M., & Chapman, C. (2006). *Computer and internet use by students in 2003* (NCES 2006–065; Statistical Analysis Report). https://nces.ed.gov/pubs2006/2006065.pdf

Elena-Bucea, A., Cruz-Jesus, F., Oliveira, T., & Coelho, P.S. (2021). Assessing the role of age, education, gender and income on the digital divide: Evidence for the European Union. *Information Systems Frontiers*, *23*(4), 1007-1021. https://doi.org/10.1007/s10796-020-10012-9

Eschachasthi, R., Purwa, T., & Cendekia, D.G. (2022). Does Palapa Ring Project infrastructure bridging connectivity and economic activity? In *Proceedings of 2021 International Conference on Data Science and Official Statistics (ICDSOS)*, 418-435. https://doi.org/DOI: https://doi.org/10.34123/icdsos.v2021i1.99

Hargittai, E. (1999). Weaving the Western web: explaining differences in Internet connectivity among OECD countries. *Telecommunications Policy*, *23*(10-11), 701-718. https://doi.org/https://doi.org/10.1016/S0308-5961(99)00050-6

Hoffman, D.L., & Novak, T.P. (1998). *Bridging the digital divide: The impact of race on computer access and internet use*. https://files.eric.ed.gov/fulltext/ED421563.pdf

Hosmer, D.W., & Lemeshow, S. (2000). *Applied logistic regression* (2nd ed.). John Wiley & Sons.

King, G., & Zeng, L. (2003). Logistic regression in rare events data. *Journal of Statistical Software*, *8*, 137-163. https://doi.org/10.18637/jss.v008.i02

Lampropoulos, G., Siakas., K., & Anastasiadis., T. (2019). Internet of things in the context of Industry 4.0: An overview. *International Journal of Entrepreneurial Knowledge, Center for International Scientific Research of VSO and VSPP*, *7*(1), 4-19. https://ideas.repec.org/a/jek/journl/v7y2019i1p4-19.html

Lera-López, F., Billon, M., & Gil, M. (2011). Determinants of internet use in Spain. *Economics of Innovation and New Technology*, *20*(2), 127–152. https://doi.org/https://doi.org/10.1080/10438590903378017

Lindblom, T., & Räsänen, P. (2017). Between class and status? Examining the digital divide in Finland, the United Kingdom, and Greece. *The Information Society*, *33*(3), 147-158. https://doi.org/https://doi.org/10.1080/01972243.2017.

1294124

Lunardon, N., Menardi, G., & Torelli, N. (2014). ROSE: A package for binary imbalanced learning. *R Journal*, *6*(1), 79-89. https://doi.org/10.32614/rj-2014-008

Martin, S.P., & Robinson, J.P. (2007). The income digital divide: Trends and predictions for levels of internet use. *Social Problems*, *54*(1), 1-22. https://doi.org/DOI:10.1525/SP.2007.54.1.1

Martínez-Domínguez, M., & Fierros-González, I. (2022). Determinants of internet use by school-age children: The challenges for Mexico during the COVID-19 pandemic. *Telecommunications Policy*, *46(1)*(102241), 1-18. https://doi.org/https://doi.org/ 10.1016/j.telpol.2021.102241

Middleton, K.L., & Chambers, V. (2010). Approaching digital equity: Is wifi the new leveler? *Information Technology and People*, *23*(1), 4-22. https://doi.org/DOI: 10.1108/09593841011022528

Ministry of Education, Culture, Research, and T. (2017). Pentujuk pelaksanaan program Indonesia pintar the 2017. https://psma.kemdikbud.go.id/data/files/ Petunjuk Pelaksanaan Program Indonesia Pintar th 2017.pdf

Mubarak, F., Suomi, R., & Kantola, S-P. (2020). Confirming the links between socio-economic variables and digitalization worldwide: The unsettled debate on digital divide. *Journal of Information, Communication and Ethics in Society*, *18*(3), 415-430. https://doi.org/https://doi.org/10.1108/JICES-02-2019-0021

Nakagawa, M., Oura, A., & Sugimoto, Y. (2022). Under-and over-investment in education: The role of locked-in fertility. *Journal of Population Economics*, *35*(2), 755-784. https://doi.org/https://doi.org/10.1007/s00148-021-00823-8

Noce, A.A., & McKeown, L. (2008). A new benchmark for internet use: A logistic modeling of factors influencing internet use in Canada, 2005. *Government Information Quarterly*, *25*(3), 462-476. https://doi.org/DOI: 10.1016/j.giq. 2007.04.006

Pénard, T., Poussing, N., Zomo Yebe, G., & Ella, N. (2012). Comparing the determinants of internet and cell phone use in Africa: Evidence from Gabon. *Communications & Strategies*, *86*, 65-83. https://econpapers.repec.org/article/idtjournl/cs8603.htm

Pick, J.B., & Nishida, T. (2015). Digital divides in the world and its regions: A spatial and multivariate analysis of technological utilization. *Technological Forecasting and Social Change*, *91*, 1-17. https://doi.org/https://doi.org/10.1016/j.techfore.2013.12.026

Purwa, T., & Cendekia, D.G. (2021). Mapping the potential use of ICT for distance learning during Covid-19 : Demand and supply-side approach. *63rd ISI World Statistics Congress 2021*, *1*, 77-80. https://www.isi-web.org/files/docs/papers-and-abstracts/23-day1-cps018-mapping-the-potential-use-of-i.pdf

Puspitasari, L., & Ishii, K. (2016). Digital divides and mobile internet in Indonesia: Impact of smartphones. *Telematics and Informatics*, *33*(2), 472-483. https://doi.org/https://doi.org/10.1016/j.tele.2015.11.001

Qomariyah, A.N. (2009). *Perilaku penggunaan internet pada kalangan remaja di perkotaan* [Universitas Airlangga]. https://repository.unair.ac.id/18241/

Rice, R.E., & Katz, J.E. (2003). Comparing internet and mobile phone usage: Digital divides of usage, adoption, and dropouts. *Telecommunications Policy*, *27*(8-9), 597-623. https://doi.org/https://doi.org/10.1016/S0308-5961(03)00068-5

Salas-Eljatib, C., Fuentes-Ramirez, A., Gregoire, T.G., Altamirano, A., & Yaitul, V. (2018). A study on the effects of unbalanced data when fitting logistic regression models in ecology. *Ecological Indicators*, *85*, 502-508. https://doi.org/http://dx.doi.org/10.1016/j.ecolind.2017.10.030

Shavkun, I., Bukharina, L., Dybchynska, Y., & Onyshchenko, O. (2021). Social economic factors of ICT use in education: Lessons from the pandemic. *CEUR Workshop Proceedings*, *3013*, 193-203. http://ceur-ws.org/Vol-3013/20210193.pdf

Singh, V. (2004). *Factors associated with household internet use in Canada, 1998–2000* (No. 28034; Agriculture and Rural Working Paper Series). https://doi.org/DOI: 10.22004/ag.econ.28034

Smith, D.T., & Graham, R. (2012). Household expenditures on information and communication technologies: A proposal for a digital practice model. *Race, Gender & Class*, *19*(3-4), 161-178. https://www.jstor.org/stable/43497494

Srivastava, A., & Mohanty, S.K. (2010). Economic proxies, household consumption and health estimates. *Economic and Political Weekly*, *45*(16), 55-63.

Sulisworo, D. (2016). The contribution of the education system quality to improve the nation's competitiveness of Indonesia. *Journal of Education and Learning*, *10*(2), 127-138. https://doi.org/DOI: 10.11591/edulearn.v10i2.3468

Szumilas, M. (2010). Explaining odds ratios. *Journal of the Canadian Academy of Child and Adolescent Psychiatry*, *19*, 227-229. https://doi.org/10.1136/bmj.c4414

The Government of Kapuas Hulu Regency. (2020). *PIP bantu pelajar pada masa pandemic Covid-19*. https://info.kapuashulukab.go.id/2020/10/13/pip-bantu-pelajar-pada-masa-pandemi-covid-19/

The Jakarta Post. (2020). *Indonesian internet users hit 196 million, still concentrated in Java: APJII survey*. https://www.thejakartapost.com/news/2020/11/11/indonesian-internet-users-hit-196-million-still-concentrated-in-java-apjii-survey.html

United Nations. (2020). *Policy brief: Education during COVID-19 and beyond*.

Usluel, Y.K., Aşkar, P., & Baş, T. (2008). A structural equation model for ICT usage in higher education. *Journal of Educational Technology & Society*, *11*(2), 262-273. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.542.4046&rep=rep1&type=pdf

Van Deursen, A.J., Van Dijk, J.A., & Peter, M. (2015). Increasing inequalities in what we do online: A longitudinal cross sectional analysis of internet activities among the Dutch population (2010 to 2013) over gender, age, education, and income. *Telematics and Informatics*, *32*(2), 259-272. https://doi.org/https://doi.org/10.1016/j.tele.2014.09.003

Vodoz, L., Reinhard, M., & Pfister, B.G. (2007). The farmer, the worker and the MP. *GeoJournal*, *68*(1), 83.

Wahid, F. (2007). Using the technology adoption model to analyze internet adoption and use among men and women in Indonesia. *The Electronic Journal of Information Systems in Developing Countries*, *32*(1), 1-8. https://doi.org/DOI: 10.1002/j.1681-4835.2007.tb00225.x

Willoughby, T. (2008). A short-term longitudinal study of internet and computer game use by adolescent boys and girls: Prevalence, frequency of use, and psychosocial predictors. *Developmental Psychology*, *44*(1), 195-204. https://doi.org/DOI: 10.1037/0012-1649.44.1.195

Yu, L. (2011). The divided views of the information and digital divides: A call for integrative theories of information inequality. *Journal of Information Science*, *37*(6), 660-679. https://doi.org/https://doi.org/10.1177/0165551511426246

Zhang, Z. (2016). Model building strategy for logistic regression: Purposeful selection. *Annals of Translational Medicine*, *4*(6), 4-10. https://doi.org/10.21037/atm.2016.02.15

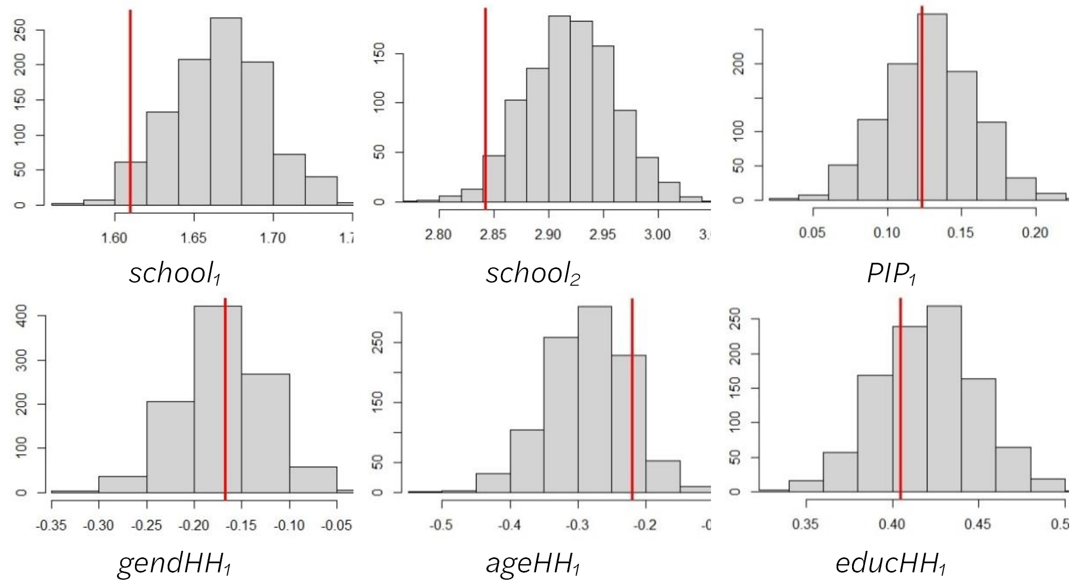
Appendix 1. Summary of Likelihood Ratio Test (LRT) Results

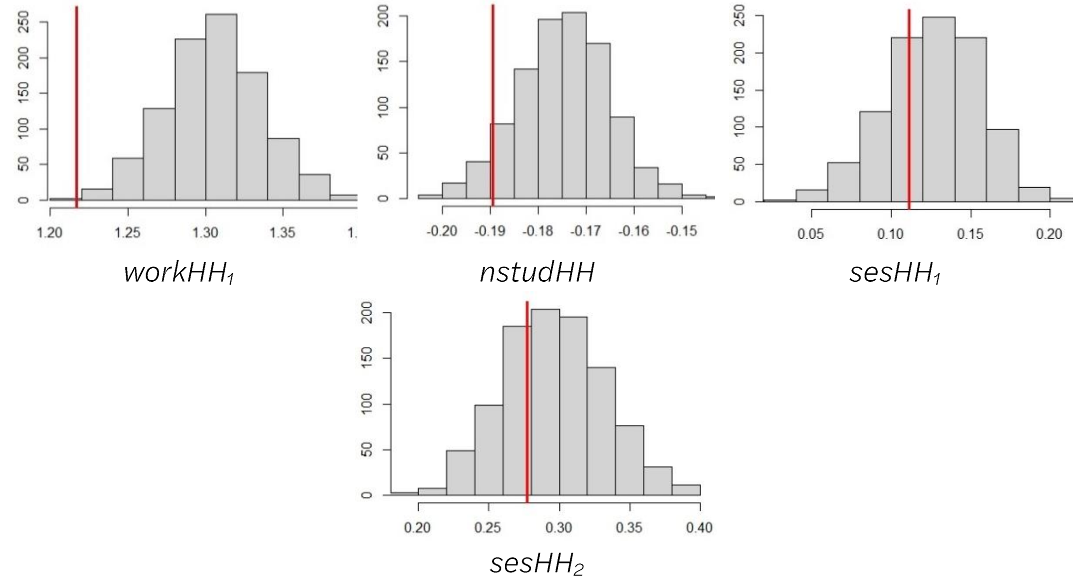
Test*	LRT statistic	p-value
Model 2 vs Model 1	2079.976	<0.001
Model 3 vs Model 1	2078.523	<0.001
Model 2 vs Model 3	1.452334	0.2282

Note: p-value<0.05 indicates that both models are significantly different, vice versa.

*the first model is more complex (has more variable) than the second one.

Appendix 2. Robustness Check of Logistic Regression Estimation using Repeated Under-Sampling Scheme





Note: vertical red line indicates estimation result using imbalance data.