The Comparison of Discrete and Continuous Survival Analysis

Sunha Kim

Dissertation submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
In
Educational Research and Evaluation

Mido Chang, Co-Chair
Kusum Singh, Co-Chair
Katherine R. Allen
Inyoung Kim

April 15, 2014
Blacksburg, Virginia

Keywords: survival analysis, discrete model, and continuous model

The Comparison of Discrete and Continuous Survival Analysis

Sunha Kim

ABSTRACT

There has been confusion in choosing a proper survival model between two popular survival models of discrete and continuous survival analysis. This study aimed to provide empirical outcomes of two survival models in educational contexts and suggest a guideline for researchers who should adopt a suitable survival model. For the model specification, the study paid attention to three factors of time metrics, censoring proportions, and sample sizes. To arrive at comprehensive understanding of the three factors, the study investigated the separate and combined effect of these factors. Furthermore, to understand the interaction mechanism of those factors, this study examined the role of the factors to determine hazard rates which have been known to cause the discrepancies between discrete and continuous survival models. To provide empirical evidence from different combinations of the factors in the use of survival analysis, this study built a series of discrete and continuous survival models using secondary data and simulated data. In the first study, using empirical data from the National Longitudinal Survey of Youth 1997 (NLSY97), this study compared analyses results from the two models having different sizes of time metrics. In the second study, by having various specifications with combination of two other factors of censoring proportions and sample sizes, this study simulated datasets to build two models and compared the analysis results. The major finding of the study is that discrete models are recommended in the conditions of large units of time metrics, low censoring proportion, or small sample sizes. Particularly, discrete model produced better outcomes for conditions with low censoring proportion (20%) and small number (i.e., four) of large time metrics (i.e., year) regardless of sample sizes. Close examination of those conditions of time metrics, censoring

proportion, and sample sizes showed that the conditions resulted into high hazards (i.e., 0.20). In conclusion, to determine a proper model, it is recommended to examine hazards of each of the time units with the specific factors of time metrics, censoring proportion and sample sizes.

Acknowledgements

Most of all, I would like to thank my co-advisor, Dr. Mido Chang, who has offered all the intellectual, emotional, spiritual, and financial support to start and complete this dissertation. She has been a great source of strength for me and has helped me to overcome all of the challenges throughout this long dissertation process. Also, I would like to extend my thankfulness to all other committee members: Dr. Kusum Singh, my co-advisor, who has provided excellent guidance and practical solutions to maintain and complete my dissertation; Dr. Katherine Allen, who has provided unceasing inspiration via warm smiles and sincere support; and Dr. Inyoung Kim, who has asked insightful questions so that I would aim higher.

At the same time, I would like to give thanks to my parents, Kwanjoong Kim and Youngsook Song, who have created the best learning environment for me to pursue my dreams and who have always trusted my potential and offered sincere and faithful prayers with love. Last but not least, it was God who started good work in me and who has just completed it.

Table of Contents

List of Tables

CHAPTER ONE

INTRODUCTION

This study deals with survival analysis, which is a statistical method to analyze categorical longitudinal data. Although survival analysis is a proper method to explore educational phenomena such as student dropouts, teacher attrition and retention, student persistence in STEM field, and attainment of important developmental milestones over the time, the use of survival analysis in the field of education has a short history.

There is another issue in the use of survival analysis for educational research. There has been confusion in adopting an appropriate survival model among the two types of available survival analyses. Despite clear conceptual and application differences between discrete- and continuous-time survival models, exiting studies have adopted either discrete or continuous survival models without clear rules or theoretical guidance.

The study has been motivated to provide clear guidelines in adopting a proper survival model in educational research and provide both empirical and theoretical evidence. In developing rationale for a survival model, this study took two methodological approaches: Empirical investigation using empirical data and theoretical inquiry using simulated data. The first approach explored the influence of time metrics which cause the discrepancies in parameter estimation and fit statistics between discrete and continuous models. As a real educational context, the study chose the events of high school student dropouts.

Importantly, this study paid focused attention to time metrics of dropouts because the time metric, the unit to measure the timing of the event, is the one that makes the conceptual difference between the two models. The first approach as an empirical study built both discrete and continuous survival models and compared the parameter estimates and fit statistics using four different time metrics resulting in eight models. The study compared the results of these eight models to explore the differences between discrete and survival model outcomes.

The second approach using simulated data was to provide comprehensive guidelines by considering hazard rate. The hazard rate is one of the main factors which leads to the differential outcomes of survival analyses and it is determined by three factors of time metrics, censoring proportions, and sample sizes. Having different combinations of the three factors, the study simulated 60 sets of data for each of two survival models (discrete and continuous models) with a total of 120 sets. Using the outcomes of data analyses, particularly parameter estimates and fit statistics, the study compared the discrete- and continuous models.

The study results are expected to inform the importance of choosing a proper model to prevent biased results to researchers in educational fields. The study will also provide organized findings of data analyses with different hazard rates and their time metrics, censoring proportions, and sample sizes. Basing on the outcomes of the current study, the goal of the study is to provide clear guidelines for the selection of the proper survival model in educational research.

Background of the Study

Survival analysis is an advanced statistical method that deals with dichotomous outcomes in longitudinal data. Survival analysis has been prevalently adopted to estimate time to events such as death, recovery from disease, or treatment responses in the field of medicine and biology (Klein & Moeschberger, 2003). Recently, survival analysis has been used in the field of social science by modeling the occurrence of events such as student dropouts or teacher attrition in a longitudinal time frame (Kelly, 2004; Kirby, Berends & Naftel, 1999; Ma & Willms, 1999; Murphy, Gaughan, Hume & Moore, 2010; Plank, DeLuca, & Estacion, 2008).

Discrete- and Continuous-Time Estimation

Survival analysis estimates a hazard function, also called a conditional risk, such that a target event will occur given that the target event has not occurred yet. It also uses two time estimation methods: discrete-time and continuous-time estimations (Singer & Willet, 2003). Discrete time estimation uses the time to a target event in a large time metric such as a year, quarter, or month. Continuous time estimation records an event time with a precise and fine metric such as an hour, day, or week.

Despite the clear distinction between discrete- and continuous-time estimation, literature reviews show that there has been confusion over the use of time estimation when building survival analysis models. Many studies have adopted either of them without clear distinctions of time metrics (Calcagno, Crosta, Bailey & Jenkins, 2007; Doyle, 2006; Donaldson & Johnson, 2010; Kahn & Schwalbe, 2010; Kelly, 2004; Kirby, Berends & Naftel, 1999; Ma & Willms, 1999; Murphy, Gaughan, Hume & Moore, 2010; Murtaugh, Burns & Schuster, 1999; Singer, 1992; Singer, Davidson, Graham & Davidson, 1998). The adoption of a survival model without an empirical and theoretical understanding can be problematic because an incorrect model results in biased estimates and incorrect conclusions.

Time Metrics

As mentioned previously, the study paid attention to time metrics, censoring proportions, and sample sizes as critical factors to determine hazard. When using large time metrics (i.e., year) to measure timing of the target event, there will be a small number of waves (i.e., 3 years). The large measurement units will cause all events to a small number of waves and many events will be recoded into the same waves, which is called ‘tied event.’ The high tied event is known to cause the discrepancy of outcomes between discrete and continuous models.

On the other hand, fine time metrics (i.e., day) will lead to a large number of waves and allow events to occur at different time points across different waves, resulting in a small number of tied events. As an example, Hofstede and Wedel (1999) showed the effect of the size of time metrics by applying different time metrics in survival models. When Hofstede and Wedel aggregated a small time metric (originally a day) into large time metrics (a week or a month), they found the discrepancies in parameter estimates from the two models. The hazard estimates in discrete models were overestimated while those in continuous models were underestimated.

Censoring Proportion

Censored data indicate samples (subjects) that do not experience a target event during the study period. Censoring is an important factor because censoring proportion influences the number of cases (subjects) that experience the target event, and changes the number of tied observations and hazard rates. According to Hertz-Picciotto and Rockhill (1997), censoring reduced the estimate biases because it kept the low tie proportion. Although a few researchers such as Colosimo, Chalita, and Demétrio (2000) and Hess (2009) paid attention to censoring proportions to explore the discrepancies between discrete and continuous survival models, there is not enough research that provides detailed information on the effects of various censoring proportions by linking them to hazard rates.

Sample Size

There are studies that examined the effect of sample size for the use of survival analyses, but they do not provide clear guidelines for the proper use of survival analysis. The results of the studies presented mixed results regarding the sample sizes that cause discrepancies between the discrete and continuous models. Hess' study (2009) found different results in parameter estimation from discrete and continuous models when the sample size was as large as 1,000 or more.

Contrastingly, the study conducted by Hertz-Picciotto and Rockhill (1997) detected the discrepancies of estimated parameters between the two models with a small sample. The mixed results of those studies can be caused by study designs without considering hazard rate. Thus, the current study attempted to investigate the effect of sample size on the discrepancy in parameter estimation between two models while keeping other factors constant and considering hazard rate.

Hazards

This study paid attention to hazard rate which has been known as a critical factor determining the discrepancy between a discrete and a continuous model. Particularly, the high hazard rate leads to the high discrepancies in outcomes from these two models as Singer and Willet (2003) evidenced significant differences in the two models. Hosmer and Lameshow (1999) also found the discrepancies between two models with time metrics with hazards higher than 0.15. Hess (2009) highlighted 0.3 hazard rate as an important factor to make a difference in parameter estimation of discrete and continuous models. In the following section, the study presented the brief description and definitions of basic terms of survival analysis.

Hazard is estimated by dividing the number of events in each period by the number of the risk set (cases that did not experience the target event until that time) as follows (Singer & Willet, 2003):

$h(t_{\mathrm{ij}})$ = n events_j / n at risk_j

where n events_j refers to the number of samples with an event occurrence in the time period j and n at risk_j indicates the number of samples who have not yet experienced the event up to the time period j.

In the discrete model, hazard is transformed into log odds (logits). In the continuous model, such as Cox proportional hazard model, hazard is estimated by an instantaneous change in the occurrence rate of the event (Singer & Willet, 2003). More specifically, Cox proportional hazard model uses complementary log-log (Clog-Log) as one of instantaneous change measuring methods. Clog-log model estimates hazard as a complementary log-log probability, the logarithm of the negated logarithm of the probability of event nonoccurrence (Allison, 2010).

First Approach

In response to the prevalent confused use of a survival model in educational fields, this study aimed to provide evidence of different parameter estimation from different survival analyses which cause researchers to arrive at different conclusions using the same data and the same research hypotheses. A long term goal of the study is to suggest a guideline in the adoption of one of two major survival analyses, discrete or continuous model. The first approach of the study is to extensively employ discrete and continuous survival models using empirical data.

To analyze important factors to influence the potential discrepancies between discrete- and continuous-time survival analyses, the study paid attention to time metrics considering a time metric is a major factor that makes the conceptual and definitional difference between discrete- and continuous-time survival models (Singer & Willet, 2003). Specifically, the study employed four different time metrics of month, quarter, half-year, and year. The choice of the four time metrics was driven by the frequent use in current education research.

Datasets and Variables

In search of proper time estimations and metrics, in the first approach an empirical investigation was conducted by analyzing the National Longitudinal Survey of Youth 1997 (NLSY97). By using 2,216 high school students from NLSY97, the study has built a series of survival analysis models to estimate a student's dropout risk and its association with contextual factors. For the outcome variable, the study chose a high-school dropout (target event), and for predictor variables the study selected two race indicators (Black, Hispanic), the English language learner (ELL) indicator, parental educational levels, students' perception towards teachers, student GPA, and student gender.

Analyses

The primary analysis of the study was survival analyses. Prior to the primary analysis, the study conducted a series of preliminary analyses to obtain essential information of the data, including descriptive statistics, and correlation analyses. Also, the study built logistic regression model to explore the effect of predictors on the final outcome of dropout status.

To compare discrete- and continuous- time survival analyses, the study built a series of survival analyses using different time metrics. By applying four different time metrics to the discrete- and continuous- time survival analyses models, this study had eight models in the first approach. For a discrete-time survival analysis, this study adopted a logit model with a general time specification, thus building four logit models: Year model, Half-Year model, Quarter model, and Month model.

For a continuous-time survival analysis, the study employed a Cox Proportional Hazard model, having four Cox models: Year model, Half-Year model, Quarter model, and Month model. To compare the survival models comprising of different time models and metrics, the study examined parameters of individual predictors and the fit statistics.

Research Questions

The first approach of the study had one overarching research question and three embedded questions to the main question.

The overarching question asks,

- How different are the outcomes of the discrete- and continuous- survival models in four time metric conditions?

The following are embedded research questions to the main question in the first approach.

- Do time metrics influence discrepancies of parameter estimation between the two survival model outcomes? If they do, what are the time metric conditions which show discrepancy?

- Do time metrics lead to discrepancies of fit statistics between two survival models? If they do, what are the time metric conditions which show discrepancy? Which model will demonstrate better performance in terms of fit statistics? Which time metric (year, half-year, quarter, and month) will be the most appropriate unit for a discrete model?
- Which time metric will be the most appropriate unit for a continuous model?

Second Approach

Followed by the first approach of using empirical data, the second approach built a series of survival analysis models using simulated data. When comparing discrete- and continuous- survival models, the second study took a more comprehensive approach of including multiple factors to determine the potential discrepancy between these two models. As factors to determine the discrepancy, the second approach included three factors of time metrics, censoring proportion, and sample size (Colosimo, Chalita, & Demétrio, 2000; Hertz-Picciotto & Rockhill, 1997; Hess, 2009; Hofstede & Wedel, 1999). The study paid special attention to hazard rates, which have been identified as an important variable to make a distinction between the two models, by influencing three factors (Hess, 2009; Hosmer & Lameshow, 1999; Singer & Willet, 2003).

Data

To explore the discrepancies between discrete- and continuous- models in various data conditions, the study simulated a series of data using SAS. The study simulated 60 sets of data by combining the following conditions of three factors: three time metrics (4, 12, and 48), five censoring proportions (0%, 20%, 40%, 60%, and 80%), and four sample sizes (50, 100, 500, and 1,000).

Analyses

The study employed both discrete- and continuous- survival analyses to 60 sets of simulated data resulting from combining different levels of three factors of the interest, thus building a total of 120 models. This study adopted a logit model with a general time specification for a discrete-time survival analysis, having 60 logit models.

The study chose a clog-log model for continuous-time survival analysis, building 60 clog-log models. The clog-log model is one of sub-models of Cox proportional hazard model which features the most conservative method to treat tied observations and calculates a clog-log hazard for each time period, which enables the direct comparison with the discrete model. The parameter estimates and fit statistics resulting from the two model analyses were compared to investigate the discrepancies between two estimation methods.

Research Questions

To achieve the research goal to provide a guideline for researchers to adopt a proper survival model in educational fields, the second approach asked the following overarching research question:

- How different are the outcomes of the discrete- and continuous- survival models in various data conditions?

The study also addressed the next four embedded research questions in answering the main research question:

- Do time metrics influence discrepancies between the two model outcomes? If they do, in which condition of time metrics does a discrepancy show?
- Do censoring proportions associate with discrepancies between the two model outcomes? If they do, in which condition of censoring proportions does a discrepancy appear?
- Are sample sizes related to discrepancies between the two model outcomes? If they are, in which condition of sample sizes does a discrepancy reveal?
- Do hazard rates lead to discrepancies between the two model outcomes? If they do, in which condition of hazard rates does a discrepancy appear?

Organization of the Study

Followed by Introduction, this study in Chapter two provides comprehensive literature review on two major types of survival analysis and other important factors of survival analysis. The literature review includes the sections on time metrics, censoring proportions, sample sizes, and hazards which are the main factors to determine the discrepancies between discrete and continuous survival models in the study.

Chapter three describes two methodological approaches of the study particularly exploring the factors which cause the discrepancies between discrete and continuous survival models. The first half of Chapter three explains empirical dataset for the first approach, preliminary analyses, steps of building eight survival models to explore the impact of time metrics, and methods to compare analyses results from two models.

The second half of Chapter three discusses simulated data, steps of building survival models to investigate the influence of time metrics, censoring proportions and sample sizes for the second approach. It ends with the method descriptions to compare discrete and continuous models using 120 data and survival analyses.

Chapter four presents analyses results of two methodological approaches. First approach results include the comparison of the results between discrete- and continuous survival analyses across different time metrics. Results of the second approach explain the comparison of results between two models across various time metrics, censoring proportions and sample sizes.

Chapter five expounds summary of the findings, discussion and implications of the study findings, and study limitations.

CHAPTER TWO

LITERATURE REVIEW

Survival analysis predicts whether and when an event happens and what factors are related to the event happening. In order to conduct survival analysis, a researcher needs to pay attention to important features of survival analysis including target event, initial condition at a beginning of time, and time metrics. At the initial stage of analysis, a researcher needs to have a target event which is the main interest of the research. The researcher also needs to include only the samples (subjects) who have not experienced the target event at the beginning of the study. The researcher needs to determine the size of time metrics to measure the timing of the event. The researcher can use either a discrete time unit in a large scale or a continuous time unit in a fine scale (See Allison, 2010; Singer & Willet, 2003).

The time metric is important as its unit makes conceptual and operational difference between two major models (discrete vs. continuous models) of survival analysis. Discrete survival analysis measures the target events in comparatively larger time units assuming that two or more samples (subjects) will experience the target events at the same time. On the other hand, continuous-time survival analysis uses very fine time metrics in measuring the target events. In theory, all samples will experience the target events at different time points when using a truly continuous time metric. Thus, continuous survival analysis is very sensitive to the events that two or more samples (subjects) experience the target events simultaneously, which are called, tied events (ties) (See Allison, 2010; Hosmer & Lemeshow, 1999; Singer & Willet, 2003).

Survival analysis estimates the event occurrence as hazard, a conditional probability that an individual (subject) will experience a target event in a specific time period. The discrete survival analysis adopts hazard as log odds (logits), while continuous survival analysis estimates hazard as instantaneous change in the event occurrence rate. There are several types of continuous survival models including Cox proportional hazards model and clog-log model. Cox proportional hazards model, which is one of the most popular continuous survival models because of its efficiency, produces parameter estimates without calculating a base hazard function. The parameter estimation without producing hazard can appeal the empirical researchers who wish to deal with final analysis results and reach their research conclusions directly. However, the convenience of Cox proportional model can be problematic for researchers who study the methodological issues of survival analysis. Without the hazard function, the efficiencies of a survival analysis cannot be properly determined.

As stated in the goal of the study, this study planned to examine the hazards of the survival analyses. In order to generate hazard, the study chose complementary log-log (clog-log) model which produces relative hazard complementing Cox proportional model. In other words, clog-log model produces clog-log hazards for a respective time period which is equivalent to the values of the baseline hazard function for Cox model. In sum, the study compared the hazard of logits from the discrete survival and relative hazard of clog-log from the continuous survival model.

Despite clear distinction between discrete- and continuous- time survival analyses, there has been confusion over adopting a proper model in education research fields. Researchers have used either discrete or continuous survival analyses without empirical or theoretical considerations. Inappropriate models, however, can end up with poor fit statistics, and biased parameter estimates. For example, inflated values of log-likelihood statistic (LL), Bayesian Information Criterion (BIC) or Akaike Information Criterion (AIC) can be caused by a wrongly chosen survival model. It is imperative to come up with a clear guideline for researchers who should adopt a suitable model for their research and help them to reach valid research conclusions.

This study identified hazard as a critical factor to determine the discrepancies between discrete and continuous survival models. Studies showed that discrete and continuous survival models tended to show discrepancies for the time periods with high hazards. Furthermore, as factors to influence hazard, this study specified different conditions of time metrics, censoring proportions, and sample sizes. When using larger time metrics (i.e., year) to measure timing of the target event, there were a smaller number of waves (i.e., 1). When using smaller time metrics such as month, there were a large number of waves (i.e., 12). The large measurement units forced many events into a small number of waves, causing many tied events.

The high tied events would cause high hazards as many samples (subjects) experienced the event simultaneously within respective time periods. Censored data indicate samples (subjects) that do not experience a target event during the study period. Censoring proportion is an important factor because censoring proportion influences the number of cases (subjects) that experience the target event. High censoring proportion leads to low hazards as there are a small number of tied events. The sample size is important as the same size influences the number of tied events which determine hazards.

Survival Analysis

Using longitudinal data, survival analysis estimates whether, when, and why an event of interest (target event) occurred (Singer & Willet, 2003). To conduct survival analysis, researchers should consider not only the target event and the onset of the event but also the time metrics, tied events, and censoring proportion. A target event refers to a phenomenon whose occurrence a researcher is interested in. An event occurrence indicates an individual's experience of a change from one condition to another. A beginning of time refers to a time when every individual in the population demonstrates the same initial condition prior to a transition into another condition (an event occurrence) (See Allison, 2010; Singer & Willet, 2003).

A metric of time refers to units of measuring the passage of time between the initial time and the time when an individual experiences a target event (event time). The time metrics can be measured in either discrete or continuous time units. Discrete time units record the time passage in a coarse (broader) metric (i.e., semester or year), while continuous time units record the time passage in a fine (more precise) metric (i.e., day, hour, or minute). The time metrics also determine the tied observations, which are defined as two or more observations that occur at the same time. In other words, when there are smaller time units, there will be fewer tied observations (See Allison, 2010; Singer & Willet, 2003).

Censoring proportion refers to the cases to which the target event does not occur during the data collection of the study. It is important to note that survival analysis takes censoring proportion into account in its modeling, which makes its design stronger than other longitudinal analyses (See Allison, 2010; Hosmer & Lemeshow, 1999; Singer & Willet, 2003).

Discrete-Time Hazard Models

Discrete-time survival analysis estimates the risk (probability) of the target event's occurrence in comparatively larger time units. The risk is estimated as a conditional probability that the event of interest will occur. The discrete-time survival analysis assumes that two or more observations will occur simultaneously, as it uses broad time metrics. Thus, the discrete-time survival analysis is recommended in the case of heavy ties, which does not lead to biased estimates for those conditions (Singer & Willet, 2003).

The risk is addressed as a hazard ($h(t_{ij})$): the conditional probability that the event of interest will occur to an individual *i* in time period *j*, assuming that the event of interest has not occurred to that individual up to that time. The hazard of discrete- time survival analysis is estimated as below:

$$h(t_{ij}) = \text{n events}_j \,/\, \text{n at risk}_j$$

where n events$_j$ refers to the number of samples with an event occurrence in time period *j* and n at risk$_j$ indicates the number of samples who have not yet experienced the event occurrence up to the time period *j*.

By transforming a hazard into log odds (logits), a discrete hazard model with a completely general time specification is specified as below:

$$\text{Logit } h(t_{ij}) = [\alpha_1 D_{1ij} + \alpha_2 D_{2ij} + \ldots + \alpha_J D_{Jij}] + [\beta_1 X_{1ij} + \beta_2 X_{2ij} + \ldots + \beta_J X_{Jij}]$$

where intercept parameters $\alpha_1, \alpha_2, \ldots, \alpha_J$ indicate the logit hazard of respective time periods for individuals whose values of all predictors are null. And slope parameters $\beta_1, \beta_2, \ldots, \beta_J$ represent the effects of respective predictors with controlling for other predictors.

As an alternative specification of a hazard model, through an inverse transformation, logit $h(t_{ij})$ is re-represented into the hazard as below:

$$h(t_{ij}) = 1/\ 1 + e^{-\{[\alpha 1D1ij + \alpha 2D2ij + \ldots + \alpha JDJij] + [\beta 1X1ij + \beta 2X2ij + \ldots + \beta JXJij]\}}$$

A discrete hazard model estimates parameters using maximum likelihood estimates. A log-likelihood function is specified as below:

$$LL = \Sigma\Sigma\, EVENT_{ij} \log h(t_{ij}) + (1- EVENT_{ij}) \log (1- h(t_{ij}))$$

where $EVENT_{ij}$ represents whether an event occurs to an individual i in time period j.

To express the effect of the predictors, parameter estimates are anti-logged into an odds ratio. Odds are computed in two groups – one group whose value of that specific predictor is 0 and another group whose value of that specific predictor is 1. To compute odds, the ratio of estimates of the two groups is taken as below:

$$\begin{aligned}\text{Estimated odds ratio} &= e^{\alpha JDJ + \beta 1} / e^{\alpha JDJ} \\ &= e^{\alpha JDJ}\, e^{\beta 1} / e^{\alpha JDJ} \\ &= e^{\beta 1}\end{aligned}$$

By taking out time indicators and their parameters (D_J, α_J), antilogging of a coefficient (β_1) will lead to the comparison of the odds of event occurrence between one group that takes on 0 as a value of β_1 and another group that takes on 1 as a value of β_1 (See Allison, 2010; Hosmer & Lemeshow, 1999; Singer & Willet, 2003).

Continuous-Time Hazard Models

Continuous- time estimation records the occurrence of events in fine units such as minutes, hours, or days. Among continuous survival models, the most popular model is the Cox proportional hazards model (Singer & Willet, 2003). Typically, based on the Cox proportional hazards model, a hazard is estimated as an instantaneous change in the occurrence rate of the event using partial maximum likelihood. Continuous- time survival analysis estimates hazard function, which is transformed into log function through logarithmic transformation. One of the important features of the Cox proportional hazards model is that the value of the baseline hazard function is not estimated, assuming an underlying baseline hazard as a completely general one. The Cox model estimates the hazard as below:

$$h(t_{ij}) = h_0(t_j)\, e^{[\beta 1X1ij + \beta 2X2ij + \ldots + \beta PXPij]}$$

where $h_0(t_j)$ represents the baseline hazard function.

The above equation is re-expressed through the antilog as below:

$$\log h(t_{ij}) = \log h_0(t_j) + [\beta 1X1ij + \beta 2X2ij + \ldots + \beta PXPij]$$

Raw coefficients represent the effect of the respective predictors on log hazards while the anti-logged coefficients ($e^{(\text{coefficient})}$) indicate the effect of respective predictors on raw hazards. The antilog of raw coefficient can also represent hazard ratios; the hazard function ratio that conforms to the effect of the respective predictor showing the difference in the hazard between one group with the value of 0 for the respective predictor and another group with the value of 1 for that predictor (See Allison, 2010; Hosmer & Lemeshow, 1999; Singer & Willet, 2003).

Partial Likelihood Method of Estimation

While discrete survival analysis uses maximum likelihood estimation, continuous survival analysis uses partial likelihood method of estimation. As with full maximum likelihood estimation, partial maximum likelihood estimation builds a likelihood function to estimate the likelihood that the samples (subjects) are observed. Then, the partial estimation continues to evaluate the effectiveness of parameter estimates and iterates until it identifies parameter estimates which maximize the likelihood function (Singer & Willet, 2003).

The difference between maximum and partial likelihood estimation lies in a probability that a maximum likelihood function seeks to estimate. A full maximum likelihood function attempts to estimate the probability that a target event will occur to an individual *i* in time period *j* given that the event does not occur to that individual until then.

On the other hand, a partial likelihood function estimates the probability that it would be an individual *i*, given that an event occurs to someone in time period *j*. Thus, only individuals who experience the events contribute to the partial likelihood function. Contrastingly, in the full maximum likelihood function, every individual makes a contribution to the function. The log likelihood function for a Cox proportional hazards model is specified as follows (See Allison, 2010; Hosmer & Lemeshow, 1999; Singer & Willet, 2003):

Partial log likelihood =

$$\sum_{\substack{noncensored\\individuals}} \left[\quad (\beta_1 x_{1ij} + \cdots + \beta_j x_{jij}) - \log \sum_{\substack{risk\ set\\at\ t_{ij}^*}} (\beta_1 x_{1ij} + \cdots + \beta_j x_{jij}) \quad \right]$$

Where $\beta's$ denotes population parameters

Ties

Ties, or tied events refer to the events that more than one sample (subject) experience the target event at the same time. The extra caution is required when building a Cox proportional hazards model because the Cox model is very sensitive to ties and often times leads to invalid analysis outcomes. Common methods to manage ties in continuous survival analysis are the exact method, the Breslow-Peto approximation, and the Efron approximation (Hertz-Picciotto & Rockhill, 1997). The exact method treats all the possible combinations of observations in ranking tied events. The Breslow-Peto approximation randomly posits a sequential occurrence of tied observations. The Efron approximation assumes all the possible rankings of tied events but adopts simple computations.

Among the three methods, the exact method is the most recommended, followed by the Efron approximation (Prentice & Gloeckler, 1978; Singer & Willet, 2003). However, due to the computational difficulty, the exact method is not always feasible. When the exact method is not feasible, the Efron approximation is preferred to the Breslo-Peto approximation.

However, when there are more tied observations than untied observations in the data, none of the three tie-handling methods would work. Singer and Willet (2003) suggested that discrete-time methods should be used for such cases instead of continuous-time survival analysis. Similarly, Hess and Persson (2010) compared estimates from a Cox proportional model and a discrete model, and showed that a Cox proportional model resulted in biased coefficients and standard errors from the data with heavy ties.

Complementary Log-Log (Clog-Log) Model

For the continuous hazard model, this study, particularly for the second approach, adopted a clog-log model, following the recommendations of prior studies for the case of heavy ties when building a Cox proportional model (Allison, 2010; Hosmer & Lameshow, 1999). The clog-log model was equivalent to the exact method of the Cox proportional model, which is known to generate the most precise parameter estimates in the case of heavy ties. In particular, the clog-log model was suitable for this study, as the clog-log model produced a clog-log hazard for each time period, which can be compared with those from the discrete model.

The clog-log model transforms a hazard into the complementary log-log probability, the logarithm of the negated logarithm of the probability of event nonoccurrence as follows (Singer & Willet, 2003):

The hazard of clog-log = log(-log(1-probability))

The study adopted the clog-log hazard by Allison (2010) as follows:

$$Clog\text{-}log\ h(t_{ij}) = \alpha_1 D_{1ij} + \alpha_2 D_{2ij} + \ldots + \alpha_J D_{Jij}$$

where intercepts $\alpha_1, \alpha_2, \ldots, \alpha_J$ indicate the clog-log hazard for each time period.

The specifications of the clog-log model are very similar to those of the discrete model. The only difference is that the clog-log model estimates a clog-log hazard while the discrete model estimates a logit hazard.

Model Comparison: Goodness-of-Fit

Log-Likelihood Statistic (LL)

A log-likelihood statistic (LL) is used for comparing nested models of the survival analyses. The larger-numbered LL statistic indicates the better fit. The LL statistic is converted into -2LL as a basis to evaluate a model fit. Thus, the smaller the deviance statistic is, the better fit a model demonstrates (Hosmer & Lemeshow, 1999; Singer & Willet, 2003).

To compare nested models, a likelihood ratio test is used based on -2LL from two competing models. The procedure begins with getting the difference between -2LL's of two competing models, and then the difference is to be compared with a magnitude of the critical value of X^2 distribution on *k* degrees of freedom (*k* refers to the difference between the degrees of the freedom of these two competing models). If the difference is greater than the critical value of those degrees of freedom, then there is significant reduction in -2LL between competing models (See Allison, 2010; Singer & Willet, 2003).

AIC and BIC

To compare non-nested models fitted to the same dataset, AIC (Akaike Information Criterion) and Bayesian Information Criterion (BIC) are recommended. The smaller these values, the better fit the model demonstrates. Both AIC and BIC are calculated by being penalized based on the number of parameters. In addition, a BIC calculation involves the number of events (number of participants who experienced the event). The equations of AIC and BIC are as follows (See Allison, 2010; Singer & Willet, 2003):

$$\mathrm{AIC} = -2\mathrm{LL} + 2p$$

$$\mathrm{BIC} = -2\mathrm{LL} + [\{\ln(\mathrm{E})\}p]$$

Where LL denotes log-likelihood statistic, p stands for number of parameters and E presents number of events

Current Status on Adopting Survival Analysis Models

Target Event

This study reviewed 13 studies that utilize survival analysis. As target events, these studies explored various events including attrition issues, dropout, the re-occurrence of a phenomena and others: departures from employment (Donaldson, 2010; Kelly, 2004; Kirby, 1999; Singer, 1992; 1998); college graduation (Calcagno, 2007; Jacobs, 2002); dropout from schools (Murphy, 2010; Plank, 2008) or courses (Ma, 1999); re-report of child maltreatment (Kahn, 2010); and program adoption (Doyle, 2006). These studies calculated the time to the occurrence of these target events from a beginning time: the start time of employment; the enrollment time of school or courses; the first time of reporting child maltreatment; and the time prior to the adoption of the program in the area of the interest.

Time Metrics

Thirteen studies adopted either discrete- or continuous- time estimation. Seven of them employed a discrete-time estimation, using a year, a semester, or a month as the time metric for their studies: three studies using year as a metric of time (Donaldson, 2010; Ma, 1999; Singer, 1992); three studies using semester as a time metric (Calcagno, 2007; Kahn, 2010; Singer, 1998); and one study using month (Jacobs, 2002).

The other six studies used continuous-time estimation. Just like the discrete-time survival analyses, these continuous-time survival analyses adopted year, semester or month as time metrics for the studies: year (Doyle, 2006; Kelly, 2004; Kirby, 1999; Murphy, 2010); semester (Murtaugh, 1999); and month (Plank, 2008). The table below presents the time metrics of these 13 discrete- and continuous-time survival analyses.

Table 1. *Use of Time Estimations and Metrics in Research*

Time Metric	Discrete-Time Survival Analysis	Continuous-Time Survival Analysis
Year	Singer (1992), Donaldson (2010), Ma (1999)	Doyle (2006), Kelly (2004), Kirby (1999), Murphy (2010)
Semester (Trimester, Quarter)	Singer (1998), Kahn (2010), Calcagno (2007)	Murtaugh (1999)
Month	Jacobs (2002)	Plank (2008)

Discrete-Time Estimation: Model Specification and Analysis

This section presents the model specification and the interpretation of results of seven studies adopting discrete-time survival analysis. Out of these studies, two studies did not specify the equation of the model. The study done by Singer (1992) used 13-year longitudinal data on 6,600 beginning teachers in Michigan and North Carolina and derived hazard functions to identify contextual factors that influence the risk of a teacher's departure from teaching. Singer plotted the hazard functions, providing graphical representations of these hazard functions. However, Singer did not present the results using an odds ratio along with associated significance levels.

Singer and her colleagues (1998) explored the retention of 2,654 primary care physicians in Community Health Centers. The authors estimated a hazard model to link the likelihood of a physician's departure to characteristics of the physician themselves and the centers that they were working for. The authors presented the parameter estimates in odds and associated standard errors.

Unlike the above studies which did not specify the equations of hazards, Calcagno, Crosta, Bailey, and Jenkins (2007) used logit specification to build a discrete hazard model. Based on 42,641 students who first matriculated in Florida community colleges in the fall term between 1998 and 1999, researchers modeled the probability of college completion by including various demographic and academic characteristics of college students as important contextual factors.

To fit hazard functions on the probabilities of college completion, the study adopted logit specification, which expressed a linear relationship between independent variables and college completion using maximum likelihood function. The study specified a time as a fully general spline using 17 dummy time variables denoting each time period (trimester) of 17 trimesters. The study presented the parameter estimates in odds with respective standard errors.

Like the time specification by Calcagno, Crosta, Bailey, and Jenkins (2007), Ma and Willms (1999) adopted a fully general spline for time specification, but they utilized a hazard model through an inverse transformation. Using 6-wave data on 3,116 7th graders from the Longitudinal Study of American Youth (LSAY), Ma and Willms explored the dropout chance of taking advanced math courses by these students and the effects of school characteristics on students' dropout.

To represent the time indicators, the authors incorporated five dummy variables, representing each grade, by excluding an intercept. The study built a model without an intercept to facilitate Hierarchical Linear Modeling (HLM) analysis in constructing the school-level effects. Five coefficients of five dummy variables represented the time effects of Grades 8 through 12. The conditional probability of a dropout was estimated using the hazard model. The study presented odd ratios, standard errors, the p-values and the effect sizes.

Beyond the general specification of time indicators, some studies adopted more parsimonious specification for their discrete hazard model specification. These studies presented odds ratio, standard errors and associated p-values when examining the effects of respective predictors. Jacobs and King (2002) estimated the likelihood of college completion among 8,535 women aged from 15 to 44. They explored the association of the likelihood of college completion with women's age and other variables. The study included the measure on enrollment periods expressed in both linear and quadratic forms in order to explore nonlinear relationship between enrollment duration and the attainment of degrees. The results demonstrated a curvilinear relationship, revealing that women with long enrollment durations were less likely to graduate.

Similarly, using nationally representative longitudinal data on 5,501 children, aged from birth to 14 years, Kahn and Schwalbe (2010) investigated the timing of the re-report of child maltreatment and associated contextual factors. The authors specified the main effect of time as a quadratic representation.

Finally, Donaldson and Johnson (2010) analyzed 2,029 Teach for America (TFA) teachers to explore when and why the teachers left teaching. Instead of the general specification of time, the researchers used a more parsimonious time specification, with each time variable taking on a different representation depending on its relationship with the outcome variable.

Continuous-Time Estimation: Model Specification and Analysis

Studies using continuous- time survival analysis adopted a Cox model to estimate the rate of event occurrence. Three of the studies employed a Cox proportional hazards model and presented analysis results using hazard ratios, associated standard errors and p values. Based on data on 98,951 newly-hired teachers in Texas from 1979 to 1996, Kirby and her colleagues (1999) examined the attrition issues of minority teachers in high-risk districts.

Doyle (2006) employed a Cox proportional hazards model to explore the relationship between state characteristics and the chance of adopting a merit-aid program based on longitudinal data from the 48 continental states. By analyzing 2,200 minority students at Georgia Tech (GT), Murphy and his colleagues (2010) examined the relationship between the graduation rate and an intervention program after controlling for other contextual variables.

While adopting a Cox proportional hazards model like the studies above, Murtaugh (1999) estimated a hazard function based on a risk score for each individual. By analyzing data on 8,867 college students at Oregon State University, Murtaugh developed a proportional hazards model to estimate student attrition probability associated with academic and demographic factors. The author adopted the Cox proportional hazards regression model, which estimated a risk score for each individual as below:

$$R = B_1X_1 + B_2X_2 + \ldots + B_kX_k$$

where $X_1, X_2, \ldots, X_k$ denoted k independent variables (risk factors) and $B_1, B_2, \ldots, B_k$ represented regression coefficients using maximum likelihood method.

Different from the above studies utilizing Cox proportional hazards models, Kelly (2004) employed Stratified Cox Proportional Hazards models to handle the non-proportionality issues of certain variables. Kelly used continuous survival analysis to model the attrition rate of teachers by analyzing the 1990-1991 Schools and Staffing Survey and the 1992 Teacher Follow-up Survey. The study used the retrospective data on 7,200 teachers who had left their schools. This study adopted Stratified Cox Proportional Hazards models to represent three phases of teacher career path identified as a result of the hazard function plot over the time: the initial years, the middle years, and the retirement years.

The study designated Year 5 (T1) as a dividing point between the early and middle years, having Year 31 (T3) as the entry point into the retirement years. Using these two time points, the study estimated two baseline hazards for these two strata. The study further examined the proportionality of variable effects across the strata. While most of the variables were found to be proportional, certain variables were not proportional. To treat non-proportionality issues, after fitting un-proportional variables in each strata, these variables were converted into interaction terms by multiplying these variables by two time dummy variables (T1, T3).

Similarly, Plank and his colleagues (2008) adopted a non-proportional hazards model by including time-varying variables. By analyzing 856 high school students from the National Longitudinal Survey of Youth 1997 (NLSY97), the authors modeled the relationship between dropout and Career and Technical Education (CTE). The authors specified the hazard function of non-proportional hazards model as below:

$h(t_{ij}) = \lambda_0(t)\, e^{[\beta 1 Xi + \beta Xi(t)]}$
where $\lambda_0(t)$ denotes a baseline hazard function; $\beta_1 X_i$ represents a set of time-invariant covariates while $X_i(t)$ presents series of time-varying covariates.

Model Comparisons

-2LL. Based on -2LL, many of the studies evaluated the goodness-of-the-fit of the nested models to determine a better fit. Some studies compared nested models with different sets of independent variables. Calcagno, Crosta, Bailey, and Jenkins (2007) tested -2LL of six models which have different sets of independent variables. For example, Model 1 consisted of 17 time dummy variables and the age variable. The study tested the goodness-of-the-fit of this model by comparing its deviance reduction (-2LL) with the null model only with time dummy variables. The study developed subsequent five models by adding more independent variables to the variables of the proceeding models.

Similarly, Doyle (2006) constructed four models: three subsets of the model and the full model. Three subset models included only covariates of interest at one time. The full model included all covariates from the three subset models simultaneously. The author compared the model fit of the four models based on a likelihood ratio test.

On the similar vein, the reduction of -2LL was evaluated in parallel but separate models that fit the data to different subgroups of samples. Kirby and her colleagues (1999) conducted four parallel but separate models: three subset models that included each of three ethnicities (Non-Hispanic White, Hispanic, and Black) at one time; and the full model with all three ethnicities.

Murtaugh (1999) examined the goodness-of-fit of the final model by comparing observed attrition rate with the predicted rate by categorizing students into four subgroups depending on risk score: lowest risk; lower risk; higher risk; and highest risk. In addition, Jacobs and King (2002) conducted a separate but parallel multivariate analysis on both full- and part-time students. The model with full-time students demonstrated a better fit than the model with part-time students as the model with full-time students showed greater reduction in -2LL.

AIC and BIC. While the above studies used -2LL to compare nested models, other studies employed AIC or BIC to assess fit statistics of non-nested models. Calcagno, Crosta, Bailey, and Jenkins (2007) specified a time as a fully general spline for discrete survival models. Using AIC, the authors evaluated the fit statistics of a general specification by comparing alternative time specifications such as a linear, quadratic, or cubic function of time. Similarly, Kahn and Schwalbe (2010) specified the main effect of time in three different behaviors: linear, quadratic and log-time. The authors chose the model with quadratic representation as the model with squared quarter revealed the smallest BIC.

The Mixed Use of Discrete or Continuous Models in Literature

As reviewed in the above section on the metric of time, most of the studies indicated confusion over two estimation methods regardless of time metrics. The six studies among the thirteen adopted a continuous time estimation method by building a Cox regression model for relatively large time metrics for their choice of the continuous model (Doyle, 2006; Kelly, 2004, Kirby, Berends & Naftel, 1999; Murphy, Gaughan, Hume, & Moore, 2010; Murtaugh, Burns & Schuster, 1999; Plank, DeLuca, & Estacion, 2008).

Moreover, the seven studies that used a year as a time metric adopted either discrete (Donaldson & Johnson, 2010; Ma & Willms, 1999; Singer, 1992) or continuous survival analysis (Doyle, 2006; Kelly, 2004, Kirby, Berends, & Naftel, 1999; Murphy, Gaughan, Hume, & Moore, 2010). Four studies that chose a semester as a time metric used discrete (Calcagno, Crosta, Bailey & Jenkins 2007; Kahn & Schwalbe, 2010; Singer, Davidson, Graham & Davidson, 1998) or continuous analysis (Murtaugh, Burns & Schuster, 1999). Two studies used a month as a time metric for discrete (Jacobs & King, 2002) or continuous analysis (Plank, DeLuca, & Estacion, 2008).

While most of these studies did not indicate a proper rationale for using either a continuous or a discrete model, some studies provided the rationale in various ways. Three studies provided the rationale for the choice of discrete- time estimation over continuous- time estimation. Instead of precise career duration measured in days or weeks, Singer (1998) chose to utilize the service length in quarterly intervals based on results of preliminary analysis. Physicians showed a tendency to leave on an annual basis, revealing needs to aggregate the employment durations in larger units.

To find an appropriate aggregated unit, Singer conducted separate analyses using four different time units: 1-, 2-, 3- and 6- month units. Singer chose a 3-month (quarter) interval, which could emphasize the pronounced annual effects during a 21-month measurement period.

Also, according to Singer, a 3-month unit was specific enough to show the decrease in the likelihood of departure immediately followed by the anniversary quarter.

Ma and Willms (1999) chose a discrete-time estimation method, following not only the results from preliminary analysis but also theoretical guidelines. As the time interval measuring math participation was large (a year), the authors adopted discrete- time estimation, not continuous- time estimation, according to the recommendation as cited by Allison (1984). The preliminary analysis supported such treatment of discrete time as the dropout risk differed greatly from one grade to another. Thus, it was impossible to estimate a constant hazard rate according to the authors.

Kahn and Schwalbe (2010) conducted a discrete-time survival analysis using quarters as a time metric, instead of conducting a continuous-time survival analysis using days. The authors provided logical and practical reasons as follows: the database that the authors used usually measures variables in larger units such as quarters, half years or years; prior studies adopted discrete time intervals including 30 day-, month- or six month- period in examining the occurrence or re-occurrence of child maltreatment; analysis by day would not provide meaningful, parsimonious interpretation in the difference between a re-report at day 175 and 180 while analysis by quarters would provide clear, wieldy interpretation; and Stata would prefer a discrete survival analysis using quarters when handling the study's complex database with weights.

Contrary to the above studies that provided the rationale for the choice of discrete-time survival analysis, Doyle (2006) explained his choice of continuous-time survival analysis with the year as a time metric for the following reasons: no expectation on the underlying shape of the hazard rate of the policy adoption is consistent with the Cox model's assumption of no specific shape of the hazard function; and prior studies on political phenomena used the Cox model extensively to estimate the occurrence of political events.

Furthermore, Doyle conducted another continuous survival analysis using the date of the adoption as an alternative time specification to the year-time specification of his study. He found that continuous survival analysis using days produced similar results to those using years as the time-specification.

In addition, some studies attempted to handle tied observations by adopting the most preferred one – the exact method - out of three potential methods: the exact method; the Breslow-Peto approximation; and the Efron approximation. These studies needed to treat ties as these studies used relatively large time units for continuous survival analysis: a year as a time metric (Murphy, Gaughan, Hume, & Moore, 2010); and a month as a time metric (Plank, DeLuca & Estacion, 2008).

Another problem of prior studies is the fact that the researchers adopted either a continuous or a discrete model without considering important factors for survival model-building. For example, Plank, DeLuca, and Estacion (2008) chose a continuous model with an exact method because the authors did not find any difference in the parameter estimates between continuous and discrete models without consideration of hazard rates. Obviously, a guideline that suggests the proper use of either a discrete or a continuous model is needed.

Discrepancies between Discrete (Logit) and Continuous (Clog-Log) Models

Studies have been conducted to compare a logit model with a clog-log model and found no difference between them. Allison (2010) empirically investigated the difference between these two methods by using the same data and found the similar parameter estimates. The p-values of covariate estimates of two methods were similar while parameter estimates of a logit model were slightly larger than those of a clog-log model. Allison argued that parameter estimates from a discrete model tend to be larger than those of a clog-log model.

Similar results were also noted in other studies. Colosimo, Chalita, and Demétrio (2000) compared logit and clog-log models for data from various conditions of 12 time metrics, a sample size of 198, 0.005 to 0.44 hazard rates, and a 25% censoring proportion. Using likelihood ratio tests, the researchers did not detect discrepancies between these two models. Corrente, Chalita, and Moreira (2003) attempted to provide a guideline for a clog-log or a logit model using data for 286 samples with high tied events during 52 intervals. The authors were not able to identify a discrepancy between the two models by using residuals and fit statistics.

In sum, the study findings on the difference between two models are not conclusive. This is partly due to the fact that these studies have not incorporated important factors into their studies. In response to the findings, the second approach of the study compared the two models taking important factors into account. The important factor that the study considered is a hazard. In addition, the second approach also paid attention to the time metric, censoring proportion, and sample size, as these factors influence the hazard.

Factors to Determine the Discrepancies between Discrete and Continuous Models

Time Metrics

The major difference between discrete and continuous survival models is the unit used to measure the timing of an event. A discrete model uses comparatively a large time metric, resulting in a small number of waves (frequency). A continuous model uses a continuous time metric, leading to a large number of waves. A small number of waves will force many events into the same wave even though the events occurred at different time points. Thus, the presence of many tied observations will lead to a high hazard rate, increasing the potential discrepancy between the logit and clog-log models, while a large number of waves will lead to a low chance of a discrepancy by having a low hazard rate.

The study conducted by Hofstede and Wedel (1999) showed the effect of the size of time metrics on the potential discrepancy between the logit and clog-log models. By building both continuous- and discrete-time hazard models using different sizes of time metrics, Hofstede and Wedel found that, when time metrics were aggregated into a large unit, the discrepancy in parameter estimates between continuous and discrete-time hazard models was large. In other words, when a small time metric (originally a day) was aggregated into a week or a month, the hazard estimates in discrete models were overestimated while those in continuous models were underestimated.

Censoring Proportion

Censored data represents sample data for which a target event does not occur during the data collection period (Singer & Willet, 2003). Censoring is an important factor when comparing discrete and continuous survival models because the censoring proportion influences the number of cases that experience the target event and changes the number of tied observations and the hazard rates. Despite the importance of the censoring proportion in distinguishing between discrete and continuous models, there is little or no research that paid attention to the different proportions of censoring as a main interest.

Hertz-Picciotto and Rockhill (1997) discussed the importance of censoring in survival analysis, although the authors did not include the censoring proportion in their model.

According to the authors, censoring reduced the estimate biases because it kept the tie proportion low. A study conducted by Colosimo, Chalita, and Demétrio (2000) generated data with censoring proportions of 0%, 30%, and 60% to examine the effects of censoring on the two methods. However, the researchers explored the effects of censoring proportions in a variety of sample sizes, so they did not solely explore the effect of the censoring proportion. Hess (2009) also paid attention to low (16%-20%) and high censoring proportions (24%-27%). Hess found that increasing the censoring proportions increased the standard deviations of the parameter estimates but did not result in a discrepancy in the parameter estimates.

Sample Size

Given the fact that a large sample size leads to a more robust statistical analysis, there is a lack of research on the effect of sample size relating to the proper use of survival analysis. To make things complicated, there are mixed results regarding the effect of sample size for discrete and continuous hazards models. In Hess' study (2009), the logit and clog-log models displayed a difference when the sample size was equal to or more than 1,000 because a large sample has more tied observations.

In contrast, the study conducted by Hertz-Picciotto and Rockhill (1997) showed that the discrepancy between the two models was larger with a small sample. The researchers compared three methods (the Breslow, Kalbflesch-Prentice, and Efron methods) of treating ties for the Cox proportional model by differentiating between sample sizes. Hertz-Picciotto and Rockhill used four different sample sizes of 50, 100, 500, and 1,000 and found that all three methods treating ties displayed biased estimates when the sample size was 50.

However, the discrepancy was due to the design of the study. The researchers fixed the number of ties across different sample sizes, resulting in the creation of high hazards when the sample was small. Therefore, their study emphasized the tied observations more than the influence of sample size. Unfortunately, there are few studies that have examined the effect of sample size while keeping other factors constant. This study sought to contribute information in regard to the sample size and its effect on survival model outcomes.

Hazards

Hazards has been identified as an important factor that determines the discrepancy between a logit and a clog-log model. According to Hosmer and Lameshow (1999), while the outcome of a clog-log model is similar to that of a logit model when a hazard is smaller than 0.15, the difference between the two models becomes notable when a hazard is greater than 0.15.

Similarly, Singer and Willet (2003) found a difference between a logit and a clog-log model outcomes during periods with a high hazard. However, the authors did not notice differences in parameter estimates from the two models. In a similar vein, Hess (2009) found a great difference in the estimated parameters from clog-log and logit models when a hazard was as high as 0.3. However, no difference was noted in the parameter estimates for the periods with a hazard lower than 0.3.

CHAPTER THREE

METHODOLOGY

This study is motivated by the fact that the current published studies use discrete and continuous survival models without having distinctions and rules despite there exist clear conceptual and application differences between the two survival models. This study compared two survival analyses using two different methodological approaches. The first methodology is empirical investigation which employs the four different time metrics of year, half-year, quarter, and month to a nationally representative data. The study in the first approach built eight models by adopting two survival models for four different time metrics. The use of different time metrics is one of the critical factors in the use of survival analysis because the time metrics determine the hazards of the survival model. The findings of the empirical analysis aimed to compare the parameter estimates and fit statistics of the discrete method and continuous methods using the same data and explore the impact of different time metrics in application of survival analyses.

For the goal, this study built survival models using multiple waves of dropout status of high school students and coding the dichotomous dependent variables (1=dropout; 0= graduation). As predictor variables, the study included student race, student language status, parental educational levels, student perceptions towards teachers, student GPA, and student gender. Prior to the main analysis, this study conducted preliminary analysis including descriptive statistics, correlation analysis and logistic regression in order to obtain essential information on the data. Followed by the preliminary analyses, this study designed a series of discrete- and continuous-time survival analyses. The study adopted a logit model as a discrete-time estimation method, while building Cox proportional hazard model as a continuous- time estimation method. By employing the two survival analyses using four different time metrics, the study built eight survival models. Using the results from the eight models, the study explored the discrepancies between these two models using four different time metrics.

For the second approach, the study explored the important factors that determined the discrepancies between discrete- and continuous-time survival analyses. As important factors to determine discrepancies, the study identified time metrics, censoring proportions, and sample sizes. To analyze the impact of the time metrics, censoring proportions, and samples sized on the hazards, the study simulated data in 60 conditions by combining three factors: three time metrics (4, 12, and 48); five censoring proportions (0%, 20%, 40%, 60%, and 80%); and four sample sizes (50, 100, 500, and 1,000) for one survival model condition. Using the simulated data, a series of survival analyses were built, by adopting a logit model as discrete model and a clog-log model of Cox proportional hazard model as a continuous model. By building both logit and clog-log models with 60 conditions, the study had 120 models. For the comparison of the two survival models, the study compared the parameter estimates and fit statistics.

First Approach

Data Sources

To build a series of survival analyses using empirical data, the study used data from the National Longitudinal Survey of Youth 1997 (NLSY97), a nationally representative database (Center for Human Resource Research, 2006). The NLSY97 contained the information on 8,984 youths who were born between the years of 1980 and 1984. In 1997, the NLSY97 started collecting the data on youths' family characteristics, education, and labor market experiences through interviews with these youths and their parents. Since 1997, the NSLY97 tracked these youths annually until now through survey interviews. In addition to the interviews, the NLSY97 collected high school transcripts on a subsample of youths from respective high schools. The current study used a transcript subsample of youths whose transcript data was available. This study traced 2,216 students for 48 months from when they were enrolled in 8th grade until they reached 12th grade. This study included 8th grade as a starting time for the high school period to get a comprehensive perspective, as some high schools enrolled 8th graders along with 9th graders. Based on the available transcript data, the study analyzed when and whether these 2,216 youths dropped out from schools within the study time frame (48 months).

Variables

The outcome variable in the study was the dropout event among high school students. By exploring the reasons for leaving schools from transcript data (LEFT_REASON: R97928), this study traced the dropout and censored data. A dropout event was referring to the status that students left school for the following reasons: dropout, expulsion, withdrawal, or the discharge from high school due to reaching the age limit. Students were censored if they left school because they graduated from high school or they obtained a General Education Diploma (GED) within the study timeframe. Additionally, students were treated as missing data if they left school for other reasons: transfer, other reasons, home schooling, and unknown reason. Based on this categorization, the event and censor variable were created to be used for discrete and continuous hazard models, respectively. The event variable indicated the dropout status by coding 0 for graduation and 1 for dropout. The censor variable represented the censored status, by reversing the coding of the event variable: 0 for the dropout; and 1 for censoring.

To measure the duration to the dropout, this study tracked down variables on the calendar month and year when a student was enrolled in the school at the first time and when a student left the school: SCH_START_DATE (R97890.-R97890.01) and LEFT_DATE (R97927.-R97927.01). When the variable on the time of leaving school was not available, this study imputed the school end date variable: the time when students were enrolled at the last time. Furthermore, to ensure these variables indicated the dropout status during the high school period, this study examined variables on the grades of students when students were enrolled in the schools and left schools: TERM_GRADE.xx:R97784-R97801. This study included only time variables, identified as time periods during high school. Then, this study subtracted school leaving time from starting time variables to create the duration variable. The duration variable represented the number of months that passed between the enrollment in 8th grade until a student graduated or dropped out within the study timeframe. As this study considered that 48 months is an appropriate timeframe to measure whether respective 8th graders graduated or dropped out, this study did not include the dropout events whose durations represented more than 48 months.

Using a duration variable with a month basis scale, different time variables were created to represent four time metrics of this study. Four time variables were developed, indicating a month, a quarter, a half-year, or a year, respectively: time_month variable (each scale represents one month, ranging from 1 to 48); time_quarter variable (each scale indicates one quarter, ranging from 1 to 16); time_half_year variable (each scale represents half year, ranging from 1 to 8); and time_year variable (each scale represents one year, ranging from 1 to 4). These four time indicators were used for continuous survival analysis.

Using these four time variables in month, quarter, half-year and year metrics, this study computed respective dummy time variables to indicate month, quarter, half year and year, in order to be used for discrete survival analysis models: 48 time dummy variables to indicate each of 48 months (DM1-DM48); 16 time dummy variables to indicate each quarter (DQ1-DQ16); 8 time dummy variables to represent each half-year (DHY1-DHY8); and 4 time dummy variables to represent each year (DY1-DY4).

As important contextual variables, this study used the following variables: student race; student language status; parental educational level; students' perceptions towards teachers; student GPA; and student gender. To identify student races in four categories, this study combined two race variables (KEY!ETHNICITY:R05386 and KEY!RACE:R05387) and identified four racial backgrounds of the students as follows: Caucasian; Hispanic; Black; and Asian. This study created race dummy variables by having Caucasian and Asian students as a reference group; due to the small number of Asian students, Asian students were combined with Caucasian students. To represent two dominant racial groups (Black and Hispanic), two racial dummy variables were created: Black (0= Non-Black; 1=Black); and Hispanic (0=Non-Hispanic; 1= Hispanic).

In addition, this study included the linguistic status of students by focusing on English Language Learners (ELL), those who are not native speakers of English. This study identified ELL students as those who had ever taken English as a Second Language (ESL) courses (TRANS_CRS_CODE L66 1999: r9721800) or participated in bilingual programs (TRANS_BILING_ED: R97887). The ELL dummy variable indicated the linguistic background of students by coding 0 for Non-ELL and 1 for ELL.

To represent parental educational level, this study combined the mother's (P2-03: R05545) and father's educational levels (P2-034: R05548). The responses for this item range from 0 (No Education) to 7 (Grad/professional degree). By choosing a maximum level out of father's and mother's educational levels, this study created a variable of parental educational level (Parental Education Level).

This study also included students' perception towards teachers by averaging three items that examine how much students agreed with the following statements: "the teachers are good" (YSCH-36400: R00694); "Teachers are interested in students (R0069500)"; and "students are graded fairly (r0069700)." The responses were recoded to represent the following four ranges: 1. strongly disagree; 2. disagree; 3. agree; and 4. strongly agree.

This study also examined students' overall GPA (TRANS_CRD_GPA_OVERALL: R98719), which represents the average GPA of all courses that students had taken, weighted by Carnegie credits; the overall GPA range extended from 0 to 500.

In addition, this study included the gender of students (R05363: KEY!SEX), by creating a gender dummy variable with the following coding: 0 = male; and 1 = female students.

Analysis

Preliminary Analysis. To have fundamental understanding of the data, this study conducted several preliminary analyses having descriptive statistics, and correlation analyses. Descriptive statistics showed the student group composition in terms of races, and languages genders. Simultaneously, descriptive statistics examined both dropout and completion rates of these different student groups. Correlation analyses examined the association between dropouts and three contextual variables: Parental education level, student-teacher relationship, and student GPA. Furthermore, this study compared mean differences of parental educational level, student perception of teachers, and the student GPA across different races, linguistic status, and genders.

Logistic Regression. Prior to the main analysis of building survival models, this study conducted logistic regression to explore the effect of variables of the interest on the graduation rates. Logistic regression is a proper regression analysis using dichotomous dependent variable (Allison, 1999). The difference between logistic regression and survival analysis lies in the fact that survival analysis uses multiple waves of dichotomous dependent variables collected at multiple time points while logistic regression uses only one dichotomous dependent variable measured at one time point (Allison, 2010), mostly the final outcome. In other words, the survival analysis deals with several incidents of drop-out incidents while logistic regression analyzes the final outcome of drop-out at the final wave. To run logistic regression, this study used SPSS and built logistic regression model by incorporating dropout status as one time dependent variable and including other predictor variables such as student races (Hispanic and Black), student language status, parental educational levels, students' perceptions towards teachers, student GPA, and student gender.

Survival Analysis. The primary analysis of this study was a survival analysis designed to investigate the occurrence of and the timing of events using longitudinal data (Singer & Willett, 2003). Particularly, this study dealt with two time estimation methods of survival analysis: discrete- and continuous-time estimation. Furthermore, this study paid attention to the use of four different time metrics when conducting these two survival analyses (discrete and continuous): year; half-year; quarter; and month. By adopting two different time estimation methods using four different time metrics, this study built eight models. To construct these eight models, this study used SPSS. The use of SPSS led this study to deal with issues with ties for continuous time estimation by adopting the Breslow method, which is a default option in SPSS.

Discrete Models. Using a discrete time estimation method, this study built four models: the Year model; Half-Year model; Quarter model; and Month model. As an outcome variable, these four discrete models used an event variable. These models contained the same predictors of interest but had different time indicator variables. As predictors of interest, these models included the following variables: student races (Hispanic and Black); student language status; parental educational level; students' perceptions towards teachers; student GPA; and student gender.

Time variables that these discrete models adopted were as follows: the Year model included four time indictors, each of which represented a year; the Half-Year model included eight time dummy variables, each of which indicated a half-year; the Quarter model includes 16 time dummy variables, each of which denoted a quarter; and the Month model included 48 time dummy variables, each of which represented a month. These time variables were so included into the models because of the use of a completely general time specification for this study. This study did not transform these time indicators into linear or quadratic functional forms to have parsimonious modeling. Moreover, these time specifications went beyond the scope of this study. The equations of these four discrete models were as below:

Year model

$$h(t_{ij}) = 1/\ 1 + e^{-\{[\alpha_1 DY_{1ij} + \alpha_2 DY_{2ij} + \alpha_3 DY_{3ij} + \alpha_4 DY_{4ij}] + [\beta_1(\text{Black})_{1ij} + \beta_2(\text{Hispanic})_{2ij} + \beta_3(\text{ELL})_{3ij} + \beta_4(\text{Parental\_Edu})_{4ij} + \beta_5(\text{Teachers})_{5ij} + \beta_6(\text{GPA})_{6ij} + \beta_7(\text{Gender})_{7ij}]\}}$$

where *i* denotes an individual student in year *j*; intercept parameters $\alpha_1, \alpha_2, \ldots, \alpha_4$ indicate the "baseline hazard" that a student, who takes on null values for all predictors, will dropout for the first time in a given year, assuming that he or she had been enrolled at the high school until that year; slope parameter β_1 indicates the effect of being a Black student on a student's risk of dropout; slope parameter β_2 indicates the effect of being a Hispanic student on a student's risk of dropout; slope parameter β_3 indicates the effect of being an ELL student on a student's risk of dropout; slope parameter β_4 indicates the effect of parents' educational level on a student's risk of dropout; slope parameter β_5 indicates the effect of student's positive perception towards teachers on a student's risk of dropout; slope parameter β_6 indicates the effect of student GPA on a student's risk of dropout; and slope parameter β_7 indicates the effect of being a female student on a student's risk of dropout.

Half- Year model

$$h(t_{ij}) = 1/\ 1 + e^{-\{[\alpha_1 DHY_{1ij} + \alpha_2 DHY_{2ij} + \ldots + \alpha_7 DHY_{7ij} + \alpha_8 DHY_{8ij}] + [\beta_1(Black)_{1ij} + \beta_2(Hispanic)_{2ij} + \beta_3(ELL)_{3ij} + \beta_4(Parental\_Edu)_{4ij} + \beta_5(Teachers)_{5ij} + \beta_6(GPA)_{6ij} + \beta_7(Gender)_{7ij}]\}}$$

where *i* denotes an individual student in year *j*; intercept parameters $\alpha_1, \alpha_2, \ldots, \alpha_8$ indicate the "baseline hazard" that a student, who takes on null values for all predictors, will dropout for the first time in each six month, assuming that he or she had been enrolled at the high school until that time; slope parameter β_1 indicates the effect of being a Black student on a student's risk of dropout; slope parameter β_2 indicates the effect of being a Hispanic student on a student's risk of dropout; slope parameter β_3 indicates the effect of being an ELL student on a student's risk of dropout; slope parameter β_4 indicates the effect of parents' educational level on a student's risk of dropout; slope parameter β_5 indicates the effect of student's positive perception towards teachers on a student's risk of dropout; slope parameter β_6 indicates the effect of student GPA on a student's risk of dropout; and slope parameter β_7 indicates the effect of being a female student on a student's risk of dropout.

Quarter model

$$h(t_{ij}) = 1/\ 1 + e^{-\{[\alpha_1 DQ_{1ij} + \alpha_2 DQ_{2ij} + \ldots + \alpha_{15} DQ_{15ij} + \alpha_{16} DQ_{16ij}] + [\beta_1(Black)_{1ij} + \beta_2(Hispanic)_{2ij} + \beta_3(ELL)_{3ij} + \beta_4(Parental\_Edu)_{4ij} + \beta_5(Teachers)_{5ij} + \beta_6(GPA)_{6ij} + \beta_7(Gender)_{7ij}]\}}$$

where *i* denotes an individual student in year *j*; intercept parameters $\alpha_1, \alpha_2, \ldots, \alpha_{16}$ indicate the "baseline hazard" that a student, who takes on null values for all predictors, will dropout for the first time in a given quarter, assuming that he or she had been enrolled at the high school until that quarter; slope parameter β_1 indicates the effect of being a Black student on a student's risk of dropout; slope parameter β_2 indicates the effect of being a Hispanic student on a student's risk of dropout; slope parameter β_3 indicates the effect of being an ELL student on a student's risk of dropout; slope parameter β_4 indicates the effect of parents' educational level on a student's risk of dropout; slope parameter β_5 indicates the effect of student's positive perception towards teachers on a student's risk of dropout; slope parameter β_6 indicates the effect of student GPA on a student's risk of dropout; and slope parameter β_7 indicates the effect of being a female student on a student's risk of dropout.

Month model

$h(t_{ij}) = 1/\ 1 + e^{-\{[\alpha_1 DM_{1ij} + \alpha_2 MQ_{2ij} + \ldots + \alpha_{47} DM_{47ij} + \alpha_{48} DM_{48ij}]\ +\ [\beta_1(\text{Black})_{1ij} + \beta_2(\text{Hispanic})_{2ij} + \beta_3(\text{ELL})_{3ij} + \beta_4(\text{Parental\_Edu})_{4ij} + \beta_5(\text{Teachers})_{5ij} + \beta_6(\text{GPA})_{6ij} + \beta_7(\text{Gender})_{7ij}]\}}$

where *i* denotes an individual student in year *j*; intercept parameters $\alpha_1, \alpha_{2,\ldots,}\ \alpha_{16}$ indicate the "baseline hazard" that a student, who takes on null values for all predictors, will dropout for the first time in a given month, assuming that he or she had been enrolled at the high school until that month; slope parameter β_1 indicates the effect of being a Black student on a student's risk of dropout; slope parameter β_2 indicates the effect of being a Hispanic student on a student's risk of dropout; slope parameter β_3 indicates the effect of being an ELL student on a student's risk of dropout; slope parameter β_4 indicates the effect of parents' educational level on a student's risk of dropout; slope parameter β_5 indicates the effect of student's positive perception towards teachers on a student's risk of dropout; slope parameter β_6 indicates the effect of student GPA on a student's risk of dropout; and slope parameter β_7 indicates the effect of being a female student on a student's risk of dropout.

Continuous Models. By adopting continuous-time survival analysis, this study has built four models: the Year model; Half-Year model; Quarter model; and Month model. As the outcome, these four models used a censoring variable. As main predictors, these models incorporated the same six variables as the discrete models above. These models included different time variables: a year for the Year model (which has four scales, each of which represented one year); a half-year for the Half-Year model (which has eight scales representing each half-year); a quarter for the Quarter model; and a month for the Month model. However, because of the time specification feature of the Cox Proportional Regression model, the use of a different time metric for each model did not lead to a different equation for these four models. The Cox Proportional model did not estimate the value of the baseline hazard function, but rather specified it as an underlying hazard function. Thus, the equations of the four continuous models were the similar to discrete models except using different time variables for underlying hazard functions as below:

Year model

$$h(t_{ij}) = h_0(t_j)\, e^{[\beta_1(\text{Black})_{1ij} + \beta_2(\text{Hispanic})_{2ij} + \beta_3(\text{ELL})_{3ij} + \beta_4(\text{Parental\_Edu})_{4ij} + \beta_5(\text{Teachers})_{5ij} + \beta_6(\text{GPA})_{6ij} + \beta_7(\text{Gender})_{7ij}]}$$

where i denotes an individual student in time j; $h_0(t_j)$ represents the baseline hazard function using a year time variable; slope parameter β_1 indicates the effect of being a Black student on a student's risk of dropout; slope parameter β_2 indicates the effect of being a Hispanic student on a student's risk of dropout; slope parameter β_3 indicates the effect of being an ELL student on a student's risk of dropout; slope parameter β_4 indicates the effect of parents' educational level on a student's risk of dropout; slope parameter β_5 indicates the effect of student's positive perception towards teachers on a student's risk of dropout; slope parameter β_6 indicates the effect of student GPA on a student's risk of dropout; and slope parameter β_7 indicates the effect of being a female student on a student's risk of dropout.

Half-Year model

$$h(t_{ij}) = h_0(t_j)\, e^{[\beta_1(\text{Black})_{1ij} + \beta_2(\text{Hispanic})_{2ij} + \beta_3(\text{ELL})_{3ij} + \beta_4(\text{Parental\_Edu})_{4ij} + \beta_5(\text{Teachers})_{5ij} + \beta_6(\text{GPA})_{6ij} + \beta_7(\text{Gender})_{7ij}]}$$

where i denotes an individual student in time j; $h_0(t_j)$ represents the baseline hazard function using a half-year time variable; slope parameter β_1 indicates the effect of being a Black student on a student's risk of dropout; slope parameter β_2 indicates the effect of being a Hispanic student on a student's risk of dropout; slope parameter β_3 indicates the effect of being an ELL student on a student's risk of dropout; slope parameter β_4 indicates the effect of parents' educational level on a student's risk of dropout; slope parameter β_5 indicates the effect of student's positive perception towards teachers on a student's risk of dropout; slope parameter β_6 indicates the effect of student GPA on a student's risk of dropout; and slope parameter β_7 indicates the effect of being a female student on a student's risk of dropout.

Quarter model

$$h(t_{ij}) = h_0(t_j)\, e^{[\beta_1(\text{Black})_{1ij} + \beta_2(\text{Hispanic})_{2ij} + \beta_3(\text{ELL})_{3ij} + \beta_4(\text{Parental\_Edu})_{4ij} + \beta_5(\text{Teachers})_{5ij} + \beta_6(\text{GPA})_{6ij} + \beta_7(\text{Gender})_{7ij}]}$$

where i denotes an individual student in time j; $h_0(t_j)$ represents the baseline hazard function using a quarter time variable; slope parameter β_1 indicates the effect of being a Black student on a student's risk of dropout; slope parameter β_2 indicates the effect of being a Hispanic student on a student's risk of dropout; slope parameter β_3 indicates the effect of being an ELL student on a student's risk of dropout; slope parameter β_4 indicates the effect of parents' educational level on a student's risk of dropout; slope parameter β_5 indicates the effect of student's positive perception towards teachers on a student's risk of dropout; slope parameter β_6 indicates the effect of student GPA on a student's risk of dropout; and slope parameter β_7 indicates the effect of being a female student on a student's risk of dropout.

Month model

$$h(t_{ij}) = h_0(t_j)\, e^{[\beta_1(\text{Black})_{1ij} + \beta_2(\text{Hispanic})_{2ij} + \beta_3(\text{ELL})_{3ij} + \beta_4(\text{Parental\_Edu})_{4ij} + \beta_5(\text{Teachers})_{5ij} + \beta_6(\text{GPA})_{6ij} + \beta_7(\text{Gender})_{7ij}]}$$

where i denotes an individual student in time j; $h_0(t_j)$ represents the baseline hazard function using a month time variable; slope parameter β_1 indicates the effect of being a Black student on a student's risk of dropout; slope parameter β_2 indicates the effect of being a Hispanic student on a student's risk of dropout; slope parameter β_3 indicates the effect of being an ELL student on a student's risk of dropout; slope parameter β_4 indicates the effect of parents' educational level on a student's risk of dropout; slope parameter β_5 indicates the effect of student's positive perception towards teachers on a student's risk of dropout; slope parameter β_6 indicates the effect of student GPA on a student's risk of dropout; and slope parameter β_7 indicates the effect of being a female student on a student's risk of dropout.

The details of this study's eight models are found in the table below, featuring the population number, number of events, number of parameters and time indicator variables of each model.

Table 2. *Overview of Eight Models of This Study*

	Year	Half-Year	Quarter	Month
Discrete-Time				
N	6341	12801	24564	72481
n events	209	204	204	203
n parameters	11	15	23	55
time indicators	DY1-DY4	DHY1-DHY8	DQ1-DQ16	DM1-DM48
Continuous-Time				
N	1860	1842	1842	1840
n events	209	204	204	203
n parameters	7	7	7	7
time indicators	Time_Year	Time_Half-Year	Time_Quater	Time_Month

Maximum and Partial Maximum Likelihood Estimation

As an estimation method, this study adopted the maximum likelihood method for the discrete model (Singer & Willet, 2003). A log-likelihood function is specified as below:

$$LL = \Sigma\Sigma\, EVENT_{ij} \log h(t_{ij}) + (1 - EVENT_{ij}) \log (1 - h(t_{ij}))$$

where $EVENT_{ij}$ represents whether an event occurs to an individual i in time period j.

As an estimation method for the continuous model (Cox proportional model), this study chose the partial maximum likelihood estimation method with the following function (Singer & Willet, 2003):

Partial log likelihood =

$$\sum_{\substack{noncensored\\ individuals}} \left[\quad (\beta_1 x_{1ij} + \cdots + \beta_j x_{jij}) - \log \sum_{\substack{risk\ set\\ at\ t_{ij}^{*}}} (\beta_1 x_{1ij} + \cdots + \beta_j x_{jij}) \quad \right]$$

Model Comparison

After building series of discrete and continuous models, the study examined the parameter estimates with significance levels along with model fit statistics. For fit statistics, the study used the goodness-of-fit statistics using -2LL. At the same time, the study adopted the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) (see Allison, 2010; Singer & Willet, 2003). The equations of AIC and BIC are as follows:

$$AIC = -2LL + 2p$$

BIC = -2LL + [{ln(Event numbers)}p]

where p indicates the number of parameters.

Second Approach

Data Generation

The data were generated using SAS with a variety of time metrics, censoring proportions, and sample sizes, and these three factors determine hazards. Based on the total sample size and censoring proportions, the number of cases that experienced the target event was determined for each data set. By combining three time metrics, five censoring proportions, and four sample sizes, the study generated 60 combinations of data.

Time Metrics. The study specified three types of time metrics of 4, 12, and 48 which emulated educational conditions. For example, for the case of 4 years of high school, one can choose the base unit as a year, half-year, quarter, or month. When a year is chosen, 4 is the time metric; when a quarter or half-year is chosen, 12 is the time metric; and when a month is chosen, 48 is the time metric.

Censoring Proportion. The study included five censoring proportions to make a possible range of 0%, 20%, 40%, 60%, and 80%. The censoring proportion indicates cases that did not experience the target event. The study assumed the right-hand censoring, meaning that the censoring occurred at the end of the data collection period.

Sample Size. This study selected the following sample sizes: 50, 100, 500, and 1,000, as prior researchers such as Hertz-Picciotto and Rockhill (1997) used.

The raw hazard was calculated by dividing the number of the risk set (cases that did not experience the target event until that time) with the number of events in each period. First, the total number of events (cases that experienced the target event) was obtained by multiplying the total sample size by an uncensored proportion (100%-censoring proportion). The number of events (tied observations) per period was calculated by dividing the total number of events by the number of periods. The hazard is estimated as a proportion of the number of tied observation (events) per each period out of the number of the risk set (the total number of cases in which the target event did not occur to up to that period). This study assumed right-hand censoring and no missing data.

The Model: Logit Model vs. Clog-Log Model

As this study used three time metrics, the three equations for the logit and clog-log models were specified as follows:

Using 4 time metrics:

$$Logit\ h(t_{ij}) = \alpha_1 D_{1ij} + \alpha_2 D_{2ij} + \alpha_3 D_{3ij} + \alpha_4 D_{4ij}$$

$$Clog\text{-}Log\ h(t_{ij}) = \alpha_1 D_{1ij} + \alpha_2 D_{2ij} + \alpha_3 D_{3ij} + \alpha_4 D_{4ij}$$

where *i* denotes an individual student in a specific time period *j*; intercept parameters $\alpha_1, \alpha_2, \ldots, \alpha_4$ indicate the "baseline hazard" that a sample data will experience the event, assuming that the sample data did not experience the event until that time

Using 12 time metrics:

$$Logit\ h(t_{ij}) = \alpha_1 D_{1ij} + \alpha_2 D_{2ij} + \alpha_3 D_{3ij} + \ldots + \alpha_{10} D_{10ij} + \alpha_{11} D_{11ij} + \alpha_{12} D_{12ij}$$

$Clog\text{-}Log\ h(t_{ij}) = \alpha_1 D_{1ij} + \alpha_2 D_{2ij} + \alpha_3 D_{3ij} + \ldots + \alpha_{10} D_{10ij} + \alpha_{11} D_{11ij} + \alpha_{12} D_{12ij}$

where *i* denotes an individual student in a specific time period *j*; intercept parameters $\alpha_1, \alpha_2, \ldots, \alpha_{12}$ indicate the "baseline hazard" that a sample data will experience the event, assuming that the sample data did not experience the event until that time

Using 48 time metrics:

$Logit\ h(t_{ij}) = \alpha_1 D_{1ij} + \alpha_2 D_{2ij} + \alpha_3 D_{3ij} + \ldots + \alpha_{46} D_{46ij} + \alpha_{47} D_{47ij} + \alpha_{48} D_{48ij}$

$Clog\text{-}Log\ h(t_{ij}) = \alpha_1 D_{1ij} + \alpha_2 D_{2ij} + \alpha_3 D_{3ij} + \ldots + \alpha_{46} D_{46ij} + \alpha_{47} D_{47ij} + \alpha_{48} D_{48ij}$

where *i* denotes an individual student in a specific time period *j*; intercept parameters $\alpha_1, \alpha_2, \ldots, \alpha_{48}$ indicate the "baseline hazard" that a sample data will experience the event, assuming that the sample data did not experience the event until that time

Maximum Likelihood Estimation

As an estimation method, this study adopted the maximum likelihood method, which estimates population parameters by maximizing the probability that the sample data will be observed. The likelihood function stands for the likelihood of observing the pattern of event occurrence or non-occurrence in a dataset. In the case of the discrete (logit) model, the likelihood function was specified as follows:

$$\text{Likelihood} = \prod_{i=1}^{n} \prod_{j=1}^{Ji} h(t_{ij})^{EVENTij} (1 - h(t_{ij}))^{(1-EVENTij)}$$

where $h(t_{ij})$ refers to the probability that the event will occur to an individual *i* in the *j* period. $EVENTij$ indicates whether the event happens to an individual *i* in the *j* period; 1 indicates an event occurrence, while 0 indicates no event occurrence.

The likelihood function shown above was simplified into the following log-likelihood

(LL) function.

$$LL = \sum_{i=1}^{n} \sum_{j=1}^{Ji} Event_{ij}\, log\, h\big(t_{ij}\big) \, + \, \big(1 - Event_{ij}\big) log\, (1 - h\big(t_{ij}\big))$$

(Singer & Willet, 2003)

The maximum likelihood estimation function for the clog-log model is as follows:

$$\mathrm{LL} = \sum_{\mathrm{i}=1}^{\mathrm{n}} \sum_{\mathrm{j}=1}^{\mathrm{Ji}} \mathrm{Event_{ij}}\, \mathrm{log}\, \big(1 - \exp(-\exp \mathrm{h}(\mathrm{t_{ij}}))\big) \, + \, (1 - \mathrm{Event_{ij}}) \mathrm{log}\, \big(-\exp(-\exp \mathrm{h}(\mathrm{t_{ij}}))\big)$$

(Franklin, 2005)

Model Comparison

After building two models, the study compared the parameter estimates with significance levels (the hazard of the respective period) and model fit statistics. The hazard estimates from the two models were also compared with significance levels of 0.05 and 0.01.

For fit statistics, the study used the goodness-of-fit statistics for these two models for 60 conditions. The goodness-of-fit statistics used -2LL, which was converted from a log-likelihood statistic (LL). In particular, the study adopted the Akaike Information Criterion (AIC) because the models under comparison are non-nested models (see Allison, 2010; Singer & Willet, 2003). The smaller the values are, the better fit the model demonstrates. The AIC is calculated by being penalized based on the number of parameters as follows:

$$AIC = -2LL + 2p$$

where p indicates the number of parameters.

CHAPTER FOUR

RESULTS

Despite the potential of survival analysis to deal with research areas of student dropout and teacher attrition and retention, there has been confusion in the use of a proper model out of two popular models (discrete and continuous model). In response, this study has compared analysis results from discrete and continuous survival models by taking two methodological approaches: Empirical inquiry using empirical data and theoretical investigation using simulated data.

Using empirical data sets, the first approach built a series of discrete and continuous models by using four different time metrics (year, half-year, quarter, and month), resulting in eight models. First, the study results provided empirical evidence regarding a proper model based on the discrepancies in the parameter estimates and fit statistics across these eight models. Clearly, the study results encouraged researchers to build the discrete model rather than the continuous model when using comparatively large time metrics, particularly, year. However, considering the current practice of collecting education data using large time units, a month is the most suitable time metric for researchers to build either a continuous or discrete model interchangeably among year, half-year, quarter and month.

In order to provide rich findings regarding the discrepancies between discrete and continuous survival models, the second approach paid attention to the effect of censoring proportion and sample size in addition to time metrics. Moreover, to promote deeper understanding of the roles of these three factors, this study linked these three factors to hazards which are known to cause the discrepancies between discrete and continuous models. By combining different levels of the factors, this study simulated 60 datasets and built discrete and continuous models to the datasets.

In terms of time metrics, the study results showed pronounced discrepancies in large units of time metrics (i.e., year of the first approach). On the other hand, in using fine units (i.e., 48), either discrete or continuous model can be built. Regarding censoring proportion, greater discrepancies was demonstrated in a small censoring proportion, less than 20%. For the sample size, smaller samples demonstrated greater differences between two models. With regards hazard rates, two models revealed discrepancies with higher hazard rates (i.e., greater than 0.21).

First Approach

Preliminary Analyses

To obtain essential understanding of the data, this study conducted descriptive statistics and correlation analyses shown in Table 3. The students with the information on dropout status totaled 2,216. The students comprised of 57.6% of Caucasian or Asian students (*N*=1,248), 20.5% of Black students (*N*=444), and 21.9% of Hispanic students (*N*=475) as shown in the Table 3. The racial groups showed different dropout rates. Only 10.3% of the Caucasian/Asian group were dropouts comparing with 19.4% of the Black group dropouts and 17.1% of the Hispanic group dropouts. The language groups revealed similar dropout rates, with 13.7% of the Non-ELL group and 12.8% of the ELL group. The gender groups demonstrated different dropout rates, having 17.0% male students 10.3% female students.

Table 3. Dropout/Completion Rates of Racial, Language and Gender Groups

Group	Dropout	Completion	Subtotal
Race			
Caucasian/Asian	129 (10.3%)	1,119 (89.7%)	1,248 (57.6%)
Black	86 (19.4%)	358 (80.6%)	444 (20.5%)
Hispanic	81 (17.1%)	394 (82.9%)	475 (21.9%)
Subtotal	296 (13.7%)	1,871 (86.3%)	2,167
Language			
Non-ELL	278 (13.7%)	1,751 (86.3%)	2,029 (91.6%)
ELL	24 (12.8%)	163 (87.2%)	187 (8.4%)
Subtotal	302 (13.6%)	1,914 (86.4%)	2,216
Gender			
Male	189 (17.0%)	925 (83.0%)	1,114 (50.3%)
Female	113 (10.3%)	989 (89.7%)	1,102 (49.7%)
Subtotal	302 (13.6%)	1,914 (86.4%)	2,216

Table 4 presents the results of correlation analyses to show the relationship of dropout rates with three contextual variables: Parental education level, student's perception of teachers, and student GPA. The parental education level had a significant correlation with $r = -0.084$ ($p < 0.01$), having higher parental educational levels being associated with lower dropout rates of students. The parent education level was significantly higher in the Caucasian or Asian group ($r = 0.351$, $p < 0.01$), and significantly lower in the Black group ($r = -0.131$, $p < 0.01$) and Hispanic group ($r = -0.311$, $p < 0.01$). The parental education in the ELL group was significantly lower than the Non-ELL group, with $r = -0.109$ ($p < 0.01$). The parental education level in the female group was significantly lower ($r = -0.064$, $p < 0.01$) compared with the male group, demonstrating no significant difference.

Table 4. Bivariate Correlation

	Dropout	Parent Education	Teachers	GPA
Dropout	__	-.084**	-.125**	-.365**
Caucasian/Asian	-.109**	.351**	.085**	.182**
Black	.084**	-.131**	-.105**	-.107**
Hispanic	.052*	-.311**	.001	-.114**
ELL	-.007	-.109**	.017	-.033
Parent Education	-.084**	__	.028	.157**
Teachers	-.125**	.028	__	.203**
GPA	-.365**	.157**	.203**	__
Gender	-.098**	-.064**	.062**	.195**

* denotes a significance level at .05

** denotes a significance level at .01

The student perception of teacher had a significant correlation with r= -0.125 ($p < 0.01$), indicating that when students had a positive relationship with their teachers, they were less prone to drop out. The student perception of teacher in the Caucasian/Asian group was significantly higher ($r = 0.085$, $p < 0.01$) and in the Black group was significantly lower ($r = -0.105$, $p < 0.01$) when compared to that of the other groups. The student perception of teachers in the Hispanic group was not significantly different from the other groups ($r = 0.001$, $p > 0.05$). There was no significant difference in the student perception of teachers between the ELL and the Non-ELL group ($r = 0.017$, $p > 0.05$). The student perception of teachers in the female group was significantly higher ($r=0.062$, $p < 0.01$) compared with the male group.

The student GPA variable showed a significant correlation ($r = -0.365$, $p < 0.01$), showing that students who had a higher GPA tended to less likely drop out. The student GPA was significantly higher in the Caucasian/Asian group ($r = 0.182$, $p < 0.01$). However, the student

GPA was significantly lower in the Black group ($r = -0.107$, $p < 0.01$) and Hispanic group ($r = -0.114$, $p < 0.01$). The student GPA in the ELL group was slightly lower than the Non-ELL group, having no significant difference ($r = -0.033$, $p > 0.05$). The student GPA was significantly higher in the female group compared with the male group ($r = 0.195$, $p < 0.01$).

Also, this study compared parental educational levels, student perception towards teachers, and the student GPA across various groups as shown in Table 5. Compared with other racial groups, Caucasian/Asian students showed a notable difference in the student GPA and the parental education levels. Compared with the Caucasian/Asian group (Mean= 3.48), parents of the other groups had lower educational levels (Black, Mean= 2.55; and Hispanic, Mean= 1.99). Also, compared with the Caucasian/Asian students (Mean=300.15), other students showed a lower GPA (Black, Mean= 278.04; and Hispanic, Mean= 278.26). The notable difference between the Non-ELL and ELL students was detected only in GPA (Non-ELL, Mean= 291.56; ELL, Mean=284.47). Similar patterns were detected between the female and male groups. Compared with female students (Mean= 302.36), male students showed a lower GPA (Mean= 279.63).

Table 5. Means (Standard Deviation)

	Parent Education	Teachers	GPA
Race			
Caucasian/Asian	3.48 (1.49)	3.10 (.50)	300.15 (58.54)
Black	2.55 (1.38)	2.95 (.52)	278.04 (55.63)
Hispanic	1.99 (1.70)	3.06 (.46)	278.26 (55.56)
Language			
Non-ELL	2.89 (1.61)	3.06 (.50)	291.56 (57.94)
ELL	2.46 (1.74)	3.09 (.49)	284.47 (63.64)

Gender			
Male	3.11 (1.69)	3.09 (.50)	279.63 (58.47)
Female	2.89 (1.60)	3.03 (.50)	302.36 (56.17)

Logistic Regression Results

As shown in Table 6, the model chi-square value is 259.779, which is the difference between the constant-only model and the full model. The null hypothesis is rejected because the significance is less than .01. Thus, the set of independent variables improves prediction of the outcome.

Table 6. Omnibus Tests of Model Coefficients

		Chi-square	Df	Sig.
Step 1	Step	259.779	7	.000
	Block	259.779	7	.000
	Model	259.779	7	.000

The Cox & Snell R-Square is .132 and the Negelkerke pseudo R^2 is .262 (See Table 7). The Cox & Snell R-Square is a generalized coefficient of determination to estimate the proportion of variance in the dependent variable, which is explained by the predictor variables. Thus, about 26 percent of variance was explained by the model.

Table 7. Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square

1	1022.543	.132	.262

As shown in Table 8, the goodness-of-fit statistic is 7.070, distributed as a chi-square value, and is associated with a p value of .520, indicating an acceptable match between predicted and observed probabilities.

Table 8. Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	7.070	8	.529

The output in Table 9 divides the data into approximately 10 equal groups. The first group (column) represents those participants least likely to dropout. The second group (column) represents those participants most likely to dropout.

Table 9. Contingency Table for Hosmer and Lemeshow Test

		Dropout = No		Dropout = Yes		Total
		Observed	Expected	Observed	Expected	
Step 1	1	181	182.366	3	1.634	184
	2	183	180.884	1	3.116	184
	3	180	178.982	4	5.018	184
	4	175	177.051	9	6.949	184
	5	175	174.221	9	9.779	184
	6	169	170.573	15	13.427	184
	7	161	165.530	23	18.470	184
	8	163	158.325	21	25.675	184

9	149	144.074	35	39.926	184
10	102	105.996	84	80.004	186

According to Table 10, the overall predictive accuracy is 89.8%. The model predicted better for the graduation but, not the dropout (99.1% and 14.7%, respectively).

Table 10. Classification Table

			Predicted		
			Dropout		Percentage Correct
	Observed		No	Yes	
Step 1	Dropout	No	1624	14	99.1
		Yes	174	30	14.7
	Overall Percentage				89.8

According to Table 11, among the predictors of interest, statistically significant predictors included student GPA and gender. One unit increase in student GPA is related to an approximately 2% lower chance of a dropout ($e^{(coefficient)}$= .980, p<.01). A female student is expected to be 0.676 times as likely to drop out as compared to a male student in a given time ($e^{(coefficient)}$=.676, p<05). Other predictors were not significant but, showed the following tendencies. Compared with Caucasian/Asian students, Black students were 1.390 times as likely to drop out. Similarly, Hispanic students were 1.390 times as likely to dropout. Students with higher parental education are less likely drop out. Compared with non-ELL students, ELL students were less likely to drop out. Students with more positive perception toward teachers tended to drop out less.

Table 11. Variables in the Equation

		B	S.E.	Wald	Df	Sig.	Exp(B)
Step 1	Black	.304	.212	2.052	1	.152	1.356
	Hispanic	.329	.215	2.336	1	.126	1.390
	ELL	-.508	.328	2.401	1	.121	.602
	Parental Edu	-.056	.056	.982	1	.322	.946
	Teachers	-.273	.159	2.943	1	.086	.761
	GPA	-.020	.002	158.429	1	.000	.980
	Gender	-.391	.172	5.179	1	.023	.676
	Constant	4.242	.623	46.343	1	.000	69.577

Survival Analysis Results

This section presents the results of building eight models: four discrete-time survival analysis models and four continuous-time survival analysis models. This section begins with the results from one discrete model and one continuous model to provide an overview of the survival analysis results of the first approach. Then, this section presents the comparison results between discrete and continuous time estimations, encompassing all eight models. This overall comparison is followed by the model comparisons based on different time units: year, half-year, quarter and month. This study provides these comparison results of different time units, first for discrete models, and then for continuous models.

The analysis results of all eight models are presented in below the Table 12. When presenting results, this study did not include estimated parameters of time metrics. In case of continuous time estimation, the survival analysis modeling did not estimate the effect of time, because of its assumption on the baseline hazard model. On the contrary, in case of discrete time estimation, the survival analysis estimated the effects of the time metrics. However, this study did not include these parameters of time metrics from discrete time estimation for the following reasons: the primary focus of the first approach was to compare models by exploring potential discrepancies in parameters of identical predictors across models as a result of the changes in time metrics. Thus, this study was not interested in parameters of time metrics that differed from model to model.

Table 12. *Survival Analysis Results of Eight Models*

	Discrete Models				Continuous Models			
	Year	Half Year	Quarter	Month	Year	Half Year	Quarter	Month
Black	1.426 (.187)	1.370 (.187)	1.415 (.185)	1.438* (.184)	1.396 (.179)	1.368 (.182)	1.422 (.183)	1.452* (.183)
Hispanic	1.381 (.190)	1.368 (.187)	1.388 (.185)	1.402 (.185)	1.327 (.179)	1.327 (.181)	1.346 (.182)	1.358 (.182)
ELL	.636 (.298)	.671 (.291)	.674 (.287)	.680 (.284)	.607 (.281)	.629 (.282)	.619 (.282)	.616 (.283)
Parental Edu	.969 (.050)	.981 (.050)	.983 (.049)	.989 (.049)	.953 (.049)	.959 (.049)	.959 (.050)	.963 (.050)
Teachers	.725* (.050)	.753* (.139)	.738* (.138)	.731* (.138)	.769* (.131)	.791 (.138)	.772 (.138)	.762* (.138)
GPA	.984** (.001)	.985** (.001)	.985** (.001)	.985** (.001)	.985** (.001)	.985** (.001)	.985** (.001)	.984** (.001)
Gender	.732* (.001)	.718* (.154)	.716* (.153)	.718* (.152)	.733* (.148)	.717* (.151)	.711* (.151)	.712* (.151)
-2LL	1587.781	1844.718	2096.371	2435.635	3074.020	2990.177	2966.394	2920.282
N parameters	11	15	23	55	7	7	7	7
AIC	1609.781	1874.718	2142.371	2545.635	3088.02	3004.177	2980.394	2934.282
BIC	1646.547	1924.489	2218.688	2727.861	3111.416	3027.404	3003.621	2957.474

* denotes a significance level at .05
** denotes a significance level at .01

Discrete Year Model and Continuous Month Model

This subsection starts with the discrete year and continuous month models. Considering the four time units of this study, year and month were the most appropriate time metrics for discrete- and continuous- time survival analysis, respectively. Year was the largest discrete time unit, while month was the smallest time unit of this study.

The discrete Year model included ELL, Black, Hispanic, parental educational level, students' perception towards teachers, students' GPA and gender as variables of interest, along with four time dummy variables. Out of the predictors of interest, statistically significant predictors included students' positive perceptions towards teachers, GPA and gender. One unit increase in positive perception towards teachers was associated with a 27% reduction in the hazard of a dropout, having all other variables constant ($e^{(coefficient)}$= .725, $p<.05$). One unit increase in student GPA is related to an approximately 1% lower chance of a dropout ($e^{(coefficient)}$=.984, $p<.01$). A female student is expected to be 0.732 times as likely to drop out compared to a male student in a given year ($e^{(coefficient)}$=.723, p<05). This model had deviance statistics as follows: -2LL =1587.781; AIC = 1609.781; and BIC =1646.547.

The continuous month model incorporated the same variables of interest as the discrete year model featured a different time variable. As significant predictors, the analysis results identified the Black variable, together with the same three variables as the discrete year model: students' positive perceptions of teachers; GPA; and gender. The likelihood of a dropout is nearly one and half times as high for Black students in comparison with Caucasian and Asian students when controlling for other effects ($e^{(coefficient)}$= 1.452, $p<.05$). However, students' perception of teachers, GPA and being a female student were found to reduce the hazard of a dropout: teachers ($e^{(coefficient)}$= .762, $p<.05$); GPA ($e^{(coefficient)}$= .984, $p<.01$); and Gender

($e^{(coefficient)}$=.712, $p<.05$). Compared with the discrete model, the continuous model demonstrated a worse fit: a -2LL of - 2920.282, while having an AIC of 2934.282 and a BIC of 2957.474.

Comparison of Discrete and Continuous Models

This subsection presents a summary of magnitudes and statistical significances of individual coefficients across eight models to compare discrete models with continuous models. Overall, magnitudes and significances of predictors were similar between discrete and continuous models. However, some difference was found in terms of the significance of students' perception towards teachers. The coefficient of students' perception towards teachers was significant in all discrete models: year model ($e^{(coefficient)}$= .725, $p<.05$), half year model ($e^{(coefficient)}$= .753, $p<.05$), quarter model ($e^{(coefficient)}$= .738, $p<.05$), and month model ($e^{(coefficient)}$= .731, $p<.05$). But, it was not always significant in continuous models. The coefficient of students' perception towards teachers was significant only in the year model ($e^{(coefficient)}$= .769, $p<.05$) and month model ($e^{(coefficient)}$= .762, $p<.05$).

According to the fit-statistics, discrete models demonstrated better fit than continuous models. In terms of -2LL, the largest value among discrete models was smaller than the smallest among continuous models: the discrete month model (2435.635); and the continuous month model (2920.282). The worst fit was found in a continuous year model (3074.020), while the best fit was found in a discrete year model (1587.781). Similarly, for the AIC, the largest value among discrete models was smaller than the smallest one among continuous models: The discrete month model (2545.635) and the continuous month model (2934.282). The worst fit was found in a continuous month model (3088.02), while the best fit was found in a discrete year model (1609.781). BIC values demonstrated the same patterns, revealing better fit in discrete models than continuous models. The smallest BIC value among continuous models is greater than the largest BIC value among discrete models: Continuous month model (2957.474) and the discrete month model (2727.861). The largest BIC of the Continuous Models is almost two times larger the than smallest BIC of the discrete models: the continuous year model (3111.416) and the discrete year model (1646.547). The best fit was found in a discrete year model (1646.547). The worst fit was found in a continuous month model (3111.416).

The Comparison of Discrete Models

This subsection presents the model comparison results of discrete models that adopted different time units. This study constructed four discrete models by including respective time indicators to represent four different time metrics: year, half-year, quarter, and month. These four models had identical predictors of interest: ELL, Black, Hispanic, parental educational level, students' positive perception towards teachers, student's GPA, and gender. On the other hand, these models had different sets of time indicators: The year model (four time indicators), the half-year model (eight time indicators), the quarter model (16 time indicators), and the month model (48 time indicators).

By comparing parameter estimates of the discrete models, this study found that the statistical significance of the Black variable varied from model to model. The coefficient of Black was significant in the month model ($e^{(coefficient)}$= 1.438, $p<.05$), but it was not significant in the other three models: The year Model ($e^{(coefficient)}$= 1.426, $p>.05$), the half year model ($e^{(coefficient)}$= 1.370, $p>.05$), and the quarter model ($e^{(coefficient)}$= 1.415, $p>.05$). Other than this difference, overall, these four models demonstrated similar magnitudes of individual coefficients at the same significant levels. For example, across all four models, three variables (teacher perceptions, GPA, and gender) were found to be significant predictors with similar odd ratios with the same statistical significance. That is, parameter estimates of the year and month models demonstrated the following results: teacher perception with the year model ($e^{(coefficient)}$= .725, $p<.05$), the half year model ($e^{(coefficient)}$= .753, $p<.05$), the quarter model ($e^{(coefficient)}$= .738, $p<.05$) and the month model ($e^{(coefficient)}$= .731, $p<.05$); GPA with the year model ($e^{(coefficient)}$= .984, $p<.01$), the half year model ($e^{(coefficient)}$= .985, $p<.01$), the quarter model ($e^{(coefficient)}$= .985, $p<.01$) and the month model ($e^{(coefficient)}$= .985, $p<.01$); and the gender with the year model ($e^{(coefficient)}$=.732, $p<.05$), the half year model ($e^{(coefficient)}$= .718, $p<.05$), the quarter model ($e^{(coefficient)}$= .716, $p<.05$) and the month model ($e^{(coefficient)}$= .718, $p<.05$).

The examination of fit statistics of discrete models resulted in the identification of significant discrepancies; a discrete model with a larger discrete time unit demonstrated a better fit than with a more time unit. The year model demonstrated the best fit, while the month model revealed the worst fit: the year model (-2LL: 1587.781 with 11 df), and the month model (-2LL: 2435.635 with 55 df). The half year and quarter models demonstrated the intermediate fit level between these two models: the half year model (-2LL: 1844.718 with 15 df); and the quarter model (-2LL: 2096.371 with 23 df).

These four discrete models were nested models, as these models included identical predictors other than different sets of time indicators. The year model was nested within the half year model, which included more time indicators than the year model while sharing all the same other variables. In the same way, the quarter model was nested within the month model. Because these models were nested models, this study conducted a likelihood ratio test, detecting significant Chi-square improvement among these four models. The discrete models with fewer time indicators demonstrated a better fit than the models with more time indicators, as seen in the greater reduction in -2LL. The Chi-square improvements between models were as follows: the year and half year models (-2LL difference value of 256.937 on 4 df, 0.05 critical value of 9.5 for a X^2 distribution on 4 df), the half year and quarter models (-2LL difference value of 251.653 on 8 df, 0.05 critical value of 20.1 for a X^2 distribution on 8 df), and the quarter and month models (-2LL difference value of 339.264 on 4 df, 0.05 critical value of 46.2 for a X^2 distribution on 32 df).

The values of AIC and BIC had patterns similar to -2LL. These values were greatest in the month model, followed by the quarter and half year models: the month model (AIC=2545.635, BIC =2727.861), the quarter model (AIC=2142.371, BIC =2218.688), and half year model (AIC=1874.718, BIC =1924.489). The year model had the smallest values: AIC (1609.781); and BIC (1646.547). In sum, discrete models with larger time units showed the better fit.

The Comparison of Continuous Models

In addition to four discrete models, this study built four continuous models by fitting the same model to the same data four times, once for each of four time indicators: a year model with four scales; a half-Year model with eight scales; a quarter model with 16 scales; and a month model with 48 scales. Other than the difference in the time metric, these four models incorporated identical predictors: ELL, Black, Hispanic, parental educational level, students' perception towards teachers, GPA and gender.

First, this study paid attention to the magnitude of individual coefficients and associated statistical significance levels across continuous models. This study found different statistical significance of coefficients on not only Black, but also on students' perceptions towards teachers across models. The hazard ratio of Black was statistically significant only in the month model ($e^{(coefficient)}$= 1.452, $p<.05$), while it was not significant in the other models: the year model ($e^{(coefficient)}$= 1.396, $p>.05$); the half year model ($e^{(coefficient)}$= 1.368, $p>.05$); and the quarter model ($e^{(coefficient)}$= 1.422, $p>.05$). The hazard ratios of student perceptions towards teachers were significant in the month and year models, but those were not significant in the half-year and quarter models: the month model ($e^{(coefficient)}$=.762, $p<.05$), the year model ($e^{(coefficient)}$=.769, $p<.05$), the half year model ($e^{(coefficient)}$= .791, $p>.05$), and the quarter model ($e^{(coefficient)}$= .772, $p>.05$). Other than these two predictors, the rest of the predictors demonstrated similar magnitudes and statistical significance across the four continuous models.

Next, this study examined the fit statistics of these four models. The smaller time metric the model employed, the better fit the model demonstrated. The month model demonstrated better fit than other models, while the year model demonstrated worse fit than other models. The values of -2LL, AIC and BIC of the month and year models are as follows: the month model (-2LL: 2920.282, AIC: 2934.282, BIC: 2957.474) and the year model (-2LL: 3074.020, AIC: 3088.02, BIC: 3111.416). The values of -2LL, AIC and BIC of the quarter and half year models are as follows: the quarter model (-2LL: 2966.394, AIC: 2980.394, BIC: 3003.621) and the half year model (-2LL: 2990.177, AIC: 3004.177, BIC: 3027.404).

Second Approach

This section presents the results of building discrete- and continuous-time survival analysis models: a series of logit and clog-log models under 60 conditions of time metrics, censoring proportions, and sample sizes. First, the comparison results between logit and clog-log models focusing on three time metrics of 4, 12 and 48 time metrics are presented. Second, the comparison results in five censoring proportions of 0%, 20%, 40%, 60%, and 80%. Third, the model comparisons having sample sizes of 50, 100, 500, and 1,000. Last, the raw hazard rates of the parameter estimates having discrepancies between the clog-log and logit models are discussed.

Time Metrics

With regard to time metrics, the small frequency of time metrics (i.e., 4 time metrics) was associated with more discrepancies as shown in Table 13. The use of four time metrics revealed the most discrepancies. Among all 20 conditions (five censoring conditions* four sample sizes) of the four time metrics, eight conditions displayed discrepancies between two models in terms of the significance levels of the estimates between the two models: all four sample sizes of 0% and 20% censoring conditions except 0% censoring with 100 sample size; and 40% censoring with 50 sample size. The use of 12 time metrics revealed discrepancies in seven conditions: all four sample sizes of 0% and 20% censoring conditions other than 20% censoring with 1,000 sample size. The use of 48 time metrics showed discrepancies in only two conditions: sample sizes of 50 and 100 with 0% censoring status.

Table 13. Hazard Estimates across Different Time Metrics

	Logit		Clog-Log			Logit		Clog-Log		Discrepancy Frequency	AIC Difference
Parameter	Estimate	SE	Estimate	SE	Parameter	Estimate	SE	Estimate	SE		
4 Time Metrics										8	517
α_1	-2.06	0.06	-1.97	0.23	α_3	-1.68	0.23	-1.53	0.23		
α_2	-1.88	0.12	-1.77	0.23	α_4	-1.18	5.63	-0.66	12.51		
12 Time Metrics										7	1,935
α_1	-3.96	14.82	-3.68	12.28	α_7	-2.94	0.13	-2.74	0.41		
α_2	-3.16	0.00	-3.03	0.39	α_8	-2.86	0.15	-2.64	0.40		

α_3	-3.14	0.04	-3.00	0.40	α_9	-2.71	0.18	-2.48	0.39		
α_4	-3.10	0.05	-2.95	0.40	α_{10}	-2.68	0.25	-2.40	0.40		
α_5	-3.06	0.08	-2.89	0.41	α_{11}	-2.46	0.31	-2.11	0.34		
α_6	-2.96	0.09	-2.79	0.40	α_{12}	-2.08	9.08	-1.26	9.76		
48 Time Metrics										2	8,567
α_1	-4.46	0.17	-4.35	0.65	α_{25}	-3.98	0.32	-3.83	0.57		
α_2	-4.49	0.19	-4.37	0.66	α_{26}	-4.13	0.39	-3.93	0.62		
α_3	-4.42	0.17	-4.32	0.65	α_{27}	-4.05	0.37	-3.86	0.60		
α_4	-4.34	0.15	-4.25	0.63	α_{28}	-4.08	0.40	-3.87	0.62		
α_5	-4.39	0.17	-4.28	0.64	α_{29}	-3.91	0.34	-3.73	0.57		
α_6	-4.42	0.19	-4.30	0.66	α_{30}	-4.01	0.40	-3.80	0.62		
α_7	-4.29	0.15	-4.19	0.62	α_{31}	-4.05	0.34	-3.81	0.58		
α_8	-4.38	0.20	-4.26	0.66	α_{32}	-4.05	0.36	-3.79	0.59		
α_9	-4.34	0.18	-4.22	0.65	α_{33}	-3.88	0.30	-3.65	0.54		
α_{10}	-4.27	0.17	-4.16	0.63	α_{34}	-3.97	0.35	-3.71	0.58		
α_{11}	-4.32	0.19	-4.19	0.65	α_{35}	-3.94	0.36	-3.66	0.58		
α_{12}	-4.32	0.21	-4.19	0.66	α_{36}	-3.89	0.36	-3.62	0.58		
α_{13}	-4.18	0.16	-4.07	0.61	α_{37}	-3.78	0.33	-3.52	0.55		
α_{14}	-4.31	0.22	-4.17	0.66	α_{38}	-3.86	0.38	-3.56	0.58		
α_{15}	-4.23	0.20	-4.10	0.64	α_{39}	-3.82	0.38	-3.51	0.58		
α_{16}	-4.25	0.22	-4.12	0.66	α_{40}	-3.75	0.38	-3.43	0.57		
α_{17}	-4.13	0.18	-4.01	0.62	α_{41}	-3.70	0.33	-3.35	0.53		
α_{18}	-4.23	0.23	-4.09	0.66	α_{42}	-3.70	0.36	-3.32	0.55		
α_{19}	-4.18	0.22	-4.04	0.65	α_{43}	-3.68	0.37	-3.28	0.55		
α_{20}	-4.18	0.23	-4.03	0.66	α_{44}	-3.61	0.38	-3.19	0.55		
α_{21}	-4.07	0.32	-3.92	0.57	α_{45}	-3.47	0.36	-3.05	0.52		
α_{22}	-4.21	0.38	-4.03	0.62	α_{46}	-3.49	0.41	-3.01	0.54		
α_{23}	-4.15	0.36	-3.97	0.61	α_{47}	-3.43	0.44	-2.91	0.55		
α_{24}	-4.13	0.38	-3.95	0.62	α_{48}	-2.94	14.98	-1.77	9.23		

Censoring Proportions

The examination of censoring proportions revealed that fewer censoring proportions were associated with greater discrepancies between the two models as demonstrated in Table 14. When the censoring was 0%, the number of estimates that showed differences between the two models was the highest. Out of 12 conditions (three time periods * four sample sizes) of 0% censoring, the parameter estimates of nine conditions revealed different significance levels. The three conditions that showed matched results from the two models were the data with large samples (500 and 1,000) and large time metrics (48).

The non-matching results continued to show with a 20% censoring proportion, in which seven out of 12 conditions produced non-matching estimate results: four time periods with four sample sizes; and 12 time periods with three sample sizes (50, 100, and 500). With a 40% censoring proportion, however, a smaller number of discrepancies was detected; only one condition displayed a discrepancy (40% censoring status, 50 sample size, and four time periods). For the 60% and 80% censoring proportions, no discrepancies were detected in any of 12 cases.

Table 14. Hazard Estimates across Different Censoring Proportions

Parameter	**Logit** Estimate	SE	**Clog-Log** Estimate	SE	Parameter	**Logit** Estimate	SE	**Clog-Log** Estimate	SE	**Discrepancy Frequency**	**AIC Difference**
0% Censoring Proportion										9	5,581
α_{1}	-2.29	0.36	-2.35	0.35	α_{25}	-2.58	0.57	-2.59	0.56		
α_{2}	-2.13	0.36	-2.21	0.35	α_{26}	-2.54	0.57	-2.56	0.56		
α_{3}	-1.88	0.37	-2.01	0.35	α_{27}	-2.51	0.57	-2.53	0.56		
α_{4}	2.93	100.82	-0.97	20.82	α_{28}	-2.47	0.57	-2.49	0.56		
α_{5}	-2.58	0.45	-2.61	0.44	α_{29}	-2.44	0.57	-2.45	0.56		
α_{6}	-2.50	0.45	-2.54	0.44	α_{30}	-2.40	0.58	-2.42	0.56		
α_{7}	-2.41	0.46	-2.46	0.44	α_{31}	-2.36	0.58	-2.38	0.56		
α_{8}	-2.31	0.46	-2.36	0.44	α_{32}	-2.31	0.58	-2.33	0.56		
α_{9}	-2.18	0.46	-2.25	0.44	α_{33}	-2.27	0.58	-2.29	0.56		
α_{10}	-2.01	0.47	-2.10	0.44	α_{34}	-2.22	0.58	-2.24	0.56		
α_{11}	-1.74	0.49	-1.88	0.44	α_{35}	-2.17	0.58	-2.19	0.56		
α_{12}	5.14	267.06	-0.40	21.64	α_{36}	-2.11	0.58	-2.14	0.56		
α_{13}	-2.89	0.57	-2.90	0.56	α_{37}	-2.05	0.58	-2.08	0.56		
α_{14}	-2.87	0.57	-2.88	0.56	α_{38}	-1.99	0.58	-2.02	0.56		
α_{15}	-2.84	0.57	-2.86	0.56	α_{39}	-1.92	0.59	-1.95	0.56		
α_{16}	-2.82	0.57	-2.83	0.56	α_{40}	-1.84	0.59	-1.88	0.56		
α_{17}	-2.80	0.57	-2.81	0.56	α_{41}	-1.76	0.59	-1.80	0.56		
α_{18}	-2.77	0.57	-2.78	0.56	α_{42}	-1.66	0.60	-1.71	0.56		
α_{19}	-2.75	0.57	-2.76	0.56	α_{43}	-1.56	0.60	-1.61	0.56		
α_{20}	-2.72	0.57	-2.73	0.56	α_{44}	-1.43	0.61	-1.50	0.56		
α_{21}	-2.69	0.57	-2.71	0.56	α_{45}	-1.29	0.62	-1.36	0.56		
α_{22}	-2.67	0.57	-2.68	0.56	α_{46}	-1.11	0.63	-1.20	0.56		
α_{23}	-2.64	0.57	-2.65	0.56	α_{47}	-0.88	0.65	-1.00	0.56		

α_{24}	-2.61	0.57	-2.62	0.56	α_{48}	12.16	648.95	2.15	33.16		
20% Censoring Proportion										7	6,564
α_{1}	-2.44	0.38	-2.52	0.36	α_{25}	-2.71	0.59	-2.77	0.58		
α_{2}	-2.31	0.38	-2.48	0.38	α_{26}	-2.69	0.59	-2.88	0.65		
α_{3}	-2.13	0.39	-2.28	0.35	α_{27}	-2.65	0.58	-2.71	0.57		
α_{4}	-0.85	20.60	-2.15	0.38	α_{28}	-2.63	0.59	-2.83	0.65		
α_{5}	-2.75	0.47	-2.80	0.44	α_{29}	-2.60	0.59	-2.67	0.58		
α_{6}	-2.68	0.47	-2.84	0.48	α_{30}	-2.55	0.59	-2.77	0.65		
α_{7}	-2.62	0.47	-2.70	0.44	α_{31}	-2.53	0.59	-2.61	0.58		
α_{8}	-2.55	0.48	-2.73	0.48	α_{32}	-2.49	0.59	-2.72	0.65		
α_{9}	-2.46	0.48	-2.56	0.44	α_{33}	-2.44	0.59	-2.54	0.57		
α_{10}	-2.36	0.48	-2.56	0.48	α_{34}	-2.42	0.59	-2.52	0.58		
α_{11}	-2.23	0.49	-2.39	0.44	α_{35}	-2.38	0.59	-2.48	0.58		
α_{12}	-0.35	102.52	-2.31	0.47	α_{36}	-2.32	0.59	-2.43	0.57		
α_{13}	-2.99	0.58	-3.01	0.58	α_{37}	-2.28	0.60	-2.40	0.58		
α_{14}	-2.97	0.58	-3.13	0.65	α_{38}	-2.23	0.60	-2.36	0.58		
α_{15}	-2.94	0.58	-2.97	0.57	α_{39}	-2.17	0.60	-2.30	0.57		
α_{16}	-2.93	0.58	-3.10	0.65	α_{40}	-2.12	0.60	-2.27	0.58		
α_{17}	-2.91	0.58	-2.94	0.58	α_{41}	-2.06	0.60	-2.22	0.58		
α_{18}	-2.87	0.58	-3.05	0.65	α_{42}	-1.98	0.60	-2.15	0.57		
α_{19}	-2.86	0.58	-2.90	0.58	α_{43}	-1.92	0.61	-2.10	0.58		
α_{20}	-2.84	0.58	-3.02	0.65	α_{44}	-1.83	0.61	-2.03	0.58		
α_{21}	-2.81	0.58	-2.85	0.57	α_{45}	-1.72	0.62	-1.95	0.57		
α_{22}	-2.79	0.59	-2.98	0.65	α_{46}	-1.61	0.63	-1.88	0.58		
α_{23}	-2.77	0.59	-2.82	0.58	α_{47}	-1.47	0.65	-1.78	0.58		
α_{24}	-2.73	0.58	-2.92	0.65	α_{48}	1.92	204.68	-1.65	0.57		
40% Censoring Proportion										1	7,718
α_{1}	-2.75	0.42	-2.79	0.41	α_{25}	-2.97	0.61	-2.98	0.60		
α_{2}	-2.71	0.43	-2.75	0.42	α_{26}	-3.07	0.68	-3.08	0.67		
α_{3}	-2.57	0.42	-2.62	0.41	α_{27}	-3.07	0.68	-3.08	0.67		
α_{4}	-2.42	0.41	-2.48	0.39	α_{28}	-3.00	0.67	-3.01	0.66		
α_{5}	-3.08	0.52	-3.10	0.51	α_{29}	-2.89	0.61	-2.90	0.60		
α_{6}	-3.08	0.53	-3.10	0.52	α_{30}	-2.99	0.68	-3.00	0.67		
α_{7}	-2.92	0.48	-2.94	0.47	α_{31}	-2.99	0.57	-3.00	0.57		
α_{8}	-2.98	0.53	-3.00	0.52	α_{32}	-2.92	0.56	-2.93	0.55		
α_{9}	-2.91	0.52	-2.94	0.51	α_{33}	-2.79	0.47	-2.80	0.47		
α_{10}	-2.82	0.50	-2.85	0.49	α_{34}	-2.93	0.57	-2.93	0.56		
α_{11}	-2.81	0.52	-2.83	0.51	α_{35}	-2.93	0.57	-2.94	0.57		
α_{12}	-2.77	0.53	-2.80	0.52	α_{36}	-2.86	0.56	-2.86	0.55		
α_{13}	-3.17	0.61	-3.17	0.60	α_{37}	-2.72	0.48	-2.73	0.47		
α_{14}	-3.27	0.68	-3.28	0.67	α_{38}	-2.86	0.57	-2.87	0.56		
α_{15}	-3.28	0.68	-3.28	0.67	α_{39}	-2.86	0.57	-2.87	0.57		
α_{16}	-3.21	0.67	-3.22	0.66	α_{40}	-2.78	0.56	-2.79	0.55		
α_{17}	-3.11	0.61	-3.11	0.60	α_{41}	-2.65	0.48	-2.66	0.47		

α_{18}	-3.21	0.68	-3.22	0.67	α_{42}	-2.78	0.57	-2.79	0.56		
α_{19}	-3.21	0.68	-3.22	0.67	α_{43}	-2.78	0.57	-2.79	0.57		
α_{20}	-3.15	0.67	-3.16	0.66	α_{44}	-2.71	0.56	-2.72	0.55		
α_{21}	-3.04	0.61	-3.05	0.60	α_{45}	-2.57	0.48	-2.58	0.47		
α_{22}	-3.15	0.68	-3.15	0.67	α_{46}	-2.70	0.57	-2.71	0.56		
α_{23}	-3.15	0.68	-3.15	0.67	α_{47}	-2.70	0.57	-2.71	0.57		
α_{24}	-3.08	0.67	-3.09	0.66	α_{48}	-2.62	0.56	-2.63	0.55		
60% Censoring Proportion										0	8,857
α_{1}	-3.10	0.47	-3.13	0.46	α_{25}	-3.12	0.51	-3.13	0.50		
α_{2}	-3.05	0.47	-3.07	0.46	α_{26}	-3.34	0.62	-3.35	0.62		
α_{3}	-2.96	0.47	-2.99	0.46	α_{27}	-3.30	0.62	-3.31	0.61		
α_{4}	-2.87	0.45	-2.91	0.44	α_{28}	-3.32	0.62	-3.33	0.62		
α_{5}	-3.39	0.57	-3.40	0.56	α_{29}	-3.14	0.53	-3.15	0.52		
α_{6}	-3.33	0.56	-3.35	0.56	α_{30}	-3.22	0.60	-3.22	0.59		
α_{7}	-3.29	0.54	-3.30	0.53	α_{31}	-3.29	0.62	-3.30	0.62		
α_{8}	-3.33	0.58	-3.34	0.57	α_{32}	-3.28	0.62	-3.29	0.62		
α_{9}	-3.25	0.57	-3.27	0.56	α_{33}	-3.07	0.52	-3.07	0.51		
α_{10}	-3.17	0.53	-3.18	0.53	α_{34}	-3.26	0.62	-3.26	0.62		
α_{11}	-3.32	0.56	-3.33	0.55	α_{35}	-3.19	0.60	-3.20	0.60		
α_{12}	-3.23	0.54	-3.25	0.54	α_{36}	-3.21	0.62	-3.21	0.61		
α_{13}	-3.35	0.65	-3.36	0.64	α_{37}	-3.05	0.53	-3.06	0.52		
α_{14}	-3.48	0.72	-3.49	0.71	α_{38}	-3.21	0.62	-3.22	0.62		
α_{15}	-3.40	0.70	-3.41	0.70	α_{39}	-3.17	0.62	-3.18	0.61		
α_{16}	-3.46	0.72	-3.46	0.71	α_{40}	-3.13	0.61	-3.14	0.60		
α_{17}	-3.30	0.65	-3.31	0.64	α_{41}	-2.94	0.43	-2.94	0.43		
α_{18}	-3.40	0.71	-3.41	0.71	α_{42}	-2.89	0.42	-2.90	0.42		
α_{19}	-3.42	0.72	-3.42	0.71	α_{43}	-2.92	0.43	-2.93	0.43		
α_{20}	-3.36	0.71	-3.36	0.70	α_{44}	-2.91	0.43	-2.92	0.43		
α_{21}	-3.19	0.52	-3.19	0.51	α_{45}	-2.87	0.42	-2.87	0.42		
α_{22}	-3.38	0.62	-3.39	0.62	α_{46}	-2.90	0.43	-2.90	0.43		
α_{23}	-3.37	0.62	-3.38	0.62	α_{47}	-2.89	0.43	-2.89	0.43		
α_{24}	-3.34	0.61	-3.34	0.61	α_{48}	-2.84	0.42	-2.84	0.42		
80% Censoring Proportion										0	10,712
α_{1}	-4.64	26.21	-4.62	22.16	α_{25}	-3.61	0.60	-3.61	0.60		
α_{2}	-3.52	0.51	-3.53	0.51	α_{26}	-3.61	0.60	-3.61	0.60		
α_{3}	-3.64	0.58	-3.65	0.57	α_{27}	-3.60	0.60	-3.60	0.60		
α_{4}	-3.50	0.54	-3.52	0.53	α_{28}	-3.60	0.60	-3.60	0.60		
α_{5}	-4.00	0.69	-4.00	0.69	α_{29}	-3.60	0.60	-3.60	0.60		
α_{6}	-3.88	0.65	-3.88	0.64	α_{30}	-3.59	0.60	-3.60	0.60		
α_{7}	-3.98	0.69	-3.98	0.69	α_{31}	-3.59	0.60	-3.59	0.60		
α_{8}	-3.84	0.64	-3.85	0.64	α_{32}	-3.51	0.58	-3.51	0.58		
α_{9}	-3.84	0.65	-3.85	0.64	α_{33}	-3.58	0.60	-3.59	0.60		
α_{10}	-3.84	0.65	-3.85	0.64	α_{34}	-3.58	0.60	-3.58	0.60		
α_{11}	-3.85	0.59	-3.86	0.59	α_{35}	-3.58	0.60	-3.58	0.60		
α_{12}	-3.82	0.59	-3.83	0.58	α_{36}	-3.57	0.60	-3.58	0.60		

α_{13}	-3.79	0.72	-3.80	0.72	α_{37}	-3.57	0.60	-3.57	0.60		
α_{14}	-3.84	0.74	-3.85	0.74	α_{38}	-3.57	0.61	-3.57	0.60		
α_{15}	-3.84	0.74	-3.84	0.74	α_{39}	-3.57	0.61	-3.57	0.60		
α_{16}	-3.78	0.72	-3.78	0.72	α_{40}	-3.49	0.58	-3.49	0.58		
α_{17}	-3.83	0.74	-3.83	0.74	α_{41}	-3.56	0.61	-3.56	0.60		
α_{18}	-3.82	0.74	-3.83	0.74	α_{42}	-3.56	0.61	-3.56	0.60		
α_{19}	-3.82	0.74	-3.82	0.74	α_{43}	-3.55	0.61	-3.55	0.60		
α_{20}	-3.81	0.74	-3.82	0.74	α_{44}	-3.55	0.61	-3.55	0.60		
α_{21}	-3.62	0.60	-3.62	0.60	α_{45}	-3.41	0.54	-3.41	0.54		
α_{22}	-3.62	0.60	-3.62	0.60	α_{46}	-3.41	0.54	-3.41	0.54		
α_{23}	-3.62	0.60	-3.62	0.60	α_{47}	-3.40	0.54	-3.40	0.54		
α_{24}	-3.54	0.58	-3.54	0.58	α_{48}	-3.32	0.51	-3.32	0.51		

Sample Sizes

Among four different sample sizes (50, 100, 500, and 1,000), the smaller samples showed more discrepancies as revealed in Table 15. Out of fifteen conditions (three time periods* five censoring status) of each sample size, six conditions of 50 sample sizes revealed discrepancies. In the samples of 100 and 500 cases, four conditions out of fifteen conditions showed discrepancies. In the 1000-case samples, only three conditions showed discrepancies.

Table 15. Hazard Estimates Across Different Sample Sizes

	Logit		**Clog-Log**			**Logit**		**Clog-Log**		**Discrepancy**	**AIC**
Parameter	Estimate	SE	Estimate	SE	Parameter	Estimate	SE	Estimate	SE	**Frequency**	**Difference**
50 Sample Size										6	330
α_{1}	-3.68	20.84	-3.69	17.64	α_{25}	-3.22	1.02	-3.24	1.00		
α_{2}	-2.73	0.68	-2.78	0.66	α_{26}	-3.18	1.02	-3.20	1.00		
α_{3}	-2.64	0.71	-2.71	0.68	α_{27}	-3.14	1.02	-3.16	1.00		
α_{4}	-1.51	23.86	-2.38	8.24	α_{28}	-3.09	1.02	-3.11	1.00		
α_{5}	-3.22	0.83	-3.25	0.82	α_{29}	-3.04	1.02	-3.07	1.00		
α_{6}	-3.22	0.85	-3.24	0.83	α_{30}	-3.00	1.02	-3.02	1.00		
α_{7}	-3.12	0.84	-3.14	0.82	α_{31}	-2.94	1.03	-2.97	1.00		
α_{8}	-3.10	0.85	-3.13	0.83	α_{32}	-2.89	1.03	-2.92	1.00		
α_{9}	-2.98	0.84	-3.02	0.82	α_{33}	-2.83	1.03	-2.86	1.00		
α_{10}	-2.94	0.86	-2.99	0.83	α_{34}	-2.77	1.03	-2.80	1.00		
α_{11}	-2.68	0.85	-2.75	0.81	α_{35}	-2.71	1.03	-2.74	1.00		
α_{12}	-0.40	117.47	-2.22	10.09	α_{36}	-2.64	1.04	-2.67	1.00		
α_{13}	-3.61	1.01	-3.62	1.00	α_{37}	-2.56	1.04	-2.60	1.00		

α_{14}	-3.58	1.01	-3.60	1.00	α_{38}	-2.48	1.04	-2.53	1.00		
α_{15}	-3.56	1.01	-3.57	1.00	α_{39}	-2.40	1.04	-2.44	1.00		
α_{16}	-3.53	1.01	-3.54	1.00	α_{40}	-2.30	1.05	-2.35	1.00		
α_{17}	-3.50	1.02	-3.51	1.00	α_{41}	-2.20	1.05	-2.25	1.00		
α_{18}	-3.47	1.02	-3.48	1.00	α_{42}	-2.08	1.06	-2.14	1.00		
α_{19}	-3.43	1.02	-3.45	1.00	α_{43}	-1.95	1.07	-2.01	1.00		
α_{20}	-3.40	1.02	-3.42	1.00	α_{44}	-1.79	1.08	-1.87	1.00		
α_{21}	-3.37	1.02	-3.38	1.00	α_{45}	-1.61	1.10	-1.70	1.00		
α_{22}	-3.33	1.02	-3.35	1.00	α_{46}	-1.39	1.12	-1.50	1.00		
α_{23}	-3.30	1.02	-3.31	1.00	α_{47}	-1.10	1.15	-1.25	1.00		
α_{24}	-3.26	1.02	-3.28	1.00	α_{48}	7.25	578.26	0.89	30.02		
100 Sample Size										4	742
α_{1}	-3.13	0.60	-3.17	0.58	α_{25}	-3.43	0.72	-3.44	0.71		
α_{2}	-3.05	0.60	-3.10	0.59	α_{26}	-3.92	0.94	-3.94	0.93		
α_{3}	-2.89	0.60	-2.96	0.58	α_{27}	-3.72	0.86	-3.74	0.85		
α_{4}	-1.78	16.73	-2.59	5.83	α_{28}	-3.88	0.94	-3.89	0.93		
α_{5}	-3.61	0.74	-3.63	0.73	α_{29}	-3.32	0.72	-3.34	0.71		
α_{6}	-3.55	0.74	-3.58	0.73	α_{30}	-3.81	0.94	-3.83	0.93		
α_{7}	-3.38	0.69	-3.40	0.67	α_{31}	-3.61	0.87	-3.63	0.85		
α_{8}	-3.48	0.75	-3.51	0.73	α_{32}	-3.76	0.94	-3.77	0.93		
α_{9}	-3.30	0.71	-3.33	0.69	α_{33}	-3.20	0.72	-3.22	0.71		
α_{10}	-3.18	0.69	-3.22	0.67	α_{34}	-3.51	0.87	-3.53	0.85		
α_{11}	-3.15	0.73	-3.20	0.70	α_{35}	-3.47	0.87	-3.49	0.85		
α_{12}	-1.57	82.38	-2.85	7.22	α_{36}	-3.44	0.87	-3.46	0.85		
α_{13}	-3.84	0.77	-3.85	0.77	α_{37}	-3.04	0.72	-3.07	0.71		
α_{14}	-4.24	0.95	-4.25	0.94	α_{38}	-3.35	0.87	-3.37	0.85		
α_{15}	-4.09	0.89	-4.10	0.88	α_{39}	-3.30	0.87	-3.32	0.85		
α_{16}	-4.21	0.95	-4.22	0.94	α_{40}	-3.25	0.87	-3.28	0.85		
α_{17}	-3.77	0.77	-3.78	0.77	α_{41}	-2.74	0.73	-2.77	0.71		
α_{18}	-4.17	0.95	-4.18	0.94	α_{42}	-2.90	0.83	-2.93	0.80		
α_{19}	-4.01	0.89	-4.02	0.88	α_{43}	-2.83	0.83	-2.86	0.80		
α_{20}	-4.13	0.95	-4.14	0.94	α_{44}	-2.75	0.84	-2.79	0.81		
α_{21}	-3.52	0.72	-3.54	0.71	α_{45}	-2.41	0.74	-2.47	0.71		
α_{22}	-4.02	0.94	-4.03	0.93	α_{46}	-2.54	0.85	-2.59	0.81		
α_{23}	-3.83	0.86	-3.84	0.85	α_{47}	-2.40	0.86	-2.47	0.81		
α_{24}	-3.98	0.94	-3.99	0.93	α_{48}	3.07	272.88	-1.12	14.48		
500 Sample Size										4	4,376
α_{1}	-3.18	0.28	-3.22	0.27	α_{25}	-4.13	0.45	-4.14	0.45		
α_{2}	-3.08	0.28	-3.13	0.27	α_{26}	-4.15	0.46	-4.16	0.46		
α_{3}	-2.94	0.28	-3.01	0.27	α_{27}	-4.11	0.46	-4.12	0.45		
α_{4}	-1.76	19.94	-2.70	2.63	α_{28}	-4.08	0.46	-4.09	0.45		
α_{5}	-3.66	0.34	-3.68	0.34	α_{29}	-4.09	0.46	-4.10	0.46		
α_{6}	-3.60	0.34	-3.63	0.34	α_{30}	-4.00	0.45	-4.01	0.44		
α_{7}	-3.59	0.35	-3.61	0.34	α_{31}	-4.04	0.46	-4.05	0.46		
α_{8}	-3.51	0.34	-3.54	0.34	α_{32}	-3.98	0.46	-4.00	0.45		

α_{9}	-3.43	0.34	-3.47	0.34	α_{33}	-3.97	0.46	-3.98	0.45		
α_{10}	-3.36	0.34	-3.40	0.34	α_{34}	-3.96	0.46	-3.98	0.46		
α_{11}	-3.27	0.35	-3.33	0.34	α_{35}	-3.89	0.45	-3.90	0.45		
α_{12}	-1.64	28.91	-2.91	2.61	α_{36}	-3.85	0.45	-3.86	0.45		
α_{13}	-4.39	0.46	-4.39	0.46	α_{37}	-3.87	0.46	-3.89	0.46		
α_{14}	-4.37	0.46	-4.38	0.46	α_{38}	-3.84	0.46	-3.86	0.46		
α_{15}	-4.29	0.45	-4.29	0.44	α_{39}	-3.78	0.46	-3.80	0.45		
α_{16}	-4.31	0.45	-4.32	0.45	α_{40}	-3.69	0.45	-3.71	0.44		
α_{17}	-4.32	0.46	-4.33	0.46	α_{41}	-3.73	0.46	-3.75	0.46		
α_{18}	-4.28	0.46	-4.29	0.45	α_{42}	-3.66	0.46	-3.68	0.45		
α_{19}	-4.29	0.46	-4.30	0.46	α_{43}	-3.64	0.47	-3.66	0.46		
α_{20}	-4.19	0.44	-4.20	0.44	α_{44}	-3.56	0.46	-3.58	0.45		
α_{21}	-4.23	0.46	-4.24	0.45	α_{45}	-3.42	0.44	-3.45	0.43		
α_{22}	-4.23	0.46	-4.24	0.46	α_{46}	-3.38	0.44	-3.41	0.43		
α_{23}	-4.22	0.46	-4.22	0.46	α_{47}	-3.30	0.45	-3.34	0.43		
α_{24}	-4.14	0.45	-4.15	0.45	α_{48}	0.05	73.42	-2.47	4.09		
1000 Sample Size										3	8,739
α_{1}	-3.17	0.19	-3.21	0.19	α_{25}	-4.16	0.32	-4.17	0.32		
α_{2}	-3.07	0.19	-3.11	0.19	α_{26}	-4.12	0.32	-4.14	0.32		
α_{3}	-2.95	0.20	-3.01	0.19	α_{27}	-4.11	0.32	-4.12	0.32		
α_{4}	-1.67	23.12	-2.68	1.86	α_{28}	-4.08	0.32	-4.09	0.32		
α_{5}	-3.67	0.24	-3.69	0.24	α_{29}	-4.08	0.32	-4.09	0.32		
α_{6}	-3.60	0.24	-3.62	0.24	α_{30}	-4.02	0.32	-4.04	0.32		
α_{7}	-3.57	0.24	-3.60	0.24	α_{31}	-4.03	0.32	-4.04	0.32		
α_{8}	-3.48	0.24	-3.51	0.23	α_{32}	-3.94	0.31	-3.95	0.31		
α_{9}	-3.44	0.24	-3.47	0.24	α_{33}	-3.96	0.32	-3.98	0.32		
α_{10}	-3.35	0.24	-3.39	0.24	α_{34}	-3.93	0.32	-3.95	0.32		
α_{11}	-3.25	0.25	-3.31	0.24	α_{35}	-3.92	0.33	-3.93	0.32		
α_{12}	-1.65	21.68	-2.92	1.94	α_{36}	-3.86	0.32	-3.88	0.32		
α_{13}	-4.33	0.31	-4.34	0.31	α_{37}	-3.86	0.33	-3.88	0.32		
α_{14}	-4.34	0.32	-4.35	0.32	α_{38}	-3.81	0.32	-3.83	0.32		
α_{15}	-4.33	0.32	-4.34	0.32	α_{39}	-3.78	0.32	-3.80	0.32		
α_{16}	-4.26	0.31	-4.27	0.31	α_{40}	-3.69	0.31	-3.71	0.31		
α_{17}	-4.31	0.32	-4.32	0.32	α_{41}	-3.71	0.33	-3.73	0.32		
α_{18}	-4.26	0.32	-4.27	0.32	α_{42}	-3.64	0.32	-3.67	0.32		
α_{19}	-4.27	0.32	-4.28	0.32	α_{43}	-3.62	0.33	-3.65	0.32		
α_{20}	-4.24	0.32	-4.25	0.32	α_{44}	-3.56	0.33	-3.58	0.32		
α_{21}	-4.23	0.32	-4.24	0.32	α_{45}	-3.50	0.33	-3.53	0.32		
α_{22}	-4.20	0.32	-4.21	0.32	α_{46}	-3.43	0.33	-3.46	0.32		
α_{23}	-4.20	0.32	-4.21	0.32	α_{47}	-3.36	0.33	-3.40	0.32		
α_{24}	-4.11	0.31	-4.12	0.31	α_{48}	0.00	51.93	-2.52	2.90		

Interestingly, even when the hazard in the respective time period of different sample sizes was the same, discrepancies were detected only in the smallest sample size. For example, in the case of 0% censoring and 48 time periods, the hazard of the time period 36 was 0.07 in all four sample sizes. The discrepancy between discrete and continuous estimation was detected only in 50 sample size. Furthermore, even when the hazard of with the smaller samples was lower compared with the larger samples, the discrepancy was detected only in the smaller sample sizes. For instance, in the case of 40% censoring and four time periods, the hazard of the time period 4 was 0.25 for 50 samples while that of the larger samples was 0.27. Out of these four sample sizes, only 50 sample size showed the discrepancy in the time period four.

Hazard Rates of Parameter Estimates

The study also calculated and compared the raw hazard rates of the parameter estimates that showed discrepancies between the clog-log and logit models. The results showed that the hazard rates of the estimates ranged from 0.12 to 0.50, with 14 cases having hazard rates higher than 0.21. The important findings of the study regarding the hazard rates and discrepancies of the two models were that, when the parameter estimates showed a discrepancy, they displayed high hazard rates. However, the reverse was not true. In other words, the high hazard rates did not always lead to discrepancies.

Overall, the magnitudes of the parameter estimates of the logit and clog-log models were similar, while the magnitudes of the logit model were larger than those of the clog model (Allison, 2010). The average hazard estimates from both the logit and clog-log models are included in Tables 13, 14, and 15. It is important to note that a big discrepancy in the hazard estimates was detected in the case of 0% censoring. The hazard estimates of the logit models ranged from 14.2029 to 17.2029, while those of the clog-log models ranged from 2.6824 to

2.8746.

AIC Differences

The study found that in the models with logit, a small number of time metrics, a small proportion of censoring, and a small sample tended to show a low AIC value (See the tables for details). For example, among 180 models, the smallest AIC value (177) was found in the logit model with 0% censoring, four time metrics, and 50 sample cases. The largest AIC (86,831) was found in the clog-log model with 80% censoring, 48 time metrics, and 1,000 cases.

In all conditions, the logit models showed better fit statistics with lower values of AIC than the clog-log models for the same conditions, showing a difference ranging from 24 to 26,350. The smallest AIC difference (24) was from the comparison of a logit model (AIC =177) with a clog-log model (AIC=202) with 0% censoring, four time metrics, and 50 cases. The largest difference (26,350) resulted from the comparison of a logit model (AIC= 6,011) with a clog-log model (AIC= 86,361) with 80% censoring, 48 time metrics, and 1,000 cases.

The AIC differences relating to censoring conditions revealed the tendency that higher censoring was associated with more differences between discrete and continuous models. In case of sample sizes of 500 and 1,000, higher censoring led to more differences in all three time periods (4, 12, and 48 time periods). In case of small sample sizes such as 50 and 100, more censoring resulted in greater deviances in four and twelve time periods, but not in the 48 time period. In the 48 time period, there was no clear association between censoring status and discrepancies.

In case of time periods, more time periods led to greater AIC values in both discrete and continuous models, thus making discrepancies between the two models greater. When having the censoring status and sample size constant, AIC difference was the least in four time periods while it was the greatest in 48 time periods, with that in 12 time periods in-between. For example, in case of 80% censoring and 1,000 sample size, the deviance between the discrete and continuous estimation was 2,054 in four time periods while it was 26,357 in 48 time periods, having 6,479 in 12 time periods.

In case of sample size, the smaller sample was related to the smaller AIC value of discrete and continuous survival analysis, having the smaller difference in AIC values between discrete and continuous survival models. Holding censoring status and time periods constant, the difference was the greatest in the sample size of 1,000, having it the smallest in the sample size of 50. For instance, in case of 0% censoring and four time periods, the difference in the 50 sample size was 24 while it was 452 in the 1,000 sample size, having 45 and 226 for the 100 and 500 sample sizes, respectively.

Difference in the Magnitudes of Parameter Estimates

In case of 0% censoring, the big discrepancy in hazard estimates was detected in the last time period where the hazard rate was one in all sample sizes and time periods. The hazard of discrete survival analysis ranged from 14.2029 to 17.2029 while that of continuous survival analysis ranges from 2.6824 to 2.8746.

Detailed Difference in Terms of 0% Censoring Proportion

When there is no censoring, the most of discrepancies in terms of the significance level in the parameter estimates were noted. Out of four sample sizes (50, 100, 500, and 1,000), the smaller sample size revealed more of discrepancies. In case of the sample size of 50, the significance level difference was detected in all three time periods (4, 12, and 48). With more time periods, there were more discrepancies in the significance levels. In case of four time periods, only one discrepancy was detected in the time period 2. The hazard in discrete survival analysis was significant at the .05 level while that in clog-log survival analysis was significant at less than .01 level. In 12 time periods, there were two discrepancies in time periods of 9 and 10. The hazard in the time period 9 in discrete survival revealed the significance at .05 level contrasting with the .01 level in the continuous survival analysis. The hazard in the time period 10 of discrete survival analysis was not significant while that of continuous survival analysis was significant at .05 level. With the 48 time period, significant differences were detected in three hazards out of 48 hazards. The hazards in the time periods 36 and 37 in discrete survival analysis were significant at .05 level, while those in continuous survival analysis were significant at less than .01 level. In case of time period 43, the hazard in discrete survival analysis was not significant but that in continuous survival analysis was significant at .05 level.

In case of sample size 100, the significance level difference was noted in the two time periods of 12 and 48. With 12 time period of 12, one difference was detected in the time period 10 at the .05 significance level in discrete survival analysis and at the .01 significance level in continuous survival analysis. With 48 time period, the difference was detected in the hazards in the time periods of 43, 44 and 46. In time periods of 43 and 44, the hazards in discrete estimation were significant at .05 level but those in continuous estimation were significant at .01 level. In the time period of 46, the hazard of continuous estimation was significant at .05 level, in contrast with no significance level of discrete estimation.

In case of sample sizes of 500 and 1,000, the same pattern was noted in the discrepancy in terms of the significance level. In the case of four time periods, the difference in the hazard significant level was noted in the time period 3. In the same way, with 12 time periods, the time difference in the significance level was found in the time period 11. The hazards of discrete analysis were not significant, but those of continuous analysis were significant at less than .01 level in both time periods of 3 and 11.

Detailed Difference in Terms of 20% Censoring Proportion

In case of 20% censoring, the difference in the significance levels was detected in all the sample sizes but not in the four and 12 time periods. Compared with the 0% censoring, the less number of discrepancies were detected. For example, 0% censoring condition revealed the discrepancies in the 50 and 100 sample sized in the 48 time period. However, in the 20% censoring condition, no discrepancy was found in 48 time periods.

In the sample size of 50, the difference in the hazard of the time period 3 was detected in 4 and 12 time periods. In the four time period, the hazard of discrete analysis in the time period 3 was not significant while that of continuous analysis was significant at .01 level. In the 12 time periods, the hazard of the discrete analysis in the time period 12 was significant at .05 level while that of the continuous analysis was significant at less than .01 level.

With 100 and 500 sample sizes, the same pattern was noted: the discrepancy was detected in the time period 3 in the 4 time periods while the discrepancy was found in the time periods of 12 in 12 time periods. In the discrete analysis with the 100 sample size, the hazard of the time periods 3 out of 4 time periods was significant at .05 level while that in the continuous analysis was significant at less than .01 level. With 500 sample sizes in the 4 time period, the hazard of the time period 3 in the discrete estimation was not significant, but that in the continuous estimation was significant at less than .01 level. With 100 and 500 sample sizes in twelve time periods, the hazard of the time period 12 of the discrete estimation was significant at .05 level while that of the continuous estimation was significant at less than .01 level.

With 1,000 sample sizes, out of 3 cases of the time periods, the discrepancy between the discrete and continuous estimation was detected only in four time periods. The hazard of discrete analysis in the time period 3 out of 4 time periods was not significant, but that of continuous analysis was significant at less than .01 level.

Detailed Difference in Terms of 40%, 60% and 80% Censoring Proportions

When 40% censoring occurred, the discrepancy was noted in only one case out of 12 cases (3 time periods*4 sample sizes). With the 4 time intervals and 50 sample sizes, there was only one discrepancy in the time period of 4. The hazard of the discrete analysis was significant at .05 level but that of continuous analysis was significant at less than .01 level.

In case of 60% censoring, no discrepancy was detected in terms of the significance level of the parameter estimates. As noted in the 60% censoring condition, the 80% censoring condition did not demonstrate any discrepancy in significance levels of the parameter estimates.

CHAPTER FIVE

DISCUSSION

In order to provide empirical results in adopting a suitable model for survival analysis for educational research, this study compared two popular models of discrete and continuous survival analysis. As important factors to differentiate the use of two models, this study considered three factors of time metrics, censoring proportion, and sample sizes. To explain the importance of each of the factors, this study included hazards as a part of the study because hazards considering the three factors have not yet been examined despite the importance of the factors in choosing a survival model for research. To examine the effects of the factors, this study provided both empirical and simulated evidence by building a series of discrete and continuous survival models and presented summarized results from the analyses. The study was able to reach practical recommendations specific to each of the three factors in building and applying a discrete or continuous model. Importantly, this study was able to provide model recommendations in consideration of each factor in the contexts of the other two factors. In other words, the major contribution of the study is in providing recommendations for the model specifications for combined effects of the factors in the educational contexts, differing from existing studies that studied the separate effects of each of the three factors.

First Approach

Four Time Metrics

In response to the lack of studies on the adoption of an appropriate time estimation and metric when building survival models, this study aimed to provide guidelines in this regard based on empirical evidence. This study focused attention on 1) two time estimation methods (discrete and continuous time estimation); and 2) four time metrics (year, half-year, quarter, and month). Using large data sets from NLSY97, this study constructed eight survival models to review these different time estimation methods and metrics. These eight models incorporated the same predictors and outcome variable, differing only in the time estimation and metrics. To identify potential differences across models, this study examined the absolute magnitudes and statistical significance of individual parameter estimates and model fit statistics.

Importance of Choosing a Proper Model

The study results emphasized the proper choice of respective time estimation method and time metric for a survival analysis. When fitting the models with the same components to the same data, the survival analyses using different time estimations and metrics yielded different results across models: different statistical significance of some predictors and different model fit statistics. These differences were found not only between discrete and continuous models, but also among different time metrics within the same time estimation method (discrete- or continuous- time estimation). The resulting discrepancies strongly suggest that researchers need to be careful to adopt an appropriate time estimation and metrics in order to avoid biased results of survival analyses.

Discrete vs. Continuous Models

In terms of the model comparison between discrete and continuous time estimation, the findings of this study supported the use of discrete models to fit the data with relatively large time metrics: month, quarter, half-year and year. The fit statistics supported discrete models, which demonstrated far better fit than continuous models. These results were aligned with recommendations by Singer and Willet (2003): a discrete-time estimation method was preferred when the event time was measured in discrete time units such as months, quarter, or year.

Time Metrics for Discrete Models

This study identified a year as the most appropriate time metric out of the four time metrics for discrete models of this study. The year unit was supported on not only empirical but also theoretical (substantive) basis. Empirically, the year model demonstrated better fit than other time metrics. Theoretically, this year model was appropriate to capture the annual effects of the dropout (Plank, 2008); a lot of dropouts occurred at the beginning of entry into high school and right before high school graduation.

Time Metrics for Continuous Models

This study revealed that a month was the most proper time metric out of four time metrics when building continuous models. To begin with, continuous models were not recommendable for use with time metrics such as month, quarter, half-year, and year (Singer & Willet, 2001). The time metric for continuous time survival analysis should be as fine as possible, such as hours, days, or weeks. However, the reality is that many studies have employed a continuous survival analysis when measuring the event occurrences in large time units such as month, quarter, half-year and year (Doyle, 2006; Kelly, 2004; Kirby, 1999; Murphy, 2010; Murtaugh, 1999; Plank, 2008).

While some of these studies should have adopted the discrete survival analysis, especially in the case of the year metric, this study investigated the adoption of continuous survival analysis to accommodate the case; continuous time survival analysis would be the only option. Sometimes, it is not possible to build discrete models with limited numbers of data, as the discrete model should include the number of time indicators equal to the number of time units of the study. For example, examining 10 waves of data using the month metric would lead to the inclusion of 120 time indicators in the model: the multiplication of 12 months with 10 years would be 120. The huge number of time indicators would result in no model convergence or in biased results even when converged. Therefore, in the case when continuous time estimation is the only option, a month is the most recommendable unit out of four time metrics: year, half-year, quarter and month.

Future Studies

The findings of this study might be further augmented by future studies, which will fit the models to data with different numbers of events. Future studies, which will adopt smaller or larger numbers of events than the 200 events of the current study, will lead to different conclusions on a proper time estimation and time metrics. Studies with smaller numbers of event times would have difficulty when fitting discrete models with a large number of time indicators as a result of adopting smaller time units. In this case, a proper time metric for discrete models might differ from this study's month unit. Furthermore, a future study with a smaller number of events might find better effectiveness of continuous models as smaller number of events will lead to smaller numbers of tied events. A future study might reveal fewer discrepancies across continuous models, as well as between discrete and continuous models.

On the other hand, future studies that simulate the current data to derive a greater number of events would provide another insight on time estimation methods and time units. Discrete models with more events would have more power to treat large numbers of time indicators. Thus, a future study might find better fit and more pronounced effects of some predictors, favoring the discrete models with finer time unit than this study's year unit. At the same time, more events with the same event times would aggravate the issues with ties, thus causing more problems with fitting continuous survival analysis. A future study might find more discrepancies across continuous models, as well as between discrete and continuous models.

In addition to considering different numbers of event times, further study is needed for using different time specifications such as linear or quadratic functions in building discrete models. The current study developed discrete-time models using the general time specification, which was similar to a general spline of time estimation of continuous models. Different time specifications in constructing discrete models might lead to more pronounced differences between discrete and continuous models. Future study might identify different magnitudes of individual predictors along with different statistical significance between discrete and continuous models, beyond the statistical significance and goodness-of-fit that the current study found.

Second Approach

Three Factors

To compare discrete and continuous survival models, the second approach adopted a logit model as a discrete model and a clog-log model as a continuous model. Just like the first approach, the second approach confirmed the need for a guideline through the study's results, which indicated discrepancies between the discrete and continuous models in many conditions. The study also verified the importance of censoring proportions, sample size, and in addition to time metrics in choosing survival models. At the same time, this study confirmed that the three factors of time metrics, censoring proportions, and sample sizes influence hazards and in turn determine the discrepancies between the discrete and continuous models. Furthermore, the study addressed the interaction effects of the three factors affecting the discrepancy of the outcomes of a logit model and a clog-log model in the same condition.

To examine the effect of the three factors, this study generated 60 sets of data by combining different levels of the factors: time metrics (4, 12, and 48); censoring proportions (0%, 20%, 40%, 60%, and 80%); and sample size (50, 100, 500, and 1,000). After employing two methods to each of 60 simulation conditions, the study compared the parameter estimates and fit statistics to evaluate the performance of the two models.

Time Metrics

In terms of the time metrics, a small number of time metrics was associated with greater discrepancies. A small number of time metrics resulted from large units of time metrics (i.e., year of the first approach). Notably, these results were aligned with not only empirical findings on the time metrics of the first approach but also the study by Hofstede and Wedel (1999).

Furthermore, this study enriched existing findings on the time metrics by relating to two other factors of censoring proportions and sample sizes. This discrepancy pattern in a small number of time metrics was particularly detected with small censoring proportions because, when there is a small number of time metrics, there will be a large number of tied observations and a high hazard rate for a period with a fixed number of cases.

Thus, when the data are measured using large time metrics such as 4 or 12 times (i.e., year, half-year, quarter) with small censoring proportions (less than or equal to 20%), it is recommended to build a discrete model regardless of sample size. For these conditions, a continuous model would produce biased estimates. When the data are measured in fine units such as 48 time metrics (i.e., month), the building of either continuous or discrete models should be considered.

Censoring Proportion

The study results indicated that a small censoring proportion was associated with greater discrepancies because low censoring indicates more events during each time period. In particular, discrepancies occurred in the conditions with less than 40% censoring. This finding is aligned with the study results done by Hertz-Picciotto and Rockhill (1997).

The discrepancies are more pronounced when low censoring was combined with small sample sizes and small time metrics. Thus, this study recommends the following: when the censoring proportions are equal to or less than 40%, it is desirable to build a discrete model. In particular, with small censoring proportions in small samples (equal to or less than 500) with small time metrics, it is recommended to build a discrete model. The finding on the marked effect of low censoring proportion in the small sample and the small number of time metrics are new addition to the existing studies on censoring proportion (Colosimo, Chalita, & Demétrio , 2000; Hess, 2009).

Sample Sizes

With regards to sample size, smaller samples were associated with greater discrepancies between two models. This finding provided empirical data to conclude the mixed effect of sample sizes (Hertz-Picciotto & Rockhill, 1997; Hess, 2009). Also, this study showed the importance of considering hazard rates, proportion of tied events, in discussing the effect of sample sizes. However, compared with other factors of censoring proportions and time metrics, the sample size did not lead to much discrepancy between the two models.

Hazard Rates

For hazard rates, as suggested by Singer and Willet (2003), higher hazard rates were associated with greater discrepancies. Higher hazards included ones greater than 0.15, particularly, 0.30 as a critical point (Hess, 2009; Hosmer & Lemeshow, 1999). In this context, the study finding also provided empirical evidence on hazards greater than 0.20. The study suggests that a choice of model based on hazard rates should be made in consideration of other three factors.

Discrete vs. Continuous Models

Overall, the discrete models showed better fit statistics than the continuous models. In all 60 conditions, the discrete models indicated comparatively small deviances. Singer and Willet (2003) suggested that discrete models should be used when there are more tied observations than unique observations.

Future Studies

By employing both continuous and discrete models to the simulated data, the study estimated hazard rates for the parameter estimates, exploring discrepancies between two models. The main focus of the study was on detecting the effects of the three factors of time metrics, censoring proportions, and sample size in addition to hazard rates. However, the study did not pay attention to other covariates (other predictor variables). The study suggests that future studies should extend the scope by including other important covariates and estimate hazards in order to examine interactional effects.

REFERENCES

Allison, P. (2010). *Survival Analysis Using SAS® : A Practical Guide Second Edition*. Cary, NC: SAS Publishing.

Allison, P. (1999). *Logistic Regression Using SAS® : Theory and Application*. Cary, NC: SAS Publishing.

Calcagno, J. C., Crosta, P., Bailey, T., & Jenkins, D. (2007). Does Age of Entrance Affect Community College Completion Probabilities? Evidence From a Discrete-Time Hazard Model. *Educational Evaluation and Policy Analysis*, 218-235.

Center for Human Resource Research. (2010). *NLSY97 User's Guide: A Guide to Rounds 1-8 Data*. Columbus, OH: Author.

Colosimo, E. A., Chalita, L. V. A. S., & Demétrio, C. G. B. (2000). Tests of Proportional Hazards and Proportional Odds Models for Grouped Survival Data. *Biometrics, 56*(4), 1233-1240.

Corrente, J. E., Chalita, L., & Moreira, J. A. (2003). Choosing between Cox proportional hazards and logistic models for interval- censored data via bootstrap. *Journal of Applied Statistics*, *30*(1), 37 - 47.

Donaldson, M. L., & Johnson, S. M. (2010). The Price of Misassignment: The Role of Teaching Assignments in Teach For America Teachers' Exit From Low-Income Schools and the Teaching Profession. *Educational Evaluation and Policy Analysis 32*(2), 299-323.

Doyle, W. R. (2006). Adoption of Merit-Based Student Grant Programs: An Event History Analysis. *Educational Evaluation and Policy Analysis*, 259-285.

Hertz-Picciotto, I., & Rockhill, B. (1997). Validity and Efficiency of Approximation Methods for Tied Survival Times in Cox Regression. *Biometrics, 53*(3), 1151-1156.

Hess, W. (2009). A flexible hazard rate model for grouped duration data. *Working papers series (Tartu Ü likool. Majandusteaduskond)*.

Hess, W., & Persson, M. (2010). *The duration of trade revisited: Continuous-time vs. discrete-time hazards*. Unpublished manuscript.

Hofstede, F. t., & Wedel, M. (1999). A Monte Carlo study of time aggregation in continuous-time and discrete-time parametric hazard models. *Economics Letters, 58*(2), 149-156.

Hosmer, D. W., & Lemeshow, S. (1999). *Applied survival analysis: regression modeling of time to event data*. New York, N.Y.: Wiley.

Jacobs, J. A., & King, R. B. (2002). Age and College Completion: A Life-History Analysis of Women Aged 15-44. *Sociology of Education, 75*(3), 211-230.

Kahn, J. M., & Schwalbe, C. (2010). The timing to and risk factors associated with child welfare system recidivism at two decision-making points. *Children and Youth Services Review, 32*(7), 1035-1044.

Kelly, S. (2004). An event history analysis of teacher attrition: Salary, teacher tracking, and socially disadvantaged schools. *Journal of Experimental Education, 72*(3), 195-220.

Kirby, S. N., Berends, M., & Naftel, S. (1999). Supply and Demand of Minority Teachers in Texas: Problems and Prospects. *Educational Evaluation and Policy Analysis*, 47-66.

Klein, J. P., & Moeschberger, M. L. (2003). *Survival analysis: Techniques for censored and truncated data*. New York: Springer.

Kleinbaum, D. G., & Klein, M. (2011). *Survival analysis: A self-learning text.* New York: Springer-Verlag.

Ma, X., & Willms, J. D. (1999). Dropping Out of Advanced Mathematics: How Much Do Students and Schools Contribute to the Problem? *Educational Evaluation and Policy*

Analysis, 365-383.

Murphy, T. E., Gaughan, M., Hume, R., & Moore, S. G. (2010). College Graduation Rates for Minority Students in a Selective Technical University: Will Participation in a Summer Bridge Program Contribute to Success? *Educational Evaluation and Policy Analysis*, 70-83.

Murtaugh, P. A., Burns, L. D., & Schuster, J. (1999). Predicting the retention of university students. *Research in Higher Education, 40*(3), 355-371.

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. [Article]. *Sociology of Education, 81*(4), 345-370.

Prentice, R. L., & Gloeckler, L. A. (1978). Regression Analysis of Grouped Survival Data with Application to Breast Cancer Data. *Biometrics, 34*(1), 57-67.

Singer, J. D. (1992). Are special educators' career paths special? Results from a 13-year longitudinal study. *Exceptional Children*, *59*(3), 262-279.

Singer, J. D., Davidson, S. M., Graham, S., & Davidson, H. S. (1998). Physician Retention in Community and Migrant Health Centers: Who Stays and for How Long? *Medical Care, 36*(8), 1198-1213.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: modeling change and event occurrence*. Oxford; New York: Oxford University Press.