The Influence of Variability in Measurement Occasions in Discrete-Time Survival Analysis

Thesis Pedagogical Sciences

Student: **Lieke Hesen**

Student code: **3988252**

Supervisor: **Dr. M. Moerbeek**

Second reviewer: **Dr. M. Safarkhani**

Hand-in date: **26th of June, 2015**

Faculty: **Methods and Statistics**

Abstract

Within-wave variability is common in longitudinal studies. Previous studies show this does not influence slope parameters and standard error in latent growth models. The current study aimed to extend these findings for discrete time-survival analysis in an experimental setting. In an experimental setting, people randomly assigned to either a treatment or control condition are measured to establish a treatment effect. By doing a simulation study, parameters for discrete-time survival analysis are estimated both for a fixed time points distribution and for varying time points distributions: a truncated normal, uniform and truncated chi-square distribution. Using an underlying Weibull function, data was simulated for these distributions, while changing width of the measurement interval (0, 0.25, 0.5, 1). These simulations created 270 scenarios. Results for treatment effect bias showed how only one scenario exceeded the criterion of Hoogland and Boomsma (1998) presenting more than 5% parameter bias when width of a truncated normal distribution was set to 1. In addition, the percentage standard error bias did not exceed the criterion of 10%. Overall, it can be concluded that increasing measurement intervals and varying shape of measurement distribution do not substantially influence discrete-time survival analysis. This may lead to new standards in the practice of survival analysis.

Keywords: discrete-time survival analysis, within-wave variability, Weibull distribution, simulation, measurement, occasions, interval

The Influence of Variability in Measurement Occasions in Discrete-Time Survival Analysis

For decades empirical researchers have been interested in investigating change. Both studies on natural change, such as learning to walk, and change caused by targeted interventions, for instance medication, can reveal information about development (Singer & Willet, 2003). Over the past 25 years, more attention has been given to the development of methodology appropriate for longitudinal data (Mehta & West, 2000). Moreover, since 1980 researchers have been able to investigate change adequately, given that they had developed appropriate statistical methods, such as individual growth modeling (Singer & Willet, 2003).

More specifically than change, psychologists are often interested in whether and when events occur (Singer & Willet, 1991). Research questions for instance focus on the premature termination of psychotherapy (see Corning & Malofeeva, 2004). Also studies investigating onset, cessation, relapse and recovery focus on the occurrence and timing of an event (see Willet & Singer, 1993). The when question introduces the need for the outcome value, time, to be known for every individual. However this is rarely the case (Singer & Willet, 1993). These types of studies therefore indicate the need for a special kind of statistical analysis: Survival analysis.

Survival methods, especially based on continuous-time, have become more popular since Cox's 1972 publication on them. After being introduced in biostatistics, mainly areas such as sociology and economics, have widely picked up survival analysis (Singer & Willet, 1993; Willet & Singer, 1991). Survival analysis is powerful in use and easily applicable to psychological phenomena. (Singer & Willet, 1991). For example, questions concerning the effect of prevention programs for smoking can be studied or studies can determine the effect of home visit programs (See Le Roux et al., 2010; Storr, Ialogo, kellam, & Anthony, 2002).

Survival analysis can be defined as a 'class of techniques designed for studying the occurrence of events in longitudinal data; in short, it estimates models to predict the risk of

occurrence of an event given that the event has not yet occurred' (Corning & Malofeeva, 2004, p. 354). These techniques enable researchers to give a description of temporal patterns and compare these among groups (Willet & Singer, 1991). Moreover, it creates the possibility of developing statistical models describing the risk of event occurrence (Willet & Singer, 1991). Events can be specified as a change from one state to another (Singer & Willet, 2003). These transitions can be studied in continuous or discrete time. Continuous-time survival analysis requires the timing of an event to be presented on a continuous scale (Willet & Singer, 1993). Time indicators are in this case specific moments, for instance days or minutes (Corning & Malofeeva, 2004). In contrast, discrete-time survival analysis is based only on the information of the period in which the event occurred (Willet & Singer, 1993). For this type of analysis bounded intervals, such as grades in school, are used to specify time (Corning & Malofeeva, 2004). Continuous-time survival analysis used to be popular, however Singer and Willet (1993) put forward how discrete-time survival analysis is a more suitable framework for describing the relationship between predictors and events. The current study will focus on discrete-time survival analysis.

Discrete-time survival analysis can be applied within an experimental study. An experiment is a: ' study in which an intervention is deliberately introduced to observe its effects ' (William, Shadish, Cook, & Campbell, 2002, p. 13). There is both a control and one or more treatment conditions. In the case of a randomized experiment, every participant is randomly assigned to either a control or treatment condition. All participants have exactly the same chance of being in either group (William et al., 2002). By conducting an experiment the effect of a treatment condition can be assessed: the treatment effect. The treatment is also sometimes referred to as a predictor. In discrete-time survival analysis study, there can be one or more predictors, dichotomous or continuous. In the current study the focus will be on dichotomous predictors, for instance a smoking prevention program in a study investigating

smoking behavior in adolescence. This predictor, the prevention program, creates two conditions: adolescents who are assigned to the prevention program and those who are not. When adolescents are assigned to the prevention program, this can be called the treatment condition, and the other group can be called the control condition. When doing discrete-time survival analysis the conditions can be compared and the effect of the predictor can be determined, resulting in a treatment effect. In this case the prevention program can be viewed as the treatment or predictor. The statistical theory behind the use of discrete-time survival analysis in experimental settings will be explained further under the heading statistical model.

Furthermore, when using discrete-time survival analysis for the mentioned purposes, it overcomes four main problems that traditional methods applied to event occurrence faced (Corning & Malofeeva, 2004). First of all, discrete-time survival analysis enables straightforward parameter interpretation and it creates comprehensible results, for instance the hazard as a probability instead of as a rate (Singer & Willet, 1993).

Moreover, when using discrete-time survival analysis the results are more informative than provided by traditional longitudinal methods. For example in premature termination studies, the outcome variable is often categorized, providing only two options: either the event occurred or it did not. Using this method all participants are treated the same, ignoring the aspect of time. Important information on relationships between variables is lost (Corning & Malofeeva, 2002). In discrete-time survival analysis the shape of the hazard function can be studied. This provides more information on the actual risks in relation to time and enables researchers to more easily answer the question when the event is most likely to occur (Singer & Willet, 1993).

In addition to the informative benefit of discrete-time survival analysis, researchers are also able to correctly incorporate time-varying predictors into the analysis. Time-varying predictors have different, varying effects on the risk of event occurrence (Corning &

Malofeeva, 2004). Furthermore discrete-time survival analysis offers the possibility of including multiple types of outcomes, introducing the competing-risks structure. This is applicable when multiple types of events can occur (Corning & Malofeeva, 2004; Singer & Willet, 1991).

Finally, the greatest advantage of discrete-time survival analysis is the way censored cases can be handled. As mentioned at the beginning in longitudinal data not for every individual the specific time at which the event occurs is known. Very often participants do not experience an event during the entire study or they leave the risk set prematurely. These types of cases are called censored (e.g. Corning & Malofeeva, 2004; Singer & Willet, 1993). Traditional methods, such as imputation, yield biased results (Singer & Willet, 1993). Therefore discrete-time survival analysis incorporates these censored cased into the calculation of the hazard. Censored cases are excluded from the risk set when they are no longer available for study (Corning & Malofeeva, 2004).

In conclusion to the above mentioned advantages, discrete-time survival analysis can be viewed as an important statistical method available to scientists studying event occurrence. Nevertheless, the use of discrete-time survival analysis also introduces the existence of new problems. One of the main challenges faced, is the variability in measurement occasions during longitudinal studies. As Aydin, Leite and Algina (2014) conclude: Researchers tend to assume that during multi-wave studies, individuals are always measured at exactly the same point in time. However during lengthy longitudinal studies this rarely happens due to practical issues. As a result, measurement of multiple individuals often happens during several months (Aydin, Leite, & Algina, 2014). This causes fixed time points to become measurement intervals, which may bias the results of longitudinal studies (Aydin et al., 2014; Blozis & Cho, 2008; Coulombe, Selig, & Delaney, 2015; Mehta & West, 2000).

Where previous studies focused mainly on latent growth models characterized by continuous event measurement and a continuous outcome variable (see Aydin et al., 2014), the current study will investigate the influence of within-wave variability on discrete-time survival analysis, where the outcome variable is dichotomous: the occurrence of an event. This aim results in the main research question: What is the effect of within-wave variability on the treatment effect and standard error resulting from discrete-time survival analysis? First the theoretical framework and the statistical background will be more thoroughly discussed. Hereafter the methods will be elucidated, also shedding light on what results are expected and what factors are included in the simulation and analysis. Next the results will be presented, leading to a discussion and conclusion. Finally limitations will be acknowledged and future recommendations will be made.

Theoretical framework: The causes and consequences of within-wave variability

In longitudinal studies there are two types of datasets: time-structured and time-unstructured. Time-structured indicates that every participant is assessed at exactly the same measurement occasions, whereas time-unstructured indicates heterogeneity in times of assessment (Coulombe et al., 2015). Longitudinal investigations tend to be characterized by this heterogeneity in measurement timing (Blozis & Cho, 2008). This variability can be seen in either different ages of individuals at the measurement occasion or varying lengths of time periods within discrete-time survival analysis. The first is caused by for instance varying ages of participants at the beginning of the study. Even though this results in similar calendar dates for individuals, the response can sometimes be influenced by differences in age. This heterogeneity in age can therefore cause bias (Mehta & West, 2000). In addition, the mentioned varying time intervals are mainly caused by practical issues (Aydin et al., 2014). In this case, some people are measured for instance a week before or after the aimed measurement occasion, extending or decreasing the time period in which the event could have

occurred. Specifically when the planning of data collection differs per individual, measurement periods can vary. Daily diary studies or data-gathering based on responses initiated by the participant also increase within-wave variability (Walls & Schafer, as cited in Sterba, 2014). Furthermore, when measurement starts at the onset of a certain condition, such as an illness, the time since onset until the event can vary. Also when using retrospective reports the event may have taken place at varying times (Blozis & Cho, 2008). During the current study the focus will be on within-wave variability concerning varying measurement occasions.

The above mentioned causes of within-wave variability have several important consequences. To date studies have mainly investigated the effect when using latent growth models. For instance Aydin and colleagues (2014) showed how this within-wave variability influences the estimates of longitudinal analysis by doing a simulation study. Factors, such as measurement interval range, sample size and type of distribution may influence the bias that is produced by within-wave variability. First, a larger range in assessment dates, when assuming fixed points, increases the amount of unexplained variance in the dependent variable. In this case an interaction effect between distribution and range exists: The larger range in combination with a uniform distribution affects residual variances the most. Range also has a significant influence on the estimation of slope variances (Aydin et al., 2014; Coulombe et al., 2015). In addition, this bias increases when distributions become more skewed. Overall the intercept and slope variances tend to be underestimated when ignoring variability (Aydin et al., 2014). In contrast, Mehta and West (2000) found these variances to be overestimated. However Mehta and West only used one dataset in which there was no systematic variation in factors such as distribution. Furthermore, Blozis and Cho (2008) found slight biases in parameter estimates across datasets with and without assumed age heterogeneity. However due to the use of two existing real datasets they were unable to

conclude the magnitude or existence of bias when ignoring age heterogeneity. Expanding on these inconclusive findings, Coulombe, Selig and Delaney (2015) concluded that skewness does contribute to bias in slope variance, but in contrast to Aydin and colleagues (2014) this bias is positive. This inconsistency could be the result of including improper solutions by Aydin and colleagues (2014) in the analysis (Coulombe et al., 2015). Moreover, sample size does not seem to create bias in parameter estimates, except for when sample size drops to 200. Lastly, in contrast to the above mentioned influenced parameters, mean slope and intercepts seem to be unbiased both in fixed time and time varying conditions. (Aydin et al., 2014; Coloumbe et al., 2015).

Furthermore, standard error may also be influenced by within-wave variability. Varying time points do produce negative bias in standard error. In contrast, the specific type of distribution and range are unrelated to standard error bias. Finally, sample size also does not seem to largely influence standard error bias. Only in the case of variable measurement occasions, standard error was influenced by an increasing sample size (Aydin et al., 2014).

The mentioned studies, all based on latent growth models, use a wide array of methods to examine the effect of within-wave variability. Mehta and West (2000) and Blozis and Cho (2008) focused on accounting for age heterogeneity at each measurement occasion. Mehta and West (2000) accomplished their study by using a simulated dataset with fixed time points ignoring age heterogeneity and comparing this to three models accounting for age heterogeneity. Blozis and Cho (2008) investigated two existing datasets, where they modeled age heterogeneity by using age at assessment for time coding and using fixed codes. Both studies do not create opportunities for examining factors of influence on the effect of within-wave variability. Therefore Aydin and colleagues (2014) used a Monte Carlo simulation study to create different scenarios, based on different distributions, ranges and sample sizes. In addition, Coulombe and colleagues (2015) also used a Monte Carlo study, but extended this

by also examining factors such as small sample sizes. They were able to compare both the time-structured and time-unstructured analyses of longitudinal data. These previous studies indicate that using a simulation study is an adequate method to examine within-wave variability.

In conclusion, previous studies suggest that variability in measurement occasions during a longitudinal study, assuming a linear growth model, influences results on some aspects (See Aydin et al., 2014; Blozis & Cho, 2008; Mehta & West, Singer & Willet, 2003). However standard error and mean slope estimates seem free of relevant bias. To the knowledge of the researcher there has not yet been a study investigating this within-wave variability for discrete-time survival analysis. The current research therefore fills this gap by aiming to extend the findings of Aydin and colleagues (2014) and Coulombe and colleagues (2015) for discrete-time survival analysis with a dichotomous outcome variable. The aim is to elucidate the effects of measurement interval and type of distribution, among other factors, on discrete-time survival analysis. It is expected that as in previous studies, the slope parameter and standard error remain free of substantial bias. A simulation study based on a six-period design will be executed, creating four different distributions for measurement variability: a truncated normal, uniform, truncated chi-square and a distribution with fixed measurement occasions.

Statistical model

The current study will be based on a few statistical assumptions, which will be explained in this section. First of all, discrete-time survival analysis will be explained more in-depth. When using discrete-time survival analysis in an experimental design, there are a few ways to describe the results. First, the population survival function can be used to describe the change that a random person will not experience the event in or before a measured time period. When a study starts the proportion of nonoccurrence is 1.00, which

will drop as the study progresses. Due to censoring or the type of event the survival function usually does not reach zero (Singer & Willet, 1991). For instance in the case of early treatment termination, not everyone will terminate the treatment early, resulting in a higher than zero survival function. The survival function can be written as (Singer & Willet, 2003):

$$S(T_{ij}) = \Pr[T_i > j]$$

In this formula T_i is specified as the time period in which individual *i* experiences the event. The measured time period is represented by *j*. The estimated survival function can be written as (Singer & Willet, 2003):

$$\hat{S}(t_j) = \frac{n \text{ who have not experienced the event at the end of time period } j}{n \text{ in the dataset}}$$

In addition to the survival function, results can be described by the population hazard function. This function provides information on the risk of a random person experiencing the event within a time period, given that this person has not yet experienced the event before this time period (Willet & Singer, 1993). This makes the hazard a conditional probability, since it depends on whether the person has already experienced the event. The group of people who are still at risk for experiencing an event are defined as the risk set (Singer & Willet, 2003). In contrast to the survivor function, the shape of the hazard function can take on various forms, starting high or low, remaining constant or showing peaks throughout the various time-periods, since risk only depends on the risk set and the people experiencing the event. The formula the hazard function looks like (Singer & Willet, 2003):

$$h(t_{ij}) = \Pr[T_i = j] \mid j\ [T_i \geq j]$$

Per time period the following formula can be used to estimate the hazard (Singer & Willet, 2003):

$$\hat{h}(t_j) = \frac{n\ events\ j}{n\ at\ risk\ j}$$

Furthermore, using the hazard function can be helpful to both identify the period in which individuals are most at risk to experience the event and to find out what the shape of risk over time is. For instance whether risk decreases, increases or remains stabile over time. In addition, by using the last mentioned function, the problems caused by censoring are eliminated. Everyone who is censored, is removed from the risk set. Therefore the hazard can be calculated for each period, independent of censoring. In addition the survivor function can be written as a function of the hazard (Singer & Willet, 2003):

$$\hat{S}(t_j) = \hat{S}(t_{j-1})\,[1 - \hat{h}(t_j)].$$

$\hat{S}(t_{j-1})$ represents the survival function for the time period before the time period in which the event occurs, whereas $\hat{h}(t_j)$ refers to the hazard function of the time period in which the event occurred. Using this formula enables researchers to also calculate the survivor function in the case of censoring (Singer & Willet, 2003).

Moreover, the mentioned hazard function, and also the survival function, have bounded scales. To transform the hazard into an unbounded scale of event occurrence over time, the complementary log-log link will be applied. This link is represented by the following formula (Singer & Willet, 2003):

$$\text{clog-log} = \log(-\log(1 - \text{probability}))$$

As can be seen in the formula, the clog-log is the 'logarithm of the negated logarithm of the probability of event nonoccurrence' (Singer & Willet, 2003, p. 421). The benefit of the clog-log link, in contrast to other links, is that the function more closely resembles a process that is actually measured by a continuous time scale. The current simulated data represents six time periods. However the event underlying these time periods is actually one that could be measured in a continuous time manner. The clog-log link is more accurate in representing this fact, because it has a built-in proportional hazards assumption which shows similarities to models for continuous-time survival analysis (Singer & Willet, 2003). The proportional

hazards assumption, states how all hazard functions of different conditions are magnifications or diminutions of each other, therefore creating parallel functions (Willet & Singer, 1993).

All mentioned functions can be provided for both of the two conditions that are being examined in an experimental study. The clog-log function for the one condition, usually the control condition, is the baseline hazard function. The other function of the treatment condition is represented by the same shape as the baseline function, but vertically shifted. Both hazard functions, for the treatment and control condition, can be described by the following formula (Willet & Singer, 1993):

$$\text{clog-log } h(t_j) = [\alpha_1 D_1 + \alpha_2 D_{2\ +\ \ldots\ +}\ \alpha_j Dj] \ + \beta\text{treatment}$$

The first value, $\alpha_j Dj$, describes the intercept, which is different per time period. Therefore dummy variables are created per time period, which are multiplied by the specific intercepts (α_j) per time period . The latter part of the formula, consists of a slope parameter (β) and a variable indicating whether there is a treatment condition or not. In the case of the treatment condition 1 is filled in, whereas for the control condition 0 is filled in. In the current study both the treatment effect slope parameter and the intercepts for the dummy variables will be estimated via a simulation study.

Furthermore, the survival functions that will be examined in the simulation study are based on an underlying continuous Weibull distribution. The current simulation study focuses on discrete-time survival analysis, measuring an underlying continuous process. The formulas that stem from this Weibull distribution are the following survival function and hazard function (Jóźwiak, 2013):

$$S(t) = \mathrm{e}^{-\lambda t^{\tau}}$$

$$h(t) = \ \lambda\tau t^{\tau-1}$$

In the current study λ will be constituted by $-\log(1-\omega)$ with $\omega \in [0,1]$, resulting in the following formulas respectively for survival function, hazard function and clog-log hazard function (Jóźwiak, 2013):

$$S(t) = (1-\omega)^{j^{\tau}}$$

$$h(t)=\frac{(s_t - s_{t-1})}{s_t}$$

$$clog - log(t) = LN(-1 * LN(1-ht)$$

These formulas are useful for simulating data in different scenarios on a continuous scale, which for the purpose of this study will then be presented on a discrete-time scale.

Lastly, the final statistical information regards the use of a person-period format in the current study. A person-period format consists of several lines per person, describing each time period of being at risk in a different row (Singer & Willet, 2003). In the current study, in order to create useful results, a person-period converter is used, to create data in the required format.

Method

To make comparisons between various distributions possible, the use of multiple scenarios was required. Therefore the current study used simulated data created by the program R. The latest version of R 3.2.0 was used (R Development Core Team, 2008). In total 270 scenarios were created based on averages of 2000 datasets with a sample of 1000 per dataset. All scripts can be retrieved by contacting the researchers of the current study.

Factors in the Simulation Study

In the current study simulations of a six-period discrete-time survival analysis with one predictor were executed. First of all, four factors, ω, τ, β and width of the measurement interval were varied to examine the influences on discrete-time survival analysis. In addition, width varied per type of distribution of measurement occasions, which could be a fixed time

points, a truncated normal, a truncated chi-square and a uniform distribution. Next, these factors will be explained in the mentioned order.

The parameters τ, ω and β are factors that determine what the survival function will look like. First, $\tau \in [0, +\infty]$ describes the shape of the survival and hazard function. For τ >1 the risk of event occurrence is lowest at the beginning of the study, which leads to a high survival function at first. The opposite is true for $\tau < 1$, where the risk of event occurrence is highest at the beginning. Furthermore, for $\tau = 1$ hazard is constant throughout the study, resulting in a horizontal hazard function and linear survival function. To create diversity in simulated scenarios, the current study examined 0.5, 1 and 2 for τ. The corresponding clog-log hazard functions are shown in figure 1 in appendix A.

Next, the ω value represents the proportion of the people in the baseline condition who will experience the event by the end of the final time period. When choosing higher values this moves the clog-log hazard upwards and the opposite is true for lower values for ω. Therefore high, low and medium values were chosen for ω to create diversity: 0.25, 0.5 and 0.75. This results in identically shaped, but vertically shifted functions as is shown in figure 2 in appendix A.

Furthermore, β describes the difference between the two clog-log hazards of the two conditions: the treatment effect. β was set to -0.125, -0.25 and -0.5, based on what seem plausible values for treatment effects. The diverse corresponding clog-log hazard functions are shown in figure 3 in appendix A.

Finally, width of the measurement interval is the main factor in the current study. The measurement interval is defined as the range of measurement occasions centering around the actual aimed measurement occasion. This actual aimed measurement occasion is referred to as a fixed time point. It should be emphasized that the concept measurement interval is different from the term time interval or period, which refers to the interval in which the event may or

may not have taken place. When applying width in the current study, the value of width represents the measurement interval, which is a proportion of the time period. For instance when width is set to 0.5 when the time period is 1 month, the measurement interval is two weeks, of which one week before and one week after the actual aimed measurement occasion.

First, width was set to 0 to investigate a distribution with fixed time points. Hereafter the width of the measurement interval was increased to 0.25 $[-0.125, +0{,}125]$, 0.5 $[-0.25, +0{,}25]$ and 1 $[-0.5, +0{,}5]$ for the other three distributions: truncated normal, uniform and truncated chi-square. During the uniform distribution it is assumed that all participants are measured equally distributed around an aimed measurement occasion. The truncated normal distribution describes a distribution in which most people are measured on or around the actual aimed measurement occasions, described by a middle peak, and some people are measured further from the actual measurement occasion resulting in tails on either side of the peak. The normal distribution is cut off at -2 and 2, hence the name truncated. Finally, the chi-square distribution describes a measurement interval starting with a peak and ending in a tail. Often during a wave, participants tend to fill in the questionnaire at the beginning and a few follow creating the tail. The chi-square distribution is also truncated.

Furthermore, in the simulation study a few different outcome variables were examined. Per simulation the model for discrete-time survival analysis was estimated 2000 times; every parameter was estimated 2000 times, resulting in an average per scenario. Finally, also the convergence rate is provided, indicating the amount of the 2000 datasets that reached estimation within a maximum of 25 iterations.

Criteria for evaluation

Analysis consisted of calculating the percentage treatment effect bias and the percentage bias in standard error of treatment effect. This was done via the following formulas based on calculations provided by Hoogland and Boomsma (1998):

$$\text{percentage bias treatment effect} = \frac{\overline{estimated\ \beta} - \beta}{\beta} * 100\%$$

$$\text{percentage bias standard error} = \frac{\overline{SE} - SD}{SD} * 100\%$$

Hereafter tables presenting the data per width and per type of distribution were created to describe the effect of the factors on discrete-time survival analysis.

Next, the established biases were be evaluated. It should be mentioned how there is only limited agreement concerning the criteria for biases, since these are often characterized as subjective (Boomsma & Hoogland, 2001). Based on the same criteria Aydin and colleagues (2014) used for a similar study, the percentage parameter biases were considered acceptable when approximately 5%. For the standard error bias a larger, 10% was considered reasonable (Hoogland & Boomsma, 1998).

Results

Via simulation 270 different scenarios were created and analyzed via R and SPSS (See appendix C for all data). The convergence rate throughout all scenarios was 2000, which means that for all simulated scenarios all datasets were estimated within 25 iterations. Overall, only one case showed unacceptable bias in treatment effect (see table 1 in appendix B). Furthermore, standard error bias was acceptable for all scenarios (see table 2 in appendix B). First treatment effect bias will be more thoroughly discussed. Hereafter, the bias of the standard error of treatment effect will be described.

Percentage treatment effect bias

Overall, there was no critical bias in treatment effect, except for one scenario. This scenario is described by a truncated normal distribution where width=1, ω=0.25, τ=0.5 and β=-0.125. Although bias in other cases does not exceed the criteria of 5%, it does vary over the various scenarios. This diversity in bias of treatment effect is partly caused by the different values for ω, τ, β, width of the measurement interval and type of distribution

involved. The descriptive statistics for percentage treatment effect bias separated by these factors can be found in table 3 in appendix B.

First, for lower values of ω the bias is more widespread and covers a bigger interval $[-3.14;\ 5.16]$ than for the highest value of ω $[-2.56;\ 4.16]$. The same effect is established for β: The lowest value -0.125 shows the largest range in bias $[-3.97;\ 5.61]$, whereas the highest value of -0.5 has the smallest bias interval $[-1.18; 2.19]$. In contrast, this effect is opposite for τ, where the range of percentage treatment effect bias seems to increase when τ increases from 0.5 $[-2.99; 4.21]$ to the largest value of τ $[-3.97; 5.61]$.

The trends for ω, τ and β were further examined by separating between width and type of distribution. First for ω, it is remarkable how for the chi-square distribution more bias is negative when width is set to 1. All other bias from the different types of distributions seems to mainly overlap. Next, for β is seems that for different types of distributions and various widths the same trend remains visible. Only one remarkable small effect is seen for β -0.125: When width of the measurement interval increases, the bias of the chi-square distribution becomes mainly negative, whereas the uniform and truncated normal distribution show mainly positive, high bias. Lastly, the effect of τ varies in many ways between the different types of distributions and per value of width; No clear pattern is presented in the data.

Furthermore, specifically for width, percentage treatment effect bias seems to slightly increase along with the measurement interval. Highest bias is found for a measurement interval of 1. This is graphically represented in figure 4 in appendix B. Most scenarios do remain clustered around 0 and the increase in range of bias is caused mainly by a few higher values, which are described by low ω and β values.

When width is 0, indicating fixed time points, there are two outliers, 4.21,and 3.31, whereas the other scenarios all lie within an $[-1.03; 1.28]$ interval. Both outliers are described by small values of ω and β, whereas τ differs: both 0.5 and 1. Given the previously

described effects of ω and β on bias, this seems a logical cause of higher bias in these two cases. This is also in line with the previous statement that mainly a small ω and β cause more percentage bias. Surprisingly, these two outliers exceed the highest bias of width 0.25 and width 0.5, which seems contrasting to previous statements about the effect of width. However these two cases seem outliers and the overall upwards trend for width remains supported.

Moreover, separating width per type of distribution indicates how for a smaller width, the truncated chi-square distribution shows mainly positive bias, whereas the uniform and normal distribution shows large negative bias. For the highest value of width this effect is reversed, where the chi-square shows large negative bias, the truncated normal and uniform distribution show larger positive bias These effects can be considered small and can be ignored.

Finally, the type of distribution only explains very small differences in percentage bias. The truncated chi-square, uniform and truncated normal distributions show comparable and overlapping bias for all widths. Only small differences between scenario's, less than 1%, are shown. Therefore these can be considered irrelevant. Overall all mentioned effects of factors on percentage bias in treatment effect can be considered negligible.

Percentage standard error bias

The percentage standard error bias yields similar results as the previously discussed percentage treatment effect bias. Overall, all bias in standard error is acceptable according to the criteria by Hoogland and Boomsma (1998). Bias does not exceed the criteria of 10%. As the bias in treatment effect, the bias in standard error also varies across scenarios. The differences are less clear than for treatment effect bias. However there are small trends visible for various values of ω, τ, β, width and for the various distributions as is shown in figure 5 and table 4 in appendix B.

First of all, a larger ω increases the range in percentage standard error bias from $[-3.62;\ 3.21]$ to $[-4.13;\ 3.92]$. Furthermore, for τ there seems no distinct pattern, where the range first increases for τ is 1 and then decreases for τ is 2. Also range of percentage bias does not seem to differ more than 1% per τ value. For β there is also no distinct trend visible.

When width of the measurement interval and type of distribution are taken into account some patterns present themselves in the data. For ω the mentioned overall effect of increasing range percentage treatment effect bias only holds for the chi-square distribution. For the chi-square distribution the range of percentage standard error bias for ω 0.75 $[-3.09; 3.92]$ is larger than the range for ω 0.25 $[-3.31; 2.09]$ and ω 0.5 $[-2.51; 1.98]$. In contrast, the truncated normal distribution shows least bias for the highest value of ω. Furthermore, the other two types of distributions, uniform and fixed time points, show minimal differences in bias per value of ω. All mentioned differences between bias values are very small, therefore it is only a minimal trend.

In contrast, for τ bias of all four types of distribution seem to overlap mainly. When only examining the truncated normal distribution, the range in bias does seem to slightly decrease when τ increases. Moreover, as width of the measurement interval is also taken into account, τ is 1 and width is 0.5 indicate the most widespread percentage standard error bias, mainly caused by high bias for the uniform distribution. Bias for width 0.5 and 1 seems larger than for a smaller width but this seems to overlap mostly for different types of distributions. As mentioned for ω, differences for τ do remain small, less than 1%, and overall no bias exceeds the limit of 10%.

Moreover, for β, separating per type of distribution and for width of the measurement interval, indicates a pattern for the truncated normal distribution. When width is set to 0.25 for the truncated normal distribution percentage bias decreases in range and becomes more

negative as β approaches -0,5. When width further increases this trend fades. The other distributions overlap and do not show a trend. Differences per β value, even for the described pattern, remain between 1 and 2%, and can therefore be ignored.

When examining width and type of distribution separate from ω, τ and β, there is still no substantial influence on the percentage standard error bias. Range remains in all cases of width around 7%, only minimally shifting vertically, indicating that negative bias increases minimally along with width. As in the other cases, there remains overlap in the majority of the values of the different distributions. Overall there are only minimal trends and differences between the percentage standard error bias of the various scenarios.

Discussion and Conclusion

The current study examined how variability in measurement occasions influences discrete-time survival analysis. The results are in line with the hypothesized acceptable bias in standard error and treatment effect. Both percentage standard error bias and percentage treatment effect bias yielded minimal bias, only exceeding the criteria for percentage treatment effect bias once.

Overall width of the measurement interval and type of distribution do not seem important factors in determining standard error and treatment effect bias. For treatment effect bias a larger measurement interval minimally increases bias. In addition, for standard error bias only negative bias seems to increase, where positive bias decreases. However the differences that were found, were so minimal they can possibly be ignored. This is an interesting finding, indicating that it may not be of importance whether you are measured almost half the time interval before or after the aimed measurement occasion. Furthermore the type of distribution seems to also have minimal influence, mainly showing overlap in bias for different types of distributions. Therefore it does not seem of importance how the measurement occasions are distributed throughout the measurement interval. In addition to

width and distribution, also ω and β show minimal trends for bias in treatment effect. The range of treatment effect bias seems to decrease slightly for higher values of ω and β, centering more closely around 0.

The one scenario that exceeded the criteria of 5% fits into the mentioned trends, being described by low values of β and ω and a large measurement interval. This indicates that studies may be slightly more at risk for bias when examining events for which the treatment effect is small and only a small proportion of the participants experience However, according to the criteria, bias does not change relevantly in other scenarios when the underlying survival function changes.

When comparing the current results to previous findings there are similarities between results. Where both Aydin and colleagues (2014) and Coulombe and colleagues (2015) found minimally biased mean slopes varying slightly across scenarios, the current study found similar results for treatment effect. This indicates that also for a discrete time survival model bias remains acceptable. Furthermore, the current results on standard error, also partly replicate the results of Aydin et al. (2014) for discrete-time survival analysis. There was only a minimal, negligible influence of width of the measurement interval on standard error bias in discrete-time survival analysis.

In addition, the current study presented some contrasting findings. Where Coulombe et al. (2015) describe bias below 2%, the current study found bias up to 4% in some scenarios. This may be caused by differences between the underlying models or choices made during the simulation study and analysis. Coulombe et al. (2015) used 5 times more iterations per scenario, 10.000 instead of 2000. Also they only examined width 0.4 and 1.0. The current study examined multiple widths and instead of only examining skewed or non-skewed, also focused on the truncated normal, truncated chi-square and uniform distribution. Moreover, it should be strongly emphasized how previous studies focused on linear growth models,

whereas the current study assumes a discrete-time survival model. This changes the underlying assumptions of the simulated model and may therefore have caused large differences in bias between the studies.

Concluding from the current results, there is no relevant effect of ignoring variability and varying ω, τ, β and width on discrete time survival analysis. This leads to some practical implications, since this indicates that longitudinal analysis can still be accurate when participants are measured half the time period later than the aimed measurement occasion. Time is of great concern in longitudinal analysis and making decisions about time concerns money and statistical benefits (Singer & Willet, 2003). As Singer and Willet (2003) indicate: "Time is a fundamental predictor in every study of change; it must be measured reliably and validly in a sensible metric" (p. 10, Singer & Willet, 2003). When there is variability in measurement occasions in some cases specifying time differently will solve problems (Singer & Willet, 2003). In the case of discrete-time survival analysis this remains difficult. However, the current study indicates how for discrete-time survival analysis, within-wave variability does not cause bias in treatment effect and standard error. This relieves some of the pressure on longitudinal researchers when it comes to within-wave variability. For instance, it enables researchers to conduct larger studies by themselves. Given that it may take time for a researcher to gather data from one person, the researcher can more easily spread the data gathering and thereby increasing the measurement interval without introducing critical bias in the results. The current results therefore also present researchers with new possibilities in their studies.

However there are some limitations of the current study. First of all, other parameters than the ones in the current study can be influenced by variability in measurement occasions, for instance power and model fit could be further investigated. Furthermore, more factors, for instance sample size, can be involved in determining the effect, maybe in interaction with

width of the measurement interval. The current study focused on ω, τ and β. However factors such as sample size may for example influence the convergence (Coulombe et al., 2015), which in turn may influence the bias.

However, in addition to these limitations there are also several strong characteristics of this study. First of all, the current study seems to be the only one to date focusing on within-wave variability effects for discrete-time survival analysis. This fills a large gap in existing research. Second, the current study also examined a wide array of different scenarios of discrete time survival analysis using an underlying Weibull distribution, which makes the current results widely applicable and representative of various types of events. In the current study both events for which risk is stable or high or low at the beginning is taken into account. It may still be interesting for future studies to investigate varying risk over time. Lastly, the simulation study only provided converged results, indicating that the model was estimated within the maximum of 25 iterations. This indicates complete solutions on which the results of the current study are based.

In conclusion, the results suggest that variability in measurement occasions does not cause substantial bias in standard error or treatment effect estimates. However due to the mentioned limitations it is recommendable to per study examine whether large measurement intervals are considered acceptable. Future studies should examine more conditions to determine whether the standard of fixed measurement occasions in longitudinal studies may indeed be omitted.

References

Aydin, B., Leite, W. L., & Algina, J. (2014). The consequences of ignoring variability in measurement occasions within data collection waves in latent growth models. *Multivariate Behavioral Research, 49,* 149-160. doi:10.1080/00273171.2014.887901

Blozis, S. A., & Cho, Y. I. (2008). Coding and centering of time in latent curve models in the presene of interindividual time heterogeneity. *Structural Equation Modeing: A Multidisciplinairy Journal, 15,* 413-433. doi:10.1080/10705510802154299

Boomsma, A., & Hoogland, J. J. (2001). The robustness of LISREL modeling revisited. *Structural equation models: Present and future. A Festschrift in honor of Karl Jöreskog,* 139-168. retrieved from http://www.gmw.rug.nl.proxy.library.uu.nl/~boomsma/ssi.pdf

Corning, A. F., & Malofeeva, E. V. (2004). The application of survival analysis to the study of psychotherapy termination. *Journal of Counseling Psychology, 51,* 354-367. doi:10.1037/0022-0167.51.3.354

Coulombe, P., & Selig, J. P., & Delaney, H. D. (2015). Ignoring individual differences in times of assessment in growth curve modeling. *International Journal of Behavioral Development, volume number,* 1-11. doi:10.1177/0165025415577684'

Hoogland, J. J., & Boomsma, A. (1998). Robustness studies in covariance structure modeling: An overview and meta-analysis. *Sociological Methods & Research, 26,* 329-367. doi:10.1177/0049124198026003003

Jóźwiak, K (2013). *Improving statistical power in studies on event occurrence by using an optimal design.* (Unpublished doctoral thesis). University of Utrecht, Utrecht, The Netherlands.

Lauterbach, E., Felber, W., Müller-Oerlinghausen, B., Ahrens, B., Bronisch, T., Meyer, T., ... Hohagen, F. (2008). Adjunctive lithium treatment in the prevention of suicidal behaviour in depressive disorders: A randomised, placebo-controlled, 1-year trial. *Acta Psychiatrica Scandinavia, 118,* 469-479. doi:10.1111/j.1600-0447.2008.01266.x

Le Roux, I. M., Le Roux, K., Comoluda, W. S., Greco, E. M., Desmond, K. A., Mbewu, N., Rotheram-Borus, M. J. (2010). Home visits by neighborhoud Mentor Mothers provide timely recovery from childhood malnutrition in South Africa: Results from a randomized controlled trial. *Nutrition Journal, 9*(56), 1-10. Retreived from http:/www.nutritionj.com/content/9/1/56

Mehta, P. D., & West, S. G. Putting the individual back into individual growth curves. *Psychological methods, 5,* 23-43. doi:10.1037/10.1037//1082-989X.5.1.23/1082-989X.5.1.23

R Development Core Team (2008). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Singer, J. D., & Willet, J. B. (1991a). Modeling the days of our lives: Using survival analysis when designing and analyzing longitudinal studies of duration and the timing of events. *Psychological Bulletin, 110,* 268-290. doi:10.1037/0033-2909.110.2.268

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

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. Oxford, United Kingdom: Oxford university press.

Sterba, S. K. (2014). Fitting nonlineair latent growth curve models with individually varying time points. *Structural Equation Modeling: A Multidisciplinary Journal, 21,* 630-647. doi:10.1080/10705511.2014.919828

Storr, C. L., Ialongo, N. S., Kellam, S. G., & Anthony, J. C. (2002). A randomized controlled trial of two primary school intervention strategies to prevent early onset tobacco smoking. *Drug and Alcohol Dependence, 66,* 51-60. doi:10.1016/S0376-8716(01)00184-3

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

Willet, J. B., & Singer, J. D. (1993). Investigating onset, cessation, relapse and recovery Why you should, and how you can, use discrete-time survival analysis to examine event occurence. *Journal of Consulting and Clinical Psychology, 61,* 952-965. doi:10.1037/0022-006X.61.6.952

William R.. Shadish, Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Hampshire, United Kingdom: Wadsworth Cengage learning.

Appendix A: cloglog hazard functions of ω, τ and β

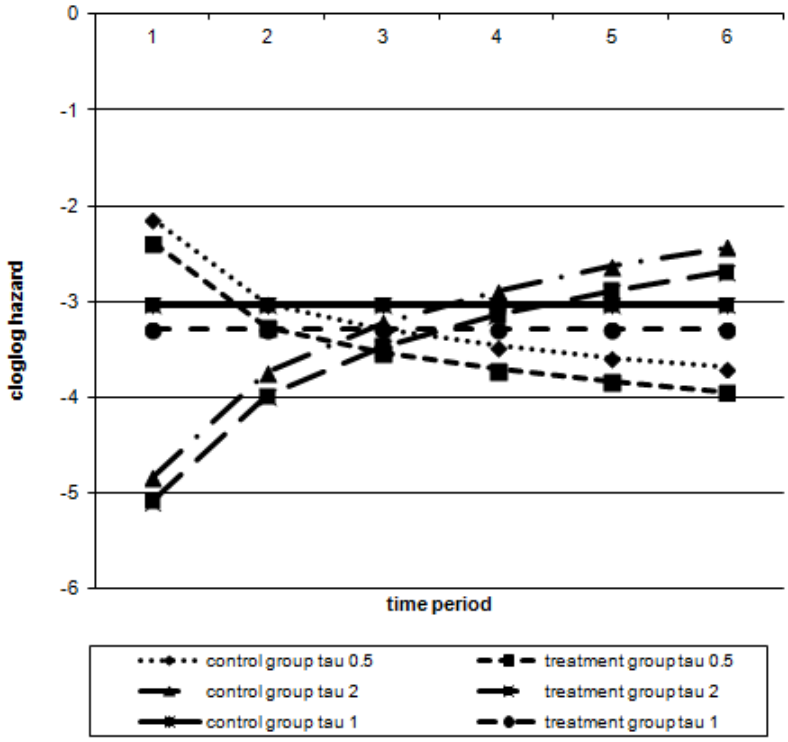

Figure 1. Distribution of clog-log hazard for treatment and control group against time period, for three different τ values. Also ω is 0.25 and β is -0.25.

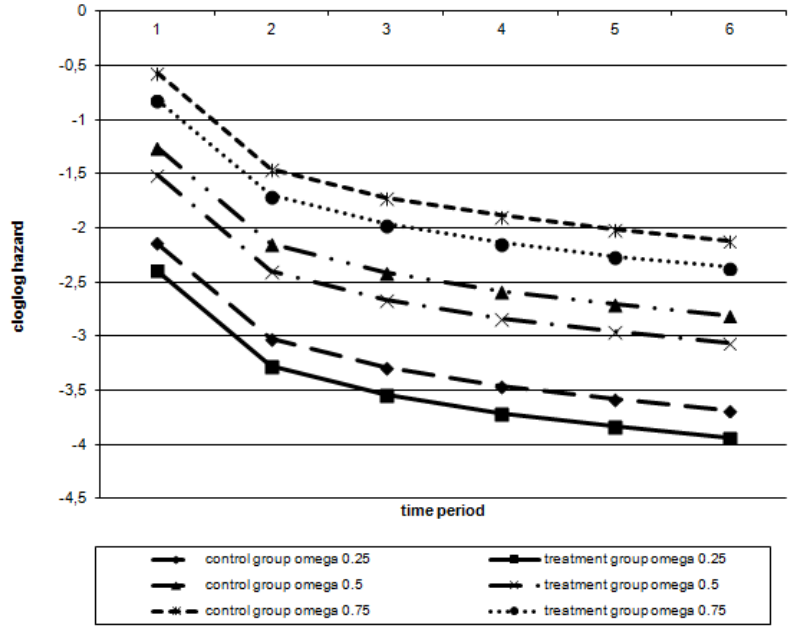

Figure 2. Distribution of clog-log hazard of the treatment and control group against time period, for three different values of ω. Also τ is 0.5 and ω is 0.25.

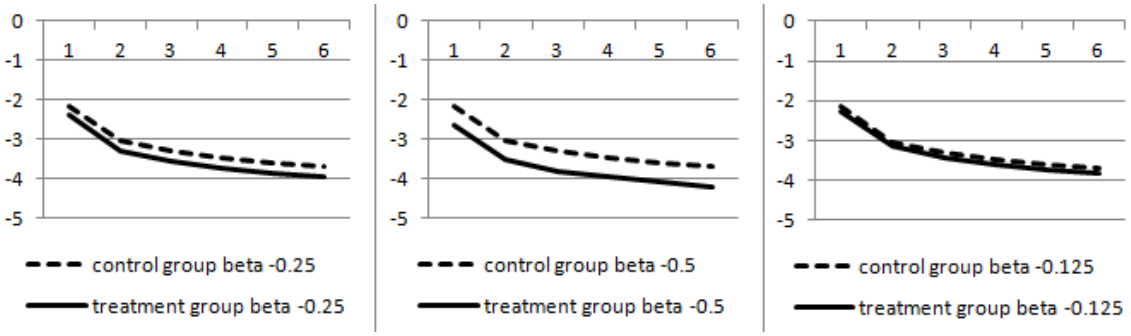

Figure 3. Three distributions of clog-log hazard for treatment and control condition against time period, for three different β values. Also τ is 0.5 and ω is 0.25.

Appendix B: Simulation data

Table 1.

Bias in percentage of treatment effect ($\bar{\beta}$) for a fixed time points, truncated normal, uniform and truncated chi-square distribution (width=0.25, width=0.5, width=1) per τ, ω, β value. Substantial bias exceeding the criteria of 5% is presented bold faced.

Fixed 0			Normal 0.25	Normal 0.5	Normal 1	Uniform 0.25	Uniform 0.5	Uniform 1	Chi 0.25	Chi 0.5	Chi 1
ω	**β**	**bias $\bar{\beta}$**	**bias $\bar{\beta}$**	**bias $\bar{\beta}$**	**bias $\bar{\beta}$**	**bias $\bar{\beta}$**	**bias $\bar{\beta}$**	**bias $\bar{\beta}$**	**bias $\bar{\beta}$**	**bias $\bar{\beta}$**	**bias $\bar{\beta}$**
0.5	0.25	-0.125	4.20776	0.62992	-1.12792	-2.9892	-2.31464	-2.012	3.55184	-0.97608	-0.92488	0.66056
0.5	0.25	-0.25	-0.93844	0.39684	2.50044	-0.30032	-0.25584	-2.35708	0.69816	-0.45508	0.24732	0.66988
0.5	0.25	-0.5	1.28446	0.41872	-0.4782	0.42132	0.56742	0.19938	0.85632	0.91352	-0.2703	1.02028
0.5	0.5	-0.125	1.10288	1.43272	-2.52576	-0.04648	-0.52152	3.63984	0.3156	1.55008	-1.18384	-1.46288
0.5	0.5	-0.25	0.58864	1.3198	-0.29368	-1.2824	-0.09052	-0.51804	-0.264	-0.23624	0.44568	-0.99632
0.5	0.5	-0.5	-0.08746	-0.45136	0.73278	0.29788	0.44486	0.59818	0.36776	-0.06896	0.3267	0.68868
0.5	0.75	-0.125	0.3036	0.53536	-0.04112	1.35744	-0.18896	-0.99152	-0.68912	1.49496	1.98216	2.27352
0.5	0.75	-0.25	0.2834	0.33416	1.40724	0.68244	0.62096	0.35392	-0.03824	0.22848	-0.86816	0.10632
0.5	0.75	-0.5	0.0421	0.70988	-1.01906	0.06758	0.21262	-0.21512	-0.1343	0.78242	-0.11712	-0.19224
1	0.75	-0.125	-0.79608	1.79288	1.7856	1.68072	-0.81784	1.30568	0.86896	4.1632	-0.2288	1.1372
1	0.75	-0.25	0.28936	-0.68828	1.45168	-0.2154	-0.35056	-0.23072	0.23116	0.4332	-0.89964	0.31396
1	0.75	-0.5	-0.07058	0.24248	0.52028	0.08822	-0.6485	0.44092	0.12796	-0.22504	0.47286	0.00882
1	0.5	-0.125	-0.08016	-0.43648	1.15216	3.28552	-2.38512	-0.78096	0.93144	1.48848	2.15096	-2.82656
1	0.5	-0.25	-0.51996	0.4912	0.49732	1.30976	-0.65676	1.30588	0.55288	0.19336	-0.44004	1.47832
1	0.5	-0.5	-0.14008	0.14544	-0.7915	-0.32292	0.59078	0.9149	0.0304	0.09382	0.47242	0.38418
1	0.25	-0.125	0.672	-0.37808	1.30256	4.79456	0.16112	1.8036	4.97056	3.79008	0.15248	-0.83824
1	0.25	-0.25	1.24984	2.8792	-0.33524	0.04556	1.33716	2.36316	0.93776	0.3064	1.06676	-1.39564

1	0.25	-0.5	-0.23986	-0.15686	1.46034	2.19308	1.4781	1.4931	0.75188	0.6373	0.73476	-1.18296
2	0.25	-0.125	3.30568	0.42672	2.82776	**5.6144**	3.2392	2.43024	1.25504	3.70568	-1.98416	-3.14152
2	0.25	-0.25	-0.02008	0.63736	-0.5994	-0.7638	0.04052	-0.08364	1.86292	-0.82976	2.5644	0.60012
2	0.25	-0.5	0.6398	1.24384	0.92832	0.54608	0.37936	0.51656	-0.0942	-0.3409	0.91016	0.13998
2	0.5	-0.125	0.51368	0.36016	0.41784	1.62744	0.77192	-3.96744	1.3644	0.7812	-1.17048	-0.37232
2	0.5	-0.25	-1.39608	1.461	0.65744	1.31388	-1.43932	-1.23748	-1.27244	0.99284	-0.8792	0.42904
2	0.5	-0.5	0.35828	0.09192	0.44436	-0.19868	-0.21406	0.07434	0.2659	0.30578	-0.61934	0.61334
2	0.75	-0.125	-1.02976	0.44192	-0.7976	-0.99128	-0.45624	-2.56248	2.1616	0.54704	-0.46024	1.65304
2	0.75	-0.25	-0.315	-1.65652	-1.1264	0.07584	0.72424	0.86396	0.68852	0.0794	0.07208	0.16008
2	0.75	-0.5	0.09946	0.41814	0.62828	0.2932	1.0409	0.07378	-0.3242	0.62492	0.5735	-0.79254

Table 2.

Bias in percentage of standard error($\overline{SE}$) for a fixed time points, truncated normal, uniform and truncated chi-square distribution (width=0.25, width=0.5, width=1) per τ, ω, β value.

		Fixed 0		Normal 0.25	Normal 0.5	Normal 1	Uniform 0.25	Uniform 0.5	Uniform 1	Chi 0.25	Chi 0.5	Chi 1
τ	**ω**	**β**	**bias $\overline{SE}$**	**bias $\overline{SE}$**	**bias $\overline{SE}$**	**bias $\overline{SE}$**	**bias $\overline{SE}$**	**bias $\overline{SE}$**	**bias $\overline{SE}$**	**bias $\overline{SE}$**	**bias $\overline{SE}$**	**bias $\overline{SE}$**
0.5	0.25	-0.125	1.20336	1.69955	0.76938	-2.24364	1.49938	0.03224	-0.459	-2.49361	-3.31465	-2.6907
0.5	0.25	-0.25	-1.83276	0.554	3.20846	0.91808	-2.71333	1.53134	0.74748	1.92031	0.43988	-0.5544
0.5	0.25	-0.5	-1.44988	-3.41683	0.71781	1.55521	-2.24317	1.2473	-0.21777	-2.21639	0.67175	-0.61025
0.5	0.5	-0.125	0.53592	-2.9837	-1.47516	2.41045	-1.85138	-2.10539	-0.48634	-0.31302	-1.9505	-2.51355
0.5	0.5	-0.25	-0.50718	0.05013	-0.32521	-0.45621	1.85758	-1.41265	-0.1566	-0.98875	1.51527	0.46458
0.5	0.5	-0.5	0.47387	0.15279	-0.65392	-3.62144	1.83655	0.78752	-1.90024	1.29372	-0.02441	-1.11945
0.5	0.75	-0.125	-0.33503	0.92938	0.35501	-0.08573	1.50068	-0.47179	1.06434	-2.39092	-0.1206	-1.18215
0.5	0.75	-0.25	0.65754	0.06527	0.51825	1.78313	0.97613	2.87908	-4.13035	1.06051	-0.48454	0.55238
0.5	0.75	-0.5	2.461	-0.79945	-0.24411	-0.21424	1.16524	-0.46968	-1.53472	0.74855	-0.29657	-3.09125
1	0.75	-0.125	-0.92588	0.93748	0.32262	-0.05743	0.5364	1.64945	1.37149	-2.58992	0.57536	-0.68326
1	0.75	-0.25	-0.45589	0.13086	0.35517	-0.90352	1.73183	-2.2606	1.28785	3.91659	1.93806	-0.07938
1	0.75	-0.5	-1.49591	0.05744	-2.83689	0.32392	2.30649	-2.37014	-1.69328	-0.88764	-0.47678	-1.2881
1	0.5	-0.125	0.44694	2.29894	-0.50798	-2.61749	-0.2459	-3.70204	1.19127	0.47977	-0.02357	-1.56501
1	0.5	-0.25	4.08022	-1.4707	-0.33415	-1.38876	-0.35729	3.77689	-1.90342	1.7125	-1.34136	1.97852
1	0.5	-0.5	1.77322	0.46831	2.18859	-2.28205	-1.51917	0.71022	-0.48159	-1.2607	-0.09655	-0.0849
1	0.25	-0.125	-1.56365	-1.9479	1.93423	0.60543	0.90691	0.26755	1.83683	-1.90196	-1.33315	-2.46505
1	0.25	-0.25	-2.7533	-0.96613	2.40023	-0.65401	1.10541	0.01149	1.27992	0.71459	-0.74657	1.03622
1	0.25	-0.5	-1.54397	-1.85695	0.1637	-2.33616	0.35818	-0.79422	-1.18425	2.09188	0.26114	-1.04899
2	0.25	-0.125	1.71041	2.08635	-1.17449	0.97709	2.19396	-3.623	-0.5264	1.552	1.26773	-0.65536
2	0.25	-0.25	-1.6687	-1.40267	0.34058	0.71323	1.18741	0.38762	-3.31344	-2.01671	1.70853	0.53776
2	0.25	-0.5	-0.00925	0.04628	1.06844	1.96374	-0.35772	-2.31876	-0.44552	-1.81129	-2.34096	1.09732
2	0.5	-0.125	1.14558	-1.56699	-1.42861	0.43473	-1.13843	0.609	-2.55662	0.96162	-2.16252	-1.64142

2	0.5	-0.25	-2.34797	-2.05774	0.30993	-2.39245	2.45067	-1.71304	0.13941	-1.51073	1.85545	-1.81941
2	0.5	-0.5	1.55724	-1.61445	0.07133	-0.54857	-0.65039	-1.11599	-1.77495	-0.90528	-0.82433	0.29796
2	0.75	-0.125	2.37872	-0.91044	-1.15708	-0.52498	-2.13694	2.72316	2.53724	1.28459	-0.23288	1.28854
2	0.75	-0.25	-1.44013	-1.03316	-1.26052	-1.40035	-2.05205	-1.35492	1.28267	0.87134	2.68817	2.82603
2	0.75	-0.5	-2.2088	2.28902	0.1537	-0.30462	-0.01656	-1.66634	-2.01748	-0.80344	-1.18217	1.18049

Table 3
The descriptive statistics of percentage treatment effect bias for the different values of ω, τ, β, width and for all four different examined distributions.

	min	max	range	Mean	SD
ω 0.25	-3.14	5.16	8.76	0.69	1.67
ω 0.5	3.97	3.64	7.61	0.13	1.15
ω 0.75	-2.56	4.16	6.73	0.25	0.95
τ 0.5	-2.99	4.21	7.20	0.19	1.19
τ 1	-2.83	4.97	7.80	0.58	1.32
τ 2	-3.97	5.61	9.58	0.30	1.38
β -0.125	3.97	5.61	9.58	0.59	1.95
β -0.25	-2.36	2.88	5.24	0.18	0.98
β -0.5	-1.18	2.19	3.38	0.29	0.57
width 0.0	-1.40	4.21	5.60	0.34	1.19
width 0.25	-2.39	4.16	6.55	0.42	1.14
width 0.5	-3.97	3.64	7.61	0.19	1.32
width 1.00	-3.14	5.61	8.76	0.47	1.48
fixed time points	-1.40	4.21	5.60	0.34	1.19
normal	8.60	-2.99	5.61	0.50	1.33
uniform	7.90	-4.13	3.78	0.30	1.39
chi-square	7.30	-3.14	4.16	0.26	1.24

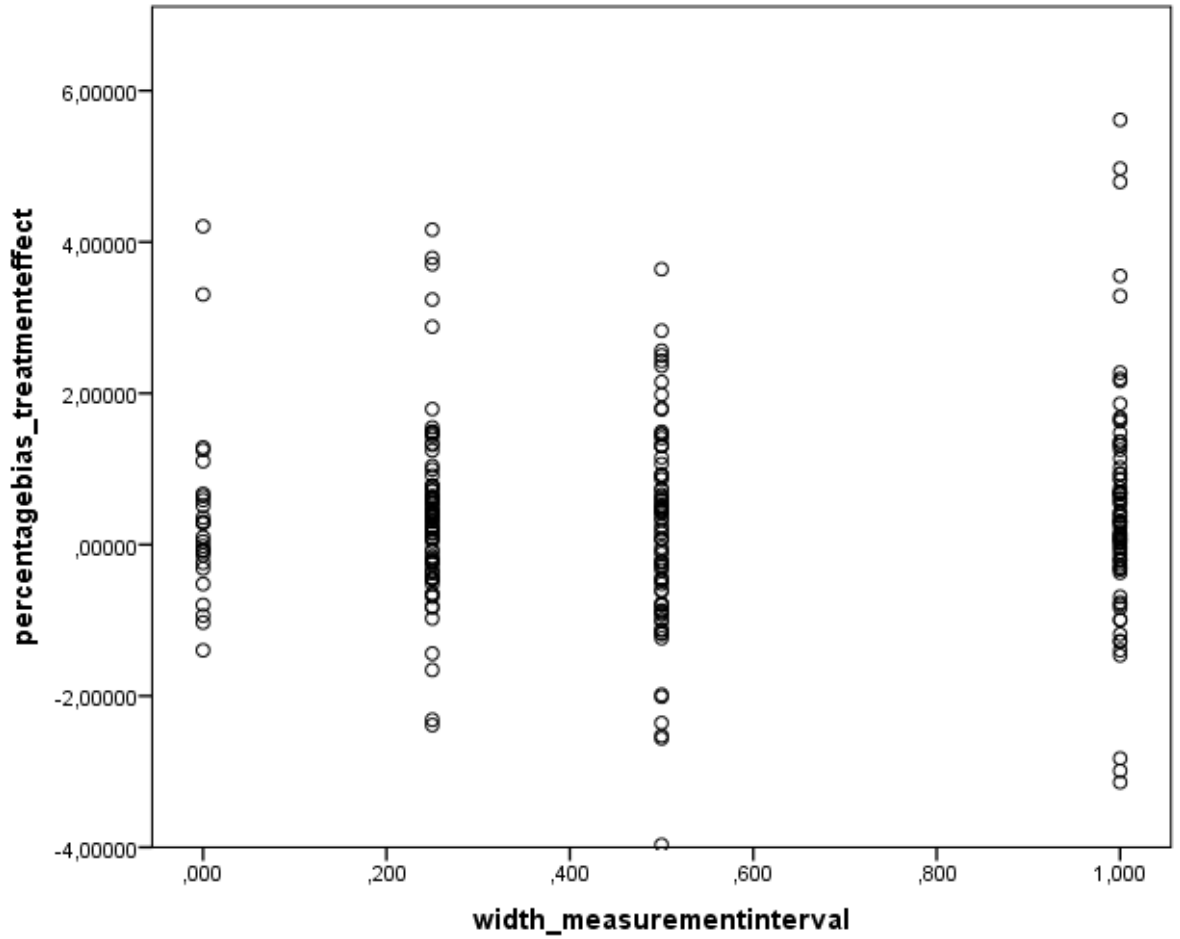

Figure 4. The distribution of percentage bias in treatment effect (percentagebias_TE) against width of measurement interval (width_measurementinterval), for all different distributions, values of ω, τ and β.

Table 4

The descriptive statistics of percentage standard error bias for the different values of ω, τ, β, width and for all four different examined distributions.

	min	max	range	Mean	SD
ω 0.25	-3.62	3.21	6.83	-0.21	1.62
ω 0.5	-3.70	4.08	7.78	-0.37	1.57
ω 0.75	-4.13	3.92	8.05	-0.01	1.56
τ 0.5	-4.13	3.21	7.34	-0.23	1.59
τ 1	-3.70	4.08	7.78	-0.11	1.61
τ 2	-3.62	2.83	6.45	-0.26	1.58
β -0.125	-3.70	2.72	6.43	-0.25	1.62
β -0.25	-4.13	4.08	8.21	0.09	1.70
β -0.5	-3.62	2.46	6.08	-0.43	1.40
width 0.0	-2.75	4.08	6.83	-0.08	1.73
width 0.25	-3.42	3.92	7.33	-0.09	1.61
width 0.5	-3.70	3.78	7.48	-0.12	1.56
width 1.00	-4.13	2.83	6.96	-0.42	1.54
fixed time points	-2.75	4.08	6.83	-0.08	1.73
normal	-3.62	3.21	6.83	-0.21	1.47
uniform	-4.13	3.78	7.91	-0.18	1.72
chi-square	7.23	-3.31	3.92	-0.24	1.54

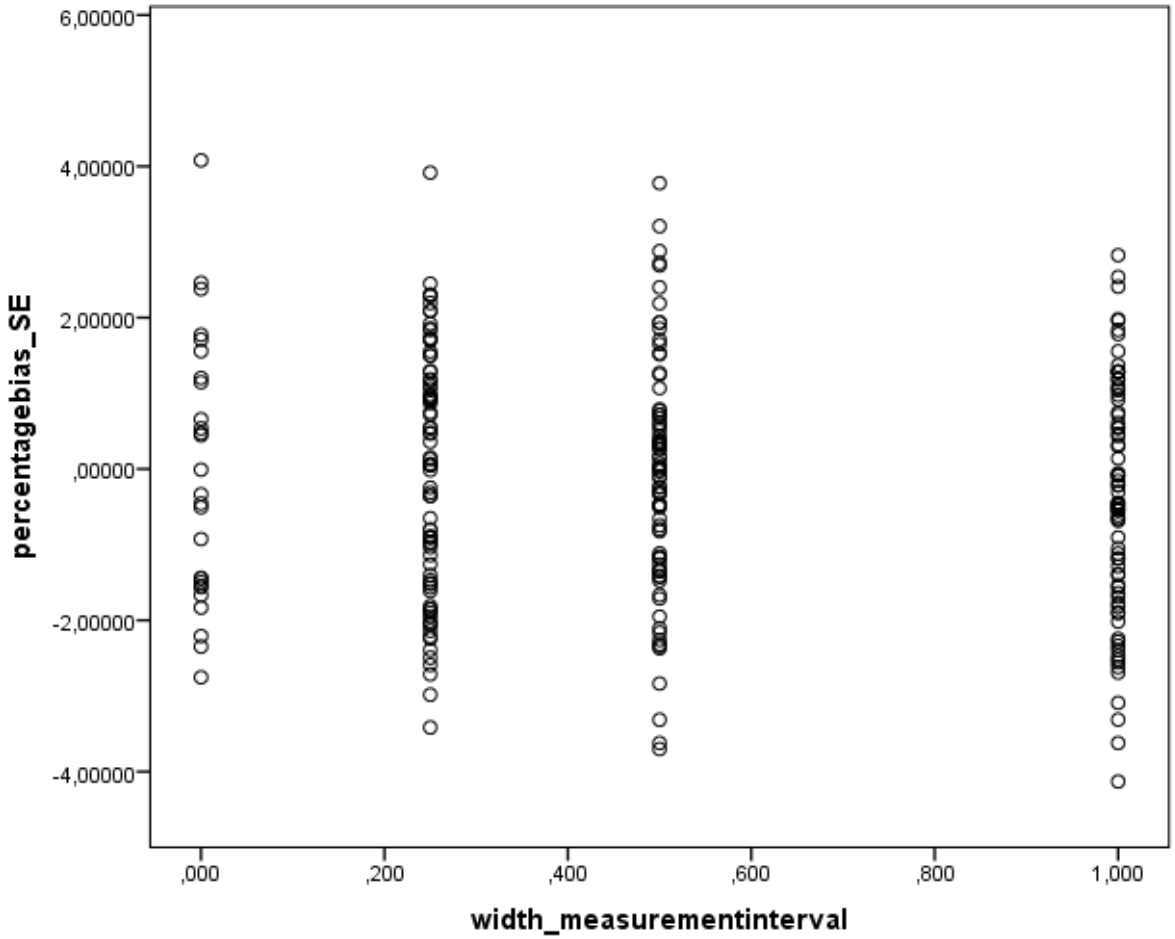

Figure 5. The distribution of percentage bias in standard error (percentagebias_SE) against width of measurement interval (width_measurementinterval). for all different distributions. values of ω. τ and β.