

Abstract

The time until the occurrence of an event of interest is a quantity that is often sought after by clinicians, engineers, and numerous other scientists for understanding lifetimes. Through scientific knowledge, it can be assumed that many distributions are stochastically ordered. Including a stochastic ordering constraint on estimators can vastly improve the bias and mean squared error properties of an estimator. Rojo (2004) developed a pair of estimators for the case of estimating two stochastically ordered survival functions. The goal of this research is to develop a generalization of Rojo's estimators to accommodate for the case of estimating more than two survival functions. The quality of the estimators developed in this study is assessed through simulations, testing the estimators against a variety of scenarios. The estimator proposed in this research is shown to have better mean squared error properties than estimators proposed by Barmi and Mukerjee (2005) for cases with equal censoring rates. However, for cases with unequal censoring between distributions, Barmi and Mukerjee's estimator performs better than the estimator proposed in this study.

Table of Contents

1	Introduction	1
1.1	Survival Analysis	1
1.2	Literature Review	2
1.2.1	Censored Estimators	3
1.2.2	Stochastically Ordered Estimators	6
1.3	Examples and Applications	13
1.3.1	Engineering	14
1.3.2	Ecology	14
1.3.3	Economics	15
1.3.4	Business	16
1.3.5	Additional Applications	17
2	Developing Estimators	18
3	Simulations	21
3.1	Comparing Estimators	21
3.2	Equal Censoring Rates	23
3.3	Unequal Censoring Rates	24
4	Conclusions	46
4.1	Future Directions	46
	References	47

List of Figures

1.1	Survival curves for diabetic and non-diabetic populations undergoing CABG and PTCA in the BARI study (Brooks et al., 2007)	7
1.2	The comparison between ESPs with and without emulsion (Pflueger, 2011).	14
1.3	The survival curves representing the time until an insect arrives at a male (σ) or female (φ) flower (Muenchow, 1986).	15
1.4	Empirical survival functions of the times of unemployment among different education levels in Romania (Ciuca Matei, 2010).	16
3.1	An exponentially distributed survival curve with red lines representing 0.2 increments and blue lines representing the corresponding x values.	22
3.2	MSE for (1.7) tends to do better than (2.1) for high censoring rates.	25
3.3	MSE for (1.7) tends to do worse than (2.1) for low censoring rates.	26
3.4	MSE for (1.7) tends to do better than (2.1) for high censoring rates.	27
3.5	MSE for (1.7) tends to do worse than (2.1) for low censoring rates.	28
3.6	MSE for (1.7) tends to do worse than (2.1) for high censoring rates.	29
3.7	MSE for (1.7) tends to do worse than (2.1) for low censoring rates.	30
3.8	MSE for (1.7) tends to do worse than (2.1) for low censoring rates.	31
3.9	MSE for (1.7) tends to do worse than (2.1) for low censoring rates.	32
3.10	MSE for (1.7) tends to do worse than (2.1) for low censoring rates.	33
3.11	MSE for (1.7) tends to do worse than (2.1) for low censoring rates.	34
3.12	MSE for (1.7) does better than (2.1) when the stochastically smallest distribution has a higher censoring rate than the other distributions.	36
3.13	MSE for (1.7) does better than (2.1) when the stochastically largest distribution has a higher censoring rate than the other distributions.	37

3.14	MSE for (1.7) does better than (2.1) when the stochastically smallest distribution has a higher censoring rate than the other distributions.	38
3.15	MSE for (1.7) does better than (2.1) when the stochastically largest distribution has a higher censoring rate than the other distributions.	39
3.16	MSE for (1.7) does better than (2.1) when the stochastically smallest distribution has a higher censoring rate than the other distributions.	40
3.17	MSE for (1.7) does better than (2.1) when the stochastically largest distribution has a higher censoring rate than the other distributions.	41
3.18	MSE for (1.7) does better than (2.1) when the stochastically smallest distribution has a higher censoring rate than the other distributions.	42
3.19	MSE for (1.7) does better than (2.1) when the stochastically largest distribution has a higher censoring rate than the other distributions.	43
3.20	MSE for (1.7) does better than (2.1) when the stochastically smallest distribution has a higher censoring rate than the other distributions.	44
3.21	MSE for (1.7) does better than (2.1) when the stochastically largest distribution has a higher censoring rate than the other distributions.	45

Chapter 1

Introduction

1.1 Survival Analysis

Survival analysis is the area in statistics that studies the times until the occurrence of an event of interest. Survival analysis has widespread use in multiple fields and is referred to by different names in the various fields. For example, survival analysis is called reliability analysis by engineers, event history analysis by sociologists, and duration analysis by economists. Although the word "survival" in survival analysis refers to the event of interest as death, the event of interest can be a variety of different events such as a heart attack, the waiting times until the next telephone call, or a mechanical failure. Studying the times until an event time is essential for prediction and inference. Comparing the survival times between populations is the key to understanding that one medical treatment yields higher survival than another medical treatment, or understanding that one mechanical part is more reliable than another mechanical part.

A survival function describes the probability of surviving beyond a given time. Survival functions may take on different shapes depending on the underlying probability distributions of the stochastic phenomenon in question. Useful survival functions include the exponential, Weibull, gamma, and log-normal distributions. Each of these distributions may be used to model the stochastic behavior of the times at which the event occurs. The true survival functions are rarely known and must be estimated.

1.2 Literature Review

An estimator is a mathematical rule that uses data to estimate a parameter. In survival analysis, the parameter of interest is a survival function. Therefore, estimators use data to estimate the underlying survival function. Estimators are important because they are the basis for testing for the equality of two or more survival functions. It is possible that two equal distributions may look different due to random noise.

To best understand the context of this research, it is necessary to understand the standard estimators used in survival analysis, since many of the stochastically ordered estimators rely on these basic estimators to estimate survival functions.

A random variable is a variable whose value is not known absolutely, but whose value is one that varies under a probability distribution. A probability distribution maps values to probabilities of obtaining such a value. For example, the probability of obtaining a 3 on a six sided die is $1/6$. Random variables are denoted by a capital letter such as X .

The simplest estimator for survival curves is the empirical survival function. Since the empirical survival function is so basic, it is the most popular estimator when trying to estimate survival functions with uncensored data. Let n be the number of observations, t be the time, $X_i, i = 1, \dots, n$ be the survival times, and I be the indicator function that takes on the value of 1 when X_i is less than t and a value of 0 when X_i is greater than t . The empirical survival function is defined as

$$\hat{F}_n(t) = 1 - \frac{1}{n} \sum_{i=1}^n I_{X_i \leq t}. \quad (1.1)$$

As shown above, the empirical survival function is simply the fraction of the surviving sample at a given time.

1.2.1 Censored Estimators

It is common for a dataset to have censored observations. Data are censored whenever the data are known to a certain degree, but not known exactly. For instance, the data may be known to be greater than a certain value, smaller than a certain value, or in between two values. More formally, for random variables T and C , a data point X is censored if for a survival time T and a censoring time C ,

$$X = \min(T, C) = C. \quad (1.2)$$

The variable C may be constant or random. C may be constant when no more event times may be observed because a study ends or an instrument breaks. Handling censored observations is a challenging topic in statistics because the problem is to extract as much information from the data as possible. Although the exact event time of censored data is unknown, censored data contains useful information that can be taken advantage of in the data analysis.

There are several types of censoring. Left censoring is knowing that the time of event of an observation is below a certain value. For example, if a study is trying to determine the time students start studying for an exam, students who have started studying for the exam before the study began are left censored because the students started studying before the start of the study, and the exact time is unknown. Conversely, right censoring is knowing that the survival time is above a certain value, which may be a constant or random variable. Consider a trial whose goal is to model the times of death of people with Alzheimer's disease. If the study loses contact with an 80-year-old participant, the participant is right censored. It is only known that the participant is at least 80 years old, but the exact age the participant dies is unknown. Lastly, interval censoring is knowing that the survival time is between two values. An example of interval censoring is knowing that a patient of a dentist developed a cavity between two check-ups, but it is unknown exactly when the patient got a cavity. For the purpose of the research described in this thesis, the data are assumed to be

right censored.

Censoring is an important aspect and a prevalent feature of many survival analysis studies. Henceforth, it is assumed that knowing the censoring time of a data point does not imply any further knowledge about the data point's event time. For example, if participants who have poorer health in a study tend to withdraw from the study more often, then the assumption that the censoring time does not provide any further information about the participants' survival time is violated. Since it is known that people with poorer health tend to withdraw from the study more often than healthy participants, then the person with poorer health is more likely to experience an event sooner than a healthy person.

To understand the significance of censoring, take the example of a participant in a heart disease study. The clinicians in the study are interested in the survival of participants that undergo a heart disease treatment. A 75-year-old participant in the study moves away and the clinicians are unable to follow up with the participant. Therefore, the clinicians have a limited knowledge of how long the participant survived with the treatment. The clinicians know that the participant lived to at least 75 years old since the participant was still alive when the participant moved away. The estimated survival times would be negatively biased if a participant was marked dead every time the participant left a study. In contrast, removing the data point completely from the data set would be a waste of money since the information contained in the data point would be lost. Therefore, the data is handled through the use of specialized estimators that are developed to handle the case of censored data. By using specialized estimators that can handle censored data, the data point is discarded, instead, the data point provides information that is known about the participant.

There is a larger variety of estimators to use when dealing with censored data. There are many philosophies (Kaplan & Meier, 1958; Efron, 1967; Gill, 1980; Moeschberger & Klein, 1985) on how to handle censored data. There is an ambiguity as to what to do with a data point when it is censored. As a result, there are many flavors of estimators, all of which take on different ranges of biases.

The most popular estimator to calculate survival curves with censoring is the Kaplan-Meier estimator, proposed by Kaplan and Meier (1958), defined as

$$\hat{F}(t) = \begin{cases} 1 & : t < t_1 \\ \prod_{t_i \leq t} [1 - \frac{d_i}{Y_i}] & : t \geq t_1 \end{cases} \quad (1.3)$$

such that t_i is the time of the i th event, d_i represents the number of events at time t_i , and Y_i represents the number of data points that are able to experience an event at time t_i - the number of items at risk prior to t_i . For the case of no censoring ($d_i = 1$), the Kaplan-Meier estimator reduces to the empirical survival function. However, when there is censoring, the Kaplan-Meier estimator treats each censored observation as a reduction in population, but does not count the censored observation as a death.

It is possible that estimators that handle censored data produce an estimated survival curve that never reaches zero. For example, Gill (1980) proposes an estimator that defines the estimated survival beyond the last observation to be the same as the survival at the last observation when the observation is censored. The idea behind this estimator is that since the time of death of the last surviving observation is unknown, it is better to not make any assumptions about the time of death of the last censored observation. However, a survival curve that never reaches zero is problematic because the estimator suggests that there is a possibility that the event time will never occur. Of course, the survival curve should always reach zero since all event times should occur eventually.

The Efron estimator (1967) is one approach that ensures that the estimated survival function converges to zero. As previously stated, the Gill estimator allows for the possibility for an estimated survival curve to never reach a survival of zero. The Efron estimator addresses this problem by redistributing the weights of censored observations to all observations to the right of a censored observation. At the final censored or uncensored observation, the estimated probability of survival is always zero. Therefore, the Efron estimator tends to have a negative bias since the estimator forces the tail to zero to avoid ambiguity of survival. Since the Gill and Efron estimators differ so greatly in philosophies of how to

handle cenroing, other statisticians (Brown, Hollander, and Kowar, 1974; Moeschberger & Klein, 1985) have explored less extreme approaches. Brown, Hollander, and Kowar (1974) suggested the following estimator for the estimator beyond the last observation

$$\hat{F}(t) = \exp \left\{ t \frac{\ln[\hat{F}(t_{\max})]}{t_{\max}} \right\}, \quad (1.4)$$

such that t_{\max} represents the time of the last observation. This estimator creates an exponential tail to the end of the survival distribution function after the last observation. This approach is less extreme than the Gill and Efron estimators since the approach does not take on stagnant values such as zero or the last observation's survival until infinity. Since it is common to have Weibull distributed data, Moeschberger and Klein (1985) developed an analogous approach that makes the distribution after the last event follow a Weibull distribution.

1.2.2 Stochastically Ordered Estimators

Through scientific knowledge, some populations may be assumed to have higher survival probabilities than other populations. Such a population is described as being stochastically ordered since the underlying survival functions are ordered in terms of survival times. However, the ordering may be stochastic (synonymous to random) due to noise. Although one survival function may be stochastically larger than than another survival function, the inherent noise in data may make one distribution appear to have a larger survival function at some points. For survival distribution functions $F(t)$ and $G(t)$, Lehmann (1955) defined $F(t)$ to be stochastically larer than $G(t)$ if $F(t) \geq G(t)$ for all t .

To illustrate stochastic ordering further, it is best to analyze an example. During Bypass Angioplasty Revascularization Investigation (BARI), the survival times for diabetic and non-diabetic populations were studied for two different cardiovascular treatments (Brooks et al., 2007). The investigation concluded that the diabetic participants in the study who underwent the coronary artery bypass grafting (CABG) treatment had better overall sur-

vival rates than the participants who underwent the percutaneous transluminal coronary balloon angioplasty (PTCA) treatment as shown in figure 1.1. Therefore, in this case, the participants who underwent CABG are said to have stochastically higher survival times than participants who underwent PTCA since the probability of survival was higher for CABG than PTCA for all times. Figure 1.1 illustrates that the non-diabetic population who underwent revascularization has overall higher survival rates than the diabetic population who underwent revascularization. Hence, the non-diabetic population is said to have stochastically higher survival times than the diabetic population.

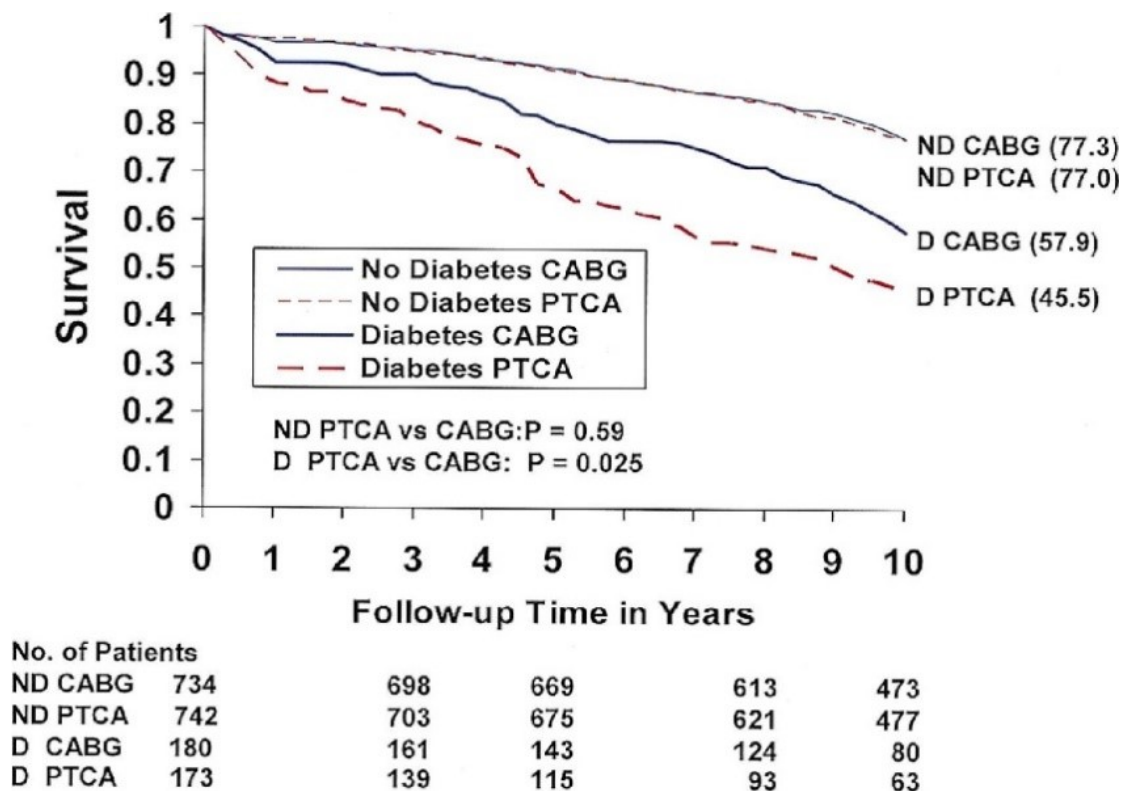


Figure 1.1: Survival curves for diabetic and non-diabetic populations undergoing CABG and PTCA in the BARI study (Brooks et al., 2007)

When the assumption of stochastic ordering holds, incorporating this constraint in the estimator improves estimation. Including a stochastic ordering constraint improves estimation because a survival function F_1 that is known to be stochastically larger than an-

other survival function F_2 should not have estimates that are stochastically smaller than F_2 . Therefore, upholding stochastic ordering for estimators yields better estimates when it is known that distributions are stochastically ordered.

Besides the estimators that handle censored and uncensored data, estimators that are constrained by stochastic ordering have also been considered. If the underlying distributions that the data follows are stochastically ordered, then the estimators should also be stochastically ordered. The empirical distributions are not always stochastically ordered due to inherent noise in the data, allowing for empirical survival curves to violate stochastic ordering. Stochastically ordered estimators solve this problem by forcing the distributions to obey stochastic ordering.

Nonparametric likelihood estimation (NPMLE) was the original focus of research in stochastically ordered estimators. Dykstra (1982) developed maximum likelihood estimators for two stochastically ordered distributions. These results were extended to multiple stochastically ordered distributions for uniform, linear, and partial stochastic orderings (Dykstra, Kochar, and Robertson, 1991; Feltz & Dykstra, 1985; Dykstra & Feltz, 1989). Despite the success of NPMLEs, it became clear that the distribution theory was difficult to obtain (Dykstra, Kochar, and Robertson, 1991). In response to the difficulty in obtaining the distribution theory for NPMLEs, research began to shift to new approaches to obtain the distribution theory.

For example, Lo (1987) proposed the following estimators for the case of having two populations. Let F_1 and F_2 be survival functions, such that F_1 is stochastically smaller than F_2 . It follows that Lo's estimators are

$$\begin{aligned}\hat{F}_1(t) &= \min(F_1(t), F_2(t)) \\ \hat{F}_2(t) &= \max(F_1(t), F_2(t)).\end{aligned}\tag{1.5}$$

Lo forces stochastic ordering of the survival functions F_1 and F_2 by using the min and max functions. Using this estimator ensures that F_1 is stochastically smaller than F_2 . Rojo (1995) determined the weak asymptotic convergence of these estimators considered by Lo.

Rojo (2004) proposed a similar estimator for two populations, but relies on the construction of a benchmark function. For each of the distributions F_1 and F_2 , there are the corresponding sample sizes n_1 and n_2 . The estimator that Rojo suggests is

$$\begin{aligned}\hat{F}_1(t) &= \min\left(F_1(t), \frac{n_1}{n_1 + n_2}F_1(t) + \frac{n_2}{n_1 + n_2}F_2(t)\right) \\ \hat{F}_2(t) &= \max\left(F_2(t), \frac{n_1}{n_1 + n_2}F_1(t) + \frac{n_2}{n_1 + n_2}F_2(t)\right).\end{aligned}\tag{1.6}$$

The benchmark function is the weighted average of the two survival functions. If stochastic ordering is violated at any t , both of the estimators will take the value of the benchmark function. Rojo (2004) shows that equation 1.6 outperforms equation 1.5. This is an expected result since the Rojo estimator deviates from the original empirical survival function less than the Lo estimator during violations of stochastic order.

Although there is a variety of estimators that can handle the stochastic ordering constraint, not all of the estimators are able to handle multiple distributions. For example, equation 1.6 can only handle two distributions. However, other estimators are able to handle multiple distributions. Despite the large literature on stochastically ordered estimators for two samples, the literature on stochastically ordered estimators for several populations is far less developed. Many of the stochastically ordered estimators for multiple samples rely on algorithmic techniques rather than closed-form expressions.

For example, Feltz and Dykstra (1985) suggest estimating survival functions for multiple populations by treating the task as an optimization problem. Feltz and Dykstra used the maximum likelihood estimator to obtain the estimated survival functions. Although Feltz, and Dykstra were able to obtain estimates for the survival functions, their approach did not yield a closed-form expression. The advantage that closed-form expressions have over algorithms is that closed-form expressions tend to be more computationally efficient as well as easier to interpret. Furthermore, mathematical results are more tractable for closed form expressions than algorithms. In contrast, Barmi and Mukerjee (2005) developed the following closed-form expression to describe the estimators for multiple stochastically or-

dered populations

$$\hat{F}_i(t) = \max_{r \leq i} \left(\min_{s \geq i} \left(\sum_{j=r}^s \frac{n_j}{\sum_{j=r}^s n_j} F_j(t) \right) \right), \quad (1.7)$$

where s is the total number of survival functions, n_j is the size of population j , and F_j is the survival function. Equation 1.7 reduces to equation 1.6 for the case of two populations (Barmi and Mukerjee, 2005).

The proof that (1.7) reduces to (1.6) for the case of two samples is as follows. Since there are only two populations for the 2-sample case of Barmi and Mukerjee's estimator, $i = 1$ or $i = 2$. As a result, for $i = 1$ we obtain

$$\hat{F}_1(t) = \max_{r \leq 1} \left(\min_{s \geq 1} \left(\sum_{j=r}^s \frac{n_j}{\sum_{j=r}^s n_j} F_j(t) \right) \right)$$

It follows that $r = 1$ and $s = 1$ or $s = 2$. Thus, by substituting the newfound values of r and s , we get

$$\begin{aligned} \hat{F}_1(t) &= \max_{1 \leq 1} \left(\min_{s \geq 1} \left(\sum_{j=1}^s \frac{n_j}{\sum_{j=1}^s n_j} F_j(t) \right) \right) \\ &= \min \left(\left(\sum_{j=1}^1 \frac{n_j}{\sum_{j=1}^1 n_j} F_j(t) \right), \left(\sum_{j=1}^2 \frac{n_j}{\sum_{j=1}^2 n_j} F_j(t) \right) \right) \\ &= \min \left(\left(\frac{n_1}{\sum_{j=1}^1 n_j} F_1(t) \right), \left(\sum_{j=1}^2 \frac{n_j}{\sum_{j=1}^2 n_j} F_j(t) \right) \right) \\ &= \min \left(F_1(t), \left(\sum_{j=1}^2 \frac{n_j}{\sum_{j=1}^2 n_j} F_j(t) \right) \right) \\ &= \min \left(F_1(t), \left(\frac{n_1}{n_1 + n_2} F_1(t) + \frac{n_2}{n_1 + n_2} F_2(t) \right) \right) \end{aligned}$$

We obtain the expression for $i = 2$ similarly. Starting from equation 1.7, and letting $i = 2$, we get

$$\hat{F}_2(t) = \max_{r \leq 2} \left(\min_{s \geq 2} \left(\sum_{j=r}^s \frac{n_j}{\sum_{j=r}^s n_j} F_j(t) \right) \right)$$

Since $r \leq 2$ and $s \geq 2$, then $r = 1$ or $r = 2$ and $s = 2$. Using this, we get

$$\begin{aligned}
\hat{F}_2(t) &= \max_{r \leq 2} \left(\min \left(\sum_{j=r}^2 \frac{n_j}{\sum_{j=r}^2 n_j} F_j(t) \right) \right) \\
&= \max \left(\min \left(\sum_{j=1}^2 \frac{n_j}{\sum_{j=1}^2 n_j} F_1(t) \right), \min \left(\sum_{j=2}^2 \frac{n_j}{\sum_{j=2}^2 n_j} F_2(t) \right) \right) \\
&= \max \left(\left(\sum_{j=1}^2 \frac{n_j}{\sum_{j=1}^2 n_j} F_1(t) \right), \left(\sum_{j=2}^2 \frac{n_j}{\sum_{j=2}^2 n_j} F_2(t) \right) \right) \\
&= \max \left(\left(\sum_{j=1}^2 \frac{n_j}{\sum_{j=1}^2 n_j} F_1(t) \right), F_2(t) \right)
\end{aligned}$$

Expanding the terms, we obtain

$$\hat{F}_2(t) = \max \left(\left(\frac{n_1}{n_1 + n_2} F_1(t) + \frac{n_2}{n_1 + n_2} F_2(t) \right), F_2(t) \right)$$

In conclusion, the estimators for $i = 1$ and $i = 2$ are

$$\begin{aligned}
\hat{F}_1(t) &= \min \left(F_1(t), \left(\frac{n_1}{n_1 + n_2} F_1(t) + \frac{n_2}{n_1 + n_2} F_2(t) \right) \right) \\
\hat{F}_2(t) &= \max \left(\left(\frac{n_1}{n_1 + n_2} F_1(t) + \frac{n_2}{n_1 + n_2} F_2(t) \right), F_2(t) \right)
\end{aligned} \tag{1.8}$$

Therefore, (1.7) reduces to (1.6) for two populations. Although (1.7) can be described as a closed-form expression, the process of finding an estimate can be long and tedious for obtaining the estimates of many distributions since the estimator contains nested max and min functions.

An alternative to Barmi and Mukerjee's estimator is the pointwise nonparametric maximum likelihood estimator proposed by Park et al. (2012). Similar to (1.7), the pointwise nonparametric maximum likelihood estimator is another estimator that is obtained through an algorithm. Park shows that a weakness in (1.6) is that it does not perform well when the censoring distributions between different groups are vastly different.

The goal of the research presented in this thesis is to develop a generalization of Rojo's estimator to the case of estimating more than two populations. More importantly, the goal of this research is to develop an estimator that can handle multiple stochastically ordered populations that can outperform existing estimators that can handle multiple stochastically ordered populations.

Generalizing Rojo's estimator for two or more populations begins by developing a mathematical expression that satisfies stochastic ordering and is able to handle multiple populations. To begin this process, it is best to start simply and try to satisfy stochastic ordering for three or four populations. Afterwards, the estimator for three or four populations can be generalized to any number of populations.

Simulations are run to understand how the new estimator compares to other existing estimators. In brief, a simulation generates random data that follow a known distribution and inputs the data into the estimator. Since the estimator estimates a survival function, the estimated survival function is compared to the real survival function that the data is generated from. The estimator can be studied to see how accurately the estimated survival function follows the actual survival function. In particular, the mean squared error (MSE) and bias are obtained to see how well the estimator performs. MSE is defined as

$$MSE = E(\hat{F}(x) - F(x))^2. \quad (1.9)$$

Similarly, bias is defined as

$$Bias = E(\hat{F}(x)) - F(x). \quad (1.10)$$

It is also important to study how the estimator behaves under different scenarios. For example, the estimator may change depending on the underlying distribution. Perhaps the estimator performs better under a Weibull distribution than an exponential distribution. Furthermore, the estimator will be tested with varying sample sizes. Small sample sizes often cause violations in stochastic ordering, so running simulations with varying sample

sizes will illuminate how well the estimator behaves under small sample sizes. The estimators developed should also be tested using different types of underlying estimators such as the Kaplan-Meier estimator as well as variations of the Kaplan-Meier estimator, such as the Efron estimator. By using different types of underlying estimators, it will show how the estimator behaves under different approaches.

In the current state of research, there are only a few estimators that can handle stochastically ordered survival functions. There are even fewer estimators that can handle stochastically ordered survival functions for more than two populations. The Examples and Applications section illustrates how prevalent survival analysis is in a wide variety of fields. Furthermore, it was shown by Rojo and Ma (1996) that the stochastic ordering restriction may help in improving bias. In addition, some theoretical results from NPMLEs are difficult to obtain. Consequently, there has been a growing need for alternative approaches to stochastically ordered estimators. Therefore, not only is it important to improve estimators, but to improve stochastically ordered estimators in particular. As a result, improving estimation of survival functions may help in finding better treatments for disease, or developing better hardware for machinery. All in all, since estimators are the tools used by clinicians and engineers to study the time until an event, the incremental improvement of estimators is necessary to produce higher quality inferences and predictions.

1.3 Examples and Applications

To better understand the context of this research, it is vital to understand the impact survival analysis has had on various disciplines. Although survival analysis had its origins in clinical trials, its applications have spread to a variety of disciplines. To name a few examples, survival analysis has had influence on disciplines such as medicine, engineering, economics, sociology, political science, and business. In many of these cases, stochastically ordered estimators may be used to improve MSE and bias properties.

1.3.1 Engineering

In engineering, the survival time that is of interest is the time of failure of a piece of equipment. For example, Pflueger (2011) compares the lifetimes of electrical submersible pumps (ESP) for cases with emulsion and without emulsion. Emulsion is the dispersion of a liquid in another liquid. In the context of petroleum engineering, water droplets are often dispersed in oil, which causes a variety of problems. It is clear from figure 1.2 that ESPs without emulsion last longer than ESPs with emulsion since the probability of survival for ESPs without emulsion is consistently higher than ESPs with emulsion. Therefore, ESPs are stochastically ordered with respect to emulsion. The ESPs with emulsion tend to do worse than the ESPs without emulsion. As a result, better estimates can be found by using stochastically ordered estimators.

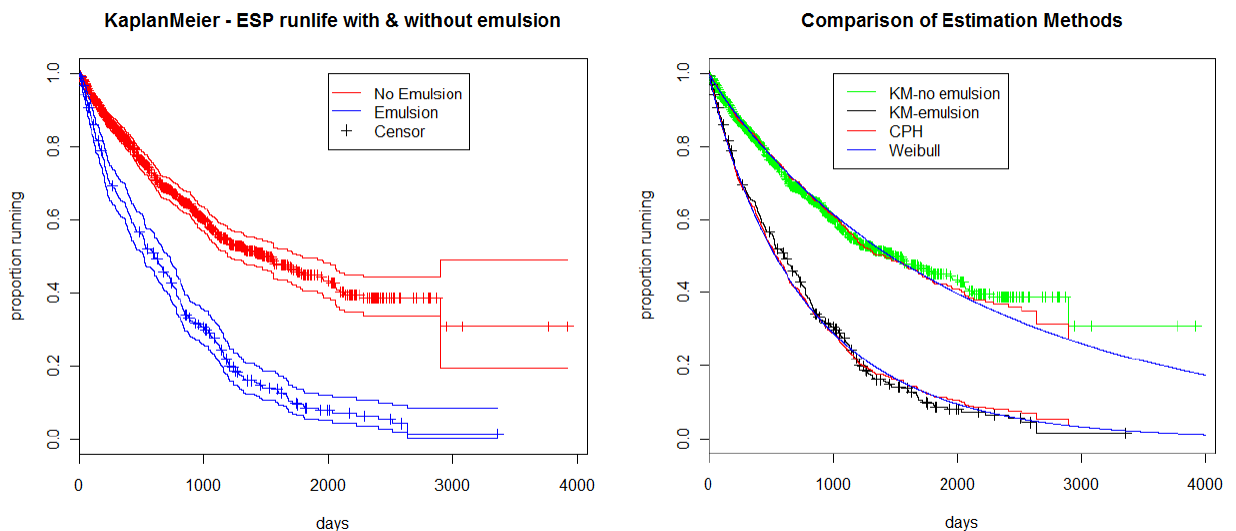


Figure 1.2: The comparison between ESPs with and without emulsion (Pflueger, 2011).

1.3.2 Ecology

Applications of survival analysis have also been found in ecology. Muenchow (1986) studied the times that male and female flowers of *Clematis ligusticifolia* were visited by insects.

Muenchow concluded that male flowers were visited faster than female flowers; see figure 1.3. Therefore, to estimate a survival curve, a stochastically ordered estimator is desirable since the event times that male flowers are visited are stochastically smaller than female flowers.

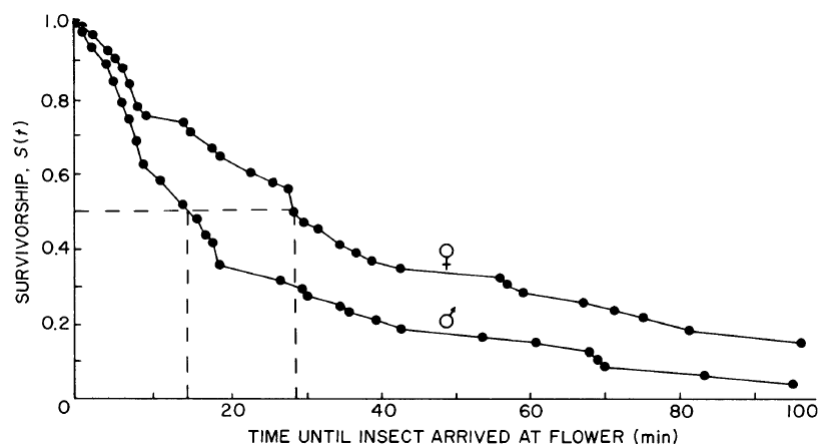


Figure 1.3: The survival curves representing the time until an insect arrives at a male (σ) or female (φ) flower (Muenchow, 1986).

1.3.3 Economics

It has been clear in the previous examples that the separate groups are likely to be stochastically ordered. However, the use of stochastically ordered estimators is most important in cases where the ordering is less clear. For example, Ciuca and Matei (2010) studied unemployment in Romania and the influences that different variables have on unemployment survival curves. Figure 1.4 below shows the survival functions of unemployment for different educational levels in Romania. Level 1 denotes the population with an elementary school level education, level 2 represents the population with vocational school education, level 3 represents the population with a high school education, and level 4 represents the population with higher education. In this example, a trend that the population with more education tends to be less unemployed. However, the data suggests that the population with

level 1 education does better than the other levels of education at three to four months. If it is known that people with more education tend to have shorter times of unemployment, then it would imply that the stochastic ordering violation at three to four months is due to the inherent variability in the data and the estimators should respect stochastic ordering to minimize MSE and bias.

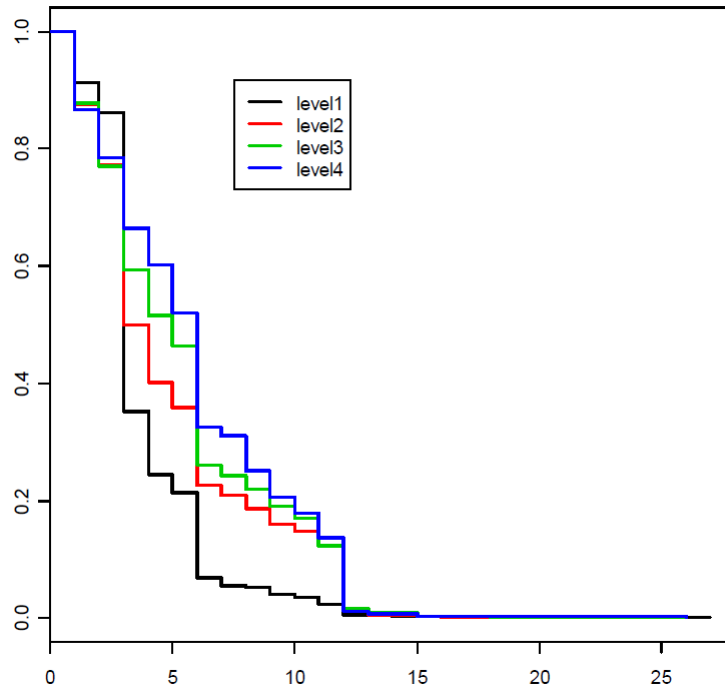


Figure 1.4: Empirical survival functions of the times of unemployment among different education levels in Romania (Ciuca Matei, 2010).

1.3.4 Business

Although stochastic ordering is common among many applications, it is important to avoid making wild claims of stochastic ordering. There are cases where intuitively there is reason to believe there is stochastic ordering between groups. However, incorrect conclusions may be obtained when taking such assumptions. For example, Kalleberg and Leicht (1991) studied the difference between businesses headed by men and businesses headed by

women. Kalleberg and Leicht assert that there is a stigma that “businesses owned by men are more successful than those owned by women.” Therefore, under this notion, one would believe that there is stochastic ordering between the survival times of businesses headed by men and businesses headed by women. However, in the study, it was determined that businesses headed by women performed just as well as businesses headed by men. As a result, it is a mistake to use a stochastically ordered estimator in this scenario since the survival times of businesses led by men and women are not actually stochastically ordered.

1.3.5 Additional Applications

Additional applications of survival analysis include the usage in economics, sociology, and political science. Helmers and Rogers (2007) studied the survival of startups in Britain. They concluded that failure rates differed greatly between different regions in Britain as a result of different competitive conditions. Survival analysis was applied to sociology by Logubayom and Luguterah (2013) as they studied the time from marriage to the birth of the first child of women in Ghana. Diermeier, Keane, and Merlo (2002) developed a model to study the effects of career decisions and probabilities of reelection using survival analysis. Therefore, it is clear that the influence of survival analysis is widespread. It has touched areas ranging from medicine to political science.

Chapter 2

Developing Estimators

This chapter presents the estimator developed to generalize (1.6) and shows that the estimator reduces down to (1.6) in the case of two populations.

For k distributions, assume the distributions $F_i(t)$, $i = 1, 2, \dots, k$ are stochastically ordered such that $F_1(t) < F_2(t) < \dots < F_k(t)$. For each distribution $F_i(t)$, there is the corresponding sample size n_i . The estimator developed in this study is defined as

$$\hat{F}_i(t) = \frac{\sum_{j=1}^i (n_j F_j(t)) I_{\{F_j(t) > F_i(t)\}} + \sum_{j=i}^k (n_j F_j(t)) I_{\{F_j(t) < F_i(t)\}} + n_i F_i(t)}{\sum_{j=1}^i (n_j) I_{\{F_j(t) > F_i(t)\}} + \sum_{j=i}^k (n_j) I_{\{F_j(t) < F_i(t)\}} + n_i} \quad (2.1)$$

Similar to (1.6), (2.1) takes the weighted average of distributions that violate stochastic ordering at t . The estimator creates an individual estimate for each of the k distributions by iterating through each of the distributions. While iterating through each of the distributions, the estimator also checks if any other distribution violates stochastic ordering. At each instance of stochastic ordering violation, the estimator makes a weighted average of all violating distributions.

This estimator is shown to be reducible to (1.6) as follows. Let $k = 2$. For F_1 , it follows

that

$$\hat{F}_1(t) = \frac{\sum_{j=1}^1 (n_j F_j(t)) I_{\{F_j(t) > F_1(t)\}} + \sum_{j=1}^2 (n_j F_j(t)) I_{\{F_j(t) < F_1(t)\}} + n_1 F_1(t)}{\sum_{j=1}^1 (n_j) I_{\{F_j(t) > F_1(t)\}} + \sum_{j=1}^2 (n_j) I_{\{F_j(t) < F_1(t)\}} + n_1}$$

$$\hat{F}_1(t) = \frac{(n_1 F_1(t)) I_{\{F_1(t) > F_1(t)\}} + \sum_{j=1}^2 (n_j F_j(t)) I_{\{F_j(t) < F_1(t)\}} + n_1 F_1(t)}{(n_1) I_{\{F_1(t) > F_1(t)\}} + \sum_{j=1}^2 (n_j) I_{\{F_j(t) < F_1(t)\}} + n_1}$$

$$\hat{F}_1(t) = \frac{(n_1 F_1(t)) I_{\{F_1(t) > F_1(t)\}} + \sum_{j=1}^2 [n_j F_j(t)] I_{\{F_j(t) < F_1(t)\}} + n_1 F_1(t)}{[n_1] I_{\{F_1(t) > F_1(t)\}} + \sum_{j=1}^2 (n_j) I_{\{F_j(t) < F_1(t)\}} + n_1}$$

$$\hat{F}_1(t) = \frac{(n_2 F_2(t)) I_{\{F_2(t) < F_1(t)\}} + n_1 F_1(t)}{(n_2) I_{\{F_2(t) < F_1(t)\}} + n_1}$$

When there is no stochastic ordering violation ($F_1(t) < F_2(t)$), the expression results in

$$\hat{F}_1(t) = \frac{n_1 F_1(t)}{n_1}$$

$$\hat{F}_1(t) = F_1(t)$$

When there is stochastic ordering violation ($F_2(t) > F_1(t)$), the expression results in

$$\hat{F}_1 = \frac{n_2 F_2(t) + n_1 F_1(t)}{n_1 + n_2}$$

$$\hat{F}_1 = \frac{n_1 F_1(t)}{n_1 + n_2} + \frac{n_2 F_2(t)}{n_1 + n_2}$$

In a similar manner, the result for $\hat{F}_2(t)$ can also be found.

$$\hat{F}_2(t) = \frac{\sum_{j=1}^2 (n_j F_j(t)) I_{\{F_j(t) > F_2(t)\}} + \sum_{j=2}^2 (n_j F_j(t)) I_{\{F_j(t) < F_2(t)\}} + n_2 F_2(t)}{\sum_{j=1}^2 [n_j] I_{\{F_j(t) > F_2(t)\}} + \sum_{j=2}^2 [n_j] I_{\{F_j(t) < F_2(t)\}} + n_2}$$

$$\hat{F}_2(t) = \frac{\sum_{j=1}^2 (n_j F_j(t)) I_{\{F_j(t) > F_2(t)\}} + n_2 F_2(t)}{\sum_{j=1}^2 (n_j) I_{\{F_j(t) > F_2(t)\}} + n_2}$$

$$\hat{F}_2(t) = \frac{(n_1 F_1(t)) I_{\{F_1(t) > F_2(t)\}} + (n_2 F_2(t)) I_{\{F_2(t) > F_2(t)\}} + n_2 F_2(t)}{(n_1) I_{\{F_1(t) > F_2(t)\}} + (n_2) I_{\{F_2(t) > F_2(t)\}} + n_2}$$

$$\hat{F}_2(t) = \frac{(n_1 F_1(t)) I_{\{F_1(t) > F_2(t)\}} + n_2 F_2(t)}{(n_1) I_{\{F_1(t) > F_2(t)\}} + n_2}$$

When there is no stochastic ordering violation ($F_1(t) < F_2(t)$), the resulting equation is

$$\hat{F}_2(t) = \frac{n_2 F_2(t)}{n_2}$$

$$\hat{F}_2(t) = F_2(t)$$

Similarly, when there is a stochastic ordering violation ($F_1(t) > F_2(t)$), it follows that

$$\hat{F}_2(t) = \frac{n_1 F_1(t) + n_2 F_2(t)}{n_1 + n_2}$$

$$\hat{F}_2(t) = \frac{n_1 F_1(t)}{n_1 + n_2} + \frac{n_2 F_2(t)}{n_1 + n_2}$$

Notice that when there is no distribution that violates stochastic ordering of $F_n(t)$, it follows that $\hat{F}_n(t) = F_n(t)$. Therefore, it is clear that (2.1) reduces to (1.6) for two samples.

Chapter 3

Simulations

One primary objective for an estimator is to estimate a survival function as closely as possible. Bias and MSE are used to measure how close an estimate is to the survival function. Since the exact survival function of real life survival data is unknown, simulations are run to imitate real life applications and to study how the estimator behaves under different scenarios.

Since survival functions differ in terms of shape and size, bias and MSE must be compared at the underlying distribution probability quantiles. Survival curves are monotonically decreasing functions F , such that $F(0) = 1$, and $F(\infty) = 0$. As a result, to normalize all comparisons, bias and mean squared error are only calculated for certain probability quantiles. In particular, the quantiles needed are the points at which the blue and red lines in Figure 3.1 intersect. The red lines represent 20% intervals in probability, whereas the blue lines represent the corresponding x value for each 20% interval.

Before studying the properties of the estimator given by (2.1), it is important to check that all simulations are conducted correctly. All simulations were run in R, the programming language for statistical computing. Bias and MSE curves are known for Rojo's estimators because they were studied in previous simulations studies (Rojo, 2004; Rojo & Ma, 1996). Therefore, the goal is to imitate previous studies to ensure that the results obtained from this simulation study can easily be compared with past simulation studies.

3.1 Comparing Estimators

The estimator provided in (2.1) must be implemented in R to study how it performs against other estimators. Since every estimator has its own unique strengths and weaknesses, running simulations under a variety of conditions reveal the strengths and weaknesses of (2.1).

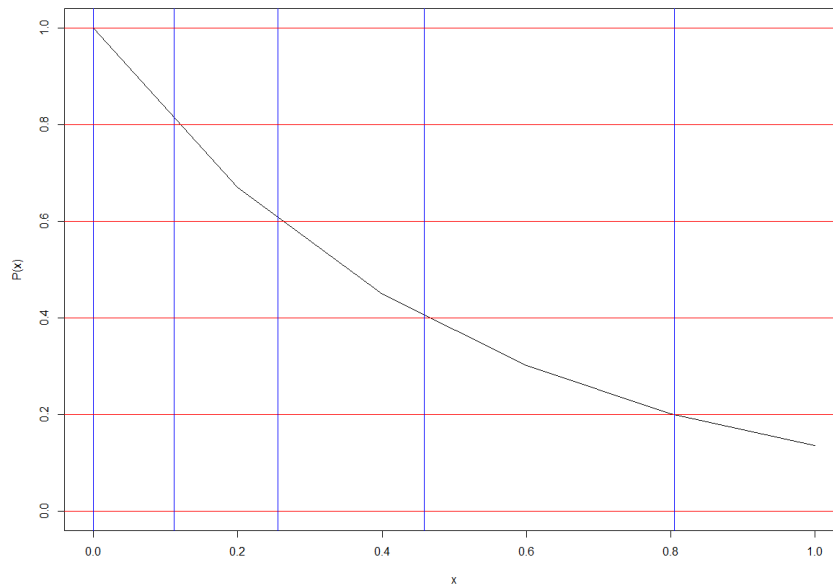


Figure 3.1: An exponentially distributed survival curve with red lines representing 0.2 increments and blue lines representing the corresponding x values.

Conditions that may change the performance of an estimator include the type of survival distribution, parameters of that survival distribution, sample size, and censoring.

Censoring rates are estimated through simulations. An observation is censored if the censoring time occurs before the event time as described in (1.2). By randomly generating censoring times and event times, the censoring rate can be calculated as the fraction of observations that are censored over the total number of observations.

The estimator developed in this study was run under various parameters to study under what conditions the estimator behaves better than Barmi and Mukerjee's estimator. (2.1) and (1.7) were both studied for varying underlying survival distributions, population sizes, and censoring rates. Each simulation was run 1000 times.

3.2 Equal Censoring Rates

As censoring rates increase, (1.7) has lower MSE than (2.1) for the exponential and heavy tail distributions as shown in Figures 3.2 - 3.5. However, Figures 3.6 - 3.11 show that (1.7) has higher MSE for the gamma, log-normal, and Weibull distributions for lower censoring rates. Often times, the estimate for (2.1) performs approximately equally as well as (1.7) or does significantly better for all types of censoring rates. Therefore, (2.1) performs better than (1.7) for most cases of equal censoring.

Equal Censoring Figures				
Figure #	Distribution	Sample Size	Censoring Rate	Parameters
3.2	Exponential	30, 30, 30	72%, 72%, 72%	rate = 1.4, 1.2, 1.0
3.3	Exponential	30, 30, 30	5%, 4%, 3%	rate = 1.4, 1.2, 1.0
3.4	Heavy Tail	30, 30, 30	75%, 76%, 76%	rate = 1.4, 1.2, 1.0
3.5	Heavy Tail	30, 30, 30	5%, 5%, 6%	rate = 1.4, 1.2, 1.0
3.6	Gamma	30, 30, 30	75%, 75%, 75%	scale = 1.4, 1.2, 1.0; shape = 2.0, 2.0, 2.0
3.7	Gamma	30, 30, 30	5%, 5%, 5%	scale = 1.4, 1.2, 1.0; shape = 2.0, 2.0, 2.0
3.8	Log-Normal	30, 30, 30	75%, 75%, 74%	mean = 1.0, 1.2, 1.4; SD = 2.0, 2.0, 2.0
3.9	Log-Normal	30, 30, 30	5%, 5%, 5%	mean = 1.0, 1.2, 1.4; SD = 2.0, 2.0, 2.0
3.10	Weibull	30, 30, 30	75%, 75%, 75%	scale = 1.0, 1.2, 1.4; shape = 2.0, 2.0, 2.0
3.11	Weibull	30, 30, 30	5%, 5%, 5%	scale = 1.0, 1.2, 1.4; shape = 2.0, 2.0, 2.0

3.3 Unequal Censoring Rates

Throughout the numerous simulations conducted illustrated by Figures 3.12 through 3.21, it is clear that (2.1) performs significantly poorer than (1.7) when the stochastically largest survival function has a very high censoring rate and the other survival functions have small censoring rates. Similarly, (2.1) performs significantly poorer than (1.7) when the stochastically smallest survival function has a very high censoring rate and the other survival func-

Estimators Under Exponentially Distributed Survival Times with High Censoring Rates

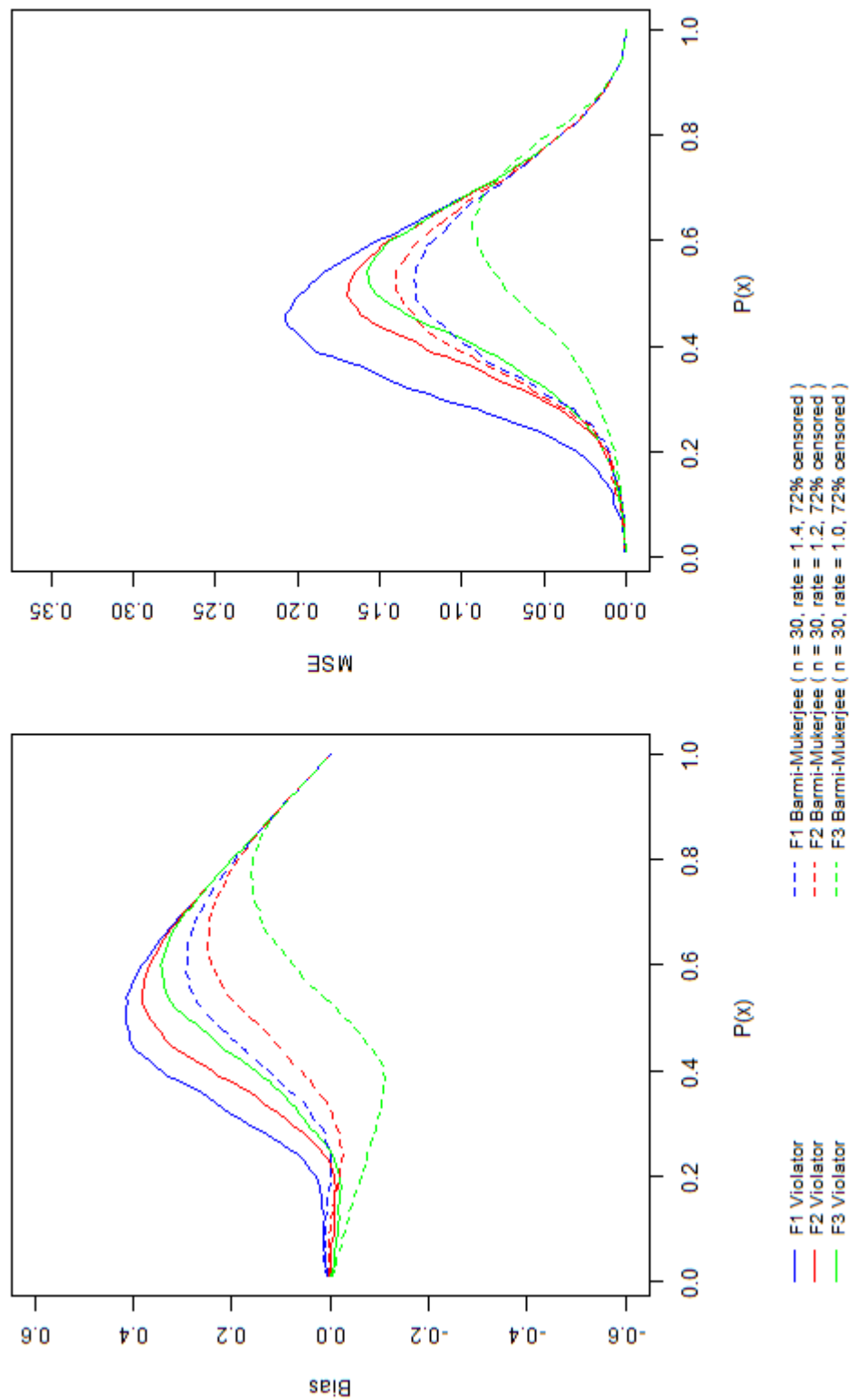


Figure 3.2: MSE for (1.7) tends to do better than (2.1) for high censoring rates.

Estimators Under Exponentially Distributed Survival Times with Low Censoring Rates

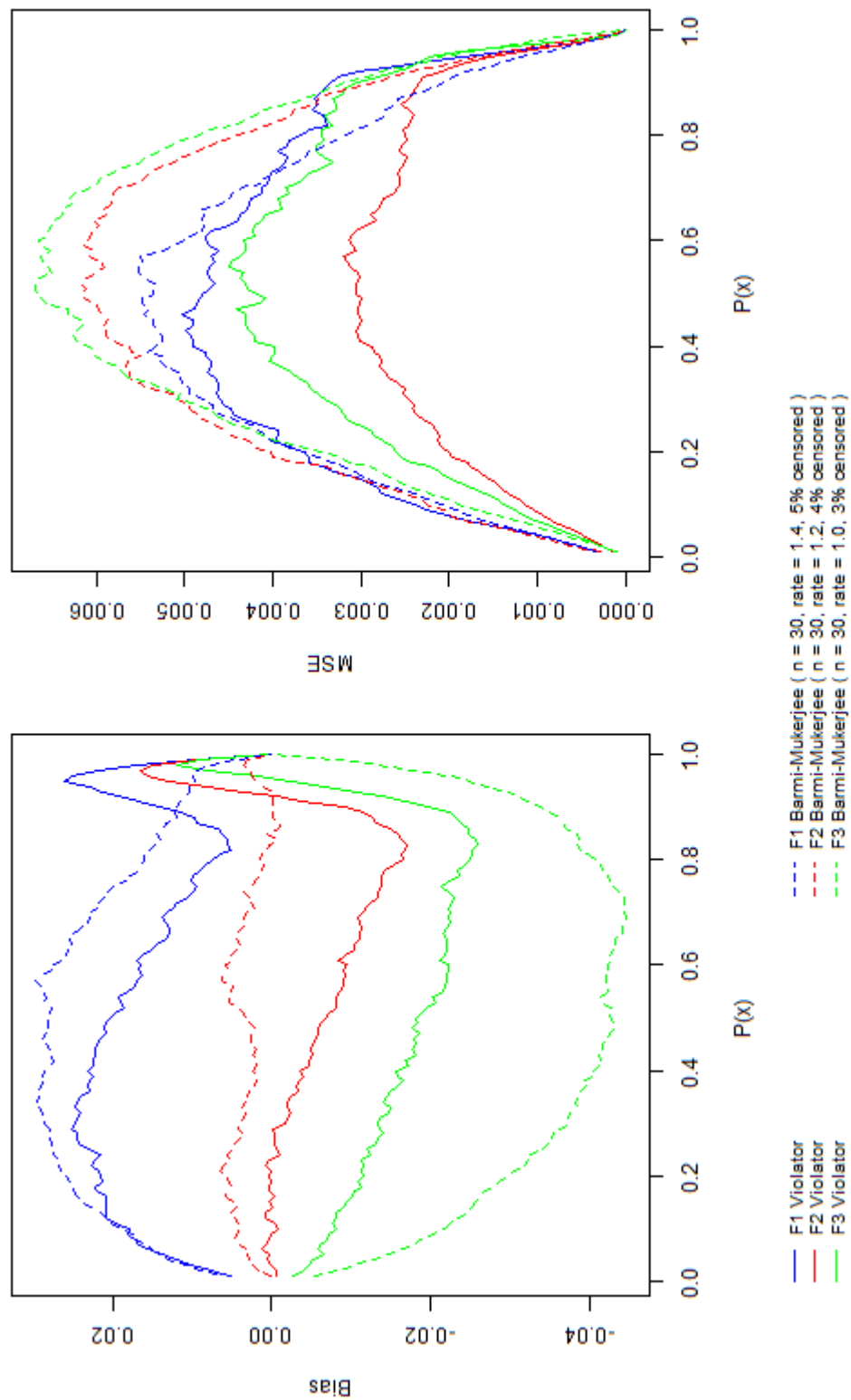


Figure 3.3: MSE for (1.7) tends to do worse than (2.1) for low censoring rates.

Estimators Under a Heavy Tail Distribution with High Censoring Rates

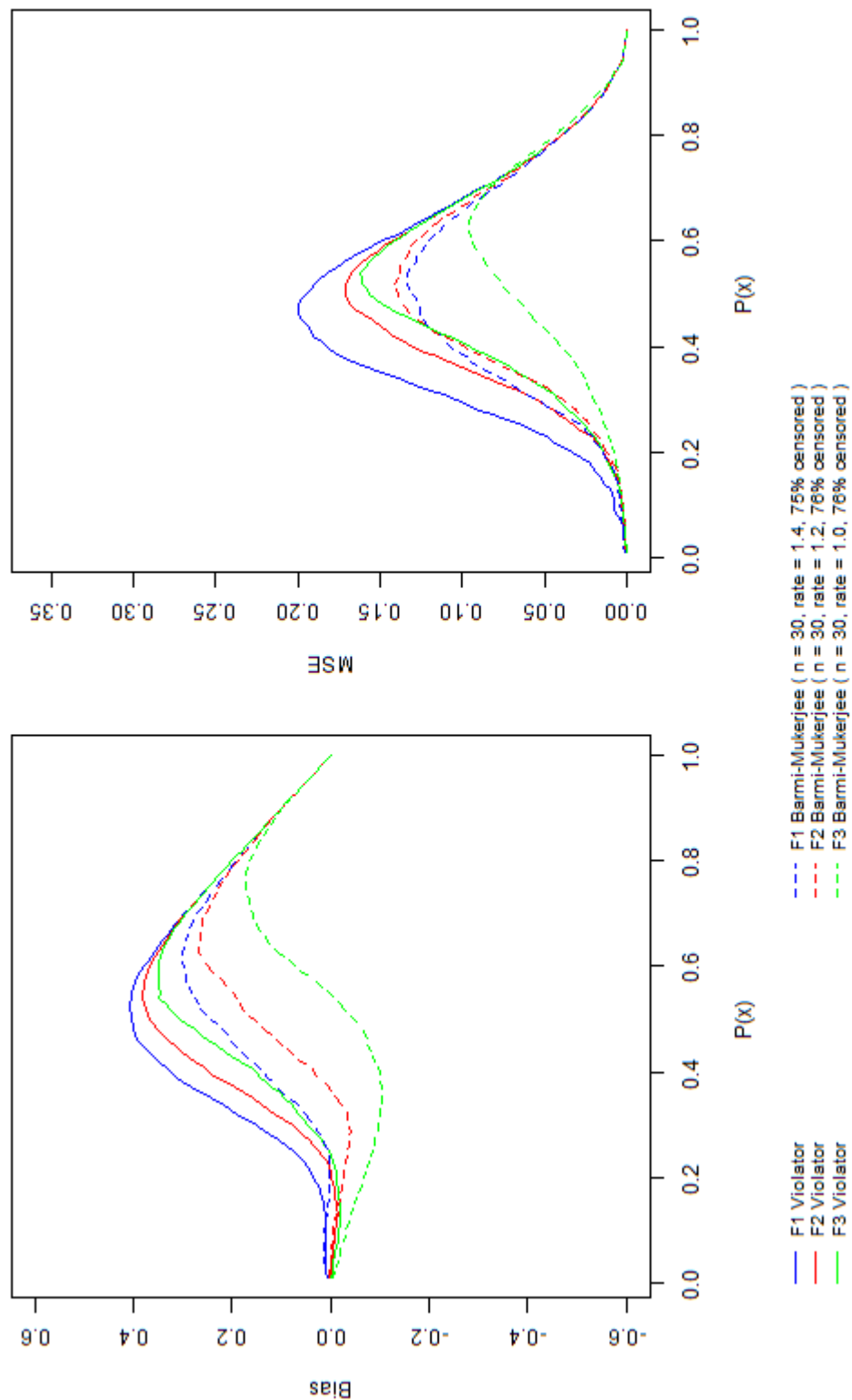


Figure 3.4: MSE for (1.7) tends to do better than (2.1) for high censoring rates.

Estimators Under a Heavy Tail Distribution with Low Censoring Rates

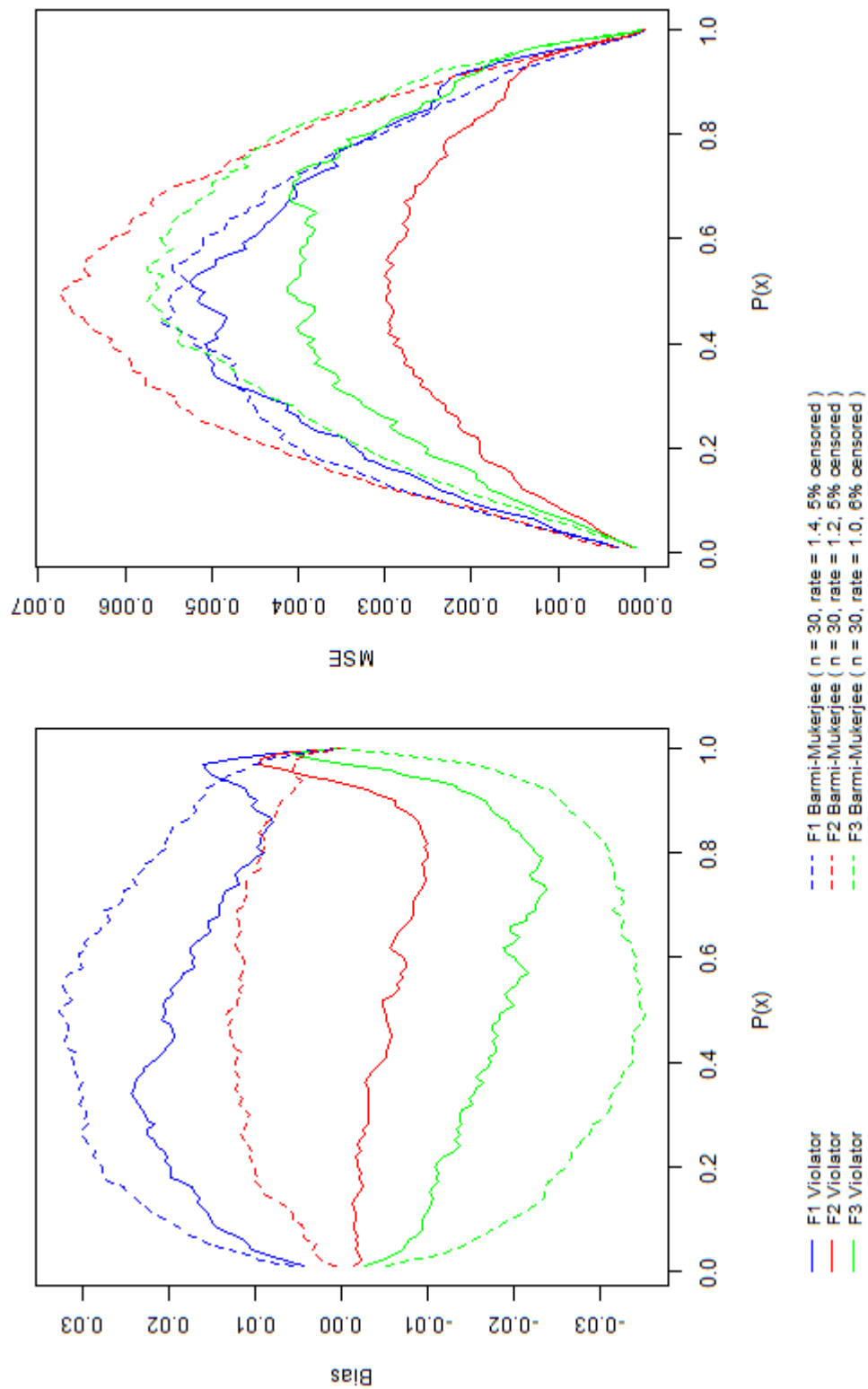


Figure 3.5: MSE for (1.7) tends to do worse than (2.1) for low censoring rates.

Estimators Under a Gamma Distribution with High Censoring Rates

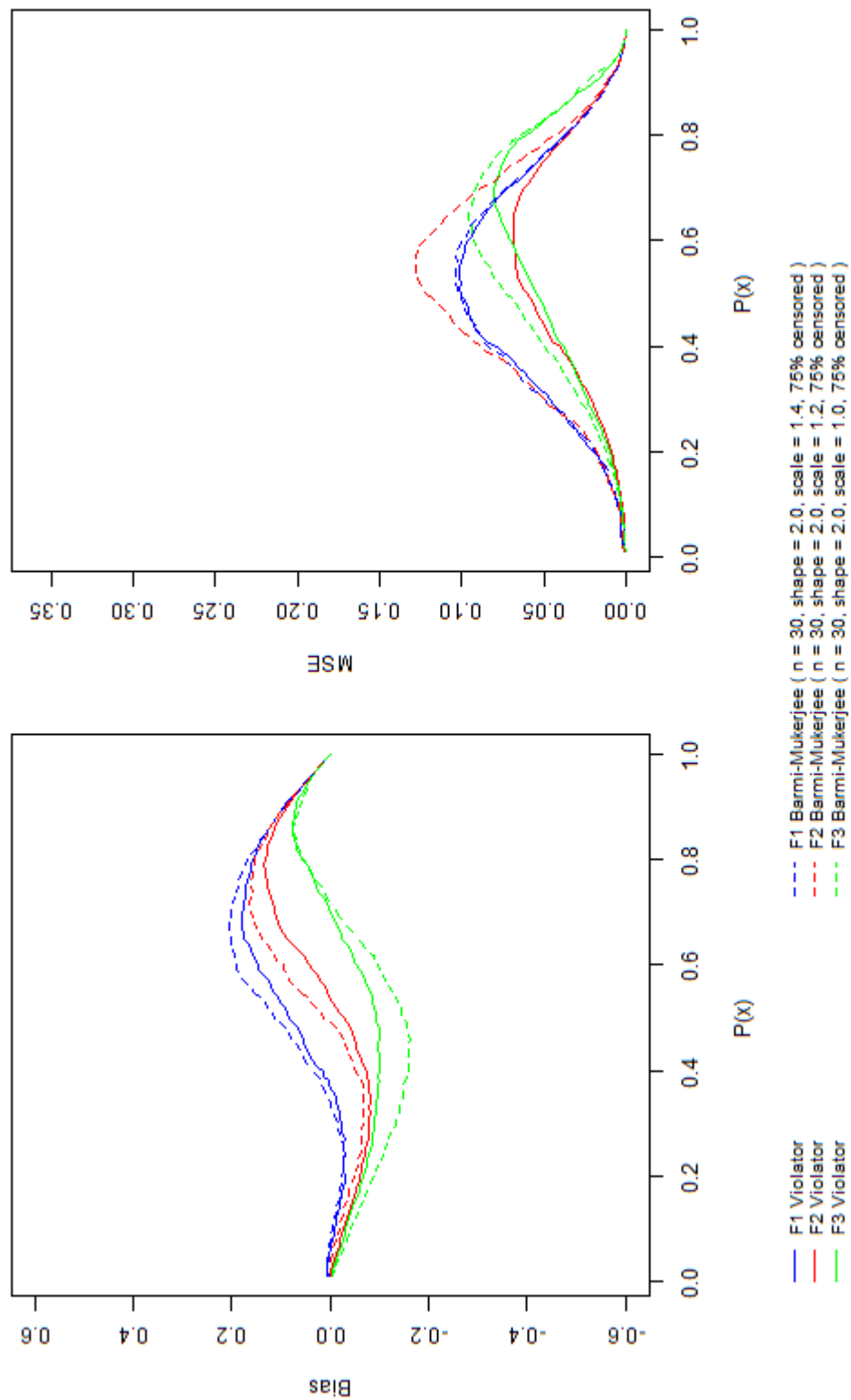


Figure 3.6: MSE for (1.7) tends to do worse than (2.1) for high censoring rates.

Estimators Under a Gamma Distribution with Low Censoring Rates

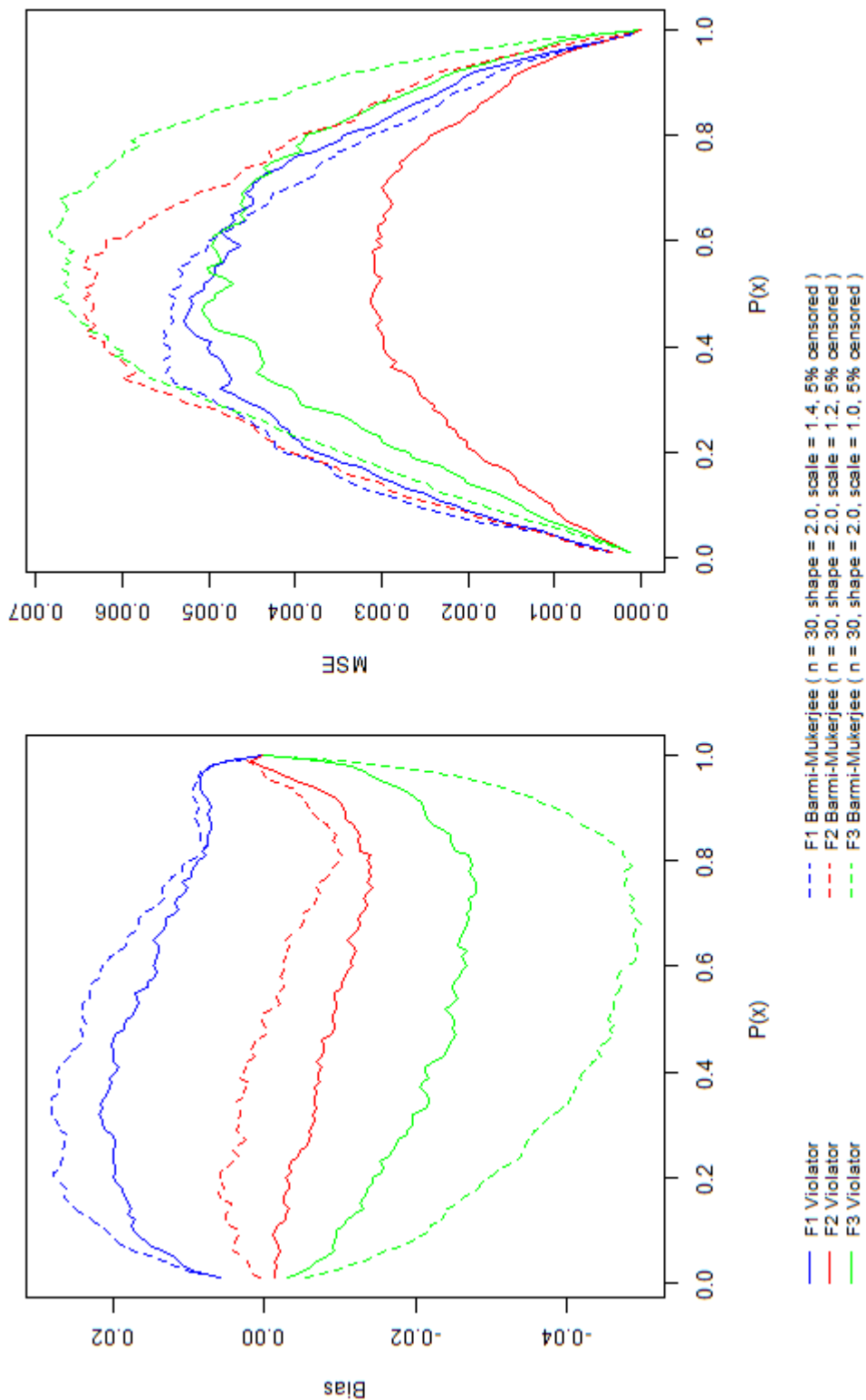


Figure 3.7: MSE for (1.7) tends to do worse than (2.1) for low censoring rates.

Estimators Under a Log-Normal Distribution with High Censoring Rates

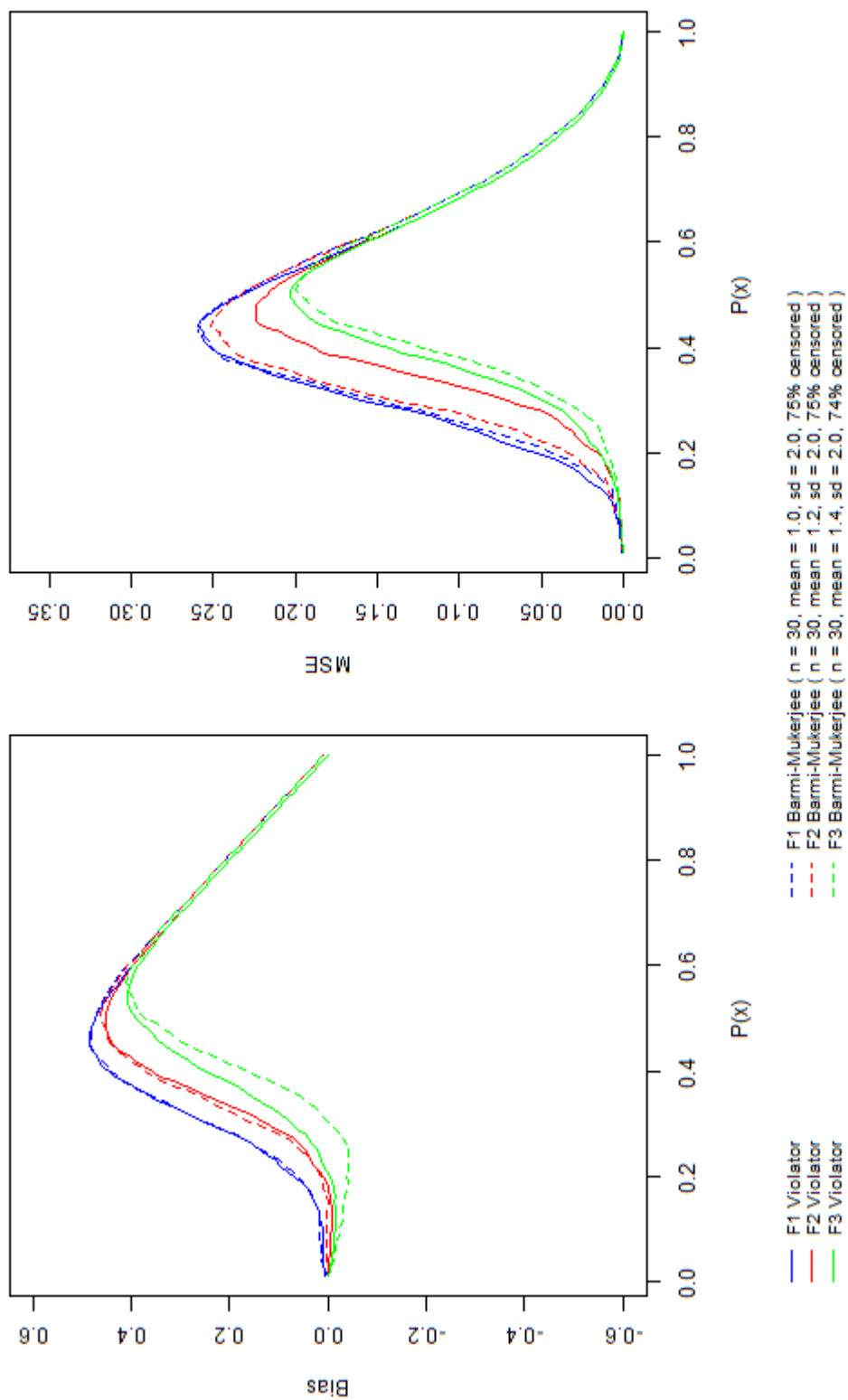


Figure 3.8: MSE for (1.7) tends to do worse than (2.1) for low censoring rates.

Estimators Under a Log-Normal Distribution with Low Censoring Rates

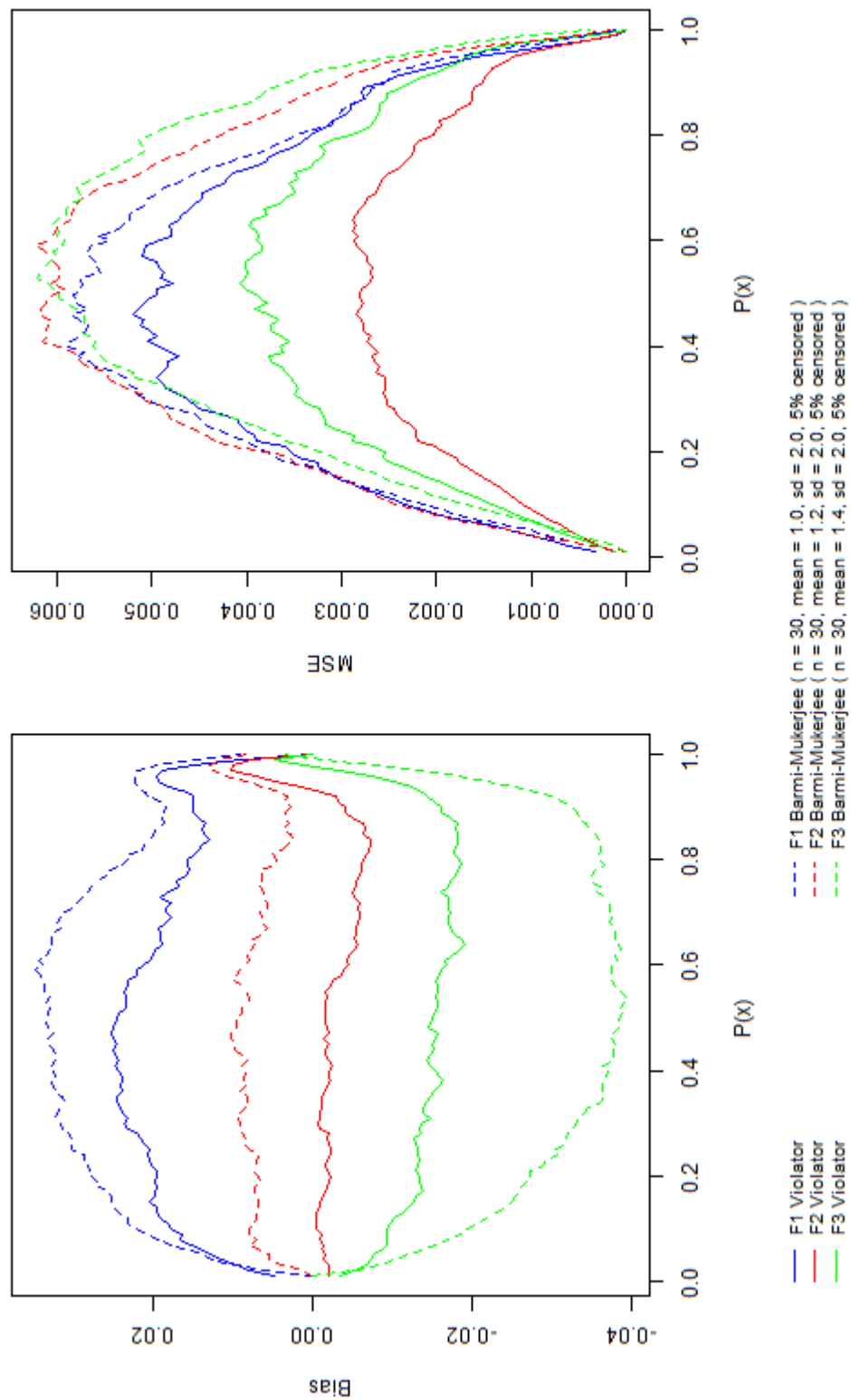


Figure 3.9: MSE for (1.7) tends to do worse than (2.1) for low censoring rates.

Estimators Under a Weibull Distribution with High Censoring Rates

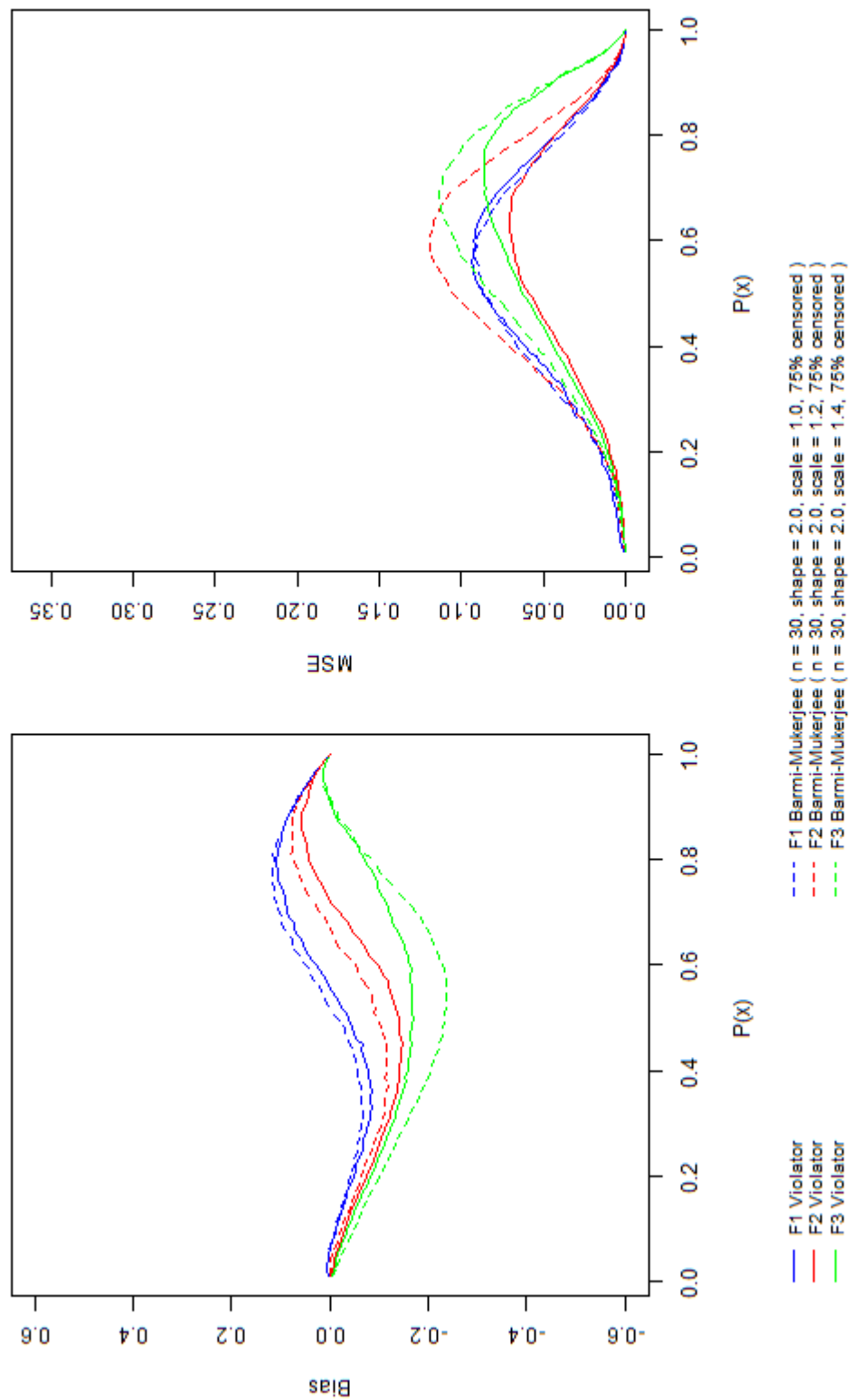


Figure 3.10: MSE for (1.7) tends to do worse than (2.1) for low censoring rates.

Estimators Under a Weibull Distribution with Low Censoring Rates

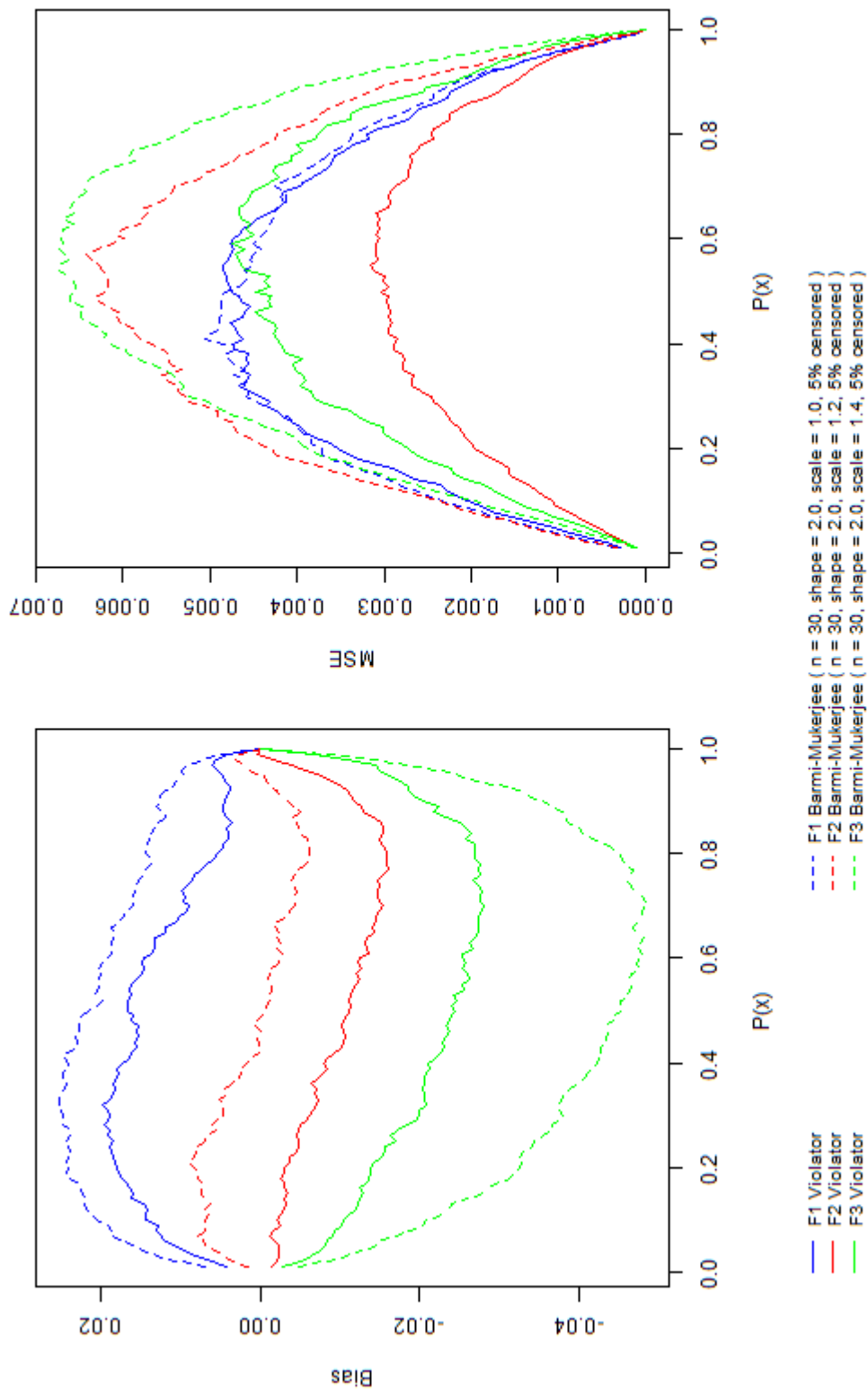


Figure 3.11: MSE for (1.7) tends to do worse than (2.1) for low censoring rates.

tions have small censoring rates. In both of the previously listed scenarios, the bias and MSE of the middle stochastically ordered survival function consistently matches the bias and MSE of the stochastically smallest or stochastically largest survival function.

Unqual Censoring Figures				
Figure #	Distribution	Sample Size	Censoring Rate	Parameters
3.12	Exponential	300, 50, 25	83%, 7%, 3%	rate = 1.4, 1.2, 1.0
3.13	Exponential	25, 50, 300	5%, 7%, 87%	rate = 1.4, 1.2, 1.0
3.14	Heavy Tail	300, 50, 25	85%, 15%, 10%	rate = 1.4, 1.2, 1.0
3.15	Heavy Tail	25, 50, 300	10%, 16%, 89%	rate = 1.4, 1.2, 1.0
3.16	Gamma	300, 50, 25	85%, 15%, 10%	scale = 1.4, 1.2, 1.0; shape = 2.0, 2.0, 2.0
3.17	Gamma	25, 50, 300	10%, 16%, 89%	scale = 1.4, 1.2, 1.0; shape = 2.0, 2.0, 2.0
3.18	Log-Normal	300, 50, 25	85%, 15%, 10%	mean = 1.0, 1.2, 1.4; SD = 2.0, 2.0, 2.0
3.19	Log-Normal	25, 50, 300	10%, 15%, 85%	mean = 1.0, 1.2, 1.4; SD = 2.0, 2.0, 2.0
3.20	Weibull	300, 50, 25	85%, 15%, 11%	scale = 1.0, 1.2, 1.4; shape = 2.0, 2.0, 2.0
3.21	Weibull	25, 50, 300	10%, 15%, 85%	scale = 1.0, 1.2, 1.4; shape = 2.0, 2.0, 2.0

Estimators Under a Exponential Distribution with Unequal Censoring Rates

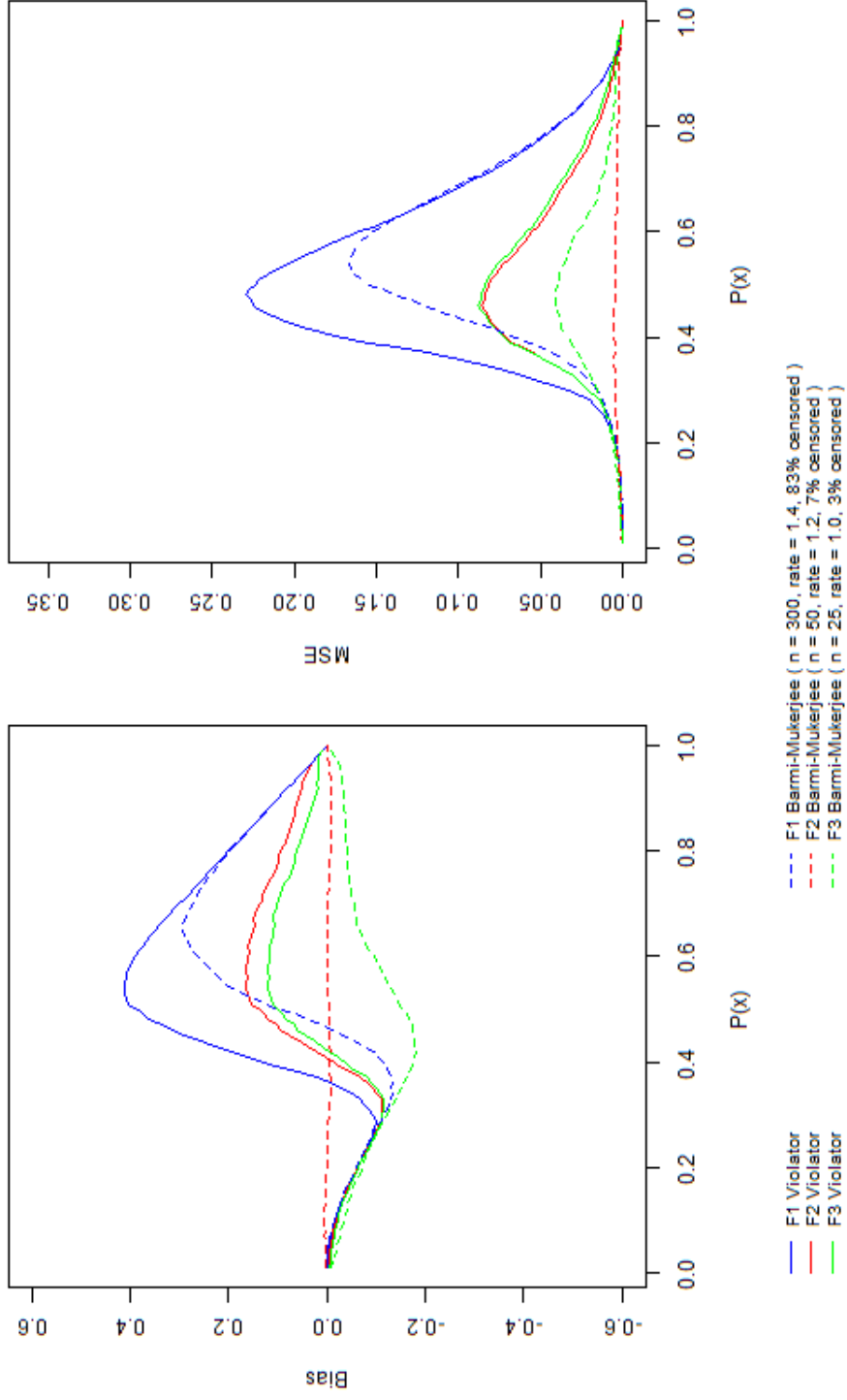


Figure 3.12: MSE for (1.7) does better than (2.1) when the stochastically smallest distribution has a higher censoring rate than the other distributions.

Estimators Under a Exponential Distribution with Unequal Censoring Rates

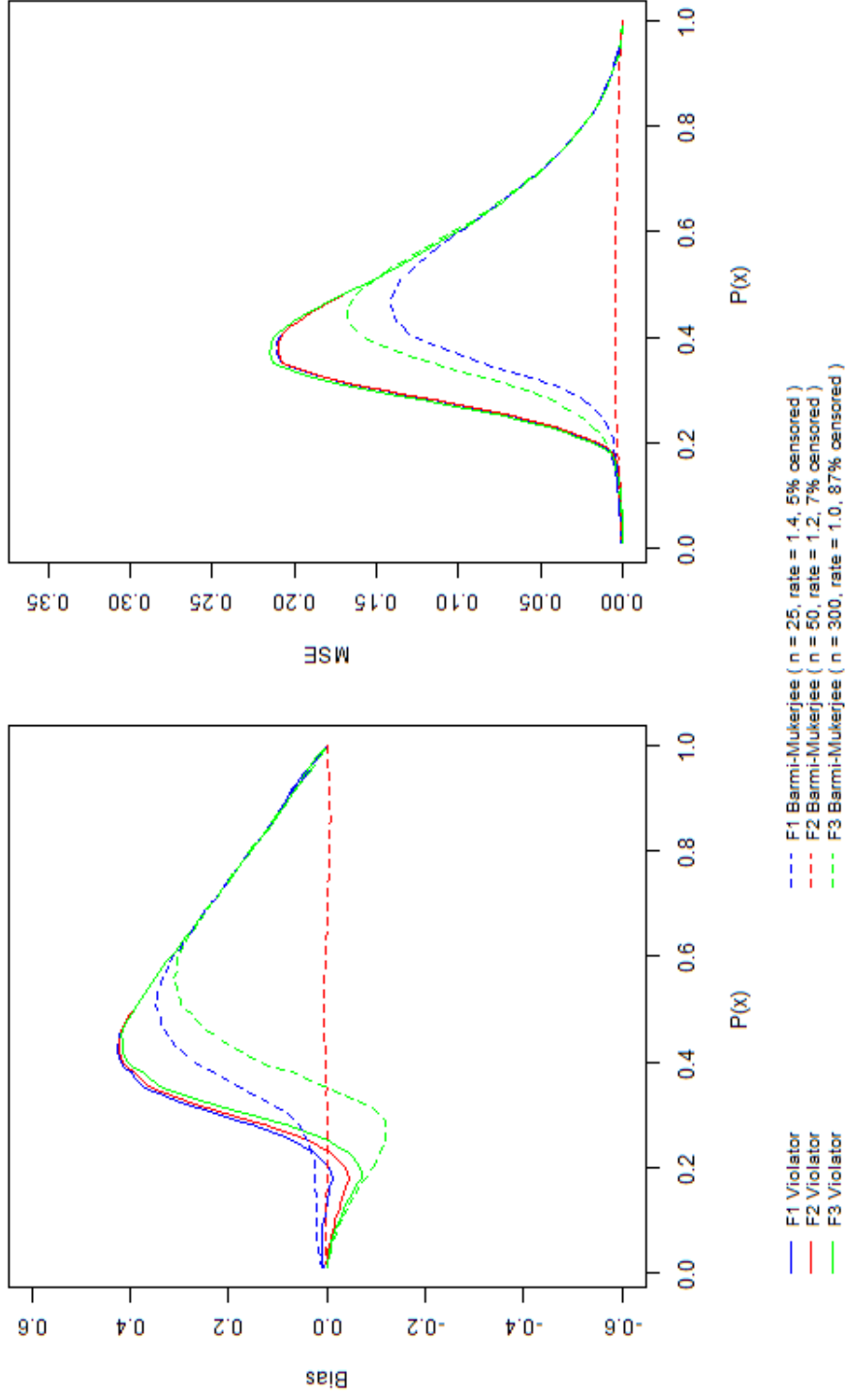


Figure 3.13: MSE for (1.7) does better than (2.1) when the stochastically largest distribution has a higher censoring rate than the other distributions.

Estimators Under a Heavy Tail Distribution with Unequal Censoring Rates

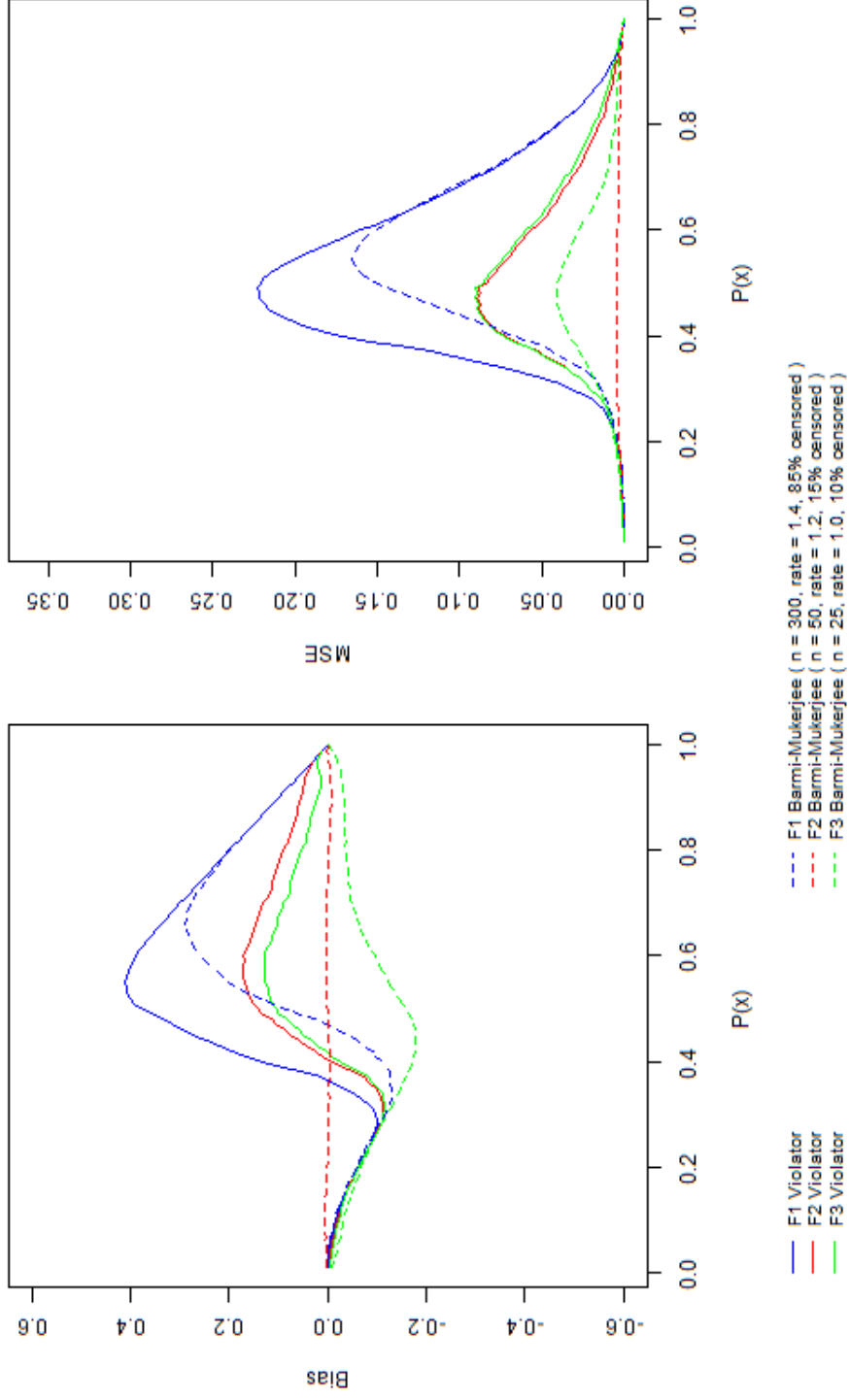


Figure 3.14: MSE for (1.7) does better than (2.1) when the stochastically smallest distribution has a higher censoring rate than the other distributions.

Estimators Under a Heavy Tail Distribution with Unequal Censoring Rates

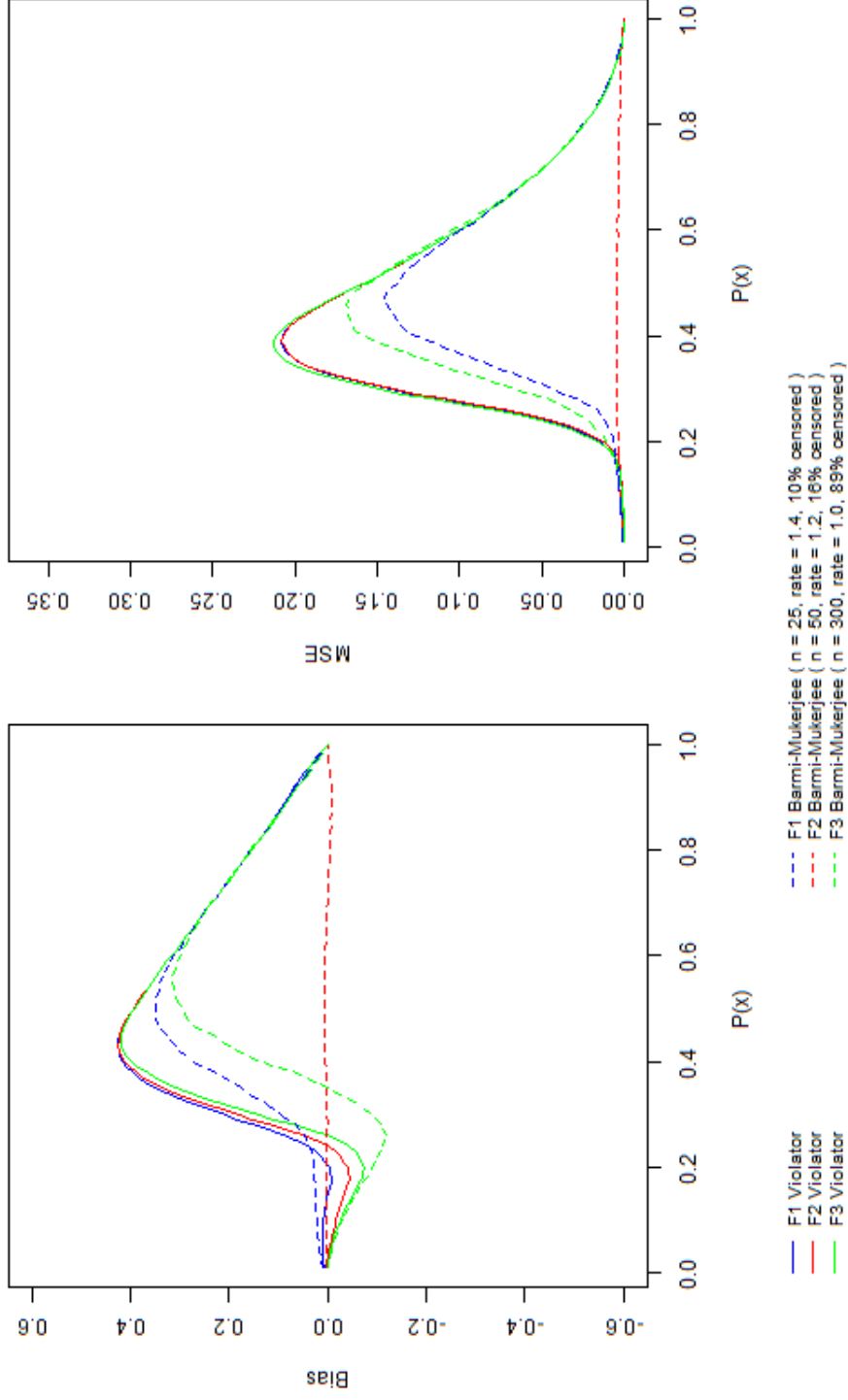


Figure 3.15: MSE for (1.7) does better than (2.1) when the stochastically largest distribution has a higher censoring rate than the other distributions.

Estimators Under a Gamma Distribution with Unequal Censoring Rates

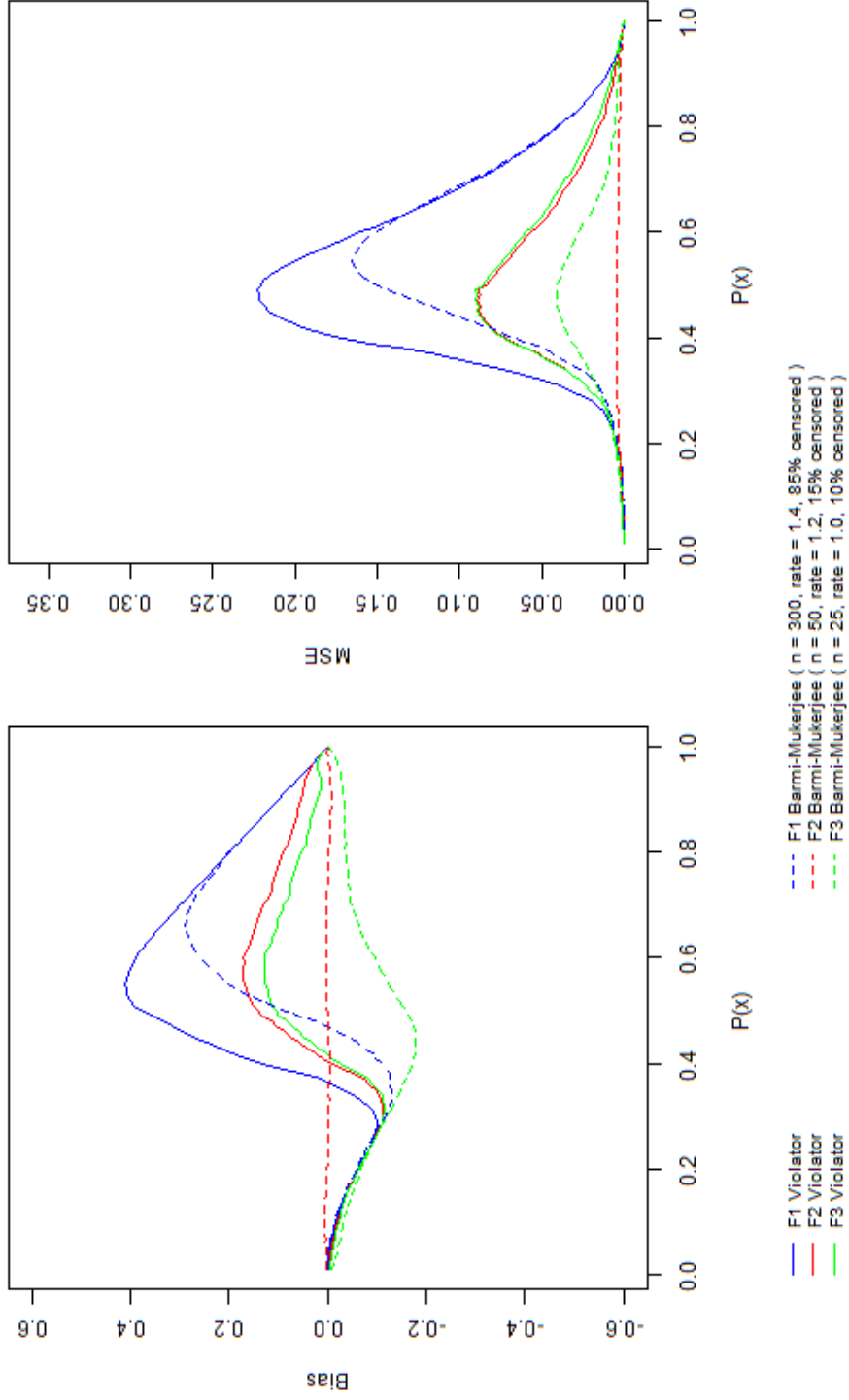


Figure 3.16: MSE for (1.7) does better than (2.1) when the stochastically smallest distribution has a higher censoring rate than the other distributions.

Estimators Under a Gamma Distribution with Unequal Censoring Rates

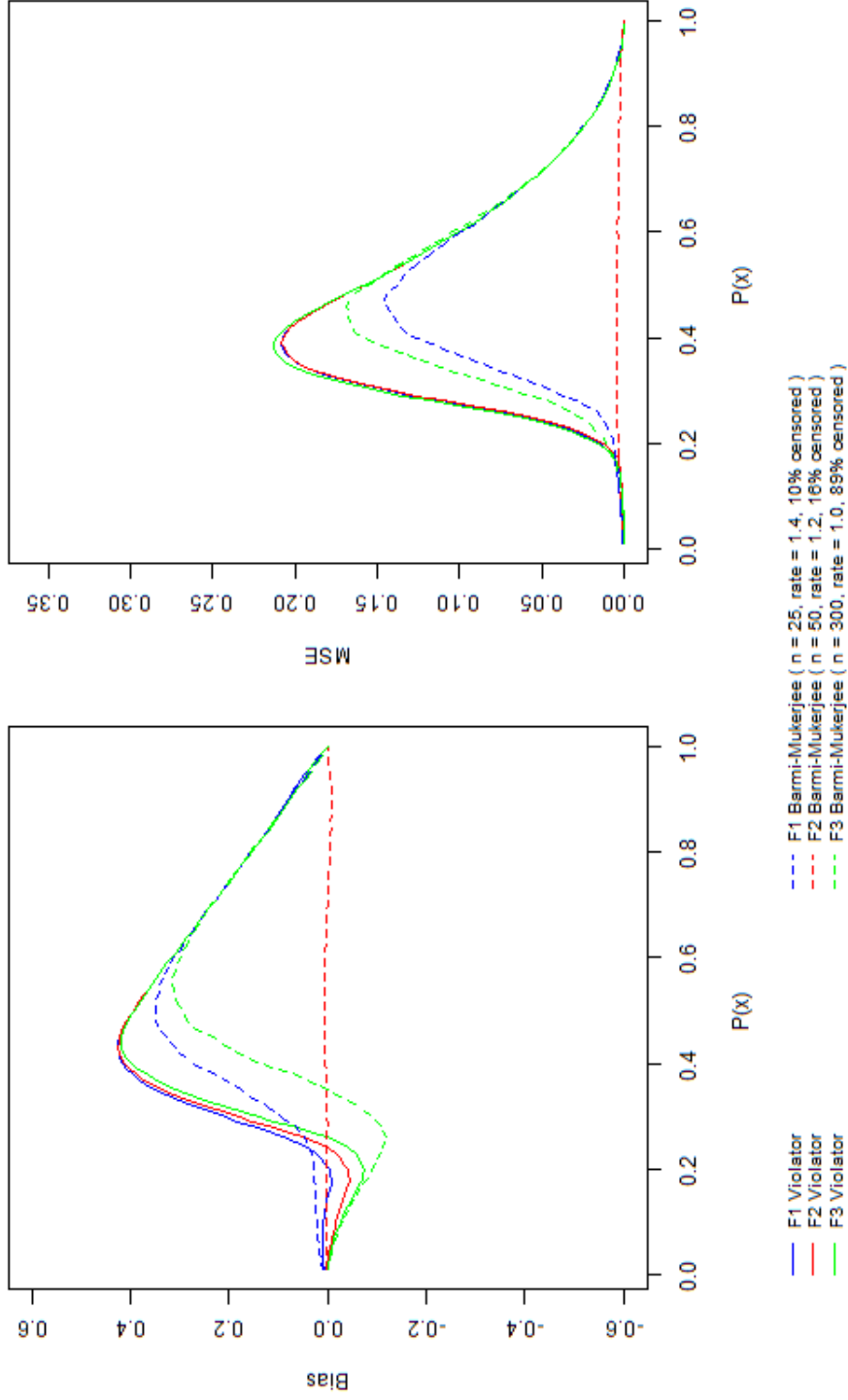


Figure 3.17: MSE for (1.7) does better than (2.1) when the stochastically largest distribution has a higher censoring rate than the other distributions.

Estimators Under a Log-Normal Distribution with Unequal Censoring Rates

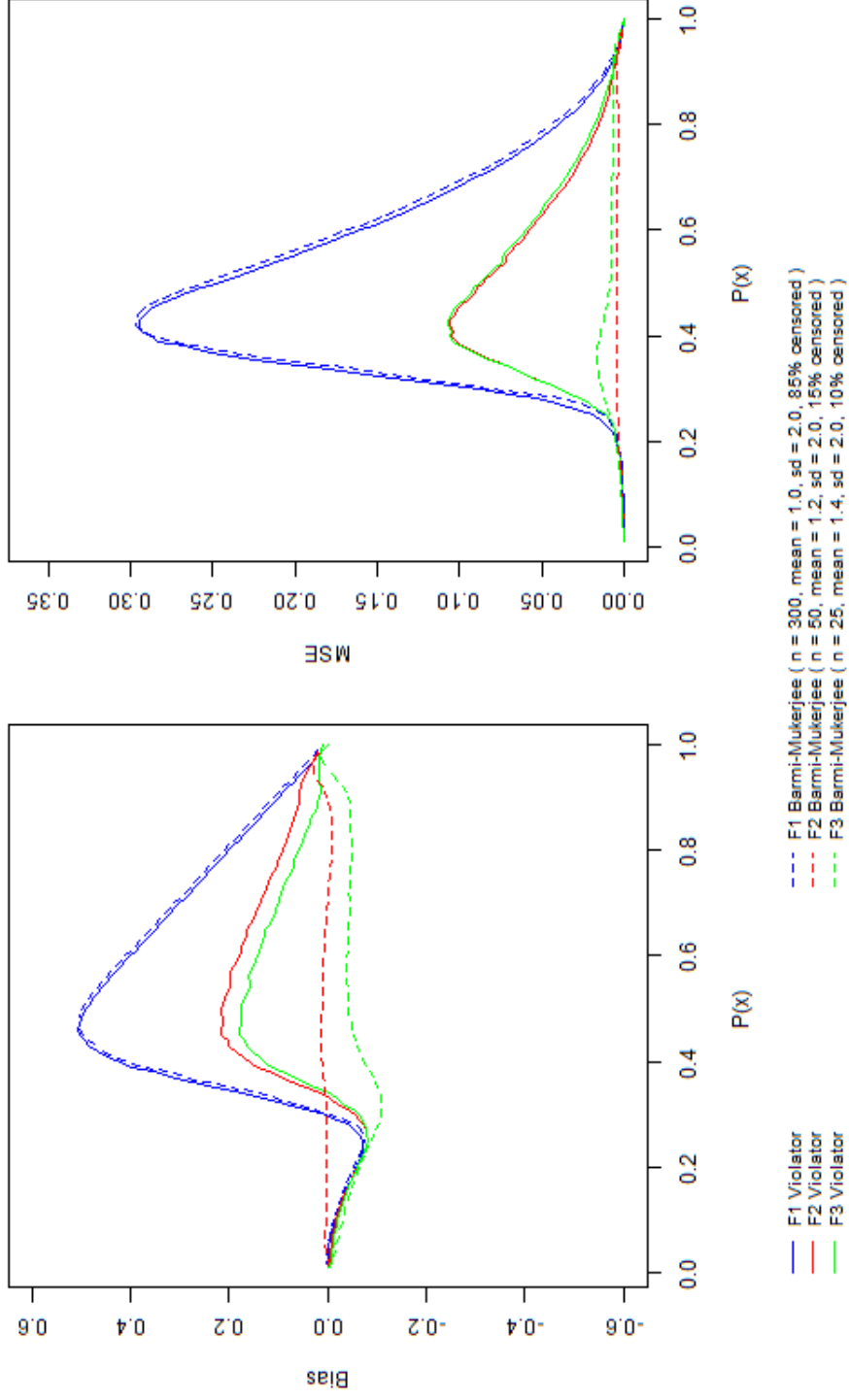


Figure 3.18: MSE for (1.7) does better than (2.1) when the stochastically smallest distribution has a higher censoring rate than the other distributions.

Estimators Under a Log-Normal Distribution with Unequal Censoring Rates

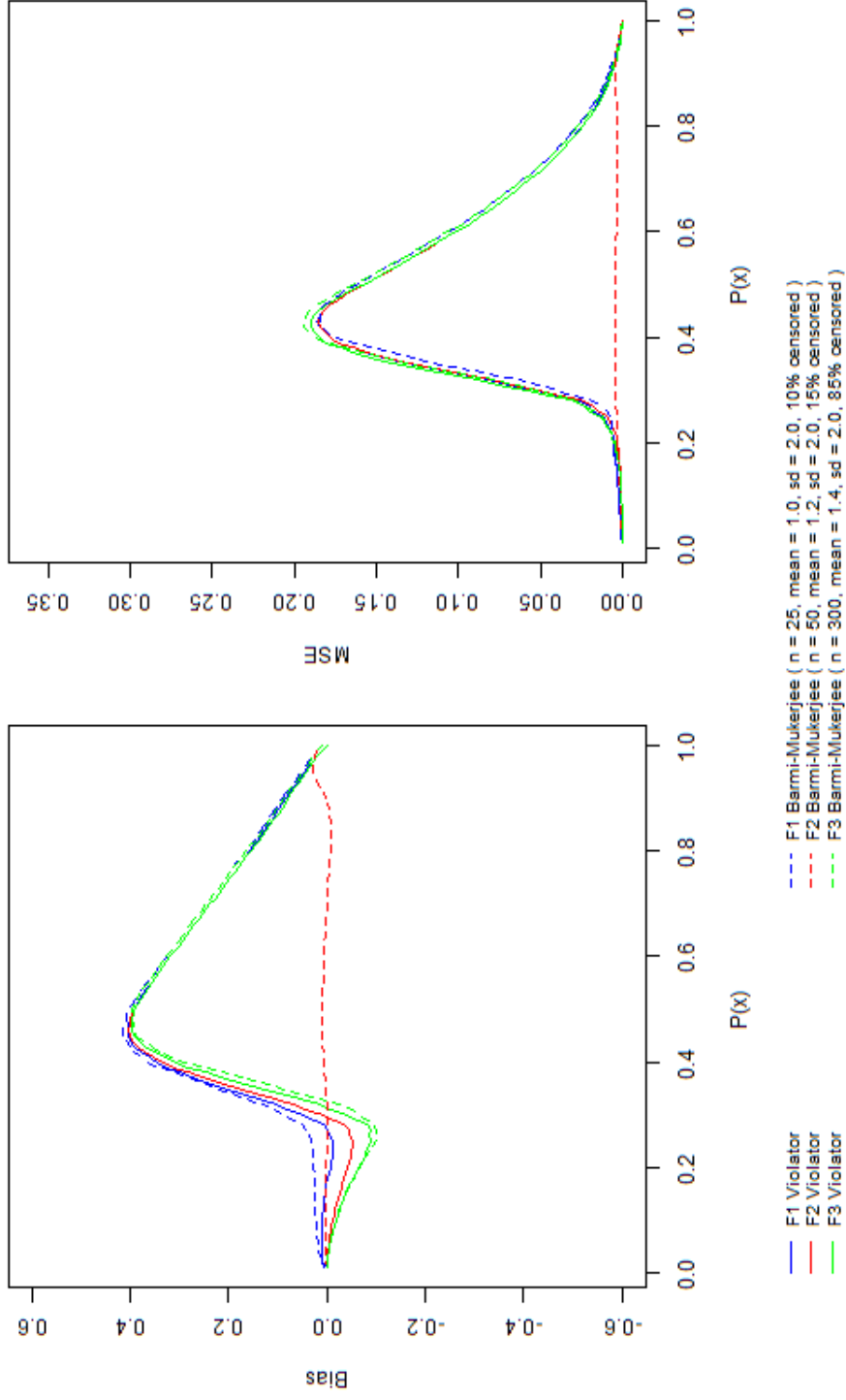


Figure 3.19: MSE for (1.7) does better than (2.1) when the stochastically largest distribution has a higher censoring rate than the other distributions.

Estimators Under a Weibull Distribution with Unequal Censoring Rates

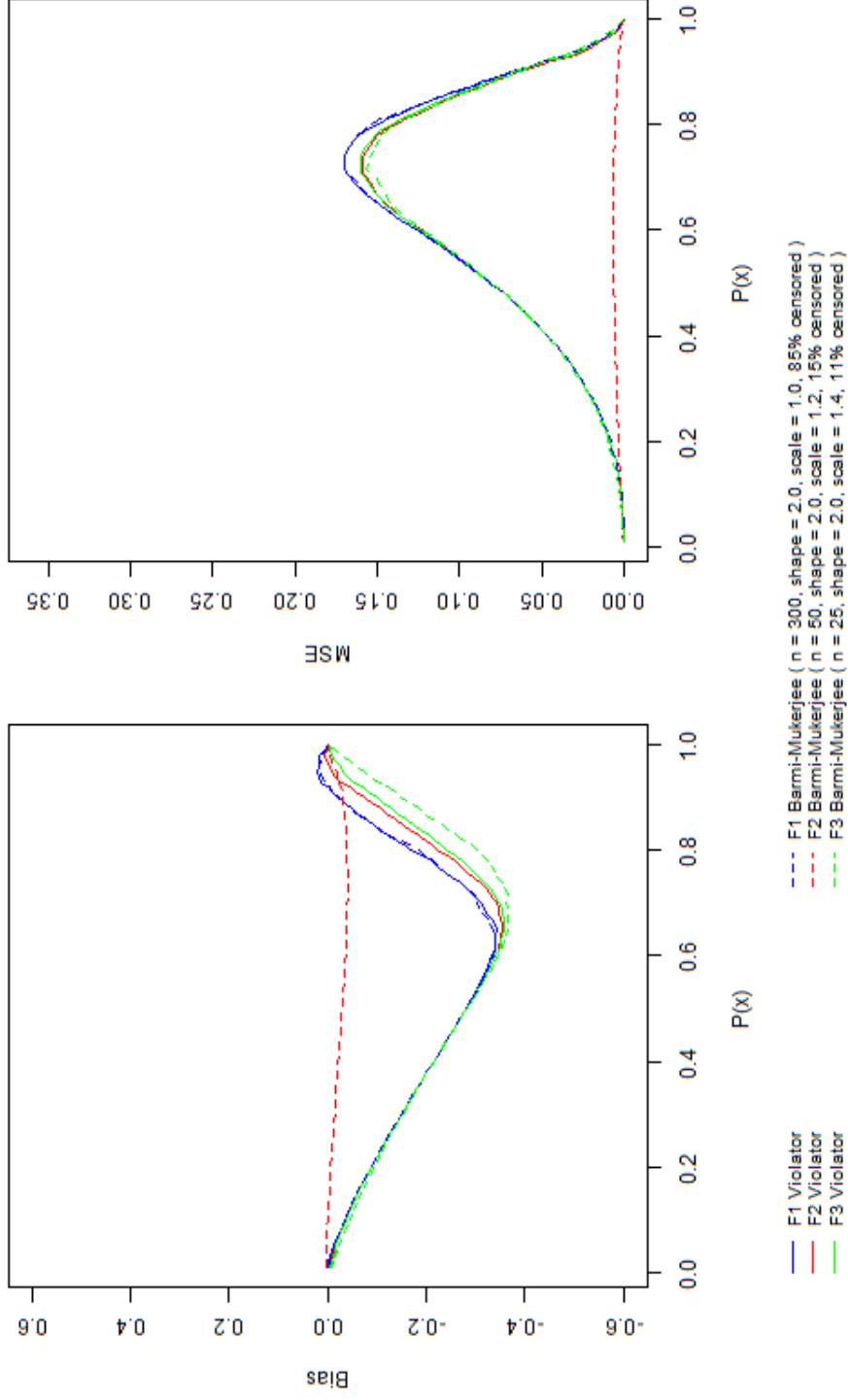


Figure 3.20: MSE for (1.7) does better than (2.1) when the stochastically smallest distribution has a higher censoring rate than the other distributions.

Estimators Under a Weibull Distribution with Unequal Censoring Rates

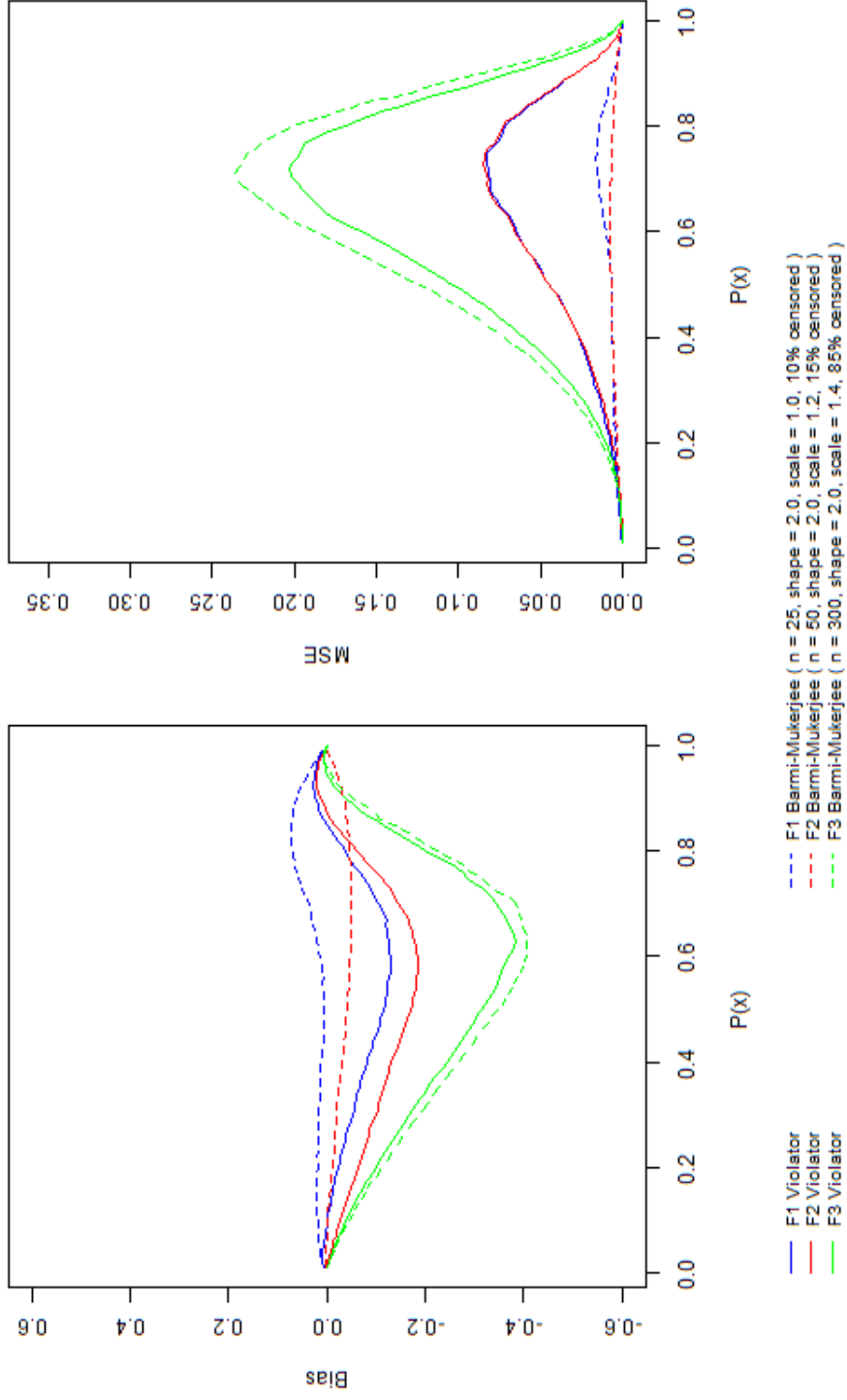


Figure 3.21: MSE for (1.7) does better than (2.1) when the stochastically largest distribution has a higher censoring rate than the other distributions.

Chapter 4

Conclusions

All in all, the simulations conducted in this study reveal that the estimator proposed performs well under equally censored distributions. In particular, (2.1) performs better than (1.7) for low censoring rates of exponential and heavy tail distributed survival times. However, (2.1) performs worse than (1.7) for high censoring rates for exponential and heavy tail distributed survival times. The estimator proposed in this study also outperforms Barmi and Mukerjee's estimator for distributions with equal censoring rates that are distributed under the log-normal, gamma, and Weibull distributions, regardless of censoring rate.

Although (2.1) has superior performance over (1.7) for equal censoring, (2.1) performs significantly poorer than (1.7) for many cases of unequal censoring. Unequal censoring often causes violation in stochastic ordering. Since (2.1) takes the weighted average of all distributions violating stochastic ordering, heavy censoring rates on the stochastically smallest or largest survival functions result in constant stochastic ordering violation.

4.1 Future Directions

Although the estimator presented has been simulated in numerous environments varying in terms of distribution, censoring rate, and sample size, there are still many situations to consider. Exploring new scenarios to test (2.1) may illuminate exactly why it performs better than (1.7) in some situations and performs worse in other situations.

The estimator presented in this thesis was mainly compared against Barmi and Mukerjee's estimator. It would be interesting to compare (2.1) against Park's estimator (Park et al., 2012). It is known how Park's estimator performs against Rojo's estimators for the two sample case. However, it is unknown if Park's estimators outperform the generalized case of Rojo's estimators presented in this document.

BARI Investigators. (2007). The final 10-year follow-up results from the BARI randomized trial. *Journal of the American College of Cardiology*, 49(15), 1600-1606. doi:10.1016/j.jacc.2006.11.048.

Barmi, H. E., Mukerjee, H. (2005). Inferences under a stochastic ordering constraint: The k-sample case. *Journal of the American Statistical Association*, 100(469), 252-261. doi:10.1198/016214504000000764.

Brown, J. B. W., Hollander, M., and Korwar, R. M. Nonparametric Tests of Independence for Censored Data, with Applications to Heart Transplant Studies. In *Reliability and Biometry: Statistical Analysis of Lifelength*, F. Proschan and R. J. Serfling, eds. Philadelphia: SIAM, 1974, pp. 327-354.

Ciuca, V., Matei, M. (2010). 5th WSEAS International Conference on Economy and Management Transformation. In *Survival Analysis for the Unemployment Duration*. Retrieved from <http://www.wseas.us/e-library/conferences/2010/TimisoaraW/EMT/57.pdf>

Diermeier, D., Keane, M., Merlo, A. (2005). A political economy model of congressional careers. *The American Economic Review*, 95(1), 347-373. doi:10.1257/0002828053

Dykstra, R. L. (1982). Maximum likelihood estimation of the survival functions of stochastically ordered random variables. *Journal of the American Statistical Association*, 77(379), 621-628.

Dykstra, R. L., Feltz, C. J. (1989). Nonparametric maximum likelihood estimation of survival functions with a general stochastic ordering and its dual. *Biometrika*, 76(2), 331. doi:10.2307/2336667

Dykstra, R., Kochar, S., Robertson, T. (1995). Likelihood based inference for cause specific hazard rates under order restrictions. *Journal of Multivariate Analysis*, 54(2), 163-174. doi:10.1006/jmva.1995.1050

- Efron, B. The Two Sample Problem with Censored Data. In *Proceedings of the Fifth Berkeley Symposium On Mathematical Statistics and Probability*. New York: Prentice-Hall, (1967): 4, 831-853.
- Feltz, C. J., Dykstra, R. L. (1985). Maximum likelihood estimation of the survival functions of N stochastically ordered random variables. *Journal of the American Statistical Association*, 80(392), 1012-1019.
- Gill, R.D. Censoring and Stochastic Integrals. *Mathematical Centre Tracts*. Amsterdam: Mathematisch Centrum, 1980, 124.
- Helmers, C., Rogers, M. (2008). *Innovation and the survival of new firms across British regions*.
- Kalleberg, A. L., Leicht, K. T. (1991). Gender and organizational performance: Determinants of small business survival and success. *The Academy of Management Journal*, 34(1), 136-161.
- Kaplan, E. L., Meier, P. Nonparametric Estimation from Incomplete Observations. *Journal of the American Statistical Association* 53 (1958): 457-481.
- Klein, J. P., 1950-2013, Moeschberger, M. L. (1997). *Survival analysis: Techniques for censored and truncated data*. New York: Springer.
- Lehmann, E. L. (1955). Ordered families of distributions. *The Annals of Mathematical Statistics*, 26(3), 399-419. doi:10.1214/aoms/1177728487.
- Lo, S. H. (1987). Estimation of distribution functions under order restrictions, *Statistics and Decisions*, 5, 252-262.
- Logubayom, I. A., Luguterah, A. (2013). Survival Analysis of Time to First Birth after Marriage. *Research on Humanities and Social Sciences*, 3, 117-125.

- Moeschberger, M. L., Klein, J. P. (1985). A comparison of several methods of estimating the survival function when there is extreme right censoring. *Biometrics*, 41(1), 253-259.
- Muenchow, G. (1986). Ecological use of failure time analysis. *Ecology*, 67(1), 246-250.
- Park, Y., Jeremy M G Taylor, Kalbfleisch, J. D. (2012). Pointwise nonparametric maximum likelihood estimator of stochastically ordered survivor functions. *Biometrika*, 99(2), 327-343. doi:10.1093/biomet/ass006
- Pflueger, M. (2011). *Electrical Submersible Pump Survival Analysis* (Unpublished master's thesis). Department of Statistics, Texas AM, College Station.
- Rojo, J. (1995). On the weak convergence of certain estimators of stochastically ordered survival functions. *Journal of Nonparametric Statistics*, 4(4), 349-363. doi:10.1080/10485259508832626
- Rojo, J. (2004). On the estimation of survival functions under a stochastic order constraint. *Lecture Notes-Monograph Series*, 44, 37-61.
- Rojo, J., Ma, Z. (1996). On the estimation of stochastically ordered survival functions. *Journal of Statistical Computation and Simulation*, 55(1-2), 1-21.