DETERMINANTS OF SECONDARY SCHOOL DROPOUT AND GENDER GAP IN EDUCATION IN PAKISTAN

Adiqa Kiani
Corresponding author: adiqakian@gmail.com
Federal Urdu University, Islamabad
&
Mukaram Khan
Federal Urdu University, Islamabad

Abstract

The main purpose of this study is to explore the determinants of school dropout at the completion of primary level and shed light on reasons behind gender gap in secondary education by using micro level data of Pakistan Social and Living Standard Measurement (PSLM) for the period 2013 to 14. Logit regression model is used to investigate the determinants of secondary school dropout while separate regression analysis is made for both boys and girls. The results show that age, household size and distance to the nearest school lead to school dropout while household income and parental education reduce the probability of school dropout with gender and place of residency also being important variables. Furthermore, girls as compared to boys are more likely to quit schooling after primary level. Similarly, a child belonging to rural areas is more likely to drop out. On the other hand, separate regression results for boys and girls reveal that distance to the nearest school matters only for girls, leaving us with conclusion that distance to the nearest secondary school is one the reasons which creates gender gap in education.

Keywords: Drop out, Secondary School, Gender Gap

INTRODUCTION

Education is considered to be one of the most important components of human capital. Every government tries its best to enhance the school enrollment ratio in order to increase literate population in the country. For this purpose, many governmental and non-governmental organizations are campaigning for improving enrollment ratio. Unfortunately, Pakistan is one of those countries having low literacy rates, low enrollment rates and higher dropout rates. Dropout of secondary level is particularly quite high in the country. Dropping out of

secondary school after the completion of primary school education is considered to be one of the most serious social problems for the developing countries including Pakistan. Early school dropout creates many other social problems like child labor, child abusing, street crimes etc. There is high disparity between boys' and girls' dropout rates. The ratio of dropping out of school at any level is highest among girls than in boys and in rural areas than in urban areas. Moreover, the latest statistics provided by Pakistan Social and Living Standard Measurements (PSLM) reveal that dropout rates at primary level are almost same across gender however, after the completion primary level the gap between boys' and girls' enrollments in secondary education becomes wider [PSLM, 2013-14]. Therefore, it is the need of the hour to focus not only on enhancing the enrollment ratio but also to lower early school dropout.

There are a number of factors responsible for gender gap in education and high dropout of girls is one the reasons. This study tries to find out those factors which are affecting secondary school dropout rate. The study further focuses on the gender disparity of secondary education and tries to investigate the reasons of higher girls' dropout rate.

LITERATURE REVIEW

Innumerable studies are available in the literature exploring the socio-economic determinants of school dropout. In Pakistan, most of the studies are conducted for specific region. Factors affecting school dropout vary across countries and regions. Most of the studies have identified a wide range of factors responsible for early school dropout. The factors can be categorized as demographic, household related and school related.

Child Characteristics

The effect of age and gender on school dropout is well documented. Increase in the age of child leads to early dropout. Thus older children are more likely to drop out (Arulampalam et al., 2004; Cardoso & Verner, 2007). Parents consider boys as the future insurance therefore, households face higher opportunity cost when enrolling a girl than when enrolling a boy. Thus gender is also found to be one of the important variables. Boys are more likely to stay in school as compared to girls whose probability of dropout increases after completion of primary education. [See for example, Cairns et al. (1989) Rumberger (1995), Connelly and Zheng (2003), Brown and Park (2002), Reynolds (2001), Cardoso and Verner (2007), Hasnain and Krantz (2010)]. However, exploring the effects of child gender and age on

school dropout some researchers came with insignificant results. For instance, Mikeet al., (2008) found that gender of child and gender of household head were not significant variables. They further found that girls belong to rural area were more likely to dropout at older age. Contrary to these studies Alspaugh (2000) and Arulampalam et al., (2004) found that boys were more likely than girls to withdraw.

Parental Education

Parental education level also plays significant role in explaining early school dropout. Children whose parental education level is high ar**e** less likely to drop out after primary level. Reynolds (2001) found that parental education is one of the significant variables affecting the school enrollment and school dropout.

Household Related Factors

Household size, poverty, and household income level are among those variables which significantly affect school dropout [Reynolds (2001), Song and Appleton (2006), Cardoso and verner (2006)]. Increase in household size increases the probability of early school dropout. In contrast to this general belief, Mike et al. (2008) found that children to larger households were less likely to dropout. It was due to fact that larger households substitute other children for child labor.

Extreme poverty and lower household income adversely affect school dropout. Household demand for high schooling is positively related with household income level. Children belonging to low income households are more likely to drop out after completion of primary standard. Song and Appleton (2006) opined that girls' schooling was more income elastic than that of boys' schooling.

Distance to Nearest School

Distance to nearest school is one of the important supply side variables. This variable matters more for girls. Most of the studies have ignored this variable especially in the studies conducted in Pakistan. The effect of this variable is not significant at primary level and for boys. Study of Mike et al. (2008) found that distance to school, gender of child and gender of household head were not significant variables at primary level. Their results showed that household size and parental education were the significant factors influencing dropout. This study was conducted in Uganda and sample population was taken from urban area where

distance to the nearest school did not significantly affect drop out. Stratton et al. (2008) showed that financial grants increased the probability of school enrollment and reduced the dropout out rate.

RESEARCH GAP AND SIGNIFICANCE OF THE STUDY

Literature on school dropout shows that most of the studies in Pakistan are conducted on specific region with very small sample size. Distance to nearest school is one of the most important variables. However, most of these studies ignore this variable. The current study tries to investigate the factors that significantly affect dropout rates in Pakistan. The study uses micro level data from HIES/PSLM for the period 2014-15. To increase school enrollment and literacy rate has been the main objective of every government. This study explores those factors which are responsible for high dropout rate after primary level. The findings of the current study may helpful for policy makers and educationists while formulating any policy to lower dropout rates in the country.

Model

Dependent variable in our model is school dropout which is a binary response variable having value 1 if student is still enrolled after the completion of primary level and zero if student dropouts after primary level. Due to categorical dependent variable, we use logit regression model.

Log it model has the following mathematical form,

$$P(Y_i = 1) = \frac{1}{1 + e^{-(a+X_iB)}} \qquad (1.1)$$

To make equation 1.1 more simpler we suppose,

$$W = a + X_iB \qquad (1.2)$$

Where, “a” is vector of intercept, “B” is vector of coefficients of independent variables and “X” is the vector of all independent variables included in the model.

By substituting equation 1.2 in equation 1.1 we obtain the following equation,

$$P(Y_i = 1) = \frac{e^W}{1+e^W} \qquad (1.3)$$

Equation 1.3 shows the probability that an event will occur. Therefore, the probability that an event will not occur can be expressed by equation 1.4, $P(Y_i = 0) = \frac{1}{(1+e^W)}$ (1.4)

By dividing equation (1.3) by equation (1.4) we get the following equation,

$$P(Y_i = 1)/P(Y_i = 0) = \left(\frac{\frac{e^W}{1+e^W}}{\frac{1}{(1+e^W)}} \right)$$

OR,

$$\frac{Pi}{1 - Pi} = e^W \qquad (1.5)$$

Now put $W = a + X_i B$ in equation 1.5 we get,

$$\frac{Pi}{1 - Pi} = e^{a+X_i B}$$

By taking natural log of the above equation we get,

$$Y_i = ln\left\{\frac{P_i}{1-P_i}\right\} = a + X_i B \qquad (1.6)$$

Our regression model has the following form,

$$Y_i = \alpha + X_i \beta + \epsilon \ldots\ldots. (1.7)$$

In equation (1), "Y" is the log of odd ratio while "X" is the vector of independent variables, β is vector of coefficients and ε is error term. Independent variables include the following four categories. First category is household related factors which includes household size, household income, and place of residency. Second category of independent variables is child-related factors which include age and gender of child. Third category of explanatory variables includes parental education level. Fourth category of independent variables uses distance to nearest school. Thus, we estimate the following regression model,

$$DO_i = \beta_0 + \beta_1 Age_i + \beta_2 Gend_i + \beta_3 HI_i + \beta_4 HS_i + \beta_5 Durb_i + \beta_6 Fedu_i + \beta_7 Medu_i + \beta_8 Dist2_i + \beta_9 Dist3_i + \beta_{10} Dist4_i + \beta_{11} Dist5_i + \boldsymbol{\varepsilon_i} \qquad (\mathbf{1.8})$$

Explanation of Variables Used in the Model

DO= Dummy variable having value one if child dropouts after the completion of primary level and zero otherwise.

Age= Child's age in years.

Gend= Gend is dummy variable having value one for male and zero otherwise.

HI= Household per capita income in PKR.

HS= Household size

Durb= Dummy variable having value one if child belongs to urban area and zero otherwise.

Fedu= Father's education is continuous variable measured in number of years of schooling.

Medu= Mother's education is continuous variable measured in number of years of schooling.

Dist1= Dummy variable having value one if distance to nearest high school is 1 to 14 minutes and zero otherwise.

Dist2= Dummy variable having value one if distance to nearest high school is 14 to 29 minutes and zero otherwise.

Dist3= Dummy variable having value one if distance to nearest high school is 30 to 44 minutes and zero otherwise.

Dist4= Dummy variable having value one if distance to nearest high school is 45 to 59 minutes and zero otherwise.

Dist5= Dummy variable having value one if distance to nearest high school is 60 minutes and above and zero otherwise.

We include five dummy variables for distance to nearest high school, however to avoid dummy variable trap we have included four dummy variables in the model.

DATA SOURCE AND WORKING SAMPLE

This study uses micro level data from nationally representative survey of PSLM for the period of 2014-15. PSLM survey provides essential information about education, household income, employment, assets ownership, health, household consumption expenditures, etc.

The current study covers children with age less than 16 years, who have completed their primary level education and are now currently enrolled in any class or drop out of school.

RESULTS AND DISCUSSION

Coefficients of equation (1.8) measure the change in the log of odds for one unit change in the value of independent variable, holding other variables constant. By taking antilog of the coefficient, we can get odd ratio or the relative ratio of the probability that an event will occur. We have run three regression models first regression is run for overall data. In order to examine the differential effects of explanatory variables on gender's dropout, we have run separate regression for boys and girls. Similarly, we have run separate regressions for urban and rural areas. Regression results of our first logit model are given in table 1.

Coefficients of all explanatory variables have expected signs and are consistent with the prior thoughts. Coefficient of age is positive and significant which suggests that probability of dropout increases with increases in age. In other words, older children are more likely to drop out after completing the primary level. These results are consistent with the earlier results of Cardoso and Verner (2007). Gender is one of the most important variables responsible for dropout. Negative and significant sign of gender shows that dropout rate is significantly different across gender. Girls as compared to boys are 100 percent more likely to drop out after the completion of primary level.

Negative and significant coefficient of dummy variable for urban area suggests that a child belonging to rural area is more likely to quit schooling after primary level. Results of our first regression model show that increase in household income and father's education reduce the probability of school dropout. But mother education and household size are not significant.

Also, we have included four distance dummies for five categories. All distance dummy variables are significant (except Dist4) and have positive signs. Thus our results support the idea that distance to the nearest school increases the probability of school dropout. Variable Dist5 has the highest coefficient among all distance dummies. Odd ratio of this variable shows that, if a child is at the distance of 60 minutes and above from secondary school, he/she will be 3.5 times more likely to quit schooling. Probability of LR statistic is very low showing that our overall model is significant at 1% level of significance.

Table 1: Regression results of logit model

Variable	*Coefficient*	*Odd ratio*	*P- value*
Constant	-6.70	0.0012	0.000
Age	0.42	1.52	0.000
Gender (male=1)	-0.67	0.51	0.000
Durban (urban=1)	-0.37	0.69	0.000
HI (household income)	-0.000095	0.99	0.000
HS (household size)	0.01	1.01	0.420
Fedu	-0.04	0.96	0.000
Medu	-0.01	0.99	0.316
Dist2	0.30	1.34	0.020
Dist3	0.32	1.38	0.010
Dist4	0.17	1.18	0.432
Dist5	1.25	3.50	0.000
Macfedden R-square	0.11		
Prob(LR statistics)	0.000		

Gender Wise Regression Analysis

In order to compare the differential effects of explanatory variables on dropout of boys and girls, we have made separate regression analysis for boys and girls. The results of our second regression model are listed in table 2. Results of the two models show very interesting facts. For example, place of residency does not matter for boys as it is insignificant. On other hand, this variable is important for girls. Girls belonging to rural area are more likely to drop out of school after completing primary education. Similarly, mother's education level has more pronounced negative effect on girls' dropout as compared to boys. Distance to the nearest high school does not affect boys' dropout as all distance dummies are insignificant except Dist5. However, distance to high school significantly affect girls' dropout.

Our results clearly show that distance is one of the reasons for higher girls' dropout rates and thus distance to nearest school is responsible for gender gap in the secondary level in Pakistan. Most of the parents are reluctant to send their daughters to school for secondary education especially when girls' high school is far away. There are many reasons behind this

positive association between distance to school and dropout. Firstly, parents become more cautious about their daughters when their daughters reach the age of puberty. Secondly, parents do not send their daughters to faraway school because they want them at home to help in domestic work. Last but not least, parents consider girls' education as wastage of time and money. Other variables like household income, household size and age have almost similar effects on school dropout for both boys and girls.

Table- 2 Regression Results of logit Model for Boys and Girls

	Boys	*Girls*
Variable	Coefficient	Coefficient
Constant	-7.4*	6.95*
Age	0.41*	0.42*
Durban (urban=1)	-0.10	-0.62*
HI (household income)	-0.000085*	-0.0002*
HS (household size)	0.01	0.016
Fedu	-0.07*	-0.02**
Medu	-0.03**	-0.05*
Dist2	0.17	0.44*
Dist3	0.30	0.40**
Dist4	0.45	0.28
Dist5	0.99*	1.80*
Macfedden R-square	0.10	0.14
Prob(LR statistics)	0.000	0.000

Note: * and** show significant at 1 percent, 5 percent significance level respectively

INTER-REGIONAL ANALYSIS

For direct comparison, we have estimated separate regressions for urban and rural areas. Results of the both equations are reported in table 3. Separate results show that in rural areas, gender has more negative effects on school dropout. Girls belonging to rural areas are more likely to dropout as compared to girls from urban areas. All distance dummies are insignificant in urban areas and highly significant in rural areas. It is due to the fact that high

schools are situated at the nearest distance in urban areas while transport facility is also available to urban children. Because of these reasons, distance to high school matters only in rural areas.

Table 3 regression results of logit model for urban and rural areas

	Urban	*Rural*
Variable	Coefficient	Coefficient
Constant	-7.62*	6.42*
Age	0.50*	0.40*
Gender (male=1)	-0.25**	-0.90*
HI (household income)	-0.0001*	-0.0002*
HS (household size)	0.02	0.015
Fedu	-0.07*	-0.03*
Medu	-0.07*	0.05*
Dist2	-027	0.52*
Dist3	-0.02	0.52**
Dist4	0.62	0.23
Dist5	-0.20	1.37*
Macfedden R-square	0.13	0.10
Prob(LR statistics)	0.000	0.000

*Note: * and** show significant at 1 percent, 5 percent significance level respectively*

CONCLUSION AND POLICY IMPLICATION

Results of our regression analysis reveal that older children are more likely to drop out after the completion of primary education while the gender of child has also significant effect on dropout with girls more likely to quit education after primary level as compared to boys. On the other hand, household size is not an important variable in context of dropout from the school. But household income negatively affects the probability of dropout with increase in household income decreasing the probability of dropout. Also, children whose parents are more educated are less likely to quit education while distance to the nearest school is one of the important variables explaining the probability of dropout.

Separate regressions for boys and girls, on the one hand, and for urban and rural areas, on the other, reveal some important and interesting results. Distance to nearest school is more important for girls in rural areas with increase in distance to the nearest school significantly enhancing chances of their dropout. However, this variable is insignificant for urban areas as well as for school going boys, leaving us with the conclusion that distance is one of the main variables responsible for gender gap in secondary education since girls who are far away from secondary school quit schooling after completion of their primary education.

The study identifies a good number of factors responsible for higher dropout rate after primary education. One of them like distance to nearest school as one of the key determinants of school dropout is very important for policymakers. Distance to the nearest school as an important variable in context of girls belonging to rural areas should be focused on to help decrease girls high dropout rate apart from addressing the issue of gender gap in secondary education by building more secondary schools.

REFERENCES

Lspaugh, J. (2000). The Effect of Transition Grade to High School, Gender, and Grade Level Upon Dropout Rates. *American Secondary Education*, *29*(1), 2-9. Retrieved from http://www.jstor.org/stable/41064411

Arulampalam, W., Naylor, R. A., & Smith, J. P. (2004). A hazard model of the probability of medical school drop-out in the UK. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *167*(1), 157-178.

Brown, P. H., & Park, A. (2002). Education and poverty in rural China. *Economics of education review*, *21*(6), 523-541.

Cairns, R. B., Cairns, B. D., &Neckerman, H. J. (1989). Early school dropout: Configurations and determinants. *Child development*, 1437-1452.

Cardoso, A. R., & Verner, D. (2007). School drop-out and push-out factors in Brazil: The role of early parenthood, child labor, and poverty.

Connelly, R., & Zheng, Z. (2003). Determinants of school enrollment and completion of 10 to 18 year olds in China. *Economics of education review*, *22*(4), 379-388.
Farid-ul-Hasnain, S., &Krantz,

G. (2011). Assessing reasons for school/college dropout among young adults and implications for awareness about STDs and HIV/AIDS: findings from a population-based study in Karachi, Pakistan. *International journal of behavioral medicine*, *18*(2), 122-130.

Gujrati, Damodar "Basic Econometrics Fourth edition" The McGraw-Hill, 2004.

Okumu, I. M., Nakajjo, A., &Isoke, D. (2008). Socioeconomic determinants of primary school dropout: the logistic model analysis.

Reynolds, A. J., Temple, J. A., Robertson, D. L., & Mann, E. A. (2001). Long-term effects of an early childhood intervention on educational achievement and juvenile arrest: A 15-year follow-up of low-income children in public schools. *Jama*, *285*(18), 2339-2346.

Rumberger, R. W. (1995). Dropping out of middle school: A multilevel analysis of students and schools. *American educational Research journal*, *32*(3), 583-625.

Stratton, L. S., O'Toole, D. M., & Wetzel, J. N. (2008). A multinomial logit model of college stopout and dropout behavior. *Economics of education review*, *27*(3), 319-331.