ÇUKUROVA UNIVERSITY
INSTITUTE OF NATURAL AND APPLIED SCIENCES

PHD THESIS

RESAMPLING METHODS AND THEIR APPLICATIONS IN LIU REGRESSION

Mustafa DENİZ

Department Of Statistics

August, 2024

ÇUKUROVA UNIVERSITY

INSTITUTE OF NATURAL AND APPLIED SCIENCES

PhD THESIS ACCEPTANCE

RESAMPLING METHODS AND THEIR APPLICATIONS IN LIU REGRESSION

We certify that the thesis titled above was reviewed and approved by unanimity / the majority of votes for the award of the degree of the Philosophy of Doctorate by the board of jury on 19/08/2024.

Jury	Prof. Dr. Selahattin KAÇIRANLAR (Supervisor):	…….
	Prof. Dr. Gülesen ÜSTÜNDAĞ ŞİRAY:	…….
	Prof. Dr. Hüseyin GÜLER:	…….
	Assoc. Prof. Dr. Murat GENÇ:	…….
	Assoc. Prof. Dr. Erkut TEKELİ:	…….

PhD Thesis is written at the Statistics Department of Institute of Natural and Applied Sciences of Çukurova University.

Registration Number:

Prof. Dr. Sadık DİNÇER

Director

Institute of Natural and Applied Sciences

CONTENTS

ÇUKUROVA UNIVERSITY

INSTITUTE OF NATURAL AND APPLIED SCIENCES

PhD THESIS

RESAMPLING METHODS AND THEIR APPLICATIONS IN LIU REGRESSION

Mustafa DENİZ

Supervisor: Prof. Dr. Selahattin KAÇIRANLAR

Department Of Statistics

ABSTRACT

Resampling methods are used to provide estimates relying on the underlying data, which helped make estimates for statistics that could not have been before, or double check the estimates obtained using other estimation methods.

In this thesis, the application of resampling methods on the linear regression, especially the Liu estimator (LE), is explored. First, diverse resampling methods are discussed. Next, the application of the resampling methods as a tool to estimate the confidence interval (CI) of the Liu regression parameters (LPs) is proposed. After that, the Bootstrap method and cross-validation are used to find the optimal value of the Liu biasing parameter, and in the last chapter the resampling methods are employed to show the effects of different methods of estimating the Liu parameter (LP) on the CI that are estimated using different CI estimation methods. In the conclusion, a comprehensive discussion of the key findings and conclusions is provided.

Key Words: Liu estimator, multicollinearity, Bootstrap, Jackknife, resampling methods.

ÇUKUROVA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

DOKTORA TEZİ

LIU REGRESYONDA YENİDEN ÖRNEKLEME METOTLARI VE UYGULAMALARI

Mustafa DENİZ

Danışman: Prof. Dr. Selahattin KAÇIRANLAR
İstatistik Anabilim Dalı

ÖZ

Yeniden örnekleme yöntemleri, daha önce tahmin yapamadığımız istatistikler için tahminler yapmamıza veya diğer tahmin yöntemlerini kullanarak elde ettiğimiz tahminleri kontrol etmemize yardımcı olan temel verilere dayanan tahminler elde etmek için kullanılır.

Bu tezde yeniden örnekleme yöntemlerinin lineer regresyona, özellikle de Liu tahmin edicisine uygulanması üzerinde çalışılmıştır. Öncelikle farklı yeniden örnekleme yöntemlerinden bahsedilmiştir. Ardından Liu regresyon parametrelerinin güven aralıklarını tahmin etmek için yeniden örnekleme yöntemlerinin uygulanması önerilmiştir. Daha sonra, Liu yanlılık parametresinin en uygun değerini bulmak için önyükleme ve çapraz doğrulama yöntemi kullanılmıştır ve son bölümde, Liu parametresini tahmin etmeye yönelik farklı yöntemlerin, farklı güven aralık tahmin yöntemleri kullanılarak tahmin edilen güven aralıkları üzerindeki etkilerini göstermek için yeniden örnekleme yöntemleri kullanılmaya çalışılmıştır. En sonunda, temel bulgular ve sonuçlara ilişkin bir tartışmaya yer verilmiştir.

Anahtar Kelimeler: Liu tahmin edici, Çoklu iç ilişki, Önyükleme, Jackknife, Yeniden örnekleme metotları.

EXTENDED ABSTRACT

The resampling methods are considered a very useful tool for nonparametric estimation. Many literatures discussed their application in various fields of statistics, which also included the application of the Bootstrap and Jackknife methods for linear regression. Many other ideas of resampling methods were proposed, such as the cross-validation method which helped evaluate and develop some estimates, and the permutation tests which are used to conduct nonparametric statistical tests of statistical hypotheses. These applications established the resampling methods as one of the crucial proposed statistical methods in recent years.

In this thesis, the application of the resampling methods in the linear regression, especially for the LE, is addressed, which is applied by adopting the proposed Bootstrap and Jackknife methods, and their former applications in other fields of statistics to be applicable for the LE. The thesis concentrates on the estimation of the CI of the regression parameters using several methods that are used for estimating the CIs and finding the optimal value of the Liu regularization parameter (LP) using the Bootstrap method and the mean squared error (MSE) criterion, in addition to showing the effect of different methods of estimating the LP on the CIs that are estimated using various estimation methods.

The first chapter starts by introducing the concept of resampling methods and their application in statistics and linear regression. This introduction is the result of a literature review that shows some previous papers and academic works that used Bootstrap and Jackknife methods along with other resampling methods in the linear regression.

The second chapter gives fundamental information about resampling methods, their history, and their applications in the numerous fields of statistical science, like the estimation of some statistics, the estimation of bias, and its use to extract the bias-corrected estimator, the estimation of standard error (SE), the CI estimation, and the use of the Bootstrap method for hypotheses testing, with a comparison between the permutation tests and Bootstrap's statistical tests. After that, the advantages and disadvantages of the Bootstrap and Jackknife methods are discussed and summarized in a clear comparison between these two resampling methods. The fundamental information that is proposed in this chapter is crucial for our application of the resampling methods in the next chapters.

The third chapter aims to develop and compare different CIs for regression coefficients (RCs) based on the LE using Bootstrap and Jackknife methods. The used CI estimation methods are; the normal method, the Studentized-t method, the Jackknife method, the Percentile method, and the bias corrected and accelerated (BC_a) method. For comparison purposes, two criteria are used, which are the confidence width (CW) and the coverage probability (CP), which are calculated using the Bootstrap and Jackknife methods in a simulation study for the data that suffers from the problem of multicollinearity. This chapter concludes with a discussion of the conclusions, as it

concludes that according to the CW criteria, the Normal method demonstrates superiority over the ordinary least squares estimator (OLSE), while according to the CP criteria, there is at least one Bootstrap method (Normal, Percentile, Studentized-t and BC_a) that gives better results than the OLSE. On the other hand, the Jackknife method does not give good results compared to the other resampling methods.

Chapter four develops and compares different techniques for the selection of the LP using Bootstrap and cross-validation methods. The methods used for estimating the LP are the CL (first letters of Colin Lingwood Mallows's name) method, the CLh (an adaptation of CL method) method, and the PR method, in addition to proposing the Bootstrap method to find the optimal value, and the results of which are compared to the previously mentioned methods and the OLSE method. The procedure is illustrated by application to published real data sets, and a simulation study comparing the several methods is also described. Analysis strategies are suggested for data with various degrees of multicollinearity and varying levels of signal-to-noise ratio (SNR). It is concluded that the selection of the LP seriously affects the results of the LE, and the proposed Bootstrap method to select the optimal value of the LP guarantees selecting a better d value. The Bootstrap estimate of the LP emerges as a better value, because even if it gives the best results according to one criterion (MSE criterion in this chapter), it does not necessarily mean that these results are the best results according to another criterion, which is presented in chapter five.

Chapter five aims to show the effect of various estimation methods of the LP on the CI of the regression parameters that are estimated using different CIs estimation methods, which are the Normal, the Studentized-t, the percentile, the BC_a, and the Bootstrap methods. The outputs are compared using two criteria, the average CW, and the CP of those CIs. The simulation study uses a variation of several factors to show the different outputs according to a variety of cases, these factors were the *SNR*, the $\beta'\beta$ value, and the multicollinearity level. It has been found that there are many factors that affect the regression estimates, where the effects of the *SNR* and $\beta'\beta$ values are significantly high, and the multicollinearity level's effect is medium. Also, it is found that the different Liu estimates according to different methods of calculating the LP show different reactions to these three factors ($\beta'\beta$, *SNR* and ρ^2). In addition to that, it has been shown that using different methods of estimating the CI may give different results, knowing that some estimates (CLh and PR) have an advantage over the other estimates in most cases and for most methods used for estimating the CI.

The thesis is finalized by a discussion of the results of the applications that have been proposed in the previous chapters, in order to give the overall findings and provide recommendations for the application of the resampling methods in the LE, which is proposed in chapter six.

GENİŞLETİLMİŞ ÖZET

Yeniden örnekleme yöntemleri parametrik olmayan tahminler için çok yararlı bir araçtır. Literatürdeki birçok çalışmada, lineer regresyon için önyükleme ve Jackknife yöntemlerinin uygulanmasını da içeren, istatistiğin birçok alanında yeniden örnekleme yöntemlerinin uygulamaları incelenmiştir. Bazı tahminlerin hesaplanmasına ve geliştirilmesine yardımcı olan çapraz doğrulama yöntemi ve istatistiksel hipotezlerin parametrik olmayan istatistiksel testini yürütmek için kullanılan permütasyon testleri gibi yeniden örnekleme yöntemlerine ilişkin birçok fikir önerilmiştir. Bu uygulamalar, yeniden örnekleme yöntemlerini son yıllarda önerilen en kullanışlı istatistiksel yöntemlerden biri haline getirmiştir.

Bu tezde, bilinen önyükleme ve Jackknife yöntemlerini uyarlayarak, lineer regresyonda özellikle Liu tahmin edicisi için yeniden örnekleme yöntemlerinin uygulanması üzerinde çalışılmıştır. Tezde, güven aralık tahmin etmek için çeşitli yöntemler kullanarak regresyon parametrelerinin güven aralıklarının tahmin edilmesine odaklanılmıştır. Önyükleme ve ortalama kare hatası kriterlerinin kullanılmasıyla Liu tahmin edicisinin yanlılık parametresinin optimal tahmininin bulunması ve ek olarak birçok tahmin metodunun kullanılmasıyla yanlılık parametresinin güven aralıkları üzerindeki etkileri incelenmiştir.

İlk bölüme yeniden örnekleme yöntemleri ve bunların lineer regresyondaki uygulamalarına bir giriş yaparak başlıyoruz. Lineer regresyonda önyükleme, Jackknife ve diğer yeniden örnekleme yöntemlerini kullanan önceki bazı makaleler ve akademik çalışmaları gösteren bir literatür taraması yoluyla yapılmıştır.

İkinci bölümde yeniden örnekleme yöntemleri, bunların tarihçesi ve istatistik biliminin farklı alanlarındaki uygulamaları hakkında temel bilgiler verilmektedir. Örneğin, bazı istatistiklerin tahmini, yanlılık tahmini ve düzeltilmiş yanlılık tahmin edicinin bulunması, standart hatalar, güven aralık tahmini ve hipotez testi için önyükleme yöntemlerinin kullanımı, permütasyon testleri ile önyükleme istatistiksel testleri arasında bir karşılaştırma sunar. Daha sonra önyükleme ve Jackknife yöntemlerinin avantajları ve dezavantajları bu iki yeniden örnekleme yöntemi arasında net bir karşılaştırma yapılarak tartışılmış ve özetlenmiştir. Bu bölümde önerilen temel bilgiler, sonraki bölümlerdeki yeniden örnekleme yöntemlerini uygulamamız açısından önemli olmuştur.

Üçüncü bölüm, önyükleme ve Jackknife yöntemlerinin kullanılmasıyla Liu tahmin edicisine dayalı regresyon katsayıları için farklı güven aralıklarını geliştirmeyi ve karşılaştırmayı amaçlamaktadır. Kullanılan güven aralık tahmin yöntemleri: normal dağılım yöntemi, studentized-t yöntemi, Jackknife yöntemi, yüzdelik yöntemi ve BC_a yöntemidir. Karşılaştırma amacıyla, çoklu bağlantı problemi olan veriler için bir simülasyon çalışmasında önyükleme ve Jackknife yöntemleri kullanılarak hesaplanan CW ve CP olmak üzere iki kriter kullanılmıştır. Bu bölümün sonunda, Normal yöntemin OLS tahmin ediciye göre bir üstünlüğü olduğu sonucuna varılmış, ancak diğer önyükleme yöntemlerinin (Yüzdelik, studentized-t ve BC_a) OLSE'ye göre bir avantajı olmadığı

sonucuna varılmıştır. Jackknife yöntemi de diğer yeniden örnekleme yöntemlerine göre iyi sonuçlar vermemiştir.

Dördüncü bölüm, önyükleme ve çapraz doğrulama yöntemlerini kullanarak Liu regresyon parametresinin seçimine yönelik farklı teknikleri geliştirir ve karşılaştırır. Liu parametresini tahmin etmek için kullanılan yöntemler CL yöntemi, CLh yöntemi ve PR yöntemidir; ek olarak optimal değeri bulmak için önyükleme yöntemi önerilmiş, önermenin yanı sıra daha önce bahsedilen yöntemler ve OLSE yöntemi karşılaştırılmıştır. Prosedür, yayınlanmış gerçek veri setlerine uygulanarak örneklendirilmiş, ek olarak bir simülasyon çalışması ile çeşitli yöntemler karşılaştırılmıştır. İç ilişkinin çeşitli dereceleri ve değişen sinyal-gürültü oranına sahip veriler için analiz stratejileri önerilmiştir. Liu parametresinin seçiminin Liu tahmincisinin sonuçlarını ciddi şekilde etkilediği ve Liu parametresinin optimal değerini seçmek için önerilen önyükleme yönteminin daha iyi d değerinin seçilmesini garanti ettiği sonucuna varılmıştır. Liu parametresinin önyükleme tahmini daha iyi bir değer olarak adlandırılmıştır. Çünkü bu bölümde ele alınan MSE kriterine göre en iyi sonuçları verse bile, başka bir kritere göre en iyi sonuçların elde edilemeyeceği bilinmektedir. Detaylar beşinci bölümde gösterilmiştir.

Beşinci bölüm, normal, studentized-t, yüzdelik ve önyükleme yöntemleri gibi farklı güven aralık tahmin yöntemlerinin kullanılmasıyla, tahmin edilen güven aralıkları üzerinde, Liu tahmin edicisinin yanlılık parametresinin farklı tahmin yöntemlerinin etkisini göstermeyi amaçlamaktadır. Sonuçlar, ortalama güven genişliği ve bu güven aralıklarının kapsanma olasılığı olmak üzere iki kriter kullanılarak karşılaştırılmıştır. Simülasyon çalışmasında, çeşitli durumlara göre farklı çıktıları göstermek için birçok faktörün değişimi kullanmıştır; bu faktörler SNR, $\beta'\beta$ değeri ve iç ilişkinin düzeyidir. Regresyon tahminlerini etkileyen birçok faktörün olduğu, SNR ve $\beta'\beta$ değerlerinin etkisinin anlamlı düzeyde yüksek, iç ilişkinin düzeyinin etkisinin ise orta düzeyde olduğu tespit edilmiştir. Ayrıca, Liu parametresini hesaplamak için kullanılan farklı yöntemlere göre elde edilen farklı Liu tahminlerinin bu üç faktöre ($\beta'\beta$, SNR ve ρ^2) farklı tepkiler gösterdiği görülmüştür. Buna ek, güven aralıklarını tahmin etmek için kullanılan farklı yöntemlerin farklı sonuçlar verebileceği gösterilmiştir.

Tez, yeniden örnekleme yöntemlerinin Liu tahmin edicide uygulanmasına yönelik öneriler sunmak amacıyla önceki bölümlerde önerilen uygulamaların sonuçlarının tartışılmasıyla sonuçlandırılmıştır.

ACKNOWLEDGEMENT

I begin by thanking ALLAH the Almighty for granting me success in all stages of my PhD study, and previous stages of my education journey.

Then I extend my sincere gratitude to my supervisor, Dr. Selahattin Kaçıranlar, who provided all the necessary support in all stages of working on my PhD and was patient with me in all the circumstances and challenges I faced while working on my PhD.

I also thank my mother and father, without whom I would not have reached where I am. They provided me with full support and intellectual awareness to persevere to complete all my academic stages.

I conclude by thanking my dear wife, who stood by my side during the nights I spent studying and was patient with my negligence towards her due to my preoccupation with my education.

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

OLSE	:	Ordinary Least Squares Estimator
LE	:	Liu Estimator
LP	:	Liu Regression Parameter
RE	:	Ridge Estimator
RRE	:	Ridge Regression Estimator
RC	:	Regression Coefficient
MSE	:	Mean Squared Error
MSEP	:	Mean Squared Error Prediction
EMSE	:	Estimated Mean Squared Error
AURE	:	Almost Unbiased Ridge Estimator
JRE	:	Jackknife Ridge Estimator
BC_a	:	The Bias Corrected and Accelerated Method
ABC	:	Approximated Bootstrap Confidence Interval
JGLE	:	Jackknifed Generalized Liu Estimator
GCV	:	Generalized Cross-Validation
CW	:	Confidence Width
CP	:	Coverage Probability
CI	:	Confidence Interval
SE	:	Standard Error
SNR	:	Signal-to-noise Ratio
IDF	:	Inverse Distribution Function
AULE	:	Almost unbiased Liu Estimator
PR	:	Predicted Residual Sum of Squares method
CL	:	First letters of Colin Lingwood Mallows's name
CLh	:	An adaptation of CL method
CV	:	Coefficient of Variation

MAIN CHAPTERS OF THE THESIS

1. INTRODUCTION

This chapter gives an introduction to the ideas that is discussed the theses and what is the motivation to work on this topic.

2. RESAMPLING METHODS

This chapter gives fundamental information about the resampling methods, their history, and their applications in the numerous fields of statistical science, and the advantages and disadvantages of the Bootstrap and Jackknife methods, and finally gives a comparison of these two resampling methods. This fundamental information is important for our application of the resampling methods in the next chapters.

3. CI OF THE LE'S RCs USING JACKKNIFE AND BOOTSTRAP METHODS

In this chapter, different CIs for RCs based on the LE are developed and compared using Bootstrap and Jackknife methods. For comparison, CW and the CP are calculated through a simulation study of the data that suffers from the problem of multicollinearity. The last section presents a discussion of the conclusions.

4. BOOTSTRAP SELECTION OF THE LP

In this chapter, various techniques for the selection of the LP using Bootstrap and cross-validation methods are developed and compared. The procedure is illustrated by application to published real data sets. A simulation study comparing several techniques is also described. Analysis strategies are suggested for data with various degrees of multicollinearity and varying levels of SNR.

5. SELECTION OF LP USING THE CI AND BOOTSTRAP METHODS

This chapter aims to propose a method that helps choose the optimal value of the LP of the LE, by showing the effect of different estimation methods on the CI of the regression parameters. The outputs will be compared using two criteria; the average CW, and the CP of several methods of estimating the CI (Normal, Studentized-t, Percentile, and Bootstrap methods).

6. RESULTS AND SUMMARY

The thesis will be finalized by a discussion of the results of the applications that have been proposed in the previous chapters, to give the overall findings and provide recommendations for the application of the resampling methods in the LE.

REFERENCES

CURRICULUM VITAE

1. INTRODUCTION

The estimation of the regression parameters faces the uncertainty issue just like all other statistical methods. Statistical methods use the available data as a sample to make estimates or to provide expectations of a studied phenomenon, or the data of the past to forecast the values of the parameters in the future. This uncertainty issue is caused by the fact that not all the data of the studied society is usually available. Therefore, a sample that mostly represents it is used, which means whatever the sample is used, it would not be 100% representing the society, or when forecasting the future is attempted, means whatever the type of the data of the past is, it also would not give 100% accuracy.

All the efforts exerted by statistical scientists were centered on increasing the accuracy level of the statistical methods that are used to calculate the anticipated measures. Increasing the accuracy level will never mean 100% representation, but it means finding methods that give results that converge to 100%.

Another paradox that makes it more complicated is that even the criteria that are used to evaluate the statistical methods or to compare different methods do not give 100% accuracy of results, as they also face the uncertainty issue.

The above discussed dilemma means that perfect methods will never be achieved in statistics. which is the reason why our work is limited to developing the methods that already exist, or inventing new methods, striving to have better results and trying to reach the situation which gives the best possible results.

One of the methodologies that has been used to improve the statistical methods and the criteria to evaluate and compare statistical methods is Bootstrapping. It is like you are stuck in a lake and you want to get up from the water, but you do not have anything to use, so you hold the bootstraps of your shoes to pull yourself up. The Bootstrap method generates an almost unlimited number of samples from the original sample by withdrawing with replacement. It uses the sample itself to know more about the sample, like having better statistical estimators and more information about the accuracy of the estimators (Efron and Tibshirani, 1993).

Just like the Bootstrap method, the Jackknife method was also used for application in statistics. The Jackknife relies on the idea of taking out one observation and conducting the estimation, then repeating the same process for all sample observations.

These two methods (Bootstrap and Jackknife) are known as the resampling methods, which are getting more applicable with the development of computing power in recent years, and their applications in statistics have become very popular and are getting wider over time.

The first noticed use of the resampling methods was by Fisher in 1935, who used the permutation tests, which helped improve hypothesis testing. Later, the Jackknife method was proposed by Quenouille (1949), who used it to improve an estimator by correcting its bias.

Afterward, Tukey (1958) proposed wider uses of the Jackknife, the most important of which was to estimate the SE of an estimator. Tukey was the first to give this method the name Jackknife.

In 1979 in his published paper "Bootstrap Methods: Another Look at the Jackknife", Efron (1979) proposed the use of resampling methods in a more practical manner, whereas he used the Monte Carlo samples to take several samples of size n (the same size of the original sample) from the original sample. While as an idea, it has been used before Efron by Simon (1969), who can be called the inventor of the Bootstrap, Simon's purpose was to use it as a teaching tool for statistics and probability that did not rely on the abstract concept of a parametric probability model to produce the original sample (Chernick and LaBudde, 2011). Efron was the first to use Bootstrap as a replacement, or as a new method instead of the Jackknife method to estimate the SE, and then Efron recognized the broad use of the Bootstrap method in estimating the CI, the bias, hypothesis testing and more complex problems. Which was proposed by Efron (1982), and Efron and Tibshirani (1986).

Gong (1986) discussed the use of Bootstrapping methods in the logistic regression's subset selection. Efron and Tibshirani (1993) discussed the use of the Bootstrapping in the linear regression and compared the Bootstrapping of the residuals and the Bootstrapping of the observations, then proposed the use of the resampling methods to estimate the CIs for some parameters, as one of them was the CI of the parameters of the linear regression estimator.

In linear regression, the use of biased estimation techniques brings the obligation of selecting the biasing parameters of these estimators, as well. Consequently, numerous methods for selecting the biasing parameters have been studied, (see for example; Hoerl et al., 1975; Mc Donald and Galerneau, 1975; Hoerl and Kennard, 1976; Lawless and Wang, 1976; Hocking, 1976; Dempster et. Al, 1977; Wichern and Curchill, 1978; Golub et al., 1979; Gibbons, 1981; Nordberg, 1982; Tekeli et al., 2019). One important method that has been given by Delaney and Chatterjee (1986) uses the Bootstrap method to choose the optimal ridge parameter, which demonstrated that the Bootstrap choice of the ridge parameter leads to a smaller mean squared error of prediction than the ridge trace method. It also concluded that the benefits of the usage of the Bootstrap method to choose the ridge's biasing parameter include its less subjective nature, ease of implementation, and robustness. Furthermore, Chaubey et al. (2018) proposed the use of resampling methods to estimate the CI of the RE in the linear regression. They used residuals Bootstrapping and compared several methods to estimate the CI of the linear regression's parameters; the normal, the studentized-*t*, the Jackknife, the percentile, the BC_a, and the approximated Bootstrap CI (ABC) methods. After using the Monte Carlo simulation, they discussed the differences between these methods, by comparing the results relying on two criteria; the CW, and the CP.

Crivelli et al (1995) compared the effects of different methods of calculating the ridge's biasing parameter k on the accuracy of the RE by showing the results of many methods of estimating the CI of the regression parameters that are estimated by the RE and tested it using two

criteria. Their study combined the Bootstrap method with the Edgeworth expansion to estimate the CI. Following that, Özkale and Altuner (2022) expanded the study to be conducted on 11 methods of estimating k, while Crivelli et al (1995) conducted it on 3 methods.

Our aim in this thesis is to widen the application of the resampling methods in the linear regression, precisely for the LE, which will help obtain estimates for some statistics that cannot be obtained using the traditional statistical methods and have more accurate estimates using the resampling methods for cross-validation of the results obtained using the other estimation methods.

2. RESAMPLING METHODS

2.1. Introduction

To make an estimation in the traditional statistical methods, the distribution of the data should be previously known or assumed, and then the estimate of the statistic of interest is simulated according to the properties of this distribution. One challenge with this method is that enough information may not be available to make the assumption of the distribution, and a wrong assumption may be made, leading to the selection of a distribution that is far from the distribution of the data. This means that all the estimates extracted will not be reliable. The other challenge is that the traditional methods do not propose estimates for many statistics like the SE of many statistics, which places restrictions on the expansion of statistical applications.

The resampling methods are a good solution for the previously discussed challenges, knowing that they were not proposed at the beginning as a solution for these challenges, they were developed with time to become the way are used nowadays. In addition, the development of computer capacities in the last couple of years made the resampling methods more applicable, which is why their wide range of applications in various fields can be noticed.

The main proposed resampling methods are the Monte Carlo simulation method, the permutation method, the Jackknife method, the Bootstrap method, and the cross-validation method. These methods are used as a solution for many statistical issues like the estimation of SE, bias, CI, hypothesis testing, percentiles and quantiles, and reliability testing and validation, in addition to many other estimates. Knowing that there are two main types of resampling methods, which are the parametric methods and the nonparametric methods, where, as can be understood from their names, the first one uses the distribution of the data to generate the subsamples, while the latter uses the sample itself to generate them.

The key advantage of the resampling methods is that they helped make estimates that were impossible or difficult to have before. The non-parametric methods provided an additional advantage which is they do not need information about the distribution of the underlying data, which is not available in most cases in real life. In addition, the brilliant feature of the resampling methods is that they are applicable for a very wide range of cases, with simplicity of application, making them useful for statistical application.

2.2. Jackknife Resampling Method

The Jackknife method was first proposed by Maurice Quenouille in 1949. It can be said that it is the first proposed resampling method. Quenouinlle used it as a method to improve an estimator by applying bias correction. It suggests estimating the bias and then extracting the improved estimator by taking out the bias from the estimator (Chernick and LaBudde, 2011). After that, the Jackknife method was used to give estimates of many other statistics.

The advantages of this method include that it helps estimate statistics that were difficult or impossible to estimate, like the bias, the SE, and the CI of any parameter. Another advantage of this method is that it does not need any information or prior assumptions about the distribution of the data.

The Jackknife method is very simple, it depends on the idea of leave-one-out, where it sequentially deletes the observations of the original sample to have sub-samples of size $n-1$, then uses them to estimate the statistic of interest. Later, Efron and many other researchers proposed more advanced Jackknife methods, like the delete-d-method which is proposed to overcome the disadvantages of the leave-one-out Jackknife method.

2.2.1. Leave-one-out Jackknife Method

Suppose a sample $x = (x_1, x_2, \dots, x_n)$ and an estimator $\hat{\theta} = s(x)$. The leave-one-out Jackknife method uses the sample to calculate the bias of the estimator $\hat{\theta}$ by extracting the sub-samples by deleting the i^{th} observation to have $x_{(i)}, i = 1,2, \dots, n$ that are called the Jackknife samples, where:

$$x_{(i)} = (x_1, x_2, \dots, x_{i-1}, x_{i+1}, \dots, x_n) \tag{2.1}$$

Then calculates the estimator $\hat{\theta}$ from the Jackknife samples to get the Jackknife estimators $\hat{\theta}_i$ (Efron, 1996). This means:

$$\hat{\theta}_{(i)} = s(x_{(i)}) \tag{2.2}$$

Is the i^{th} Jackknife estimation of the estimator $\hat{\theta}$.

As an illustration, let θ be the sample mean that needs to be estimated using the Jackknife method, and let the sample be given as follows:

$$x_1 = 2, x_2 = 4, x_3 = 4, x_4 = 5, x_5 = 6, x_6 = 9 \tag{2.3}$$

The mean of the original sample is $\bar{x} = 5$.

Now let the Jackknife samples be taken, where the first one is taken by deleting the first observation $x_1 = 2$, to have the sample:

$$x_1^* = 4, x_2^* = 4, x_3^* = 5, x_4^* = 6, x_5^* = 9 \tag{2.4}$$

The mean of this Jackknife sample is $\bar{x}_1^* = 5.6$.

To get the second Jackknife sample, the second observation $x_2 = 4$ is deleted, and then the Jackknife estimate of the mean for the second sample is calculated.

By continuing through all $n = 6$ sample observations, all Jackknife samples and their estimates of the mean are obtained, which is given in the following table:

Table 2.1. Mean values of Jackknife samples

Jackknife sample	x_1^*	x_2^*	x_3^*	x_4^*	x_5^*	$\bar{x}_i^*$
1	4	4	5	6	9	5.6
2	2	4	5	6	9	5.2
3	2	4	5	6	9	5.2
4	2	4	4	6	9	5
5	2	4	4	5	9	4.8
6	2	4	4	5	6	4.2

Finally, the overall Jackknife estimate of the mean is obtained by taking the average of the Jackknife estimates of all Jackknife samples, which is:

$$\bar{x}_{(.)}^* = \sum_{i=1}^{6} \bar{x}_i^* = 5 \tag{2.5}$$

In this example, the sample mean, it is found that the Jackknife estimate of the mean equals the mean of the original sample. This case is not necessarily the same for all statistics, but eventually, the good Jackknife estimate should converge to the value of the statistic of interest.

2.2.2. Jackknife Estimate of Bias

The bias of the estimate $\hat{\theta}$ is defined as:

$$Bias = E_F\theta(\hat{F}) - \theta(F) \tag{2.6}$$

Where F is the data distribution function.

While usually the value of the statistic of interest is unknown, the bias of the estimate may not be able to be calculated, therefore, as mentioned above, Quenuille (1949) proposed the Jackknife method as a solution, by using it to give an estimate of the statistic of interest θ, which is proposed as the average of the Jackknife estimates $\hat{\theta}_i$:

$$\hat{\theta}_{(.)} = \sum_{i=1}^{n} \hat{\theta}_{(i)}/n \tag{2.7}$$

Now by replacing $\theta(F)$ with $\hat{\theta}_{(.)}$ in (2.6), the Jackknife estimate of the bias becomes as follows:

$$\widehat{bias}_{jack} = (n-1)\left(\hat{\theta}_{(.)} - \hat{\theta}\right) \tag{2.8}$$

Finally, this leads to the calculation of the bias-corrected Jackknife estimate of θ (Efron, 1996):

$$\tilde{\theta} = \hat{\theta} - \widehat{bias} = n\hat{\theta} - (n-1)\hat{\theta}_{(.)} \tag{2.9}$$

2.2.3. Jackknife's Pseudo Values

The pseudo value of Jackknife represents the value that comes against the i^{th} observation as follows:

$$\tilde{\theta}_i = n\hat{\theta} - (n-1)\hat{\theta}_{(i)} \tag{2.10}$$

Where $\hat{\theta}_{(i)}$ is the estimate of θ within the deletion of the i^{th} observation.

Simply, the Jackknife's pseudo values can be described as the estimates of each individual observation that is calculated by subtracting the value of the Jackknife estimate corresponding to the i^{th} observation from the overall Jackknife estimate of the parameter of interest.

By taking $\hat{\theta} = \bar{X}$, With simple calculations, it can be found that the sample average is:

$$\tilde{\theta}_i = n\hat{X} - (n-1)\hat{X}_{(i)} =$$
$$n\left(\sum_{j=1}^{n} x_j/n\right) - \frac{(n-1)(x_1, x_2, \dots, x_{i-1}, x_{i+1}, \dots, x_n)}{n-1} = x_i \tag{2.11}$$

The average is a special case, where the Jackknife's pseudo values are equal to the sample observations themselves:

$$\tilde{\theta}_i = x_i, \quad i = 1,2, \dots, n \tag{2.12}$$

The importance of the pseudo values comes from the fact that the bias corrected estimate of any parameter equals the average of the pseudo values, where:

$$\frac{1}{n}\sum \tilde{\theta}_i = n\hat{\theta} - (n-1)\bar{\theta}$$
$$= \hat{\theta} - \widehat{\text{bias}} \quad (2.13)$$

Hence:

$$\tilde{\theta} = \frac{1}{n}\sum \tilde{\theta}_i \quad (2.14)$$

2.2.4. Jackknife Estimate of SE

Tukey (1958) proposed the Jackknife estimate of the SE from the Jackknife's bias corrected estimate and the Jackknife's pseudo values of the parameter of interest $\tilde{\theta}_i$ which is:

$$\widehat{se}_{jack} = \sqrt{\widehat{var}} = \sqrt{\frac{n-1}{n}\sum_{i=1}^{n}(\tilde{\theta}_i - \tilde{\theta})^2} \quad (2.15)$$

It is clear that the Jackknife estimate of SE uses the same concept of the normal standard deviation, with the replacement of the sample observations by the Jackknife pseudo values $\tilde{\theta}_i, i = 1,2, \dots, n$, and the average of these pseudo values instead of the sample average (Miller, 1974).

It can be noticed that the normal standard deviation is multiplied by $1/(\mathrm{n}-1)$, while the Jackknife estimate of the SE is multiplied by $(\mathrm{n}-1)/\mathrm{n}$. This difference comes from the fact that the size of each Jackknife sample is $n-1$ not n, and the deviation of the Jackknife samples is smaller than the deviation of the Bootstrap samples (Efron and Tibshirani, 1993). Next, it will be shown that the Bootstrap estimate of the SE is multiplied by $1/(\mathrm{n}-1)$ just like the normal standard deviation.

2.2.5. Jackknife Estimate of CI

While the Jackknife estimates of SE and the Jackknife estimate of the parameter of interest are presented, it is very easy to estimate the CI using the Jackknife method, which is:

$$\left[\tilde{\theta} - t_{n-1}^{(1-\alpha)}\widehat{se}_{jack}, \tilde{\theta} + t_{n-1}^{(1-\alpha)}\widehat{se}_{jack}\right] \quad (2.16)$$

Where t represents the student distribution with $n-1$ degrees of freedom, which is given as $t = (\tilde{\theta}_i - \tilde{\theta})/\widehat{se}_{jack}$, and $t_{n-1}^{(1-\alpha)}$ is the percentile value of the probability corresponding to $(1-\alpha)$ of the student distribution (Efron and Tibshirani, 1993).

2.2.6. Hypothesis Testing Using Jackknife Method

The statistical tests and the CI reflect the same concept. It can be considered that if the value of the statistic of interest is out of the CI, this means the null hypothesis is rejected and the alternative hypothesis is accepted. This gives the advantage that while the CI can be estimated using the Jackknife method, it can also be used for statistical testing. However, after the literature review, it seems that the Jackknife method is not discussed as a tool for hypothesis testing. On the other hand, the Bootstrap method is presented as a hypothesis testing tool.

2.2.7. Weakness of the Jackknife

With a deeper look at the Jackknife method, it can be noticed that it suffers from a catastrophic issue, which appears in a specific type of statistics; the non-smooth statistics, although it gives better results for the smooth statistics.

As a definition of the smoothness terminology, it can be said that the smooth statistic is the statistic that will have a small change when making small changes in the data, while the non-smooth statistic does not change when making small changes in the data. In other words, smoothness is all about the differentiability of the function.

The clearest examples of smooth statistics and non-smooth statistics are the mean and the median. The mean is a smooth statistic because any small change in the data causes a small change in its value, on the other hand, the median does not act the same way, as its value does not change when making small changes in the data (Shao and Tu, 1995).

For illustration, let the sample be given as follows:

7, 9, 12, 20, 21, 30, 34, 38, 41, 63, 78, 102, 104

Sample mean: $\bar{x} = 43$

Sample median: $\tilde{x} = 34$

The estimates of the mean and the median that correspond to each Jackknife sample are:

Table 2.2. Comparison of smooth and non-smooth statistics

Statistic	Deleted observation												
	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}	x_{13}
$\bar{x}$	46	45.8	45.6	44.9	44.8	44.1	43.8	43.4	43.2	41.3	40.1	38.1	37.9
$\tilde{x}$	36	36	36	36	36	36	34	32	32	32	32	32	32

And the pseudo values are:

Table 2.3. Pseudo values comparison of smooth and non-smooth statistics

Statistic	Pseudo value												
	$\tilde{x}_1$	$\tilde{x}_2$	$\tilde{x}_3$	$\tilde{x}_4$	$\tilde{x}_5$	$\tilde{x}_6$	$\tilde{x}_7$	$\tilde{x}_8$	$\tilde{x}_9$	$\tilde{x}_{10}$	$\tilde{x}_{11}$	$\tilde{x}_{12}$	$\tilde{x}_{13}$
$\bar{x}$	7	9	12	20	21	30	34	38	41	63	78	102	104
$\tilde{x}$	10	10	10	10	10	10	34	58	58	58	58	58	58

It can be noticed that the Jackknife estimates of the median and the pseudo values have only three values:

- Almost half of them $(\frac{n}{2} - 1)$ that are under the value of the sample median have the same value as the Jackknife estimate of the median.
- The observation in the middle has a unique corresponding value of the Jackknife estimate of the median.
- The rest of the values $(\frac{n}{2} - 1)$ have only one corresponding value of the Jackknife estimate of the median.

On the other hand, the Jackknife estimates of the mean differentiate for each Jackknife sample corresponding to the observation that was deleted from the original sample. This will cause the Jackknife estimation of the bias, the SE, and the CI of the non-smooth statistics to be significantly inconsistent, hence, the Jackknife is not recommended as an estimation method for the non-smooth statistic.

Finally, it is important to mention that this issue is not the same for all non-smooth statistics, as the smoothness level significantly varies from one statistic to another.

2.2.8. The Jackknife Delete-d-method

To overcome the weakness of the Jackknife method with the non-smooth statistics, it has been proposed that instead of deleting one observation at a time, d observations are deleted, where $1 \leq d \leq n$ (Efron and Tibshirani, 1993). This idea helped improve the differentiability of the Jackknife estimates for the non-smooth statistics (Shao and Tu, 1995).

Efron and Tibshirani (1993) discussed that if $n^{1/2}/d \to 0$ and $n - d \to 0$ a good estimate of the median can be obtained from the Jackknife delete-d method. In addition to the need to take $\sqrt{n} < d < n$ to have a good Jackknife estimate of the SE. In the meantime, the optimum value of $d > 1$ needed to obtain a consistent Jackknife estimate depends on the degree of non-smoothness

of the statistic, which means the more non-smoothness of the statistic the more observations are needed to be deleted in the Jackknife delete-d method (Shao and Tu, 1995).

This method has only one challenge; when the sample size n is large and the number of observations to be deleted d is also large, a huge number of Jackknife samples is obtained, which means a huge number of calculations. To address this challenge, it is proposed to take a sample from the available Jackknife samples, which returns to be almost the same as the Bootstrap method (Efron and Tibshirani, 1993), with one difference; that the Jackknife sample size is $n - d$, while the size of the Bootstrap samples remains the same as the original sample size.

2.3. Bootstrap Resampling Method

The Bootstrap is also a resampling method as the Jackknife, but it relies on a different way of resampling, which mainly uses a resampling method that is almost the same as the Monte Carlo simulation method (Efron, 1994), since the parametric Bootstrap method generates the samples from the assumed distribution of the original sample, the non-parametric Bootstrap method generates the samples directly from the original sample by withdrawing with replacement, and without the need to having any previous knowledge or making any assumptions about the distribution of the data (Davison and Hinkley, 1997). By using the Bootstrap method, as many samples as needed can be taken, while the number of available Bootstrap samples is very huge, as will be shown, this means that only a sample of the Bootstrap samples from the overall available Bootstrap samples is taken. The main benefit that makes the Bootstrapping method a brilliant idea is that the Bootstrap method delves into the dataset in order to extract more information that helps approximate the distribution of the data (Chernick and LaBudde, 2011).

Let a sample $x = (x_1, x_2, \dots, x_n)$ of size n be considered, and an estimator $\hat{\theta} = s(x)$, the Bootstrap sampling gives all original sample observations the same likelihood to be selected, which means each observation of the sample has the probability to be selected equals to $1/n$. The method generates many Bootstrap samples, let the number of the samples be B, all Bootstrap samples are of the same size n, which is the size of the original sample. The estimates of the statistic of interest θ are calculated for all Bootstrap samples $\hat{\theta}_i = s\left(x^*_{(i)}\right),\ i = 1, 2, \dots, B$, these estimates help get estimates for many statistics, like the CI, the bias, the SE…etc.

Let $x^*_1 = x_7, x^*_2 = x_5, x^*_3 = x_3, x^*_4 = x_{22}, \dots, x^*_n = x_7$, where $x^*_i, i = 1, 2, \dots, n$ refers to the i^{th} observation of the Bootstrap sample, which may be any observation of the original sample, in the meantime, the Bootstrap sample may have some of the observations that can be selected once, twice or more, or not be selected at all, since it uses withdrawing with replacement.

As an illustration, let the same example of the Jackknife method proposed in section 2.2.1 be considered. By generating six random Bootstrap samples:

Table 2.4. Mean values of Bootstrap samples

Bootstrap sample	x_1^*	x_2^*	x_3^*	x_4^*	x_5^*	x_6^*	$\bar{x}_i^*$
1	5	9	2	4	4	6	5
2	6	2	5	6	2	2	3.83
3	9	9	2	2	2	4	4.67
4	4	6	4	4	2	4	4
5	4	9	9	5	9	6	7
6	2	4	6	2	6	5	4.17

The last column shows the values of the Bootstrap estimates of the mean, from which the Bootstrap estimate of the mean is calculated, which is the average of all Bootstrap estimates:

$$\bar{x}_{(.)}^* = \sum_{i=1}^{6} \bar{x}_i^* = 4.78 \qquad (2.17)$$

It can be noticed that it is very close to the mean of the original sample, taking into consideration the small number of Bootstrap samples ($B = 6$). It has been proved that the Bootstrap estimate of a statistic converges to its value by taking $B \to \infty$ (Efron and Tibshirani, 1993). In our example by taking bigger numbers of Bootstrap samples, the following table shows the results:

Table 2.5. Convergence of estimates for different numbers of Bootstrap samples

Bootstrap sample size	$\bar{x}_i^*$
6	4.78
100	4.94
1,000	5.0065
10,000	4.988567
100,000	4.999047
1,000,000	5.000073

It is important to mention that the above shown Bootstrap method is called the non-parametric Bootstrap method, because it withdraws the Bootstrap samples directly from the original sample.

2.3.1. The Parametric Bootstrap Method

The parametric Bootstrap method can be easily described as a method that uses the available information about the distribution of the data to generate the Bootstrap samples. In other words, it does not generate the Bootstrap samples from the original sample, instead, it makes an assumption about the distribution of the data, and then generates the Bootstrap samples from the distribution itself. On the other hand, the nonparametric Bootstrap method uses the original sample to generate the Bootstrap samples, which means it is used without the need for any previous knowledge or assumptions about the distribution of the data.

Let:

$$se_{\hat{F}_{par}}(\hat{\theta}^*) \tag{2.18}$$

Indicates the parametric Bootstrap estimate of the SE of θ, where $\hat{F}$ is the estimate of F derived from a parametric model of the data. Also let:

$$se_B(\hat{\theta}^*) \tag{2.19}$$

Be the nonparametric estimate of the SE of θ, where B is the number of Bootstrap samples generated using the underlying data.

The difference between these two methodologies in calculating $se_{\hat{F}_{par}}(\hat{\theta}^*)$ and $se_B(\hat{\theta}^*)$, in the parametric and nonparametric estimates of the SE of θ appears only in the first steps, where after generating the Bootstrap samples, they follow the same steps.

The following table illustrates the two algorithms:

Nonparametric method	**Parametric method**
1. $\hat{F}$ the empirical distribution. 2. Generate B Bootstrap samples from the original sample $X^{*1}, X^{*2}, \dots, X^{*B}$. 3. Calculate the Bootstrap estimates of the parameter of interest (the mean in our example), $\hat{\theta}^*(1) = \bar{X}^{*1}, \hat{\theta}^*(2) = \bar{X}^{*2}, \dots, \hat{\theta}^*(B) = \bar{X}^{*B}$ 4. Calculate the Bootstrap estimate of SE: $\widehat{se}_B = [\sum_{b=1}^{B} \frac{[\hat{\theta}^*(b) - \hat{\theta}^*(.)]^2}{B-1}]^{1/2}$ where: $\hat{\theta}^*(.) = \sum_{b=1}^{B} \frac{\hat{\theta}^*(b)}{B}$	1. Make an assumption (or have previous information) about the distribution of the data. 2. Generate B Bootstrap samples from the distribution $X^{*1}, X^{*2}, \dots, X^{*B}$. 3. Calculate the Bootstrap estimates of the parameter of interest (the mean in our example), $\hat{\theta}^*(1) = \bar{X}^{*1}, \hat{\theta}^*(2) = \bar{X}^{*2}, \dots, \hat{\theta}^*(B) = \bar{X}^{*B}$. 4. Calculate the Bootstrap estimate of SE: $\widehat{se}_{par} = [\sum_{b=1}^{B} \frac{[\hat{\theta}^*(b) - \hat{\theta}^*(.)]^2}{B-1}]^{1/2}$ where: $\hat{\theta}^*(.) = \sum_{b=1}^{B} \frac{\hat{\theta}^*(b)}{B}$

It can be noticed that the nonparametric Bootstrap method has an advantage over the parametric method, as the first one does not need to have any assumptions about the distribution of the underlying data. This advantage can be noticed more in case of making wrong assumption of the distribution, which means errors in the generated Bootstrap samples.

It should be mentioned here that the parametric method is more useful in the case of the availability of information about the distribution of the data. However, when using traditional statistical methods, parametric assumptions need to be made to have the estimates of the statistics of interest. By using the nonparametric method, traditional statistical methods do not need to be used, which means the restriction of having parametric assumptions can be avoided (Efron and Tibshirani, 1993). In the meantime, even if information about the distribution of the data is available, the nonparametric method helps double-check the results of the parametric method (Chernick and LaBudde, 2011).

2.3.2. Bootstrap Estimate of SE

Let:

$$\hat{\theta}_{(i)} = s\left(x_{(i)}\right), \quad i = 1, \dots, B \tag{2.20}$$

Be the Bootstrap estimates of the parameter θ calculated from the Bootstrap samples. Using the same concept of the Jackknife estimate of SE, the Bootstrap estimate of SE will be given as:

$$\widehat{se}_B = \left[\frac{1}{B-1}\sum_{i=1}^{B}\left(\hat{\theta}_{(i)} - \hat{\theta}_{(.)}\right)^2\right]^{1/2} \tag{2.21}$$

where:

$$\hat{\theta}_{(.)} = \sum_{i=1}^{B}\hat{\theta}_{(i)}/B \tag{2.22}$$

is the average of the Bootstrap sample estimates of the statistic of interest.

2.3.3. Bootstrap Estimate of Bias

When estimating a parameter θ using the statistic $\hat{\theta} = s(X)$, the bias of an estimate is defined to be the difference between the expectation of the estimate $\hat{\theta}$ and the value of the parameter θ.

$$bias_F(\hat{\theta}, \theta) = E_F[s(X)] - t(F) \tag{2.23}$$

Where F is an unknown probable distribution of the sample $X = (x_1, x_2, \dots, x_n)$.

Using the Bootstrap, $bias_F$ can be estimated, this can be given by substituting $\hat{F}$ with F in (2.23), resulting in:

$$bias_{\hat{F}}(\hat{\theta}, \theta) = E_{\hat{F}}[s(X^*)] - t(\hat{F}) \tag{2.24}$$

The formula above shows that in order to get the Bootstrap estimate of the bias, the expectation of the estimate of the statistic of interest is needed, here comes the average of the Bootstrap estimates of the statistic of interest, which can be calculated by generating the Bootstrap samples $X^{*1}, X^{*2}, \dots, X^{*B}$, calculating the Bootstrap estimate for each sample $\hat{\theta}^* = s(X^{*b}), b = 1,2, \dots, B$, and then calculating the average of the Bootstrap estimates:

$$\hat{\theta}^*(\cdot) = \frac{\sum_{b=1}^{B} \hat{\theta}^*(b)}{B} = \frac{\sum_{b=1}^{B} s(X^{*b})}{B} \tag{2.25}$$

Finally, the Bootstrap estimate of the bias is taken by substituting $\hat{\theta}^*(\cdot)$ with $E_{\hat{F}}[s(X^*)]$:

$$\widehat{bias}_{\hat{F}} = \hat{\theta}^*(\cdot) - t(\hat{F}) \tag{2.26}$$

2.3.4. Bootstrap Estimate of CI

In many cases, the point estimate does not help that much since it only gives an estimate of the parameter without any further information. There are many statistics that give more information, that are usually used as complementary to the information provided by the point estimate. The CI comes as one of these statistics, as it gives the boundaries between which the parameter of interest may range according to a specific confidence level (Thomas and Efron, 1996).

Estimating the CI requires much information; the point estimate, the SE, and the inverse distribution function (IDF) of the data distribution. But as is known for most of the statistics, it is difficult or impossible to estimate the SE using the traditional statistical methods, this means the CI cannot be estimated too. To overcome this issue the Bootstrap method has been used to estimate se as shown before.

Let $\hat{\theta}$ and $\widehat{se}$ be the Bootstrap estimates of the parameter of interest θ and its SE respectively, the Bootstrap estimate of the CI of $\hat{\theta}$ is given as:

$$\left[\hat{\theta} - z_{\left(1-\frac{\alpha}{2}\right)}.\widehat{se}, \hat{\theta} + z_{\left(1-\frac{\alpha}{2}\right)}.\widehat{se}\right] \tag{2.27}$$

Where $z_{\left(1-\frac{\alpha}{2}\right)}$ is the IDF of the standard normal distribution, which represents the assumed distribution of θ, knowing that the assumption of the distribution of θ may be any other statistical distribution.

The problem with the above method is that sometimes enough information about the distribution of the data is not available, or an error is made by using a distribution that is far from its original distribution. That is why a different way to estimate the IDF is needed. Here comes the benefit of using the Bootstrap method, where the nonparametric Bootstrap method can be used to estimate the IDF without the need to rely on any assumptions about the distribution of the data. One of the suggested methods is known as the percentile-*t* CI, where the IDF estimate can be calculated by:

$$\#\frac{\{z^*(\mathrm{b}) \leq \hat{t}^{(\alpha)}\}}{B} = \alpha \tag{2.28}$$

For example, if 2000 Bootstrap samples are generated, the estimate of 95$^{\text{th}}$ percentile will be the 1900$^{\text{th}}$ largest value of the $z^*(\mathrm{b})$s, as well as the 5$^{\text{th}}$ percentile will be the 100$^{\text{th}}$ largest value.

$z^*(\mathrm{b})$ in (2.28), which is given as follows:

$$z^*(\mathrm{b}) = \frac{\hat{\theta}^*(b) - \hat{\theta}}{\widehat{se}^*(b)} \tag{2.29}$$

Where $z^*(\mathrm{b})$ is the z value of the Bootstrap sample of number b, and $\hat{\theta}^*(b)$ is the Bootstrap estimate of θ of the Bootstrap sample X^{*b}, and $\widehat{se}^*(b)$ is the estimate of the SE of $\hat{\theta}^*$ of the Bootstrap sample X^{*b}.

It is important to mention that there are many Bootstrap methods to estimate the CI, such as:

1. The normal theory method.
2. The percentile method.
3. The studentized-*t* method.
4. The bias corrected and accelerated ($\mathrm{BC_a}$) method.
5. The approximate Bootstrap CI (ABC) method.

The percentile and BC_a and ABC methods are nonparametric methods, while the normal theory and the studentized-*t* methods are parametric methods.

2.3.5. Hypotheses Testing Using Bootstrap

Permutation Tests

The permutation hypothesis testing method can be demonstrated as follows:

1. Calculate the observed value of the test T (i.e., the difference between the means of the samples).
2. Combine the observations of both samples to become as one sample $z = \{x_1, \dots, x_n, y_1, \dots, y_m\}$ with sample size $n + m$.
3. Randomly reorder the observations of the sample z.
4. For the new reordered sample, take the first n observations as the sample X and the rest m observations as the sample Y.
5. Calculate the parameter of interest t_i (the mean difference between these two samples).
6. Repeat the steps from 3 to 5 for a big number of times.
7. The p-value of the test is the proportion of parameters of interest t_i of the permutation samples that are greater than the observed parameter T (proportion of sample means that are greater than the observed mean).

$$P(T \leq t_i | H_0) = 1 - \alpha$$

Relationship With Bootstrap Hypothesis Testing

Efron and Tibshirani (1993) demonstrated the method of Bootstrap hypothesis testing as follows:

1. Calculate the observed value of the test T (the difference between the means of the samples).
2. Combine the observations of both samples to become as one sample $z = \{x_1, \dots, x_n, y_1, \dots, y_m\}$ with sample size $n + m$.
3. Draw a sample of size $n + m$ with replacement from z.
4. For the new bootstrap sample, take the first n observations as the sample X and the rest m observations as the sample Y.
5. Calculate the parameter of interest t_i (the mean difference).
6. Repeat steps from 3 to 5 for B times.
7. The p-value of the test is the proportion of parameters of interest t_i that are greater than the observed parameter T (proportion of sample means that are greater than the observed mean).

$$P(T \leq t_i | H_0) = 1 - \alpha$$

The demonstrated Bootstrap hypothesis testing method goes through the exact same steps as the permutation tests demonstrated in the previous section, the only difference can be noticed in step 3, where the permutation tests use resampling without replacement from the original sample, while the Bootstrap tests use resampling with replacement. This means that some observations in the Bootstrap hypothesis tests may be selected many times, and some others may not appear at all, while this does not happen in the permutation tests.

The advantage of the Bootstrap tests over the permutation tests is that the latter has a limited number of samples, which is not the case in the Bootstrap tests, the thing that made them widely applicable (Efron and Tibshirani, 1993).

Relation Between CI and Bootstrap Hypothesis Tests

In resampling methods there is a strong connection between hypothesis testing and CI, many literatures describe it as 1 to 1 relation (look at Efron and Tibshirani, 1993; Chernich and LaBudde, 2011). For a CI with the confidence level $(1 - \alpha)\%$, to test any hypothesis, i.e., $H_0: \theta = \theta_0$, it can be done just by looking at the CI, so if θ_0 is laying outside the CI, then this gives a strong reason to reject the null hypothesis H_0. This relation is applicable for both one-sided tests and two-sided tests, just like the one-sided CI and the two-sided CI respectively. At the same time, for any Bootstrap method that is used for estimating the CI, there is a corresponding Bootstrap method for hypothesis testing (Chernic and LaBudde, 2011).

2.3.6. Number of Bootstrap Samples

Number of Available Bootstrap Samples

It is important to know how many distinct Bootstrap samples can be obtained, it could easily be proved that it is equal to $\binom{2n-1}{n}$, where n is the size of the original sample. i.e., let $n = 2$, resulting in 3 available samples, which are $(x_1, x_1), (x_2, x_2)$, and (x_1, x_2). Since the order does not matter, this means (x_1, x_2) is the same as (x_2, x_1). Taking the order into consideration, it can be found that the number of the available Bootstrap samples will be n^n, the difference here is that the exchangeability of the observations is not considered (Chernick and LaBudde, 2011).

The available number of Bootstrap samples does not mean that there is a limit to the number of Bootstrap samples, since it does not matter if the Bootstrap sample occurs more than once, in other words, there are no conditions that the generated Bootstrap samples should be distinct. The overall discussion leads to the conclusion that an unlimited number of Bootstrap samples can be considered available.

What is amazing with the Bootstrap method is that even if the distinct samples are taken, a very big number of Bootstrap samples is available, i.e., if $n = 10$, the number of available Bootstrap samples will be $n^n = 10^{10}$ which is equal to 10 billion (Chernic and LaBudde, 2011). This thing gives the Bootstrap method a huge advantage over the Jackknife method, where in its case, only 10 Jackknife samples are available. On the other hand, this advantage creates a challenge, which is choosing the optimal number of Bootstrap samples.

The Optimal Number of Bootstrap Samples

In theory, the best estimate of a statistic of interest should be the one that is calculated using all possible Bootstrap samples, which should give the exact estimate of the statistic, means that when $B \to \infty$ (Efron, 1994). But it can be easily realized that in most cases it is impossible to take this choice, because the number of the available Bootstrap samples is n^n (putting the ordering into consideration), that is why it is more applicable to use a sample from the available Bootstrap samples (Chernic and LaBudde, 2011). On the other hand, taking a small number of samples gives an inaccurate estimate, this leads to one of the challenging questions when using the Bootstrap method, which is, what is the optimal number of Bootstrap samples?

The main constraint when choosing the number of Bootstrap samples is the computing capacity, because ideally, B should be as big as possible to guarantee a good estimate.

Literature review confirms that there is no clear answer to this question, where many suggested numbers have been proposed, but it can be noticed that:

1. They vary in a very wide range.
2. The proposed numbers did not rely on any theoretical background.

3. The proposed numbers were for specific cases and statistics, not as something that could be generalized for all statistics and cases.

Efron and Tibshirani (1993) suggested using 200 Bootstrap samples to estimate the SE, they saw that the number is good enough, and in the case of time and computational capacity constraints, they mentioned that 25 samples is an acceptable number. While for the CI, they suggested a significantly bigger number, which is 1000. Devidson and Mackinnon (2000) suggested the optimal number of Bootstrap samples for hypothesis tests to be 599.

As a literature method to choose the number of Bootstrap samples, the coefficient of variation (CV) has been proposed. CV is the standard deviation divided by the expectation of the statistic that is estimated using the Bootstrap methods. It supports our decision about the optimal number of Bootstrap samples, which is the one that has CV small enough and with no significant change in its value when taking the bigger number of samples. What is good with the CV is that it measures two types of errors, the original sampling error and the Bootstrap sampling error (Efron and Tibshirani, 1993). But it could be also noticed that the use of CV does not give a definite decision, which raises another question: how small should the CV be?

Mainly the selection of the optimal number of Bootstrap samples should consider many factors, which are:

1. The used Bootstrap method: whether parametric or non-parametric. In the latter, it is better to use a bigger number of Bootstrap samples to guarantee excluding the effect of having a low representative sample, while in the parametric Bootstrapping case, a smaller number of samples is sufficient, especially in the case when a previous knowledge about the distribution of the data is available, which is better than the case of assuming the distribution according to the underlying data.
2. The statistic of interest: the number of required Bootstrap samples significantly varies according to the statistic of interest, whether it is the mean or the standard deviation. The standard deviation is more complicated, this is why it requires a bigger number of samples.
3. The needed estimate: in addition to the statistic of interest, the required estimate makes a difference in the number of required Bootstrap samples. It is enough to have $B = 200$ for the SE, while a bigger number of samples is needed for the CI.
4. Complexity of the problem: for two variable cases, there is a need for a bigger number of Bootstrap samples compared to one variable case.
5. Required confidence level: where the increase in the required confidence level means an increase in the number of Bootstrap samples.

6. Smoothness of the statistic of interest: especially for the Bootstrap methods that are more sensitive to the smoothness.
7. Data distribution: where the skewed distribution and the distributions with lower kurtosis require a bigger number of Bootstrap samples to give good estimates.
8. Occurrence of outliers: here, the increase in the number of Bootstrap samples helps using the trimmed estimates to decrease the negative effect of the outliers.

At the end, the development of computational capacity in recent years allowed increasing the number of Bootstrap samples in a level that exceeds the numbers that were suggested by many literatures, that is why it is recommended to increase the number of Bootstrap samples as much as possible, this way, the Bootstrap method achieves the optimal number of samples.

2.4. Challenges of Resampling Methods

Mainly the resampling methods have two sources of error, the first one comes from the fact that the sample is not 100% representative of the distribution of the data, while what makes the issue bigger is that if the underlying sample is not a good representative of the distribution, which increases the magnitude of error (Efron and Tibshrani, 1993). This type of issue could come in many ways, like using a small sample that is not big enough to represent the distribution, using wrong sampling method, having missing data, and/or having outliers in the sample. This demonstrates how critical is choosing the original sample for the resampling methods, because it affects all the estimations that will be made using the Bootstrap or the Jackknife methods, that is why it is important to try to have a sample that highly represents the distribution of the data.

The second source of error can be called resampling variability, which means that the subsamples themselves do not give a good representation of the original sample, this issue especially occurs in the Bootstrap method, because all available Bootstrap samples cannot be taken, which necessitates the use a sample from the available Bootstrap samples (Shao and Tu, 1995). This issue can be resolved by increasing the number of the used Bootstrap samples which reduces the magnitude of this error.

In addition to the errors that are mentioned above, there are many advantages and disadvantages of the Bootstrap and the Jackknife methods that were discussed. The challenges are related to many areas, as there are challenges related to usability, while other challenges are related to the accuracy and reliability of the results. The most discussed challenges are related to the issues with the non-smooth statistics, the magnitude of the needed computations, the usability for hypothesis testing, the existence of missing data or outliers, the convergence of the results, the accuracy of the results when using it in skewed distributions, and the selection of the number of subsamples.

Starting with the available number of subsamples, it is clear that the number of Jackknife samples is limited to the sample size, while the number of Bootstrap samples is significantly bigger, as presented before. This helps giving the Bootstrap method an significant advantage over the Jackknife method, which is the scalability of the number of the subsamples, which is also reflected to the convergence performance, where the Bootstrap converges to the value of the statistic of interest when increasing the number of the withdrawn samples, but it is not the same with the Jackknife method, because the number of the Jackknife samples is the same as the original sample size, this means that the Jackknife estimates are not necessary to converge to the value of the statistic on interest (Shao and Tu, 1995). But what is considered an advantage for the Bootstrap in a certain way becomes a disadvantage in another way, this can be noticed when coming to the choice of the optimal number of Bootstrap samples that are needed to have a good estimate, as for the Jackknife it is very clear and simple, which is n the original sample size, but for the Bootstrap, it is a very difficult and complicated decision, where it is significantly varied according to the statistic that is to be estimated.

The issue presented above also overlaps with another challenge, which is the usability and accuracy of results when having a small sample. Since the resampling methods use the original sample to generate the subsamples, this poses a challenge, this challenge is more significant with the Jackknife method, because as mentioned above, the number of available Jackknife samples is limited to the size of the original sample, on the other hand, it can be noticed that the Bootstrap method is of a better performance in this case, which helps having accurate results.

This difference between the resampling methods is reflected by another advantage of the Bootstrap, which can be called the differentiability, while the Jackknife takes specific samples, and the Bootstrap takes a sample from the available Bootstrap samples, this means that each time the experiment is repeated, the same results are obtained in the Jackknife case, but different results are obtained in the Bootstrap case. This variability in the Bootstrap case helps in repeated verifications and multi-experimental use (Shao and Tu, 1995).

The scalability of the number of Bootstrap samples presents a disadvantage against the Jackknife, where the Bootstrap is more computationally intensive and needs more computational capacity. Sure, this challenge is becoming lower with time, because of the increase of computers capacity (Severiano et al, 2011).

There are two additional challenges related to the original sample, which are noticed when facing a sample with missing values or outliers. In the case of the missing values, both Bootstrap and Jackknife methods do not perform well, unless data cleaning is conducted or imputation methods are applied (Shao and Tu, 1995). While in the case of the outliers, the Bootstrap method performs better than the Jackknife, this happens because, i.e., when one outlier is present, that means this outlier will certainly appear in n-1 Jackknife samples, which is not the same case in the Bootstrap method. This advantage of the Bootstrap method appears more clearly when using it for

repeated verification, that, as mentioned before, gives different results in the Bootstrap method, but it gives the same result in the Jackknife method.

The Jackknife shows weak performance in the case of the non-smooth statistics (Efron and Tibshrani, 1993), which is discussed before, in addition to the weak performance for the skewed distributions, which is also related to the advantage of repeated verification of the Bootstrap, which helps testing the results by using the multi-experimental approach. Finally, it has been shown that the Jackknife method has not been proposed as a hypothesis testing tool, while the Bootstrap is used to test hypotheses.

The discussion presented above shows the advantage of the Bootstrap over the Jackknife method in general, but as shown, the case is not the same for all situations.

The table below provides a brief comparison between the Bootstrap and the Jackknife methods:

Table 2.6. Comparison between the Bootstrap and the Jackknife methods

Case	**Jackknife**	**Bootstrap**
Estimates of non-smooth statistics	Inaccurate results	Better results
Needed computation capacity	Lower computations	Intensive computations
The use for repeated verifications/multi-experimental use	Not usable	Usable
Performance with small samples	Lower performance	Good performance
Scalability of the number of subsamples (B)	Not scalable	Highly scalable
Selecting the number of the subsamples (B)	Simply the same size as the original sample n.	Difficult and complicated decision to choose the optimal number of Bootstrap samples B. Varies too much according to the needed estimate.
Convergence performance	Not necessary to converge to the value of the statistic of interest	Converges to the value of the statistic of interest
Skewed distributions	Weak performance	Better performance
Missing values	Does not perform well	Does not perform well
Outliers	Weak performance	Performs better when taking a big number of samples then use a trimmed estimate, or using a multi-experimental approach
The use for hypothesis testing	Not proposed as a tool for hypothesis testing	Many approaches proposed as hypothesis testing tools

To overcome the discussed issues, many developments on the Bootstrap and Jackknife methods were proposed, like the Jackknife delete-d method which has been discussed above, which has been proposed to solve the issue of non-smooth statistics.

Finally, it is important to mention that there are more challenges that should be discussed for other types of estimations, like the estimations using resampling methods in time series, which have different data structures.

2.5. Conclusion

The resampling methods made a big step in the statistical estimation, where the Bootstrap and Jackknife helped solve statistical issues that was impossible to solve before proposing them, this is why their use increased significantly in the recent years because of the development of computational capacity of computers, the thing that solved the biggest challenge that was facing the resampling methods, which is being computationally intensive compared to traditional statistical methods.

A key advantage of the resampling methods is that no assumptions about the distribution of the underlying data need to be made for their use, which is a significant challenge faced when using traditional statistical methods.

The nonparametric Bootstrap has an advantage over the parametric Bootstrap, This advantage is that the need for previous information or making assumptions about the data distribution is relieved. On the other hand, even if the parametric Bootstrap is used for estimation, the nonparametric Bootstrap can be used as a validation method of the results obtained using the parametric Bootstrap.

Finally, it is important to mention that, just like any other method, there are some weaknesses and disadvantages of the resampling methods, which are mainly related to the usability and accuracy, like the usability for non-smooth statistics and dirty data, and the accuracy when having a small sample and skewed data… etc. To overcome these weaknesses, many advanced Bootstrap and Jackknife methods have been developed.

3. CI OF THE LE'S RCs USING JACKKNIFE AND BOOTSTRAP METHODS

In multiple regression analysis, the use of the LE over the OLSE was suggested by Kejian (1993) to overcome the problem of multicollinearity. In this chapter, different CIs for RCs based on the LE using Bootstrapping and Jackknife methods are developed and compared. For comparison, CWs and the coverage probabilities are calculated through a simulation study for the data that suffer from the problem of multicollinearity. Estimators related to the LE and traditional estimation methods of the LP d are discussed in Section 3.2. In Section 3.3, CIs based on Bootstrap and Jackknife methods using the LE in the linear regression model are presented. The methods are applied to simulated data, and their performance is compared based on the CWs and the coverage probabilities in Section 3.4. Finally, the discussion and conclusions are given in Section 3.5.

3.1. The Biased Estimators

Consider the multiple linear regression model

$$y = X\beta + \varepsilon, \tag{3.1}$$

where y is an $n\times 1$ vector of observations of the dependent variable, X is an $n\times p$ model matrix of observations of p non-stochastic explanatory variables, β is a $p\times 1$ vector of unknown parameters and ε is an $n\times 1$ vector of disturbances with expectation $E(\varepsilon)=0$ and dispersion matrix $Cov(\varepsilon)=\sigma^2 I_n$. The OLSE of the parameter β in model (3.1) is given by

$$\hat{\beta} = (X'X)^{-1}X'y. \tag{3.2}$$

This estimator is widely used for estimating linear regression models. The main reason for focusing on the OLSE is because it is unbiased and has the minimum variance among all the linear unbiased estimators. But in the presence of multicollinearity, the OLSE becomes unstable and often gives misleading information. Alternative estimators designed to combat multicollinearity yield biased estimators. The usage of biased estimators results in significant improvement in the estimation, prediction or forecasting performance of the linear regression model. One of the most popular estimators that are proposed to deal with multicollinearity is ridge regression estimator (RRE) which is proposed by Hoerl and Kennard (1970) and is defined as

$$\hat{\beta}(k) = (X'X + k\,I)^{-1}X'y, \quad k \geq 0 \tag{3.3}$$

Where the parameter k is known as the biasing parameter. The RRE is biased, but it is superior to the OLSE according to the criterion of estimation variance, where Hoerl and Kennard (1970) showed that there is always an RRE with a smaller MSE than the OLSE has. However, the use of biased estimation techniques brings the obligation of selecting the biasing parameters of these estimators, as well. Consequently, numerous methods for selecting the biasing parameters have been studied, see for example, (Hoerl et al., 1975; Mc Donald and Galerneau, 1975; Hoerl and Kennard, 1976; Lawless and Wang, 1976; Hocking, 1976; Dempster et. Al, 1977; Wichern and Curchill, 1978; Golub et al., 1979; Gibbons, 1981; Nordberg, 1982; Tekeli et al., 2019). One of the important methods is the one that has been given by Delaney and Chatterjee (1986), which uses the Bootstrap method to choose the optimal Ridge parameter, where it demonstrates that the Bootstrap choice of the ridge parameter leads to a smaller MSE of prediction than the ridge trace method. It also concluded that the benefits of the usage of the Bootstrap method to choose ridge's biasing parameter include its less subjective nature, ease of implementation, and robustness. Furthermore, different CIs for RCs based on the RRE were given using Bootstrap and Jackknife methodologies by Chaubey et al (2018). Crivelli et al (1995) compared the effects of different methods of calculating the Ridge's biasing parameter k on the accuracy of the RE by showing the results of many methods of estimating the CI of the regression parameters that are estimated by the RE and tested using two criteria. Their study combined the Bootstrap method with the Edgeworth expansion to estimate the CI. Following that, Özkale and Altuner (2022) expanded the study to be conducted on 11 methods of estimating k, whereas Crivelli et al (1995) conducted it on 3 methods.

Using the same idea of the RE, Kejian (1993) proposed a new estimator that is then known as the LE (see Akdeniz and Kaçıranlar, 1995), which is given as follows

$$\hat{\beta}(d) = (X'X + I)^{-1}(X'y + d\hat{\beta}),\ 0 < d < 1 \tag{3.4}$$

It appears to combine the advantages of the RRE and the Stein-type estimator $\hat{\beta}_s = s\hat{\beta}$, $0 < s < 1$. $\hat{\beta}(k)$ is effective in practice, but it is a complicated function of k. The advantage of $\hat{\beta}(\mathrm{d})$ over $\hat{\beta}(k)$ is that $\hat{\beta}(\mathrm{d})$ is a linear function of d. Kejian (1993) showed that $\hat{\beta}(\mathrm{d})$ has a similar good property as $\hat{\beta}(k)$. He has shown that there always exists $0 < d < 1$ such has a smaller MSE of LE than the OLSE has. He also gave some other estimates of d by analogy with the estimates of k in the RE. Later, many authors have worked on the selection of the d parameter in the LE by analogy with the estimates of k in the RE (see for example, Özkale and Kaçıranlar, 2007; Shukur et al., 2015; Tekeli et al., 2019; Qasim et al., 2019). Many authors have also examined the properties of the LE, its applications in different models, and compared it with other known estimators. For example, Sakallıoğlu et al (2001) gave a comparison between the LE

and the RE and the iteration estimator. These authors showed that in some conditions, the LE is better than the Ridge's and the iteration estimator. Özbey and Kaçıranlar (2015) discussed the predictive performance of the LE compared to the OLSE, as well as to two other popular biased estimators, namely, the principal components and the RREs. Furthermore, LE plays an important role in many fields, such as chemistry, social sciences, economics and biostatistics, etc. It also has very common applications in various disciplines and various regression models such as the RRE.

For ease of computation, model (3.1) can be considered in canonical form. Using the spectral decomposition of $X'X$, it is written as:

$$y = Z\alpha + \varepsilon \tag{3.5}$$

where $\alpha = Q'\beta$, $Z = XQ$, $Q'X'XQ = Z'Z = \Lambda$, $\Lambda = diag(\lambda_1, \lambda_2, ..., \lambda_p)$ is a diagonal matrix of the eigenvalues of $X'X$ and $Q = diag(q_1, q_2, ..., q_p)$ is an orthogonal matrix with the eigenvectors of $X'X$ as its columns. Under the canonical form, the OLSE, ridge and LEs can be given respectively as follows

$$\hat{\alpha} = \Lambda^{-1}Z'y. \tag{3.6}$$

$$\hat{\alpha}(k) = (\Lambda + kI)^{-1}Z'y, \quad k \geq 0 \tag{3.7}$$

and

$$\hat{\alpha}(d) = (\Lambda + I)^{-1}(Z'y + d\hat{\alpha}) = (\Lambda + I)^{-1}(\Lambda + dI)\hat{\alpha}, \quad 0 < d < 1. \tag{3.8}$$

The estimates of regression parameters, as any other estimate, are preferable to be given in conjunction with the estimates of their CIs, that is why many methods of estimating CI of the ridge regression parameter have been presented like the normal theory method, the studentized-t method, in addition to other methods that depend on resampling methods like the percentile method and the Jackknife method. The Bootstrap and Jackknife are known as powerful resampling methods for variance and CI estimation even for non-normal errors. The results of these methods are affected by many criteria like the level of correlation between the explanatory variables, the sample size, and the value of the parameter k. As far as is known, no similar studies of the CI of the RE have been made for the LE. In addition to using the LE to create proper estimates of regression parameters with lower MSE compared to the OLSE, where it is expected that the LE gives better results on CI.

3.2. Related Estimators and the Choice of the LP

3.2.1. Related Estimators

The generalized LE proposed by Kejian (1993) is given by

$$\begin{aligned}\hat{\beta}(\mathrm{D}) &= (X'X+I)^{-1}(X'y+D\hat{\beta}), \\ &= (X'X+I)^{-1}(X'\mathrm{X}+D)\hat{\beta} \\ &= \left[I-(X'X+I)^{-1}(I-D)\right]\hat{\beta} \\ &= F_D\hat{\beta}\end{aligned} \tag{3.9}$$

where D is a positive definite diagonal matrix with nonzero elements d_i, $0<d_i<1$, $i=1,2,...,p$ and $F_D = (X'X+I)^{-1}(X'\mathrm{X}+D) = \left[I-(X'X+I)^{-1}(I-D)\right]$.

When $d_1=d_2=....=d_p=d, 0<d<1$, the LE in (3.4) is obtained. The LE can be written as

$$\begin{aligned}\hat{\beta}(\mathrm{d}) &= (X'X+I)^{-1}(X'\mathrm{X}+dI)\hat{\beta},\ 0<d<1 \\ &= \left[I-(1-d)(X'X+I)^{-1}\right]\hat{\beta} = \left[I-(1-d)\,\mathrm{A}^{-1}\right]\hat{\beta}\end{aligned} \tag{3.10}$$

where $\mathrm{A}=(X'X+I)$. To reduce the bias, Kadiyala (1984) proposed a class of almost unbiased estimators, which include a bias corrected RE as a special case. But unfortunately, his estimator is not operational since the bias corrected term includes unknown parameters. Then, Ohtani (1986) replaced the unknown parameters by their REs and the resulting estimator is called almost unbiased ridge estimator (AURE). Singh et al (1986) used the Jackknife procedure to reduce the bias of the RE. They called their estimator Jackknife ridge estimator (JRE). They also demonstrated that JRE can be simplified to AURE. Moreover, they showed that the Jackknife technique offers a very simple method for providing CI for β. Later, many articles about Jackknife ridge or the AURE were published, such as, (Özkale, 2008; Khurana et al., 2014; and Chaubey et al., 2018).

Similarly, following Kadiyala (1984), Ohtani (1986), Singh et al (1986) and Nomura (1988). Akdeniz and Kaçıranlar (1995) firstly introduced the almost unbiased generalized LE (AUGLE) and unbiased estimation of the bias and MSE. AUGLE is given by

$$\tilde{\beta}(\mathrm{D}) = \left[I + (X'X + I)^{-1}(I - D) \right] \hat{\beta}(D), \quad (3.11)$$
$$= \left[I - (X'X + I)^{-2}(I - D)^2 \right] \hat{\beta}.$$

Later, Duran and Akdeniz (2012) used the Jackknife procedure to reduce the bias of the generalized LE. They called their estimator Jackknifed generalized LE (JGLE). But it is easy to show that JGLE will be the same as the AUGLE defined by Akdeniz and Kaçıranlar (1995). When $d_1 = d_2 = ... = d_p = d$, $0 < d < 1$, an almost unbiased LE (AULE) is obtained. AULE or Jackknifed LE can be written as

$$\hat{\beta}_{JLE}(\mathrm{d}) = \left[I + (1-d)(X'X + I)^{-1} \right] \hat{\beta}(\mathrm{d}), \quad (3.12)$$
$$= \left[I - (1-d)^2 (X'X + I)^{-2} \right] \hat{\beta}.$$

Furthermore, Hu and Xia (2013) introduced a generalized LE and Jackknifed LE in a linear regression model with correlated or heteroscedastic errors. Wu and Yang (2013) proposed an almost unbiased two parameter estimator. Later, Duran and Akdeniz (2012) and Yildiz (2018) proposed a modified Jackknifed Liu-type estimator in linear regression model.

3.2.2. The Choice of the LP

Several methods have been proposed in the literature for finding the biasing parameter of the LE, some of these are $\hat{d}_{MG}$ (Mcdonald-Galarneau method), $\hat{d}_{ITR}$ (iterative method), $\hat{d}_{MM}$ (minimum mse method), $\hat{d}_{MMH}$ (minimum mse method), $\hat{d}_{CL}$ (Mallows C_p method) and $PRESS_d$ (Özkale and Kaçıranlar, 2007). Golub et al (1979) suggested the generalized cross-validation (GCV) statistic for choosing k for RE. Similarly, Kejian (1993) gave $\hat{d}_{GCV}$ for the LE. He showed that $\hat{d}_{GCV}$ is just $\hat{d}_{CL}$. Thus, the formulas for $\hat{d}_{MM}$, $\hat{d}_{MMH}$ and $\hat{d}_{CL}$ are given below (for details see Kejian, 1993):

$$\hat{d}_{MM} = 1 - \hat{\sigma}^2 \left[\sum_{i=1}^{p} \frac{1}{\lambda_i(\lambda_i+1)} / \sum_{i=1}^{p} \frac{\hat{\alpha}_i^2}{(\lambda_i+1)^2} \right], \quad (3.13)$$

$$\hat{d}_{MMH} = 1 - h\hat{\sigma}^2 \left[\sum_{i=1}^{p} \frac{1}{\lambda_i(\lambda_i+1)} / \sum_{i=1}^{p} \frac{\hat{\alpha}_i^2}{(\lambda_i+1)^2} \right], \quad (3.14)$$

$$\hat{d}_{CL} = 1 - \hat{\sigma}^2 \left[\sum_{i=1}^{p} \frac{1}{(\lambda_i + 1)} / \sum_{i=1}^{p} \frac{\lambda_i \hat{\alpha}_i^2}{(\lambda_i + 1)^2} \right] \quad (3.15)$$

where,

$$0 < h < \frac{2(n-p)}{(n-p+2)} \left(1 - \frac{2}{\sum_{i=1}^{p} \lambda_p(\lambda_p+1)/\lambda_i(\lambda_i+1)} \right). \quad (3.16)$$

3.3. CI Based on Bootstrap and Jackknife Methodologies Using the LE

In this chapter, CIs based on Bootstrap and Jackknife methodologies using the LE are presented. The approach of (Chaubey et al., 2018) for RRE is extended to LE. Now for Bootstrapping, for the model defined in (3.1), the OLSE for full sample is fitted, the standardized residuals $\hat{e}_i$ are calculated and then a Bootstrap sample of size n with replacement $(\hat{e}_1^b, \hat{e}_2^b, ..., \hat{e}_n^b)$ is selected from the residuals $\hat{e}_i$'s giving $\frac{1}{n}$ probability to each $\hat{e}_i$. After this, the Bootstrap y values are obtained using the resampled residuals keeping the design matrix fixed as shown below

$$y^{(b)} = X\hat{\beta} + \hat{e}^{(b)}, \quad (3.17)$$

where $\hat{e}^{(b)} = \left(\hat{e}_1^b, \hat{e}_2^b, \ldots, \hat{e}_n^b\right)'$. These Bootstrapped y values are then regressed on the fixed design matrix to obtain the Bootstrap estimates of the RCs. So, the LE from the first Bootstrap sample is given by

$$\hat{\beta}_{d(b_1)} = \left[I - (1-d)(X'X + I)^{-1} \right] (X'X)^{-1} X' y_{(b_1)}, \quad (3.18)$$

where $y_{(b_1)}$ is the Bootstrapped y values for the first Bootstrap sample. Repeating the above steps *K* times, where *K* is the number of Bootstrap samples, the Bootstrap LE will be given by

$$\bar{\hat{\beta}}_{d(b)} = \frac{1}{K} \sum_{r=1}^{K} \hat{\beta}_{d(b_r)}. \quad (3.19)$$

The estimated bias and variance-covariance matrix are respectively given by

$$\widehat{Bias} = \bar{\hat{\beta}}_{d(b)} - \hat{\beta}(d), \quad (3.20$$

)

And

$$\widehat{Var} = \sum_{r=1}^{K} \left(\hat{\beta}_{d(b_r)} - \bar{\hat{\beta}}_{d(b)}\right)\left(\hat{\beta}_{d(b_r)} - \bar{\hat{\beta}}_{d(b)}\right)' / (K-1). \quad (3.21)$$

Now, based on these estimates, the CI for β is constructed.

3.3.1. CI for LPs Based on Resampling Methods

There are several methods for constructing the Bootstrap CI based on the estimate of variance given in (3.21). In this chapter, the normal theory, Jackknife, percentile, studentized-t and BC_a methods that are given in the literature for different estimators, are preferred for use.

Normal theory method

If $\hat{\theta}$ is an estimator for a parameter θ, it is known that $\hat{\theta} - bias(\hat{\theta})$ is a bias corrected estimator of θ (Kadiyala, 1984). Following Ohtani (1984) showed that if θ is replaced by the biased estimator $\tilde{\theta}$, the almost unbiased estimator is obtained. The resulting estimator will be

$$\hat{\theta} - \widehat{bias}(\hat{\theta}) \quad (3.22)$$

where $\widehat{bias}(\hat{\theta})$ is a suitable estimate of the bias of $\hat{\theta}$. The first method for constructing Bootstrap CI assumes that the sampling distribution of $\hat{\beta}(d)$ is normal. Using (3.20) and (3.22), the Bootstrap almost unbiased estimator for β is obtained as follows

$$\hat{\beta}(d) - \widehat{Bias} = \hat{\beta}(d) - (\bar{\hat{\beta}}_{d(b)} - \hat{\beta}(d)) = 2\hat{\beta}(d) - \bar{\hat{\beta}}_{d(b)}. \quad (3.23)$$

A 95% CI for β_j based on $\hat{\beta}(d)$ is given by

$$(2\hat{\beta}(d) - \bar{\hat{\beta}}_{d(b)})_j - z_{1-\frac{\alpha}{2}}\sqrt{\widehat{Var}_j} < \beta_j < (2\hat{\beta}(d) - \bar{\hat{\beta}}_{d(b)})_j + z_{1-\frac{\alpha}{2}}\sqrt{\widehat{Var}_j} \quad (3.24)$$

where $(2\hat{\beta}(d) - \bar{\hat{\beta}}_{d(b)})_j$ is the j_th element of $2\hat{\beta}(d) - \bar{\hat{\beta}}_{d(b)}$ and $\widehat{Var}_j$ is the j_th diagonal element of the Bootstrap estimate of the variance of $\hat{\beta}(d)$ as defined in (3.21), $\alpha = 0.05$ and $z_{1-\frac{\alpha}{2}}$ is the $1-\frac{\alpha}{2}$ quantile of the standard normal distribution.

Jackknife method

As discussed above, the Jackknife technique is used to reduce the bias of estimators and to estimate the variances. For this method, the variance estimate of LE is given by

$$\widehat{Var}_J(\hat{\beta}(d)) = \frac{1}{n(n-p)}\sum_{i=1}^{n}(Q_i - \hat{\beta}_{JLE}(d))(Q_i - \hat{\beta}_{JLE}(d))'$$

where Q_i's are pseudo values defined as

$$Q_i = \hat{\beta}(\mathrm{d}) + \mathrm{n}(1-\mathrm{w}_i)(\hat{\beta}(\mathrm{d}) - \hat{\beta}_{(-i)}(\mathrm{d}))$$

where $\hat{\beta}_{(-i)}(\mathrm{d})$ is the LE calculated by deleting the i^{th} row from the data and $w_i = x_i^{'}A^{-1}x_i$, $\mathrm{A} = (X'X + I)$ is as defined earlier and $x_i^{'}$ is the i^{th} row of X matrix (see also Hu and Xia, 2013). A 95% Jackknife CI for β based on $\hat{\beta}(d)$ is

$$\hat{\beta}_{JLE(j)}(d) - t(1-\tfrac{\alpha}{2};n-p)\sqrt{\widehat{Var}_j} < \beta_j < \hat{\beta}_{JLE(j)}(d) + t(1-\tfrac{\alpha}{2};n-p)\sqrt{\widehat{Var}_j}, \qquad (3.25)$$

where $\hat{\beta}_{JLE(j)}(d)$ is the j_th element of $\hat{\beta}_{JLE}(d)$ and $\widehat{Var}_j$ is the j_th diagonal element of $\widehat{Var}_J(\hat{\beta}(d))$.

Percentile method

In this method, the empirical quantiles of the Bootstrap of LE ($\hat{\beta}_{d(b)}$) are used to construct the CI for β. A 95% CI for *K* Bootstrap resamples is

$$\hat{\beta}_{d(0.025\times K)} < \beta < \hat{\beta}_{d(0.975\times K)} \qquad (3.26)$$

where $\hat{\beta}_{d(\mathrm{r})}$ is the r^{th} observation in the ordered Bootstrap replicates of $\hat{\beta}_{d(b)}$ (see Efron and Tibshirani, 1994).

Studentized-t method

The studentized-t Bootstrap CI takes the same form as the normal CI, except that instead of using the quantiles from the normal distribution. The quantiles are calculated using the Bootstrapped t distribution (see Efron and Tibshirani, 1994; Chaubey et al., 2018). The t statistic for the j^{th} element of β is calculated using the following formula

$$t = (\hat{\beta}_j(d) - \beta_j)/se_j$$

where se_j is the sample estimate of the SE of $\hat{\beta}_j(d)$ and $\hat{\beta}_j(d)$ is the j_th element of $\hat{\beta}(d)$. The Bootstrap form of t is given by

$$t_b = (\hat{\beta}_{d(b),j} - \hat{\beta}_j(d))/se_{b,j}$$

where $se_{b,j}$ is the Bootstrap estimate of the SE. Denote the $100s^{th}$ Bootstrap percentile of t_b by b_s and consider the statement that $b_{0.025} < t < b_{0.975}$, and after substituting t, the following CI at a 95% confidence level is obtained

$$\hat{\beta}(d) - (se)\, b_{0.975} < \beta < \hat{\beta}(d) + (se)\, b_{0.025}. \tag{3.27}$$

BC_a Method

The Bias Corrected and accelerated method (BC_a), is more efficient and accurate compared to the other methods (basic, percentile, studentized-*t*, Bootstrap, and standard normal), while it is more accurate in the asymmetric data (Efron and Tibshirani, 1994). This method depends on using a different way of calculating the endpoints of the CI, by using different α value, which is calculated as follows:

$$\alpha_1 = \Phi\left(\hat{z}_0 + \frac{\hat{z}_0 + Z^{(\alpha)}}{1 - \hat{a}(\hat{z}_0 + Z^{(\alpha)})}\right) \tag{3.28}$$

$$\alpha_2 = \Phi\left(\hat{z}_0 + \frac{\hat{z}_0 + Z^{(1-\alpha)}}{1 - \hat{a}(\hat{z}_0 + Z^{(1-\alpha)})}\right) \tag{3.29}$$

where $\Phi(.)$ is the standard normal cumulative distribution function and $\hat{z}_0$ is called the bias correction parameter that can be easily calculated from the proportion of Bootstrap replications less than the original estimate $\hat{\theta}$ as follows:

$$\hat{z}_0 = \Phi^{-1}\left(\frac{\#\{\hat{\theta}^*(b)<\hat{\theta}\}}{K}\right). \tag{3.30}$$

Where the symbol # represents the number of the cases where $\hat{\theta}^*(b) < \hat{\theta}$.

It can be said that $\hat{z}_0$ helps overcome the asymmetry issue by making some changes in the calculation of the parameter according to the level of bias.

The acceleration parameter $\hat{a}$ is given as follows:

$$\hat{a} = \frac{\sum_{i=1}^{n}(\hat{\theta}_{(.)} - \hat{\theta}_{(i)})^3}{6\{\sum_{i=1}^{n}(\hat{\theta}_{(.)} - \hat{\theta}_{(i)})^2\}^{3/2}} \tag{3.31}$$

where $\hat{\theta}_{(i)}$ is the Jackknife value of the statistic $\hat{\theta}$ obtained by deleting the i_th point from the original sample and $\hat{\theta}_{(.)}$ is the average of all Jackknife values of the statistic defined as follows:

$$\hat{\theta}_{(.)} = \sum_{i=1}^{n}\hat{\theta}_{(i)}/n. \tag{3.32}$$

The BC_a method represents a kind of improved percentile Bootstrap method, where the parameter $\hat{z}_0$ plays the role of bias correction, while the acceleration parameter $\hat{a}$ helps in improving the convergence accuracy. When $\hat{a}$ and $\hat{z}_0$ are equal to zero, the percentile CI is obtained (Chaubey et al., 2018). By applying the BC_a method described above to the LE, one can obtain the Bootstrap CI for regression parameters.

In the next section, the previously shown methods will be compared to test their accuracy in estimating the CI. Two parameters will be used, the CP which represents the proportion of the CI that include the true parameter which is calculated from the repeated sampling from the underlying population, in addition to the CW, which is defined as the difference between the upper and lower confidence limits.

3.4. Monte Carlo Simulation

In this section, the CP and the CW of the CI for the LE will be compared based on the Normal Theory method, the Jackknife method, the Percentile method, the Studentized-t method, and the BC_a method using a simulation study. The linear regression model used in the simulations

is described in (3.1), where the error terms are generated independently and identically from a standard normal distribution. The explanatory variables are generated using the following formula:

$$x_{ij} = (1-\rho^2)^{1/2} w_{ij} + \rho w_{i(p+1)}, i = 1,2,...,n; j = 1,2,...,p \tag{3.33}$$

where w_{ij} are independent standard normal pseudo-random numbers and ρ^2 is the correlation between any pair of the explanatory variables (see Mc Donald and Galerneau, 1975). In our example two values of ρ are taken as $\rho = 0.90$ and 0.99 to discover the effects of different multicollinearity degrees. By following (Newhouse and Oman, 1971), β is taken as the normalized eigenvector which corresponds to the largest eigenvalue of the matrix $X'X$. According to the sample size, three different sizes are taken, $n = 25, 50$ and 100. $p = 3$ is used in the simulation.

The recommended value of LE's biasing parameter d is given as shown in section (3.2.2), where it takes $\hat{d} = \hat{d}_{MM}$, if $0 < \hat{d}_{MM} < 1$, otherwise it takes $\hat{d} = \hat{d}_{MMH}$, if $0 < \hat{d}_{MMH} < 1$, else, it takes $\hat{d} = \hat{d}_{CL}$.

For the LE, the package "boot.ci" in R program is used to calculate the CIs of the methods that presented above, while the Jackknife confidence limits are calculated using the formula (3.21). The CP and the average CW are calculated using 1999 Bootstrap samples, while the experiment has been repeated 1000 times.

The CP is calculated as follows:

$$CP = \frac{\#(\beta_L < \beta < \beta_U)}{N}$$

Where the symbol # represents the number of the cases where $\hat{\beta}_L < \beta < \hat{\beta}_U$.

and the average CW is given as:

$$CW = \frac{\sum_{i=1}^{N} (\hat{\beta}_{U,i} - \hat{\beta}_{L,i})}{N}$$

where N is the simulation size, which equals 1000 in our example, $\hat{\beta}_{L,i}$ and $\hat{\beta}_{U,i}$ are the lower and upper limits of the CI for the i_th replication of simulations, respectively.

The results of the CP and the average CW at two confidence levels, which are 99% and 95%, are reported in Table 3.1 to Table 3.4.

As a result of the first screening of Tables 3.1 and 3.2, it can be noticed that the coverage probabilities of the different methods significantly vary, where the highest CPs are of the OLSE, Studentized, Percentile and Normal methods, but the BC_a and Jackknife are giving low CP values, which means the latter methods are less accurate than the others. It also looks like the different

multicollinearity degrees (0.90 and 0.99) do not make a significant difference in the CPs. It is also noticed that the increase in the sample size does not cause a significant difference in the value of the CP.

In Tables 3.3 and 3.4 of the average coverage width (CW), the results seem different compared to the result of the CP, as while the Jackknife method gives close results to the other methods for the 0.90 of multicollinearity degree, it becomes worse with the higher multicollinearity, while it gives almost double wider CW compared to the other methods when $\rho = 0.99$. The other methods give close CWs with small differences that can be noticed, where the normal method gives the smallest CW, and the OLSE and the Percentile methods give close results then come the BC_a and the Studentized- methods respectively with a bit wider CWs.

Tables 3.3 and 3.4 also show that a bigger sample size gives a smaller CW, which can be illustrated by saying that the CW of the 100 sample size is less than the half of the one of $n = 25$. Seemingly the same effect for the different degrees of multicollinearity, the CW when $\rho = 0.99$ becomes almost the double or three times wider for $\rho = 0.90$.

Table 3.1. Coverage probabilities of different methods at 95% confidence level

n		*Coeff*	*OLSE*	*Normal*	*Percentile*	*Studentized*	BC_a	*Jackknife*
25	0.9	β_1	0.9500	0.9490	0.9500	**0.9610**	0.9390	0.9460
		β_2	0.9530	0.9460	0.9540	**0.9570**	0.9300	0.9110
		β_3	0.9360	0.9150	0.9400	0.9430	0.9000	**0.9530**
	0.99	β_1	0.9490	0.9510	0.9510	**0.9640**	0.9410	0.9620
		β_2	0.9390	0.9290	0.9550	0.9400	**0.9850**	0.9700
		β_3	0.9820	0.9570	0.9440	**0.9830**	0.9820	0.9820
50	0.9	β_1	0.9560	0.9530	0.9520	**0.9600**	0.9490	0.9460
		β_2	**0.9610**	0.9540	0.9530	0.9600	0.9450	0.9150
		β_3	**0.9650**	0.9640	0.9580	**0.9650**	0.9600	0.9600
	0.99	β_1	**0.9590**	0.9540	0.9530	0.9580	0.9540	0.9540
		β_2	0.9570	0.9490	0.9570	**0.9650**	0.9150	0.9400
		β_3	0.9780	0.9570	0.9640	0.9790	0.9160	**0.9860**
100	0.9	β_1	0.9410	0.9440	0.9420	**0.9480**	0.9450	0.9400
		β_2	0.9530	0.9490	0.9370	**0.9570**	0.9400	0.9080
		β_3	0.9570	0.9490	0.9450	**0.9600**	0.9440	0.9320
	0.99	β_1	0.9440	0.9430	0.9430	**0.9490**	0.9460	0.9450
		β_2	**0.9620**	0.9550	0.9410	**0.9620**	0.9140	0.9180
		β_3	0.9640	0.9510	0.9440	0.9650	0.9120	**0.9720**

Table 3.2. Coverage probabilities of different methods at 99% confidence level

n		*Coeff*	*OLSE*	*Normal*	*Percentile*	*Studentized*	BC_a	*Jackknife*
25	0.9	β_1	0.9870	0.9860	0.9850	**0.990**	0.9805	0.9855
		β_2	**0.9935**	0.9845	0.9870	0.9925	0.9775	0.9735
		β_3	0.9940	0.9850	0.9875	**0.9950**	0.9625	0.9890
	0.99	β_1	0.9860	0.9885	0.9875	**0.9925**	0.9835	0.9905
		β_2	0.9865	0.9635	0.9885	0.9870	0.9350	**0.9935**
		β_3	**0.9985**	0.9925	0.9890	**0.9985**	0.9545	0.9950
50	0.9	β_1	0.9925	0.9920	**0.9930**	0.9925	0.9910	0.9900
		β_2	**0.9950**	0.9925	0.9885	0.9945	0.9885	0.9810
		β_3	**0.9920**	0.9895	0.9905	**0.9920**	0.9875	0.9905
	0.99	β_1	0.9905	0.9910	**0.9925**	0.9910	0.9905	0.9920
		β_2	**0.9960**	0.9880	0.9890	**0.9960**	0.9640	0.9820
		β_3	0.9990	0.9930	0.9910	**0.9995**	0.9715	0.9960
100	0.9	β_1	0.9895	0.9895	0.9900	**0.9910**	0.9900	0.9870
		β_2	**0.9920**	0.9900	0.9910	0.9910	0.9900	0.9690
		β_3	0.9950	**0.9980**	0.9910	0.9970	0.9950	0.9890
	0.99	β_1	0.9890	0.9900	**0.9910**	0.9900	**0.9910**	0.9900
		β_2	**0.9960**	0.9840	0.9890	0.9940	0.9680	0.9640
		β_3	0.9970	0.9930	0.9940	**0.9980**	0.9800	0.9970

Table 3.3. Average CW of different methods at 95% confidence level

n		*Coeff*	*OLSE*	*Normal*	*Percentile*	*Studentized*	BC_a	*Jackknife*
25	0.9	β_1	2.391818	**2.262103**	2.301718	2.369512	2.373546	2.829685
		β_2	2.266746	**2.243227**	2.266746	2.310391	2.271629	2.523521
		β_3	2.376209	**2.358398**	2.376209	2.424493	2.397483	2.784828
	0.99	β_1	**6.576237**	6.597696	6.679237	6.615172	6.70064	10.41938
		β_2	6.360067	**6.209885**	6.360067	6.509726	6.514881	10.100110
		β_3	6.761487	**6.607427**	6.761487	6.952358	6.890519	10.748250
50	0.9	β_1	1.520936	**1.350361**	1.533366	1.379252	1.473646	1.440304
		β_2	1.423707	**1.416494**	1.423707	1.448140	1.423807	1.449212
		β_3	1.549024	**1.536868**	1.549024	1.574883	1.552204	1.577136
	0.99	β_1	4.207209	4.132672	**4.117029**	4.271159	4.237864	5.794998
		β_2	3.918245	**3.817233**	3.918245	4.005388	3.961890	5.394279
		β_3	4.272753	**4.174830**	4.272753	4.353670	4.343077	6.173637
100	0.9	β_1	1.200137	1.099692	1.246077	1.162913	1.087877	**0.993889**
		β_2	1.1070680	**1.1049590**	1.1070680	1.1261430	1.1072070	1.1064920
		β_3	1.0747450	**1.0698050**	1.0747450	1.0902820	1.0749470	1.0783050
	0.99	β_1	3.095477	**2.868377**	2.887837	3.000262	3.063632	4.161706
		β_2	3.039237	**2.980257**	3.039237	3.099931	3.058685	4.054548
		β_3	2.962096	**2.919236**	2.962096	3.008812	2.987658	3.947704

Table 3.4. Average CW of different methods at 99% confidence level

n		*Coeff.*	*OLSE*	*Normal*	*Percentile*	*Studentized*	BC_a	*Jackknife*
25	0.9	β_1	3.162719	3.042074	**3.004419**	3.335236	3.0425	3.61399
		β_2	3.062997	**2.953371**	3.062997	3.089409	3.081329	3.441042
		β_3	3.199281	**3.103537**	3.199281	3.222922	3.225771	3.767778
	0.99	β_1	9.208246	**8.480646**	8.979106	9.101246	9.182618	14.49355
		β_2	8.767475	**8.200895**	8.767475	8.902586	8.778653	14.139590
		β_3	9.452596	**8.698177**	9.452596	9.617346	9.437943	14.784770
50	0.9	β_1	2.094543	**1.840617**	1.862544	1.982551	2.072696	2.040974
		β_2	1.898971	**1.861498**	1.898971	1.920025	1.902575	1.940650
		β_3	2.067215	**2.020396**	**2.020396**	2.089556	2.071636	2.123598
	0.99	β_1	5.762919	**5.119859**	5.573769	5.694366	5.560966	7.843586
		β_2	5.411404	**5.016694**	5.411404	5.504343	5.442553	7.198365
		β_3	5.923213	**5.497523**	5.923213	6.021348	5.927179	8.452887
100	0.9	β_1	**1.401126**	1.483682	1.453336	1.545871	1.54875	1.482187
		β_2	1.4702110	**1.4521620**	1.4702110	1.4972630	1.4711580	1.4648250
		β_3	1.4290400	**1.4059620**	1.4290400	1.4508380	1.4301210	1.4275090
	0.99	β_1	4.055026	**3.888134**	4.092746	4.189925	4.220521	5.251085
		β_2	4.1637620	**3.9167220**	4.1637620	4.2597010	4.1888450	5.3675970
		β_3	4.0324100	**3.8365260**	4.0324100	4.1057080	4.0581560	5.2261520

3.5. Conclusions

In this chapter, several methods to estimate the CI of the regression parameters using the resampling methods have been proposed. Then, the CIs of different Bootstrap methods and the Jackknife method are compared with that of the OLSE using the CP and the average CW. The simulation study showed different results according to the used criterion, where it showed, according to the CP criterion, that there is always at least one Bootstrap methods (Normal, Percentile, Studentized or BC_a) that gives results better than the OLSE. While according to CW, it showed the superiority of the Normal method over the OLSE and the other Bootstrap methods. While the Jackknife method did not give good results compared to the other resampling methods. All of this leads to recommending the use of the Normal theory method to estimate the regression parameters of the LE.

Finally, it should be mentioned that the noticed behavior of the different resampling methods in this chapter may differ according to many factors, like using the resampling of the observations (pairs resampling) instead of the resampling of the errors, or using a different method of calculating the LP d, like using the resampling methods to estimate it instead of the formulas that are used to estimate the value of d in this chapter.

4. BOOTSTRAP SELECTION OF THE LP

In this chapter, different techniques for the selection of the LP using Bootstrap and cross-validation are intended to be developed and compared. The procedure is illustrated by application to published real data sets. A simulation study comparing several techniques is also described. Analysis strategies are suggested for data with various degrees of multicollinearity and varying levels of SNR. The traditional choices for LP d are given in Section 4.1. In Section 4.2, an optimal choice for the LE parameter d is given using the Bootstrap resampling method. In Section 4.3, the methods are applied to simulated data and their performance is compared based on EMSE. Analysis of real data is given in Section 4.4. Discussion and conclusions in Section 4.5 complete this chapter.

4.1. The Traditional Choices of a LP

Several methods have been proposed in the literature to find the biasing parameter of the LE. For this purpose, Kejian (1993) first obtained some of them similar to the k finding methods in the RRE. These are:

$\hat{d}_{MM}$ or $\hat{d}_{MMH}$ (minimum MSE method), $\hat{d}_{MG}$ (Mcdonald-Galarneau method) and $\hat{d}_{ITR}$ (iterative method).

In such cases, prediction-oriented criterion for choosing the k value should be considered. Allen (1971) introduced the PRESS statistic for good prediction and suggested to find such that the PRESS statistic is minimum. Then Golub et al. (1979) proposed the GCV statistic as a rotation-invariant version of Allen's PRESS statistics. The C_p statistic given by Mallows (1973) is another prediction-oriented criterion to evaluate the prediction performance of a model.

Similarly, the prediction-oriented criterion for choosing the d value should be first considered by Kejian (1993). These are:

$\hat{d}_{CL}$ (Mallows C_p method) and $\hat{d}_{GCV}$ (Golub et al. GCV statistic). Then Özkale and Kaçıranlar (2007) proposed Allen's PRESS statistic for choosing the d value and they obtained $\hat{d}_{PRESS}$. For the reason that will be explained later, the prediction-oriented criterion will be considered for choosing the d value.

First, the C_L criterion is considered. The C_L is defined as

$$C_L = \frac{SS_{\mathrm{Res},d}}{\hat{\sigma}^2} + 2tr(H_d) - (n-2)$$

where $H_d = X(X'X+I)^{-1}(X'X+dI)X'$ and $SS_{\mathrm{Re}s,d}$ is the residual sum of squares using $\hat{\alpha}(\mathrm{d})$. d can be chosen by minimizing C_L.Thus, the formula for $\hat{d}_{CL}$ is given below in canonical form:

$$\hat{d}_{CL} = 1 - \hat{\sigma}^2 \left[\sum_{i=1}^{p} \frac{1}{(\lambda_i + 1)} \Big/ \sum_{i=1}^{p} \frac{\lambda_i \hat{\alpha}_i^2}{(\lambda_i + 1)^2} \right]. \tag{4.1}$$

The generalized version of $\hat{d}_{CL}$ is also given as $\hat{d}_{CLh}$

$$\hat{d}_{CLh} = 1 - h\hat{\sigma}^2 \left[\sum_{i=1}^{p} \frac{1}{(\lambda_i + 1)} \Big/ \sum_{i=1}^{p} \frac{\lambda_i \hat{\alpha}_i^2}{(\lambda_i + 1)^2} \right]. \tag{4.2}$$

Assume h satisfy

$$0 < h < \frac{2(n-p)}{n-p+2} \left(\sum_{i=1}^{p} \frac{\lambda_p}{\lambda_i(\lambda_i+1)} - \frac{2}{(\lambda_p+1)} \right) \Big/ \sum_{i=1}^{p} \frac{1}{(\lambda_i+1)}. \tag{4.3}$$

then for all α and σ^2 , $MSE(\hat{\alpha}(\hat{d}_{CLh})) < MSE(\hat{\alpha})$. $\hat{\sigma}^2$ given above is an unbiased estimate of σ^2 in canonical form.

GCV criterion is another good method to choose d . GCV is defined $GCV_L = \frac{SS_{Res,d}}{(n-1-tr(H_d))}$. d can be chosen by minimizing GCV_L. Kejian (1993) showed that $\hat{d}_{GCV}$ is just $\hat{d}_{CL}$.

The PRESS statistic for the LE is defined as

$$PRESS_d = \sum_{i=1}^{n} (\hat{e}_{d(i)})^2 = \sum_{i=1}^{n} \left[(1-d) \frac{\hat{e}_{Ri(1)}}{1-h_{1-ii}} + d \frac{\hat{e}_i}{1-h_{ii}} \right]^2. \tag{4.4}$$

The estimation of the value that minimizes statistic is given by Özkale and Kaçıranlar (2007) as follows:

$$\hat{d}_{PR} = \frac{\sum_{i=1}^{n} \frac{\hat{e}_{Ri}(1)}{1-h_{1-ii}} \left(\frac{\hat{e}_{Ri}(1)}{1-h_{1-ii}} - \frac{\hat{e}_i}{1-h_{ii}} \right)}{\sum_{i=1}^{n} \left(\frac{\hat{e}_{Ri}(1)}{1-h_{1-ii}} - \frac{\hat{e}_i}{1-h_{ii}} \right)^2}. \tag{4.5}$$

Here the following definitions and abbreviations are provided:

$$\hat{e}_{(i)} = y_i - \hat{y}_{(i)} = \frac{y_i - \hat{y}_i}{1-h_{ii}} = \frac{\hat{e}_i}{1-h_{ii}}, \; i = 1, 2, ..., n,$$

$\hat{y}_{(i)}$ is found by withholding the ith observation and fitting the regression model to the remaining observations and using this equation to predict y_i,

$\hat{e}_i = y_i - \hat{y}_i$ is the ordinary residual and h_{ii} is the ith diagonal element of the hat matrix $H = X(X'X)^{-1}X'$,

and $\hat{e}_{Ri}(k) = y_i - x_i'\hat{\beta}(k)$ is the ith ridge residual for a specific value of k and h_{k-ii} is the ith diagonal element of $H_k = X(X'X + kI)^{-1}X'$.

It is seen that $\hat{d}_{CL}$ needs the unbiased estimate of σ^2, however, $\hat{d}_{PR}$ does not. The analyst considers the purpose of the regression. If estimation is important $\hat{d}_{MM}, \hat{d}_{MG}, \hat{d}_{ITR}$ or other methods should be used. If prediction is important $\hat{d}_{CL}$ and $\hat{d}_{PR}$ should be attempted.

4.2. The Bootstrap Choice of a LP

4.2.1. The Bootstrap

Bootstrapping is a resampling method that proposes to solve some statistical problems that are difficult or impossible to solve using the traditional statistical methods. It helps in calculating the bias, the SE, the CI, and statistical testing of very wide types of problems. The Bootstrap was first proposed by Efron (1979), who later proposed wide applications of the resampling methods. Delaney and Chatterjee (1986) proposed a Bootstrap method to select the ridge regression parameter.

In regression analysis two Bootstrap methods were proposed, the Bootstrap of residuals, which withdraws the samples from the residuals, and the Bootstrap of pairs, which withdraws the samples directly from the observations (see Efron, 1979; and Efron and Tibshirani, 1994).

In this chapter, the method of Bootstrap of residuals is used. Now for Bootstrapping, for the model defined in (3.1), the OLSE is fitted for full sample, the residuals $\hat{e}_i$ are generated from the normal distribution with zero average and σ^2 of sample of size n, and this is repeated to have B Bootstrap samples $(\hat{e}_1^b, \hat{e}_2^b, ..., \hat{e}_n^b)$. After this, the Bootstrap y values are obtained using the resampled residuals keeping the design matrix fixed as shown in (3.17).

These Bootstrapped y values are regressed on the fixed to obtain the Bootstrap estimates of the RCs. So, the OLSE from the first Bootstrap sample is given by

$$\hat{\beta}_{(b_1)} = (X'X)^{-1}X'y_{(b_1)}. \tag{4.6}$$

Repeating the above steps B times, where B is the number of Bootstrap samples, the Bootstrap OLSE is given by

$$\bar{\hat{\beta}}_{(b)} = \frac{1}{B}\sum_{r=1}^{B}\hat{\beta}_{(b_r)}. \tag{4.7}$$

4.2.2. The Bootstrap Choice of the LP

Following Delaney and Chatterjee (1986), a combination of Bootstrap and cross-validation methods has been used to select the optimal biasing parameter of the LE based on MSE prediction (MSEP), which can be illustrated by the following algorithm:

A population of n observations on p independent variables and one dependent variable is available. A Bootstrap random sample of size n is chosen with replacement from this population. For this Bootstrap sample, the Liu estimates of the parameter vector are computed for each value of a selected set of d values. Suppose that G values of d are used for each Bootstrap sample. The unchosen observations are predicted from the estimates obtained. Let the prediction vector for the unchosen be $\hat{y}_{K_j}(d_g)$, where the subscript K_j indicates the number of unchosen observations when the j^{th} sample was selected and the subscript on d represents the g^{th} value of d used. Let y_{K_j} be the vector of the actual values of the response variable for this unchosen set. The MSEP for the j^{th} Bootstrap sample with $d = d_g$ will be

$$MSEP_j(d_g) = (\hat{y}_{K_j}(d_g) - y_{K_j})'(\hat{y}_{K_j}(d_g) - y_{K_j})/K_j,\ g = 1,2,...,G. \tag{4.8}$$

The foregoing procedure is repeated for B Bootstrap samples and final measures of prediction errors and spread are computed. A final MSEP is obtained as follows:

$$MSEP(d_g) = \sum_{j=1}^{B} (MSEP_j(d_g))(K_j) / \sum_{j=1}^{B} K_j \ , \ g = 1, 2, ..., G. \tag{4.9}$$

The optimal value of d is $d_{boot} = \min(MSEP(d_g))$.

4.3. Monte Carlo Simulation

This section will use generated data to test the proposed methodology of choosing the optimal value of the Liu biasing parameter. The simulation will take a variety of levels of many factors: condition number, parameter vector length, and SNR.

The simulation starts with determining the preliminary factors which are the sample size $n = 50$, the number of explanatory variables $p = 5$, the number of Bootstrap samples $B = 100$, and the levels of multicollinearity $\rho = 0.85, 0.90, 0.95, 0.99$. Then simulate the explanatory variables using the formula (3.33). While the parameter vector is computed as:

$$\beta_j = (\|\beta\| / r_u) u_j \tag{4.10}$$

which is taken as Delaney and Chatterjee (1986), where $\|\beta\|^2 = \beta'\beta$, which measures the strength of the signal and is taken for different values 4, 16, 100, and 900, u_j is generated from the uniform distribution of the interval $[-1,1]$ and $r_u^2 = \sum_{j=1}^{P} u_j^2$.

The errors ε_i is taken from the normal distribution with zero expectation and σ^2 variance. The values of the variance are computed from $SNR = \frac{\beta'\beta}{\sigma^2}$, which means they are calculated using SNR and $\|\beta\|^2$. The values of SNR are taken as 1, 4, 9, 25, 400 which respectively decrease poor fit regions for the OLSE.

According to the above generated values of $X, \beta,$ and ε, the observations of the response variable can be calculated using the linear model (3.1).

Finally, the steps above are repeated 500 times, which represents the Bootstrap samples of errors, by regenerating the errors, and calculate the estimated MSE (EMSE) to evaluate the estimators:

$$EMSE(\hat{\beta}) = \frac{1}{500} \sum_{b=1}^{500} (\hat{\beta}_b - \beta)'(\hat{\beta}_b - \beta)$$

where $\hat{\beta}_b$ is the estimator of β of the b^{th} Bootstrap sample.

Tables 1 to 4 show the values of EMSE of the OLSE and LEs of different methods of calculating LP in addition to Bootstrap method for the first regression coefficient β_1.

It can be noticed that:

1. The LE based on Bootstrap method in general gives significantly better results than the LE based and the other methods and the OLSE estimator. While the other methods give better results in very few cases, which appear when $SNR = 1\ or\ 400$, and specific values of ρ.
2. In most of the cases, for different values of the $\beta'\beta$, *SNR*, the value of EMSE significantly increases with the increasing of ρ. This shows the high effect of the multicollinearity on the estimation accuracy for the LE.
3. The LE based on d_{PR} generally gives worse results than the other methods, even, in some cases, worse than the OLSE estimator. But in very few cases it gives better results than the others, which is noticed when SNR becomes lower than 9.
4. There exist a few cases where the OLSE estimator gives better results than the LE based on different methods, even the Bootstrap method, which shows the significant superiority of LE methods over the OLSE estimator.

Table 4.1. EMSE values of OLSE and LEs with $\beta'\beta = 900$ for β_1

SNR	$\boldsymbol{\rho}$	***OLSE***	$\boldsymbol{d_{boot}}$	$\boldsymbol{d_{CL}}$	$\boldsymbol{d_{CLh}}$	$\boldsymbol{d_{PR}}$
1	0.99	6146.572	**1007.893**	7309.443	4537.999	3815.259
	0.95	1151.537	595.8612	932.3586	925.3448	**578.2413**
	0.90	549.741	384.9189	400.1752	466.705	**377.2942**
	0.85	354.0663	279.9427	**263.9361**	311.9277	276.9946
4	0.99	1536.643	**403.4332**	1091.489	1246.214	1606.696
	0.95	287.8843	**169.2858**	211.1126	259.4708	180.2975
	0.90	137.4352	**102.9938**	114.3717	129.2337	104.1874
	0.85	88.51659	**73.30137**	78.75393	84.9759	73.60027
9	0.99	682.9525	**286.5027**	453.1298	596.1303	1372.261
	0.95	127.9486	**91.12466**	106.8803	121.1675	106.6467
	0.90	61.08233	**50.77517**	55.96345	59.34752	52.40043
	0.85	39.34071	**34.94794**	37.43641	38.64213	35.42874
25	0.99	245.8629	**161.9136**	195.8635	230.9408	680.9874
	0.95	46.06149	**40.11314**	42.97611	45.1175	62.6744
	0.90	21.98964	**20.55715**	21.37628	21.77346	27.92421
	0.85	14.16265	**13.61856**	13.96282	14.08055	16.38885
400	0.99	15.36643	**14.96543**	15.12345	15.30088	34.42649
	0.95	2.878843	**2.868349**	2.871825	2.876129	7.332757
	0.90	**1.374352**	1.375089	1.374412	1.37398	3.642672
	0.85	0.8851659	0.8869082	0.8857244	**0.8851039**	2.32605

Table 4.2. EMSE values of OLSE and LEs with $\beta'\beta = 100$ for β_1

SNR	$\boldsymbol{\rho}$	***OLSE***	$\boldsymbol{d_{boot}}$	$\boldsymbol{d_{CL}}$	$\boldsymbol{d_{CLh}}$	$\boldsymbol{d_{PR}}$
1	0.99	682.9525	**111.9882**	812.1603	504.2221	423.9177
	0.95	127.9486	66.2068	103.5954	102.8161	**64.24903**
	0.90	61.08233	42.76877	44.46391	51.85612	**41.92157**
	0.85	39.34071	31.10474	**29.32623**	34.65864	30.77718
4	0.99	170.7381	**44.82591**	121.2766	138.4683	178.5217
	0.95	31.98715	**18.80953**	23.45695	28.83009	20.03305
	0.90	15.27058	**11.44376**	12.70797	14.3593	11.57638
	0.85	9.835176	**8.144596**	8.750437	9.441766	8.177808
9	0.99	75.88361	**31.83363**	50.34776	66.2367	152.4734
	0.95	14.21651	**10.12496**	11.87559	13.46305	11.84964
	0.90	6.786926	**5.641685**	6.218161	6.594169	5.82227
	0.85	4.371189	**3.883105**	4.159602	4.293569	3.936526
25	0.99	27.3181	**17.9904**	21.76261	25.66008	75.66527
	0.95	5.117944	**4.457015**	4.775123	5.013056	6.963822
	0.90	2.443293	**2.284127**	2.375142	2.419274	3.10269
	0.85	1.573628	**1.513174**	1.551425	1.564505	1.820983
400	0.99	1.707381	**1.662826**	1.680384	1.700098	3.825165
	0.95	0.3198715	**0.3187054**	0.3190916	0.3195699	0.8147507
	0.90	0.1527058	0.1527877	0.1527125	**0.1526644**	0.4047413
	0.85	0.09835176	0.09854536	0.09841382	**0.09834488**	0.2584501

Table 4.3. EMSE values of OLSE and LEs with $\beta'\beta = 16$ for β_1

SNR	$\boldsymbol{\rho}$	***OLSE***	$\boldsymbol{d_{boot}}$	$\boldsymbol{d_{CL}}$	$\boldsymbol{d_{CLh}}$	$\boldsymbol{d_{PR}}$
1	0.99	109.2724	**17.9181**	129.9457	80.67554	67.82683
	0.95	20.47178	10.59309	16.57526	16.45057	**10.27984**
	0.90	9.773173	6.843004	7.114226	8.296978	**6.707452**
	0.85	6.294513	4.976758	**4.692197**	5.545382	4.924348
4	0.99	27.3181	**7.172145**	19.40426	22.15492	28.56348
	0.95	5.117944	**3.009525**	3.753112	4.612814	3.205288
	0.90	2.443293	**1.831002**	2.033275	2.297488	1.85222
	0.85	1.573628	**1.303135**	1.40007	1.510683	1.308449
9	0.99	12.14138	**5.093381**	8.055641	10.59787	24.39575
	0.95	2.274642	**1.619994**	1.900094	2.154088	1.895942
	0.90	1.085908	**0.9026696**	0.9949058	1.055067	0.9315633
	0.85	0.6993903	**0.6212968**	0.6655363	0.6869711	0.6298442
25	0.99	4.370896	**2.878465**	3.482018	4.105614	12.10644
	0.95	0.818871	**0.7131224**	0.7640197	0.8020889	1.114211
	0.90	0.3909269	**0.3654604**	0.3800227	0.3870838	0.4964304
	0.85	0.2517805	**0.2421078**	0.248228	0.2503209	0.2913573
400	0.99	0.273181	**0.2660521**	0.2688614	0.2720157	0.6120264
	0.95	0.05117944	**0.05099287**	0.05105466	0.05113119	0.1303601
	0.90	0.02443293	0.02444603	0.02443399	**0.02442631**	0.06475861
	0.85	0.01573628	0.01576726	0.01574621	**0.01573518**	0.04135201

Table 4.4. EMSE values of OLSE and LEs with $\beta'\beta = 4$ for β_1

SNR	$\boldsymbol{\rho}$	***OLSE***	$\boldsymbol{d_{boot}}$	$\boldsymbol{d_{CL}}$	$\boldsymbol{d_{CLh}}$	$\boldsymbol{d_{PR}}$
1	0.99	27.3181	**4.479526**	32.48641	20.16889	16.95671
	0.95	5.117944	2.648272	4.143816	4.112643	**2.569961**
	0.90	2.443293	1.710751	1.778557	2.074245	**1.676863**
	0.85	1.573628	1.24419	**1.173049**	1.386346	1.231087
4	0.99	6.829525	**1.793036**	4.851064	5.538731	7.140869
	0.95	1.279486	**0.7523813**	0.9382781	1.153203	0.8013221
	0.90	0.6108233	**0.4577504**	0.5083189	0.574372	0.4630551
	0.85	0.3934071	**0.3257839**	0.3500175	0.3776707	0.3271123
9	0.99	3.035344	**1.273345**	2.01391	2.649468	6.098937
	0.95	0.5686604	**0.4049985**	0.4750235	0.5385221	0.4739854
	0.90	0.271477	**0.2256674**	0.2487264	0.2637667	0.2328908
	0.85	0.1748476	**0.1553242**	0.1663841	0.1717428	0.157461
25	0.99	1.092724	**0.7196161**	0.8705045	1.026403	3.026611
	0.95	0.2047178	**0.1782806**	0.1910049	0.2005222	0.2785529
	0.90	0.09773173	**0.09136509**	0.09500568	0.09677095	0.1241076
	0.85	0.06294513	**0.06052694**	0.06205699	0.06258021	0.07283933
400	0.99	0.06829525	**0.06651303**	0.06721535	0.06800391	0.1530066
	0.95	0.01279486	**0.01274822**	0.01276367	0.0127828	0.03259003
	0.90	0.00610823	0.00611150	0.00610849	**0.00610657**	0.0161896
	0.85	0.00393407	0.00394181	0.00393655	**0.00393379**	0.0103380

4.4. Numerical Examples

This section will be used to illustrate the application of Bootstrap selection of LP on real data. The used data is the one that was first published by McDonald and Schwing (1973), which is the pollution data that contains measurements of death rates and 15 explanatory variables, which was measured in 1960 for 60 cities. This data is very useful for our example because of the high multicollinearity between the explanatory variables.

To apply the Bootstrap choice of the LP, 20 random values of d are taken from 1000 values ranging between 0 and 1, with 0.001 steps starting from 0. To calculate the MSEP 1000 Bootstrap samples are taken from the original data.

Table 4.5 shows the results taking many measures, the average MSEP, the 10% trimmed mean and the median, as measures to show the central tendency, in addition to the standard deviation, and the interquartile range, as measures to show the disturbance. It looks clear that the optimal value of the LP is 0.059, where it gives the minimum average MSEP, the minimum trimmed mean, and the minimum median. Although the lowest standard deviation of MSEP was for $d = 0.006$, and the lowest interquartile range of MSEP was for $d = 0.234$, but still $d = 0.059$ gives a very small standard deviation, and interquartile range.

It is clear that as the optimal value of d in the pollution data is close to zero, and MSEP is increasing as the d value increases.

Table 4.5. Results of Bootstrap selection of LP of pollution data

d	*Average MSEP*	*Trimmed mean (10%)*	*Median*	*Standard deviation*	*Interquartile range*
0.003	0.3707446	0.3303411	0.3044844	0.2822042	0.2772547
0.006	0.3538599	0.3185576	0.2947091	**0.2464653**	0.2729612
0.059	**0.3523860**	**0.3093311**	**0.2852495**	0.2678338	0.2702080
0.075	0.3587693	0.3212027	0.2920950	0.2552163	0.2802074
0.178	0.3764608	0.3317682	0.3057383	0.3226190	0.3023846
0.234	0.3660490	0.3200702	0.2955919	0.3127215	**0.2680876**
0.268	0.3634056	0.3253655	0.3025716	0.2568784	0.2811946
0.320	0.3780066	0.3299956	0.2991301	0.3184698	0.2902588
0.391	0.3815020	0.3329022	0.2942947	0.3106944	0.3098436
0.456	0.3807705	0.3314086	0.3030178	0.3236999	0.2816544
0.481	0.3850935	0.3323185	0.3022022	0.3321053	0.2998378
0.513	0.3869669	0.3434636	0.3117986	0.2993364	0.3158572
0.572	0.3929036	0.3326452	0.3088122	0.3815432	0.2911468
0.580	0.3946519	0.3424693	0.3095212	0.3582413	0.2943462
0.601	0.4259916	0.3597993	0.3130004	0.4159597	0.3121357
0.724	0.4317582	0.3497542	0.3124139	0.4973111	0.3183867
0.733	0.4081662	0.3522636	0.3199372	0.3438287	0.3115297
0.751	0.4132067	0.3527873	0.3148328	0.3714076	0.3175184
0.860	0.4558233	0.3583377	0.3236607	0.6313077	0.3112038
0.903	0.4621061	0.3814656	0.3449560	0.4715295	0.3629556

In order to double check the results obtained in Table 4.5, the experiment above is repeated for 1000 times by using different seed numbers randomly selected from the range [1,100000], to show the convergence of the optimal d value. Chart 4.1 shows the histogram of the d values that give the minimum average MSEP of 1000 different experiments corresponding to 1000 different seed numbers of the pollution data. The chart shows that there is a convergence to the values ranging between [0.026, 0.076], which includes the d value previously obtained.

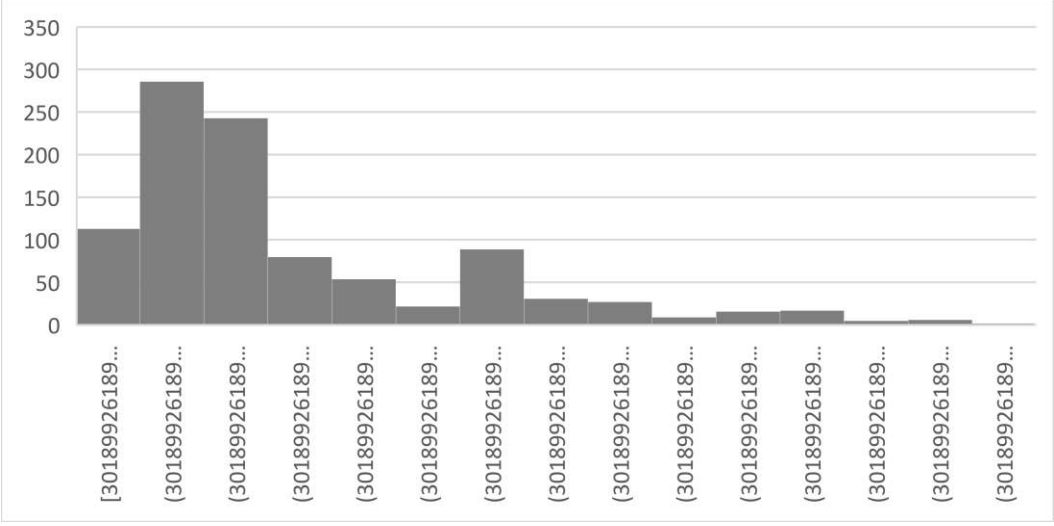

Figure 4.1. Values of LP corresponding to minimum average MSEP

4.5. Discussion and Conclusions

It is clear that the selection of the LP is seriously affecting the results of the LE, and while many literature methods to calculate the LP have been proposed, it is not that easy to discover which one of these methods gives the optimal value. The proposed Bootstrap method to select the optimal value of the LP in this chapter uses the advantage of the Bootstrapping which can give many measures to compare the accuracy of the results, which guarantees selecting better d value as shown in our both examples (Monte Carlo simulation, and real data).

It is concluded that the values of $\beta'\beta$, multicollinearity and SNR do not have that much effect on the d selection methods, as the Bootstrap method in most cases gives better results than the other methods. Also d_{PR} gives results worse than the other methods, but in very few cases it gives better results than the others. While all LE methods give better results compared to the OLSE.

5. SELECTION OF LP USING THE CI AND BOOTSTRAP METHODS

The OLSE tends to be inaccurate in the occurrence of multicollinearity, its estimates give a large variance. Biased estimates have been proposed as a solution for this issue, one of which is the LE. This chapter aims to propose a method that helps choose the optimal value of the LP of the LE, which will be through using two criteria, the average CW, and the CP of the CI that is estimated using the Bootstrap methods. Section 5.1 talks about the CI estimation methods that will be used in this chapter. Following that, section 5.2 includes a simulation study to apply the proposed methodology. Finally, the chapter ends with a discussion of the conclusion in section 5.3.

5.1. The LP and the CI

The CI reflects in some way the concept of the statistical tests, this fact gives a useful application of the CI as a tool of testing and comparing different statistical methods, in addition to that, the availability of the Bootstrap methods that have been explained first by (Efron, 1979), made an extraordinary step in the use of CI in statistical application.

In chapter (3), the use of the Bootstrap method to estimate the CI of the Liu RCs is discussed, where many methods of estimating the CI were compared, such as the Normal theory, the Jackknife, the Percentile, the Studentized-t, and the BC_a methods. The regression errors Bootstrapping method was used to generate replications that helped calculate the CP and average CW for different methods of estimating the CI.

Our work uses the Bootstrap of the residuals that explained in section 4.2.1, to get the Bootstrap variance estimation given in Section 3.3 which helps estimate the CI for the RCs of the LE. Many methods of estimating the CI of the RCs were proposed, such as the Normal theory method, the studentized-t method, the percentile method, the ABC method…etc. In this chapter, the Normal theory method, the studentized-t method, the percentile method, and the BC_a method that was explained in section 3.3.1, are used.

5.2. Simulation Study

The simulation study is used to show the differences between the different methods of estimating the LP that are discussed before, $\hat{d}_{CL}$, $\hat{d}_{CLh}$, $\hat{d}_{PR}$, and $\hat{d}_{boot}$, these will be compared by studying their effect on the CI of the regression parameters, where four different methods to estimate the CI were used, which are the Normal method, the Percentile method, the Studentized-t method, and the BC_a method. To compare the CIs, two criteria were used; the first one is the average width, which compares the widths of the CIs, where it considers that the narrower width of the CI is, the better the estimate method is, while the second criterion is the CP, which says that the higher percentage of parameters occur inside the CI the better estimate method will be.

The simulation study will use the same methods that explained in section 4.3. where the correlation will be ρ, which its values were used are 0.85, 0.90, 0.95, and 0.99. and SNR takes many values that are 4, 9, 25, and 100, and $\|\beta\|^2 = \beta'\beta$ will take the values 4, 16, and 100.

Finally, the values of the dependent variable y are calculated using (5.17), where the number of explanatory variables is $p = 3$, the sample size $n = 100$, the number of the Bootstrap samples $B = 500$, and the number of repetitions is 300.

The following tables from table 5.1 to table 5.24 show the results for the CP and CW for the first regression coefficient β_1:

Coverage Probability:

Table 5.1. Normal Theory's CP for OLSE and LEs with $\beta'\beta = 4$ for β_1

SNR	$\boldsymbol{\rho}$	***OLSE***	$\boldsymbol{d_{boot}}$	$\boldsymbol{d_{CL}}$	$\boldsymbol{d_{CLh}}$	$\boldsymbol{d_{PR}}$
4	0.99	**0.9100000**	0.9000000	0.8933333	**0.9100000**	0.9066667
	0.95	**0.9100000**	0.8933333	0.9000000	**0.9100000**	0.9066667
	0.90	**0.9100000**	0.8966667	0.9000000	**0.9100000**	**0.9100000**
	0.85	0.9100000	0.8966667	0.9000000	**0.9133333**	0.9100000
9	0.99	0.8800000	0.8733333	0.8666667	**0.9000000**	0.8933333
	0.95	0.8800000	0.8733333	0.8633333	**0.8966667**	**0.8966667**
	0.90	0.8800000	0.8733333	0.8666667	**0.8966667**	**0.8966667**
	0.85	0.8800000	0.8700000	0.8666667	**0.8966667**	**0.8966667**
25	0.99	0.8033333	0.7466667	0.7233333	**0.8366667**	0.8066667
	0.95	0.8033333	0.7666667	0.7533333	**0.8266667**	0.8066667
	0.90	0.8033333	0.7700000	0.7600000	**0.8266667**	0.8066667
	0.85	0.8033333	0.7766667	0.7633333	**0.8266667**	0.8066667
100	0.99	0.4633333	0.3733333	0.3700000	**0.4933333**	0.4700000
	0.95	0.4700000	0.4100000	0.4066667	**0.4866667**	0.4733333
	0.90	0.4700000	0.4166667	0.4133333	**0.4866667**	0.4733333
	0.85	0.4700000	0.4266667	0.4233333	**0.4866667**	0.4733333

Table 5.2. Studentized-t theory's CP for OLSE and LEs with $\beta'\beta = 4$ for β_1

SNR	$\boldsymbol{\rho}$	***OLSE***	$\boldsymbol{d_{boot}}$	$\boldsymbol{d_{CL}}$	$\boldsymbol{d_{CLh}}$	$\boldsymbol{d_{PR}}$
4	0.99	0.9033333	0.9133333	0.9000000	**0.9300000**	0.9133333
	0.95	0.9033333	0.9066667	0.9100000	**0.9333333**	0.9133333
	0.90	0.9033333	0.9066667	0.9133333	**0.9300000**	0.9133333
	0.85	0.9033333	0.9066667	0.9133333	**0.9300000**	0.9133333
9	0.99	0.8933333	0.8700000	0.8800000	**0.8966667**	0.8866667
	0.95	0.8933333	0.8800000	0.8800000	**0.8933333**	0.8866667
	0.90	0.8933333	0.8833333	0.8800000	**0.8933333**	0.8866667
	0.85	0.8933333	0.8833333	0.8800000	**0.8933333**	0.8866667
25	0.99	0.8133333	0.7666667	0.7633333	**0.8566667**	0.8100000
	0.95	0.8133333	0.7833333	0.7866667	**0.8533333**	0.8133333
	0.90	0.8133333	0.7833333	0.7900000	**0.8500000**	0.8133333
	0.85	0.8133333	0.7866667	0.7966667	**0.8466667**	0.8133333
100	0.99	0.4833333	0.4100000	0.4100000	**0.5100000**	0.4900000
	0.95	0.4900000	0.4433333	0.4533333	**0.5066667**	0.4933333
	0.90	0.4900000	0.4566667	0.4600000	**0.5066667**	0.4966667
	0.85	0.4900000	0.4566667	0.4666667	**0.5066667**	0.5000000

Table 5.3. Percentile's CP for OLSE and LEs with $\beta'\beta = 4$ for β_1

SNR	$\boldsymbol{\rho}$	***OLSE***	$\boldsymbol{d_{boot}}$	$\boldsymbol{d_{CL}}$	$\boldsymbol{d_{CLh}}$	$\boldsymbol{d_{PR}}$
4	0.99	**0.9133333**	0.9066667	0.8933333	**0.9133333**	0.9100000
	0.95	**0.9133333**	0.9033333	0.8966667	**0.9133333**	**0.9133333**
	0.90	**0.9133333**	0.9033333	0.8966667	**0.9133333**	**0.9133333**
	0.85	**0.9133333**	0.9033333	0.8966667	**0.9133333**	**0.9133333**
9	0.99	0.8733333	0.8633333	0.8566667	**0.9000000**	0.8766667
	0.95	0.8733333	0.8633333	0.8700000	**0.8966667**	0.8766667
	0.90	0.8733333	0.8600000	0.8700000	**0.8966667**	0.8800000
	0.85	0.8733333	0.8600000	0.8733333	**0.8966667**	0.8800000
25	0.99	0.8033333	0.7366667	0.7266667	**0.8300000**	0.8233333
	0.95	0.8033333	0.7566667	0.7566667	0.8200000	**0.8233333**
	0.90	0.8066667	0.7566667	0.7666667	0.8166667	**0.8233333**
	0.85	0.8066667	0.7633333	0.7733333	0.8100000	**0.8233333**
100	0.99	0.4800000	0.4000000	0.3833333	**0.5066667**	0.4500000
	0.95	0.4800000	0.4233333	0.4100000	**0.4966667**	0.4533333
	0.90	0.4800000	0.4300000	0.4233333	**0.4966667**	0.4533333
	0.85	0.4800000	0.4333333	0.4233333	**0.4966667**	0.4533333

Table 5.4. BC_a's CP for OLSE and LEs with $\beta'\beta = 4$ for β_1

SNR	$\boldsymbol{\rho}$	***OLSE***	$\boldsymbol{d_{boot}}$	$\boldsymbol{d_{CL}}$	$\boldsymbol{d_{CLh}}$	$\boldsymbol{d_{PR}}$
4	0.99	0.9133333	0.9033333	0.8966667	**0.9266667**	0.9166667
	0.95	0.9133333	0.9066667	0.8933333	**0.9266667**	0.9166667
	0.90	0.9133333	0.9066667	0.8966667	**0.9233333**	0.9166667
	0.85	0.9100000	0.9066667	0.8933333	**0.9233333**	0.9166667
9	0.99	0.8800000	0.8666667	0.8566667	0.8966667	**0.9000000**
	0.95	0.8833333	0.8766667	0.8766667	**0.9033333**	0.9000000
	0.90	0.8833333	0.8766667	0.8800000	**0.9000000**	**0.9000000**
	0.85	0.8833333	0.8766667	0.8800000	**0.9000000**	0.8966667
25	0.99	0.8066667	0.7500000	0.7400000	**0.8266667**	0.8033333
	0.95	0.8066667	0.7733333	0.7666667	**0.8200000**	0.8066667
	0.90	0.8066667	0.7733333	0.7766667	**0.8133333**	0.8066667
	0.85	0.8100000	0.7833333	0.7800000	**0.8133333**	0.8066667
100	0.99	0.4800000	0.3900000	0.3733333	**0.5033333**	0.4700000
	0.95	0.4800000	0.4200000	0.4066667	**0.4966667**	0.4733333
	0.90	0.4800000	0.4300000	0.4166667	**0.4933333**	0.4766667
	0.85	0.4800000	0.4333333	0.4266667	**0.4933333**	0.4766667

Table 5.5. Normal Theory's CP for OLSE and LEs with $\beta'\beta = 16$ for β_1

SNR	$\boldsymbol{\rho}$	***OLSE***	$\boldsymbol{d_{boot}}$	$\boldsymbol{d_{CL}}$	$\boldsymbol{d_{CLh}}$	$\boldsymbol{d_{PR}}$
4	0.99	**0.9400000**	0.9366667	0.9233333	**0.9400000**	**0.9400000**
	0.95	0.9400000	0.9366667	0.9366667	**0.9433333**	0.9400000
	0.90	0.9400000	0.9400000	0.9366667	**0.9433333**	0.9400000
	0.85	0.9400000	0.9400000	0.9366667	**0.9433333**	0.9400000
9	0.99	0.9233333	**0.9266667**	0.9166667	0.9233333	**0.9266667**
	0.95	0.9233333	0.9233333	0.9133333	**0.9266667**	**0.9266667**
	0.90	0.9233333	0.9233333	0.9133333	**0.9266667**	**0.9266667**
	0.85	0.9233333	0.9233333	0.9133333	**0.9266667**	**0.9266667**
25	0.99	0.8933333	0.8900000	0.9000000	0.8966667	**0.9000000**
	0.95	0.8933333	0.8900000	**0.9000000**	**0.9000000**	**0.9000000**
	0.90	0.8933333	0.8900000	**0.9000000**	**0.9000000**	**0.9000000**
	0.85	0.8933333	0.8933333	**0.9000000**	**0.9000000**	**0.9000000**
100	0.99	0.8033333	0.7800000	0.7666667	**0.8233333**	0.8066667
	0.95	0.8033333	0.8000000	0.7866667	**0.8200000**	0.8066667
	0.90	0.8033333	0.8033333	0.7933333	**0.8200000**	0.8066667
	0.85	0.8033333	0.8066667	0.7966667	**0.8200000**	0.8066667

Table 5.6. Studentized-t theory's CP for OLSE and LEs with $\beta'\beta = 16$ for β_1

SNR	$\boldsymbol{\rho}$	***OLSE***	$\boldsymbol{d_{boot}}$	$\boldsymbol{d_{CL}}$	$\boldsymbol{d_{CLh}}$	$\boldsymbol{d_{PR}}$
4	0.99	0.9400000	0.9400000	0.9266667	0.9533333	**0.9566667**
	0.95	0.9400000	0.9400000	0.9400000	**0.9566667**	**0.9566667**
	0.90	0.9400000	0.9366667	0.9366667	**0.9600000**	0.9566667
	0.85	0.9400000	0.9366667	0.9366667	**0.9566667**	**0.9566667**
9	0.99	0.9166667	0.9266667	0.9200000	**0.9400000**	0.9366667
	0.95	0.9166667	0.9266667	0.9200000	**0.9400000**	0.9366667
	0.90	0.9166667	0.9266667	0.9200000	**0.9366667**	**0.9366667**
	0.85	0.9166667	0.9266667	0.9233333	**0.9366667**	**0.9366667**
25	0.99	0.9000000	0.8900000	0.9066667	**0.9100000**	0.9000000
	0.95	0.9000000	0.9000000	**0.9100000**	0.9066667	0.9000000
	0.90	0.9000000	0.9000000	**0.9100000**	0.9066667	0.9000000
	0.85	0.9000000	0.9000000	**0.9100000**	0.9066667	0.9000000
100	0.99	0.8133333	0.7966667	0.8066667	**0.8333333**	0.8100000
	0.95	0.8133333	0.8100000	0.8100000	**0.8333333**	0.8133333
	0.90	0.8133333	0.8100000	0.8100000	**0.8333333**	0.8133333
	0.85	0.8133333	0.8100000	0.8100000	**0.8333333**	0.8133333

Table 5.7. Percentile's CP for LS and LEs with $\beta'\beta = 16$ for β_1

SNR	$\boldsymbol{\rho}$	***OLSE***	$\boldsymbol{d_{boot}}$	$\boldsymbol{d_{CL}}$	$\boldsymbol{d_{CLh}}$	$\boldsymbol{d_{PR}}$
4	0.99	**0.9433333**	0.9400000	0.9333333	0.9333333	0.9300000
	0.95	**0.9433333**	0.9400000	0.9333333	0.9333333	0.9300000
	0.90	**0.9433333**	0.9400000	0.9333333	0.9333333	0.9300000
	0.85	**0.9433333**	0.9400000	0.9333333	0.9333333	0.9300000
9	0.99	0.9300000	**0.9366667**	0.9233333	0.9233333	0.9266667
	0.95	0.9300000	**0.9333333**	0.9166667	0.9266667	0.9266667
	0.90	0.9300000	**0.9333333**	0.9166667	0.9266667	0.9266667
	0.85	0.9300000	**0.9333333**	0.9166667	0.9266667	0.9266667
25	0.99	0.8933333	0.8933333	0.8933333	**0.9000000**	0.8900000
	0.95	0.8933333	0.8933333	0.9000000	**0.9033333**	0.8900000
	0.90	0.8966667	0.8933333	0.9000000	**0.9033333**	0.8900000
	0.85	0.8966667	0.8933333	0.9000000	**0.9033333**	0.8900000
100	0.99	0.8033333	0.7733333	0.7800000	0.8000000	**0.8233333**
	0.95	0.8033333	0.7800000	0.7833333	0.8000000	**0.8233333**
	0.90	0.8066667	0.7800000	0.7900000	0.8000000	**0.8233333**
	0.85	0.8066667	0.7900000	0.7933333	0.8000000	**0.8233333**

Table 5.8. BC_a's CP for OLSE and LEs with $\beta'\beta = 16$ for β_1

SNR	$\boldsymbol{\rho}$	***OLSE***	$\boldsymbol{d_{boot}}$	$\boldsymbol{d_{CL}}$	$\boldsymbol{d_{CLh}}$	$\boldsymbol{d_{PR}}$
4	0.99	0.9400000	**0.9466667**	0.9233333	0.9400000	0.9366667
	0.95	0.9400000	**0.9466667**	0.9233333	0.9333333	0.9366667
	0.90	0.9400000	**0.9466667**	0.9233333	0.9333333	0.9366667
	0.85	0.9400000	**0.9466667**	0.9233333	0.9333333	0.9366667
9	0.99	**0.9333333**	0.9300000	0.9233333	0.9300000	0.9266667
	0.95	**0.9333333**	**0.9333333**	0.9200000	0.9300000	0.9266667
	0.90	**0.9366667**	0.9300000	0.9200000	0.9300000	0.9266667
	0.85	**0.9366667**	0.9300000	0.9200000	0.9300000	0.9266667
25	0.99	0.8966667	0.9000000	0.8900000	**0.9100000**	0.9033333
	0.95	0.8966667	0.8966667	0.8900000	**0.9100000**	0.9033333
	0.90	0.8966667	0.8966667	0.8900000	**0.9100000**	0.9033333
	0.85	0.8966667	0.8966667	0.8933333	**0.9100000**	0.9033333
100	0.99	**0.8066667**	0.7866667	0.7866667	**0.8066667**	0.8033333
	0.95	**0.8066667**	0.7966667	0.7966667	0.8033333	**0.8066667**
	0.90	**0.8066667**	0.7966667	0.8033333	0.8033333	**0.8066667**
	0.85	**0.8100000**	0.8000000	0.7966667	0.8033333	0.8066667

Table 5.9. Normal Theory's CP for OLSE and LEs with $\beta'\beta = 100$ for β_1

SNR	$\boldsymbol{\rho}$	***OLSE***	$\boldsymbol{d_{boot}}$	$\boldsymbol{d_{CL}}$	$\boldsymbol{d_{CLh}}$	$\boldsymbol{d_{PR}}$
4	0.99	0.9433333	0.9466667	0.9433333	**0.9500000**	0.9466667
	0.95	0.9433333	**0.9500000**	0.9433333	**0.9500000**	0.9466667
	0.90	0.9433333	**0.9500000**	0.9433333	**0.9500000**	0.9466667
	0.85	0.9433333	0.9500000	0.9433333	**0.9533333**	0.9466667
9	0.99	0.9400000	**0.9466667**	0.9400000	0.9433333	**0.9466667**
	0.95	0.9400000	**0.9500000**	0.9433333	0.9433333	0.9466667
	0.90	0.9400000	**0.9500000**	0.9433333	0.9433333	0.9466667
	0.85	0.9400000	**0.9500000**	0.9433333	0.9433333	0.9466667
25	0.99	0.9400000	0.9400000	0.9333333	**0.9433333**	0.9400000
	0.95	0.9400000	0.9400000	0.9333333	**0.9433333**	0.9400000
	0.90	0.9400000	**0.9433333**	0.9333333	**0.9433333**	0.9400000
	0.85	0.9400000	**0.9433333**	0.9333333	**0.9433333**	0.9400000
100	0.99	**0.9100000**	0.9066667	0.9000000	0.9000000	0.9066667
	0.95	**0.9100000**	0.9066667	0.9000000	0.8966667	0.9066667
	0.90	**0.9100000**	0.9066667	0.9000000	0.8966667	0.9100000
	0.85	**0.9100000**	**0.9100000**	0.9000000	0.8966667	**0.9100000**

Table 5.10. Studentized-t theory's CP for OLSE and LEs with $\beta'\beta = 100$ for β_1

SNR	$\boldsymbol{\rho}$	***OLSE***	$\boldsymbol{d_{boot}}$	$\boldsymbol{d_{CL}}$	$\boldsymbol{d_{CLh}}$	$\boldsymbol{d_{PR}}$
4	0.99	0.9433333	0.9500000	0.9533333	0.9500000	**0.9566667**
	0.95	0.9433333	0.9433333	0.9533333	0.9500000	**0.9566667**
	0.90	0.9433333	0.9433333	0.9533333	0.9500000	**0.9566667**
	0.85	0.9433333	0.9433333	0.9533333	0.9500000	**0.9566667**
9	0.99	0.9433333	0.9433333	0.9500000	0.9500000	**0.9533333**
	0.95	0.9433333	0.9433333	**0.9533333**	0.9500000	**0.9533333**
	0.90	0.9433333	0.9433333	**0.9533333**	0.9500000	**0.9533333**
	0.85	0.9433333	0.9433333	**0.9533333**	0.9500000	**0.9533333**
25	0.99	0.9400000	0.9400000	0.9400000	0.9533333	**0.9566667**
	0.95	0.9400000	0.9366667	0.9400000	**0.9566667**	**0.9566667**
	0.90	0.9400000	0.9366667	0.9400000	**0.9566667**	**0.9566667**
	0.85	0.9400000	0.9366667	0.9400000	**0.9566667**	**0.9566667**
100	0.99	0.9033333	0.9166667	0.9200000	**0.9233333**	0.9133333
	0.95	0.9033333	0.9166667	0.9200000	**0.9233333**	0.9133333
	0.90	0.9033333	0.9166667	0.9200000	**0.9233333**	0.9133333
	0.85	0.9033333	0.9166667	0.9200000	**0.9233333**	0.9133333

Table 5.11. Percentile's CP for OLSE and LEs with $\beta'\beta = 100$ for β_1

SNR	$\boldsymbol{\rho}$	***OLSE***	$\boldsymbol{d_{boot}}$	$\boldsymbol{d_{CL}}$	$\boldsymbol{d_{CLh}}$	$\boldsymbol{d_{PR}}$
4	0.99	0.9400000	0.9400000	0.9433333	0.9433333	**0.9466667**
	0.95	0.9400000	0.9433333	0.9400000	0.9433333	**0.9466667**
	0.90	0.9400000	0.9433333	0.9400000	0.9433333	**0.9466667**
	0.85	0.9400000	0.9433333	0.9400000	0.9433333	**0.9466667**
9	0.99	**0.9433333**	0.9400000	0.9366667	0.9400000	**0.9433333**
	0.95	**0.9433333**	0.9400000	0.9300000	0.9400000	**0.9433333**
	0.90	**0.9433333**	0.9400000	0.9300000	0.9400000	**0.9433333**
	0.85	**0.9433333**	0.9400000	0.9300000	0.9400000	**0.9433333**
25	0.99	**0.9433333**	0.9400000	0.9333333	0.9333333	0.9300000
	0.95	**0.9433333**	0.9400000	0.9333333	0.9333333	0.9300000
	0.90	**0.9433333**	0.9400000	0.9333333	0.9333333	0.9300000
	0.85	**0.9433333**	0.9400000	0.9333333	0.9333333	0.9300000
100	0.99	0.9133333	**0.9166667**	0.9033333	0.9066667	0.9100000
	0.95	0.9133333	**0.9166667**	0.9033333	0.9066667	0.9133333
	0.90	0.9133333	**0.9166667**	0.9066667	0.9066667	0.9133333
	0.85	0.9133333	**0.9166667**	0.9066667	0.9066667	0.9133333

Table 5.12. BC_a's CP for OLSE and LEs with $\beta'\beta = 100$ for β_1

SNR	$\boldsymbol{\rho}$	***OLSE***	$\boldsymbol{d_{boot}}$	$\boldsymbol{d_{CL}}$	$\boldsymbol{d_{CLh}}$	$\boldsymbol{d_{PR}}$
4	0.99	**0.9466667**	0.9433333	**0.9466667**	0.9433333	0.9433333
	0.95	**0.9466667**	0.9433333	**0.9466667**	0.9400000	0.9433333
	0.90	**0.9466667**	0.9433333	**0.9466667**	0.9400000	0.9433333
	0.85	**0.9466667**	0.9433333	**0.9466667**	0.9400000	0.9433333
9	0.99	0.9400000	**0.9433333**	0.9366667	0.9366667	0.9400000
	0.95	0.9400000	**0.9433333**	0.9400000	0.9366667	0.9400000
	0.90	**0.9433333**	**0.9433333**	**0.9433333**	0.9366667	0.9400000
	0.85	**0.9433333**	**0.9433333**	0.9400000	0.9366667	0.9400000
25	0.99	0.9400000	**0.9466667**	0.9233333	0.9366667	0.9366667
	0.95	0.9400000	**0.9466667**	0.9233333	0.9366667	0.9366667
	0.90	0.9400000	**0.9466667**	0.9233333	0.9366667	0.9366667
	0.85	0.9400000	**0.9466667**	0.9233333	0.9366667	0.9366667
100	0.99	0.9133333	0.9133333	**0.9166667**	0.9100000	**0.9166667**
	0.95	0.9133333	0.9133333	0.9133333	0.9100000	**0.9166667**
	0.90	0.9133333	0.9133333	**0.9166667**	0.9100000	**0.9166667**
	0.85	0.9100000	0.9133333	**0.9166667**	0.9100000	**0.9166667**

Confidence width:

Table 5.13. Normal Theory's CW for OLSE and LEs with $\beta'\beta = 4$ for β_1

SNR	$\boldsymbol{\rho}$	***OLSE***	$\boldsymbol{d_{boot}}$	$\boldsymbol{d_{CL}}$	$\boldsymbol{d_{CLh}}$	$\boldsymbol{d_{PR}}$
4	0.99	0.3058510	0.3032388	**0.3006294**	0.3090215	0.3067610
	0.95	0.3058067	0.3038697	**0.3031221**	0.3076173	0.3067153
	0.90	0.3057730	0.3039776	**0.3035573**	0.3073099	0.3066805
	0.85	0.3057463	0.3039911	**0.3036911**	0.3071503	0.3066529
9	0.99	0.1359338	0.1352671	**0.1348720**	0.1366273	0.1363382
	0.95	0.1359141	0.1355633	**0.1354717**	0.1363553	0.1363179
	0.90	0.1358991	0.1356038	**0.1355553**	0.1362872	0.1363024
	0.85	0.1358873	0.1356088	**0.1355763**	0.1362492	0.1362902
25	0.99	0.0489362	0.0488587	**0.0488301**	0.0490603	0.0490818
	0.95	0.0489291	0.0489149	**0.0489110**	0.0490215	0.0490745
	0.90	0.0489237	**0.0489187**	0.0489188	0.0490092	0.0490689
	0.85	0.0489194	**0.0489177**	0.0489188	0.0490015	0.0490645
100	0.99	**0.0122340**	0.0122387	0.0122389	0.0122522	0.0122704
	0.95	**0.0122323**	0.0122416	0.0122428	0.0122484	0.0122686
	0.90	**0.0122309**	0.0122409	0.0122423	0.0122466	0.0122672
	0.85	**0.0122299**	0.0122400	0.0122415	0.0122453	0.0122661

Table 5.14. Studentized-t theory's CW for OLSE and LEs with $\beta'\beta = 4$ for β_1

SNR	$\boldsymbol{\rho}$	***OLSE***	$\boldsymbol{d_{boot}}$	$\boldsymbol{d_{CL}}$	$\boldsymbol{d_{CLh}}$	$\boldsymbol{d_{PR}}$
4	0.99	0.3150719	0.3141911	**0.3133154**	0.3185834	0.3162635
	0.95	0.3150248	**0.3145828**	0.3162048	0.3169527	0.3162158
	0.90	0.3149883	**0.3146259**	0.3166729	0.3166279	0.3161786
	0.85	0.3149593	**0.3146202**	0.3168104	0.3164593	0.3161491
9	0.99	**0.1400319**	0.1400708	0.1406645	0.1407847	0.1405615
	0.95	**0.1400110**	0.1403081	0.1413262	0.1404913	0.1405404
	0.90	**0.1399948**	0.1403390	0.1414109	0.1404199	0.1405238
	0.85	**0.1399819**	0.1403398	0.1414322	0.1403805	0.1405107
25	0.99	**0.0504115**	0.0505713	0.0509399	0.0505496	0.0506022
	0.95	**0.0504040**	0.0506213	0.0510241	0.0505088	0.0505945
	0.90	**0.0503981**	0.0506235	0.0510316	0.0504960	0.0505886
	0.85	**0.0503935**	0.0506219	0.0510312	0.0504876	0.0505839
100	0.99	**0.0126029**	0.0126655	0.0127676	0.0126241	0.0126505
	0.95	**0.0126010**	0.0126678	0.0127715	0.0126202	0.0126486
	0.90	**0.0125995**	0.0126670	0.0127708	0.0126182	0.0126471
	0.85	**0.0125984**	0.0126661	0.0127700	0.0126167	0.0126460

Table 5.15. Percentile's CW for OLSE and LEs with $\beta'\beta = 4$ for β_1

SNR	$\boldsymbol{\rho}$	***OLSE***	$\boldsymbol{d_{boot}}$	$\boldsymbol{d_{CL}}$	$\boldsymbol{d_{CLh}}$	$\boldsymbol{d_{PR}}$
4	0.99	0.3079512	0.3055914	**0.3036178**	0.3116582	0.3089352
	0.95	0.3079040	0.3063386	**0.3059424**	0.3103127	0.3088892
	0.90	0.3078685	0.3064735	**0.3064042**	0.3100076	0.3088541
	0.85	0.3078407	**0.3064847**	0.3065479	0.3098487	0.3088260
9	0.99	0.1368672	0.1363570	**0.1361217**	0.1378216	0.1373045
	0.95	0.1368462	**0.1366764**	0.1367437	0.1375540	0.1372841
	0.90	0.1368304	**0.1367149**	0.1368318	0.1374852	0.1372685
	0.85	0.1368181	**0.1367189**	0.1368548	0.1374456	0.1372560
25	0.99	0.0492722	**0.0492601**	0.0492872	0.0494919	0.0494296
	0.95	**0.0492646**	0.0493155	0.0493719	0.0494526	0.0494223
	0.90	**0.0492590**	0.0493190	0.0493801	0.0494397	0.0494167
	0.85	**0.0492545**	0.0493178	0.0493803	0.0494316	0.0494122
100	0.99	**0.0123181**	0.0123389	0.0123542	0.0123600	0.0123574
	0.95	**0.0123162**	0.0123417	0.0123582	0.0123560	0.0123556
	0.90	**0.0123147**	0.0123410	0.0123577	0.0123541	0.0123542
	0.85	**0.0123136**	0.0123400	0.0123569	0.0123528	0.0123530

Table 5.16. BC_a's CW for OLSE and LEs with $\beta'\beta = 4$ for β_1

SNR	$\boldsymbol{\rho}$	***OLSE***	$\boldsymbol{d_{boot}}$	$\boldsymbol{d_{CL}}$	$\boldsymbol{d_{CLh}}$	$\boldsymbol{d_{PR}}$
4	0.99	0.3081964	0.3056971	**0.3041260**	0.3118503	0.3092694
	0.95	0.3081583	**0.3058943**	0.3059050	0.3106603	0.3092414
	0.90	0.3081324	**0.3059255**	0.3063475	0.3103705	0.3092172
	0.85	0.3080617	**0.3059266**	0.3064153	0.3101881	0.3091687
9	0.99	0.1369762	0.1362366	**0.1360670**	0.1379603	0.1374530
	0.95	0.1369593	**0.1364349**	0.1367086	0.1377134	0.1374406
	0.90	0.1369477	**0.1364934**	0.1367807	0.1376155	0.1374299
	0.85	0.1369163	**0.1364973**	0.1367794	0.1375785	0.1374083
25	0.99	0.0493114	**0.0491738**	0.0492834	0.0495417	0.0494831
	0.95	0.0493053	**0.0492334**	0.0493429	0.0494977	0.0494786
	0.90	0.0493012	**0.0492459**	0.0493461	0.0494897	0.0494748
	0.85	0.0492899	**0.0492472**	0.0493479	0.0494748	0.0494670
100	0.99	0.0123279	**0.0123188**	0.0123464	0.0123716	0.0123708
	0.95	0.0123263	**0.0123246**	0.0123509	0.0123683	0.0123697
	0.90	0.0123253	**0.0123234**	0.0123500	0.0123657	0.0123687
	0.85	0.0123225	**0.0123219**	0.0123492	0.0123634	0.0123668

Table 5.17. Normal Theory's CW for OLSE and LEs with $\beta'\beta = 16$ for β_1

SNR	$\boldsymbol{\rho}$	***OLSE***	$\boldsymbol{d_{boot}}$	$\boldsymbol{d_{CL}}$	$\boldsymbol{d_{CLh}}$	$\boldsymbol{d_{PR}}$
4	0.99	1.2234040	1.2126580	**1.2022730**	1.2362510	1.2270440
	0.95	1.2232270	1.2149950	**1.2122350**	1.2305940	1.2268610
	0.90	1.2230920	1.2152100	**1.2138950**	1.2293940	1.2267220
	0.85	1.2229850	1.2150860	**1.2143400**	1.2287860	1.2266120
9	0.99	0.5437351	0.5409526	**0.5394254**	0.5465398	0.5453529
	0.95	0.5436563	0.5421343	**0.5418356**	0.5454455	0.5452716
	0.90	0.5435964	0.5422615	**0.5421546**	0.5451794	0.5452097
	0.85	0.5435490	0.5422511	**0.5422211**	0.5450333	0.5451607
25	0.99	0.1957446	0.1954113	**0.1953114**	0.1962449	0.1963270
	0.95	0.1957163	0.1956436	**0.1956373**	0.1960890	0.1962978
	0.90	0.1956947	**0.1956556**	0.1956664	0.1960407	0.1962755
	0.85	0.1956777	**0.1956460**	0.1956644	0.1960106	0.1962579
100	0.99	**0.0489362**	0.0489527	0.0489549	0.0490088	0.0490818
	0.95	**0.0489291**	0.0489651	0.0489707	0.0489938	0.0490745
	0.90	**0.0489237**	0.0489620	0.0489685	0.0489866	0.0490689
	0.85	**0.0489194**	0.0489588	0.0489652	0.0489814	0.0490645

Table 5.18. Studentized-t theory's CW for OLSE and LEs with $\beta'\beta = 16$ for β_1

SNR	$\boldsymbol{\rho}$	***OLSE***	$\boldsymbol{d_{boot}}$	$\boldsymbol{d_{CL}}$	$\boldsymbol{d_{CLh}}$	$\boldsymbol{d_{PR}}$
4	0.99	1.2602880	1.2565150	**1.2529640**	1.2745130	1.2650540
	0.95	1.2600990	**1.2578800**	1.2645470	1.2679390	1.2648630
	0.90	1.2599530	**1.2578010**	1.2663390	1.2666720	1.2647140
	0.85	1.2598370	**1.2575910**	1.2668000	1.2660280	1.2645960
9	0.99	**0.5601278**	0.5601797	0.5625903	0.5631721	0.5622462
	0.95	**0.5600440**	0.5611153	0.5652515	0.5619903	0.5621615
	0.90	**0.5599792**	0.5612020	0.5655740	0.5617107	0.5620953
	0.85	**0.5599276**	0.5611721	0.5656412	0.5615594	0.5620429
25	0.99	**0.2016460**	0.2022635	0.2037505	0.2022025	0.2024086
	0.95	**0.2016159**	0.2024696	0.2040893	0.2020384	0.2023781
	0.90	**0.2015925**	0.2024748	0.2041173	0.2019881	0.2023543
	0.85	**0.2015740**	0.2024624	0.2041134	0.2019552	0.2023354
100	0.99	**0.0504115**	0.0506600	0.0510699	0.0504966	0.0506022
	0.95	**0.0504040**	0.0506700	0.0510855	0.0504809	0.0505945
	0.90	**0.0503981**	0.0506666	0.0510828	0.0504730	0.0505886
	0.85	**0.0503935**	0.0506632	0.0510791	0.0504671	0.0505839

Table 5.19. Percentile's CW for OLSE and LEs with $\beta'\beta = 16$ for β_1

SNR	$\boldsymbol{\rho}$	***OLSE***	$\boldsymbol{d_{boot}}$	$\boldsymbol{d_{CL}}$	$\boldsymbol{d_{CLh}}$	$\boldsymbol{d_{PR}}$
4	0.99	1.2318050	1.2220330	**1.2142360**	1.2467950	1.2357410
	0.95	1.2316160	1.2248470	**1.2235040**	1.2413750	1.2355570
	0.90	1.2314740	**1.2251840**	1.2252760	1.2401840	1.2354160
	0.85	1.2313630	**1.2250560**	1.2257620	1.2395810	1.2353040
9	0.99	0.5474689	0.5453013	**0.5444227**	0.5513162	0.5492181
	0.95	0.5473849	**0.5465865**	0.5469221	0.5502400	0.5491364
	0.90	0.5473217	**0.5467053**	0.5472592	0.5499717	0.5490740
	0.85	0.5472723	**0.5466905**	0.5473342	0.5498195	0.5490239
25	0.99	0.1970888	**0.1970156**	0.1971395	0.1979712	0.1977185
	0.95	**0.1970586**	0.1972459	0.1974809	0.1978137	0.1976891
	0.90	**0.1970358**	0.1972565	0.1975118	0.1977629	0.1976666
	0.85	**0.1970180**	0.1972460	0.1975101	0.1977312	0.1976486
100	0.99	**0.0492722**	0.0493537	0.0494161	0.0494403	0.0494296
	0.95	**0.0492646**	0.0493657	0.0494324	0.0494244	0.0494223
	0.90	**0.0492590**	0.0493623	0.0494302	0.0494168	0.0494167
	0.85	**0.0492545**	0.0493589	0.0494269	0.0494114	0.0494122

Table 5.20. BC_a's CW for OLSE and LEs with $\beta'\beta = 16$ for β_1

SNR	$\boldsymbol{\rho}$	***OLSE***	$\boldsymbol{d_{boot}}$	$\boldsymbol{d_{CL}}$	$\boldsymbol{d_{CLh}}$	$\boldsymbol{d_{PR}}$
4	0.99	1.2327850	1.2224910	**1.2163420**	1.2475580	1.2370770
	0.95	1.2326330	**1.2229960**	1.2234080	1.2428060	1.2369650
	0.90	1.2325300	**1.2230200**	1.2250870	1.2416650	1.2368690
	0.85	1.2322470	**1.2228380**	1.2252820	1.2409970	1.2366750
9	0.99	0.5479047	0.5448872	**0.5441881**	0.5518630	0.5498122
	0.95	0.5478370	**0.5456293**	0.5467852	0.5508791	0.5497624
	0.90	0.5477909	**0.5457896**	0.5470728	0.5504940	0.5497195
	0.85	0.5476653	**0.5457685**	0.5470280	0.5503503	0.5496333
25	0.99	0.1972457	**0.1966750**	0.1971252	0.1981776	0.1979324
	0.95	0.1972213	**0.1969200**	0.1973675	0.1979963	0.1979145
	0.90	0.1972047	**0.1969608**	0.1973770	0.1979635	0.1978990
	0.85	0.1971595	**0.1969647**	0.1973768	0.1979040	0.1978680
100	0.99	0.0493114	**0.0492724**	0.0493847	0.0494866	0.0494831
	0.95	0.0493053	**0.0492981**	0.0494031	0.0494732	0.0494786
	0.90	0.0493012	**0.0492921**	0.0493996	0.0494631	0.0494748
	0.85	0.0492899	**0.0492864**	0.0493960	0.0494539	0.0494670

Table 5.21. Normal Theory's CW for OLSE and LEs with $\beta'\beta = 100$ for β_1

SNR	$\boldsymbol{\rho}$	***OLSE***	$\boldsymbol{d_{boot}}$	$\boldsymbol{d_{CL}}$	$\boldsymbol{d_{CLh}}$	$\boldsymbol{d_{PR}}$
4	0.99	7.6462750	7.5782280	**7.5136870**	7.7269210	7.6690240
	0.95	7.6451670	7.5920820	**7.5758600**	7.6915170	7.6678830
	0.90	7.6443250	7.5927490	**7.5860310**	7.6840920	7.6670120
	0.85	7.6436590	7.5912230	**7.5886020**	7.6803620	7.6663230
9	0.99	3.3983440	3.3805300	**3.3712750**	3.4159390	3.4084550
	0.95	3.3978520	3.3879100	**3.3863480**	3.4090940	3.4079480
	0.90	3.3974780	3.3886170	**3.3883050**	3.4074460	3.4075610
	0.85	3.3971820	**3.3884220**	3.3886790	3.4065460	3.4072540
25	0.99	1.2234040	1.2212390	**1.2206770**	1.2265390	1.2270440
	0.95	1.2232270	**1.2227160**	1.2227170	1.2255640	1.2268610
	0.90	1.2230920	**1.2227800**	1.2228940	1.2252640	1.2267220
	0.85	1.2229850	**1.2227100**	1.2228760	1.2250780	1.2266120
100	0.99	**0.3058510**	0.3059499	0.3059666	0.3063057	0.3067610
	0.95	**0.3058067**	0.3060283	0.3060659	0.3062114	0.3067153
	0.90	**0.3057730**	0.3060092	0.3060519	0.3061668	0.3066805
	0.85	**0.3057463**	0.3059857	0.3060307	0.3061346	0.3066529

Table 5.22. Studentized-t theory's CW for OLSE and LEs with $\beta'\beta = 100$ for β_1

SNR	$\boldsymbol{\rho}$	***OLSE***	$\boldsymbol{d_{boot}}$	$\boldsymbol{d_{CL}}$	$\boldsymbol{d_{CLh}}$	$\boldsymbol{d_{PR}}$
4	0.99	7.8767970	7.8524770	**7.8303860**	7.9660910	7.9065870
	0.95	7.8756190	**7.8601920**	7.9027630	7.9249260	7.9053950
	0.90	7.8747070	**7.8589500**	7.9137670	7.9170870	7.9044650
	0.85	7.8739830	**7.8568370**	7.9164370	7.9131340	7.9037280
9	0.99	3.5007990	**3.5007410**	3.5160450	3.5198970	3.5140390
	0.95	**3.5002750**	3.5065470	3.5326930	3.5125000	3.5135090
	0.90	**3.4998700**	3.5069930	3.5346690	3.5107670	3.5130960
	0.85	**3.4995480**	3.5066680	3.5350450	3.5098360	3.5127680
25	0.99	**1.2602880**	1.2640720	1.2734210	1.2637750	1.2650540
	0.95	**1.2600990**	1.2653790	1.2755410	1.2627480	1.2648630
	0.90	**1.2599530**	1.2654010	1.2757110	1.2624350	1.2647140
	0.85	**1.2598370**	1.2653110	1.2756820	1.2622320	1.2645960
100	0.99	**0.3150719**	0.3166208	0.3191853	0.3156044	0.3162635
	0.95	**0.3150248**	0.3166841	0.3192831	0.3155059	0.3162158
	0.90	**0.3149883**	0.3166627	0.3192663	0.3154566	0.3161786
	0.85	**0.3149593**	0.3166386	0.3192429	0.3154199	0.3161491

Table 5.23. Percentile's CW for OLSE and LEs with $\beta'\beta = 100$ for β_1

SNR	$\boldsymbol{\rho}$	***OLSE***	$\boldsymbol{d_{boot}}$	$\boldsymbol{d_{CL}}$	$\boldsymbol{d_{CLh}}$	$\boldsymbol{d_{PR}}$
4	0.99	7.6987810	7.6367220	**7.5884740**	7.7928200	7.7233800
	0.95	7.6976010	7.6535900	**7.6462570**	7.7588970	7.7222310
	0.90	7.6967110	**7.6550450**	7.6571450	7.7515240	7.7213530
	0.85	7.6960160	**7.6535250**	7.6599730	7.7478290	7.7206490
9	0.99	3.4216810	3.4076700	**3.4025030**	3.4457900	3.4326130
	0.95	3.4211560	**3.4157350**	3.4181360	3.4390590	3.4321030
	0.90	3.4207610	**3.4163900**	3.4202050	3.4373980	3.4317120
	0.85	3.4204520	**3.4161660**	3.4206330	3.4364610	3.4314000
25	0.99	1.2318050	**1.2312610**	1.2321020	1.2373280	1.2357410
	0.95	**1.2316160**	1.2327300	1.2342390	1.2363440	1.2355570
	0.90	**1.2314740**	1.2327860	1.2344270	1.2360280	1.2354160
	0.85	**1.2313630**	1.2327090	1.2344120	1.2358320	1.2353040
100	0.99	**0.3079512**	0.3084566	0.3088493	0.3090025	0.3089352
	0.95	**0.3079040**	0.3085320	0.3089513	0.3089028	0.3088892
	0.90	**0.3078685**	0.3085113	0.3089374	0.3088554	0.3088541
	0.85	**0.3078407**	0.3084867	0.3089166	0.3088218	0.3088260

Table 5.24. BC_a's CW for OLSE and LEs with $\beta'\beta = 100$ for β_1

SNR	$\boldsymbol{\rho}$	***OLSE***	$\boldsymbol{d_{boot}}$	$\boldsymbol{d_{CL}}$	$\boldsymbol{d_{CLh}}$	$\boldsymbol{d_{PR}}$
4	0.99	7.7049090	7.6395660	**7.6011350**	7.7980270	7.7317340
	0.95	7.7039580	**7.6427600**	7.6457410	7.7675660	7.7310340
	0.90	7.7033100	**7.6418270**	7.6561090	7.7603080	7.7304300
	0.85	7.7015430	**7.6396710**	7.6569300	7.7566800	7.7292180
9	0.99	3.4244040	3.4050780	**3.4010980**	3.4492620	3.4363260
	0.95	3.4239810	**3.4096390**	3.4173690	3.4431270	3.4360150
	0.90	3.4236930	**3.4106010**	3.4191170	3.4407100	3.4357470
	0.85	3.4229080	**3.4101980**	3.4188050	3.4397710	3.4352080
25	0.99	1.2327850	**1.2291580**	1.2320190	1.2386180	1.2370770
	0.95	1.2326330	**1.2306780**	1.2335230	1.2374740	1.2369650
	0.90	1.2325300	**1.2309110**	1.2335670	1.2372880	1.2368690
	0.85	1.2322470	**1.2309220**	1.2336230	1.2369110	1.2366750
100	0.99	0.3081964	**0.3079479**	0.3086519	0.3092916	0.3092694
	0.95	0.3081583	**0.3081077**	0.3087683	0.3092081	0.3092414
	0.90	0.3081324	**0.3080738**	0.3087460	0.3091448	0.3092172
	0.85	0.3080617	**0.3080291**	0.3087232	0.3090875	0.3091687

Through an in-depth look at the results in tables 5.1 to 5.24, the key findings can be summarized as follows:

1. For both criteria (CP and CW) there is always at least one LE from the proposed different methods of calculating the LP (Bootstrap LP calculation method, CL method, CLh method, and PR method) that shows superiority over the OLSE, where a LE that is with higher CP and narrower CI than the OLSE can always be found.
2. The results widely vary according to the different used values of the statistics, especially for the different values of *SNR* and $\beta'\beta$. On the other hand, the results do not vary that much by changing the multicollinearity level between 0.85 and 0.99 compared to the changes caused by the SNR and $\beta'\beta$.
3. The different methods of estimating CI (Normal method, Studentized-t method, Percentile method, and BC_a method) show some differences in the results, which means in one method of estimating the CI one of the methods of estimating the LP gives better results compared to the other methods, but for another method of estimating the CI it may not give better results than the other methods of estimating the LP. Noting that, in general, there are some methods that show superiority over

other methods used to estimate the LP, which is for most cases and for different methods of estimating the CI.

4. In general, the CLh and PR methods give better results compared to the other methods, as their coverage probabilities are higher and CIs are narrower in most cases compared to the OLSE method, and the Bootstrap and CL methods.
5. For all used values of $\beta'\beta$ and multicollinearity level, as the *SNR* value increases, the coverage width decreases, which is the same case for all used estimators (OLSE and LE with different methods for estimating the LP).
6. For all used values of *SNR* and multicollinearity level, the greater $\beta'\beta$ value is used the wider CI can be obtained for all used regression estimate methods, which means the value of $\beta'\beta$ affects the estimates negatively.
7. Mainly it is noticed that increasing the value of multicollinearity between the independent variables increases the coverage width, which means a wider CI. This case is not always the same for all *SNR* and $\beta'\beta$ values. In the meantime, there are some methods that give better results, like CLh and PR, which are less affected by the increase of multicollinearity compared to the other methods.
8. Increasing the *SNR* decreases the CP of all used estimates, which when taking $\beta'\beta$ and multicollinearity level constant.
9. There is a positive relation between the value of the CP and the value of $\beta'\beta$, which means when increasing the value of $\beta'\beta$ the coverage provability also increases, which is noticed for all estimates. This finding is noticed for all different values of *SNR* and multicollinearity level.
10. In most cases the increase of the multicollinearity value between the dependent variables will cause a decrease in the CP of the estimates. Knowing that the values of *SNR* and $\beta'\beta$ are constant.

5.3. Conclusion

There are many factors that affect the regression estimates including the Liu estimates, where the effects of the *SNR* and $\beta'\beta$ values are significantly high, and the multicollinearity level's effect is medium when its value varies between 0.85 and 0.99. Also, it was found that the different Liu estimates according to different methods of calculating the LP show different reactions and changes to these three factors ($\beta'\beta$, *SNR* and ρ^2). In addition to that, it has been shown that using different methods of estimating the CI may give different results, which means, for one method of estimating the CI, it has been found that one estimate has superiority over the other estimates, while the use of another method may show that this regression estimate does not have superiority over the

other estimates. Knowing that there is an advantage of some estimates (CLh and PR) over the other estimates in most cases and for most used methods of estimating the CI.

Finally, it is important to mention that at first glance, it is expected that the Bootstrap method of estimating the LP will demonstrate superiority over the other methods of estimating the LP, because it tests all available values of d to choose the optimal one, but the results were not as expected. This result was because the Bootstrap method used a different criterion to find the optimal value of d which was the MSE, that means this Bootstrap LE will give the best results according to this criterion, but it is not necessary to be the best according to other criteria, which is the main challenge in statistical estimation.

6. RESULTS AND SUMMARY

The advent of resampling methods has constituted a significant milestone in statistical estimation. Notably, techniques such as the Bootstrap and Jackknife have played a pivotal role in resolving impossible statistical challenges. Their widespread adoption in recent years is attributable to the increased computational capacity of computers. This development has effectively addressed the computational intensity challenge historically associated with resampling methods, which gave the resampling methods an advantage over traditional statistical approaches.

A key distinguishing advantage of resampling methods, precisely the nonparametric resampling methods, lies in their capacity to free statisticians from the necessity of pre-assumptions regarding the distributional characteristics of the underlying data, which is the inherent challenge encountered in traditional statistical methods. This advantage helped the use of resampling methods in very wide applications in statistics.

The parametric Bootstrap method is also proposed, but it is not widely used compared to the nonparametric Bootstrap, because it demonstrates a notable advantage over the parametric Bootstrap by eliminating the requirement for having information or assumptions related to the data distribution. Remarkably, even when using the parametric Bootstrap for estimation purposes, the nonparametric Bootstrap stands as a valuable tool for validating results derived from the parametric Bootstrap method.

Despite these advantages, it is important to mention that, like the other statistical methods, resampling methods are not immune to certain limitations and challenges. Mainly, the concerns when using Bootstrap methods are related to usability and accuracy, especially in contexts involving non-smooth statistics and unclean data. Challenges are further compounded when dealing with a limited sample size and skewed data distributions. In response to these inherent disadvantages, many literatures proposed developed resampling methods to overcome these disadvantages.

In this thesis, the Bootstrap and Jackknife methods for many applications in the Liu regression estimator were used. Many methods of estimating the CI of the regression parameters using the resampling methods were adopted, and then compared with the OLSE using the CP and the average CW as criteria of evaluation. The simulation study showed different results according to the used criterion, where it showed, according to the CP criterion, that there is always at least one Bootstrap methods (Normal, Percentile, Studentized or BC_a) that gives better results than the OLSE. While according to CW, it showed the superiority of the Normal method over the OLSE and the other Bootstrap methods. While the Jackknife method did not give good results compared to the other resampling methods. All of this leads to recommend the use of the Normal theory method to estimate the CI of the regression parameters of the LE. Knowing that this application

included a simulation study using different numbers of sample sizes and two confidence levels and using the traditional methods of estimating the LP.

It should be mentioned here that the noticed behavior of the different resampling methods in the estimation of the CI differs according to many factors, like using the resampling of the observations (pairs resampling) instead of the resampling of the errors or using a different method to estimate the LP d, like using the resampling methods to estimate it instead of the literature methods.

The second application of the Bootstrap method was to find the optimal value of the LP relying on the MSE and the cross-validation method, which was conducted with comparison with other literature methods of estimating the LP. It is concluded that the selection of the LP significantly affects the results of the LE, and while many literature methods to estimate the LP have been proposed, it becomes challenging to decide which estimate is the optimal one. The proposed Bootstrap method for choosing the value of the LP uses the advantage of Bootstrapping, which helps find the optimal value according to the required criterion. This Bootstrap estimation method of the LP has been tested for many values of the SNR, multicollinearity and $\beta'\beta$, which revealed that the different values of these factors do not have that much effect on the preference of the Bootstrap method of estimating d, as the Bootstrap method in most cases gives better results compared to the other literature methods.

The effect of using different methods to estimate the LP is also studied, which was evaluated by studying the results on the CI of the regression parameters, but as results may differ according to the used method of estimating the CI, different CI estimation methods were used. This applied to many cases that used different levels of many factors (the SNR, the multicollinearity, and $\beta'\beta$). It has been found that there are many factors that affect the regression estimates of the LE, where the effect of the *SNR* and $\beta'\beta$ values are significantly high, and the multicollinearity level's effect is medium when its value varies between 0.85 and 0.99. Also, it was found that the different Liu estimates according to different methods of calculating the LP show different reactions and changes to these three factors ($\beta'\beta$, *SNR* and ρ^2). In addition to that, it has been shown that using different methods of estimating the CI may give different results, which means, for one method of estimating the CI, it has been found that one estimate demonstrates superiority over the other estimates, while the use of another method may show that this regression estimate does not have superiority over the other estimates. Knowing that there is an advantage of some estimates (CLh and PR) over the other estimates in most cases and for most used methods of estimating the CIs.

Finally, it is important to mention that at first glance, the Bootstrap estimation methods should show a superiority over the other methods, like the case of estimating the LP, because it tests all available values of d to choose the optimal one, but the results were not as expected. This result is because the Bootstrap method used a different criterion to find the optimal value of d,

which was the MSE, which means that this Bootstrap estimate will give the best results according to this criterion, but it is not necessary to be the best according to the other criteria, which is the main challenge in statistical estimation.

As an overall conclusion, the resampling methods have many advantages over the literature methods. Starting from the ability to estimate some statistics that could not have been estimated before, and the availability of multi-experiment, which helps conduct the cross-validation that supports achieving better estimates, in addition to the ability to have nonparametric estimates, which does not rely on the assumption about the distribution of the sample. Despite these significant advantages, resampling methods face the challenge of choosing the right criterion for cross-validation and results evaluation. This leads to obtaining the optimal estimate according to a specific criterion, but it may not be the optimal estimate according to the other criteria. This leads to recommend continuing the application of the resampling methods in the LE, taking into account the significant importance of choosing the proper criteria for evaluating the results and cross-validation, which differ according to the statistic being estimated and the field of study, in addition to the high importance of choosing the proper resampling method, which needs advanced knowledge and awareness about the various resampling methods, cases in which they are recommended to be used, their disadvantages and weaknesses, and the situation of the original sample.

REFERENCES

Akdeniz, F., Kaçıranlar, S., 1995. On the almost unbiased generalized Liu estimator and Unbiased estimation of the Bias and MSE, Communications in Statistics -Theory and Methods, 24(7), 1789-1797.

Allen, DM., 1971. Mean Square Error of Prediction as a Criterion for Selection of Variables. Technometrics. 13:469-475. https://doi.org/10.2307/1267161

Brownstone, D., 1990. Bootstrapping improved estimators for linear regression models, Journal of Econometrics, Vol. 44, Issue 1-2, 171-187.

Chaubey, Y.P., Khurana, M., Chandra, S., 2018. Confidence intervals based on resampling methods using ridge estimator in linear regression model, New Trends in Mathematical Sciences, 6, No.4, 77-86.

Chernic, R. M, A., LaBudde, R., 2011. An Introduction to Bootstrap Methods with Applications to R, A John Wiley & Sons, INC,. Publication.

Crivelli, A., Firinguetti, L., Montaño, R., Muñóz, M., 1995. Confidence intervals in ridge regression by Bootstrapping the dependent variable: a simulation study, Communications in Statistics - Simulation and Computation. 24:3, 631-652. DOI: 10.1080/03610919508813264.

Davison, A.C., Hinkley, D.V., 1997. Bootstrap Methods and their Application, Cambridge University Press.

Delaney, N.J., Chatterjee, S., 1986. Use of the Bootstrap and cross-validation in ridge regression, Journal of Business and Economic Statistics, Vol.4, No.2, 255-262.

Delaney, N.J., Chatterjee, S., 1987. A note on collinearity, Bootstrapping, and cross-validation, Economic Letters, Vol.25, Issue 4, 335-338.

Dempster, A.P., Schatzoff, M., Wermuth, N., 1977. A simulation study of alternatives to ordinary least squares, Journal of the American Statistical Association, 72, 77-93.

Devidson, R., Mackinnon, J., 2000. Bootstrap test: how many Bootstraps?, Econometric Reviews, 19:1, 55-68.

DiCiccio, J., Efron, B., 1996. Bootstrap Confidence Intervals, Statistical Science, Vol. 11, No. 3, 189-228.

Duran, E.A., Akdeniz, F., 2012. Efficiency of the modified Jackknifed Liu-type estimator, Statistical Papers, 53, 265-280.

Efron, B., 1979. Bootstrap methods: Another look at the Jackknife. Annals of Statistics. 7,1-26. DOI: 10.1214/aos/1176344552

Efron, B., 1982. The Jackknife the Bootstrap and Other Resampling Plans, Library of Congress, V. 6.

Efron, B., Tibshirani, R.J., 1986. Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy. Statistical Science. 1 (1) 54 – 75. DOI: 10.1214/ss/1177013815.

Efron, B., Tibshirani, R.J., 1993. An introduction to the Bootstrap, CRC press. https://doi.org/10.1201/9780429246593

Gibbons, D.G., 1981. A simulation study of some ridge estimators, Journal of the American Statistical Association, 76, 131-139.

Golub, G.H., Heath, C.G., Wahba, W., 1979. Generalized cross-validation as a method for choosing a good ridge parameter, Technometrics, 21, 215-223.

Gong, G., 1986. Cross-validation, the Jackknife, and the Bootstrap: Excess error in forward logistic regression. J. Am. Statist. Assoc. 81, 108-113.

Hall, P., R. Wilson, S., 1991. Two Guidelines for Bootstrap Hypothesis Testing, Center for Mathematics and Their Applications, Australian National University.

Hocking, R.R., 1976. The analysis and selection of variables in linear regression, Biometrics, 32, 1-50.

Hoerl, E., Kennard, R.W., 1970. Ridge regression: Biased estimation for non-orthogonal problems, Technometrics, 12 (1), 55-67.

Hoerl, E., Kennard, R.W., 1976. Ridge regression: Iterative estimation of the biasing parameter, Communications in Statistics, A5, 77-88.

Hoerl, E., Kennard, R.W., Baldwin, K.F., 1975. Ridge regression: some simulations, Communications in Statistics, 4 (2), 105-123.

Hu, H., Xia, Y., 2013. Jackknifed Liu estimator in linear regression models, Wuhan University, Journal of Natural Sciences, Vol.18, No.4, 331-336.

Jun, S., Tu, D., 1995. The Jackknife and Bootstrap, Springer Series in Statistics.

Kadiyala, K., 1984. A class of almost unbiased and efficient estimators of regression coefficients, Economic Letters, 16, 293-296.

Kejian, L., 1993. A new class of biased estimate in linear regression, Communications in Statistics -Theory and Methods, 22 (2), 393-402.

Khurana, M., Chaubey, Y.P., Chandra, S., 2014. Jackknifing the ridge regression estimator: A revisit, Communications in Statistics -Theory and Methods, 43 (24), 5249-5262.

Kuonen, D., 2001. An Introduction to Bootstrap Methods and their Application, CStat PStat CSci.

Lawless, J.F., Wang, P., 1976. A simulation study of ridge and other regression estimators, Communications in Statistics - Theory and Methods, 5 (4), 307-323.

Mallows, CL., 1973. Some Comments on C_P. Technometrics. 15:661-675. https://doi.org/10.2307/1267380

McDonald, G.C., Galarneau, D.I., 1975. A monte carlo evaluation of some ridge-type estimators, J. Amer. Statist. Assoc, 70, 407-416.

Miller, G.R., 1974. The Jackknife – A Review, Biometrika.

Nebebe, F., Sim, A.B., 1990. The relative performances of improved ridge estimators and an empirical bayes estimator-some Monte Carlo results, Communications in Statistics -Theory and Methods, 19 (9), 3469-3495.

Newhouse, JP., Oman, SD., 1971. An evaluation of ridge estimators. Rand Report. No. R-716-PR:1-28.

Nomura, M., 1988. On the almost unbiased ridge regression estimator, Communication in Statistics- Simulation and Computation, 17(3), 729-743.

Nordberg, L., 1982. A procedure for determination of a good ridge parameter in linear regression, Communication in Statistics- Simulation and Computation, 11(3), 285-309.

Ohtani, K., 1986. On small sample properties of the almost unbiased generalized ridge estimator, Communications in Statistics -Theory and Methods, 15, 1571-1578.

Özbey, F., Kaçıranlar, S., 2015. Evaluation of the predictive performance of the Liu estimator, Communications in Statistics Theory and Methods, Vol 44, 1981-1993.

Özkale, M. R., 2008. A Jackknifed ridge estimator in the linear regression model with heteroscedastic or correlated errors, Statistics and Probability Letters, 78, 3159-3169.

Özkale, M. R., Kaçıranlar, S., 2007. A Prediction-Oriented criterion for choosing the biasing parameter in Liu estimation", Communications in Statistics-Theory and Methods, Vol 36 (10), 1889-1903.

Özkale, MR., Altuner, H., 2021. Bootstrap selection of ridge regularization parameter: a comparative study via a simulation study. Communication in Statistics, Simulation and Computation. DOI:10.1080/03610918.2021.1948574.

Özkale, R., Altuner, H., 2022. Bootstrap confidence interval of ridge regression in linear regression model: A comparative study via a simulation study. Communications in Statistics - Theory and Methods. DOI: 10.1080/03610926.2022.2045024.

Qasim, M., Amin, M., Omer, T., 2019. Performance of some new Liu parameters for the linear regression model, Communications in Statistics-Theory and Methods, Early Access, https://doi.org/10.1080/03610926.2019.1595654.

Quenouille, M., 1949. Approximate tests of correlation in time series. J. Roy. Statists. Soc. Ser. B, 11, pp. 18-48.

Sakallıoğlu, S., Kaçıranlar, S., Akdeniz, F., 2001. Mean squared error comparisons of some biased regression estimators, Communications in Statistics, Theory and Methods, Vol. 30, No. 2, 347-361.

Severiano, A., Carrico, J., Robinson, A., Ramirez, M., Pinto, F., 2011. Evaluation of Jackknife and Bootstrap for Defining Confidence Intervals for Pairwise Agreement Measures. PLoS ONE 6(5): e19539.

Shukur, G., Mansson, K., Sjolander, P., 2015. Developing interaction shrinkage parameters for the Liu estimator with an application to the electricity retail market, Computational Economics, Vol 46, Issue 4, 539-550.

Simon, J., 1969. Basic Research Methods in Social Science. Random House, New York.

Singh, B., Chaubey, Y.P., Dwivedi, T.D., 1986. An almost unbiased ridge estimator, Sankhya: The Indian Journal of Statistics, B48, 342-346.

Tekeli, E., Kaçıranlar, S., Özbay, N., 2019. Optimal determination of the parameters of some biased estimators using genetic algorithm, Journal of Statistical Computation and Simulation. 89:18, 3331-3353

Tukey, J., 1958. Bias and confidence in not quite large samples. Abstract, Ann. Math. Statist. 29, p. 614.

Wichern, D., Churchill, G., 1978. A Comparison of Ridge Estimators, Technometrics. 20 (3), 301-311.

Wu, J., Yang, H., 2013. Efficiency of an almost unbiased two-parameter estimator in linear regression model, Statistics, Vol.47, No.3, 535-545.

Yildiz, N., 2018. On the performance of the Jackknifed Liu-type estimator in linear regression model, Communications in Statistics-Theory and Methods, 47:9, 2278-2290.

CURRICULUM VITAE

Personal Information

Surname, Name : DENİZ, Mustafa

Education

Degree	**Institution**	**Graduation Year**
Master's degree in mathematical statistics	University of Aleppo, Faculty of Sciences, Department of Statistics.	2012
Bachelor in mathematical statistics	University of Aleppo, Faculty of Sciences, Department of Statistics.	2005
Secondary education	Al-Arfi High School	2001

Work Experience

Year	**Workplace**	**Job title**
2017-Present	INDICATORS Company	CEO
2015-2017	i-APS Company	Research and M&E Consultant
20014-2015	Sabr Center for Research	General Manager
2009-2013	University of Aleppo, Faculty of Sciences, Department of Statistics.	Lecturer in mathematical statistics and coding subjects
2005-2009	Freelancer	Data Analyst

Publications

1. Statistical Analysis of Questionnaires by SPSS, 2014.
2. Non-Parametric Tests and Their Application by SPSS, 2015.
3. Field Heroes: Data collection skills guide series (4 guides), 2018.
4. Online Survey Data Collection by Google Forms, 2015.
5. Accountability to Affected Populations (AAP) Standards.