WP-2024-013

Effect of Demand and Supply Side Factors on School Education Outcomes in India

Surabhi Garg and Srijit Mishra

Indira Gandhi Institute of Development Research, Mumbai
July 2024

Effect of Demand and Supply Side Factors on School Education Outcomes in India

Surabhi Garg and Srijit Mishra

Email(corresponding author): surabhig@igidr.ac.in

Abstract

A novelty of this study is to examine the effect of both demand and supply side factors on school education outcomes in India for 6-18 age group in 2019 by combining three different data sets, viz., Time Use Survey (TUS) 2019, Unified District Information Systems for Education Plus (UDISE+) 2019-20, and sub-group consistent population for 2019 extrapolated from Census of India. The proxies for demand side factors are gender, caste, religion and parents' education and the proxies for supply side factors are pupil-teacher ratio, infrastructure facilities and number of schools per school going age population. A two-stage Heckman Selection Model is used for the analysis. The model does a probit estimation at the first stage, with the dependent variable as whether the child attended the school or not. At the second stage, a least square estimation is done with the dependent variable as instructional time of the child in school. It is observed that both demand and supply side factors affect the probability of a child attending the school. However, at the second stage, it is observed that a higher number of proxies for supply side factors turn out to be significant in affecting the instructional time spent by the child in school. Given that the two-stage Heckman regression is based on the premise that higher instructional school time is associated with better education outcomes, it is difficult to place a cap on school time, especially for children with different characteristics and capabilities. Instead, a cap on the time mandated by schools is used for a robustness check. These estimation results also resonate with the baseline analysis. To wit, while both demand and supply side factors are relevant, the role of the State in facilitating the supply side factors has important public policy implications.

Keywords: Demand Side Factors, Heckman Selection Model, India, School Education Outcomes, Supply Side Factors

JEL Code: C30, C55, C81, C82, I21, I24, I25

Acknowledgements:

This paper is based on an essay of the lead author's PhD work, being carried out under the supervision of the senior author. The authors thank Subrata Sarkar and MH Suryanarayana for discussions with them. The paper also benefitted from comments by participants at the AIEFS conference held at IIT Madras, especially Meenakshi Rajeev among others. Usual disclaimers apply.

1. Introduction

Education plays an important role in an individual's well-being. This has been recognized at both national and international level. Given the importance of education, Article 21-A of India "*provide free and compulsory education of all children in the age group of six to fourteen years*" (Government of India, 2002). India has also been working towards attaining Sustainable Development Goals (SDGs) of the United Nations, including Goal 4 that aims to "*ensure inclusive and equitable quality education and promote lifelong learning opportunities for all*" by 2030 (United Nations, 2015).

Education enhances the functioning of the individuals and of the societies (Walker, 2021), which, as Drèze and Sen (1996) have articulated, is on account of its intrinsic as well as its instrumental importance. The intrinsic importance of education means that education is itself valued (Kumar, 2017) and important as an end (Gatley, 2021) while instrumental importance holds value as a means to achieve numerous ends such as income, social status and empowerment (Walker, 2021). Education as an end is affected by various factors. Existing literature has focused on the role of demand and supply side factors or *"push and pull factors"* (Goel and Husain, 2018) in affecting education outcomes, particularly at school level.

The relationship of demand and supply side factors with education outcomes has been studied separately. The importance of demand side factors has been highlighted by Borooah and Iyer (2005), Husain (2010) and Maertens (2013). Role of supply side factors in affecting education outcomes has been observed in the work of Kingdon (1996), Sharmila and Dhas (2010), Muralidharan and Prakash (2017) and Patra and Mete (2020).

Literature shows that demand and supply side factors have also been studied together. Reports on basic education (De et al, 1999; and De et al., 2011) gave some interesting theoretical insights on the demand and supply side aspects of schooling system in parts of India. Empirical work incorporating both these factors have majorly focused on parts of India such as Tamil Nadu (Duraisamy et al, 1998), rural north India (Drèze and Kingdon, 2001), rural India (Motiram and Osberg, 2012) and West Bengal (Ghosh, 2019).

This paper studies the association of demand side i.e. individual characteristics and supply side i.e. school characteristics with school education outcomes across the rural and urban regions of all the States and Union Territories of India. The data sets which allow us to capture both these effects are individual-specific information from Time Use Survey (TUS) of 2019 (Government of India, 2023), school-specific information from Unified District Information System for Education Plus (UDISE+) of 2019-20 (Government of India, 2021), and age- and district-specific extrapolated population of India for 2019 using Census of India 2001 and 2011 (Office of the Registrar General and Census Commissioner, India (ORG&CC), c.2021) that is made sub-group consistent to the rural and urban population of the States in 2019 as per estimates by the National Commission on Population (Ministry of Health and Family Welfare (MOHFW), 2020). Further, as TUS does not provide information on the school attended by the child, using UDISE+ and Census we derive district-specific school aggregates for rural and urban regions to match with the individual information

obtained from TUS. Our work, thus, makes a major contribution to the literature in terms of effect of both demand and supply side factors together on school education outcome in the pan-India context.

The subsequent sections start with literature review on the variables of interest under demand and supply side, and on possible proxies for school education outcomes. This is followed by a brief description of the data sets used in the study. We further describe our variables of interest. Given the available information from these data sets, we use two measures of school education outcomes for our study i.e. the school attendance of the child,[1] and the instructional time spent by the child in school. We give descriptive statistics for our variables of interest. The summary statistics leads us to the empirical methodology that should be used for our analysis. We then give the empirical results with their discussion and finally the conclusion.

2. Motivation and Review of Issues

The literature review below is on the variables of interest for demand side factors, supply side factors and possible school education outcomes.

2.1. Demand Side and Individual Characteristics

The demand side factors include parents' education, gender, religion and caste; which impact the education outcomes of the child. Both illiterate and literate parents value education for various reasons. Illiterate parents value education for its provision of earning potential as well as redistribution of income and reduction in income inequality among the masses (De et al, 1999; and Krueger, 2002). Studies show that parents' education has an impact on the education outcomes of the child (Khan et al., 2015; and Dickson et al., 2016). Further there may be cases when father and mother would have differential impact on child's (boy or girl) education attainment (Ermisch and Pronzato, 2010).

Gender is the other aspect of demand side. The right and desire to be educated for girls dates long back in history (Hossain, 1905; and Sperandio, 2019). However, there are challenges to the education of girl child due to social and economic reasons such as performing household chores (De et al, 1999) or taking care of younger siblings (Ghose, 2004), high opportunity cost of going to school in terms of income forgone from working as child labour and presumption amongst parents that they would have to pay higher dowry for more educated daughters. Besides, we cannot deny that female literacy is essential not just in terms of personal gains but also positive externalities that it generates for her child's welfare. In terms of the personal gains, literature shows that investment in women's education may lead to higher returns when compared to investment in men's education (Psacharopoulos, 1994) or for certain levels of education (Duraisamy, 2002).

Social norms are another facet to demand side factors and overlap with the gender aspect to education. These norms arise from belonging to a particular community or cultural group,

[1] The terms "school attendance of the child" and "child going to school" are used interchangeably in the text.

religion or caste (Borooah and Iyer, 2005). This may lead to early withdrawal of girl from school (Kelly and Carney, 1986) or child marriage or social unacceptability of female participation in labour market (De et al, 1999). The 75th round data of National Sample Survey (NSS) shows that Muslims are at the bottom of the literacy pyramid among all the religious groups in India (NSO, 2020). Studies shows that returns to education are the least for Muslims which may result from lack of formal education (Johnson, 1997), higher female discrimination in education attainment (Borooah and Iyer, 2005) or societal discrimination against Muslims (Husain, 2005; and Rani, 2014).

Similarly, literacy figures from 75th round of NSS (NSO, 2020) also indicate that Scheduled Tribe (ST) and Scheduled Caste (SC) categories rank the lowest among all the social groups. Reasons include low school enrolment rates among them and community/socio-economic characteristics such as child labour or financial constraints (Drèze and Kingdon, 2001; and Borooah and Iyer, 2005). Another cause is limited employment opportunities and male biased social programme. In fact, national programmes such as *Sarva Shiksha Abhiyan* targeted towards education of marginalized sections is also identified with design and implementation flaws (Jeffrey et al., 2005).

2.2. Supply Side and School Characteristics

Supply side factors incorporate quality of schooling as well as availability of schooling which may impact education outcomes. Here, we discuss pupil-teacher ratio and infrastructural facilities as a measure of quality and number of schools per school going age population as a measure of availability of schooling.

The UDISE+ data on pupil-teacher ratio (PTR) for India shows that although the average PTR has been close to 30:1 (a requirement as per *New Education Policy 2020*; Government of India, 2020), there is a huge gap between highest and lowest PTR for Indian States. A low PTR ensures quality learning for a child due to teacher being effectively able to attend his/her students (quality-quantity trade off) (Solheim and Opheim, 2018). Studies also show that how a fall in PTR can improve education outcomes (Card and Krueger, 1994; Duraisamy et al, 1998; Kimani and Bhorat, 2014; and Waita et al., 2016).

Literature has shown that better infrastructural facilities are associated with better education outcomes. Studies from both developed (Australia, USA) and developing countries (Kenya, Pakistan) depict a positive link between infrastructure and student achievement (Crampton, 2009; Parnwell, 2015; and Jamil et al., 2018).

Availability of schooling is an essential component of education outcomes. The positive linkage between the two is intuitive. Literature suggests that increasing the supply of schools reduces the travel time of the child to go to school (Filmer, 2007) and reduces disparity in access to education (Bhat and Khan, 2022). It raises school enrolment (Handa, 2002), school attendance (Burke and Beegle, 2004) and years of education (Duflo, 2001).

2.3. School Education Outcomes

These demand and supply side factors may affect school education outcomes. However, literature has described school education outcomes in various ways. One proxy among the various measures is academic achievements or learning outcomes of the students **(**Duraisamy et al, 1998; Schalock 1998; Linn, 2006; Parnwell, 2015; Dickson et al., 2016; Jamil et al., 2018; Solheim and Opheim, 2018). Another measure of better education outcomes used is higher years of education of an individual (De et al, 1999; Ermisch and Pronzato, 2010; and Kimani and Bhorat, 2014). Literature also uses a fall in student dropouts as a proxy for improvement in education outcomes (Jabbi and Rajyalakshmi, 2001; Ghose, 2004; Chevalier et al., 2005; Ferguson et al., 2007; and Goel and Husain, 2018). School education outcome has also been quantified by school enrolments (Borooah and Iyer, 2005) and the number of children going to school (Drèze and Kingdon, 2001).

Time spent by children in school is also defined as an education outcome in many studies. Aronson et al. (1998) discuss the difference between instructional and non-instructional time at school and mention the importance of academic learning time (instructional time) for student achievement. Lavy (2015) shows that "*additional instructional time has a positive and significant effect on test scores*," although less for developing as compared to developed countries. This positive link has also been shown by independent studies for California (Jez and Wassmer, 2015), Denmark (Andersen et al., 2016) and Delhi (Bhatnagar et al., 2022).

Based on the literature review, the demand side factors that would be important for our analysis include parents' education, gender, religion and caste. The supply side factors of interest include pupil-teacher ratio, infrastructure facilities and number of schools. The next section gives a brief description of the data sets used to obtain these demand and supply side factors.

3. Data

Demand side characteristics and school education outcomes are obtained from TUS 2019 (Government of India, 2023) while the supply side characteristics are derived from UDISE+ 2019-20 (Government of India, 2021) and age- and district-specific extrapolated population of India for 2019 using Census of India (2001 and 2011) such that the group-specific extrapolated population is made sub-group consistent to the rural and urban population of the States in 2019 as per estimates by the National Commission on Population (MOHFW, 2020).

3.1. Time Use Survey 2019

TUS 2019 collected data on the time use of all members aged 6 years and above of the surveyed households. TUS 2019 covered a total of 1,38,799 households of which 82,897 are rural and 55,902 are urban. These households comprise of 4,47,250 people (of which 2,73,194 live in rural areas and 1,74,056 live in urban areas) in the age group of 6 years and above. The information is collected using a recall period of 24 hours starting from 4:00 am prior to the date of interview to 4:00 am on the day of interview. The listing of activities in the 24 hours period comprises of nine major divisions as given by International Classification

of Activities for Time Use.[2] This survey incorporates information on important individual and household specific characteristics such as age, sex, marital status, religion, social group, education attainment, economic status and time spent on individual activities.[3]

In the current study, the focus is on school going children i.e. children in the age group of 6-18 (greater than equal to 6 years and less than 18 years). Hence, from the said age group of TUS 2019, the reference data used in the current analysis is of 63,774 children in rural areas and 33,839 children in urban areas. The analysis uses child's background characteristics and time spent by the child in formal education in a normal day contingent on child's school attendance as the important variables.[4]

3.2. Unified District Information System for Education Plus (UDISE+), 2019-20

The Unified District Information System for Education Plus (UDISE+), 2019-20 (Government of India) has information from all the recognized and unrecognized schools in India which impart formal education till grade XII and the students studying in the school who are below 18 years of age. According to the source, there are around 25,40,00,000 students in the age group of 5-18 years[5] who are enrolled in schools. The total number of schools covered during 2019-20 are 15,07,708 of which 12,58,347 schools are in rural areas and 2,49,361 schools are in urban areas. UDISE+ incorporates information on school specific characteristics like on infrastructure, teacher-related, enrolment and examination result.

3.3. Census of India

National Commission on Population provides estimates of population for rural and urban areas of States and India for 2011-2036. From this, the state-specific rural and urban estimates of population for 2019 are taken as our benchmark. For each state, separately for rural and urban, using the district-specific population of 2001 and 2011, extrapolated population for 2019 are made sub-group consistent with our state-specific benchmark population of 2019. Further, in each district the 6-18 age group (i.e. those greater than equal to 6 years and less than 18 years) and the rest of the population are made sub-group consistent to the district-population of 2019. The method for sub-group consistency is as in Mishra (2006) and the details for the current exercise are given in Appendix A1.

---

[2] The nine major division of activities for TUS are: Employment and related activities; Production of goods for own final use; Unpaid domestic services for household and family members; Unpaid caregiving services for household and family members; Unpaid volunteer, trainee and other unpaid work; Learning; Socializing and communication, community participation and religious practice; Culture, leisure, mass media and sports practices; and Self-care and maintenance.

[3] This study also requires parents' information. However, TUS 2019 does not explicitly capture parents' information but there is data on all the members of the household and their relationship with the head. Hence, the child's information with the head of the household is used to deduce the information about parents of the child.

[4] For accuracy, we only use the data of normal days for the entire analysis. A normal day for a child is one when he/she can perform routine activities such as doing homework. If the child is sick or otherwise and is not able to perform routine activities, then it is termed as a non-normal day.

[5] The school going age population generally considered is 6-17 years of age group. However, UDISE+ data set provides information for children going to school in the age group of 5-18 years. Hence, we use the information for all the students from UDISE+ since it is important in calculating school-specific variables such as pupil-teacher ratio.

4. Variables

We now combine the three data sets to obtain all our variables of interest. The merged data set comprises of the individual level information from the TUS 2019 which has been matched to derived district-specific supply side findings from the UDISE+ 2019-20 data and subgroup consistent extrapolated district-specific school-going population of 2019. The variable names and their description have been given in Appendix A2. The dependent variables and other variables of interest are also elaborated below.

4.1. Dependent Variable: Education Outcomes

TUS 2019 provides information on whether a child has attended school and the instructional time. $Attendance$ in school is a dummy that takes the value 1 if the child has attended/gone to school and takes the value 0 otherwise. Instructional $Time$ in minutes is greater than 0 if the child attended/went to school and is 0 if the child did not attend/did not go to school. As indicated earlier, in our analysis we limit the data to attendance in normal days. Besides, it is pertinent to note that TUS 2019 does not have information of children on their academic achievement in terms of grades obtained or in terms of reading/writing skills.

4.2. Variables of Interest: Control Variables

As discussed previously, demand and supply side variables are the control variables used to study the impact on education outcomes. The demand side and identification variables are individual/household specific while the supply side variables are school-specific that is to be aggregated at the district-level. The demand-side, supply-side and identification variables are elaborated in the following sections. It may be noted that each of the variables are segregated at the regional (rural/urban) and educational level (primary: classes I-V, upper primary: classes VI-VIII, secondary: classes IX-X, and senior secondary: classes XI-XII).

4.2.1. Demand Side and Identification (Individual and Household-level)

The demand side variables accrue to a child or the household in which the child resides. TUS 2019 is the data source for these variables. Child level variables are (i) Age of child in years (ii) $Male$ (dummy for gender of the child: if male, 1, otherwise, 0); and Variables depicting parents' education in years are (iii) $Mother's\_edu$, and (iv) $Father's\_edu$.[6] Household level variables are (v) $Social\_group$ of the household (multiple dummies identified with Scheduled Tribe (ST), Scheduled Caste (SC), Other Backward Classes (OBC), and Others), and (vi) $Religion$ of the household (multiple dummies identified with Hindu, Muslim, Christian, and Other religion).[7] Some other household level information that is to be used as identification variables are (vii) $No\_of\_children$ and (viii) $Season$ (multiple dummies, one for each quarter of survey period, Jan-Dec 2019).

[6] TUS gives the level of parents' education which is converted to years of education using the description given in TUS schedule.

[7] We omit measure of economic status of the household to avert potential endogeneity since we are including parents' education and other measures of household characteristics.

4.2.2. Supply Side (School-specific District-level)

The supply side variables are computed from the UDISE+ 2019-20 data. These variables include pupil-teacher ratio, school infrastructure (proxied by pucca blocks in the school and toilet facilities in schools) and availability of schools.

It is known that the TUS 2019 gives individual level data but UDISE+ 2019-20 data is available at the school level. Hence, in order to merge the two data sets and make them comparable, the school specific characteristics are aggregated at the district level (for each level of school education) to match the individual information (through the district where the child is residing).

(i) *Pupil-Teacher Ratio*

The analytical report 2009-10 of UDISE (Mehta, 2010) gives the formula for pupil-teacher ratio (PTR) of a school as the ratio of 'total enrolment in that school' to 'total teachers in that school'. This can be extended to the education level (primary, upper primary, secondary, and senior secondary) such that PTR of a school at each (education) level is the ratio of 'total enrolment in that school in that level' to 'total teachers in that school teaching in that level'.

The *Right of Children to Free and Compulsory Education (RTE) Act, 2009* (Government of India, 2009), suggests PTR to be broadly around 30:1 for primary school students and 35:1 for upper-primary school students. However, the *New Education Policy 2020* (Government of India, 2020) suggests PTR of below 30:1 at each level of school. For ease of analysis, we may consider PTR of 30:1 as the benchmark for all levels. In each area (district-specific rural/urban sector), k, for each education level, l, the measure indicating the schools that satisfy the benchmark is:

$$PTR30 = \frac{Area\ and\ level\ specific\ number\ of\ schools\ with\ PTR \leq 30}{Area\ and\ level\ specific\ number\ of\ schools} \times 100$$

(ii) *Pucca Building Blocks*

UDISE+ 2019-20, in addition to level of education imparted in each school, has data on the number of building blocks (BBs) in a school and how many of these are pucca. This can be used as a measure to indicate school infrastructure in each area (district-specific rural/urban sector) for each education level such that:

$$Pucca = \frac{Area\ and\ level\ specific\ number\ of\ pucca\ BBs\ in\ all\ the\ schools}{Area\ and\ level\ specific\ number\ of\ BBs\ in\ all\ the\ schools} \times 100$$

(iii) *Toilets in Schools*

The Bureau of Indian Standards (1993) gives the *Code of Basic Requirements for Water Supply, Drainage and Sanitation*. For schools, the minimum requirement is one urinal per twenty pupils, or students per toilet (SPT) should be 20:1 or lower. Further, SPT≤20 per school should be provided separately for each gender, male and female. This can also be computed for each education level (if a school has more than one education level then the

SPT is the same for all the levels). In each area (district-specific rural/urban sector), for each educational level, separately for each gender, (m$ale,\ female$), the measure to indicate (male/female) toilet facility is:

$$Toilet = \frac{Area, level\ and\ gender\ specific\ number\ of schools\ with\ SPT\ \leq 20}{Area, level\ and\ gender\ specific\ number\ of schools} \times 100$$

The two gender specific variables will be $Male\_Toilet$ and $Female\_Toilet$.

(iv) *Availability of Schools*

Using number of schools in the district from UDISE+ 2019-20 and sub-group consistent age-specific extrapolated population from Census 2001 and 2011 that is adjusted with state-specific estimates of 2019 by National Population Commission we can obtain the following. The school going age population is divided into four groups to represent the four educational levels: i.e. 6-10 years (primary), 11-13 years (upper primary), 14-15 years (secondary) and 16-17 years (senior secondary). In each area (district-specific rural/urban sector), for each educational level, the measure to indicate availability of schools per population is:

$$Schools\_per\_Pop = \frac{Area\ and\ level\ specific\ number\ of\ schools\ in\ a\ district}{Area\ and\ level\ specific\ number\ of\ children} \times 1000$$

5. Descriptive Statistics

We now give summary statistics of the variables discussed in the previous section. The descriptive statistics of the merged data are done separately for rural & urban areas and further, at educational level for accuracy and better exposition.[8] Table 1 gives an overview of the percentage of children going to school on a normal day. It is observed that higher percentage of children go to school in urban areas as compared to rural areas. Also, the percentage of children going to school falls with increase in education level. This fall is the highest from secondary to senior secondary level.

The following summary statistics discuss the mean and statistical test on the difference of means (t-test) of variables of interest for children going and not going to school, on a normal day.

5.1. Rural and Urban

Data depicts that the ratio of children going to school with respect to the children not going to school is 2.69 in rural areas and 3.16 in urban areas on a normal day.

Table 2 gives the mean values of continuous variables at regional level (rural and urban). The t-test for difference of means shows a significant difference in mean values of all the variables between the children who are going to school and those who are not going to school

[8] The mean values of the variables for boys and girls do not show much difference when we analyze the genders separately or together in rural and urban areas. Hence, disaggregation at the gender level is not depicted.

during the 24-hour recall period surveyed. Both the means of parents' education is higher for children going to school as compared to the children not going to school in both rural and urban areas. However, the level of the mean is lower for mothers as compared to fathers. Also, the level of the mean is lower for rural parents as compared to urban parents.

In both rural and urban areas, mean of $Pucca$ is higher for children not going to school as compared to the children going to school. As against this, the means of $PTR30$, $Male\_Toilet$, $Female\_Toilet$ and $Schools\_per\_Pop$ is higher for children going to school as compared to children not going to school. Further, for $Toilet$ (both male and female), and $Schools\_per\_Pop$ the difference is lower in urban than in rural.

5.2. Educational Levels

Here, the descriptive statistics are given at primary, upper primary, secondary and senior secondary levels. Data shows that the ratio of children going to school as compared to children not going to school falls with increase in education levels. The ratios are 4.12, 3.55, 2.59 and 1.10 in rural areas and 4.31, 3.81, 3.39 and 1.67 in urban areas for primary, upper primary, secondary and senior secondary levels respectively.

Tables 3, 4, 5 and 6 give the descriptive statistics for primary, upper primary, secondary and senior secondary levels respectively. Each of these tables gives information for both rural and urban areas. Mean of parents' education is higher for children going to school as compared to children not going to school for both rural and urban areas. However, high difference between the mean of parents' education is seen at senior secondary levels in rural and urban areas. Also, the mean of $Father's\_edu$ is higher than $Mother's\_edu$ at each school level.

At each education level, the mean of $No\_of\_children$ is slightly higher for children not going to school as compared to children going to school. Although this difference is small, it is statistically significant.

The mean of $PTR30$ is higher in rural areas as compared to urban areas at primary and upper primary levels for children going to school. Mean of $PTR30$ is higher for children going to school as compared to children not going to school at all educational levels in both rural and urban areas. Highest difference between the mean of children going and mean of children not going to school is seen at secondary and senior secondary levels in rural areas. For urban areas, the highest gap is seen at senior secondary level.

The mean of $Female\_Toilet$ falls with increase in the level of education in both rural and urban areas. The means of both $Female\_Toilet$ and $Male\_Toilet$ are higher for children going to school as compared to children not going to school in both rural and urban areas except for mean of $Male\_Toilet$ at secondary level in urban areas. However, the difference in mean of $Female\_Toilet$ and $Male\_Toilet$ is not much between children going and not going to school at all levels.

For all educational levels in both rural and urban areas, the mean of $Schools\_per\_Pop$ is higher when children go to school as compared to when children do not go to school. In rural

areas, the mean of $Schools\_per\_Pop$ falls with increase in the level of education. However, in urban areas, this mean is higher at upper primary as compared to primary level but lower at secondary and senior secondary level.

6. Empirical Methodology

Tables 3 to 6 give univariate depiction of the variables of interest. It is observed that there is a statistically significant difference between the mean values for children going and not going to school for the above discussed variables. Therefore, we now move to a multivariate analysis to show how these demand and supply side factors together affect the child's school $Attendance$. We also show the effect of these factors on instructional $Time$ of the child spent in school.

The methodology proposed for our analysis is the Heckman (1979) selection model. This model performs a probit analysis with the dependent variable being child's school $Attendance$ (a dummy), at the first stage, and does a least square estimation at the second stage with instructional $Time$ of the child as the dependent variable. Heckman's selection model is used since this model corrects for the potential sample selection bias at the second stage, where the sub-sample of children going to school, and for whom the data on instructional time are available, could potentially be a non-random sample from amongst the entire school going age population.

The identification assumption of the model is that the error terms in the selection equation and the outcome equation are jointly normally distributed. Hence, technically, there is no selection bias even if the explanatory variables in the selection and outcome equation are identical. However, as Sartori (2003) shows that if the sample size is small the Heckman's estimates have large variances which could potentially lead to estimates not being statistically significant for theoretically important variables, even though the identification assumption is met. To overcome this problem, Sartori (2003) recommends that it is advisable to have at least one explanatory variable (i.e. the exclusion restriction) in the selection equation which does not appear in the outcome equation.[9]

Here, the exclusion restrictions are the identification variables i.e. number of children in the household and seasonal dummies. Burke and Beegle (2004) mention that a household's decision of sending their child to school depends on the cost of schooling vis-a-vis the expected returns from schooling. The costs include school fees and the opportunity cost of time. The opportunity cost of time incorporates immediate income forgone, time spent in taking care of elderly or sick in the household, performing household chores, etc. The returns from schooling include higher future wages, better standard of living, higher social acceptability, among others. Based on the costs and benefits to education, "*Parents may place an intrinsic value on education. Thus, parents may prefer different levels of education for their children*" (Burke and Beegle, 2004). Illustrations from all over the world by Banerjee and Duflo (2011) show that parents may prefer to invest in the education of the child that

[9] In a review of empirical corporate finance, Li and Prabhala (2007) discuss the role of self-selection on account of unobservable private information and that is also relevant in our context.

they consider to be the most "*promising*" rather than investing in the education of all their children. Hence, a rise in the number of children in a household may actually affect the probability of each child going to school (Kugler and Kumar, 2017).

Another exclusion restriction is the seasonal variable defined by four quarters of the year. It is a well-known fact that Indian schools have summer vacations during May-June and winter vacations from mid-December to mid-January. The data from TUS 2019 also shows that on a normal day, school attendance is low during this time of the year. Hence, the quarter when the interview was conducted for the household may show different patterns of school attendance. Table 7 gives the percentage of children going to school amongst all the children surveyed in each quarter, on a normal day.

Hence, following the above discussion and in particular the works of Heckman (1979) and Sartori (2003) we now describe the construction of selection and outcome equations. The selection mechanism represented by the selection equation i.e. whether the child is going or not going to school as the 1st stage equation is:

$$z_i^* = w_i\gamma + u_i \tag{1}$$

where

$$z_i = \begin{cases} 1 \; if \; z_i^* > 0 & \text{(Child going to school)} \\ 0 \; if \; z_i^* = 0 & \text{(Child not going to school)} \end{cases}.$$

Here, z_i^* is the latent variable which measures the underlying propensity of the child to go to school; w_i encompasses various factors (child, household and school specific characteristics) which influence whether the child goes or does not go to the school; u_i contains any unmeasured characteristic in the selection equation. Hence, the Selection Equation is a Probit model and z_i is the school $Attendance$ dummy i.e. a dichotomous variable which takes the value 1 if the child is going to school, and 0 otherwise.

Now, we want to estimate the impact of demand and supply side factors on $Time$ spent in school. This is given by the outcome equation,[10]

$$y_i(time\ spent\ in\ school)|Sample\ Selection\ Rule = (x_i\beta + \epsilon_i)|Sample\ Selection\ Rule,$$

which as the 2nd stage regression is:

$$y_i|\, x_i, z_i^* > 0 = x_i\beta + \rho\sigma_\epsilon\lambda_i + v_i. \tag{2}$$

In equation (2), y_i is time spent in school, x_i encompasses various factors (child, household and school specific characteristics and identification variable) which influence time spent in school, the term $\rho\sigma_\epsilon\lambda_i$ corrects for selection bias (similar to an omitted variable bias) in the outcome equation. The model therefore incorporates and controls for the unobservable private information which could potentially impact the time spent by children in school, thereby correcting for the resultant bias in parameter estimates (Li and Prabhala, 2007).

[10] The derivation of the outcome equation is given in Appendix A3.

Thus, the selection equation (1) and the outcome equation (2) are the Heckman's model addressing selection bias.

7. Regression Results

Heckman selection model has been used to arrive at the regression results. We obtain regression estimates for the selection and outcome equation, given below.

7.1. Selection Equation

Tables 8 and 9 give the parameter estimates from the selection equation for rural and urban areas respectively. It is observed that both demand and supply side factors affect the probability of a child going to school in rural and urban areas. The effect of mother's education and father's education on child's school attendance is different in rural and urban areas. The coefficient for $Mother's\_edu$ and $Father's\_edu$ is positive but the coefficient for $Mother's\_edu$ is slightly higher as compared to $Father's\_edu$ in rural areas. These findings corroborate with Kurosaki et al. (2006) and Maitra and Sharma (2009).[11] However, in urban areas the coefficient for $Father's\_edu$ is higher as compared to $Mother's\_edu$.

The coefficient for *Muslims* is negative with *Hindus* as the base category, at higher level of education in rural and urban India. *OBC, SC* and *ST* students have a statistically significant lower school attendance as compared to the general category students in rural areas. However, only *SC* category students have a statistically significant lower probability of going to school as compared to the general category students in urban areas.

It may be noted that the $Male$ students show a statistically significant higher probability of going to school as compared to the female students at Upper Primary level. $Male$ students have a statistically significant lower probability of going to school as compared to the female students at Secondary and Senior Secondary level in rural areas. However, the coefficient for dummy variable $Male$ is not statistically significant at any level of education in urban areas.

$Schools\_per\_Pop$ has statistically significant positive association with the probability of school attendance in only rural areas at Primary level. However, $Schools\_per\_Pop$ has statistically significant negative association with the probability of attendance at Primary level in urban areas, and at Secondary and Senior Secondary levels in rural and urban areas.

$Pucca$ has statistically significant negative association with the probability of attendance at all levels in rural and urban areas. $PTR30$ has statistically significant positive association with the probability of attendance at all levels in both rural and urban areas.

The effect of father's education on child's school attendance is significantly higher than mother's education in urban areas. This result may be attributed to higher willingness of the father to financially support child's education (Dhesi, 2000). Fathers also consider the

[11] Increase in mother's year of schooling has a higher effect on child's school enrolment in rural areas as compared to urban areas.

development of child's "*academic and social skills"* as their responsibility (Sriram and Sandhu, 2013), especially among the urban middle class households.

Among the religious groups, Muslim children have a significantly lower probability of school attendance as compared to the Hindus, at higher education level. This may be due to higher school drop-outs and prevalence of child labour among Muslims (Mukherjee and Das, 2008) leading to their significantly lower education attainment (Maitra and Sharma, 2009). Also, SC and ST children face "*intrinsic disadvantage*" and bias (Drèze and Kingdon, 2001) in the schools. This may lead to their low enrolment rates (Borooah and Iyer, 2005) and high drop-out rates (Jabbi and Rajyalakshmi, 2001) in rural areas. Our result for social category in urban areas is supported by the findings of Akhtar and Kaleem (2020) who observe that the percentage of urban SCs who complete secondary education is not increasing as much as the rest of the population. Another finding that backs our result is from Tiwari et al (2020) who show that if a child belongs to a lower caste and lower class, then his/her school enrolment also falls. This finding holds true for SC category, in our case.[12]

We observe different results for both genders in rural and urban areas, and at different levels of education. At Upper Primary level, school attendance for boys is significantly higher than girls. This may be due to high students' drop-out (Goel and Husain, 2018) and low school attendance at lower level of education (Hill and Chalaux, 2011) and in our study, this may be higher for girls as compared to boys. However, at Secondary and Senior Secondary level, boys have a statistically significant lower probability of going to school as compared to the girls in rural areas. This may be due to the recognition among parents that improvement in female education would enhance intergenerational transfer of education (Husain, 2005) and chances of marriage (Goel and Husain, 2018). In addition, policies aimed at reduction in child labour have benefited girls more than boys (Mukherjee and Das, 2008). State sponsored schemes such as "*Cycle Program*" in Jharkhand (Muralidharan and Prakash, 2017) have increased female enrolment at secondary level of education. Improvement in school connectivity through roads has been a catalyst in reducing girl drop outs (Goel and Husain, 2018). Increase in monetary returns at senior secondary and higher level of education (di Gropello, 2006) overtime may also be one of the reasons driving this result. In urban areas, there is no statistically significant difference in the probability of going to school for boys and girls. Urbanization has played a positive role in female education (Sharmila and Dhas, 2010) with fall or negligible gender-gap in education (Husain and Sarkar, 2011 and Goel and Husain, 2018).

We now move to the explanation of results for supply side variables. Increase in the number of schools provides higher accessibility to education (Kingdon, 2007) and raises the attendance rates in school. This is backed by many studies such as by Motiram and Osberg, 2008 and Goel and Husain, 2018. However, in our case, higher number of schools per school going age population has negative association with school attendance. There are three factors

[12] TUS, 2019 data shows that the mean of usual monthly consumption expenditure (for the households with children under study) is the lowest for SC category in urban areas.

that may govern this result. Firstly, the number of private schools in India[13] has risen by 12 times as much as increase in the government schools. Furthermore, the school enrolment has fallen by 14.5% in government schools and risen by 38.5% in private schools (Kingdon, 2020). This clearly shows that there has been a drastic student shift from government to private schools. Secondly, UDISE does not cover all the "*unrecognised private unaided schools*" in survey.[14] Thirdly, average annual drop outs have increased at secondary level from 14.54% in 2013-14 to 22.13% in 2016-17 (Sridevi and Nagpal, 2019). Therefore, the private schools where a large proportion of students were enrolled may not be covered by UDISE,[15] and government schools with higher number of dropouts are included in the survey. This may have led to the aforementioned negative association at all levels in urban areas and higher levels of education in rural areas.

School infrastructure ($Pucca$) has negative association with child attendance in our case. There may be two reasons to this result. Firstly, more pucca buildings may be related to high cost of schooling. This may be an entry barrier for poor and lower middle class children to attend school. Secondly, our data shows that there is a negative correlation[16] between $Pucca$ and toilet facility in schools. Given the finances, schools may make a trade-off in investing between the two infrastructure proxies: pucca buildings and availability of toilets. Clearly, toilet is a comparatively important factor in affecting child's attendance and shows a positive association with school attendance at higher levels of education.

Thirdly, higher percentage of schools with lower pupil-teacher ratio is positively associated with child's attendance. This result is intuitive and consistent with the studies on pupil-teacher ratio and school outcomes (Duraisamy et al, 1998, Solheim and Opheim, 2018).

7.2. Outcome Equation

Tables 10 and 11 give the parameter estimates from the outcome equation for rural and urban areas respectively. It is seen that higher numbers of supply side variables are statistically significant in affecting $Time$ at all levels of instruction of formal education in both rural and urban areas.

Conditional on school attendance, coefficients for parents' education are not statistically significant in affecting $Time$. However, coefficient is negative for $Father's\_edu$ at Upper Primary level in rural areas and at Secondary and Senior Secondary level in urban areas. Also, the coefficient for dummy variable $Male$ is not statistically significant at all levels of education in both rural and urban areas, conditional on school attendance. This reflects the falling gender-gap in literacy discussed earlier.

[13] This study is for 21 states of India between 2010 and 2016.
[14] Kingdon, 2019 shows that the unrecognized private schools in UDISE data "constituted less than 2% of all the elementary schools in the country in 2016-17."
[15] Given that we only use UDISE to capture the information on number of schools.
[16] Correlation is statistically significant at 5% level of significance. It holds the value of -0.25 for $Male\_Toilet$ and -0.22 for $Female\_Toilet$.

However, *Muslims* spent a lower percentage of time as compared to *Hindus* in rural areas but higher percentage of time in urban areas. The coefficients for $Social\_group$ dummies are generally insignificant at Upper Primary levels in both rural and urban areas. Further, the results show that *OBC* and *SC* category students spend statistically significant higher percentage of $Time$ as compared to General category students at Primary and Secondary and Senior Secondary level in both rural and urban areas. The coefficient for *ST* students is positive at only Primary level in rural areas.

The results show that there is a statistically significant positive association between proxies of school infrastructure i.e. $Male * Male\_Toilet$, $Female * Female\_Toilet$ and $Pucca$ with $Time$. This has been corroborated by De et al, 1999; Hussain, 2011 and Dhar, 2012. Further, $Schools\_per\_Pop$ (in rural areas) and $PTR30$ (in rural and urban areas) negatively affects $Time$ at Primary level. However, $Schools\_per\_Pop$ (in rural and urban areas) and $PTR30$ (in rural areas) positively affect $Time$ at higher levels of education.

It has been well postulated in the literature that parents' education is positively related to education outcomes for their children. Nonetheless, our results show a negative association between father's education and instructional time spent by the child in school, at lower level of education in rural areas and at higher level of education in urban areas. This may be attributed to better education outcomes in private schools (Kumar and Choudhury, 2021) and non-coverage of many private schools in the UDISE data (Kingdon, 2020). Therefore, for the schools covered under UDISE data, parents may be substituting their children's school time with private tuitions and learnings at home (Kumar and Choudhury, 2021) to ensure better learning outcomes.[17] Hence, if all the schools were covered by UDISE, we could have observed a positive association.[18] Among the social groups, children from backward classes spend more instructional time in schools. This may be due to lower competency of SC, OBC and ST children in reading, writing and problem solving (Borooah, 2012). This would enable more effort in terms of higher time in school to catch up with children with higher academic achievement.

For lower level of education, number of schools per school going population and number of schools with low PTR have a negative association with instructional time at higher levels of education. This may be due to high student enrolment at lower levels of education.[19] Therefore, if $Schools\_per\_Pop$ and $PTR30$ are low i.e. there are more students per school or per teacher, then higher instructional time would be required for each child's education, given that the child is attending school.

However, number of schools per school going population and number of schools with low PTR have a positive association with instructional time at higher levels of education. This is because student time and effort has to be much higher at higher levels of education as

[17] TUS shows that mean time for "*Homework, being tutored, course review, research and activities related to formal education"* is highest at Upper Primary level in rural areas and, at Secondary and Senior Secondary level in urban areas.
[18] Assuming that higher *Time* is associated with better learning outcomes (Lavy, 2015)
[19] Dropouts increase after primary levels of education (Goel and Husain, 2018).

compared to lower levels of education. Higher $Schools\_per\_Pop$ and $PTR30$ improves accessibility and quality of education due to lesser number of students assigned to teachers or lesser students in school. Firstly, this would lead to higher vigilance on the performance of each student, given that the child is attending school. Secondly, it would nudge the student to spend more time in school in order to improve his/her learning outcomes.

7.3. Robustness Check - Regression Results with Cap on School Time

We have performed two-stage Heckman regression with the premise that higher instructional and school time is associated with better education outcomes, as also postulated in literature (Lavy, 2015). Given this presumption, it is difficult to place a cap on school time. More so, it is tough for children with different characteristics and capabilities. Nonetheless, UDISE data does give information on mandated school time for all the schools surveyed. Hence, we generate a dependent variable i.e. the difference between *ln (School Time of UDISE) and ln (School time of TUS).*[20] Here, *School Time of UDISE* is the aggregate of mandated school time of all the schools at the district level. *School time of TUS* is the total time spent by the child in school. Therefore, a regression of this difference on demand and supply side factors shows how various factors may significantly lead to deviation from actual time that the child should actually spend in school.

Tables 12 and 13 give the parameter estimates from the selection equation for rural and urban areas respectively. Tables 14 and 15 give the parameter estimates from the outcome equation for rural and urban areas respectively. The results from the outcome equation show that most of the demand and supply side factors which positively affected the *ln (Attendance at classes/lectures at all levels of instruction of formal education),* have a negative association with the difference between *ln (School Time of UDISE) and ln (School time of TUS).* This means that factors which favour (disfavour) higher instructional time are associated with lesser (higher) difference between mandated school time and actual time spent by the child in school.

8. Summary and Conclusion

A novelty of the study has been to combine the demand and supply side factors. For this, a two-stage Heckman selection model has been used to study the impact of both demand and supply side factors on attendance in classes/lectures and further on the instructional time in classes/lectures. To enable this, the approach to integrate data from different sources has been two-fold. First, the work has matched individual's demand side variables from the TUS 2019 with the district-specific supply-side variables derived from the UDISE+ 2019-20 and Census data extrapolated for 2019, to study their effect on the education outcomes in pan-India context. Second, sub-group consistent population estimates were derived for the ages 6-17 years at district level for rural and urban India to add to the district-specific supply-side aggregates.

[20] *School time of TUS = Instructional time in school + Time spent in extracurricular activities at school + Breaks at school.*

Heckman regression results show that the coefficient for inverse mills ratio is negative and statistically significant at primary and upper primary levels but not significant at secondary and senior secondary level. This means that the unobserved factors in selection and outcome equation at both primary and upper primary levels are negatively correlated, and thereby justifying the use of Heckman selection model for the analysis.

The observation from the study reiterates our understanding of the favourable demand and supply side factors are both required to enable the child to go to school. Parents' education and optimal pupil-teacher ratio are prerequisites to enable school attendance. Insignificance of the gender coefficient in effecting school time is interesting. However, policy interventions are still required to reduce student drop-out at higher levels of education, especially for females at upper primary levels in rural areas. Effective policy formulation and implementation is also required to promote attendance for SCs, STs, OBCs and Muslims, especially at higher levels of education in rural areas. Further, improvement in school infrastructure is also important.

Time spent by the child in school can be increased by focusing on improvement in supply side factors. These include construction of more schools, pucca school buildings and toilets. Focus on quality of education is also needed by ensuring optimal pupil-teacher ratio, especially at higher levels of education.

The analysis depicts that favourable supply side factors are required to enable child's school attendance as well as more time spent by the child in school. Hence, this work gives interesting implications with respect to India's demographic dividend. Abysmal infrastructure quality, poor pupil-teacher ratio and lesser availability of schools may hinder the economic and non-economic returns to India. Therefore, policies may be targeted towards improvement in supply side factors so that India does not miss out on her demographic dividend. Appropriate policies taking the instrumental advantages of the demographic dividend will steer the economy towards greater growth. It is equally important that India acknowledges the intrinsic importance of education, as envisaged in her Right to Education Act, 2009, and in providing for equal opportunities. To wit, India cannot meet the instrumental advantages of her demographic dividend if she misses out on their intrinsic importance.

This study ensures precision in the results such that the micro-data sets were matched to the smallest unit possible.[21] However, there are certain limitations to this analysis. First, there is lack of individual specific information on supply side factors. An extension to this study could be done if individual specific information on supply side factors is also available. Second, it may be difficult to quantify the time aspect with academic achievement since this study considers all the students with equal capabilities. However, some students may require much more study time to attain the same level of academic achievement as other students. Last, but not the least, UDISE+ 2019-20 data does not capture information on all the private unrecognized schools which may have given us different results. These give some directions for future research.

[21] We derive district estimates (for rural and urban regions) from UDISE+ 2019-20 data to match with individual data from TUS 2019.

Appendix

A1 Sub-group consistent population estimates for 2019

Census data is compiled for the population of all the ages separately (lying in the age group of 6-17 years) for rural and urban regions of all the districts of India. Here, 2001 and 2011 data sets are used to arrive at compound annual growth rate.

Compound Annual Growth Rate = $r = (N_{2011}/N_{2001})^{(1/10)}-1$

Estimated Population in 2019 = $N_{2019} = (1+r)^8 * N_{2011}$

where,

N_{2011}= Population in year 2011
N_{2001}=Population in year 2001
N_{2019}= Population in year 2019

Further, this growth rate is used to arrive at the population estimates for the year 2019. These population estimates are then made sub-group consistent (Mishra, 2006) with the population projections for States of India (at rural/urban level) given by National Commission on Population. The population projections are first made consistent at the district level and further for population in the age group of 6-17 years (inclusive of both years). This is done as:

a. Annual exponential growth rate is used to estimate population for 2019 at district level for all the states at regional level (rural/urban)
   $\widehat{N}_{dr}$ is the population estimate of a particular district for a region.
   This is done for all the districts of India at regional level.
b. Population of all the districts in a state is added to arrive at the state estimate at regional level (rural/urban)
   $\widehat{N}_{sr} = \sum_{d=1}^{n} \widehat{N}_{dr}$, where $\widehat{N}_{sr}$ is the population estimate of a state (with n districts) for a region.
   This is done for all the States of India at regional level.
c. This population estimate ($\widehat{N}_{sr}$) and population projection (N_{sr}) of a state for a region is used to arrive at sub-group consistent estimates at district level in rural and urban areas for each state, i.e.

$$\boxed{\widehat{\widehat{N}}_{dr} = \frac{\widehat{N}_{dr}}{\widehat{N}_{sr}} \text{ x } N_{sr}}$$

such that, $\widehat{\widehat{N}}_{dr} = \widehat{w}_{dr}$ x N_{sr} ,
$\sum_{d=1}^{n} \widehat{w}_{dr} = 1$ for a state,
$\sum_{d=1}^{n} \widehat{\widehat{N}}_{dr} = N_{sr}$

d. Further, annual exponential growth rate formula given above is also used to estimate population for different ages in 2019 at district level for all the states at regional level. These estimates are calculated for age category 0-5 together, 6-17 separately and together for age category 18-above. Hence, a total of 1+12+1=14 age estimates are obtained for a region in a district.[22] $\widetilde{N}_{adr}$ is the population estimate for an age category of a particular district for a region.
   This is done for all the age categories in all the districts of India at regional level.

e. Population of all the age categories is added to arrive at the district estimate at regional level (rural/urban)
   $\widetilde{N}_{dr} = \sum_{a=1}^{14} \widetilde{N}_{adr}$, where $\widetilde{N}_{dr}$ is the population estimate (obtained from adding all the age estimates) of a district for a region.
   This is done for all the districts of India at regional level.

f. This population estimate ($\widetilde{N}_{dr}$) and sub-group consistent district estimate ($\widehat{\widehat{N}}_{dr}$) for a region is used to arrive at age specific sub-group consistent estimates at district level in rural and urban areas for each state, i.e.

$$\boxed{\widetilde{\widetilde{N}}_{adr} = \frac{\widetilde{N}_{adr}}{\widetilde{N}_{dr}} \text{ x } \widehat{\widehat{N}}_{dr}}$$

such that, $\widetilde{\widetilde{N}}_{adr} = \widetilde{w}_{adr}$ x $\widehat{\widehat{N}}_{dr}$,
$\sum_{a=1}^{14} \widetilde{w}_{adr} = 1$ for a district,
$\sum_{a=1}^{14} \widetilde{\widetilde{N}}_{adr} = \widehat{\widehat{N}}_{dr}$ and $\sum_{d=1}^{n} \widehat{\widehat{N}}_{dr} = N_{sr}$

Hence, for our analysis, from the 28 sub-group consistent population estimates for one district (14 rural and 14 urban) we have 24 sub-groups for children of 6-17 years (one for each of the 12 years, separately for rural and urban).

[22] It may be noted that age categories 0-5 and 18-above are not of interest in the analysis but are segregated to obtain sub-group consistent estimates.

A2 Variable Names and Description

Variable Name	Description
Dependent Variables	
Attendance	Dummy variable for school attendance of a child on a normal day, equals 1 if child goes to school, and 0 otherwise*
Time	Time spent at classes and lectures at all levels of instruction of formal education; in minutes
Child and Household Specific Characteristics (at individual level)	
Age	Age of the child; in years
Male	Dummy variable for gender, equals 1 if child is male, and 0 otherwise
Mother's_edu	Education of mother; in years
Father's_edu	Education of father; in years
Religion	
Hindu	Dummy variable, equals 1 if household is Hindu, and 0 otherwise
Muslim	Dummy variable, equals 1 if household is Muslim, and 0 otherwise
Christian	Dummy variable, equals 1 if household is Christian, and 0 otherwise
Other Religion	Dummy variable, equals 1 if household is Other Religion, and 0 otherwise
Social_group#	
ST	Dummy variable, equals 1 if household is ST, and 0 otherwise
SC	Dummy variable, equals 1 if household is SC, and 0 otherwise
OBC	Dummy variable, equals 1 if household is OBC, and 0 otherwise
Others	Dummy variable, equals 1 if household is Others, and 0 otherwise
School Specific Characteristics (at district level)	
PTR30	Schools in a district with pupil-teacher ratio of less than or equal to 30; in %
Pucca	Total pucca buildings in all the schools in a district; in %
Female_Toilet	Schools in a district where functional female toilets are more than 1 per 20 female students; in %
Male_Toilet	Schools in a district where functional male toilets are more than 1 per 20 male students; in %
Schools_per_Pop	Total number of schools per school going age population (per 1000) in a district
Identification Variables	
No_of_children	Total children in the household; number
Season	
Quarter 1	Dummy variable, equals 1 if survey in Jan, Feb or Mar, and 0 otherwise
Quarter 2	Dummy variable, equals 1 if survey in Apr, May or Jun; and 0 otherwise
Quarter 3	Dummy variable, equals 1 if survey in Jul, Aug or Sep; and 0 otherwise
Quarter 4	Dummy variable, equals 1 if survey in Oct, Nov or Dec; and 0 otherwise

*: A normal day for a student will be a day when he/she goes to school, does his/her homework and does his/her other routine activities, whereas if the student on a day could not pursue routine activities due to illness such days will be classified as 'other day'.

#: ST is Scheduled Tribes, SC is Scheduled Castes, OBC is Other Backward Classes

A3 Derivation of the Outcome Equation of Heckman Selection Equation

The outcome equation in the 2nd stage regression is given as:

$$y_i(time\ spent\ in\ school)|\ Sample\ Selection\ Rule = (x_i\beta + \epsilon_i)|Sample\ Selection\ Rule$$

$$E(y_i|\ x_i, z_i^* > 0) = E(y_i|\ x_i, w_i\gamma + u_i > 0)$$

$$E(y_i|\ x_i, z_i^* > 0) = x_i\beta + E(\epsilon_i|u_i > -w_i\gamma)$$

$$E(y_i|\ x_i, z_i^* > 0) = x_i\beta + \rho\sigma_\epsilon\lambda_i$$

Where,

$$\lambda_i = \frac{\phi(\frac{-w_i\gamma}{\sigma_u})}{1 - \Phi(\frac{-w_i\gamma}{\sigma_u})} = Inverse\ Mill's\ Ratio$$

Hence, $$y_i|\ x_i, z_i^* > 0 = x_i\beta + \rho\sigma_\epsilon\lambda_i + v_i$$

Assumptions:

(a) $u_i \sim N(0,1)$
(b) $\epsilon_i \sim N(0, \sigma^2)$
(c) $Corr(u_i, \epsilon_i) = \rho$ (the error terms follow a bivariate normal distribution)

Tables

Table 1 - Percentage of Children Going to School/ School Attendance

School Type	Rural	Urban
(1)	(2)	(3)
Primary	80.48	81.16
Upper Primary	78.03	79.21
Secondary	72.16	77.23
Senior Secondary	52.35	62.61
Overall	72.90	75.98

Source and Note: Derived from TUS, 2019 data (for a normal day).

Table 2 - Mean of the Variables in Rural and Urban India

Variables	**RURAL (Mean)@**		**URBAN (Mean)@**	
	School Dummy=1	**School Dummy=0**	**School Dummy=1**	**School Dummy=0**
(1)	(2)	(3)	(4)	(5)
Child and Household Specific Characteristics (at individual level)				
$Time$	294.90	0	299.44	0
Age	11.70	13.98	11.99	13.95
$Mother's\_edu$	6.11	3.76	10.65	7.45
$Father's\_edu$	8.29	6.18	12.04	9.22
School Specific Characteristics (at district level)				
$PTR30$	69.99	62.02	67.82	63.62
$Pucca$	82.88	88.23	88.99	92.01
$Female\_Toilet$	22.15	17.77	26.64	24.37
$Male\_Toilet$	20.86	17.27	24.00	23.18
$Schools\_per\_Pop$	9.16	6.40	5.64	4.90
Identification Variables				
$No\_of\_children$	2.10	2.23	1.87	2.06
Sample Size	34503[#]	12824[##]	18379[c]	5810[cc]

Source: Derived from merged file of the TUS, 2019 (for a normal day); UDISE, 2019 and Census data sets.
Notes: @ t-test for difference of means (with unequal variance) between columns (2) v/s (3) and between (4) v/ (5) show the difference to be significant at 1% level of significance.
Except for the variables *Mother's Education* (has 1207 missing values) and *Father's Education* (has 419 missing values).
Except for the variables *Mother's Education* (has 586 missing values) and *Father's Education* (has 1760 missing values).
[c] Except for the variables *Mother's Education* (has 883 missing values), *Father's Education* (has 2185 missing values) and all the school specific characteristics (have 178 missing values).
[cc] Except for the variables *Mother's Education* (has 331 missing values), *Father's Education* (has 817 missing values) and all the school specific characteristics (have 51 missing values).

Table 3 - Mean of the Variables in Rural and Urban India for Primary Level of Education

Variables	PRIMARY- RURAL (Mean)		PRIMARY- URBAN (Mean)		t-test for difference of means, Pr (|T| > |t|)	
	School Dummy=1	School Dummy=0	School Dummy=1	School Dummy=0	(2) v/s (3)	(4) v/s (5)
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Child and Household Specific Characteristics (at individual level)						
Time	288.93	0.00	289.09	0.00	0.00	0.00
Age	8.46	8.77	8.41	8.84	0.00	0.00
Mother's_edu	6.45	5.37	10.81	9.61	0.00	0.00
Father's_edu	8.32	7.39	11.90	10.67	0.00	0.00
No_of_children	1.98	2.14	1.79	1.92	0.00	0.00
School Specific Characteristics (at district level)						
*PTR*30	68.53	65.24	60.57	55.94	0.00	0.00
Pucca	80.59	84.88	87.89	91.05	0.00	0.00
Female_Toilet	30.64	27.04	29.60	27.33	0.00	0.00
Male_Toilet	28.66	25. 60	25.64	23.72	0.00	0.00
Schools_per_Pop	12.93	11.71	6.12	5.96	0.00	0.11
Sample Size	13722[#]	3329[##]	6589[c]	1530[cc]		

Source: Derived from merged file of the TUS, 2019 (for a normal day); UDISE, 2019 and Census data sets.
Notes: # Except for the variables *Mother's Education* (has 357 missing values) and *Father's Education* (has 1530 missing values).
Except for the variables *Mother's Education* (has 91 missing values) and *Father's Education* (has 378 missing values).
[c] Except for the variables *Mother's Education* (has 174 missing values), *Father's Education* (has 534 missing values) and all the school specific characteristics (have 45 missing values).
[cc] Except for the variables *Mother's Education* (has 52 missing values), *Father's Education* (has 121 missing values) and all the school specific characteristics (have 4 missing values).

Table 4 - Mean of the Variables in Rural and Urban India for Upper Primary Level of Education

Variables	UPPER PRIMARY-RURAL (Mean)		UPPER PRIMARY-URBAN (Mean)		t-test for difference of means, Pr (|T| > |t|)	
	School Dummy=1	School Dummy=0	School Dummy=1	School Dummy=0	(2) v/s (3)	(4) v/s (5)
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Child and Household Specific Characteristics(at individual level)						
$Time$	297.29	0.00	301.92	0.00	0.00	0.00
Age	11.89	13.05	11.80	13.05	0.00	0.00
$Mother's\_edu$	5.92	3.98	10.50	8.01	0.00	0.00
$Father's\_edu$	8.12	6.47	11.91	9.61	0.00	0.00
$No\_of\_children$	2.27	2.52	2.00	2.27	0.00	0.00
School Specific Characteristics (at district level)						
$PTR30$	75.04	70.90	73.33	71.90	0.00	0.00
$Pucca$	82.35	87.30	88.88	90.89	0.00	0.00
$Female\_Toilet$	18.55	16.45	25.51	24.80	0.00	0.02
$Male\_Toilet$	17.11	15.58	22.13	21.79	0.00	0.21
$Schools\_per\_Pop$	9.22	8.12	6.89	6.61	0.00	0.035
Sample Size	8859[#]	2495[##]	4443[c]	1166[cc]		

Source: Derived from merged file of the TUS, 2019 (for a normal day); UDISE, 2019 and Census data sets

Notes: # Except for the variables *Mother's Education* (has 356 missing values) and *Father's Education* (has 1109 missing values).

Except for the variables *Mother's Education* (has 103 missing values) and *Father's Education* (has 316 missing values.

[c] Except for the variables *Mother's Education* (has 191 missing values), *Father's Education* (has 492 missing values) and all the school specific characteristics (have 23 missing values).

[cc] Except for the variables *Mother's Education* (has 57 missing values), *Father's Education* (has 145 missing values) and all the school specific characteristics (have 10 missing values).

Table 5- Mean of the Variables in Rural and Urban India for Secondary Level of Education

Variables	SECONDARY-RURAL (Mean)		SECONDARY- URBAN (Mean)		t-test for difference of means, Pr (|T| > |t|)	
	School Dummy=1	School Dummy=0	School Dummy=1	School Dummy=0	(2) v/s (3)	(4) v/s (5)
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Child and Household Specific Characteristics (at individual level)						
Time	300.66	0.00	306.83	0.00	0.00	0.00
Age	14.41	15.55	14.28	15.57	0.00	0.00
Mother's_edu	5.69	3.11	10.41	6.63	0.00	0.00
Father's_edu	8.12	5.62	12.08	8.62	0.00	0.00
No_of_children	2.21	2.35	1.95	2.19	0.00	0.00
School Specific Characteristics (at district level)						
*PTR*30	70.85	61.54	75.37	71.34	0.00	0.00
Pucca	84. 41	89.30	89.49	92.21	0.00	0.00
Female_Toilet	15.07	14.43	24.75	23.77	0.006	0.003
Male_Toilet	14.56	14.02	23.14	23.16	0.002	0.95
Schools_per_Pop	5.92	4.55	5.43	4.86	0.00	0.00
Sample Size	7342[#]	2832[##]	4210[c]	1241[cc]		

Source: Derived from merged file of the TUS, 2019 (for a normal day); UDISE, 2019 and Census data sets.
Notes: # Except for the variables *Mother's Education* (has 318 missing values) and *Father's Education* (has 912 missing values).
Except for the variables *Mother's Education* (has 150 missing values) and *Father's Education* (has 404 missing values).
[c] Except for the variables *Mother's Education* (has 240 missing values), *Father's Education* (has 617 missing values) and all the school specific characteristics (have 25 missing values).
[cc] Except for the variables *Mother's Education* (has 74 missing values), *Father's Education* (has 183 missing values) and all the school specific characteristics (have 6 missing values).

Table 6- Mean of the Variables in Rural and Urban India for Senior Secondary Level of Education

Variables	SENIOR SECONDARY-RURAL (Mean)		SENIOR SECONDARY-URBAN (Mean)		t-test for difference of means Pr (|T| > |t|)	
	School Dummy=1	School Dummy=0	School Dummy=1	School Dummy=0	(2) v/s (3)	(4) v/s (5)
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Child and Household Specific Characteristics (at individual level)						
Time	298.95	0.00	307.75	0.00	0.00	0.00
Age	16.75	17. 62	16.70	17.62	0.00	0.00
Mother's_edu	6.10	2.75	10.86	5.81	0.00	0.00
Father's_edu	8.82	5.38	12.53	8.03	0.00	0.00
No_of_children	1.92	2.06	1.72	1.95	0.00	0.00
School Specific Characteristics (at district level)						
*PTR*30	63.29	54.47	65.06	59.61	0.00	0.00
Pucca	88.35	90.72	90.87	93.38	0.00	0.00
Female_Toilet	15.03	13.43	24.55	22.05	0.00	0.00
Male_Toilet	14.83	13.84	24.38	23.60	0.00	0.01
Schools_per_Pop	2.96	2.37	3.07	2.98	0.00	0.11
Sample Size	4580[#]	4168[##]	3137[c]	1873[cc]		

Source: Derived from merged file of the TUS, 2019 (for a normal day); UDISE, 2019 and Census data sets

Notes: # Except for the variables *Mother's Education* (has 176 missing values) and *Father's Education* (has 640 missing values).

Except for the variables *Mother's Education* (has 242 missing values) and *Father's Education* (has 662 missing values).

[c] Except for the variables *Mother's Education* (has 278 missing values), *Father's Education* (has 542 missing values) all the school specific characteristics (have 85 missing values).

[cc] Except for the variables *Mother's Education* (has 148 missing values), *Father's Education* (has 368 missing values) and all the school specific characteristics (have 31 missing values).

Table 7 - Percentage of children going to school/school attendance in each quarter

Season	Rural	Urban
(1)	(2)	(3)
Quarter 1 (Jan, Feb, Mar)	77.28	79.73
Quarter 2 (April, May, June)	52.96	55.42
Quarter 3 (July, Aug, Sep)	81.63	84.96
Quarter 4 (Oct, Nov, Dec)	80.23	85.45

Source and Note: Derived from TUS, 2019 data (for a normal day).

Table 8 - Parameter Estimates of the Selection Equation (Rural)

Dependent Variable:

Attendance at classes/lectures at all levels of instruction of formal education

Explanatory Variables	**Overall Rural**	**Primary Schools**	**Upper Primary Schools**	**Secondary and Senior Secondary Schools**
(1)	(2)	(3)	(4)	(5)
Age	0.277***	0.869***	-2.602***	-4.832***
Age Squared	-0.015***	-0.055***	0.085***	0.135***
Father's_edu	0.012***	0.002	0.008***	0.030***
Mother's_edu	0.016***	0.008***	0.010***	0.033***
Male	0.027	-0.010	0.095*	-0.104**
*Male * Male_Toilet*	-0.001*	-0.001	0.003	0.004**
*Female * Female_Toilet*	-0.001	-0.001	0.003	-0.005***
Schools_per_Pop	-0.009***	0.004**	-0.002	-0.013***
PTR30	0.005***	0.002***	0.002**	0.006***
Pucca	-0.005***	-0.005***	-0.007***	-0.005***
Religion (Base - Hindu)				
Muslim	-0.026	0.110***	-0.059	-0.275***
Christian	-0.109***	-0.155***	-0.114	0.062
Others	0.161***	0.292***	0.104	0.245***
Social_group (Base - General)				
OBC	-0.094***	-0.081***	-0.041	-0.083**
SC	-0.148***	-0.150***	0.001	-0.242***
ST	-0.245 ***	-0.191***	-0.159***	-0.365***
Identification Variables				
No_of_children	-0.042***	-0.019**	-0.043***	-0.062***
Season (Base- First)				
Second	-0.406***	-0.486***	-0.464***	-0.517***
Third	-0.048***	-0.083***	-0.046	0.165***
Fourth	-0.052***	-0.108***	0.011	0.075**
/athrho	-1.991***	-2.100***	-1.925***	0.004
/lnsigma	-1.009***	-1.047***	-1.085***	-1.157***
Rho	-0.963	-0.970	-0.958	0.004
Sigma	0.364	0.351	0.338	0.314
Lambda	-0.351	-0.340	-0.324	0.001

Source: Authors Calculation

Note: *, **, and *** indicate 10%, 5% and 1% level of significance, respectively.

Table 9 - Parameter Estimates of the Selection Equation (Urban)

Dependent Variable:

Attendance at classes/lectures at all levels of instruction of formal education

Explanatory Variables	**Overall Urban**	**Primary Schools**	**Upper Primary Schools**	**Secondary and Senior Secondary Schools**
(1)	(2)	(3)	(4)	(5)
Age	0.244***	0.930***	-2.819***	-2.482***
Age Squared	-0.013***	-0.060***	0.090***	0.066***
Father's_edu	0.018***	0.005	0.011**	0.028***
Mother's_edu	0.009***	-0.001	0.005	0.018***
Male	0.030	0.110	-0.062	0.034
*Male * Male_Toilet*	0.001	0.001	-0.001	0.002
*Female * Female_Toilet*	0.002	0.003	-0.002	0.004*
Schools_per_Pop	-0.023***	-0.009*	-0.002	-0.036***
PTR30	0.006***	0.004***	0.004**	0.002***
Pucca	-0.002***	-0.005***	0.000	-0.003***
Religion (Base - Hindu)				
Muslim	-0.159***	-0.164***	-0.077	-0.172***
Christian	0.116*	-0.056	0.020	0.257***
Others	0.150***	0.147*	0.165	0.125
Social_group (Base - General)				
OBC	-0.005	0.006	-0.041	-0.020
SC	-0.141***	-0.188***	-0.010	-0.156***
ST	-0.011	0.080	-0.003	-0.047
Identification Variables				
No_of_children	-0.042***	-0.006	-0.055**	-0.063***
Season (Base- First)				
Second	-0.431***	-0.438***	-0.585***	-0.401***
Third	0.052**	-0.041	0.032	0.068*
Fourth	0.082***	-0.039	0.042	0.134***
/athrho	-1.735***	-1.957***	-1.475***	-1.598***
/lnsigma	-0.988***	-0.986***	-1.088***	-0.991***
Rho	-0.940	-0.961	-0.900	-0.921
Sigma	0.372	0.373	0.337	0.371
Lambda	-0.350	-0.358	-0.303	-0.342

Source: Authors Calculation

Note: *, **, and *** indicate 10%, 5% and 1% level of significance, respectively.

Table 10 - Parameter Estimates of the Outcome Equation (Rural)

Dependent Variable:

ln (Attendance time at classes/lectures at all levels of instruction of formal education)

Explanatory Variables	**Overall Rural**	**Primary Schools**	**Upper Primary Schools**	**Secondary and Senior Secondary Schools**
(1)	(2)	(3)	(4)	(5)
Age	-0.046***	-0.132***	-0.117**	0.082**
Age Squared	0.003***	0.009***	0.008***	-0.003**
Father's_edu	-0.003***	-0.000	-0.003***	0.001
Mother's_edu	-0.002***	-0.001	-0.001	-0.000
Male	-0.007	0.008	-0.012	-0.014
*Male * Male_Toilet*	0.002***	0.002***	0.001	0.002***
*Female * Female_Toilet*	0.001***	0.002***	0.001	0.001***
Schools_per_Pop	0.002***	-0.001***	0.002***	0.003***
PTR30	-0.0004***	-0.001***	0.001*	0.001***
Pucca	0.002***	0.003***	0.002***	0.001***
Religion (Base - Hindu)				
Muslim	-0.017***	-0.052***	-0.037***	-0.035***
Christian	0.032***	0.068***	0.014	0.015
Others	-0.037***	-0.044***	-0.009	0.008
Social_group (Base - General)				
OBC	0.044***	0.037***	0.030***	0.050***
SC	0.047***	0.042***	0.010	0.031***
ST	0.039***	0.022**	0.016	0.014

Source: Authors Calculation

Note: *, **, and *** indicate 10%, 5% and 1% level of significance, respectively.

Table 11 - Parameter Estimates of the Outcome Equation (Urban)

Dependent Variable:

ln (Attendance time at classes/lectures at all levels of instruction of formal education)

Explanatory Variables	**Overall Urban**	**Primary Schools**	**Upper Primary Schools**	**Secondary and Senior Secondary Schools**
(1)	(2)	(3)	(4)	(5)
Age	-0.016**	-0.151***	-0.247***	0.053
Age Squared	0.002***	0.011***	0.013***	0.000
Father's_edu	-0.002***	0.000	-0.000	-0.005***
Mother's_edu	-0.000	0.001	0.001	-0.002
Male	0.016	0.009	0.027	-0.003
*Male * Male_Toilet*	0.002***	0.002***	0.002**	0.003***
*Female * Female_Toilet*	0.003***	0.002***	0.003***	0.003***
Schools_per_Pop	0.008***	0.002	0.003**	0.014***
PTR30	-0.001***	-0.001***	0.001	-0.000
Pucca	0.002***	0.003***	0.001***	0.002***
Religion (Base - Hindu)				
Muslim	0.037***	0.049***	0.022	0.025**
Christian	0.003	0.037	-0.003	-0.035
Others	-0.055***	-0.013	-0.077***	-0.071***
Social_group (Base - General)				
OBC	0.017***	0.023**	0.009	0.022**
SC	0.029***	0.031**	0.005	0.046***
ST	-0.002	-0.012	0.013	0.010

Source: Authors Calculation

Note: *, **, and *** indicate 10%, 5% and 1% level of significance, respectively.

Table 12 - Parameter Estimates of the Selection Equation (Rural)

***Dependent Variable**: Attendance*

Explanatory Variables	**Overall Rural**	**Primary Schools**	**Upper Primary Schools**	**Secondary and Senior Secondary Schools**
(1)	(2)	(3)	(4)	(5)
Age	0.324***	0.872***	-2.644***	-4.826***
Age Squared	-0.017***	-0.055***	0.087***	0.135***
Father's_edu	0.015***	0.003	0.009***	0.030***
Mother's_edu	0.019***	0.009***	0.010***	0.034***
Male	0.008	-0.037	0.126**	-0.108***
*Male * Male_Toilet*	-0.001	-0.003**	-0.001	0.004**
*Female * Female_Toilet*	-0.001	-0.003***	0.002	-0.005***
Schools_per_Pop	-0.009***	0.007***	-0.000	-0.013***
PTR30	0.004***	0.001	0.001	0.006***
Pucca	-0.004***	-0.005***	-0.008***	-0.005***
Religion (Base - Hindu)				
Muslim	-0.078***	0.081**	-0.093**	-0.284***
Christian	-0.067*	-0.197***	-0.120	0.065
Others	0.139***	0.188***	0.118	0.241***
Social_group (Base - General)				
SC	-0.156***	-0.121***	-0.024	-0.247***
OBC	-0.086***	-0.064**	-0.041	-0.089***
ST	-0.258***	-0.160***	-0.179***	-0.363***
Identification Variables				
No_of_children	-0.049***	-0.015*	-0.030**	-0.062***
Season (Base- First)				
Second	-0.527***	-0.467***	-0.476***	-0.516***
Third	0.022	-0.034	0.008	0.172***
Fourth	0.053***	-0.002	0.039	0.074**
/athrho	1.279***	1.918***	1.787***	-0.019
/lnsigma	-0.978***	-1.014***	-1.054***	-1.043***
Rho	0.856	0.958	0.945	-0.019
Sigma	0.376	0.363	0.348	0.352
Lambda	0.322	0.348	0.329	-0.007

Source: Authors Calculation

Note: *, **, and *** indicate 10%, 5% and 1% level of significance, respectively.

Table 13 - Parameter Estimates of the Selection Equation (Urban)

Dependent Variable: *Attendance*

Explanatory Variables	**Overall Urban**	**Primary Schools**	**Upper Primary Schools**	**Secondary and Senior Secondary Schools**
(1)	(2)	(3)	(4)	(5)
Age	0.301***	0.968***	-3.324***	-4.751***
Age Squared	-0.016***	-0.062***	0.110***	0.133***
Father's_edu	0.020***	0.009**	0.011**	0.032***
Mother's_edu	0.014***	-0.001	0.006	0.030***
Male	0.059	0.160**	0.007	0.027
*Male * Male_Toilet*	0.003**	-0.001	-0.004	0.002
*Female * Female_Toilet*	0.005***	0.003	-0.002	0.004*
Schools_per_Pop	-0.019***	-0.011**	0.002	-0.034***
PTR30	0.006***	0.004***	0.002	0.002**
Pucca	-0.002**	-0.007***	-0.002	-0.003**
Religion (Base - Hindu)				
Muslim	-0.142***	-0.121***	-0.082	-0.173***
Christian	0.160**	-0.162	0.015	0.462***
Others	0.208***	0.159*	0.168	0.218**
Social_group (Base - General)				
SC	-0.124***	-0.181***	-0.039	-0.123**
OBC	0.009	0.006	-0.053	-0.003
ST	-0.009	0.047	0.017	-0.061
Identification Variables				
No_of_children	-0.046***	-0.001	-0.058**	-0.080***
Season (Base- First)				
Second	-0.687***	-0.490***	-0.675***	-0.623***
Third	0.157***	-0.039	0.089	0.260***
Fourth	0.174***	-0.072*	0.087	0.291***
/athrho	0.870***	1.820***	1.271***	0.050
/lnsigma	-0.963***	-1.006***	-1.008***	-0.958***
Rho	0.701	0.949	0.854	0.050
Sigma	0.382	0.366	0.365	0.384
Lambda	0.268	0.347	0.312	0.019

Source: Authors Calculation

Note: *, **, and *** indicate 10%, 5% and 1% level of significance, respectively.

Table 14 - Parameter Estimates of the Outcome Equation (Rural)

Dependent Variable: *Difference in ln(School Time of UDISE) and ln(School time of TUS)*

Explanatory Variables	**Overall Rural**	**Primary Schools**	**Upper Primary Schools**	**Secondary and Senior Secondary Schools**
(1)	**(2)**	**(3)**	**(4)**	**(5)**
Age	0.055***	0.133***	0.131**	-0.047
$Age\ Squared$	-0.003***	-0.009***	-0.008***	0.001
$Father's\_edu$	0.002***	0.001*	0.002**	-0.001
$Mother's\_edu$	0.002***	0.001	0.001	-0.001
$Male$	0.009	0.007	0.017	0.010
$Male * Male\_Toilet$	-0.002***	0.000	-0.001**	-0.002***
$Female * Female\_Toilet$	-0.001***	0.001*	-0.001	-0.002***
$Schools\_per\_Pop$	-0.001	0.002***	-0.001	0.000
$PTR30$	0.001***	-0.001***	0.001*	-0.000
Pucca	0.0003**	0.0004**	0.0005*	0.000
Religion (Base - Hindu)				
Muslim	0.005	0.028***	0.028**	0.015
Christian	-0.043***	-0.094***	-0.021	-0.003
Others	0.034***	0.021	0.040*	-0.002
Social_group (Base - General)				
SC	-0.033***	-0.025**	-0.020	-0.016
OBC	-0.022***	-0.016*	-0.033***	-0.013
ST	-0.053***	-0.050***	-0.048***	-0.013

Source: Authors Calculation

Note: *, **, and *** indicate 10%, 5% and 1% level of significance, respectively.

Table 15 - Parameter Estimates of the Outcome Equation (Urban)

***Dependent Variable**: Difference in ln(School Time of UDISE) and ln(School Time of TUS)*

Explanatory Variables	**Overall Urban**	**Primary Schools**	**Upper Primary Schools**	**Secondary and Senior Secondary Schools**
(1)	**(2)**	**(3)**	**(4)**	**(5)**
Age	0.012*	0.108***	0.180**	-0.027
Age Squared	-0.001***	-0.008***	-0.010***	0.001
Father's_edu	0.002***	0.000	0.000	0.002
Mother's_edu	0.000	-0.000	-0.001	-0.002*
Male	0.009	0.020	0.008	0.032
*Male * Male_Toilet*	-0.004***	-0.001**	-0.001	-0.007***
*Female * Female_Toilet*	-0.003***	-0.001	-0.001	-0.006***
Schools_per_Pop	-0.009***	-0.001	-0.004***	-0.016***
PTR<30	0.002***	0.001***	0.000	0.002***
Pucca	0.001**	0.002***	0.000	-0.001**
Religion (Base - Hindu)				
Muslim	-0.032***	-0.051***	-0.020	0.002
Christian	0.063***	0.002	0.079**	0.084***
Others	0.053***	-0.006	0.080***	0.073***
Social_group (Base - General)				
SC	-0.017*	-0.015	0.004	-0.027*
OBC	0.001	0.007	-0.012	0.008
ST	-0.003	0.015	-0.017	-0.014

Source: Authors Calculation

Note: *, **, and *** indicate 10%, 5% and 1% level of significance, respectively.

References

Akhtar, S.M.J. & Kaleem, S. (2020). An Analysis of Educational Attainment among Different Social Groups in India. *International Journal of Humanities and Social Science Research*, 6(6), 92-100. https://www.socialsciencejournal.in/assets/archives/2020/vol6issue6/6-6-29-763.pdf accessed 6 Jul 2024.

Andersen, S. C., Humlum, M. K., & Nandrup, A. B. (2016). Increasing Instruction Time in School does Increase Learning. *Proceedings of the National Academy of Sciences*, 113(27), 7481-7484. https://doi.org/10.1073/pnas.1516686113 accessed 6 Jul 2024.

Aronson, J., Zimmerman, J., & Carlos, L. (1998, April 20). *Improving Student Achievement by Extending School: Is It Just a Matter of Time?* [Paper Presentation]. PACE Media/Education Writers Seminar, San Francisco, California, United States. https://www2.wested.org/www-static/online_pubs/po-98-02.pdf accessed 6 Jul 2024.

Banerjee, A., & Duflo, E. (2011). *Poor Economics*. Public Affairs.

Bhat, Z. A. & Khan, M. A. (2022). Disparities in Educational Access of Tribals: Analysing School Availability in Tribal Zones of District Ganderbal, Jammu and Kashmir. *Mazedan International Journal of Social Science and Humanities*, 3(3), 1-5. https://eric.ed.gov/?id=ED624376 accessed 6 Jul 2024.

Bhatnagar, A., Jaiswal, A., Jain, A. & Bolia, N.B. (2022). An Analysis of Key Indicators for Enhancing School Performance: Evidences from India. *Socio-Economic Planning Sciences*, 84. https://doi.org/10.1016/j.seps.2022.101429 accessed 6 Jul 2024.

Borooah, V. K. (2012). Social Identity and Educational Attainment: The Role of Caste and Religion in Explaining Differences between Children in India. *The Journal of Development Studies*, 48(7), 887-903. doi.org/10.1080/00220388.2011.621945 accessed 23 Jun 2024.

Borooah, V.K., & Iyer, S. (2005). Vidya, Veda, and Varna: The influence of Religion and Caste on Education in Rural India. *The Journal of Development Studies*, 41(8), 1369 – 1404 https://www.econ.cam.ac.uk/people-files/faculty/si105/VidyaVedaVarna.pdf accessed 6 Jul 2024.

Bureau of Indian Standards (1993). *Code of Basic Requirements for Water Supply, Drainage and Sanitation*, Fourth Revision, BIS, New Delhi.

Burke, K. & Beegle, K. (2004). Why Children Aren't Attending School: The Case of Northwestern Tanzanian. *Journal of African Economies*, 13(2), 333-355. https://doi.org/10.1093/jae/ejh011 accessed 2 Apr 2024.

Card, D., & Krueger, A. B. (1994). *The Economic Return to School Quality: A Partial Survey*, [Working Paper 334], Industrial Relations Section, Princeton University. https://dataspace.princeton.edu/bitstream/88435/dsp01r781wg02t/1/334.pdf accessed 17 May 2024.

Chevalier, A., Dolton, P. & Levacic, R. (2005). School and Teacher Effectiveness. In S. Machin & A. Vignoles (Eds.), *What's the Good of Education?: The Economics of Education in the UK* (pp. 39-56). Princeton University Press. https://doi.org/10.1515/9780691188652-006 accessed 6 Jul 2024.

Crampton, F.E. (2009). Spending on School Infrastructure: Does Money Matter? *Journal of Educational Administration*, 47(3), 305-322. https://doi.org/10.1108/09578230910955755 accessed 26 Jun 2024.

De, A., Dreze, J., Kumar, S., Noronha, C., Pushpendra, Rampal, A., Samson, M. & Sinha, A. (1999). *Public Report on Basic Education in India.* Oxford University Press. Department of School Education & Literacy, Ministry of Education, Government of India. https://dsel.education.gov.in/rte accessed 4 Aug 2023.

De, A., Khera, R., Samson, M. & Kumar, A.K.S. (2011). *Probe Revisited: A Report on Elementary Education in India.* Oxford University Press. https://cordindia.com/wp-content/uploads/2023/10/Probe-Revisited.pdf accessed 6 Jul 2024.

Dhesi, A.S. (2000). Social Capital and Educational Achievement: The Case of India. *Canadian and International Education*, 29(2), 47 – 69. https://www.researchgate.net/publication/312439848_Social_Capital_and_Educational_Achievement accessed 5 May 2024.

di Gropello, E. (2006). *Meeting the Challenges of Secondary Education in Latin America and East Asia: Improving Efficiency and Resource Mobilization*. Directions in Development, Human Development. World Bank, Washington, DC, http://hdl.handle.net/10986/7173 accessed 7 Jul 2024.

Dickson, M., Gregg, P., & Robinson, H. (2016). Early, Late or Never? When Does Parental Education Impact Child Outcomes? *The Economic Journal*, 126(596), 184-231. https://onlinelibrary.wiley.com/doi/epdf/10.1111/ecoj.12356 accessed 24 February 2024.

Drèze, J., & Kingdon, G.G. (2001). School Participation in Rural India. *Review of Development Economics*, 5(1), 1-24. https://onlinelibrary.wiley.com/doi/abs/10.1111/1467-9361.00103 accessed 29 Apr 2023.

Drèze, J. & Sen, A. (1996). India: Economic Development and Social Opportunity. In A. K. Sen & J. Drèze (1999), *The Amartya Sen and Jean Drèze Omnibus*. Oxford University Press.

Duflo, E. (2001). Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment. *The American Economic Review*, 91(4), 795-813. http://www.jstor.org/stable/2677813 accessed 5 Sep 2023.

Duraisamy, P. (2002). Changes in Returns to Education in India, 1983-94: By Gender, Age-Cohort and Location. *Economics of Education Review*, 21(6), 609-622. https://doi.org/10.1016/S0272-7757(01)00047-4 accessed 6 Jul 2024.

Duraisamy, P., James, E., Lane, J. & Tan, J (1998). Is there a quantity–quality trade-off as pupil–teacher ratios increase? Evidence from Tamil Nadu, India. *International Journal of Educational Development*, 18(5), 367-383. https://www.sciencedirect.com/science/article/abs/pii/S0738059398000224 accessed 6 Jul 2024.

Ermisch, J., & Pronzato, C. (2010). *Causal Effects of Parents' Education on Children's Education*, [ISER Working Paper Series, No. 2010-16], University of Essex, Institute for Social and Economic Research (ISER), Colchester. https://www.econstor.eu/bitstream/10419/65988/1/632258780.pdf accessed 15 Jan 2024.

Ferguson, H.B., Bovaird, S. & Mueller, M.P. (2007). The Impact of Poverty on Educational Outcomes for Children. *Paediatrics & Child Health*, 12(8), 701–706. https://doi.org/10.1093/pch/12.8.701 accessed 14 Sep 2023.

Filmer, D. (2007). If you build it, will they come? School availability and school enrolment in 21 poor countries. *The Journal of Development Studies*, 43(5), 901-928. https://doi.org/10.1080/00220380701384588 accessed 6 Jul 2024.

Gatley, J. (2021). Intrinsic value and educational value. *Journal of Philosophy of Education,* 55 (4-5), 545-904. https://doi.org/10.1111/1467-9752.12555 accessed 13 Apr 2024.

Ghose, A. K. (2004). The Employment Challenge in India. *Economic and Political Weekly*, 39(48), 5106–5116. http://www.jstor.org/stable/4415834 accessed 6 Jul 2024.

Ghosh, S. (2019). Inequalities in Demand and Access to Early Childhood Education in India. *International Journal of Early Childhood,* 51(2), 145–161. https://doi.org/10.1007/s13158-019-00241-8 accessed 5 Jul 2024.

Goel, S., & Husain, Z. (2018). Gender, caste, and education in India: A cohort-wise study of drop-out from schools. *Research in Social Stratification and Mobility*, 58, 54-68. https://www.sciencedirect.com/science/article/abs/pii/S0276562417301816 accessed 13 Apr 2024.

Government of India (2020). *New Education Policy 2020*, Ministry of Human Resource Development, New Delhi. https://www.education.gov.in/sites/upload_files/mhrd/files/NEP_Final_English.pdf accessed 16 Apr 2024.

Government of India (2009). Right of Children to Free and Compulsory Education (RTE) Act, 2009, No. 35 of 2009, *The Gazette of India Extraordinary*, Part II, Section I, Ministry of Law and Justice (Legislative Department), 27 August 2009. https://dsel.education.gov.in/sites/default/files/publication/RTE_2nd.pdf accessed 16 Apr 2024.

Government of India (2002). *The Constitution (Eighty-Sixth Amendment) Act, 2002*, Ministry of Human Resource Development (now Ministry of Education), New Delhi. https://www.education.gov.in/sites/upload_files/mhrd/files/upload_document/amend86.pdf accessed 16 Apr 2024.

Government of India (2023). *Time Use Survey (TUS), January 2019-December 2019*, DDI-IND-CSO-TUS-2019-19, National Data Archive, Ministry of Statistics and Programme Implementation. https://microdata.gov.in/nada43/index.php/catalog/170/related_materials accessed 13 Apr 2024.

Government of India (2021). *Unified District Information System for Education Plus* (*UDISE+), 2019-20,* Ministry of Education. https://udiseplus.gov.in/#/en/page/udise accessed 26 Apr 2024.

Handa, S. (2002). Raising Primary School Enrolment in Developing Countries: The Relative Importance of Supply and Demand. *Journal of Development Economics*, 69(1), 103-128. https://doi.org/10.1016/S0304-3878(02)00055-X accessed 2 Jun 2024.

Heckman, J.J. (1979, January). Sample Selection Bias as a Specification Error. *Econometrica,* 47(1), 153-162. https://www.jstor.org/stable/1912352 accessed 24 May 2024.

Hill, S. & Chalaux, T. (2011, July). *Improving Access and Quality in the Indian Education System* [Working Paper No. 885]. Economics Department, Organisation for Economic Co-operation and Development. https://dx.doi.org/10.1787/5kg83k687ng7-en accessed 2 May 2024.

Hossain, R.S. (1905). *Sultana's Dream*, The Indian Ladies' Magazine, Madras. https://digital.library.upenn.edu/women/sultana/dream/dream.html accessed 24 May 2024.

Husain, Z. (2005). Analysing Demand for Primary Education: Muslim Slum Dwellers of Kolkata. *Economic and Political Weekly*, 40(2), 137-147. https://www.jstor.org/stable/4416043 accessed 6 Jun 2024.

Husain, Z. (2010). Gender disparities in completing school education in India: Analyzing regional variations**.** https://mpra.ub.uni-muenchen.de/25748/1/MPRA_paper_25748.pdf accessed 6 Jun 2024.

Husain, Z., & Sarkar, S. (2011). Gender Disparities in Educational Trajectories in India: Do Females Become More Robust at Higher Levels? *Social Indicators Research*, 101(1), 37–56. https://www.jstor.org/stable/41476418 accessed 6 Jun 2024.

Jabbi, M.K. & C. Rajyalakshmi (2001). Education of Marginalised Social Groups in Bihar. In A. Vaidyanathan & P.R. Gopinathan Nair (Eds.), *Elementary Education in Rural India: A Grassroots View, Strategy for Human Development in India* (Vol.2, pp.395–458). Sage Publications, New Delhi.

Jamil M., Mustafa, G., & Ilyas, M. (2018). Impact of school infrastructure and pedagogical materials on its academic performance: Evidence from Khyber Pakhtunkhwa. *FWU Journal of Social Sciences,* 12(1), 42-55.

https://www.academia.edu/37296104/Impact_of_school_infrastructure_and_pedagogicalmaterials_on_its_academic_performance_Evidence_from_Khyber_Pakhtunkhwa accessed 6 Jun 2024.

Jeffrey, C., Jeffrey, P., & Jeffrey, R. (2005). When schooling fails: Young men, education and low-caste politics in rural north India. *Contributions to Indian Sociology*, 39(1), 1-38. https://doi.org/10.1177/006996670503900101 accessed 26 Apr 2024.

Jez, S.J. & Wassmer, R.W. (2015). The Impact of Learning Time on Academic Achievement. *Education and Urban Society*, *47*(3), 284-306. https://doi.org/10.1177/0013124513495275 accessed 26 Apr 2024.

Johnson, D. C. (1997). Formal education vs. religious belief: soliciting new evidence with multinomial logit modelling. *Journal for the Scientific Study of Religion, 36*(2), 231–246. 10.2307/1387555 accessed 24 May 2024.

Khan, R.M.A., Iqbal, N., & Tasneem, S. (2015). The influence of Parents Educational level on Secondary School Students Academic achievements in District Rajanpur. *Journal of Education & Practice,* 6(16), 76-79. https://files.eric.ed.gov/fulltext/EJ1079955.pdf accessed 6 Jul 2024.

Kimani, E.M., & Bhorat, H. (2014). *The effects of pupil-teacher ratio and expenditure per pupil on educational attainment in South Africa* [Working Paper]. https://www.afdb.org/fileadmin/uploads/afdb/Documents/Publications/AEC_2014_-_the_effects_pupil-teacher_ratio_and_expenditure_per_pupil_on_educational_attainment_in_south_africa_-_11_2014.pdf accessed 6 Jul 2024.

Kingdon, G.G. (1996). The Quality and Efficiency of Private and Public Education: A Case-Study of Urban India. *Oxford Bulletin of Economics and Statistics*, 58(1), 57-82. https://doi.org/10.1111/j.1468-0084.1996.mp58001004.x accessed 27 Apr 2024.

Kingdon, G.G. (2020). The Private Schooling Phenomenon in India: A Review. *The Journal of Development Studies,* 56(10), 1795-1817. https://doi.org/10.1080/00220388.2020.1715943 accessed 13 Apr 2024.

Kingdon, G.G. (2007). The Progress of School Education in India. *Oxford Review of Economic Policy,* 23(2), 168-195. https://doi.org/10.1093/oxrep/grm015 accessed 18 Mar 2024.

Krueger, A.B. (2002, April). *Inequality, Too Much of a Good Thing* [Alvin Hansen Symposium on Public Policy]. Harvard University 2002, Cambridge, Massachusetts. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.92.231&rep=rep1&type=pdf accessed 6 Jun 2024.

Kugler, A.D. & Kumar, S. (2017). Preference for Boys, Family Size, and Educational Attainment in India. *Demography,* 54 (3), 835–859. https://doi.org/10.1007/s13524-017-0575-1 accessed 24 May 2024.

Kumar, D. & Choudhury, P.K. (2021). Do private schools really produce more learning than public schools in India? Accounting for student's school absenteeism and the time spent on homework. *International Journal of Educational Development,* 83. https://doi.org/10.1016/j.ijedudev.2021.102395 accessed 26 Apr 2024.

Kumar, M. (2017). Importance of Intrinsic and Instrumental Value of Education in Pakistan. *Journal of Education and Educational Development,* 4(2), 177-199. https://files.eric.ed.gov/fulltext/EJ1161505.pdf accessed 11 Aug 2023.

Kurosaki, T., Ito, S., Fuwa, N., Kubo, K. & Sawada, Y. (2006), Child Labor and School Enrollment in Rural India: Whose Education Matters?. *The Developing Economies*, 44 (4), 440-464. https://doi.org/10.1111/j.1746-1049.2006.00024.x accessed 26 Apr 2024.

Lavy, V. (2015). Do Differences in Schools' Instruction Time explain International Achievement Gaps? Evidence from Developed and Developing Countries. *The Economic Journal*, 125 (588), F397–F424. http://www.jstor.org/stable/24737554 accessed 28 Nov 2023.

Li, K. & Prabhala, N. (2007). Self-Selection Models in Corporate Finance. In B.E. Eckbo (Ed.), *Handbook of Corporate Finance: Empirical Corporate Finance* (1st ed., Vol. 1, Chapter 2, pp. 37- 86). North-Holland, Elsevier. https://doi.org/10.1016/B978-0-444-53265-7.50016-0 accessed 24 May 2024.

Linn, R.L. (2006). Validity of Inferences from Test-based Educational Accountability Systems. *Journal of Personnel Evaluation in Education,* 19, 5–15. https://doi.org/10.1007/s11092-007-9027-6 accessed 13 Apr 2024.

Maertens, A. (2013). Social Norms and Aspirations: Age of Marriage and Education in Rural India. *World Development,* 47 (C), 1-15. https://doi.org/10.1016/j.worlddev.2013.01.027 accessed 6 Jul 2024.

Maitra, P. & Sharma, A. (2009*). Parents and Children: Education Across Generations in India*. https://api.semanticscholar.org/CorpusID:155016747 accessed 28 Feb 2024.

Mehta, A. (2010). *Elementary Education in India: Progress Towards UEE, Analytical Report 2009-10*, National University of Educational Planning and Administration, New Delhi and Department of School Education and Literacy, Ministry of Human Resource Development, Government of India. https://educationforallinindia.com/Analytical%20Report%202009-10.pdf accessed 16 Apr 2024.

MOHFW (2020). *Population Projections for India and States 2011-2036: Report of the Technical Group on Population Projections*. National Commission on Population, Ministry of Health and Family Welfare, Government of India, New Delhi. https://main.mohfw.gov.in/?q=reports-0 accessed 6 Jul 2024.

Mishra, S. (2006). Suicide Mortality Rates across States of India, 1975-2001: A Statistical Note. *Economic and Political Weekly*, 41(16), 1566-1569. https://www.jstor.org/stable/4418116 accessed 27 Aug 2023.

Motiram, S., & Osberg, L. (2012). Demand or Supply for Schooling in Rural India? *electronic International Journal of Time Use Research*, 9(1), 1-27., https://web.archive.org/web/20150906171445/http://www.eijtur.org/pdf/volumes/eIJTUR-9-1.pdf accessed 13 Apr 2024.

Mukherjee, D. & Das, S. (2008). Role of Parental Education in Schooling and Child Labour Decision: Urban India in the Last Decade. *Social Indicators Research*, 89(2), 305–322. https://link.springer.com/article/10.1007/s11205-007-9233-0 accessed 13 Apr 2024.

Muralidharan, K. & Prakash, N. (2017). Cycling to School: Increasing Secondary School Enrollment for Girls in India. *American Economic Journal: Applied Economics*, 9(3), 321-350. 10.1257/app.20160004 accessed 24 May 2024.

NSO (2020). *Household Social Consumption on Education in India, NSS 75th Round (July 2017-June 2018)*, NSS Report No 585 (75/25.2/1), National Statistical Office, Ministry of Statistics and Programme Implementation, Government of India. https://mospi.gov.in/sites/default/files/publication_reports/Report_585_75th_round_Education_final_1507_0.pdf accessed 6 Jul 2024.

O'Kelly, C. G. & Carney, L. S. (1986). *Women and Men in Society: Cross-cultural Perspectives on Gender Stratification* (2nd ed.). Belmont C.A: Wadsworth Publishing.

ORG&CC (c.2021). *Census Tables* [Web Archive], Office of the Registrar General & Census Commissioner, India, Ministry of Home Affairs, Government of India. https://censusindia.gov.in/census.website/data/census-tables accessed 26 Apr 2024.

Parnwell, R. N. (2015). *Influence of School Infrastructure on Academic Performance in Public Primary Schools in Ruiri Location-Meru County, Kenya* [Thesis]. University of Nairobi. Nairobi, Kenya. http://erepository.uonbi.ac.ke/handle/11295/92947 accessed 13 Apr 2024.

Patra, J.N. & Mete, J. (2020). The Role of ICT in Improving the Quality of School Education in India. *Ideal Research Review*, V (21), 60-65. https://www.researchgate.net/publication/292403689 accessed 13 Apr 2024.

Psacharopoulos, G. (1994). Returns to Investment in Education: A Global Update. *World Development*, 22(9), 1325-1343. https://doi.org/10.1016/0305-750X(94)90007-8 accessed 13 Apr 2024.

Rani, P.G. (2014). Disparities in earnings and education in India. *Cogent Economics & Finance*, 2(1), 1-18. https://www.tandfonline.com/doi/pdf/10.1080/23322039.2014.941510?needAccess=true accessed 13 Apr 2024.

Sartori, A.E. (2003). An Estimator for some Binary-Outcome Selection Models without Exclusion Restrictions. *Political Analysis*, 11(2), 111-138. https://www.jstor.org/stable/25791717 accessed 24 May 2024.

Schalock, H.D. (1998). Student Progress in Learning: Teacher Responsibility, Accountability, and Reality. *Journal of Personnel Evaluation in Education*, 12, 237–246. https://doi.org/10.1023/A:1008063126448 accessed 13 Apr 2024.

Sharmila, N. & Dhas, A.C. (2010). Development of Women Education in India. https://mpra.ub.uni-muenchen.de/20680/1/MPRA_paper_20680.pdf accessed 6 Jul 2024.

Solheim, O.J., & Opheim, V. (2018). Beyond class size reduction: Towards more flexible ways of implementing a reduced pupil–teacher ratio. *International Journal of Educational Research*, 96, 146-153. https://www.sciencedirect.com/science/article/pii/S0883035518315052 accessed 13 Apr 2024.

Sperandio, J. (2019). *Pioneering Education for Girls across the Globe: Advocates and Entrepreneurs*, 1760–1904, Lexington Books, London.

Sridevi, K. & Nagpal, M. (2019). Trends in school dropout rate in India. *Researchers' Guild*. 10.15503/RG.V1I3.514 accessed 6 Jun 2024.

Sriram, R. & Sandhu, G. K. (2013). Fathering to Ensure Child's Success: What Urban Indian Fathers Do? *Journal of Family Issues*, 34(2), 159–181. doi:10.1177/0192513x12461136 accessed 13 Apr 2024.

Tiwari, S.K., Paltasingh, K.R. & Jena, P.K. (2020). Caste-class association and school participation in Uttar Pradesh, India: Evidence from NSSO data. *International Journal of Educational Development*, 79. https://doi.org/10.1016/j.ijedudev.2020.102292 accessed 14 Jun 2024.

United Nations (2015). *Transforming Our World: The 2030 Agenda for Sustainable Development*, A/RES/70/1, https://sdgs.un.org/sites/default/files/publications/21252030%20Agenda%20for%20Sustainable%20Development%20web.pdf accessed 13 Apr 2024.

United Nations (2021). *International Classification of Activities for Time-Use Statistics 2016*, Statistical Papers, Series M No. 98, Statistics Division, Department of Economic and Social Affairs. https://unstats.un.org/unsd/gender/timeuse/23012019%20ICATUS.pdf accessed 4 Jul 2024.

Waita, K.J., Mulei, K.O., Mueni, K.B., Mutune, M.J., & Kalai, J. (2016). Pupil-Teacher Ratio and its Impact on Academic Performance in Public Primary Schools in Central Division, Machakos County, Kenya. *European Journal of Education Studies*, 2(14), 37-68. https://oapub.org/edu/index.php/ejes/article/view/96/233 accessed 6 Jul 2024.

Walker, M. (2021). On Education and Capabilities Expansion. In E.C. Martinetti, S. Osmani & M. Qizilbash, *The Cambridge Handbook of the Capability Approach* (pp. 505-522). Cambridge University Press.