When Learners Leave School:

A Survival Analysis of Dropout

A Basic Research

[Copy for blind review: the researchers' names, positions, and schools division are removed. The final copy lists them here.]

Conducted April to November 2026

Abstract

This basic research examined when junior high school learners left school during School Year (SY) 2025–2026 in one schools division, and how their absences in the first weeks of the year, distance from home, overage, 4Ps membership, and sex related to leaving. It used a retrospective cohort of the records of all 10,615 learners enrolled in Grades 7 to 10 of the division's 18 public secondary schools, followed through the 39 school weeks of the year. Of them, 335 (3.16%) were tagged No Longer in School and did not return; 263 transferred out. Leaving rose from September and peaked in the first week after each break: 29 learners left in the week after the mid-year break and 51 in the week after the year-end break. Among the 10,564 learners still in school at week 5, a Cox regression stratified by grade, with standard errors clustered by school, found that learners absent five days or more in the first 20 class days had a hazard of leaving 3.52 times that of learners absent once or not at all, 95% CI [2.55, 4.85]; boys 2.05 times girls' [1.58, 2.66]; overage learners 1.73 times [1.28, 2.34]; and learners living more than 5 km away 1.59 times [1.27, 1.99]. Learners from 4Ps households had a lower hazard, 0.73 [0.57, 0.95], clearer among girls than boys. The association of five or more early absences did not hold all year: in a landmark model of weeks 5 to 26 it was 5.86 [4.19, 8.21], and after the year-end break too few of these learners left to show one. Absences in the first month, read together with distance, age for grade, and sex, can mark learners for follow-up before the mid-year and year-end breaks.

Keywords: dropout, survival analysis, Kaplan-Meier, Cox regression, early warning, junior high school, absences

Acknowledgment

The researchers thank the Schools Division Superintendent, who approved the study and the use of the division's records; the Planning Officer, who prepared the coded extract from the Learner Information System; the school heads and class advisers of the 18 public secondary schools, who opened their School Form 2 binders for the first weeks of the year and answered our questions about the records; the research ethics committee that reviewed the protocol; and the members of the Schools Division Research Committee. The study was funded by the Basic Education Research Fund (BERF) for 2026.

When Learners Leave School: A Survival Analysis of Dropout

Introduction of the Research

Junior high school reported 148,222 dropouts nationally in SY 2022–2023 (Second Congressional Commission on Education [EDCOM II], 2026, p. 590), and the school leaver rate of Grades 7 to 10 rose to 5.7% in SY 2023–2024, 7.4% among boys and 3.8% among girls (Albert et al., 2025, p. 23). The Basic Education Development Plan 2030 sets the aim plainly: "DepEd schools will aim for zero dropouts and zero school leavers," and notes that "Recent data show dropouts are highest in Grade 7 and Grade 8" (Department of Education [DepEd], 2022a, pp. 131, 133).

Schools already record what an early warning needs. The class adviser marks each day's attendance on School Form 2 and visits the home of a learner "absent for 5 consecutive days and/or those at risk of dropping out" (DepEd, 2014a, back page, guideline 5); when a learner stops coming, the school tags the learner No Longer in School (NLS) in the Learner Information System (LIS), with an effective date, "date after the last day of 'actual' attendance in school" (DepEd, 2014b; Office of the Undersecretary for Strategic Management [OUSM], 2025, item 5). What schools seldom do is read these records together over time. EDCOM II found that "there is no explicit policy requiring routine, data-driven risk pattern analysis" and recommended "systematic risk profiling of learners" using "attendance records, learning assessments, and socio-economic information" (EDCOM II, 2026, p. 195).

Most studies of dropout ask whether a learner left by the end of a year. A study of the records can ask when (Willett & Singer, 1991). Knowing when

learners leave tells a division when follow-up matters most, and knowing which records mark the learners who leave early tells it whom to follow up. The division had the records but no such analysis: it knew its dropout count each year, not the weeks in which learners left or how well the first month's absences foretold leaving.

This study therefore followed every junior high school learner enrolled in the division's public secondary schools at the beginning of SY 2025–2026 through the year's school weeks, and asked when they left and how absences in the first 20 class days, distance from home, overage, 4Ps membership, and sex related to leaving. It addresses the Governance theme of the Basic Education Research Agenda, under Planning ("Data management" and "Monitoring and evaluation"), with Inclusive Education as its cross-cutting theme (DepEd, 2016, pp. 6, 13).

Literature Review

The literature reviewed here is organized around the study's questions: how much and when learners leave, what marks the learners at risk, and how survival analysis answers questions about when.

Dropout in Philippine Junior High Schools

DepEd counts leaving in two ways. The simple dropout rate is the share of a grade's beginning-of-year enrollment that did not finish the grade in that year; the school leaver rate also counts learners who finished a grade but did not enroll in the next, and is "internationally known as the dropout rate" (DepEd, 2022b, pp. 12, 18). The junior high school simple dropout rate was 2.42% in SY 2019–2020 (DepEd, 2022a, p. 180). The school leaver rate of

Grades 7 to 10 ranged from 3.8% to 6.2% between SY 2017–2018 and SY 2020–2021, and was 5.7% in SY 2023–2024, with regional rates from 3.4% in the National Capital Region to 11.6% in BARMM (Albert et al., 2025, pp. 23–25). Boys leave at about twice the rate of girls (Albert et al., 2025, p. 23).

Out-of-school children name their own reasons. Among out-of-school children aged 12 to 15 in 2022, lack of personal interest was the reason most often given by boys (33.4%), and among the reasons girls named, illness or disability (20.0%) and marriage, family matters, or pregnancy (16.6%); the accessibility of school was named by 2.0% of boys and 3.0% of girls (Albert et al., 2025, pp. 30–32). Fieldwork found that absences grow "as Christmas shopping season begins in September and October … children will begin missing more days until they end up eventually dropping out," and that older children "disappear for about 1-2 months at a time" during planting and harvest (David et al., 2018, p. 20).

Early Absences and Other Markers of Risk

Attendance is among the strongest early markers in studies that follow learners over years. In Philadelphia, attending school 80% of the time or less in Grade 6 was one of four flags that together identified 60% of the students who would not graduate; only 13% of students with that flag graduated on time (Balfanz et al., 2007, pp. 223, 227). In Chicago, 63% of freshmen who missed about one week a semester graduated in four years, against 87% of those who missed less (Allensworth & Easton, 2007, p. 6). The first month matters in itself: students who missed more than four days in September were over 16 times as likely to be chronically absent over the year as students who

missed fewer than two (Olson, 2014), though that study predicted chronic absence, not leaving.

Being older than the grade's official age is another marker. Overage children "are at a high risk of eventually dropping out of school"; they are usually overage "because they had already stopped attending at some point then came back, or they did not pass the grade level and were held back from promotion" (David et al., 2018, p. 10), and grade retention is "one of the most powerful predictors of dropout status" across 17 studies (Jimerson et al., 2002, p. 441). Balfanz et al. (2007) add a caution: overage students without other flags "graduated at the same rate as the overall cohort" (p. 228), so overage may matter less once absences are counted. Distance is seldom the reason children give, but "cost or time of transportation from a family's village to the nearest school could be prohibitive" and makes it likely that "a student will start a year but fail to complete it" (David et al., 2018, p. 26). In a study of four countries, boys had a higher hazard of dropping out in two of them, with hazard ratios from 0.72 to 1.96 (Cueto et al., 2020, p. 54).

Cash Transfers and Leaving

The Pantawid Pamilyang Pilipino Program (4Ps) Act requires children five to 18 years old to "attend elementary or secondary classes at least eighty-five percent (85%) of their time" (Pantawid Pamilyang Pilipino Program [4Ps] Act, 2019, Sec. 11[e]). Its third impact evaluation found no effect on attendance, which was already about 98%, and lower dropout among Pantawid children; for ages 12 to 17, "a 4-percentage point reduction in drop-out rates was observed among female children," with no significant effect for

boys (Orbeta et al., 2021, pp. 50, 54–55). The evaluators call this "a missed opportunity for the program because boys are at greater risk of dropping out" (Orbeta et al., 2025, p. 272).

Studying When: Survival Analysis

Survival analysis studies the time until an event. Its first tool, the Kaplan-Meier estimate, gives the share still "surviving" at each time while counting learners whose records end without the event, here those who transfer or reach the end of the year, as censored (Kaplan & Meier, 1958). The Cox model then relates the hazard, the rate of the event among those still at risk, to explanatory variables, as an unknown function of time "multiplied by" a function of the variables (Cox, 1972, p. 187), so that each variable has one hazard ratio for the whole period. That proportional-hazards assumption can be checked by plotting scaled Schoenfeld residuals against time and testing their slope (Schoenfeld, 1982; Grambsch & Therneau, 1994). Studies of dropout have used it with months of schooling (Plank et al., 2008) and, where an effect was not proportional, let the effect differ by period instead of dropping it (Cox Edwards & Ureta, 2003). Time recorded in weeks gives many learners who leave in the same week; Efron's approximation handles these ties well (Hertz-Picciotto & Rockhill, 1997), and a discrete-time model gives the same estimates in another way (Allison, 1982; Singer & Willett, 1993). A Philippine study of college students' exits used the method with years as the unit, and named its own limit: "exits are not accurately recorded (i.e., ideally, they should be in months, not years)" (Bayudan-Dacuycuy, 2025, p. 4).

Synthesis

Dropout in Philippine junior high schools is concentrated in Grades 7 and 8 and among boys. Research elsewhere marks absences, age for grade, and distance as early signs, and the 4Ps evaluation found lower dropout among girls in the program. No study in the Philippine literature reviewed here has followed a division's learners week by week through one school year with the records schools already keep, which is what a division needs to time and target its follow-up.

Research Questions

This study sought to answer the following questions:

1. How many of the division's junior high school learners left school during SY 2025–2026, and in which weeks of the year did they leave?
2. Do the shares still in school at the end of the year differ by absences in the first 20 class days, distance from home, overage, 4Ps membership, and sex?
3. With the others held constant, how is each of these associated with the hazard of leaving?
4. Does the association of early absences with leaving hold through the whole school year?

Hypotheses

The following null hypotheses were tested at the .05 level of significance:

1. H01: The survival curves of learners do not differ by absences in the first 20 class days, distance from home, overage, 4Ps membership, or sex.

2. H02: With the others held constant, none of these is associated with the hazard of leaving.
3. H03: The hazard ratio of early absences is the same before and after the year-end break.

Scope and Limitation

The study covered all learners enrolled in Grades 7 to 10 of the division's 18 public secondary schools at the beginning of SY 2025–2026, followed from June 16, 2025 to March 31, 2026 through the school's own records: School Form 2, the LIS, and the enrollment form. It did not cover senior high school, private schools, or the Alternative Learning System.

The study has limits. First, it is observational: the associations it reports are not causes, and a hazard ratio for 4Ps membership compares households that differ in more than the program. Second, it uses records, which are only as good as their keeping; a dropout recorded where dropout is a performance target may be understated (EDCOM II, 2026, p. 162). Third, distance was computed from the barangay of residence, not the home. Fourth, the year is one school year, so leaving between years is outside it. Fifth, the absences used are those of the first 20 class days only, so the study can say how well the first month foretells leaving, not how the year's absences unfold.

Research Methodology

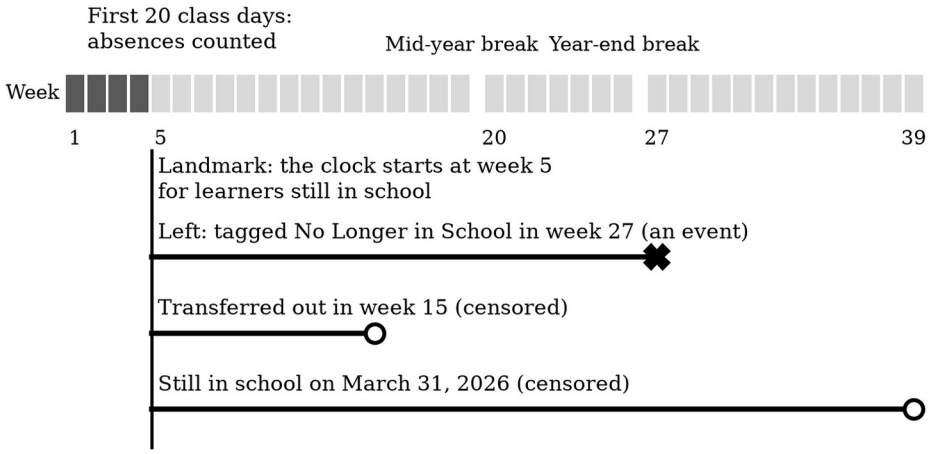
Research Design

The study used a retrospective cohort design with survival analysis. A retrospective cohort follows a defined group forward in time through records that already exist: here, every junior high school learner enrolled at the

beginning of the year, through the 39 school weeks of SY 2025–2026. The year opened on Monday, June 16, 2025 and ended on Tuesday, March 31, 2026, with 197 class days (DepEd, 2025a, para. 4). The mid-year break (October 27 to 31, 2025) and the year-end break (December 20 to 31, 2025; classes resumed on January 5, 2026) were left out of the time scale, with the three weeks that held no class days, so that time ran only while learners could leave; week 20 is November 3 to 7 and week 27 is January 5 to 9, 2026 (Figure 1).

Figure 1

The Design of the Study on the School Year



Note. Each line is a learner's record from the landmark at week 5. A record ends in an event (tagged No Longer in School and not readmitted) or is censored (transferred out, or still in school at the end of the year).

The outcome was leaving: a learner tagged No Longer in School whose tag stood at the end of the year, which by default makes the learner a dropout (DepEd, 2014b, item 4). The time of leaving was the school week containing the NLS effective date; an effective date that fell on a weekend or in a break was placed in the next school week. Learners who transferred out were

censored at the week of transfer, and learners still in school were censored at week 39. A learner tagged NLS and readmitted later in the year was not counted as leaving.

The models used a landmark at week 5. Absences in the first 20 class days (June 16 to July 11, 2025) cannot predict leaving that happens during those days, so the models included only the learners still in school at the start of week 5 and counted time from there.

A. Sampling

The study did not sample. Its population was every learner enrolled in Grades 7 to 10 of the division's 18 public secondary schools at the beginning of SY 2025–2026, 10,615 learners: 2,867 in Grade 7, 2,771 in Grade 8, 2,568 in Grade 9, and 2,409 in Grade 10. Six of the schools are in the division's town centers and twelve in rural districts.

The number of learners was fixed by the records; what decides the precision of a survival analysis is the number of events. With $\alpha = .05$, two-sided, and power .80, a Cox model needs $(z.975 + z.80)^2 / [p(1 - p)(\ln HR)^2]$ events to detect a hazard ratio HR for a trait carried by a share p of the learners (Schoenfeld, 1983). At the planning values of the proposal, rounded up, it needs 197 events to detect a hazard ratio of 2.5 for five or more early absences (5% of learners), 77 for a ratio of 1.9 between boys and girls, 299 for a ratio of 1.5 for living more than 5 km away (20%), 483 for a ratio of 1.6 for overage (8%), and 407 for a ratio of 0.75 for 4Ps membership (37%). The division's junior high school simple dropout rate in the previous year, about 3.3%, gave about 350 expected events: enough for the first three, not for

overage or 4Ps at those sizes, so their estimates would be reported with their intervals. The rule of ten events per variable (Peduzzi et al., 1995) was met many times over.

B. Data Collection

Recruitment of Participants

The study used records, not participants. The Schools Division Superintendent approved the study and the use of the division's records in April 2026, and the Planning Officer prepared a coded extract from the LIS: one row per learner, with a code in place of the name and the Learner Reference Number. No learner, parent, or teacher was asked to do anything for the study.

Data Collection Instruments

Records extract. For each learner, the LIS gave the grade, sex, birth date, status at the end of the year (in school, transferred out, or No Longer in School), and the effective date of any transfer or NLS tag. The 4Ps status came from the enrollment form, which asks, "Is your family a beneficiary of 4Ps?" (DepEd, 2025b, Basic Education Enrollment Form). The barangay of residence came from School Form 1.

School Form 2 abstraction sheet. The researchers copied each learner's absences in the first 20 class days, June 16 to July 11, 2025, from the class's School Form 2, and the reason recorded for each learner who left, using the form's own legend of reasons (DepEd, 2014a).

Derived variables. Absences in the first 20 class days were grouped as 0 or 1, 2 to 4, and 5 or more, the last matching the five consecutive absences

after which School Form 2 asks for a home visit. A learner was overage when two or more years older on the first day of school than the official age for the grade, 12 in Grade 7 (DepEd, 2022b, Annex 1, p. 2). Distance was the road distance from the center of the barangay of residence to the school, measured on a map by the researchers; a learner living more than 5 km away was coded as far.

Validity and Reliability

The records were checked rather than validated by experts, since no scale was built. A second researcher re-abstracted a random 10% of the records (1,062 learners) without seeing the first abstraction. Absences in the first 20 class days agreed closely, ICC(A,1) = .97; status at the end of the year agreed for 98.4% of learners, Cohen's κ = .86; and of the 57 learners whom both abstractions recorded as having left or transferred, 96.5% were placed in the same week. Disagreements were settled against the school's own records.

Data Collection Setting

The records were read at the division office (the LIS extract) and at each school (School Form 2), after the end of the school year, when the year's tags were final.

Data Collection Timeline

Table 1

Data Collection Timeline

Activity	Dates
Approval of the study and the data request	April 6 to 24, 2026
Review by the research ethics committee	April 13 to May 8, 2026

Activity	Dates
LIS extract prepared by the Planning Officer	May 11 to 15, 2026
School Form 2 abstraction in the 18 schools	May 18 to June 5, 2026
Record check (10% re-abstracted)	June 15 to 19, 2026
Distance from the barangay of residence	June 22 to July 3, 2026
Data analysis	July 6 to August 28, 2026

Note. SY 2025–2026 ended on March 31, 2026 (DepEd, 2025a), so every learner's status for the year was final before the records were read.

Data Recording and Storage

The extract and the abstraction sheets carried codes only. The key linking codes to learners stayed with the Planning Officer and never left the division office. The data were kept on an encrypted drive; they will be destroyed three years after the study ends.

C. Ethical Issues

The study followed DepEd's Research Management Guidelines (DepEd, 2017) and the Data Privacy Act of 2012. Education records are sensitive personal information under the Act (Data Privacy Act of 2012, 2012, Sec. 3[l]), and DepEd allows its learner data to be used for "efforts congruent to the development of education policies, programs, projects, and activities in DepEd" (DepEd, 2025b, Sec. X.A). The study used a coded extract, kept to the variables it needed. Because the study used records and could not practicably seek the consent of 10,615 families, its protocol was reviewed by a research ethics committee, which approved the use of the coded records without individual consent. The national guidelines allow such a waiver, with a committee's approval, when four conditions are met: the research presents no more than minimal risk; the waiver will not adversely affect the rights and welfare of the participants; the research cannot be practicably carried out

without it; and the participants will be given pertinent information afterward, whenever appropriate (Philippine Health Research Ethics Board [PHREB], 2022, p. 22). No result names a learner or a school, and the study's file was never used to decide anything about a learner. The researchers declared their use of an artificial intelligence (AI) tool in Appendix D (DepEd, 2026).

D. Plan for Data Analysis

The first question was answered by counts and the simple dropout rate by grade and sex, by the number of learners who left each week, and by the Kaplan-Meier estimate of the share still in school at the end of each week, with transfers out censored (Kaplan & Meier, 1958). Because a transfer ends a record before it could end in leaving, the share who left was also estimated with transfers as a competing event (the Aalen-Johansen estimate). The second question was answered, for the learners still in school at week 5, by Kaplan-Meier estimates in each group and the log-rank test. The third was answered by a Cox regression of the hazard of leaving on the five traits, stratified by grade, so that each grade had its own baseline hazard, with standard errors clustered by school and tied weeks handled by Efron's method (Cox, 1972; Hertz-Picciotto & Rockhill, 1997). As a check, the same model was fitted as a discrete-time model on learner-weeks with a complementary log-log link, whose coefficients estimate the same hazard ratios (Allison, 1982). The fourth question was answered by testing the scaled Schoenfeld residuals of each trait against time (Grambsch & Therneau, 1994) and, to see how a changing ratio changed, by two landmark models: one of weeks 5 to 26, ending every record at the last week before the year-end break,

and one of the learners still in school at week 27, followed to week 39. With 18 schools, the clustered standard errors rest on few clusters, and lifelines notes that its robust estimator does not handle tied times exactly; the model-based standard errors differed little. The analysis used Python 3.13 with lifelines 0.30.3 (Davidson-Pilon, 2019) and statsmodels 0.15; the workbook recomputes every statistic Excel can.

Discussion of Results and Recommendations

Who Left, and When

Of the 10,615 learners, 335 (3.16%) left during the year and 263 (2.48%) transferred out; 24 more were tagged No Longer in School and readmitted later in the year. Leaving fell by grade, from 4.05% in Grade 7 to 2.08% in Grade 10, and boys left at more than twice the rate of girls, 4.28% against 1.99% (Table 2). By the Kaplan-Meier estimate, 96.8% of learners were still in school at the end of week 39, 95% CI [96.4, 97.1]; counting transfers as a competing event instead, 3.15% left, almost the same.

Table 2

Learners Who Left and Transferred Out, by Grade and Sex

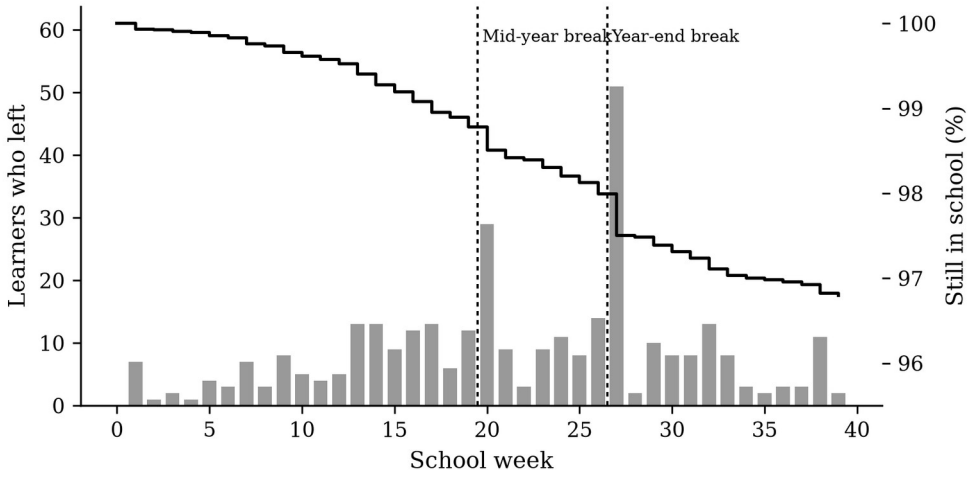
Group	Learners	Left	Transferred out	Simple dropout rate
Grade 7	2,867	116	63	4.05%
Grade 8	2,771	95	75	3.43%
Grade 9	2,568	74	64	2.88%
Grade 10	2,409	50	61	2.08%
Girls	5,221	104	125	1.99%
Boys	5,394	231	138	4.28%
All	10,615	335	263	3.16%

Note. Simple dropout rate: learners who left during the year over the beginning-of-year enrollment (DepEd, 2022b, p. 18). Transfers out are not leavers.

Leaving was not spread evenly over the year (Figure 2). It was low in June and July, rose from September, and peaked in the first week after each break: 29 learners left in week 20, the first week after the mid-year break, and 51 in week 27, the first week after the year-end break, the most of any week. January, with 71 leavers, lost the most learners of any month.

Figure 2

Learners Who Left Each School Week, and the Share Still in School



Note. Bars: learners tagged No Longer in School in each school week. Line: Kaplan-Meier share still in school, transfers censored. The breaks are left out of the time scale.

The reasons recorded on School Form 2 for the 335 learners who left are in Table 3. Lack of interest was recorded for a third of them, and child labor or work for more than a fifth; distance was recorded for 15. A recorded reason is the adviser's reading of a learner's situation, and the study did not use it as a predictor.

Table 3

Reasons Recorded on School Form 2 for the Learners Who Left

Reason (School Form 2 code)	Learners	Share
Lack of interest/Distractions (b.6)	111	33.1%

Reason (School Form 2 code)	Learners	Share
Child labor, work (e.1)	76	22.7%
Family problems (a.4)	38	11.3%
Had to take care of siblings (a.1)	26	7.8%
Poor academic performance (b.5)	24	7.2%
Illness (b.1)	19	5.7%
Early marriage/pregnancy (a.2)	15	4.5%
Distance between home and school (d.1)	15	4.5%
Overage (b.2)	11	3.3%

Note. The legend of reasons is School Form 2's (DepEd, 2014a). One reason per learner.

Leaving by Early Absences, Distance, Age for Grade, 4Ps, and Sex

Of the 10,615 learners, 10,564 were still in school at week 5; 324 of them left during the rest of the year. Table 4 gives the share still in school at the end of the year in each group. The clearest difference was by early absences: 97.6% of learners absent once or not at all in the first 20 class days were still in school at the end of the year, against 95.9% of those absent two to four days and 90.1% of those absent five days or more, log-rank $x^2(2) = 102.00$, $p < .001$ (Figure 3). The curves also differed by distance, age for grade, and sex. They did not differ significantly by 4Ps membership, $x^2(1) = 2.87$, $p = .090$. The first null hypothesis is rejected for early absences, distance, overage, and sex, and retained for 4Ps membership.

Table 4

Share Still in School at the End of the Year, by Group

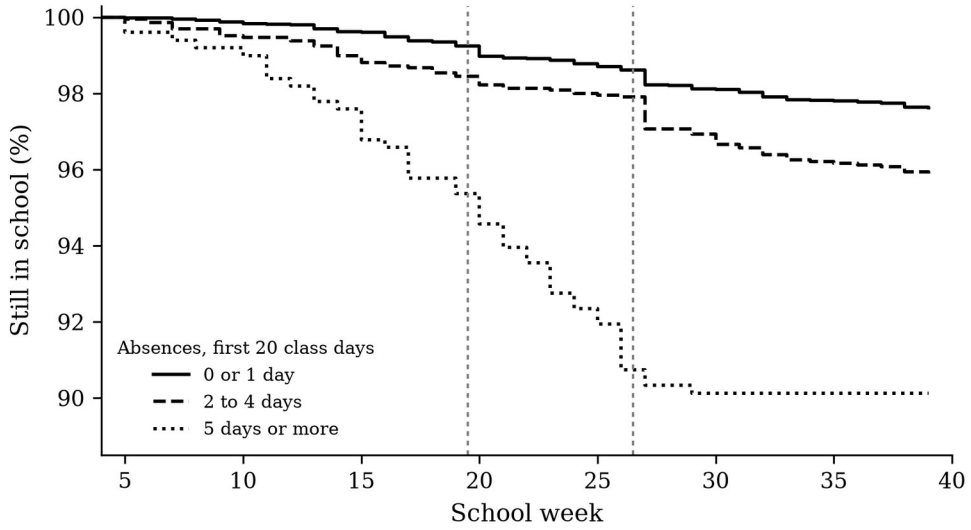
Group	Learners	Left	Still in school at week 39, % [95% CI]	Log-rank test
Absences, first 20 class days: 0 or 1 day	7,799	183	97.6 [97.3, 97.9]	$x^2(2) = 102.00$, $p < .001$
Absences, first 20 class days: 2 to 4 days	2,268	92	95.9 [95.0, 96.6]	
Absences, first 20 class days: 5 days or more	497	49	90.1 [87.2, 92.4]	

Group	Learners	Left	Still in school at week 39, % [95% CI]	Log-rank test
Distance from home: 5 km or less	8,819	241	97.2 [96.9, 97.6]	$x^2(1) = 20.06$, $p < .001$
Distance from home: More than 5 km	1,745	83	95.2 [94.1, 96.1]	
Age for grade: Official age or 1 year over	9,677	269	97.2 [96.8, 97.5]	$x^2(1) = 32.64$, $p < .001$
Age for grade: Overage by 2 years or more	887	55	93.7 [91.9, 95.1]	
4Ps: Not 4Ps	7,031	230	96.7 [96.3, 97.1]	$x^2(1) = 2.87$, $p = .090$
4Ps: 4Ps household	3,533	94	97.3 [96.7, 97.8]	
Sex: Girls	5,201	99	98.1 [97.7, 98.4]	$x^2(1) = 46.83$, $p < .001$
Sex: Boys	5,363	225	95.8 [95.2, 96.3]	

Note. Learners still in school at week 5. Kaplan-Meier estimates, transfers censored.

Figure 3

The Share Still in School by Absences in the First 20 Class Days



Note. Kaplan-Meier estimates for the learners still in school at week 5. The dotted lines mark the breaks.

The Cox Model

With the other traits held constant, every trait was associated with the hazard of leaving (Table 5, Figure 4). Learners absent five days or more in the first 20 class days had a hazard of leaving 3.52 times that of learners absent

once or not at all, 95% CI [2.55, 4.85], and learners absent two to four days 1.61 times [1.29, 2.01]. Boys had 2.05 times the hazard of girls [1.58, 2.66], overage learners 1.73 times that of learners of the official age [1.28, 2.34], and learners living more than 5 km away 1.59 times that of learners living nearer [1.27, 1.99]. Learners from 4Ps households had a lower hazard, 0.73 [0.57, 0.95]. The second null hypothesis is rejected for all five traits. The model ranked learners by risk moderately well, concordance = .67, and with 324 events for six coefficients it had 54 events per variable.

Table 5

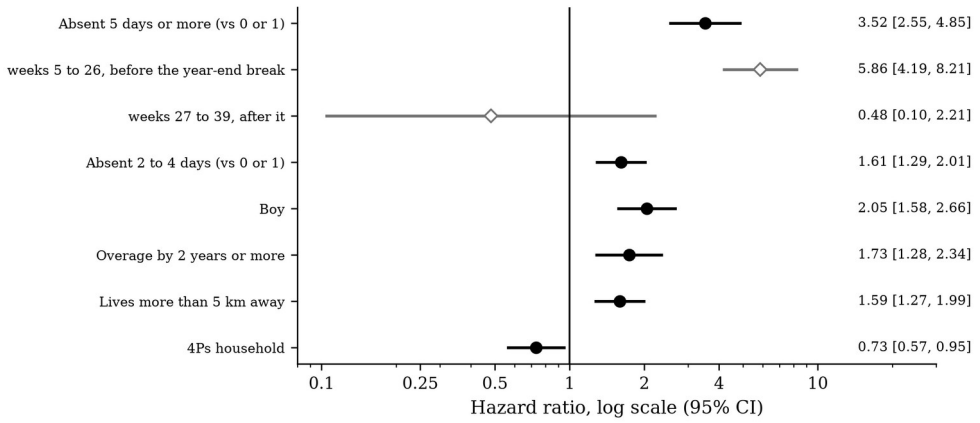
Hazard Ratios of the Cox Model

Trait	*b*	*SE*	Hazard ratio	95% CI	*p*
Absent 2 to 4 days (vs 0 or 1)	0.476	0.114	1.61	[1.29, 2.01]	< .001
Absent 5 days or more (vs 0 or 1)	1.258	0.164	3.52	[2.55, 4.85]	< .001
Lives more than 5 km away	0.465	0.113	1.59	[1.27, 1.99]	< .001
Overage by 2 years or more	0.551	0.153	1.73	[1.28, 2.34]	< .001
4Ps household	−0.311	0.131	0.73	[0.57, 0.95]	.018
Boy	0.717	0.134	2.05	[1.58, 2.66]	< .001

Note. N = 10,564 learners still in school at week 5; 324 left. Stratified by grade; standard errors clustered by the 18 schools; tied weeks by Efron's method.

Figure 4

Hazard Ratios of Leaving, With 95% Confidence Intervals



Note. Log scale. The diamonds are the hazard ratios of five or more early absences in the two landmark models of the fourth question: weeks 5 to 26 and weeks 27 to 39.

The discrete-time model, fitted on 360,601 learner-weeks, gave hazard ratios within 0.01 of the Cox model's for every trait (for five or more absences, 3.52 [2.53, 4.91]). Clustering by school changed the standard errors little.

Among girls, 4Ps membership went with a hazard of 0.51 [0.32, 0.81]; among boys, 0.85 [0.64, 1.11], not significant. The difference between the two was at the edge of significance, p = .051, so the study can say that the lower hazard was clear among girls, not that the association differed between boys and girls.

Does the Association Hold All Year?

The scaled Schoenfeld residuals of five or more early absences fell over the year, $x^2(1) = 12.89$, $p < .001$, against Kaplan-Meier time ($p = .031$ against the rank of time); no other trait's test fell below .05 (Table 6). The third null hypothesis is rejected for five or more early absences: the whole-year ratio of 3.52 was an average of something that changed. The two landmark models show how. From week 5 to week 26, 200 learners left, and those absent five

days or more had a hazard of leaving 5.86 times that of learners absent once or not at all [4.19, 8.21]. Of the 10,233 learners still in school at week 27, 124 left by week 39, only 3 of them from that group, too few to show an association: 0.48 [0.10, 2.21]. Two to four absences went with leaving in both periods, 1.42 [1.00, 2.02] and 1.85 [1.31, 2.60]. The other traits' ratios moved within their intervals; overage's was larger after the break (2.78 [1.70, 4.56]) than before it (1.27 [0.82, 1.95]), though its residual test found no change, p = .112.

Table 6

Tests of Proportional Hazards, and the Landmark Models Before and After the Year-End Break

Trait	Test, Kaplan-Meier time: $x^2(1)$	p	Test, rank of time: $x^2(1)$	p	Weeks 5 to 26: hazard ratio [95% CI]	Weeks 27 to 39: hazard ratio [95% CI]
Absent 2 to 4 days (vs 0 or 1)	0.13	.718	3.51	.061	1.42 [1.00, 2.02]	1.85 [1.31, 2.60]
Absent 5 days or more (vs 0 or 1)	12.89	< .001	4.64	.031	5.86 [4.19, 8.21]	0.48 [0.10, 2.21]
Lives more than 5 km away	0.07	.789	0.01	.904	1.58 [1.17, 2.14]	1.60 [1.15, 2.22]
Overage by 2 years or more	2.53	.112	1.73	.188	1.27 [0.82, 1.95]	2.78 [1.70, 4.56]
4Ps household	0.22	.639	0.08	.778	0.78 [0.61, 1.00]	0.66 [0.49, 0.89]
Boy	1.18	.277	2.03	.155	2.13 [1.58, 2.87]	1.93 [1.37, 2.73]

Note. Scaled Schoenfeld residuals, model without clustering (Grambsch & Therneau, 1994). The landmark models: learners in school at week 5, followed to week 26 (n = 10,564; 200 left), and learners in school at week 27, followed to week 39 (n = 10,233; 124 left); each stratified by grade, standard errors clustered by school.

Discussion

The division's learners left at a rate close to DepEd's junior high school figures: 3.16% within the year, against a simple dropout rate of 2.42% in SY 2019–2020 (DepEd, 2022a, p. 180), and more in Grades 7 and 8, as the Basic

Education Development Plan 2030 found nationally (DepEd, 2022a, p. 133). The weekly records add what a yearly rate cannot: leaving rose from September, the months in which fieldwork in other divisions found absences beginning to grow (David et al., 2018, p. 20), and peaked in the first week after each break. A learner who does not come back after a break is, in these records, the learner most likely to be lost.

Absences in the first month marked the learners who left, as attendance has marked them in studies of whole school careers (Allensworth & Easton, 2007; Balfanz et al., 2007), and the first month itself was enough (Olson, 2014). The check of proportional hazards showed what kind of marker five or more absences are: they marked the learners who left before the year-end break; after it, too few of them left to show an association, while two to four absences went with leaving in both periods. A first-month record is a short-range warning, most useful in the months before the year-end break. Being a boy and living far from school went with leaving in both periods. Overage kept its association once absences were in the model, unlike in Philadelphia, where overage students without other flags did as well as their cohort (Balfanz et al., 2007, p. 228). The 4Ps result follows the third impact evaluation, which found lower dropout among girls and no significant effect among boys (Orbeta et al., 2021, p. 55); in an observational study, though, 4Ps households differ from others in more than the program.

Recommendations

1. **Read the first month's attendance.** At the end of the first 20 class days, each school should list its learners absent five days or more, with

their distance from school, age for grade, and sex, and plan the follow-up School Form 2 already asks for: home visits and interventions for learners at risk of dropping out (DepEd, 2014a).

2. **Follow up before and after the breaks.** The weeks after the mid-year and year-end breaks lost the most learners. Schools should contact learners on their list before each break and on the first days after it, and treat a learner who has not returned by the end of the first week as urgent.
3. **Look again in January.** Five or more absences in the first month marked the learners who left before the year-end break; after it, too few of them left to tell. Schools should list learners again from the attendance of the second half of the year.
4. **Pay attention to boys, overage learners, and learners far from school.** Their higher hazard was seen over the whole year. The division should review where learners living more than 5 km away come from and what transport would help them.
5. **Study it further.** A study that follows the same learners over two or more years, with each week's attendance as a time-varying measure, would show how attendance builds up to leaving, and a trial of a follow-up after the breaks would show whether it keeps learners in school.

Dissemination and Advocacy Plans

The findings will be shared with those who can act on them, in the forms each can use (Table 7).

Table 7

Dissemination and Advocacy Plan

Audience	Purpose	Channel	Timeline
School heads and guidance designates	Act on the first month's attendance and the weeks after the breaks	Division school heads' conference; a two-page policy brief	November 2026
Division planning and monitoring units	Build the early warning list into the division's routine reports	Division management committee meeting	December 2026
Class advisers	Know when follow-up matters most	Learning action cell sessions in each school	January to February 2027
Researchers in the region	Share the method and the findings	Regional research congress; the division's research journal	March to June 2027

Note. The policy brief was paid from the grant's dissemination line (Financial Report).

References

Albert, J. R. G., Monterola, S. L. C., Cabalfin, D. L. D., & Mahmoud, M. A. (2025). *Out-of-school children in the Philippines: Post-pandemic realities and emerging challenges at the educational spectrum extremes* (PIDS Discussion Paper Series No. 2025-34). Philippine Institute for Development Studies. https://doi.org/10.62986/dp2025.34

Allensworth, E. M., & Easton, J. Q. (2007). *What matters for staying on-track and graduating in Chicago public high schools: A close look at course grades, failures, and attendance in the freshman year*. Consortium on Chicago School Research at the University of Chicago.

Allison, P. D. (1982). Discrete-time methods for the analysis of event histories. *Sociological Methodology, 13*, 61–98. https://doi.org/10.2307/270718

Balfanz, R., Herzog, L., & Mac Iver, D. J. (2007). Preventing student disengagement and keeping students on the graduation path in urban middle-grades schools: Early identification and effective interventions. *Educational Psychologist, 42*(4), 223–235. https://doi.org/10.1080/00461520701621079

Bayudan-Dacuycuy, C. G. (2025). *Hazard and incidence of exits of ever-enrolled college-age students* (PIDS Discussion Paper Series No. 2025-12). Philippine Institute for Development Studies. https://doi.org/10.62986/dp2025.12

Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society: Series B (Methodological), 34*(2), 187–202. https://doi.org/10.1111/j.2517-6161.1972.tb00899.x

Cox Edwards, A., & Ureta, M. (2003). International migration, remittances, and schooling: Evidence from El Salvador. *Journal of Development Economics, 72*(2), 429–461. https://doi.org/10.1016/S0304-3878(03)00115-9

Cueto, S., Felipe, C., & León, J. (2020). *Predictors of school dropout across Ethiopia, India, Peru and Vietnam* (Documentos de Investigación No. 109). Grupo de Análisis para el Desarrollo.

Data Privacy Act of 2012, Republic Act No. 10173 (2012).

David, C. C., Albert, J. R. G., & Vizmanos, J. F. V. (2018). *Out-of-school children: Changing landscape of school attendance and barriers to completion* (PIDS Discussion Paper Series No. 2018-25). Philippine Institute for Development Studies. https://doi.org/10.62986/dp2018.25

Davidson-Pilon, C. (2019). lifelines: Survival analysis in Python. *Journal of Open Source Software, 4*(40), 1317. https://doi.org/10.21105/joss.01317

Department of Education. (2014a). *Adoption of the modified school forms (SFs) for public elementary and secondary schools effective end of school year 2013-2014* (DepEd Order No. 4, s. 2014).

Department of Education. (2014b). *About the No Longer in School (NLS) facility* [Learner Information System user guide].

https://support.lis.deped.gov.ph/support/Manuals/NLS%20screen%20shot%20guides.pdf

Department of Education. (2016). *Adoption of the Basic Education Research Agenda* (DepEd Order No. 39, s. 2016).

Department of Education. (2017). *Research management guidelines* (DepEd Order No. 16, s. 2017).

Department of Education. (2022a). *Adoption of the Basic Education Development Plan 2030* (DepEd Order No. 024, s. 2022).

Department of Education. (2022b). *Guide in computing key performance indicators* [Memorandum of October 12, 2022, with Annex 1, Guide in computing education performance indicators].

Department of Education. (2025a). *Multi-year implementing guidelines on the school calendar and activities* (DepEd Order No. 012, s. 2025).

Department of Education. (2025b). *Revised basic education enrollment policy* (DepEd Order No. 017, s. 2025).

Department of Education. (2026). *Foundational guidelines on artificial intelligence in basic education* (DepEd Order No. 003, s. 2026).

Grambsch, P. M., & Therneau, T. M. (1994). Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika, 81*(3), 515–526. https://doi.org/10.1093/biomet/81.3.515

Hertz-Picciotto, I., & Rockhill, B. (1997). Validity and efficiency of approximation methods for tied survival times in Cox regression. *Biometrics, 53*(3), 1151–1156. https://doi.org/10.2307/2533573

Jimerson, S. R., Anderson, G. E., & Whipple, A. D. (2002). Winning the battle and losing the war: Examining the relation between grade retention and dropping out of high school. *Psychology in the Schools, 39*(4), 441–457. https://doi.org/10.1002/pits.10046

Kaplan, E. L., & Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association, 53*(282), 457–481. https://doi.org/10.1080/01621459.1958.10501452

Office of the Undersecretary for Strategic Management. (2025, March 28). *Guidance and clarifications on the School Form 10 for end of school year 2024-2025 and reiteration of senior high school status tagging* (Memorandum STR-250331-0910-PS). Department of Education.

Olson, L. S. (2014). *Why September matters: Improving student attendance* [Policy brief]. Baltimore Education Research Consortium.

Orbeta, A. C., Jr., Melad, K. A. M., & Araos, N. V. V. (2021). *Reassessing the impact of the Pantawid Pamilyang Pilipino Program: Results of the third wave impact evaluation* (PIDS Discussion Paper Series No. 2021-05). Philippine Institute for Development Studies. https://doi.org/10.62986/dp2021.05

Orbeta, A. C., Jr., Melad, K. A. M., & Araos, N. V. V. (2025). Demand intervention: Assessing the impact of the Pantawid Pamilyang Pilipino Program on education. In *Making Philippine education reforms and interventions more transformative* (pp. 249–280). Philippine Institute for Development Studies.

Pantawid Pamilyang Pilipino Program (4Ps) Act, Republic Act No. 11310 (2019).

Peduzzi, P., Concato, J., Feinstein, A. R., & Holford, T. R. (1995). Importance of events per independent variable in proportional hazards regression analysis. II. Accuracy and precision of regression estimates. *Journal of Clinical Epidemiology, 48*(12), 1503–1510. https://doi.org/10.1016/0895-4356(95)00048-8

Philippine Health Research Ethics Board. (2022). *National ethical guidelines for research involving human participants 2022*. Department of Science and Technology – Philippine Council for Health Research and Development.

Plank, S. B., DeLuca, S., & Estacion, A. (2008). High school dropout and the role of career and technical education: A survival analysis of surviving high school. *Sociology of Education, 81*(4), 345–370. https://doi.org/10.1177/003804070808100402

Schoenfeld, D. (1982). Partial residuals for the proportional hazards regression model. *Biometrika, 69*(1), 239–241. https://doi.org/10.1093/biomet/69.1.239

Schoenfeld, D. A. (1983). Sample-size formula for the proportional-hazards regression model. *Biometrics, 39*(2), 499–503. https://doi.org/10.2307/2531021

Second Congressional Commission on Education. (2026). *Turning point: A decade of necessary reform (2026–2035)* [EDCOM II final report].

Singer, J. D., & Willett, J. B. (1993). It's about time: Using discrete-time survival analysis to study duration and the timing of events. *Journal of Educational Statistics, 18*(2), 155–195. https://doi.org/10.3102/10769986018002155

Willett, J. B., & Singer, J. D. (1991). From whether to when: New methods for studying student dropout and teacher attrition. *Review of Educational Research, 61*(4), 407–450. https://doi.org/10.3102/00346543061004407

Financial Report

The study received a division-level Basic Education Research Fund grant of PhP 15,000.00, released in tranches: 40% (PhP 6,000.00) on the submission of the work plan and the instruments and 40% (PhP 6,000.00) on the submission of the data collection and initial findings; the last 20% (PhP 3,000.00) is released on the acceptance of this report (DepEd, 2017). Expenses totaled PhP 14,250.00, PhP 750.00 less than the grant. The analysis used free software, and the record check was done by the researchers.

Table 8

Statement of Expenses

Date	Particulars	Receipt or document	Amount (PhP)	Category
2026-05-12	Printing of the School Form 2 abstraction sheets (600 pages)	OR 1184	1,200.00	Reproduction
2026-05-18 to 2026-06-12	Transportation to the 18 schools for the abstraction (36 trips)	Itinerary and receipts	3,600.00	Transportation
2026-05-18 to 2026-06-12	Snacks for the abstraction days (6 persons, 18 days)	OR 2270	2,700.00	Food
2026-05-01 to 2026-09-30	Mobile data for the three researchers (five months)	Receipts	4,500.00	Communication
2026-05-12	Folders, envelopes, and an encrypted flash drive	OR 0935	1,050.00	Supplies
2026-11-09	Printing of the policy brief (60 copies)	OR 6102	1,200.00	Dissemination
	Total		14,250.00	

Note. Summary by category: communication, PhP 4,500.00; transportation, PhP 3,600.00; food, PhP 2,700.00; reproduction, PhP 1,200.00; dissemination, PhP 1,200.00; supplies, PhP 1,050.00.

Appendices

Appendix A. The Variables and Their Sources

Learner code (Planning Officer); grade, sex, and birth date (LIS); status at the end of the year and the effective date of any transfer or No Longer in School tag (LIS); 4Ps membership (Basic Education Enrollment Form); barangay of residence (School Form 1); absences in the first 20 class days, June 16 to July 11, 2025, and the reason recorded for a learner who left (School Form 2). Derived: the school week of leaving or transfer (1 to 39, breaks skipped; an effective date on a weekend or in a break placed in the next school week); absence group (0 or 1, 2 to 4, 5 or more); overage (two or more years above the official age for the grade on the first day of school); distance from the barangay center to the school by road, and far (more than 5 km).

Appendix B. The School Form 2 Abstraction Sheet

For each learner of the class, by code: the number of days absent from June 16 to July 11, 2025 (20 class days), counted from the class's School Form 2; for a learner who left, the reason recorded in the form's legend (DepEd, 2014a) and the last day of attendance. The adviser of the class was asked only to explain an entry that could not be read.

Appendix C. The School Weeks of SY 2025–2026

Week 1, June 16 to 20, 2025, to week 19, October 20 to 24; the mid-year break, October 27 to 31, skipped; week 20, November 3 to 7, to week 26, December 15 to 19; the weeks of the year-end break, December 22, 2025 to January 2, 2026, skipped (the break ran December 20 to 31, and classes

resumed January 5); week 27, January 5 to 9, 2026, to week 39, March 30 to 31, 2026 (DepEd, 2025a). The workbook's Weeks sheet lists every week.

Appendix D. Declaration of AI Use

We, the researchers, declare that we used [the name of the AI tool, as the researchers would give it] only to check the grammar of the English text of this report, for the purpose of finalizing the research study. The prompt used was: "Check the grammar of this paragraph. Do not change its meaning or add content." We have reviewed, edited, and understood all the content, and we take full accountability for it. No codes, records, or other data from the study were entered into the AI tool.