School dropout at the basic education level in rural Cambodia: Identifying its causes through longitudinal survival analysis

Fata No[1], Kyoko Taniguchi, Yukiko Hirakawa*

Graduate School of International Development and Cooperation, Hiroshima University, Japan

A B S T R A C T

This study sought to identify the causes of school dropout in Cambodia by following three student cohorts (grades 1–4, 4–7, and 7–9) over a period of 3 years. The results showed that economic status, child labor, and parents' aspirations had no significant effect in any of the cohorts. The odds of dropout increased significantly with divorce of parents, relationships with friends, and late school entry of students in grades 1–4, and with grade repetition and relative academic achievement of those in grades 4–9. In addition, dropout rates differed significantly between schools.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

School dropout that occurs before primary education completion is a significant issue in most developing countries. According to a UNESCO report, in 2010, around 75% of students who entered primary school completed the last grade (UNESCO, 2014). Cambodia is no exception. In 2010, the dropout rate from primary schools was high at 8.8% (Cambodian Ministry of Education, Youth and Sport [MoEYS], 2011). Despite gross enrolment and net enrolment figures of nearly 120% and 97%, respectively, the survival rate of students reaching grade 6 (the last grade of primary school) has remained at 66% (Cambodian MoEYS, 2013). Given the establishment of schools all over the country following the civil war, school capacities are adequate. Parent demand for education is high, as evidenced by the gross and net enrolment rates, especially after 2001 when the government abolished tuition fees for 9-year basic education (primary and lower secondary schools). The question, then, is why are many children still dropping out before completing their basic education?

Extensive research has been carried out in both developed and developing countries aimed at reducing the dropout rate at the level of basic education. In developed countries, many studies have used rich longitudinal data, available for example from the National Education Longitudinal Study (NELS) and the National Longitudinal Survey of Youth (NLSY), to analyze the causes of dropout (e.g. Rumberger, 1983; Bryk and Thum, 1989; Rumberger and Thomas, 2000). By contrast, although several studies have also been conducted in developing countries to examine factors influencing school dropout, their data and analyses have been limited.

This study aimed to examine the causes of school dropout in Cambodia by applying longitudinal survival analysis. Panel data were constructed for three periods of field research, and Cox regression was employed for the analysis.

2. Review of methods applied in dropout studies

2.1. Types of data and their analyses in developed countries

Commencing from the 1970s, studies on high school and college have been extensively conducted in developed countries, especially in the United States, and their research methods, including data collection and analysis, have been widely discussed. During the early stage of dropout research, retrospective data or two-wave prospective data have mainly been used to identify the causes of dropout. However, their use entails some limitations. First, they ignore differences among groups of people, as a result of which detailed analysis cannot be performed (Frase, 1989). Second, the data may have been inflated by respondents' intentional or unintentional self-justification (Rumberger, 1987). Third, they cannot describe the dynamic process of dropout that includes

* Corresponding author at: Graduate School of International Development and Cooperation, Hiroshima University, 1-5-1 Kagamiyama, Higashi Hiroshima City 730-8529, Japan.
 E-mail addresses: nofata@gmail.com (F. No), kyoko.taniguchi1909@gmail.com (K. Taniguchi), hirayuki@hiroshima-u.ac.jp (Y. Hirakawa).
[1] Now in Cambodian Ministry of Education, Youths and Sports.

background and chronic risks, trigger events, and signs that indicate the probability of dropout occurring within a short time period (Willet and Singer, 1991). Therefore, it is clear that multiwave data, termed panel data or longitudinal data, is currently the most appropriate data for assessing the causes of dropout. This is because this kind of data can enable observation of changes in a specific predictor that influences dropout decisions over time (Willet and Singer, 1991).

A longitudinal study that follows a single student cohort over several years is termed an event history or survival analysis. This kind of study was originally developed for biostatistical modeling of a human lifetime (Cox, 1972) and was subsequently extended to different fields, including education. In developed countries, it is widely recognized that the event history method is the most appropriate research design for studying the causes and dynamic patterns of dropout (Willet and Singer, 1991). There are two important advantages of this method compared with other methods such as logistic or probit regression. First, it can deal with problems of data censoring, enabling us to ascertain when dropout occurs. Second, it allows us to build time-varying covariates into the analysis. Event history predicts the hazard probabilities of each student across grades before they finally drop out when they reach a certain grade. Moreover, it enables changes in some independent predictors of the dropout event to be observed over the years preceding dropout.

2.2. Data types and analyses in developing countries

Several studies in developing countries, employing various types of data, have identified the causes of dropout. With the exception of Woldehanna and Hagos (2012), all of these studies have used retrospective data (Lloyd et al., 2000; Holmes, 2003; Sibanda, 2004; Wils, 2004; Crouch, 2005; Cardoso and Verner, 2007; OSI, 2007; Motala et al., 2009; Hussain et al., 2011; Shahidul, 2013) or two-wave prospective data (Hanushek and Levy, 1994; Motala, 1995; Akhtar, 1996; Hanushek et al., 2006; Lloyd et al., 2006; Barnson et al., 2013; Sabates et al., 2013; Jukes et al., 2014). One reason for this is the lack of accumulated data. In almost all developing countries, cross-sectional data, but not longitudinal data, are available. This situation contrasts with that in developing countries.

Woldehanna and Hagos (2012) used multiwave data for the "Young Lives" study in Ethiopia. In this study, students aged between 7.5 and 8.5 years in 2002 were followed in 2006 and in 2009. As described in the previous section, multiwave data are appropriate for examining the causes of dropout. Therefore, this type of data should be more widely applied in studies.

Previous studies have applied several statistical methods in their analyses. The selection of methods depended on the types of data. Studies using retrospective data applied descriptive statistics (UNESCO, 1984; Motala, 1995; Wils, 2004; Crouch, 2005; OSI, 2007; Motala and Dieltiens, 2009; Wegner et al., 2008; Hussain et al., 2011), or employed ANOVA (Hussain et al., 2011; Mzuza et al., 2014) and stepwise linear regression (Levy, 1971). Meanwhile, studies using two-wave prospective data employed logistic regression (Sibanda, 2004; Townsend et al., 2008; Sabates et al., 2013; Jukes et al., 2014), probit models (Hanushek and Levy, 1994), OLS (Holmes, 2003; Hanushek et al., 2006) and the instrumental variables approach (Cardoso and Verner, 2007; Hanushek et al., 2006).

Willet and Singer (1991) have shown that event history analysis, which predicts hazard probabilities, is the most appropriate method for examining the factors underlying dropout. However, a few studies have used hazard functions, including a discrete-time hazard model (Lloyd et al., 2000, 2006), hazard rate analysis (Akhtar, 1996), and a Cox proportional hazard model (Woldehanna and Hagos, 2012).

2.3. A review of dropout studies in Cambodia

A few dropout studies have been conducted in Cambodia, all of which have used retrospective data. Based on a study of the reasons given by girls, and their parents, for dropping out of school, Valesco (2001) concluded that poverty, housework responsibility, and cultural beliefs were the main factors underlying the earlier occurrence of school dropout among girls than among boys. Keng (2003), who also focused on girls in one rural village, conducted a causal comparative study to examine differences between drop-outs and non-dropouts. Initially, she found that there was no difference in the economic situations of the families of girls who had dropped out of school and those who had not. However, based on her subsequent analysis of the same data using logical regression and including data for boys, Keng (2004) concluded that gender (a higher dropout rate among girls), age at school entry, academic performance, and parental education all influenced the decision to drop out at a significance level of 0.01. The area of farmland owned by a student's family also influenced this decision at a significance level of 0.05.

It should be noted that the dropout situation changed drastically in 2001 when tuition was abolished nationwide. The gross enrolment rate rose from 100% to over 120% in 1 year. Nevertheless, during the data collection period for the above two studies by Velasco (2001) and Keng (2004), it was evident that more girls dropped out of school than boys. According to a UNDP report, the net primary enrolment of girls from 1995 to 1997 was 92%, while that of boys was 100% (UNDP, 2001). However, after 2001, the disparity between boys and girls in primary education was markedly reduced. In the latest data published for the academic year 2012/2013, the net enrolment rate of girls in primary school was 97%, which was the same as that of boys (Cambodian MoEYS, 2013). The number of girls in primary schools was 1,022,983, which was 47.1% of the total enrolment figure of 2,173,384. However, this phenomenon can be explained by the fact that boys' repetition of grades exceeded those of girls by 1.46 times. Girls are also positively inclined to continue their education up to the secondary school level. The advancement rate to lower secondary school was 77.0% for female students compared with a rate of 75.6% for all students (Cambodian MoEYS, 2013). After the abolishment of tuition, there have been some changes in the causes of dropout.

A World Bank study (2005) that compared dropout rates within districts across Cambodia found that poor rural districts, with low adult literacy rates, had higher dropout rates than rich, urban districts with high literacy rates. The study reported that geographic location, poverty, and low literacy rates within communities were the main causes of dropout. While it revealed the characteristics of districts in which dropout was likely to occur, it did not suggest ways of preventing dropout.

Adopting a narrative approach, Edward et al. (2014) recently noted five factors associated with continued school attendance: (1) The need for a school–family connection, (2) reduced costs of basic materials such as uniforms and school supplies, (3) proximity of school to residence, (4) awareness of the amount or level of education required, and (5) the need for improved infrastructure, health services, and scholarships. However, the study was based solely on interviews.

This study follows up on those of No et al. (2012) and No and Hirakawa (2012), which were aimed at clarifying the causes of dropout in Cambodia using the event history method. Both boys and girls were included in the analysis. The focus was only on rural areas, because dropout rates are higher in rural areas compared with urban areas, and the majority of the population lives in rural areas. To collect data on factors that possibly affect student dropout, the researchers carried out a comprehensive review of related literature on the causes of dropout, both in developed and developing countries, and from primary school extending to higher