Comparative Analysis of Non-Parametric Approaches Dealing with Endogeneity in Finite Samples

PhD Econometrics

Researcher: Gulfam Haider

Registration No. 25-SE/PhD (Et)/ S12

Supervisor: Dr. Abdul Jabbar

Co-supervisor: Dr. Asad Zaman

School of Economics

INTERNATIONAL INSTITUTE OF ISLAMIC ECONOMICS (IIIE)

INTERNATIONAL ISLAMIC UNIVERSITY ISLAMABAD (IIUI), PAKISTAN.

Comparative Analysis of Non-Parametric Approaches Dealing with Endogeneity in Finite Samples

Gulfam Haider

Registration No. 25-SE/PhD (ET)/S12

Submitted in partial fulfillment of the requirements for the

Doctor of Philosophy in Econometrics

at International Institute of Islamic Economics (IIIE)

International Islamic University Islamabad (IIUI)

Supervisor: Dr. Abdul Jabbar

Co-supervisor: Dr. Asad Zaman

October 2020

DECLARATION

I hereby declare that this thesis, neither as a whole nor as a part thereof, has been copied out from any source. It is further declared that I have carried out this research by myself and have completed this thesis on the basis of my personal efforts under the guidance and help of my supervisor. If any part of this thesis is proven to be copied out or earlier submitted, I shall stand by the consequences. No portion of work presented in this thesis has been submitted in support of any application for any other degree or qualification in International Islamic University or any other university or institute of learning.

Gulfam Haider

October 2020.

DEDICATION

To

My Beloved Father

&

My Family

APPROVAL SHEET

Comparative Analysis of Non-Parametric Approaches Dealing with Endogeneity in Finite Samples

by

Gulfam Haider

Reg. No: 25-SE/PhD(Et)/S12

Accepted by the International Institute of Islamic Economics (IIIE), International Islamic University Islamabad (IIUI), as partial fulfillment of the requirements for the award of degree of DOCTOR OF PHILOSOPHY IN ECONOMETRICS

Supervisor

Dr. Abdul Jabbar
Assistant Professor, IIIE, IIUI.

Co-supervisor

Dr. Asad Zaman
Ex-Vice Chancellor, PIDE

Internal Examiner

Dr. Arshad Ali Bhatti
Assistant Professor, IIIE, IIUI.

External Examiners

[redacted]

Dr. Tahir Mukhtar
Associate Professor/ Chairman, Department of Economics, Fatima Jannah Women University Rawalpindi.

7.12.20

Dr. Zahid Asghar
Professor/ Director, School of Economics, School of Economics, QAU, Islamabad.

Dr. Hamid Hassan
In-charge,
School of Economics

Dr. Abdul Rashid
Associate Professor,
Director General

International Institute of Islamic Economics (IIIE)
International Islamic University Islamabad (IIUI)

Date of Viva-Voce Examination: **October 14, 2020**

Acknowledgement

In the name of Allah Almighty, the Most Merciful, the Most Beneficent. I am always thankful to Allah Almighty for the knowledge, skills, capabilities, and exposure bestowed on me.

I pay gratitude to my Supervisor Dr. Abdul Jabbar, and Co-Supervisor Dr. Asad Zaman for their utmost guidance to produce this remarkable work. I would like to pay gratitude to all my teachers at International Islamic University for their support and motivation during my research. At this time, I would acknowledge the teaching and research contributions of the honorable teachers Dr. Asad Zaman, Dr. Mumtaz Ahmad, Dr. Atiq-ur-Rehman, Dr. Arshad Ali Bhatti, and all faculty members of the International Institute of Islamic Economics (IIIE) who encouraged and guided me whenever needed. I am also thankful to my loving and caring fellows including Asad-ul-Islam, Waqar Khan, Yasir Riaz, Shahid Akbar, Abdul Waheed, Mehmood Hassan, Muhammad Irfan Malik, Abdul Ghafar, Khan Buhadar, and all other classmates who helped me one way or other during the coursework and dissertation. I enjoyed their company during my stay at International Islamic University Islamabad. I also acknowledge the cooperation of the administrative staff of the International Institute of Islamic Economics (IIIE) especially Syed Niaz Ali Shah for his support during my stay at IIIE.

Without the support and encouragement of my colleagues at FAST – National University of Computer and Emerging Sciences, Chiniot -Faisalabad Campus, I was not able to achieve this milestone. I would like to thank Dr. Aftab Ahmed Maroof, Hammad Majeed, Dr. Saif Maqbool, Dr. Umar Farooq, and Syed Zeeshan Haider.

There are some special people to whom I owe a great deal of gratitude. I openly admit that the prayers of my father and mother are always present at every success in my life. I cardinally thank them for their prayers, love, and unconditional support. I am greatly indebted to my brother Ali Abbas Haider and my sisters for their moral and financial support throughout the period. I am thankful to other members of my family for standing by me throughout my long academic career. I would like to acknowledge Mr. & Mrs. Inayat Ullah for their moral and financial support during my studies. Last but not the least, I would like to pay special gratitude to my second half, my lovely wife for her patience, sacrifices, and prayers to complete this journey. I would like to thank my lovely and cute kids Muhammad Abdullah Haider and Harmain Fatima for their smiles and broken words that induced energy to move forward.

Gulfam Haider

بسم الله الرحمن الرحيم

ABSTRACT

The GMM approach is highly criticised due to its poor finite sample properties. Various authors proposed non-parametric approaches as an alternative to GMM. This dissertation primarily compares the finite sample properties of GMM and three Non-Parametric approaches, i.e., Maximum Empirical Likelihood, Maximum Exponential Empirical Likelihood, and Cressie and Read Optimal Convex Combination (CROCC) using Monte Carlo simulation analysis. The findings of this study revealed that CROCC is a more efficient estimator than all other estimators, especially from the GMM approach that has poor finite sample properties. CROOC produces less biased estimator than GMM in the exactly determined model, whereas MEL and MEEL produce lower bias than GMM and CROCC in the overdetermined model. Therefore, the study concluded that CROCC has better finite sample properties as compared to GMM.

The performance of CROCC can be improved, if the coefficient of optimal convex combination is estimated by operationalizing the minimum quadratic risk estimator appropriatly. Consequently, we introduced an alternative arbitrary method (Cressie and Read Arbitrary Convex Combination - CRACC) that is independent of the computation of the coefficient based on the minimization of quadratic risk. Our results showed that in general, CROCC is more efficient than CRACC in the exactly determined model, whereas CRACC is less biased and more efficient than CROCC in the overdetermined model. Hence, it is recommended that researchers should use CRACC for estimating the economic models having endogeneity issues.

Keywords: Endogeneity; Finite Samples; Information-Theoretic Approach; GMM; Biasedness; Non-Parametric Approaches.

Table of Contents

LIST OF FIGURES

LIST OF TABLE

LIST OF ABBREVIATIONS

DGP	Data Generating Process
MOM	Method Of Moments
GMM	Generalized Method of Moments
ELR	Empirical Likelihood Ratio
MEL	Maximum Likelihood
MEL	Maximum Empirical Likelihood
MEEL	Maximum Exponential Empirical Likelihood
CROCC	Cressie and Read Optimal Convex Combination
CRACC	Cressie and Read Arbitrary Convex Combination

CHAPTER 1. INTRODUCTION

The significant contribution of econometrists is to explore tools and techniques that could aid in the estimation of unknown parameters and their distribution of a given data set. Various techniques are effectively used to estimate different types of models based on optimization. The estimation and inference of general linear regression models face problems of unknown parameters that are functionally linked to samples moment conditions. The estimation techniques for functionally independent moment equations are linked to a number of parameters that can be classified in three ways. Firstly, if the number of moment restrictions is smaller than the number of parameters to be estimated, then the model is under-identified, which demonstrates that there is inadequate information in the model to uniquely estimate parameters.

Secondly, if the number of moment equations is exactly equal to the number of parameters, then the model will be exactly identified. In this case, the classical Pearson Method of Moment (MOM) is used to obtain a unique solution of parameters that is called the Ordinary Method of Moments. Thirdly, if the number of moment conditions exceeds the number of parameters, then the model becomes overdetermined. In this case, the additional information is available about unknown parameters. Generally, it is unlikely to set the sample average of moment conditions rigorously equal to zero. This case is associated with the Generalized Method of Moments (GMM), and its solution was proposed by Hansen (1982). A similar method to the GMM approach was discussed by Ferguson (1958), which is based on Pearson chi-squared statistics and relates to an efficient GMM estimator.

The Maximum Likelihood (MLE) technique is most extensively used to estimate the econometric model. This technique can be used in case we have complete knowledge about the model and about its probability distribution. However, the Generalized Method of Moments (GMM) approach does not need the condition of full knowledge in model estimation. It requires only a set of moment functions specification according to the model. The study of Imbens (2002) stated that the GMM approach is assumed as a combination of characteristics of various approaches, including the Instrumental Variables (IV) approach, Two Stage Least Squares (2SLS), Maximum Likelihood (MLE) and Ordinary Least Square (OLS) techniques.

Due to the nested characteristics of all the above-discussed approaches, the GMM is highly appreciated and used by authors and practitioners for the last two decades. According to Judge and Mittelhammer (2011), the GMM approach used to estimate the unknown parameters to satisfy the moment conditions, possibly close to zero. The close to zero condition can be evaluated with weighted Euclidean distance. The GMM estimator provides distinct efficient estimators concerning their weighting structure.

1.1 Empirical Likelihood (EL) Approach

The GMM based inference was proposed by Hansen (1982) for the overdetermined model, and it has good asymptotic properties. However, several authors reported that finite sample properties of GMM estimators are different from the asymptotic properties (Altonji & Segal, 1996; Hansen, Heaton, & Yaron, 1996; Pagan & Robertson, 1997). Consequently, the study of Owen (1988) proposed the Empirical Likelihood (EL) method as an alternative approach that provides improved inferences in econometric models. The EL[1] approach can be used to estimate the likelihood with the help of sample data; only in case, there is no information available about parametric functional form to construct likelihood function. Therefore, EL approach is known as non-parametric empirical likelihood approach. The EL approach with constrained optimization is used subject to the moment restrictions for estimation of unknown parameters.

Similarly, the study of Kolaczyk (1994) extended the EL approach to generalised linear models. Several other studies (Imbens & Spady, 2002; Newey & Smith, 2004; Qin & Lawless, 1994) also presented EL as an alternative approach to GMM. The efficient GMM estimator usually requires a two-step procedure for the estimation of the best weighting matrix. In contrast, the empirical likelihood approach requires only a single step that improves econometric modeling in finite samples (Judge & Mittelhammer, 2011). They discussed the EL approach as Nonparametric Maximum Likelihood (NPML) for estimating unknown parameters as the parametric functional form for the likelihood function is not available. Consistently, a number of authors proposed an alternative approach to the GMM estimator that is known as exponential tilting / Kullback-Leibler Information Criterion (KLIC)[2] to the GMM approach (Imbens, Johnson, & Spady, 1998; Kitamura & Stutzer, 1997). precisely

---

[1] Various authors reported that the GMM approach has poor finite samples properties and suggested EL approach as an alternative to GMM (Altonji & Segal, 1996; Hall & Horowitz, 1996; Imbens et al., 1998; Imbens, 1997; Kitamura & Stutzer, 1997; Qin & Lawless, 1994). The EL approach is advantageous to GMM due to the weight matrix handled internally during the estimation that improves the finite sample properties (Judge & Mittelhammer, 2011).

2 Judge and Mittelhammer (2012) discussed that minimizing KL information is apparently equal to maximizing $-\sum p_i \ln(p_i)$, that is exactly the Shannon-Jaynes information (entropy) measure. Campbell (1999) discussed that Shannon entropy function is a measuring uncertainty in a discrete probability distribution.

1.2 Cressie and Read Information Theoretic Approach

The power-divergence family of statistics was proposed by Read and Cressie (1984; 1988) for dealing with discrete data. The Cressie and Read (CR) proposed the following divergence measures:

$$I(p, q, \gamma) = \frac{1}{\gamma(\gamma+1)} \sum_{i=1}^{n} p_i \left[\left(\frac{p_i}{q_i}\right)^{\gamma} - 1\right]$$

In CR[3] divergence measure, $'\gamma'$ is a parameter of this family of statistics, which shows the different entropy measures. "p" is the probability of subjective distribution, and "q" is the probability of the reference distribution. CR comprises three main variants in his family of entropy measure as the Maximum Empirical Likelihood (MEL) approach, Maximum Exponential Empirical Likelihood (MEEL) approach, and Maximum Log Euclidean Likelihood (MLEL) based on the variation of 'gamma' coefficient. The Judge and Mittelhammer (2011) derived all three variants of CR family of statistics $I(p, q, \gamma)$ using Choices of gamma (-1, 0, 1).

They illustrated that $CR(\gamma = -1)$ represents MEL where the objective function is $\{n^{-1} \sum_{i=1}^{n} \ln(p_i)\}$. The $CR(\gamma = 0)$ defines MEEL with following objective function $\{-\sum p_i \ln(p_i)\}$. Whereas, the $CR(\gamma = 1)$ represents MLEL. The authors' Judge and Mittelhammer (2011) discussed that inference methods of CR family are better than traditional ML and GMM approaches. The authors also discussed that CR is an updated

---

[3] The authors demonstrated that over defined ranges of the divergence measures, the Cressie and Read family of information theoretic approach and entropy families are equivalent (Gorban and Judge 2010).

variant of the GMM estimator, where the unknown covariance matrix is handled internally to estimate the unknown parameter.

The Data Generating Process (DGP) is generally unknown that poses a greater impact on estimation. In order to deal with this deficiency, Judge and Mittelhammer (2012) proposed the idea of the optimum choice rule in which the information of various members of the CR family can be combined by minimizing a square error loss function. The detailed explanation of the optimum convex combination rule is given in the next section.

1.2.1 The Optimal Convex Combination of CR Family of Divergence

According to the study of Judge and Mittelhammer (2012), the optimal choice rule is based on minimizing the square error loss function. This approach requires the optimal use of the information obtained from two or more individual estimators of β from the CR family of information-theoretic approach $\{\hat{\boldsymbol{\beta}}(\gamma = -1), \hat{\boldsymbol{\beta}}(\gamma = 0), \text{etc.}\}$. The optimal use of two estimators of "$\boldsymbol{\beta}$" as follow:

$$\bar{\boldsymbol{\beta}}(\hat{\boldsymbol{\alpha}}) = \hat{\boldsymbol{\alpha}}\,\hat{\boldsymbol{\beta}}(\boldsymbol{\gamma} = -1) + (1 - \hat{\boldsymbol{\alpha}})\,\hat{\boldsymbol{\beta}}(\boldsymbol{\gamma} = 0)$$

This estimator is produced by minimizing the Quadratic Risk (QR) function that is discussed in Chapter III. Moreover, the optimal convex combination $\bar{\boldsymbol{\beta}}(\hat{\boldsymbol{\alpha}})$ is superior to the individual estimators $\hat{\boldsymbol{\beta}}(\gamma = -1)$ and $\hat{\boldsymbol{\beta}}(\gamma = 0)$ (Judge & Mittelhammer, 2012). This emphasizes that the performance of the optimal convex combination is always better than the performance of individual estimators obtain from the CR family of the information-theoretic approach. The authors discussed that the optimal choice of the estimators could be obtained from more than two individual information-theoretic estimators to enhance the efficiency of the estimator.

1.3 Endogeneity

In the estimation of the regression model, one of the main assumptions of the Gauss-Markov theorem is the orthogonality condition that must hold between independent variables and error term (i.e., exogeneity). The violation of the orthogonality assumption leads to OLS estimates biased and inconsistent. This situation is called an endogeneity problem that produces bias estimates. The failure of this assumption occurs because of simultaneity between dependent and independent variables, omitted variable bias, or measurement error in independent variables.

In this setting, we need to establish more information in the form of instrumental variables that are correlated with independent variables and orthogonal with the error term (Judge & Mittelhammer, 2011). Various techniques have been introduced that use instrumental variables to solve such an endogeneity issue in this respect. Among these techniques, the Generalized Method of Moments (GMM) is the most popular estimator. However, various studies have documented that GMM estimator has poor finite samples properties (Altonji & Segal, 1996; Hall & Horowitz, 1996; Imbens et al., 1998; Imbens, 1997; Kitamura & Stutzer, 1997; Qin & Lawless, 1994). In contrast, the information-theoretic approach (MEL, MEEL) is also used to resolve the endogeneity issue (Judge & Mittelhammer, 2011; Judge & Mittelhammer, 2012). Therefore, it is important to identify alternative techniques having good finite sample properties to solve the aforementioned problem of endogeneity.

1.4 Motivation

A number of studies reported that the finite sample properties of GMM are different from the asymptotic properties of GMM and produces biased results[4] in the presence of endogeneity problem. In contrast, the authors proposed Empirical Likelihood as an alternative approach that has lower bias and improved inference in the finite sample (Owen, 1988; Qin & Lawless, 1994). Similarly, other authors suggested Exponential Tilting/Kullback-Leibler Information Criterion (KLIC) as an alternative approach to GMM due to its poor finite sample properties (Imbens et al., 1998; Kitamura & Stutzer, 1997). In order to reduce the finite sample bias of GMM estimator Ramalho (2006) analysed different bootstrapping techniques. The efficient GMM estimator usually requires the best weighting matrix; however, the empirical likelihood approach estimates the weight matrix internally during the estimation that improves the finite sample properties of the EL estimator (Judge & Mittelhammer, 2011).

The study of Newey and Smith (2004) suggested the EL approach to reduce the finite sample bias of GMM. According to Golan et al. (1996), MEEL/KLIC and MEL are the special cases of maximum entropy[5], whereas CR family encompasses all the entropy measures, especially MEL and KLIC. The optimal choice of the parameters from the CR family of test statistics was proposed (Judge & Mittelhammer, 2011; 2012). They discussed that CR Optimal Convex Combination (CROCC) of two or more individual

---

[4]These authors (Altonji & Segal, 1996; Hall & Horowitz, 1996; Imbens et al., 1998; Imbens, 1997; Kitamura & Stutzer, 1997; Qin & Lawless, 1994), explored the poor finite samples properties of GMM approach. (Pagan & Robertson, 1997) registered a number of problems of GMM approach related with finite sample properties.

[5] Golan et al. (1996) applied maximum entopry estimation to a wide range o f econometric problems. They also extended maximum entropy methodology to estimate linear regression models.

estimators (e.g., MEL, MEEL) have lower mean square errors than the individual estimators.

The GMM approach is widely used in econometric literature, although plenty of authors stated that the GMM approach has poor finite sample properties. In addition, a number of studies suggested alternative approaches to GMM (Imbens & Spady, 2002; Kitamura, 2001; Newey & Smith, 2004). However, the fragmented literature does not offer consensus about a particular approach for various combinations of the degree of endogeneity, strength of instrumental variables, and sample size. We are going to fill the literature gap by evaluating the properties of CROCC with diverse combinations of the degree of endogeneity (weak, moderate, high, and very high degree) and strength of instrumental variables (weak, moderate, and strong strength) with variation in sample sizes. We are also interested in comparing the results of CROCC with MEL, MEEL, and GMM estimators to examine the best and the worst scenarios of all the estimators.

The study of Judge and Mittelhammer (2011) argued that the performance of CROCC could be enhanced by using appropriate operationalization of the minimum quadratic risk estimator. Therefore, we will introduce an arbitrary alternative method (Cressie and Read Arbitrary Convex Combination - CRACC) to compute the coefficient of convex combination that is independent of the computation of the coefficient based on minimization of the quadratic risk function. Additionally, we will compare the performance of CROCC with CRACC.

Furthermore, we would apply these information-theoretic estimators (MEL, MEEL, CROCC, and CRACC) on real economic data in order to get the least bias estimates in finite samples that are associated with the implementation of GMM estimators.

1.5 The objective of the Study

Following are the four main objectives of this dissertaion

- To evaluate and compare the finite samples properties of MEL, MEEL, and GMM estimators in various combinations of degree of endogeneity and strength of instrumental variables.
- To evaluate and compare the finite samples properties of Cressie and Read Optimal Convex Combination (CROCC) with GMM estimator in various combinations of degree of endogeneity and strength of instrumental variables.
- To evaluate and compare the finite samples properties of newly proposed estimator denoted as Cressie and Read Arbitrary Convex Combination (CRACC) with CROCC in various combinations of degree of endogeneity and strength of instrumental variables.
- To implement the CRACC, CROCC, MEL, MEEL and GMM estimator in two real world problems of Keynesian consumption function and money demand function.

1.6 Significance of the Study

This research study demonstrates key theoretical and practical implications in the field of inference in econometric models. As discussed earlier; the endogeneity problem and finite samples are key issues in econometric models; we are addressing these issues by comparing the information-theoretic estimators with the GMM approach in various scenarios. The significance of this study is quite evident and prevalent in the following ways.

Firstly, the study prepares substantial grounds for future researchers by offering a comparison of the properties of MEL, MEEL, CROCC and GMM estimator on various combinations of the degree of endogeneity and strength of the instrumental variables for different sample sizes. It would help researchers to choose the right estimator according to the degree of endogeneity and the sample size.

Secondly, this study investigates the properties of CROCC and also compare its performance with GMM. This comparison would help the researchers to use optimal convex combination estimator because of its supremacy on individual information theoretic estimators and GMM.

Thirdly, this study introduces the CR Arbitrary Convex Combination (CRACC) and compares its performance with CROCC. The newly proposed estimator: CRACC which has superiorty to CROCC due to its assumption free estimation of two independent samples. The introduction of this estimators and its comparison would help researchers to implement the best and the most appropriate technique in different scenarios to tackle endogeneity issues. The introduction of CRACC would open up another research stream in current literature.

1.7 Organization of Thesis

This dissertation is composed of six sections. The second chapter synthesizes the existing literature of the GMM and the information-theoretic approach. The methodological discussion is presented in the third chapter. The results and discussion of Monte Carlo simulation analysis on GMM and the information-theoretic approach have been reported in the fourth chapter. The use of MEL, MEEL, and CR information-theoretic estimators for the solution of the endogeneity problem in economics models (consumption functions and money demand functions) have been discussed in the fifth chapter. Finally, chapter six concludes the entire discussion of this dissertation.

CHAPTER 2. REVIEW OF LITERATURE

This chapter on literature includes the finite sample properties of GMM with orthogonal conditions, it prepares strong grounds based on the previous and current debate on topic to pursue further research.. This also consists of numerous studies conducted on small sample properties of the nonparametric empirical likelihood approach and other entropy criteria, that are the members of the Cressie and Read (CR) family of divergence measures by several novelists (Imbens & Spady, 2002; Kitamura, 2001; Newey & Smith, 2004). This chapter is divided into the following subsections.

Section 2.1 is focused on reviewing the GMM approach; its implication in economics and its finite sample properties are discussed respectively. In section 2.2, the non-parametric likelihood approach explored the arguments of renowned authors. Section 2.3 covers literature on Cressie and Read information divergence measures while the literature on square error loss is discussed in subsection 2.3. The final summary of this chapter is discussed in section 2.4.

2.1 GMM Approach

The MOM is commonly used to estimate the parameters of the sample average of the moments equal to zero. The number of parameters is equal to the number of moment equations referred to as an exactly identified system. Conversely, when the number of moment equations is greater than the number of parameters to be estimated, then the system becomes overdetermined, which is not feasible. However, its solution is proposed by Hansen (1982) that is known as the GMM approach. The GMM approach works with large sample properties that make sample analogues of population orthogonality condition close to zero. He further suggested that an estimator is strongly consistent and holds the asymptotic normality conditions. He examined this procedure by using the GMM estimation technique to test over-identified restriction.

2.1.1 GMM Approach, Endogeneity, and Economic Models

Several authors have reported the use of the GMM approach in the estimation of different economic models. The study of Van Beveren (2012), used different methodological problems which occurred while estimating the total factor productivity. The author of the study discussed the GMM approach as a method of estimation. Moreover, he also used GMM for the estimation of total factor productivity. The author reported, based on the study of Wooldridge (2009), that the GMM estimator is more efficient than the two step semiparametric procedures. He further argued that the GMM approach could correct the simultaneity problem theoretically. However, the results demonstrated that semiparametric estimators should be preferred as compared to the GMM approach due to its poor performance in finite samples.

Another author, Bond, Hoeffler, and Temple (2001) worked on the problem of first differenced GMM approach in the panel data growth model and reported to use an alternate to GMM estimator for estimation of panel data. They argued that a persistent time series first differenced GMM behaved poorly. Furthermore, the first difference GMM is considered as problematic for growth models. According to the results of the study, they suggested that the estimators of first differenced GMM are extremely biased.

The demand and supply-side determinants of the textile industry in Pakistan was analysed using the GMM approach to estimate simultaneous equations in the study of Latif and Javid (2016). They illustrated that the prime advantage of using the GMM approach is to obtain consistent estimates in the presence of endogeneity in the model. Whereas, the study of Kendix and Walls (2010), quantified the impact of oil industry consolidation on refined product prices using US petroleum refining industry. They have argued that the parameters can be estimated by using different techniques like least squares and instrument variables approach; however, the most appropriate technique to control the endogeneity issue is the GMM approach. The GMM was preferred to obtain an efficient estimator. They estimated the same results by Two Stage Least Squares (TSLS) that was consistent but inefficient.

The study of Ngo (2006) investigated the relationship between bank capital and profitability. The author of the study argued that the GMM estimator is more efficient than the simple instrumental variable and preferred GMM to the least squares principle. By using the GMM estimated the model and found no relation between capital and profitability. Similarly, Wooldridge (2001) worked to apply the generalized method of moments and explored the difference between the GMM estimator and the weighting matrix. The author argued that the GMM estimator minimizes the quadratic form through the weight matrix into sample moment conditions that assure the consistency of the GMM estimator. However, the study reported that a finite sample problem could occur with the GMM estimator. In conclusion, the author reported that GMM is better and more efficient than the ordinary least squares and two-stage least squares. Although theoretically, it seems that GMM will always be on preference; still, researchers prefer to use OLS and 2SLS. According to the results of this study, the author found that 2SLS and GMM yield the same results.

The study of Kapetanios and Marcellino (2010) examined the instrumental variables estimation by using factor analysis when the number of instruments increases. The authors examined that standard GMM is preferable when the regressors are the set of a large finite set of observed instruments. The theoretical findings suggested the superiority of factor GMM to the standard GMM. The author further mentioned that the accuracy of GMM estimators decreased gradually when the factors used as instruments. Hansen and West (2002), considered the role of a generalized method of moments in case of macroeconomic time series analysis. The authors stated that the GMM does not maintain strong assumptions as well as suggested to understand the behaviour of GMM estimator under particular situations or misspecifications. They explored GMM estimator in panel data studies as well as in nonlinear studies. The findings of the study concluded that the GMM is an important part in macro-econometrics and reported that GMM would be an essential estimator in near future. Additionally, they discussed nonparametric Empirical Likelihood (EL) approach that requires only sample information and moment conditions. This reported that the EL approach is free from the estimation of the optimal weight matrix.

The study of Ma (2002) re-examined the new Philips curve with two specifications that were neglected by Gali and Gertler; by applying the test statistics developed by Stock and Wright using the GMM approach. Similarly, Baum, Schaffer, and Stillman (2003) contrasted OLS, instrumental variables, and GMM approach. For this purpose, the authors applied the Durbin-Wu-Hausman test. They mentioned that the GMM approach is the most commonly used approach, especially in the case of Heteroskedasticity, as well as in empirical research for the solution of simultaneity. The authors preferred the GMM approach to the instrumental variables approach as it helps to obtain efficient estimator; however, it demonstrates poor performance in finite samples. Additionally, the study of Windmeijer & Santos Silva (1996) explored the estimation of count data models with endogenous regressors by using the GMM estimator. In their study, the authors analysed that the GMM estimator failed to give them a definite indication of endogeneity.

2.1.2 GMM Approach and Finite Sample Bias

The econometric literature demonstrates that the GMM approach has advantages in asymptotic properties in various situations. Plenty of studies also explored the numerous problems of GMM estimator and reported that the properties of the finite sample are quite different from asymptotic properties. Such as Altonji and Segal (1996) investigated the finite samples properties of the GMM estimator. The authors examined the covariance structure of the model that is associated with the weight matrix of the GMM estimator, which is labeled as an optimal minimum distance (OMD) estimator. The findings of the study displayed that OMD is highly biased in finite samples. The authors of the above-stated study suggested an alternative estimator called independently weighted optimal minimum distance (IWOMD) and compared the performance of both OMD and IWOMD. Finally, the authors concluded that the performance of the GMM estimator is poor in finite samples.

In accordance with the above study, Hall and Horowitz (1996) examined the GMM inference base study. Based on the Monte Carlo investigation, the authors exhibited the poor performance of the GMM estimator. They also developed the bootstrap critical values of the test statistics obtained from the GMM approach for overdetermined models. In relation to the above study, Hansen et al. (1996) examined the small sample properties of the GMM estimator. They focused on three types of GMM estimators in which moment conditions are weighted in different ways to differentiate these estimators. The types discussed are as follows: the first one is two step GMM estimator that used the identity matrix as a weighted matrix. Second, the Iterative method is a continuous form of two step estimator, where the weight matrix is continuously updating until the convergence of the parameters. The third one is Continuous Updating Estimator (CUE), in which the covariance matrix is repeatedly adjusted with parameter changing in minimization of the quadratic function. In addition to the above, they emphasized that overdetermined models construct the confidence regions. These authors discussed different conditions of the Consumption-Based Capital Asset Pricing Model (CCAPM) to test the small sample properties. They concluded that the CUE estimator has less bias than the other estimator.

The study of Imbens (1997) compared the GMM estimator proposed by (Hansen, 1982) and iterative GMM estimator proposed by (Hansen et al., 1996) with an Empirical likelihood estimator. Imbens (1997) discussed the advantages of the EL estimator and reported that all other conventional estimators primarily need the weight matrix estimation; however, the EL estimator has no such condition. This estimator has attracted some information-theoretic interpretation. The finding of his study reported that the EL estimator has low bias and lower root mean square error than the two step GMM and iterative GMM estimators.

2.2 Empirical Likelihood and Information-Theoretic Approach

The empirical likelihood is primarily used by Thomas and Grunkemeier (1975), with the help of a nonparametric likelihood ratio in the estimation of confidence intervals for survival function. The study of Owen (1988) initially proposed the empirical likelihood approach as an alternative approach for inference in econometric literature. He argued that the empirical likelihood ratio (ELR) approach could be used to estimate the confidence interval for a single parameter (sample mean). It is the nonparametric approach for inference that is further extension of Wilk's theorem (1938). In this context, the author (Owen, 1990) explained that the empirical likelihood could be used to estimate several means. The results were multivariate generalizations of the previous work of Owen and the nonparametric version of Wilk's theorem. It was observed that empirical likelihood intervals for a single mean are not affected significantly by skewness as compared to t-statistics by applying the Cornish Fisher expansion. The author presented an effective method for computation of empirical profile likelihoods for a vector random variable mean.

As stated above, Owen (1991) explored the properties of non-parametric empirical likelihood inference and explained that the sampling properties of empirical likelihood estimators are similar to bootstrapping. The nonparametric empirical likelihood methods from mean estimation to the estimation of regression models were explored. He considered heteroskedastic and robust regression along with fixed and random regressors. Therefore, homoscedasticity is not required to estimate the regression with the empirical likelihood approach. Three extensions of empirical likelihood method were presented, including constrained empirical likelihood for unknown distribution, examined Euclidean distance that is an alternative to Likelihood function, and presented the Triangular empirical likelihood method.

Additionally, Kolaczyk (1994) demonstrated that the Empirical likelihood approach is admissible for the class of inference in econometrics. He extended the use of the EL approach to generalized linear models. The study of Chen and Keilegom (2009) discussed the review of the EL approach to estimate the regression models and their inference problems. They considered the nonparametric, semiparametric, and parametric regression models on censored and missing data. Consequently, they suggested the EL approach can be used to construct the likelihood ratio for inference to the regression parameters.

As illustrated by Owen (1988), the parameter of unknown distribution for an empirical likelihood ratio (ELR) has a Chi-square distribution. In order to further extend, Qin and Lawless (1994) commenced a study with the prime objective to link moment equations and empirical likelihood approach. They presented how estimators of both parameters, as well as the underlying distribution, could be determined, and normal distribution can be obtained asymptotically for estimators. Moreover, they revealed that for unknown parameters, the empirical likelihood ratio has a chi-square distribution. The results found parameters to be similar for parametric likelihood inference. This demonstrates that the empirical likelihood function and the parametric likelihood function have the same properties. They recommended the EL approach as the most beneficial for the likelihood ratio approach on which to construct the confidence intervals and hypothesis testing based to some extent.

In extension to the study of (Owen, 1991; Qin & Lawless, 1994); the author's Qin and Lawless (1995) compared other approaches as well as extended EL approach in the context of constraints on parameters. Their results of the simulation indicated that ELR static approaches to their chi-square distribution are as similar as the Pseudo score statistics. In several situations, the ELR approach behaves similarly to both asymptotic and finite samples. They concluded that all the methods would be beneficial from the second order adjustments to improve the accuracy of the confidence level and to test the hypothesis for either small or moderate sample size.

A number of authors (Imbens et al., 1998; Kitamura & Stutzer, 1997; Qin & Lawless, 1994) studied the small sample properties of the GMM approach and identified that the

GMM approach has poor finite sample properties. Consequently, they suggested alternative approaches to the GMM approach based on moment restrictions. In accordance with the above, Kitamura and Stutzer (1997) proposed an alternative approach for minimizing the Kullback-Leibler the information-theoretic approach that can handle weak data sampling process. They also suggested that this approach can be used in overdetermined models. Moreover, the proposed estimator can be used to test the hypothesis of over identified restrictions as an alternative approach to the GMM approach. This approach has a similarity to the GMM approach asymptotically. The study of Kitamura (2001) explored that in the case of an overdetermined restriction test, the empirical likelihood approach overlooks other approaches in terms of large deviation principle. This study examined the size and power of moment restriction tests asymptotically. The author observed that the standard GMM has greater power than the iterative GMM test; however, it looses the power comparative to empirical likelihood ratio (ELR) test. Finally, author indicated that ELR test have value able properties than GMM approach.

In another study, an asymptotic efficient method for estimating models with conditional moment restriction was proposed (Kitamura, Tripathi, & Ahn, 2004). These authors extended the methods of Empirical likelihood approach introduced by (Owen, 1988; 1990; 1991). They also generalized the estimator of (Qin & Lawless, 1994). They conduct the Monte Carlo simulation analysis to test the efficiency of their estimator in small samples, and their results showed that their test works very well in comparison to some other estimators. They also proposed a likelihood ratio static for testing the restriction on parameters.

The study of Imbens et al. (1998) focused on an alternative approach to testing the over-identified model in spite of improving the finite sample properties of the standard GMM estimator. The authors also worked using an exponential tilting approach as it appears more appealing than the empirical likelihood approach. This study made three key contributions. The primary contribution is to introduce an appealing method for calculating one step estimators. Secondly, the study offered substitutes to the GMM approach, which are generally constructed on a quadratic form in case of the average moments. Thirdly, the authors conducted a Monte Carlo experiment and determined that the standard GMM approach gives poor results. Furthermore, they tested over-identified restrictions in the context of the cross section, and it is concluded that an exponential tilting approach can be used to test the hypothesis and to construct confidence intervals because it has finite sample properties.

The literature suggested a great number of alternative approaches in order to shrink the finite sample bias of the GMM approach. The empirical likelihood and exponential tilting approach are more valuable among all other suggested approaches. The study of Newey and Smith (2004) reported that the EL approach demonstrates lower bias than the GMM approach in small samples. The findings of the study revealed two theoretical benefits of the EL approach. First, as the number of moment restriction increases, the GMM biasedness increases while the empirical likelihood biasedness does not increase asymptotically. Secondly, they identified that the higher-order efficiency of the EL approach increases after applying bias correction in comparison to other approaches. The study also reported that the biasedness of the GMM approach increases with the square root of the number of overdetermined restrictions as compared to its asymptotic standard error, whereas the EL approach does not do so with a number of restrictions.

The authors emphasized that the efficiency of the EL approach remains controlled with the condition of bias correction among all other bias correction estimators; however, in case of elimination of the condition, the mean square error of EL approach may not be smaller. The study found that the nonparametric EL approach is equally beneficial as the parametric Maximum Likelihood Estimator (MLE) in discrete data set asymptotically. Hence, the EL is proved as the best alternative approach to the GMM estimator.

The estimation of confidence intervals with the GMM approach is based on normal approximation, as reported by (Imbens & Spady, 2002). This study used the EL approach as an alternative to the GMM approach for the construction of confidence intervals. These confidence intervals have identical asymptotic properties to the standard GMM approach; however, these properties are considerably different in small samples. These confidence intervals can be applied to both exact and overdetermined cases. In order to check small sample properties, they conducted the Monte Carlo simulation analysis in different cases. The results of the two cases suggested considerably better confidence intervals; however, they demonstrated poor intervals in the third case. The performance of all the confidence intervals increased with bootstrap experiment, but EL based confidence interval remained outstanding in bootstrapping investigation.

Golan et al. (1996) developed the entropy-based formulation that allowed them to solve a wide range of estimation and inference problems in econometrics. Similarly, Mittelhammer, Judge, and Miller (2000) placed GMM within the framework of EL and Maximum Entropy (ME) for estimation. It can be shown that many of these estimation techniques can be obtained as special cases of minimizing Cressie and Read (1984) power divergence criterion that comes directly from the Pearson (1900) chi-squared statistic. Infinite sample econometrics, the exact distribution of estimators and test statistics are usually unknown. This problem was explored in the study of Ullah (2002) that examined the application of Kullback-Leibler divergence measure to nonparametric estimation and hypothesis testing in regression models.

2.3 Cressie and Read (CR) Family of Divergence

The study of Cressie and Read (1984; 1988) proposed a family of divergence measures that comprises of three main alternatives in its family of estimators. These alternatives are known as Empirical Likelihood (EL), Kullback-Leibler (KL) – (Maximum Exponential Empirical Likelihood (MEEL) is a special case of KL information-theoretic approach) and the Euclidean Likelihood estimation techniques. The authors Judge and Mittelhammer (2011) discussed that the Kullback-Leibler divergence measure is a special case of Cressie and Read family of divergence. They proved that if we use $\gamma = 0$, the CR family of test statistics leads to Maximum Empirical Exponential Likelihood (MEEL) as: $\sum p_i \ln (p_i)$. The authors also discussed that if we use the uniform distribution as reference distribution in Kullback-Leibler (KL) information criterion, it is exactly equal to the MEEL approach. Additionally, they stated MEEL as an updated variant of the GMM estimator as it handles unknown covariance matrix internally for estimation of problem. In inference methods, the members of the CR information-theoretic approach is similar to traditional Maximum Likelihood (ML), and the GMM approaches.

The general linear models contemplated by (Akkeren, Judge, & Mittelhammer, 2002); explored a new test having attractive small and asymptotic sampling properties that are known as Data Base Information Theoretic (DBIT) estimator. In their Monte Carlo experiment, they compared three estimators, including GMM estimator, 2SLS, and DBIT. They found that DBIT is superior from both GMM and 2SLS estimators in the situation of dual balance criterion. Consequently, DBIT estimator is more appropriate in the situation of small samples having weak instrumental variables than the GMM and 2SLS estimators.

2.3.1 Optimal Convex Combinations of Cressie and Read Family

The variety of empirical studies has been conducted in social sciences to explore the basis of partial or incomplete knowledge about data structure or theoretical relationships that resulted in uncertainty in statistical models. The inappropriate model specification may negatively affect the estimation and inaccuracy in inference. In this context, Judge and Mittelhammer (2004) considered the Semi Parametric Stein Like (SPSL) estimator that reduces error in estimation on the basis of the quadratic loss function. SPSL performance is considered as superior to the least square estimator. This estimator achieves consistency and asymptotic normality. The reduction of Mean Square Error (MSE) in estimation is considered as the prime objective on which SPSL estimator is based. This objective is designed for MSE improvement in the case of a semiparametric approach. Several empirical pieces of evidence reported that the SPSL estimator performs better for both finite samples and asymptotically. They developed the formula in the context of two individual estimators, but it can be extended for more than two estimators.

Furthermore, the study of Judge and Mittelhammer (2012) proposed the optimal combination of two or more individual estimators computed using the CR family of information criterion. This optimal combination has a minimum mean square error as compared to the individual estimators. They used the concept of the CR information-theoretic approach and produced different approaches for estimation and inference. These approaches have flexible probability distributions. They conducted the Monte Carlo experiment to demonstrate small sample performance for an optimal estimator that is computed from the CR family of test statistics under the assumption of orthogonal condition that is not satisfied. They presented the more general procedure that gives the possibility to enhance the performance of the estimator.

2.4 Literature Gap

The literature frequently investigated the econometric model estimation with moment restriction, and the most popular discussed approach is the Generalized Method of Moments (GMM) due to its large sample properties (Hansen, 1982). The GMM approach is used to solve the problem of orthogonality conditions in overdetermined models. The small sample properties were ignored in his work. However, Hansen et al. (1996) explored the small sample properties of the GMM approach and compared it with other estimators. They suggested an estimator whose small sample properties are better than GMM estimator and named it "Continuous Updating Estimator (CUE)." Several authors reported that GMM estimator does not have the appropriate asymptotic properties and it has large finite sample bias (Altonji & Segal, 1996; Hall & Horowitz, 1996; Imbens et al., 1998; Imbens, 1997; Kitamura & Stutzer, 1997; Qin & Lawless, 1994).

In contrast, an alternative nonparametric approach was proposed in a study (Owen, 1988; Qin & Lawless, 1994) that is known as the Maximum Empirical Likelihood (MEL) approach, and it has improved the inference in econometric models. Numerous studies (Imbens & Spady, 2002; Kitamura, 2001; Newey & Smith, 2004) reported that the MEL approach has low bias than the GMM approach in finite samples. Conversely, the authors (Imbens et al., 1998; Kitamura & Stutzer, 1997) proposed an alternative approach that is exponential tilting / Kullback-Leibler Information Criterion (KLIC) as replacement of the GMM approach.

The study of Mittelhammer and Judge (2003) discussed that the GMM estimator usually requires two steps estimation procedure with the best weight matrix in estimation; however, its small sample performance is very poor. On the other hand, the empirical likelihood approach requires an only one-step procedure in which the weight matrix is handled within the estimation process that improves the performance of the estimator in finite samples. In parallel to these studies (Judge & Mittelhammer, 2012; Judge & Mittelhammer, 2011) discussed that the CR family of divergence encompasses all entropy measures, especially MEL and MEEL, which are the members of this family. They proposed a convex combination approach, which is based on an optimal choice of the individual estimators whose mean square error is lower than individual estimators from the CR family of the information-theoretic approach.

The above literature addressed the usage and benefits of different approaches however no one ever did the comparison of CR Optimal Convex Combination (CROCC) (Judge & Mittelhammer, 2012; Judge & Mittelhammer, 2011) of the individual estimator (MEL and MEEL estimators) with GMM approach in exact and overdetermined models when orthogonality condition isn't satisfied. This research is going to address this gap. The literature does not show the application of the information-theoretic approach in the estimation of economic models in the presence of endogeneity problem. The researchers use the GMM approach to solve the simultaneity issue, although they know that the GMM estimator has poor finite sample properties. We are going to offer a better solution for the implementation of an information-theoretic approach to solve such kind of issues in finite samples. The coefficient of Cressie and Read Optimal Convex Combination estimator (Judge & Mittelhammer, 2011) is based on assumption of two independent samples. In this context, we are introducing CR Arbitrary Convex Combinations (CRACC) which has superiorty to CROCC due to its assumption free estimation of two independent samples and compare its performance with CROCC in finite samples. This comparison would help to evaluate the finite sample properties of CROCC in different scenarios.

CHAPTER 3. METHODOLOGY

The main purpose of this study is to compare different nonparametric approaches to estimate the regression models in the presence of endogeneity. Particularly, GMM estimator is compared with information-theoretic estimators in finite samples when orthogonality conditions do not hold. For the purpose of comparison, the linear regression model is specified as:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \qquad \text{3-1}$$

Where "**Y**" is an (n × 1) vector of the dependent variable, "**X**" is an (n × k) matrix of regressors, "**β**" is a (k × 1) vector of the unknown parameters and "**ϵ**" is an (n × 1) vector of residuals. The equation "**3-1**" should satisfy the assumption of homoscedasticity $\mathbf{COV}\,(\boldsymbol{\epsilon}|\mathbf{X}) = \sigma^2\,\mathbf{I}_n$. One of the main assumptions of the Gauss-Markov Theorem is the endogeneity problem, i.e., the error term "**ϵ**" of the regression must be orthogonal to the regressors, which directly implies the orthogonality and independence of the dependent variable. The endogeneity problem occurs when the orthogonality condition of the "**ϵ**" with **X** is not satisfied (i.e., $\mathbf{E}(\mathbf{X}'\boldsymbol{\epsilon}) \neq \mathbf{0}$). The OLS estimator of (k × 1) unknown parameters is given as

$$\hat{\beta}_{ols} = (X'X)^{-1}X'Y \qquad \text{3-1A}$$

By substituting "**3-1**" in the OLS estimator "**3-1A**" and after simplifying and the expected value of "$\hat{\boldsymbol{\beta}}_{ols}$" is given as follows

$$E(\hat{\boldsymbol{\beta}}_{ols}) = \boldsymbol{\beta} + (\mathbf{X}'\mathbf{X})^{-1}E(\mathbf{X}'\boldsymbol{\epsilon}) \qquad \text{3-1B}$$

In the presence of endogeneity problem the OLS estimator produces biased and inconsistent estimates as the term "$(\mathbf{X}'\mathbf{X})^{-1}E(\mathbf{X}'\boldsymbol{\epsilon})$" shows in "**3-1B**". The violation of the exogeneity assumption of Gauss–Markov theorem (i.e., $E(\mathbf{X}'\boldsymbol{\epsilon}) \neq 0$) in "**3-1B**" leads to OLS estimates bias. Other than OLS, various techniques have been introduced that use instrumental variables to solve such an endogeneity issue in this respect. These techniques are used in such a way that satisfies the orthogonality condition of the "$\boldsymbol{\epsilon}$" with instrumental variables '**Z**' (i.e. $\mathbf{E}(\mathbf{Z}'\boldsymbol{\epsilon}) = \mathbf{0}$). Among these techniques, the Generalized Method of Moments (GMM) is the most popular estimator. However, various studies have documented that the GMM estimator has poor finite samples properties (Altonji & Segal, 1996; Hall & Horowitz, 1996; Imbens et al., 1998; Imbens, 1997; Kitamura & Stutzer, 1997; Qin & Lawless, 1994). Therefore, it is essential to identify alternative techniques having good finite sample properties to solve the problem mentioned above of endogeneity.

This dissertation argued that non-parametric approaches, including Maximum Empirical Likelihood (MEL), Maximum Exponential Empirical Likelihood (MEEL), and Cressie and Read Optimal Convex Combination (CROCC) can be among those alternatives. These estimators are the members of the Cressie and Read family of the information-theoretic approach. Hence, estimators under discussion are GMM, MEL, MEEL, and CROCC, and their finite sample properties are compared in solving the endogeneity problem. Additionally, we are introducing an estimator, Cressie and Read Arbitrary Convex Combination (CRACC) and we will test its finite samples properties. We will also compare this estimator with the CROCC estimator proposed by Judge and Mittelhammer (2011) to evaluate the performance of the CROCC and CRACC. The details of all these different approaches are given below.

3.1 Generalize Method of Moment (GMM) Approach

The GMM estimator is the most commonly used in econometric modeling when orthogonality conditions do not satisfy. The GMM is an approach used to estimate the linear regression model when the moment conditions are greater than the number of parameters proposed by (Hansen, 1982). According to "**3-1**" where $\mathbf{Y}_{(n\times 1)}$, $\mathbf{X}_{(n\times k)}$ $\boldsymbol{\beta}_{(k\times 1)}$ and the matrix of instrumental variables is $\mathbf{Z}_{(n\times \ell)}$ and $\mathbf{W}_{(\ell\times \ell)}$ is the weight matrix which will be defined later. The GMM approach is implemented by defining a weight Euclidean Distance based estimation objective function.

$$Q = \{[\mathbf{Z}' (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})]' \, \mathbf{W} \, [\mathbf{Z}' (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})]\} \quad \text{3-2}$$

The GMM estimator seeks to minimize the Weighted Euclidean distance "Q." A key problem in the implementation of the GMM estimator is its dependence upon the choice of the weight matrix, i.e., "**W**." In practice, the efficient choice of the "**W**" is not defined theoretically. Hence, in the case of the implementation of the GMM approach, the choice of "**W**" is the main issue. Though Hansen (1982) proposed a solution for efficient choice of "**W**", however, his approach is criticised due to poor finite samples properties reported by various authors (Altonji & Segal, 1996; Hall & Horowitz, 1996; Imbens et al., 1998; Imbens, 1997; Kitamura & Stutzer, 1997; Qin & Lawless, 1994). According to the study of Hansen (1982), the weight matrix "**W**" is defined as:

$$\mathbf{W} = (\mathbf{Z}'\sigma^2 \, \mathbf{I}_n \mathbf{Z})^{-1}$$

Where "σ^2" is specified in "**3-1**".

The Method of Moment (MOM) is a special case of GMM estimator when the moment conditions are equal to the number of unknown parameters, i.e., $\ell = \mathrm{k}$ known as an exactly determined model. Therefore, MOM estimator is given as follows

$$\hat{\beta}_{MOM} = [Z'X]^{-1} Z'Y \qquad 3\text{-}3$$

Conversely, for the overdetermined model, i.e., where moment conditions are greater than the number of parameters, i.e., $\ell > \mathbf{k}$, the GMM estimator will estimate the parameters by minimizing the "$\boldsymbol{Q}$" as specified in "**3-2**". Now to obtain the first-order condition, "$\boldsymbol{Q}$" is differentiated w.r.to "$\boldsymbol{\beta}$" and equated to zero. The resulting "$\hat{\boldsymbol{\beta}}$" known as "$\hat{\boldsymbol{\beta}}_{GMM}$" is given as

$$\hat{\beta}_{GMM} = [X'Z\, W\, Z'X]^{-1} X'Z\, W\, Z'Y \qquad 3\text{-}4$$

A number of studies have discussed the poor finite samples properties of the GMM approach (Altonji & Segal, 1996; Hall & Horowitz, 1996). The efficiency of the GMM estimator is based on "**W**." Conversely, alternative approaches (MEL and MEEL) have been suggested by various authors (Altonji & Segal, 1996; Hall & Horowitz, 1996; Imbens et al., 1998; Imbens, 1997; Kitamura & Stutzer, 1997; Qin & Lawless, 1994). These approaches require an only one-step procedure in which the "**W**" is handled within the estimation process, improving the performance of the estimator in finite samples (Judge & Mittelhammer, 2011; 2012).

3.2 Nonparametric Maximum Likelihood (NPML) Approach

To make the Non-parametric Maximum Likelihood (NPML) function, let us consider the case of i.i.d simple random sample $\boldsymbol{Y} = (Y_1, Y_2, Y_3, \ldots\ Y_n)'$ from the population having the probability density function (PDF) $f(y; \boldsymbol{\Theta})$ where $\boldsymbol{\Theta}$ is the parameter of the population and y_i's are random scalars. In this situation, the true likelihood function for the parameter Θ can be expressed as $L(\boldsymbol{\Theta}; \mathbf{y}) = \prod_{i=1}^{n} f(y_i; \boldsymbol{\Theta})$. In this situation, we assume that the parametric functional form $f(y; \boldsymbol{\Theta})$ is not known. However, if the parametric form is unknown, then the non-parametric maximum likelihood (NPML) function can be constructed (Judge & Mittelhammer, 2011). According to authors, consider random sample $\boldsymbol{y} = (y_1, y_2, y_3, . . ., y_n)$ to recover an estimate of PDF **y**. Generally, the objective of NPML is to estimate the entire probability density or mass function $f(.)$ instead of estimation of the unknown parameter "**ϴ**" that maximizes the likelihood function. Therefore, the NPML function is written as

$$\mathbf{L}(f; \mathbf{y}) = \prod_{i=1}^{n} f(y_i) \qquad 3\text{-}5$$

Here, the NPML function "$L(f; \boldsymbol{y})$" specifies the likelihood that the estimated function $f(.)$ is the true population distribution underlying random samples $\boldsymbol{y}$. In the NPML function, the estimates of $f(y)$ are explained as follows:

$$\hat{f}(y) = \mathbf{argmax}_f\, \mathbf{L}(f; y) = \mathbf{argmax}_f \prod_{i=1}^{n} f(y_i) \qquad 3\text{-}6$$

Judge and Mittelhammer (2011) stated that "The feasible space for this maximization problem consists of all possible functions that specify the properties of probability density/mass functions." The maximization problem in above "**3-6**" can be converted into a simple parametric form by observing it as unknown values $\boldsymbol{p_i} = f(y_i)$ (where $\boldsymbol{p_i}$ is ith probability weight allocated to the ith observation: y_i of the sample outcome) used to solve the maximum likelihood problem in "**3-6**". Afterward, we can define the empirical probability mass function of a multinomial type that represents discrete probability allocated to each number of observation of sample outcomes. In this framework, the probability weights must satisfy the condition. "$\boldsymbol{p_i} > 0$" for all sample outcomes of the value of the joint likelihood for the observed sample is $\prod_{i=1}^{n} f(y_i) = 0$, which is a minimum as opposed to a maximum value of $\mathrm{L}(f; \boldsymbol{y})$. The maximum likelihood function "**3-6**" can be denoted as parametric maximum likelihood problem of finding the optimal choice of $\boldsymbol{p_i}$'s in a multinomial-type likelihood function in which each "type" of outcome occurs only once. As a result

$$(\widehat{\boldsymbol{p}}_1, \ldots \widehat{\boldsymbol{p}}_n) = \mathbf{argmax_p}[\textstyle\prod_{i=1}^{n} \boldsymbol{p_i}] = \mathbf{argmax}_p[\mathbf{n}^{-1} \textstyle\sum_{i=1}^{n} \mathbf{ln}\ (\boldsymbol{p}_i)] \qquad 3\text{-}7$$

Where $\widehat{\boldsymbol{p}}_i > 0$ for all i.

In the maximization problem, the transformed parametric likelihood function in "**3-7**" is unbounded. It is because the probabilities $\boldsymbol{p_i}$ are unrestricted in values. In this situation, no solution will exist unless we apply the normalization condition to probability weights, as given in the following equation.

$$\textstyle\sum_{i=1}^{n}(\boldsymbol{p}_i) = \mathbf{1} \qquad 3\text{-}7A$$

Therefore, in the estimation of NPML the objective function $n^{-1} \sum_{i=1}^{n} \ln\ (\boldsymbol{p_i})$ will be maximized with the constraint of $\sum_{i=1}^{n} (\boldsymbol{p}_i) = 1$. Hence, the Lagrange function of the optimization problem can be written as follow:

$$\mathbf{ln}(\mathbf{L}(\boldsymbol{p}, \boldsymbol{\eta})) = \mathbf{n}^{-1} \textstyle\sum_{i=1}^{n} \mathbf{ln}(\boldsymbol{p_i}) - \boldsymbol{\eta}(\textstyle\sum_{i=1}^{n} \boldsymbol{p_i} - \mathbf{1}) \qquad 3\text{-}7B$$

Similarly, the parameter estimates of linear regression models can be obtained from the NPML estimation approach, which is detailed below.

3.2.1 Maximum Empirical Likelihood (MEL) Estimation for Linear Model

In the study of (Kolaczyk, 1994; Owen, 1991; Judge & Mittelhammer, 2011), the authors presented the use of the NPML/ MEL approach to estimate the linear regression models. Consider a linear regression model of the form: $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$, where "**Y**" is an $(n \times 1)$ vector of dependent variable, "**X**" is an $(n \times k)$ matrix of regressors, "$\boldsymbol{\beta}$" is a

$(k \times 1)$ vector of the unknown parameters, and "ϵ" is an $(n \times 1)$ vector of residuals. If $\mathbf{X}$ and ϵ are independent i.e. the orthogonality condition is satisfied $\mathbf{E}(\mathbf{X}'\epsilon) = \mathbf{0}$, conditional mean function for $\mathbf{Y}$ is given as $E(\mathbf{Y}|\mathbf{X} = \mathbf{X}\boldsymbol{\beta})$, and $(\mathbf{Y}_i, \mathbf{X}_{[i,.]})$ $i = 1,2,3, \dots, n$ obervations. Where $\mathbf{Y}_i$ is the ith observation of vector $\mathbf{Y}_{(n\times1)}$ and $\mathbf{X}_{[i,.]}$ is the ith row of the matrix $\mathbf{X}$.

In the estimation of the linear regression model by using MEL approach, let us consider the ith $(k \times 1)$ moment vector function $\mathbf{h}\left(\left(\mathbf{Y}_i, \mathbf{X}_{[i,.]}\right), \boldsymbol{\beta}\right) = \mathbf{X}_{[i,.]}'(\mathbf{Y}_i - \mathbf{X}_{[i,.]}\boldsymbol{\beta})$ for any $i = 1,2,3, \dots, n$. Under the assumption of orthogonality condition, we can assume that

$$E(\mathrm{h}\left(\left(\mathbf{Y}_i, \mathbf{X}_{[i,.]}\right), \boldsymbol{\beta}\right)) = E(\mathbf{X}_{[i,.]}'(\mathbf{Y}_i - \mathbf{X}_{[i,.]}\boldsymbol{\beta})) = 0, \text{ for all } i = 1,2,3, \dots, n \qquad 3\text{-}8$$

In the estimation of the linear regression model, when "$\boldsymbol{\beta}$" is the true value of the parameter vector "$\boldsymbol{\beta}$." Therefore, the moment vector function is unbiased. Moreover, the random vectors $\mathbf{M}_{[.,i]} \equiv \mathbf{X}'_{[i,.]}(\mathbf{Y}_i - \mathbf{X}_{[i,.]}\boldsymbol{\beta})$ for all $i = 1,2,3, \dots, n$ are iid random vectors under the prevailing assumptions because $(\mathbf{Y}_i, \mathbf{X}_{[i,.]})$ $i = 1,2,3, \dots, n$ are iid. $\mathbf{M}_{[.,i]}$ is a $k \times 1$ vector for each i. When we take the expected value of $\mathbf{M}_{[.,i]}$ i.e., $E(\mathbf{M}_{[.,i]}) = 0$ then $\mathbf{M}_{[.,i]}$ is summed up for all $i = 1,2,3, \dots, n$, and a $(k \times 1)$ vector is obtained known as k moment conditions for estimating k parameters.

Assuming the MEL function for "$\boldsymbol{\beta}$", multiplied by (n^{-1}) log-EL is expressed by solving constrained maximization problem. The Lagrange form of log-EL is stated as follows

$$\ln\left(\mathbf{L}_{\mathbf{EL}}(\boldsymbol{\beta}; \mathbf{Y}_i, \mathbf{X}_{[i,.]})\right) = n^{-1}\sum_{i=1}^{n}\ln(\boldsymbol{p}_i) - \mu\left(\sum_{i=1}^{n}\boldsymbol{p}_i - 1\right) - \boldsymbol{\lambda}'\sum_{i=1}^{n}\boldsymbol{p}_i\left(\mathbf{X}'_{[i,.]}(\mathbf{Y}_i - \mathbf{X}_{[i,.]}\boldsymbol{\beta})\right) \quad \text{3-9}$$

Where $\boldsymbol{\lambda}_{(k\times 1)}$, μ are Lagrange multipliers and $\boldsymbol{p}_i$ is defined as in section "**3-2**". The $\boldsymbol{p}_i > 0 \; \forall i$ is implicit in the structure of the optimization problem. To fulfill the first-order condition w.r.to the $\boldsymbol{p}_i$'s is as follows

$$\frac{1}{n\boldsymbol{p}_i} - \boldsymbol{\lambda}'\left(\mathbf{X}'_{[i,.]}(\mathbf{Y}_i - \mathbf{X}_{[i,.]}\boldsymbol{\beta})\right) - \boldsymbol{\mu} = 0, i = 1,2,3, \ldots n \quad \text{3-10}$$

The optimal value of $\boldsymbol{\mu} = 1$ because $\sum_{i=1}^{n}\boldsymbol{p}_i = 1$. Substituting the value of $\mu = 1$ in "**3-10**" as a result the solution of $\boldsymbol{p}_i$ is given as

$$\boldsymbol{p}_i(\boldsymbol{\beta}, \boldsymbol{\lambda}) = \left(n\left[\boldsymbol{\lambda}'\mathbf{X}'_{[i,.]}(\mathbf{Y}_i - \mathbf{X}_{[i,.]}\boldsymbol{\beta})\right] + 1\right)^{-1} \quad \text{3-11}$$

The first-order condition w.r.to $\boldsymbol{\lambda}$ Lagrange multiplier

$$\sum_{i=1}^{n}\boldsymbol{p}_i\left(\mathbf{X}'_{[i,.]}(\mathbf{Y}_i - \mathbf{X}_{[i,.]}\boldsymbol{\beta})\right) = \mathbf{0} \quad \text{3-12}$$

Substitute the optimal value of $\boldsymbol{p}_i$ in "**3-12**" we get the expression that is a function of $\boldsymbol{\beta}$ as follow:

$$\boldsymbol{\lambda}(\boldsymbol{\beta}) = \mathbf{arg}_{\boldsymbol{\lambda}}\left[\sum_{i=1}^{n}\left(\boldsymbol{n}\left[\boldsymbol{\lambda}'\mathbf{X}'_{[i,.]}(\mathbf{Y}_i - \mathbf{X}_{[i,.]}\boldsymbol{\beta})\right] + \mathbf{1}\right)^{-1}\left(\mathbf{X}'_{[i,.]}(\mathbf{Y}_i - \mathbf{X}_{[i,.]}\boldsymbol{\beta})\right) = \mathbf{0}\right]$$

3-13

We get expression in "**3-13,**" i.e., $\boldsymbol{\lambda}(\boldsymbol{\beta})$, which is an implicit function of the unknown parameter $\boldsymbol{\beta}$, it cannot be expressed in closed form (Judge & Mittelhammer, 2011).

According to the work of Qin and Lawless, (1994), the multinomial probability weights $\boldsymbol{p_i}$ must satisfy the condition of "$0 \leq \boldsymbol{p_i} \leq 1$" which indicates that "$\boldsymbol{\lambda}$" and "$\boldsymbol{\beta}$" necessarily satisfy the condition "$\boldsymbol{\lambda}'\left(\mathbf{X}'_{[i,.]}(\mathbf{Y}_i - \mathbf{X}_{[i,.]}\boldsymbol{\beta})\right) + 1 > 1/n$" for each i. At the fixed value of "$\boldsymbol{\beta}$," assuming that $\boldsymbol{D}_{\boldsymbol{\beta}} = \{\boldsymbol{\lambda}\text{: } (\boldsymbol{\lambda}'\left(\mathbf{X}'_{[i,.]}(\mathbf{Y}_i - \mathbf{X}_{[i,.]}\boldsymbol{\beta})\right) + 1 \geq \frac{1}{n})\}$; that shows $\boldsymbol{D}_{\boldsymbol{\beta}}$ is convex and compact. The authors also discussed that $\boldsymbol{\lambda}(\boldsymbol{\beta})$ is continuous differentiable function of "$\boldsymbol{\beta}$."

By putting the optimal solution of Lagrange multiplier "$\boldsymbol{\lambda}(\boldsymbol{\beta})$" into the optimal $\boldsymbol{p_i}$ in "**3-11**", that allows the "$\boldsymbol{p_i}$" to be denoted as $\boldsymbol{p_i}(\boldsymbol{\beta}, \boldsymbol{\lambda}(\boldsymbol{\beta})) = (\text{n}[\boldsymbol{\lambda}(\boldsymbol{\beta})'\left(\mathbf{X}'_{[i,.]}(\mathbf{Y}_i - \mathbf{X}_{[i,.]}\boldsymbol{\beta})\right) + 1])^{-1}$. According to Judge and Mittelhammer (2011), these probability weights put into the log-EL objective function which can be written as

$$\ln\left(\text{L}_{\text{EL}}(\boldsymbol{\beta}; \mathbf{Y}, \mathbf{X}_{[i,.]})\right) = -\sum_{i=1}^{n} \ln\left(\text{n}[\boldsymbol{\lambda}(\boldsymbol{\beta})'\left(\mathbf{X}'_{[i,.]}(\mathbf{Y}_i - \mathbf{X}_{[i,.]}\boldsymbol{\beta})\right) + 1]\right) \qquad \text{3-14}$$

The MEL function for "$\boldsymbol{\beta}$" is defined as

$$\widehat{\boldsymbol{\beta}}_{EL} = \arg\max_{\boldsymbol{\beta}} [\ln\left(\text{L}_{\text{EL}}(\boldsymbol{\beta}; \mathbf{Y}, \mathbf{X}_{[i,.]})\right)] \qquad \text{3-15}$$

In the linear regression model, the MEL estimator to estimate "$\boldsymbol{\beta}$," which is based on unbiased moment function in "**3-8**" leads to familiar functional form for "$\widehat{\boldsymbol{\beta}}_{EL}$". The Cressie and Read family (1984 & 1988) of test statistics contains various nonparametric likelihood approaches, including the MEL approach that estimates the unknown parameters "$\boldsymbol{\beta}$" in linear regression models.

According to the objective of the present study, our focus is to estimate the regression model having an endogeneity issue. In this case, the assumption of orthogonality condition does not hold (i.e., $\text{E}(\mathbf{X}'\epsilon) \neq 0$). So, the instrumental variables '**Z**' are required to solve such a problem that satisfies the condition of $E(\mathbf{Z}'\epsilon) = 0$. In this setting, the moment functions are as follows:

$$E\left(h\left((\mathbf{Y}_i, \mathbf{X}_{[i,.]}, \mathbf{Z}_{[i,.]}), \boldsymbol{\beta}\right)\right) = E\left(\mathbf{Z}'_{[i,.]}(\mathbf{Y} - \mathbf{X}_{[i,.]}\boldsymbol{\beta})\right) = 0 \quad \text{for } i = 1, 2, 3, \ldots, n \qquad 3\text{-}16$$

$$\ln\left(\mathbf{L}_{\mathbf{EL}}(\boldsymbol{\beta}; \mathbf{Y}_i, \mathbf{X}_{[i,.]}, \mathbf{Z}_{[i,.]})\right) = \max_p[n^{-1}\textstyle\sum_{i=1}^{n}\ln(p_i) \text{ subject to } \sum_{i=1}^{n} p_i = 1 \text{ and } \sum_{i=1}^{n} p_i\left(\mathbf{Z}'_{[i,.]}(\mathbf{Y} - \mathbf{X}_{[i,.]}\boldsymbol{\beta})\right) = 0] \qquad 3\text{-}17$$

Where $\mathbf{Y}_{(n\times 1)}, \mathbf{X}_{(n\times k)}, \mathbf{Z}_{(n\times \ell)}$ $\boldsymbol{\beta}_{(k\times 1)}$and $\boldsymbol{\lambda}_{(k\times 1)}$ and "$k = \ell$".

In the case of the exactly determined model, where "ℓ" estimating moment equations are used to estimate the vector of "k" unknown parameters that provide unique values of unknown parameter "$\boldsymbol{\beta}$."

3.3 Kullback-Leibler (KL)/ Maximum Empirical Exponential Likelihood (MEEL) Approach

Judge and Mittelhammer (2011) discussed the role of the KL information-theoretic approach to measuring the discrepancy between two probability distributions, i.e., the subjective distribution $p(y)$ and the reference distribution $q(y)$. The KL information criterion is denoted as:

$$\mathbf{KL}(p, \mathbf{q}) = \textstyle\sum_{i=1}^{n} \mathbf{p}_i \ln\left({}^{p_i}/_{q_i}\right) \qquad 3\text{-}18$$

The reference distribution is usually specified as uniform distribution "$q = (1/n)$". In the above equation, KL (p, q) is also known as the MEEL approach. Therefore, if the uniform distribution is used as a reference distribution in KL (p, q) information, then the objective function can solve the empirical moment equation of

$\sum_{i=1}^{n} p_i h(Y_i, X_{[i,.]}, Z_{[i,.]}, \beta) = \sum_{i=1}^{n} p_i \left(Z'_{[i,.]} \left(Y - X_{[i,.]}\beta\right)\right) = 0$, where instrumental variable "Z" is used due to $E(X'\epsilon) \neq 0$. Here, the objective is to minimize the KL discrepancy between a subject distribution and classical empirical distribution function on data. This KL/ MEEL criterion is also known as the maximum entropy criterion in literature.

$$\mathbf{KL}(\mathbf{p}, \mathbf{n}^{-1}\mathbf{1}_n) = \sum_{i=1}^{n} p_i \ln(n.p_i) = \sum_{i=1}^{n} p_i \ln(p_i) + \ln(n) \qquad 3\text{-}19$$

Minimizing KL information is equivalent to maximizing $-\sum_{i=1}^{n} p_i \ln(p_i)$, which is precisely the Shannon-Jaynes entropy measure stated in (Judge & Mittelhammer, 2011). Consequently, the Lagrange function is:

$$\ln\left(L_{MEEL}(\beta; Y, X_{[i,.]}, Z_{[i,.]})\right) = -\sum_{i=1}^{n} p_i \ln(p_i) - \mu\left(\sum_{i=1}^{n} p_i - 1\right) - \lambda' \sum_{i=1}^{n} p_i \left(Z'_{[i,.]}\left(Y - X_{[i,.]}\beta\right)\right) \qquad 3\text{-}20$$

Where $Y_{(n\times 1)}, X_{(n\times k)}, Z_{(n\times \ell)}$ $\beta_{(k\times 1)}$ and $\lambda_{(k\times 1)}$ and "$k = \ell$". In the case of the exactly determined model, where "ℓ" estimating moment equations are used to estimate the vector of "k" unknown parameters that provide unique values of unknown parameter "β." The "λ" and "μ" are Lagrange multipliers, and $p_i > 0 \ \forall\ i$ is implicitly structured in the above-stated log-likelihood function. The first order optimization condition applies to the above-stated function w.r.to "p_i", "λ" and "β" is as follows:

$$p_i(\beta, \lambda) = \frac{\exp[-\lambda'\left(Z'_{[i,.]}\left(Y_i - X_{[i,.]}\beta\right)\right)]}{\sum_{i=1}^{n} \exp[-\lambda'\left(Z'_{[i,.]}\left(Y_i - X_{[i,.]}\beta\right)\right)]} \qquad 3\text{-}21$$

Substitute the value of p_i, in "3-22" and "3-23" as

$$\lambda(\beta) = L_\lambda = \arg_\lambda \left[\sum_{i=1}^{n} \frac{\exp[-\lambda'(Z'_{[i,.]}(Y_i - X_{[i,.]}\beta))]}{\sum_{i=1}^{n} \exp[-\lambda'(Z'_{[i,.]}(Y_i - X_{[i,.]}\beta))]} (Z'_{[i,.]}(Y_i - X_{[i,.]}\beta)) = 0\right] \quad 3\text{-}22$$

$$L_\beta = \sum_{i=1}^{n} \frac{\exp[-\lambda'(Z'_{[i,.]}(Y_i - X_{[i,.]}\beta))]}{\sum_{i=1}^{n} \exp[-\lambda'(Z'_{[i,.]}(Y_i - X_{[i,.]}\beta))]} (-\lambda' Z'_{[i,.]} X_{[i,.]}) = 0 \quad 3\text{-}23$$

Where '$\lambda(\beta)$'s are the "k" optimal Lagrange multipliers with the "k" moment constraints $p_i > 0 \ \forall i$. The expression $\lambda(\beta)$ in "3-22", that is an implicit function of the vector of unknown parameters "β," it cannot be expressed in closed form. It is a continuously differentiable function of "β." By putting the optimal solution of Lagrange multiplier "$\lambda(\beta)$" into to the optimal probability weights in "3-21", that permits the empirical probabilities to be denoted in term of "β" as $p_i(\beta) = p_i(\beta, \lambda(\beta))$. These optimal probability weights put into the log MEEL objective function: $\ln\left(L_{MEEL}(\beta; Y_i, X_{[i,.]}, Z_{[i,.]})\right) = L_{MEEL}\left(\beta; \hat{p}_i(\beta), \hat{\lambda}(\beta)\right)$ (Judge & Mittelhammer, 2011).

3.3.1 Computation of MEL Approach

In the estimation of MEL type problems, analytical solutions are not simple. Consequently, we necessarily require numerical methods to solve these kinds of problems. There are many ways to solve such kind of estimation problems numerically. The Newton Raphson method is most commonly used. In the study of Mittelhammer et al. (2000), they discussed the procedure of the Newton-Raphson method to find the optimal point of a nonlinear function. Accordingly, the second-order Taylor expansion series is used for local approximation. Let us consider a nonlinear objective function $L(\beta)$ with the choice of parameter "β". Assumed an initial estimate of the unknown parameter "$\boldsymbol{\beta}_0$" to derive the iterative algorithm for numerical optimizations by Taylor series expansion as:

$$\mathbf{L}(\boldsymbol{\beta}) \cong \mathbf{L}(\boldsymbol{\beta}_0) + \frac{\partial \mathbf{L}(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}'}(\boldsymbol{\beta} - \boldsymbol{\beta}_0) + \frac{1}{2}(\boldsymbol{\beta} - \boldsymbol{\beta}_0)' \frac{\partial^2 \mathbf{L}(\boldsymbol{\beta})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}'}(\boldsymbol{\beta} - \boldsymbol{\beta}_0) \qquad 3\text{-}24$$

Take the derivatives of the above series, put them equal to zero, and find the optimal point. Conditional on the initial estimate. "β_0", the next iterate is

$$\boldsymbol{\beta}_1 = \boldsymbol{\beta}_0 - \left[\frac{\partial^2 \mathbf{L}(\boldsymbol{\beta})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}'}\right]^{-1} \left[\frac{\partial \mathbf{L}(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}'}\right] \qquad 3\text{-}25$$

In the iterating process, we reach an optimal point by solving the problems in which the current solution "$\boldsymbol{\beta}_t$" comes from the previous estimate "$\boldsymbol{\beta}_{t-1}$". In the Newton algorithm, the step taken during iteration "t" $\mathbf{d}_{t-1=} \boldsymbol{\beta}_t - \boldsymbol{\beta}_{t-1}$, is known as the direction. The direction is a path from the starting point to the next step in the iterative solution. The inverse of the Hessian matrix of the function is used to estimate the angle of the direction and gradient of the function is used to determine the size of the direction.

The numerical methods are used to estimate the MEL (nonparametric approaches) problems because it has no theoretical techniques to be solved. The simultaneous and sequential methods proposed by Judge and Mittelhammer (2011) to estimate the MEL problem by using the Newton-Raphson gradient-based search procedure. These authors used these methods to determine the unknown value of the population mean. They discussed that the sequential method is easier to solve the more complicated MEL (non-parametric approaches) problem. In the present study, we have adopted the sequential base numerical optimization procedure to solve the MEL and MEEL problem.

Considering the problem of log Maximum Empirical Likelihood (MEL) function with Lagrange constraint is given as follows:

$$\ln\left(L_{EL}\left(\beta; Y_i, X_{[i,.]}, Z_{[i,.]}\right)\right) = n^{-1}\sum_{i=1}^{n}\ln(p_i) - \mu\left(\sum_{i=1}^{n} p_i - 1\right) - \lambda' \sum_{i=1}^{n} p_i \left(Z'_{[i,.]}\left(Y_i - X_{[i,.]}\beta\right)\right) \quad 3\text{-}26$$

The first order optimization conditions w.r.to "p_i", "λ" and "β" are as follows:

$$p_i(\beta, \lambda) = \left(n\left[\lambda'\left(Z'_{[i,.]}\left(Y_i - X_{[i,.]}\beta\right)\right) + 1\right]\right)^{-1} \quad 3\text{-}27$$

$$L_\lambda = \sum_{i=1}^{n} p_i \left(Z'_{[i,.]}\left(Y_i - X_{[i,.]}\beta\right)\right) = 0 \quad 3\text{-}28$$

$$L_\beta = \sum_{i=1}^{n} p_i\left(-\lambda' Z'_{[i,.]} X_{[i,.]}\right) = 0 \quad 3\text{-}29$$

We restrict the equations "**3-28** and **3-29**" by putting the optimal value of p_i obtained from the equation "**3-27**" and then:

$$\lambda(\beta) = L_\lambda = \arg_\lambda\left[\sum_{i=1}^{n}\left(n\left[\lambda'\left(Z'_{[i,.]}\left(Y_i - X_{[i,.]}\beta\right)\right) + 1\right]\right)^{-1}\left(Z'_{[i,.]}\left(Y_i - X_{[i,.]}\beta\right)\right) = 0\right] \quad 3\text{-}30$$

$$L_\beta = \sum_{i=1}^{n}\left(n\left[\lambda'\left(Z'_{[i,.]}\left(Y_i - X_{[i,.]}\beta\right)\right) + 1\right]\right)^{-1}\left(-\lambda' Z'_{[i,.]} X_{[i,.]}\right) = 0 \quad 3\text{-}31$$

In the estimation of the MEL problem, the equation "**3-30**" and "**3-31**" can be solved simultaneously by using the Newton-Raphson method to minimize the expression:

$$\Omega = {L_\lambda}^2 + {L_\beta}^2 \quad 3\text{-}32$$

The above-stated expression "**Ω**" in equation "**3-32**" is numerically optimized (minimize the squared Euclidean norm for necessary conditions) by using the equations

"3-30" and "3-31". It can be solved by sequential procedure, where first "3-30" is solved for "$\boldsymbol{\lambda}$" for a given value of "$\boldsymbol{\beta}$" (initial guess "0" as starting point of $\boldsymbol{\beta}$) and then solve "3-31" for "β" at given the previous value of "$\boldsymbol{\lambda}$" to minimize the square Euclidean norm "3-32". We continue the process until the convergence achieved. The computation procedure of MEEL is the same as the computation of MEL in "3-32".

3.4 Cressie and Read (CR) Family of Divergence

The Cressie and Read family of divergence measure and maximum empirical likelihood as:

$$\mathbf{I}(\mathbf{p},\mathbf{q},\boldsymbol{\gamma}) = \frac{1}{\gamma(\gamma+1)} \sum_{i=1}^{n} p_i[(\frac{p_i}{q_i})^{\gamma} - 1] \qquad 3\text{-}33$$

In the CR criterion function, the "γ" is a parameter that indexes members of the CR family to measure the discrepancy between two probability distributions such as the subjective distribution "p" and reference distribution "q." The Lagrange function is as:

$$\ln\left(\mathbf{I}(\boldsymbol{\beta}; \mathbf{Y}_i, \mathbf{X}_{[i,.]}, \mathbf{Z}_{[i,.]})\right) = \{\frac{1}{\gamma(\gamma+1)} \sum_{i=1}^{n} p_i[(\frac{p_i}{q_i})^{\gamma} - 1]\} - \mu(\sum_{i=1}^{n}(p_i) - 1) -$$

$$\lambda' \sum_{i=1}^{n} p_i \left(\mathbf{Z}'_{[i,.]}(\mathbf{Y}_i - \mathbf{X}_{[i,.]}\boldsymbol{\beta})\right) \qquad 3\text{-}34$$

The above-discussed CR criterion function converges to the different entropy measures by varying the parameter $'\boldsymbol{\gamma}'$. As CR ($\boldsymbol{\gamma} = -1$) leads to the empirical log-likelihood, (MEL) objective $n^{-1} \sum_{i=1}^{n} \ln(p_i)$. The Specification CR ($\boldsymbol{\gamma} = 0$) leads to the empirical exponential likelihood (MEEL) which is equal to "$-\sum p_i \ln(p_i)$". CR ($\boldsymbol{\gamma} = 1$) define the log Euclidean of least squares likelihood function (MLEL) or least squares empirical likelihood stated in (Judge & Mittelhammer, 2011). Therefore, the CR family of divergence encompasses all entropy measures (non-parametric approaches). The MEL and MEEL approaches are among of them.

3.4.1 Optimal Choice and Quadratic Risk (QR) Function

We follow an approach proposed by Judge and Mittelhammer (2011) about the optimal convex combination "$\overline{\boldsymbol{\beta}}(\boldsymbol{\alpha})$" of parameters of different entropy measures "$\widehat{\boldsymbol{\beta}}(\boldsymbol{\gamma}) = \mathrm{I}(\mathbf{p}, \mathbf{q}, \boldsymbol{\gamma})$" from the CR family of test statistics to estimate unknown parameter. The optimal convex combination is defined on the basis of the minimum Quadratic Risk (QR) function. The following equation demonstrates this CR Optimal Convex Combination (CROCC), as proposed by (Judge & Mittelhammer,2011).

$$\overline{\boldsymbol{\beta}}(\boldsymbol{\alpha}) = \sum_{j=1}^{J} \alpha_j \widehat{\boldsymbol{\beta}}\left(\gamma_j\right), \text{ where } \alpha_j \geq 0 \forall_j, \text{ and } \sum_{j=1}^{J} \alpha_j = 1 \qquad 3\text{-}35$$

Here, each estimator "$\widehat{\boldsymbol{\beta}}(\boldsymbol{\gamma})$" is obtained from the CR family of test statistics by varying the coefficient "$\boldsymbol{\gamma}$" as shown in the following equation.

$$\widehat{\boldsymbol{\beta}}(\boldsymbol{\gamma}) = \mathbf{I}(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{\gamma}) = \frac{1}{\gamma(\gamma+1)} \sum \boldsymbol{p}_i \left[\left(\frac{p_i}{q_i}\right)^{\gamma} - 1\right] \qquad 3\text{-}36$$

The above equation is optimized subject to the following constraints.

$$\sum_{i=1}^{n} \boldsymbol{p}_i = 1 \text{ and } \sum_{i=1}^{n} \boldsymbol{p}_i \left(\mathbf{Z'}_{[i,.]}\left(\mathbf{Y}_i - \mathbf{X}_{[i,.]}\boldsymbol{\beta}\right)\right) = 0 \qquad 3\text{-}37$$

Following research objectives, we use two alternatives of CR family of test statistics, i.e., MEL and MEEL represented with $\widehat{\boldsymbol{\beta}}(\gamma 1)$ and $\widehat{\boldsymbol{\beta}}(\gamma 2)$ respectively. The following section explains the optimal choice of these two individual CR information-theoretic estimators using minimum QR function.

The Optimal choice of Two CR Alternatives

The following equation provides the formula of an optimal convex combination of MEL and MEEL, as proposed by (Judge & Mittelhammer, 2011).

$$\bar{\boldsymbol{\beta}}(\hat{\boldsymbol{\alpha}}) = \hat{\boldsymbol{\alpha}}\,\hat{\boldsymbol{\beta}}(\gamma_1) + (1 - \hat{\boldsymbol{\alpha}})\,\hat{\boldsymbol{\beta}}(\gamma_2) \qquad 3\text{-}38$$

In the above equation, "$\hat{\alpha}$" is an estimated coefficient of convex combination and yields as

$$\hat{\boldsymbol{\alpha}} = \frac{\rho(\hat{\beta}(\gamma_2),\ \beta) - E[(\hat{\beta}(\gamma_1) - \beta)'(\hat{\beta}(\gamma_2) - \beta)]}{\rho(\hat{\beta}(\gamma_1),\ \beta) + \rho(\hat{\beta}(\gamma_2),\ \beta) - 2\,E[(\hat{\beta}(\gamma_1) - \beta)'(\hat{\beta}(\gamma_2) - \beta)]} \qquad 3\text{-}39$$

In the equation "**3-39**" the expression "$E[(\hat{\boldsymbol{\beta}}(\gamma_1) - \boldsymbol{\beta})'(\hat{\boldsymbol{\beta}}(\gamma_2) - \boldsymbol{\beta})$" is a cross product of bias terms of $\hat{\boldsymbol{\beta}}(\gamma_1)$ and $\hat{\boldsymbol{\beta}}(\gamma_2)$. This expression is a non-zero term (Judge & Mittelhammer, 2011). The authors contended that "assuming the moment conditions are correctly specified, the $\hat{\boldsymbol{\beta}}(\gamma)$ estimators are consistent under regularity conditions no more stringent than the usual conditions imposed to obtain consistency in the generalized methods of moments (GMM) or the classical linear model". Under this assumption, "$E[(\hat{\boldsymbol{\beta}}(\gamma_1) - \boldsymbol{\beta})'(\hat{\boldsymbol{\beta}}(\gamma_2) - \boldsymbol{\beta})$" approaches to zero as "n" increases. Judge and Mittelhammer (2011) also assumed that the estimators $\hat{\boldsymbol{\beta}}(\gamma_1)$ and $\hat{\boldsymbol{\beta}}(\gamma_2)$ obtained from two entropy measures are based on independent samples of data, therefore the term "$E[(\hat{\boldsymbol{\beta}}(\gamma_1) - \boldsymbol{\beta})'(\hat{\boldsymbol{\beta}}(\gamma_2) - \boldsymbol{\beta})]$" is zero.

As a result, the equation "**3-39**" will be $\widehat{\alpha} = \frac{\rho(\widehat{\beta}(\gamma_2),\ \beta)}{\rho(\widehat{\beta}(\gamma_1),\ \beta)+\rho(\widehat{\beta}(\gamma_2),\ \beta)}$ and can be written as follow;

$$\widehat{\boldsymbol{\alpha}} = \frac{\mathrm{tr}\,(\mathrm{cov}\,(\widehat{\beta}(\gamma_1))}{\mathrm{tr}\,(\mathrm{cov}\left(\widehat{\beta}(\gamma_1)\right) + \mathrm{tr}\,(\mathrm{cov}\,(\widehat{\beta}(\gamma_2))} \qquad 3\text{-}40$$

Where

$$\mathrm{cov}\left(\widehat{\beta}(\gamma)\right) =$$

$$\left(\left[\sum_{i=1}^{n} \widehat{p}_i(\gamma) \frac{\partial h\left(Y_i, X_{[i,.]}, Z_{[i,.]}, \beta(\gamma)\right)}{\partial \beta(\gamma)}\right] \left[\sum_{i=1}^{n} \widehat{p}_i(\gamma)\, \mathbf{h}(\mathbf{Y}_i, \mathbf{X}_{[i,.]}, \mathbf{Z}_{[i,.]}, \boldsymbol{\beta})\, \mathbf{h}(\mathbf{Y}_i, \mathbf{X}_{[i,.]}, \mathbf{Z}_{[i,.]}, \boldsymbol{\beta})'\right]^{-1} \left[\sum_{i=1}^{n} \widehat{p}_i(\gamma) \frac{\partial h\left(Y_i, X_{[i,.]}, Z_{[i,.]}, \beta(\gamma)\right)}{\partial \beta(\gamma)}\right]' \right)^{-1} \qquad 3\text{-}41$$

Where $\widehat{\boldsymbol{p}}_i(\boldsymbol{\gamma})$s are estimated probability weights obtain from the solution to the estimation of $(\boldsymbol{\gamma} = -1)$ and $(\boldsymbol{\gamma} = 0)$. After estimating the "$\widehat{\boldsymbol{\alpha}}$" CROCC estimator will be as follow:

$$\overline{\boldsymbol{\beta}}(\widehat{\boldsymbol{\alpha}}) = \widehat{\boldsymbol{\alpha}}\,\widehat{\boldsymbol{\beta}}\,(\mathbf{MEL}) + (\mathbf{1} - \widehat{\boldsymbol{\alpha}})\,\widehat{\boldsymbol{\beta}}\,(\mathbf{MEEL}) \qquad 3\text{-}42$$

However, if the expression "$\mathrm{E}[(\widehat{\boldsymbol{\beta}}(\boldsymbol{\gamma}_1) - \boldsymbol{\beta})'(\widehat{\boldsymbol{\beta}}(\boldsymbol{\gamma}_2) - \boldsymbol{\beta})$" is a non-zero term, then according to Judge and Mittelhammer (2011), the aforementioned expression needs to operationalize appropriately. Therefore, we introduce an alternative approach as an arbitrary method (Cressie and Read Arbitrary Convex Combination (CRACC)) to compute the coefficient of convex combination that is independent of the computation of the coefficient of convex combination. "$\widehat{\boldsymbol{\alpha}}$". The details of CRACC are given in the following section.

3.4.2 CR Arbitrary Convex Combinations (CRACC)

Judge and Mittelhammer (2011) found that their estimated average value of "$\hat{\alpha}$" is in the neighbourhood of "0.5". Therefore, we took the arbitrary values of "$\alpha = 0.5$". Using the arbitrary value (denoted as "a"), the convex combination will be as follow

$$\bar{\beta}(a) = a\,\hat{\beta}(\gamma_1 = -1) + (1 - a)\,\hat{\beta}(\gamma_2 = 0) \qquad 3\text{-}43$$

Particularly, for the arbitrary value of a = 0.5 above equation will become

$$\bar{\beta}(a = 0.5) = 0.5\,\hat{\beta}(\mathrm{MEL}) + (1 - 0.5)\,\hat{\beta}(\mathrm{MEEL}) \qquad 3\text{-}44$$

The arbitrary convex combination is free from the assumptions required to estimate the optimal convex combination proposed by (Judge & Mittelhammer, 2011). This research also compared the performances of CRACC and CROCC to explore the best estimator. We also introduced two other arbitrary combinations on a scale of 25% and 75% to evaluate the CRACC. When the arbitrary convex combination with $CRACC_i$ is donated where i= 1, 2, and 3 representing 0.25, 0.5, and 0.75, respectively.

3.5 Overdetermined Model and Linear Transformation

In the case, if the number "ℓ" of moment equations becomes greater than the number "k" of unknown parameters to be estimated ($k < \ell$), the system of equation (moment equations) has full row rank, in this case, the system of moment equation is called "overdetermined" system.

$$E(h(Y_i, X_{[i,.]}, Z'_{[i,.]}, \beta)) = \sum_{i=1}^{n} p_i\, h(Y_i, X_{[i,.]}, Z'_{[i,.]}, \beta) = 0 \qquad 3\text{-}45$$

Where, $Y_{(n\times 1)}, X_{(n\times k)}, Z_{(n\times \ell)}, \beta_{(n\times k)}$ and ($k < \ell$) show the overdetermined model of moment equations. It is generally impossible to solve for k parameters because the system will estimate "ℓ" parameters while "$\ell - k$" moment equations are redundant.

Therefore, moment conditions will estimate the "ℓ" for "k" parameters. The functional dependence transforms the overdetermined model into the exactly determined system by reducing "ℓ" moment equations to "k" dimensions. This conversion can be done by pre-multiplying the "$\ell \times 1$" moment equations by a "$k \times \ell$" matrix denoted by "v."

$$\mathbf{E}(\mathbf{h}_v(\mathbf{Y}_i, \mathbf{X}_{[i,.]}, \mathbf{Z}'_{[i,.]}, \boldsymbol{\beta})) = \sum_{i=1}^{n} p_i[v * \mathbf{h}(\mathbf{Y}_i, \mathbf{X}_{[i,.]}, \mathbf{Z}'_{[i,.]}, \boldsymbol{\beta})] = \mathbf{0} \qquad \text{3-46}$$

The above stated "3-45" moment equations are successfully transformed to an exactly determined system that has the same estimates got from the moment conditions "3-46".The idea of linear transformation was initially proposed by (McCullagh & Nelder, 1989). These authors exhibited the optimal selection of "v" that have minimum covariance asymptotically. A number of studies, applied this idea of linear transformation to estimate the overdetermined model by MEL approach (Mittelhammer et al., 2000; Qin & Lawless, 1994). The overdetermined model transformed into just a determined model that can estimate the unique values of β having 'k' parameters. The 'ℓ' estimating moment conditions are as follows:

$$\mathbf{E}(\mathbf{h}(\mathbf{Y}_i, \mathbf{X}_{[i,.]}, \mathbf{Z}'_{[i,.]}, \boldsymbol{\beta})) = \sum_{i=1}^{n} p_i\,[\mathbf{Z}'_{[i,.]}(\mathbf{Y} - \mathbf{X}_{[i,.]}\boldsymbol{\beta})] = \mathbf{0} \qquad \text{3-47}$$

The transformation provides us the k estimates.

$$\mathbf{E}(\mathbf{h}_v(\mathbf{Y}_i, \mathbf{X}_{[i,.]}, \mathbf{Z}'_{[i,.]}, \boldsymbol{\beta})) = \sum_{i=1}^{n} p_i * v * [\mathbf{Z}'_{[i,.]}(\mathbf{Y} - \mathbf{X}_{[i,.]}\boldsymbol{\beta})] = \mathbf{0} \qquad \text{3-48}$$

Where 'v' is ($k \times \ell$) full row rank matrix as follows:

$$v = E\left[\frac{\partial \mathbf{h}(\mathbf{Y}_i, \mathbf{X}[i,.], \mathbf{Z}'[i,.], \beta)}{\partial \beta}\right]\left[cov\left(\mathbf{h}\left(\mathbf{Y}_i, \mathbf{X}_{[i,.]}, \mathbf{Z}'_{[i,.]}, \boldsymbol{\beta}\right)\right)\right]^{-1} \qquad \text{3-49}$$

Where $\left[cov\left(h\left(\mathbf{Y}_i, \mathbf{X}_{[i,.]}, \mathbf{Z}'_{[i,.]}, \boldsymbol{\beta}\right)\right)\right] = E[\left(h\left(\mathbf{Y}_i, \mathbf{X}_{[i,.]}, \mathbf{Z}'_{[i,.]}, \boldsymbol{\beta}\right)\right) * (h\left(\mathbf{Y}_i, \mathbf{X}_{[i,.]}, \mathbf{Z}'_{[i,.]}, \boldsymbol{\beta}\right))']$, numerous studies stated that the covariance of the MEL estimator comes to be more efficient as the number of moment equations increased (Mittelhammer et al., 2000; Qin & Lawless, 1994). Therefore, in the overdetermined model, we can be used the above-discussed set of moment equations $h_V(X, Y, Z, \beta)$ to estimate "k" unknown parameters in the estimating procedure of the MEL and MEEL approach. In the overdetermined model, we use the same sequential base numerical procedure to estimate the unknown parameters by using MEL and MEEL approach. We have used MATLAB software to construct the required algorithm. However, to evaluate these estimators, Monte Carlo Simulation (MCS) analysis is carried out whose description is given below.

3.6 Monte Carlo Simulation Design

The Monte Carlo Simulation (MCS) analysis is used to analyze and compare the finite sample properties of GMM and information-theoretic estimators (MEL, MEEL, and CROCC) in estimating the linear regression model having the endogeneity problem. Consider a linear regression model specified in "**3-1**".

The objective of this study is to estimate the above equation when orthogonality condition between "ϵ" and **X** is not satisfied, i.e., ($E(\epsilon|\mathbf{X}) \neq 0$). In this context, the OLS technique is not applicable to estimate the parameter "β." However, GMM, MEL, MEEL, and CROCC can be used to estimate the above model in the presence of endogeneity. To solve this kind of issue we need instrumental variables that satisfy the orthogonality condition as $E(\epsilon|\mathbf{Z}) = 0$, but have a correlation with endogenous regressor $E(\mathbf{Z}|\mathbf{X}) \neq 0$. Our analysis is based on "5000" Monte Carlo Simulations to evaluate and compare the performance of the estimators.

3.6.1 Data Generating Process (DGP)

The data generating process (DGP) used to evaluate the performance of the above discussed estimators when orthogonality condition do not satisfy, in the estimation of the regression model is defined as follows

$$\mathbf{Y}_{(n\times1)} = \mathbf{X}_{(n\times2)}\boldsymbol{\beta}_{(2\times1)} + \boldsymbol{\epsilon}_{(n\times1)} \qquad 3\text{-}50$$

Where $\boldsymbol{\beta}_{(2\times1)}$are the coefficients, $\boldsymbol{\epsilon}_{(n\times1)}$are errors of "**3-50**" and

$$\mathbf{X}_{(n\times2)} = [\boldsymbol{\gamma}\ \mathbf{x}] \qquad 3\text{-}51$$

In "**3-51**" $X_{(n\times2)}$ are the regressors consist of $\boldsymbol{\gamma} = \begin{bmatrix}1\\1\\\vdots\\1\end{bmatrix}_{(n\times1)}$ and $\mathbf{x}_{(n\times1)}$.

$$\mathbf{x}_{(n\times1)} = \mathbf{Z}_{(n\times4)}\boldsymbol{\delta}_{(4\times1)} + \boldsymbol{\upsilon}_{(n\times1)} \qquad 3\text{-}52$$

Where $\mathbf{Z}_{(n\times4)} = [Z_1\ Z_2\ Z_3\ Z_4]$ is a matrix of instrumental variables and $\boldsymbol{\delta}'_{(4\times1)} = [\delta_1 \quad \delta_2 \quad \delta_3\ \delta_4]$ are their respective coefficients and $\boldsymbol{\upsilon}_{(n\times1)}$ are residuals of "**3-52**". The $\boldsymbol{\epsilon}_{(n\times1)}$ and $\boldsymbol{\upsilon}_{(n\times1)}$ are generated through Bivariate Normal distribution as follows:

$$[\boldsymbol{\epsilon}\ \boldsymbol{\upsilon}]_{(n\times2)} \sim \mathbf{N}\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}1 & \tau\\ \tau & 1\end{bmatrix}\right) \qquad 3\text{-}53$$

Where "$\boldsymbol{\tau}$" will measure the degree of endogeneity.

First of all, $\boldsymbol{\epsilon}_{n\times1}$ and $\boldsymbol{\upsilon}_{n\times1}$are generated through "**3-53**", then four instrumental variables Z_1 Z_2 Z_3 and Z_4 each of $(n \times 1)$ order are generated. By using the fixed

$\boldsymbol{\delta}_{(4\times1)}$, $\mathbf{Z}_{(n\times4)}$ and $\boldsymbol{\upsilon}_{(n\times1)}$ that are the components of regressor i.e. $\mathrm{x}_{(n\times1)}$ are generated through "**3-52**". Then matrix of regressors i.e. $\mathbf{X}_{(\mathrm{n}\times2)}$ is generated according to "**3-51**". Finally $\mathbf{Y}_{(\mathrm{n}\times1)}$is generated according to "**3-50**" by using the matrix $\mathbf{X}_{(\mathrm{n}\times2)}$, $\boldsymbol{\beta}_{(2\times1)}$(a constant coefficients vector) and $\boldsymbol{\epsilon}_{(\mathrm{n}\times1)}$ vector of errors.

We have generated a matrix of instrumental variables that are denoted as $\mathbf{Z} = [\mathrm{Z}_1\ \mathrm{Z}_2\ \mathrm{Z}_3\ \mathrm{Z}_4]$. The instrumental variables are iid from a multivariate normal distribution, with a zero mean vector and some standard deviation. The outcomes of ϵ and $\mathbf{Z}$ will be generated independently so that $\mathrm{E}(\epsilon|\mathbf{Z}) = 0$. This fulfills the fundamental condition for $\mathbf{Z}$ to be considered as a valid instrumental variable. In equation "**3-52**" the values of $'\boldsymbol{\delta}'$ measures the strength of instrumental variables that will be determined by regressing $\mathbf{x}_{(n\times1)}$ on $\mathbf{Z}_{(n\times4)}$. Theoretically, $\mathrm{IVS} = \frac{\boldsymbol{\delta}'\boldsymbol{\delta}}{\boldsymbol{\delta}'\boldsymbol{\delta}+1}$ will determine the Strength of Instrumental Variables abbreviated as (IVS) (Judge & Mittelhammer, 2011).

Regarding the sampling scenarios of the Monte Carlo simulation experiments, sample sizes of n = 25, 50, 100, 150, 200, and 250 are used to analyze the estimators in accordance with study objectives. In all the sample sizes, we have estimated the parameter of interest $\boldsymbol{\beta}$ with all aforementioned estimation techniques by varying the values of $\boldsymbol{\tau}$ and $\boldsymbol{\delta}$. In order to analyze and compare the performances of the estimators, different combinations of sample size (n), degrees of endogeneity[6] ($\boldsymbol{\tau}$ = 0.2, 0.4, 0.6 & 0.8 with low, moderate, high and very high respectively) and strength of instrumental

---

[6]There is a linkage between endogeneity and instrumental varibles, whenever endogeneoty exist between regressors and error term we have needed IV's to resolve this issue. Various Authors used Instrumental variables to resolve the problem of endogeneity (Altonji & Segal, 1996; Hall & Horowitz, 1996; Imbens et al., 1998; Imbens, 1997; Kitamura & Stutzer, 1997; Qin & Lawless, 1994).

variables[7] ($\boldsymbol{\delta}$ = 0.25, 0.5 & 0.75 with weak, moderate and strong strength respectively) are also constructed. In case of the exactly determined model, $\delta_1 \neq 0$ *and* $\delta_2 = \delta_3 = \delta_4 = 0$, on the other hand, $\delta_1 = \delta_2 = \delta_3 = \delta_4 \neq 0$ when the model is overdetermined. The listed combination is mentioned in the table, which is as follows:

Table 3-1: Degree of Endogeneity and Instrumental Variables Strength

		Degree of Endogeneity (D.E)			
		0.2	**0.4**	**0.6**	**0.8**
Instrumental	**0.25**	0.25/0.2	0.25/0.4	0.25/0.6	0.25/0.8
Variables	**0.5**	0.5/0.2	0.5/0.4	0.5/0.6	0.5/0.8
Strength (IVS)	**0.75**	0.75/0.2	0.75/0.4	0.75/0.6	0.75/0.8

The Monte Carlo simulations on these combinations are conducted for each sample size.

3.7 Empirical Analysis

For empirical analysis, we used the real time series to analyze the performance of the information-theoretic approach to estimate the regression model in the presence of endogeneity problem. Money demand function and consumption function are two economic models that often do not satisfy the condition of orthogonality and cause simultaneous bias. We evaluated the performances of selected estimators for these two economic models. These economic models are selected because of the following reasons;

[7] Various authors have discussed method to find the strength of IV's (Camron et al 2002; Stock and Yogo 2002; Baltagai 2007). For instance, Camron et al (2002) contended that R^2 and F test can be used to find the strength of IV's.

- These models have an endogeneity issue.
- We face small samples issue in such economic models.
- No application of information-theoretic models is found to estimate the economic models.

The details of both models are provided in the following sections.

3.7.1 Money Demand Function

Keynes (1936) first proposed the idea of the money demand function, also known as liquidity preference theory, in contrast to the quantity theory of money (QTM). According to his theory, money supply and money demand depend upon the rate of interest and permanent income. The higher the rate of interest, will lower the money demand, and the higher the income will increase the money demand. Later on, many theorists added different concepts (Friedman, 1956; Tobin, 1958) and different estimating techniques Feige, (1967) regarding money demand functions.

Various studies found the problem of simultaneity bias in the estimation of the money demand function. For instance, Hsing and Jamal, (2013) estimated the money demand function using 3SLS method to solve such an endogeneity problem. Similarly, Thomas (1993) discussed the money demand functions and stated the problem of simultaneity bias in its estimation. The money demand and money supply functions and the problem of simultaneity bias are discussed below.

Money Demand function: $$\mathbf{m_d = b_o + b_1 r + b_2 y + \varepsilon_1} \qquad 3\text{-}54$$

Money supply function: $$\mathbf{m_s = a_o + a_1 r + h + \varepsilon_2} \qquad 3\text{-}55$$

For equilibrium condition, the money demand equal to the money supply

$$m_d = m_s \quad 3\text{-}56$$

Where 'h' is the stock of high-power money, 'r' is the rate of interest, and 'Y' is permanent income. The level of income and high powered money stock is assumed exogenous variables which are determined by the monetary authorities. Income appears in demand function, while high power money stock is used in supply function. However, interest is included in both functions. The consequences of simultaneity must be considered as shown by the reduced form below:

$$r = \frac{a_o - b_o}{b_1 - a_1} - \left(\frac{b_2}{b_1 - a_1}\right) y + \left(\frac{1}{b_1 - a_1}\right) h + \frac{\varepsilon_2 - \varepsilon_1}{b_1 - a_1} \quad 3\text{-}57$$

$$m = \frac{b_1 a_o - b_o a_1}{b_1 - a_1} - \left(\frac{a_1 b_2}{b_1 - a_1}\right) y + \left(\frac{b_1}{b_1 - a_1}\right) h + \frac{b_1 \varepsilon_2 - a_1 \varepsilon_1}{b_1 - a_1} \quad 3\text{-}58$$

Since, $m_s = m_d$, therefore reduced form for m_s is identical to m_d and the rate of interest is positively correlated with ε_1. This indicates that the money demand function cannot be estimated directly due to the simultaneity problem where 'r' is an endogenous variable. Money and rate of interest are endogenous variables that can be predicted by permanent income and high-power money stock. The correlation of the rate of interest with errors indicates the presence of endogeneity that needs to be solved to produce unbiased estimates. The present study intends to provide such a solution to the endogeneity problem in finite samples to estimate the money demand function.

To resolve this issue, we require instrumental variables that can be used to estimate the money demand function. The equations "**3-57**" and "**3-58**" show that permanent income and high-powered money stock variables satisfy the condition of instrumental variables. Therefore, they can be used to estimate the money demand function. A number of estimation techniques are available to estimate money demand function by using instrumental variables in this respect. However, we will apply GMM and information-theoretic estimators, as discussed earlier, to evaluate their performances in finite samples for a real economic problem. We collected annual data[8] from International Financial Statistics (IFS) and the World Bank Data Bank for several countries to estimate money demand function with a sample of 30 observations for each country.

3.7.2 Consumption Functions

The simultaneity bias can also exist in structural models such as Keynesian consumption function, which is also known as the absolute income hypothesis (Keynes, 1936). Various authors reported the endogeneity problem in estimating the Keynesian consumption function (Charemza & Deadman, 1997; Thomas, 1993). It is a two-equation model, as follow:

$$\mathbf{C_t = \mathit{a} + bY_t + \varepsilon_t} \tag{3-59}$$

$$\mathbf{Y_t = C_t + I_t} \tag{3-60}$$

Where C_t: aggregate consumption, Y_t: aggregate income and I_t: investment. Equation '3-59' shows that the effect of an unexpected change in C_t is due to change in ε_t. Both equations are showing an increase in ε_t increases the C_t while the increase in C_t also increases in Y_t. Therefore, C_t and Y_t are interdependent showing no change in C_t without the change in Y_t and vice versa. However, C_t may change with the change of I_t. Hence, both C_t and Y_t are endogenous variables in the above system. The

[8] For empirical analysis, we have collected time series data from International Financial Statistics (IFS) and the World Bank Data Bank for several countries with a sample of 30 observations for each country. The time spane was from 1972 to 2009.

endogeneity issue can be explained using equations "**3-59**" and "**3-60**" with respect to the aforementioned system.

$$C_t = \frac{a}{1-b} + \frac{b}{1-b} I_t + \frac{\varepsilon_t}{1-b} \qquad 3\text{-}61$$

$$Y_t = \frac{a}{1-b} + \frac{1}{1-b} I_t + \frac{\varepsilon_t}{1-b} \qquad 3\text{-}62$$

These equations have shown that consumption "C" and income "Y' are correlated with disturbance term "ε_t". We could not estimate the consumption function "**3-59**" directly by the OLS technique because it will produce biased estimates. Therefore, to estimate the consumption function, instrumental variables are needed to solve the model. In this problem, we used investment and government expenditures as instrumental variables. We intend to apply GMM and information-theoretic approaches to evaluate their performances for the above real economic problem having the endogeneity issue.

CHAPTER 4. RESULTS & DISCUSSIONS

This chapter provides the results of simulation analysis for both exactly determined and over-determined models, as described in the methodology chapter. A comparative analysis of different estimators is also part of this section.

4.1 Monte Carlo Simulation Analysis

This section will present the results of the Monte Carlo Simulation (MCS) in the presence of endogeneity. The simulation analysis is based on four estimators, including GMM, MEL, MEEL, and CROCC. Here CROCC is the optimal convex combination of two individual entropy measures from CR family. The simulation analysis is applied for both exactly determined and over-determined models. However, for the exactly determined model, MOM (method of the moment) as a special case of the GMM approach is applied. For both the simulation analyses, best and worst cases of the estimator are identified using different cases of finite samples.

Moreover, the performances of estimators are analysed by different combinations of the degree of endogeneity and instrumental variable strengths (IVS) for each finite sample, as shown in Table 4-1. The degree of endogeneity is segregated into a low degree, moderate, high, and very high, while the instrumental variable strengths (IVS) are labeled as weak, moderate, and strong. Subsequent sections will present the results of the entire MCS analysis based on these combinations to identify the best and worst position of the estimators for both exactly determined and over-determined models.

Table 4-1: Endogeneity and Instrumental Variables

		Degree of Endogeneity (D.E)			
		20%	**40%**	**60%**	**80%**
Instrumental variable Strength (IVS)	**25%**	0.25/0.2	0.25/0.4	0.25/0.6	0.25/0.8
	50%	0.5/0.2	0.5/0.4	0.5/0.6	0.5/0.8
	75%	0.75/0.2	0.75/0.4	0.75/0.6	0.75/0.8

4.2 Biasedness of MCS for Exactly Determined Model

The exactly determined model is a model in which the number of moment conditions is exactly equal to the numbers of unknown parameters. This section will present the results of biasedness for such a conditional system. Table 4-2 compares the GMM with information-theoretic estimators to evaluate the properties of the small sample in all the combinations of degrees of endogeneity and IVS. Various graphs for biasedness are constructed using data from Table 4-2 to explore the best and worst estimators.

These graphs are segregated into two broad categories. First, the variations in biasedness are explored for varying degrees of endogeneity and samples against different IVS. Second, the biasedness of estimators is also studied using varying IVS and samples against different degrees of endogeneity. In other words, the superiority and worseness of the estimators will be explored within different combinations of varying endogeneity, IVS, and sample sizes.

Table 4-2: Small Samples Bias in Exactly Determined Model

Samples size	Instrumental variables strength (IVS)	Degree of Endogeneity (D.E)	Estimators			
			MEL	MEEL	GMM	CROCC
25	IVS= 25%	D.E = 20%	-0.06044	-0.09479	0.612496	0.065381
		D.E = 40%	-0.04728	-0.01592	1.581684	0.157847
		D.E = 60%	0.019664	0.071691	0.122376	0.217529
		D.E= 80%	-0.00172	0.082152	0.443509	0.286005
	IV= 50%	D.E= 20%	-0.06172	-0.04714	-0.00458	0.041374
		D.E= 40%	-0.08896	-0.12494	0.080658	0.058197
		D.E= 60%	-0.1666	-0.08056	-0.37491	0.075905
		D.E= 80%	-0.24146	-0.19164	0.275808	0.067181
	IVS= 75%	D.E= 20%	-0.03351	-0.05266	-0.22652	0.003091
		D.E= 40%	-0.04985	-0.07284	-0.07731	0.005751
		D.E= 60%	-0.08503	-0.07629	-0.07718	0.011885
		D.E= 80%	-0.10757	-0.12572	-0.20937	0.01391
50	IVS= 25%	D.E= 20%	-0.03985	-0.10027	-0.64839	0.055253
		D.E= 40%	-0.09515	-0.10962	-0.54027	0.09914
		D.E= 60%	-0.11412	-0.16612	-0.7841	0.11508
		D.E= 80%	-0.17161	-0.11313	-0.54166	0.117639
	IVS= 50%	D.E= 20%	-0.01142	-0.07232	-0.0176	0.019603
		D.E= 40%	-0.04565	-0.11079	0.137229	0.027505
		D.E= 60%	-0.10529	-0.14036	-0.10204	0.033036
		D.E= 80%	-0.15453	-0.17719	-0.2707	0.030993
	IVS= 75%	D.E= 20%	1.04E-05	-0.01814	-0.01965	0.002416
		D.E= 40%	-0.01178	-0.02759	-0.0294	0.003593
		D.E= 60%	-0.01895	-0.04304	-0.04501	0.003925
		D.E= 80%	-0.03121	-0.05475	-0.05673	0.004241
100	IVS= 25%	D.E= 20%	-0.01518	-0.10518	-0.70121	0.044539
		D.E= 40%	-0.04917	-0.15121	-0.07262	0.058569
		D.E= 60%	-0.09478	-0.18375	0.036573	0.06706
		D.E= 80%	-0.17966	-0.26723	-0.185	0.058549
	IVS= 50%	D.E= 20%	0.009644	-0.03601	-0.02078	0.010184
		D.E= 40%	-0.01016	-0.05381	-0.03602	0.012135
		D.E= 60%	-0.03069	-0.07208	-0.05319	0.017219
		D.E= 80%	-0.04581	-0.07507	-0.07041	0.020696
	IVS= 75%	D.E= 20%	0.003919	-0.01344	-0.00481	0.000284
		D.E= 40%	0.005831	-0.01683	-0.00819	0.004384
		D.E= 60%	0.000452	-0.02174	-0.01324	0.004469
		D.E= 80%	-0.00642	-0.02312	-0.01461	0.004554

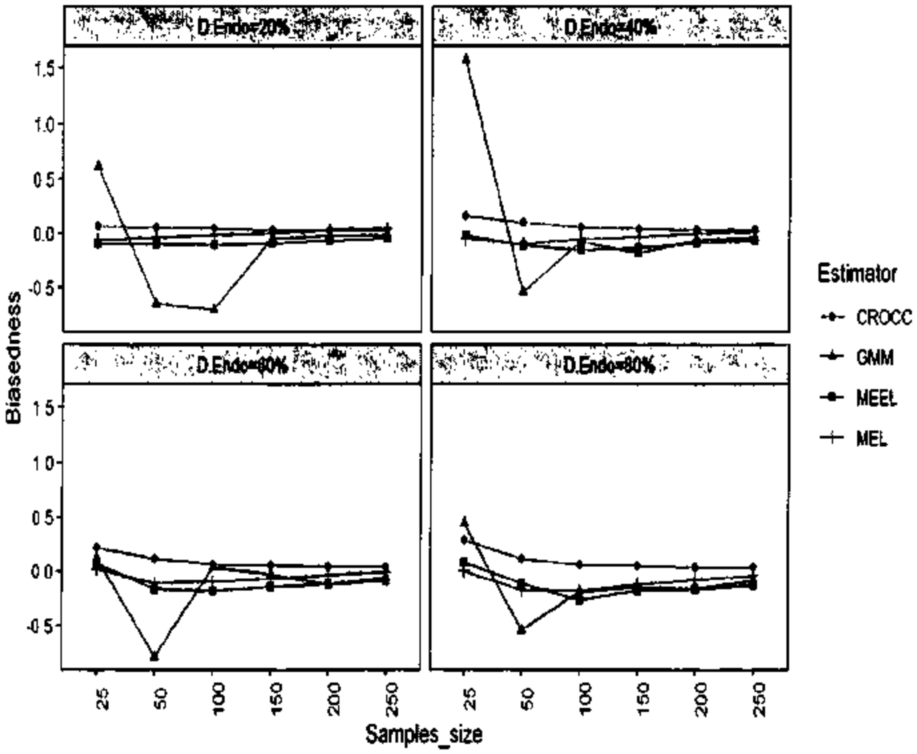
Samples size	Instrumental variables strength (IVS)	Degree of Endogeneity (D.E)	Estimators			
			MEL	MEEL	GMM	CROCC
150	IVS= 25%	D.E= 20%	-0.00104	-0.09149	-0.04498	0.031509
		D.E= 40%	-0.02977	-0.13306	-0.18092	0.043599
		D.E= 60%	-0.06964	-0.141	-0.03186	0.053063
		D.E= 80%	-0.12209	-0.17953	-0.14732	0.051465
	IVS= 50%	D.E= 20%	0.014767	-0.03043	-0.01155	0.005392
		D.E= 40%	0.00799	-0.04077	-0.02258	0.011033
		D.E= 60%	-0.00755	-0.04523	-0.02762	0.012651
		D.E= 80%	-0.01439	-0.05766	-0.041	0.010852
	IVS= 75%	D.E= 20%	0.010732	-0.01072	-0.00204	0.003268
		D.E= 40%	0.006621	-0.01265	-0.00407	0.003663
		D.E= 60%	0.004242	-0.01649	-0.00785	0.003796
		D.E= 80%	-0.0006	-0.01896	-0.01024	0.002618
200	IVS= 25%	D.E= 20%	0.032279	-0.06546	-0.01884	0.029155
		D.E= 40%	-0.00335	-0.0876	-0.05985	0.032546
		D.E= 60%	-0.03482	-0.11832	-0.12412	0.043563
		D.E= 80%	-0.08098	-0.1654	-0.16129	0.036544
	IVS= 50%	D.E= 20%	0.021588	-0.031	-0.01182	0.004395
		D.E= 40%	0.015569	-0.02989	-0.01124	0.011819
		D.E= 60%	0.008452	-0.03834	-0.02038	0.011312
		D.E= 80%	-0.00621	-0.04386	-0.02671	0.009196
	IVS= 75%	D.E= 20%	0.010961	-0.00999	-0.00136	0.002988
		D.E= 40%	0.006719	-0.01097	-0.00253	0.002851
		D.E= 60%	0.004942	-0.01766	-0.009	0.001028
		D.E= 80%	-0.00016	-0.02085	-0.01197	-0.00073
250	IVS= 25%	D.E= 20%	0.045364	-0.04703	-0.016	0.02941
		D.E= 40%	0.020263	-0.06773	-0.03798	0.036814
		D.E= 60%	-0.01101	-0.09078	-0.06093	0.037394
		D.E= 80%	-0.04154	-0.12665	-0.08718	0.040691
	IVS= 50%	D.E= 20%	0.028658	-0.02484	-0.00537	0.008856
		D.E= 40%	0.018586	-0.03148	-0.01274	0.008097
		D.E= 60%	0.011129	-0.03575	-0.01756	0.008201
		D.E= 80%	0.008881	-0.04265	-0.02478	0.009935
	IVS= 75%	D.E= 20%	0.014608	-0.00886	-0.0005	0.004753
		D.E= 40%	0.009359	-0.01353	-0.00506	0.001847
		D.E= 60%	0.007914	-0.0127	-0.00431	0.003311
		D.E= 80%	0.005427	-0.0159	-0.00734	0.002278

4.2.1 Varying Endogeneity and Samples Against Different IVS

Figure 4-1 is providing the results for the biasedness of selected estimators for weak IVS (i.e., 25%) but varying degrees of endogeneity and samples in case of the exactly determined model. Figure 4-1 is depicting that the GMM estimator has high biasedness for small samples (the sample of 25 and 50). These results are consistent for all four selected degrees of endogeneity. This indicates that when the strength of the instrument variable is weak, GMM produces high biasedness in small samples regardless of the degree of endogeneity. It is also notable that GMM is providing more variations in biasedness when the degree of endogeneity and sample size varies.

Conversely, MEL, MEEL, and CROCC are stable estimators across different sample sizes. It is found that MEL and CROCC provided less biasedness as compared to GMM and MEEL. Results also revealed that as the degree of endogeneity increases, all of the estimators provided more biasedness.

Figure 4-1: Biasedness with Weak IVS in Exactly Determined Model



The figure 4-2 evaluates the biasedness of the estimators for moderate (50%) IVS and varied degree of endogeneity and samples. It can be viewed that GMM estimators have larger businesses as compared to other estimators, especially for the sample size of 25 and 50. Moreover, all the estimators except CROCC showed high variations. Overall, the MEL estimator is found superior as compared to MEEL and the GMM, while CROCC outperformed than MEL.

Figure 4-2: Biasedness with Moderate IVS in Exactly Determined Model

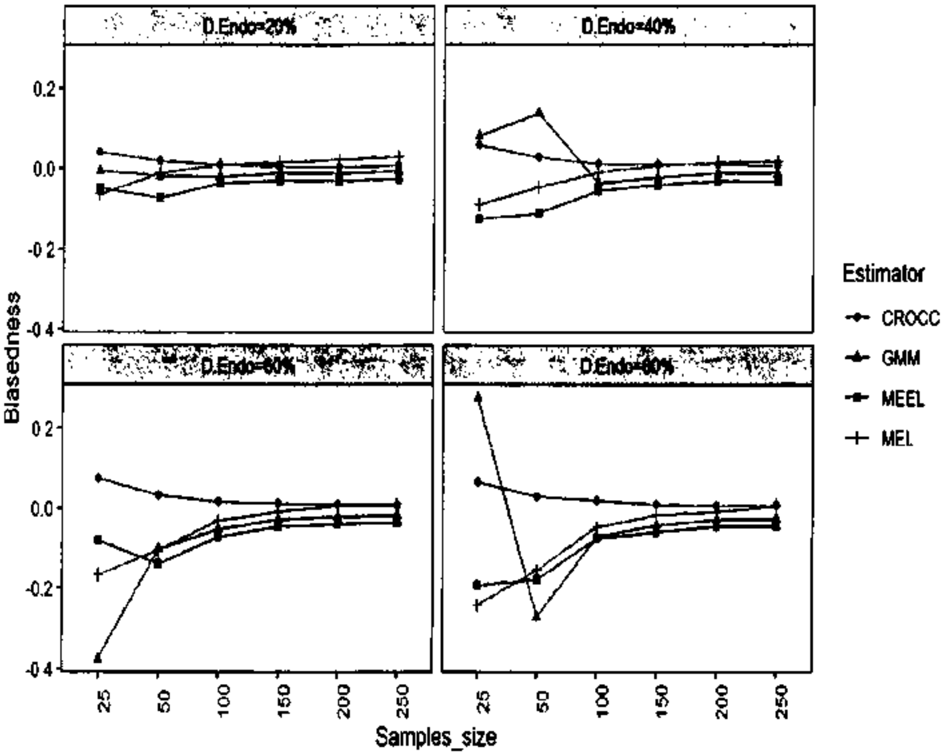
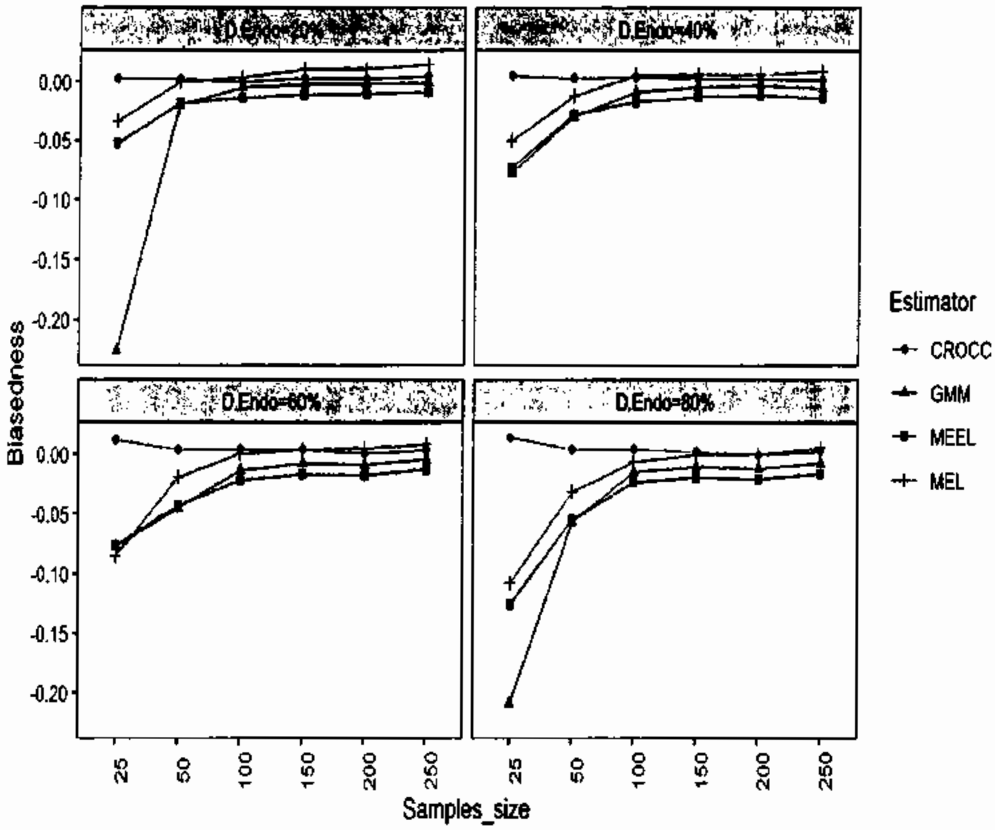


Figure 4-3 further explores the biasedness of estimators for strong IVS and varied the degree of endogeneity as well as sample sizes. All estimators, especially GMM, are showing greater bias in a small sample of 25 and 50. However, as the sample size increases, the biases of all estimators decrease. Results also showed that the CROCC estimator showed less biasedness than other estimators, especially for small samples. Hence, again CROCC estimator is concluded as a superior estimator, while GMM is found worst estimator, especially for a small sample.

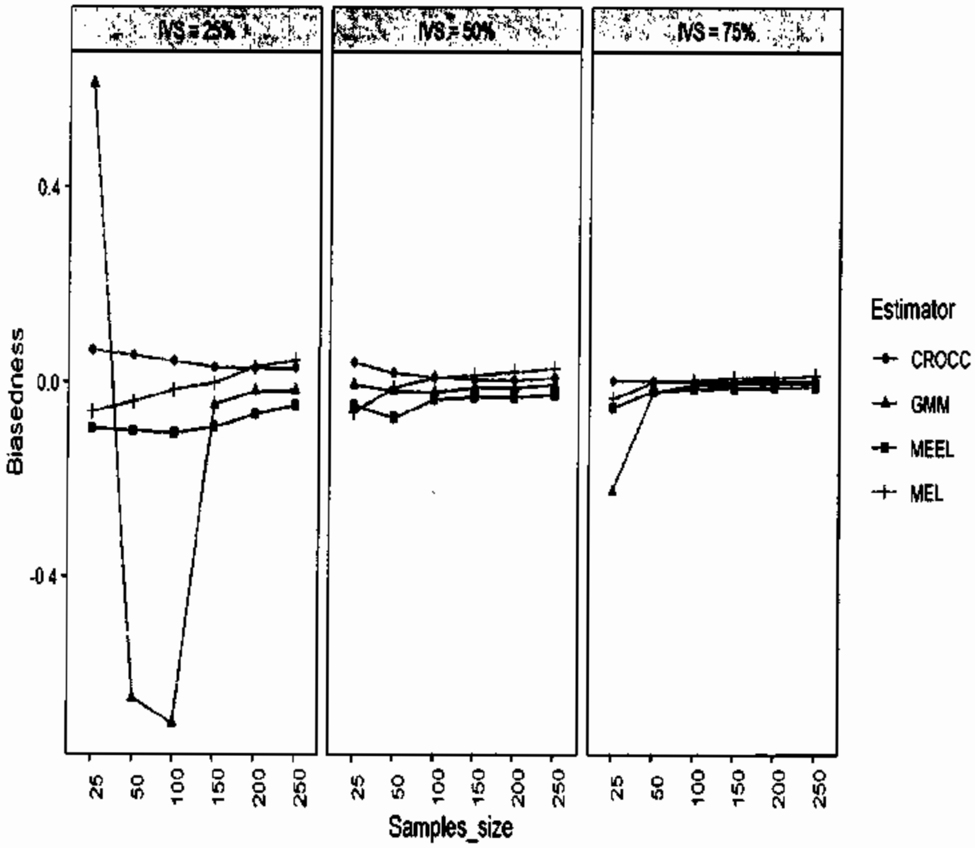
Figure 4-3: Biasedness with Strong IVS in Exactly Determined Model



4.2.2 Varying IVS and Sample Sizes Against Different Degrees of Endogeneity

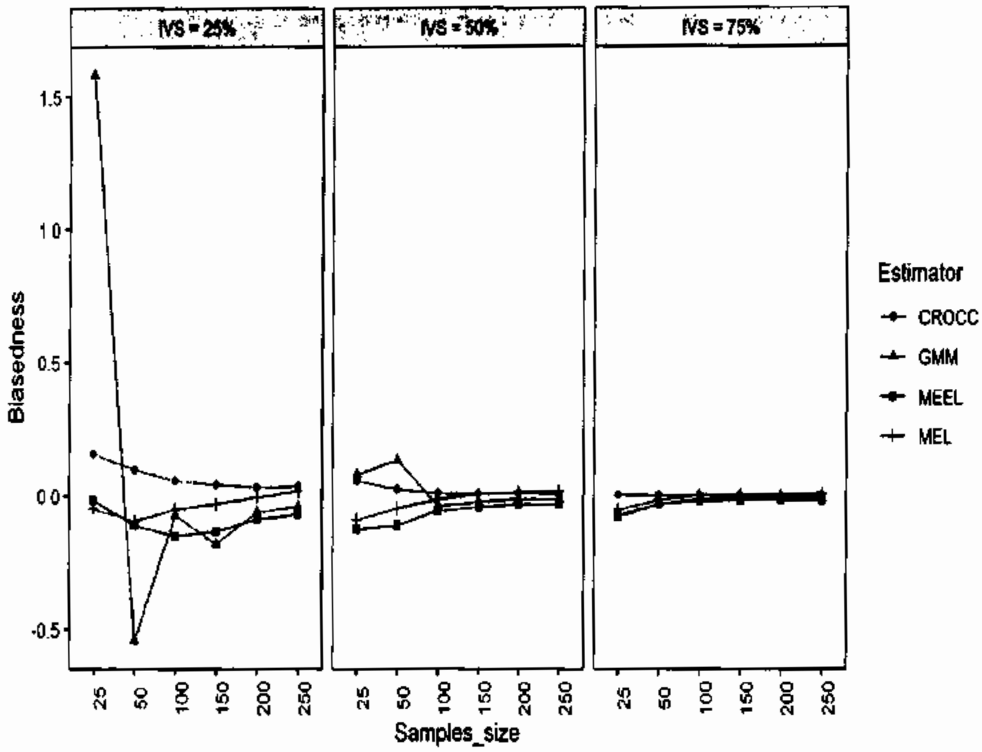
This section will provide the results of biasedness for varying IVS and sample sizes against different degrees of endogeneity in case of an exactly determined model. Figure 4-4 is showing the biasedness of four estimators in case of a low degree of endogeneity (20%) but with varying IVS and sample sizes. Graphs are depicting that the performance of the estimators increased as the strength of the instrumental variables increases from (25% to 75%) with a 25% degree of endogeneity. The biasedness of the GMM estimator is very high at the sample size of 25 to 50. Overall, Figure 4-4 is depicting that CROCC is a superior estimator comparatively in small samples.

Figure 4-4: Biasedness with Low Degree of Endogeneity in Exactly Determined Model



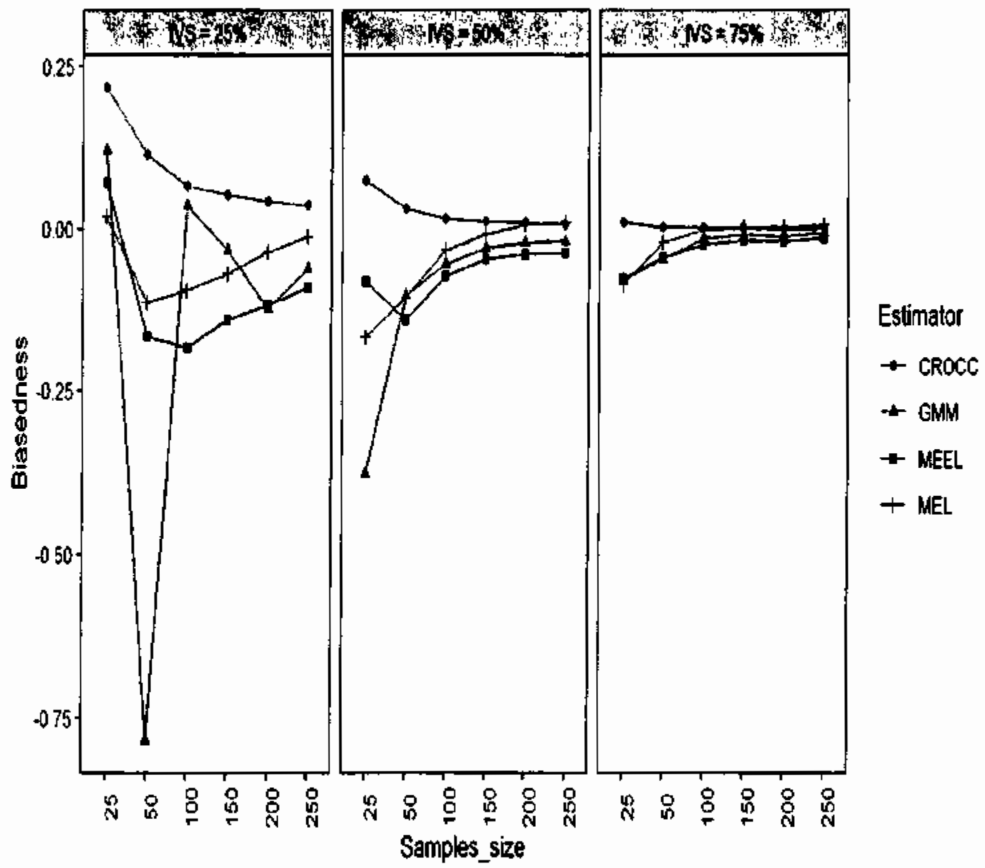
In Figure 4-5, biasedness of the estimators is explored for a moderate degree of endogeneity (40%) and varying strength of instrumental variables and sample size. Results are portraying that GMM has the highest biasedness as compared to all other estimators for sample sizes of 25 and 50. On the other hand, all the estimators have improved their performance as the strength of instrumental variables increase. Overall, among four estimators, CROCC showed the least biasedness, especially for small samples.

Figure 4-5: Biasedness with Moderate Degree of Endogeneity in Exactly Determined Model



Similarly, Figure 4-6 is providing a view of biasedness for varying IVS and samples against a high (60%) degree of endogeneity. It can be viewed that GMM produced the highest biasedness as compared to other estimators when IVS is moderate (50%) or strong (75%), and the sample is small (25). However, CROCC performed best for the same combinations of IVS and samples. Results also revealed that CROCC showed the highest bias when the IVS is weak (25%), and the sample is small (25). For all other combinations, CROCC performed better as compared to other estimators.

Figure 4-6: Biasedness with High Degree of Endogeneity in Exactly Determined Model



At last, Figure 4-7 represents the biasedness of estimators for varying IVS and samples against a very high degree of endogeneity. The graphs are showing that the GMM estimator has the highest biasedness as compared to MEL, MEEL, and CROCC in small samples. Conversely, MEEL has the highest biasedness in a large sample. Overall, the picture is depicting high performances of all the estimators as the IVS is strong. It is also notable that CROCC converges to zero for high IVS. Hence, for a small sample and reduced IVS, CROCC performed better comparatively.

Figure 4-7: Biasedness with Very High Degree of Endogeneity in Exactly Determined Model

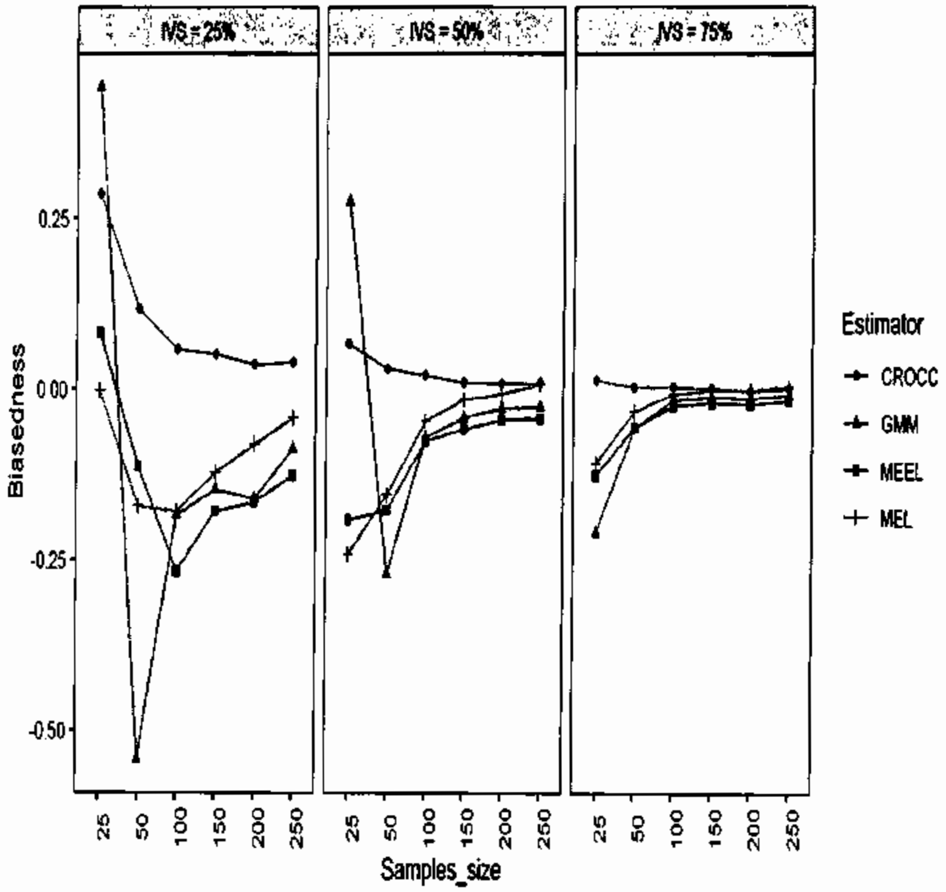
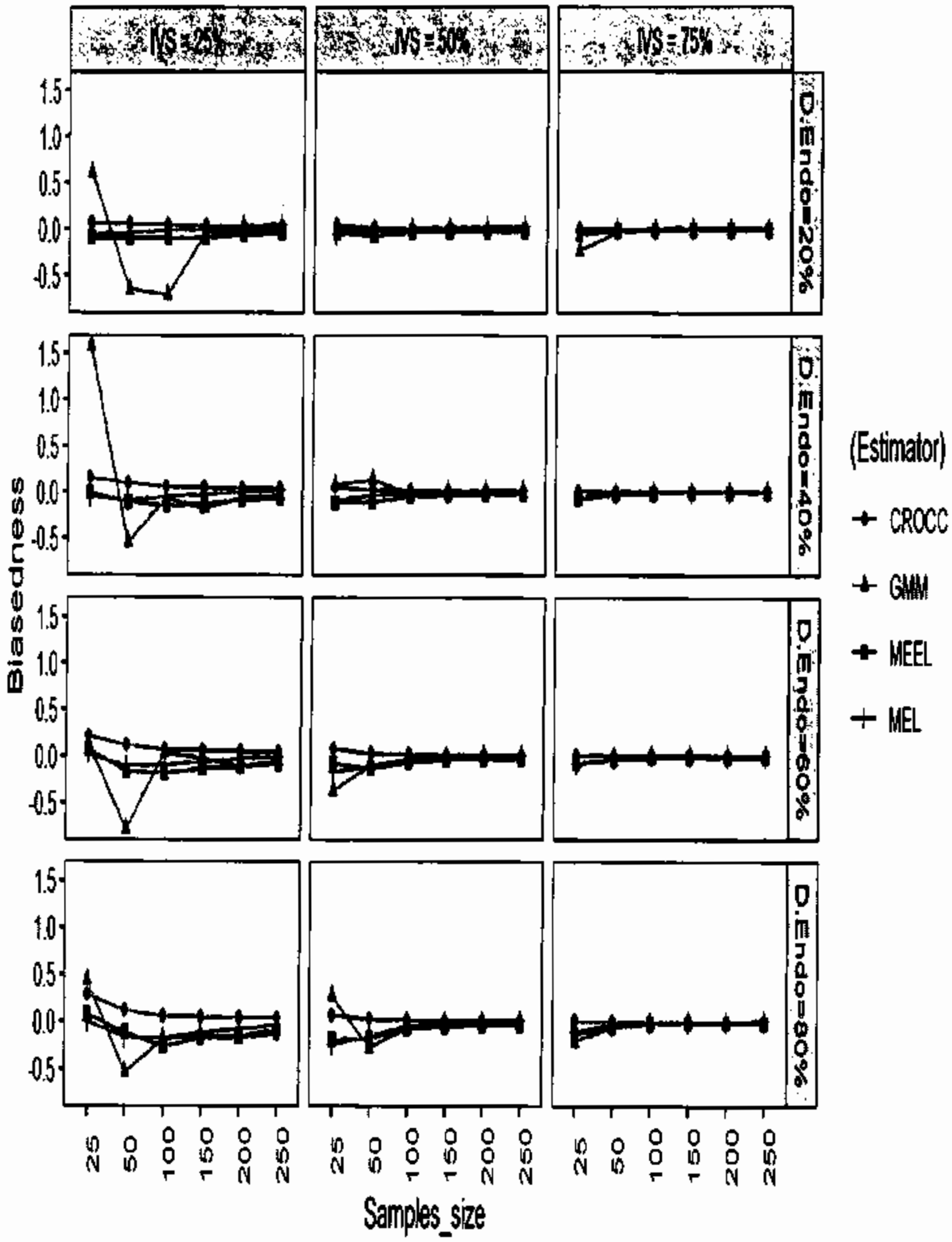


Figure 4-8: Summary of Biasedness in Exactly Determined Model



Summary of Results for Biasedness in Exactly Determined Model

Figure 4-8 is summarising the previous results and explores three useful general implications for all the estimators. First, all the estimators produced less biasedness as the IVS increased. Second, the performances of all the estimators increase as the sample size increases. Third, with the increase in the degree of endogeneity, all the estimators produced more biasedness. Table 4-3 further compares the biasedness of all the estimators for all the possible combinations of the degree of endogeneity, the strength of instrumental variables, and finite samples. Table 4-3 is showing five important implications.

First, the overall analysis showed that CROCC and MEL document the least biasedness. Second, CROCC is found better estimator when the strength of instruments is strong regardless of the degrees of endogeneity and sample sizes. Third, CROCC is also found best estimators for a very high degree of endogeneity irrespective of instrumental variables and sample sizes. Fourth, GMM is found worst estimator as compared to CROCC in all the "72" combinations except for "2" scenarios. Fifth, GMM is found worst estimators for small samples (25& 50) in all settings. It concludes that CROCC is a better estimator, especially than GMM, in estimating a regression model that does not satisfy the condition of orthogonality.

Researchers can also use Table 4-3 in selecting the best estimators. Each cell of Table 4-3 represents specific combinations where researchers can select the best estimator producing the least biasedness. For instance, if a researcher intends to estimate a model having a very high degree of endogeneity, weak strength of instrumental variables and a medium sample of 25 observations then he should follow the results of first-line from row 5 and column 2 ($[\text{MEL} < \text{MEEL} < \text{CROCC} < \text{GMM}]^{25}$).

Table 4-3: Biasedness and Practical Implications for Exactly Determined Model

	Weak Instruments	Moderate Instruments	Strong Instruments
Low Degree	[MEL < CROCC < MEEL < GMM]^{25} [MEL < CROCC < MEEL < GMM]^{50} [MEL < CROCC < MEEL < GMM]^{100} [MEL < CROCC < GMM < MEEL]^{150} [CROCC < MEL < GMM < MEEL]^{200} [CROCC < MEL < GMM < MEEL]^{250}	[CROCC < MEEL < MEL < GMM]^{25} [MEL < CROCC < MEEL < GMM]^{50} [CROCC < MEL < MEEL < GMM]^{100} [CROCC < MEL < GMM < MEEL]^{150} [CROCC < GMM < MEL < MEEL]^{200} [GMM < CROCC < MEL < MEEL]^{250}	[CROCC < MEL < MEEL < GMM]^{25} [CROCC < MEL < MEEL < GMM]^{50} [CROCC < MEL< GMM < MEEL]^{100} [CROCC < MEL< GMM < MEEL]^{150} [CROCC < GMM < MEL < MEEL]^{200} [CROCC < GMM < MEL < MEEL]^{250}
Moderate Degree	[MEEL < MEL < CROCC < GMM]^{25} [CROCC < MEL < MEEL < GMM]^{50} [CROCC < MEL < GMM < MEEL]^{100} [CROCC < MEL < MEEL < GMM]^{150} [CROCC < MEL < GMM < MEEL]^{200} [CROCC < MEL < GMM < MEEL]^{250}	[CROCC < MEEL < MEL < GMM]^{25} [CROCC < MEL < MEEL < GMM]^{50} [MEL < CROCC < GMM < MEEL]^{100} [MEL < CROCC < GMM < MEEL]^{150} [CROCC < MEL < GMM < MEEL]^{200} [CROCC < MEL < GMM < MEEL]^{250}	[CROCC < MEL < MEEL < GMM]^{25} [CROCC < MEL < MEEL < GMM]^{50} [CROCC < MEL < GMM < MEEL]^{100} [CROCC < MEL < GMM < MEEL]^{150} [CROCC < MEL < GMM < MEEL]^{200} [CROCC < MEL < GMM < MEEL]^{250}
High Degree	[MEL < MEEL < GMM < CROCC]^{25} [CROCC < MEL < MEEL < GMM]^{50} [CROCC < GMM < MEL < MEEL]^{100} [CROCC < GMM < MEL < MEEL]^{150} [CROCC < MEL < GMM < MEEL]^{200} [CROCC < MEL < GMM < MEEL]^{250}	[CROCC MEEL < < MEL < GMM]^{25} [CROCC < MEL < MEEL < GMM]^{50} [CROCC < MEL < GMM < MEEL]^{100} [MEL < CROCC < GMM < MEEL]^{150} [MEL < CROCC < GMM < MEEL]^{200} [MEL < CROCC < GMM < MEEL]^{250}	[CROCC < MEL < MEEL < GMM]^{25} [CROCC < MEL < MEEL < GMM]^{50} [MEL < CROCC < GMM < MEEL]^{100} [CROCC < MEL < GMM < MEEL]^{150} [CROCC < MEL < GMM < MEEL]^{200} [CROCC < MEL < GMM < MEEL]^{250}
Very High Degree	[MEL < MEEL < CROCC < GMM]^{25} [CROCC < MEEL < MEL < GMM]^{50} [CROCC < MEL < MEEL < GMM]^{100} [CROCC < MEL < GMM < MEEL]^{150} [CROCC < MEL < GMM < MEEL]^{200} [CROCC < MEL < GMM < MEEL]^{250}	[CROCC < MEEL < MEL < GMM]^{25} [CROCC < MEL < MEEL < GMM]^{50} [CROCC < MEL < GMM < MEEL]^{100} [CROCC < MEL < GMM < MEEL]^{150} [CROCC < MEL < GMM < MEEL]^{200} [CROCC < MEL < GMM < MEEL]^{250}	[CROCC < MEL < MEEL < GMM]^{25} [CROCC < MEL < MEEL < GMM]^{50} [CROCC < MEL < GMM < MEEL]^{100} [CROCC < MEL < GMM < MEEL]^{150} [CROCC < MEL < GMM < MEEL]^{200} [CROCC < MEL < GMM < MEEL]^{250}

The table is summarising the results of MCS (5000 simulations) for all the possible combinations of the degree of endogeneity, strength of instrumental variables, and finite sample sizes. The most left column is representing the different degrees of endogeneity while the first row is showing different strengths of instrumental variables. In the cells, all the estimators are ranked according to their biasedness for each sample size (as given in superscript).

4.3 MSE of MCS for Exactly Determined Model

This section will provide the results of mean square error (MSE) for the exactly determined model. Similar to the previous section, the results are provided for varying degrees of endogeneity, IVS, and sample sizes. The MSE is reported below in Table 4-4.

Table 4-4: Mean Square Error in Exactly Determined Model

Samples size	Instrumental variables strength (IVS)	Degree of Endogeneity (D.E)	Estimators			
			MEL	MEEL	GMM	CROCC
25	IVS=25%	D.E.=20%	5.263648	7.372908	7569.387909	0.608279
		D.E.=40%	4.913355	7.724669	5053.347323	0.605876
		D.E.=60%	5.354470	8.028549	138.997032	0.846179
		D.E.=80%	3.781407	11.889420	4633.633473	0.648154
	IVS=50%	D.E.=20%	2.044735	2.874396	36.727837	0.223713
		D.E.=40%	1.979895	3.280283	1352.915218	0.192680
		D.E.=60%	2.056596	5.415838	250.442897	0.185143
		D.E.=80%	2.405960	6.072679	302.776192	0.250193
	IVS=75%	D.E.=20%	0.351519	0.329624	1.429182	0.052027
		D.E.=40%	0.248625	0.298864	2.638258	0.056278
		D.E.=60%	0.326332	0.595453	2.631434	0.058750
		D.E.=80%	0.531379	0.995246	5.256359	0.053525
50	IVS=25%	D.E.=20%	3.325089	3.320464	714.364708	0.371587
		D.E.=40%	2.443568	4.640276	349.721256	0.362297
		D.E.=60%	2.697626	5.496462	811.753345	0.328868
		D.E.=80%	2.178263	6.498060	208.198871	0.329152
	IVS=50%	D.E.=20%	0.584771	0.445962	4.363759	0.069041
		D.E.=40%	0.372047	0.718847	98.805819	0.070282
		D.E.=60%	0.570402	0.914551	2.171688	0.078162
		D.E.=80%	0.723276	1.835039	18.702903	0.082957
	IVS=75%	D.E.=20%	0.056147	0.049879	0.049505	0.022068
		D.E.=40%	0.060866	0.052790	0.051499	0.022292
		D.E.=60%	0.055727	0.063373	0.061189	0.022746
		D.E.=80%	0.064209	0.074226	0.071389	0.024180
100	IVS=25%	D.E.=20%	0.886567	1.228483	2382.452374	0.121573
		D.E.=40%	0.837616	1.368415	95.734987	0.123083
		D.E.=60%	0.797329	2.195516	249.311614	0.120467
		D.E.=80%	1.148900	2.321780	23.437875	0.145091
	IVS=50%	D.E.=20%	0.099213	0.090654	0.163774	0.032313
		D.E.=40%	0.106259	0.102445	0.099812	0.034019
		D.E.=60%	0.139805	0.168606	0.130409	0.033908
		D.E.=80%	0.143709	0.159510	0.707938	0.035829
	IVS=75%	D.E.=20%	0.023645	0.023819	0.023259	0.011051
		D.E.=40%	0.022517	0.023945	0.022967	0.010994
		D.E.=60%	0.022729	0.025502	0.024090	0.011134
		D.E.=80%	0.022767	0.026040	0.024298	0.010956

Samples size	Instrumental variables strength (IVS)	Degree of Endogeneity (D.E)	Estimators			
			MEL	MEEL	GMM	CROCC
150	IVS= 25%	D.E.= 20%	0.609353	0.471208	1.914875	0.071177
		D.E.= 40%	0.421580	0.724948	24.234358	0.076082
		D.E.= 60%	0.444835	0.847869	24.706707	0.073157
		D.E.= 80%	0.601108	0.911651	17.182840	0.079044
	IVS= 50%	D.E.= 20%	0.053770	0.050489	0.050291	0.020884
		D.E.= 40%	0.053103	0.060548	0.058955	0.021553
		D.E.= 60%	0.057192	0.059158	0.056748	0.021699
		D.E.= 80%	0.053916	0.063646	0.060099	0.021762
	IVS= 75%	D.E.= 20%	0.015116	0.014899	0.014451	0.007111
		D.E.= 40%	0.014174	0.015762	0.015002	0.006982
		D.E.= 60%	0.014131	0.016321	0.015117	0.007057
		D.E.= 80%	0.013776	0.016788	0.015353	0.007149
200	IVS= 25%	D.E.= 20%	0.269296	0.234822	0.978227	0.051263
		D.E.= 40%	0.208236	0.319388	0.229387	0.053005
		D.E.= 60%	0.251505	0.524187	3.225318	0.053489
		D.E.= 80%	0.326377	0.531672	1.781111	0.055447
	IVS= 50%	D.E.= 20%	0.038922	0.036795	0.036443	0.016429
		D.E.= 40%	0.036075	0.037733	0.036691	0.015450
		D.E.= 60%	0.035895	0.041244	0.039315	0.016022
		D.E.= 80%	0.040885	0.042603	0.039934	0.017044
	IVS= 75%	D.E.= 20%	0.010650	0.011761	0.011329	0.005442
		D.E.= 40%	0.010511	0.011820	0.011131	0.005234
		D.E.= 60%	0.010187	0.012420	0.011327	0.005343
		D.E.= 80%	0.010412	0.013247	0.011891	0.005480
250	IVS= 25%	D.E.= 20%	0.124080	0.107231	0.109563	0.038891
		D.E.= 40%	0.130597	0.126172	0.123089	0.041460
		D.E.= 60%	0.164694	0.160929	0.140373	0.042105
		D.E.= 80%	0.187373	0.324472	7.532700	0.041137
	IVS= 50%	D.E.= 20%	0.028748	0.029165	0.028864	0.013156
		D.E.= 40%	0.028952	0.029440	0.028388	0.012787
		D.E.= 60%	0.028658	0.029982	0.028332	0.012793
		D.E.= 80%	0.026844	0.037633	0.034474	0.012590
	IVS= 75%	D.E.= 20%	0.009009	0.008879	0.008512	0.004328
		D.E.= 40%	0.008419	0.009476	0.008768	0.004244
		D.E.= 60%	0.007746	0.009802	0.008919	0.004209

4.3.1 Varying Endogeneity and Samples Against Different IVS

This section will study the MSE of estimators for varying degrees of endogeneity and sample sizes against different strengths of instrumental variables in case of an exactly determined model. Figure 4-9 provides the chart of MSE for weak IVS (25%) and varying degrees of endogeneity. It is found that all the estimators improved their performances with an increase in sample size. It is also found that the GMM estimator has larger MSE as compared to other estimators. While CROCC is found as a superior estimator in terms of MSE, especially when the sample size is small. Overall, it is concluded that CROCC is better than MEL, MEEL, and GMM in terms of MSE.

Figure 4-9: MSE with Weak IVS in Exactly Determined Model

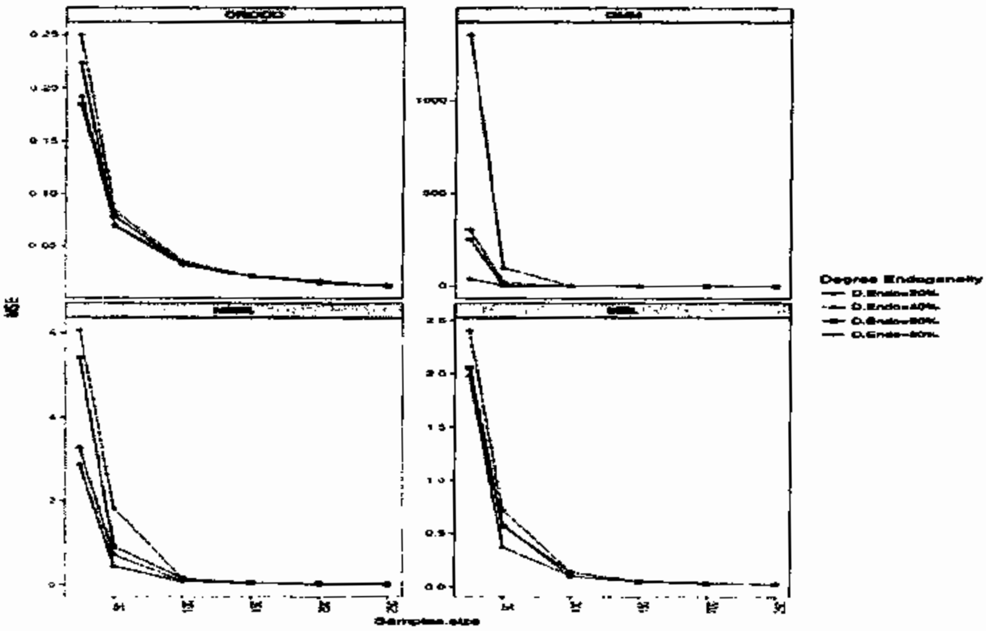
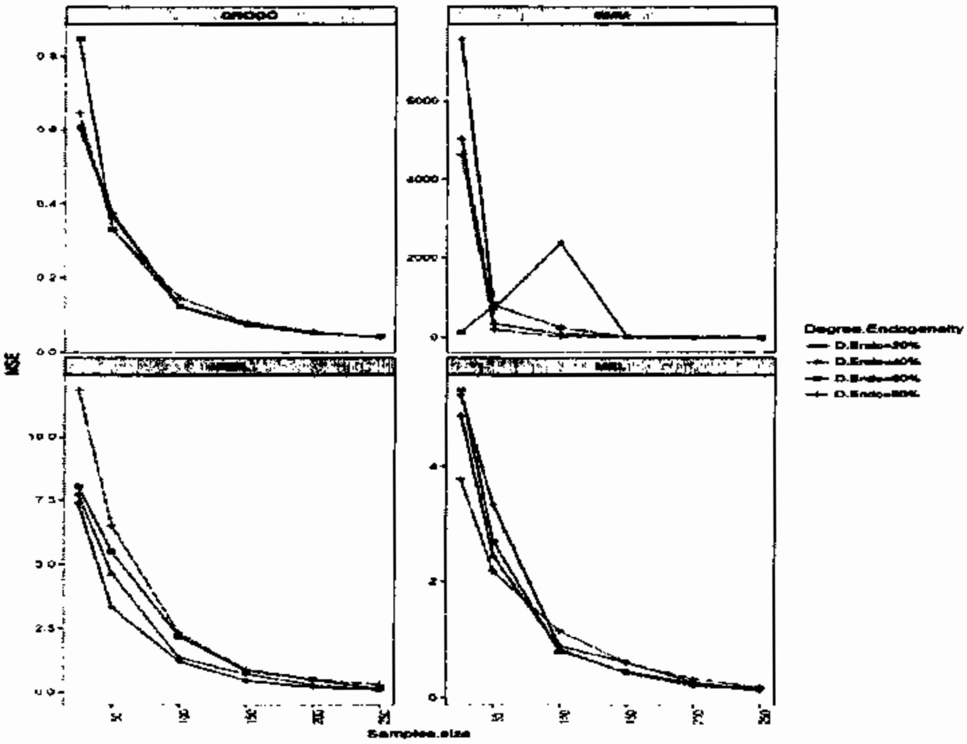


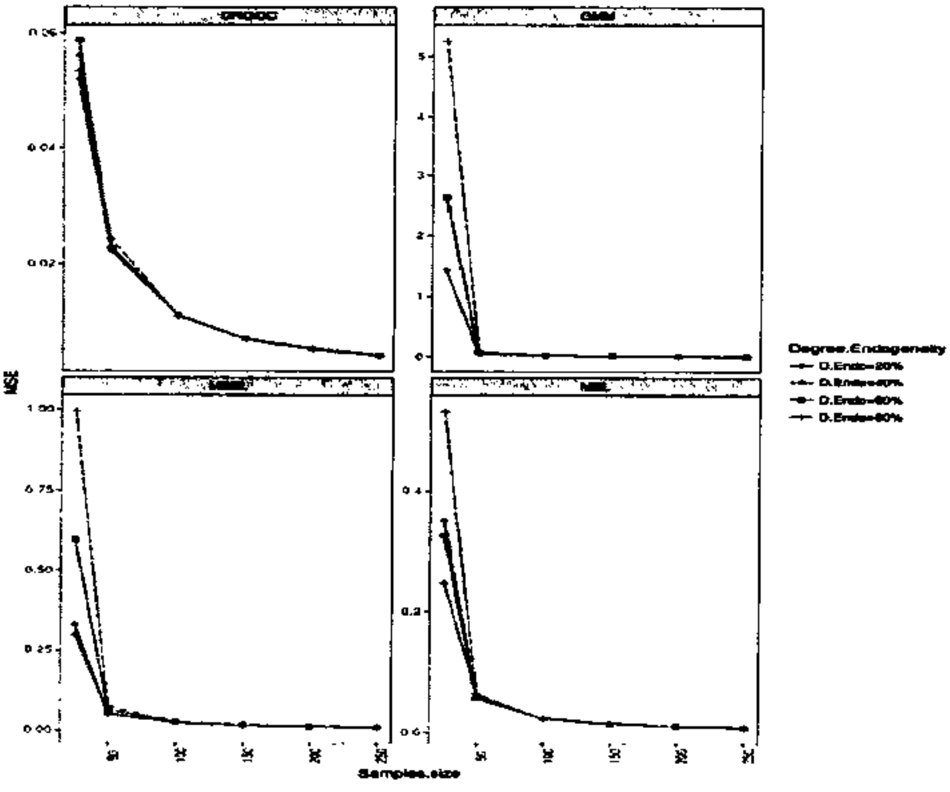
Figure 4-10 evaluates the MSE of the estimators for moderate strength of instrumental variables (50%) and varied degree of endogeneity and sample sizes. It is again found that the GMM estimator has higher MSE than other information-theoretic estimators, while CROCC documented the least MSE. Therefore, CROCC can be seen as a better estimator than MEL, MEEL, and even GMM in terms of biasedness and MSE when the strength of the instrumental variable is moderate.

Figure 4-10: MSE with Moderate IVS in Exactly, Determined Model



Similarly, Figure 4-11 provides the results of MSE for strong (75%) instrumental variable strength and varied degree of endogeneity to examine the efficiency of the estimators. The GMM estimator still has greater MSE than all other information-theoretic estimators. In small samples of 25, all the estimators except CROCC have high variations in MSE with respect to endogeneity. However, for all other samples, no variations are found for all the estimators. Overall analysis shows that CROCC has consistent least MSE as compared to MEL, MEEL, and GMM. Hence, in the case of strong instrumental variables, it can be argued that CROCC is the better estimator while GMM performs worst.

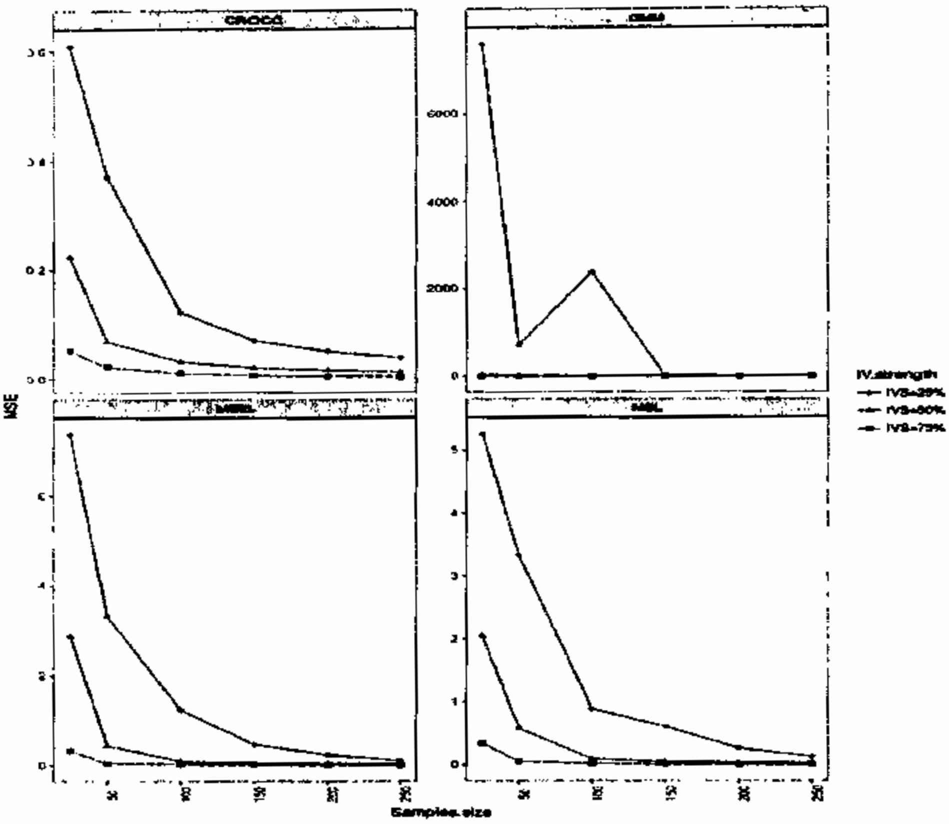
Figure 4-11: MSE with Strong IVS in Exactly, Determined Model



4.3.2 Varying IVS and Sample Sizes Against Different Degrees of Endogeneity

This section will explore the results of MSE for varying IVS and sample sizes against each category of the degree of endogeneity in the case of an exactly determined model. Figure 4-12 evaluates the MSE of the estimators for the low degree of endogeneity, varied IVS, and sample. Results are showing that CROCC has the lowest MSE as compared to all other estimators in this respect. On the other hand, the GMM estimator is showing the highest MSE. Figure 4-12 also revealed that MSE of all the estimators decreases with the increase in the sample size. Similarly, the MSE of all the estimators is less for strong IVS.

Figure 4-12: MSE with Low Degree of Endogeneity in Exactly, Determined Model



Conversely, Figure 4-13 is portraying the efficiency of the estimators in terms of MSE, for a moderate degree of endogeneity (40%) and varied IVS and sample sizes. Figure 4-13 is showing that the GMM has higher MSE than other estimators in all cases. While CROCC showed lower MSE than other estimators, overall, the performance sequence of MSE of four estimators can be written as CROCC < MEL < MEEL < GMM.

Figure 4-13: MSE with Moderate Degree of Endogeneity in Exactly, Determined Model

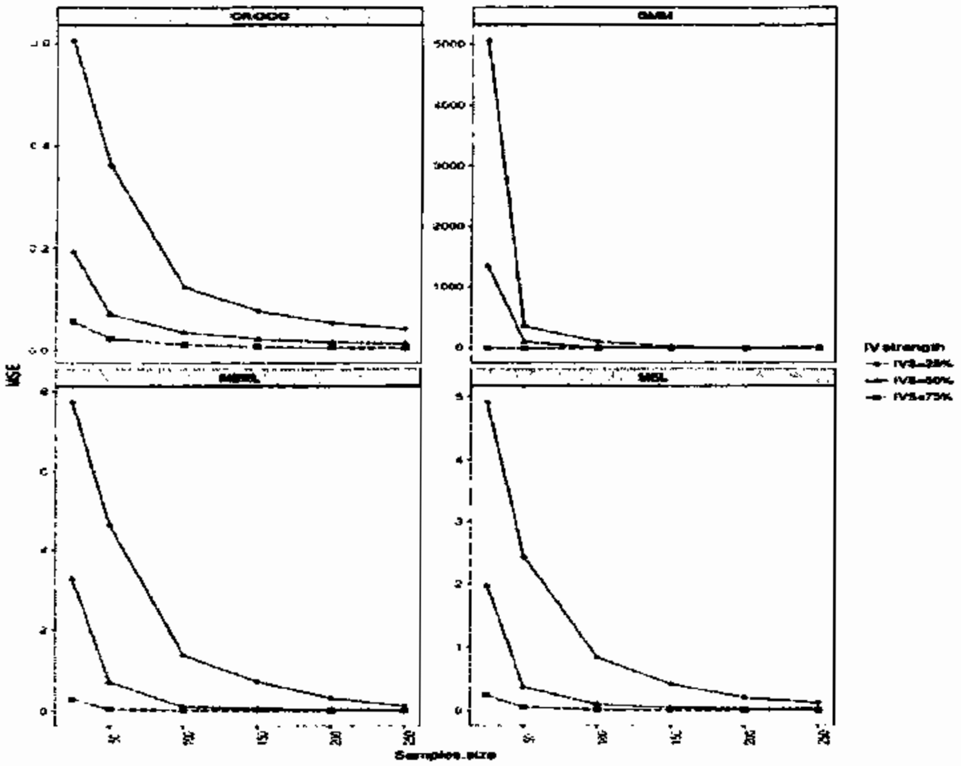
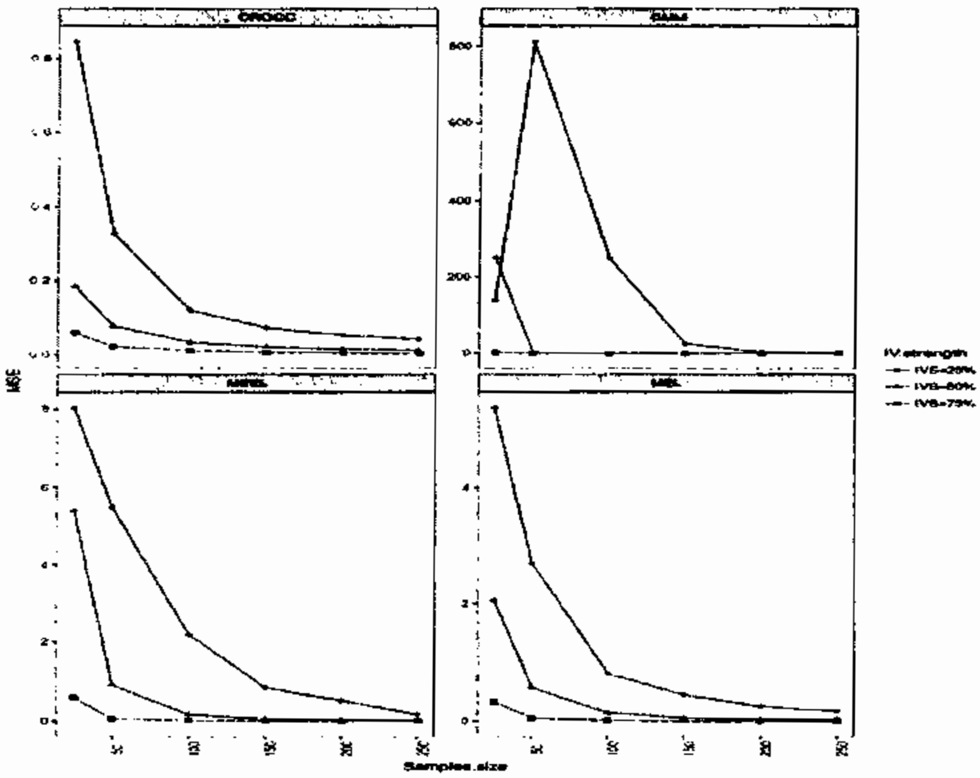


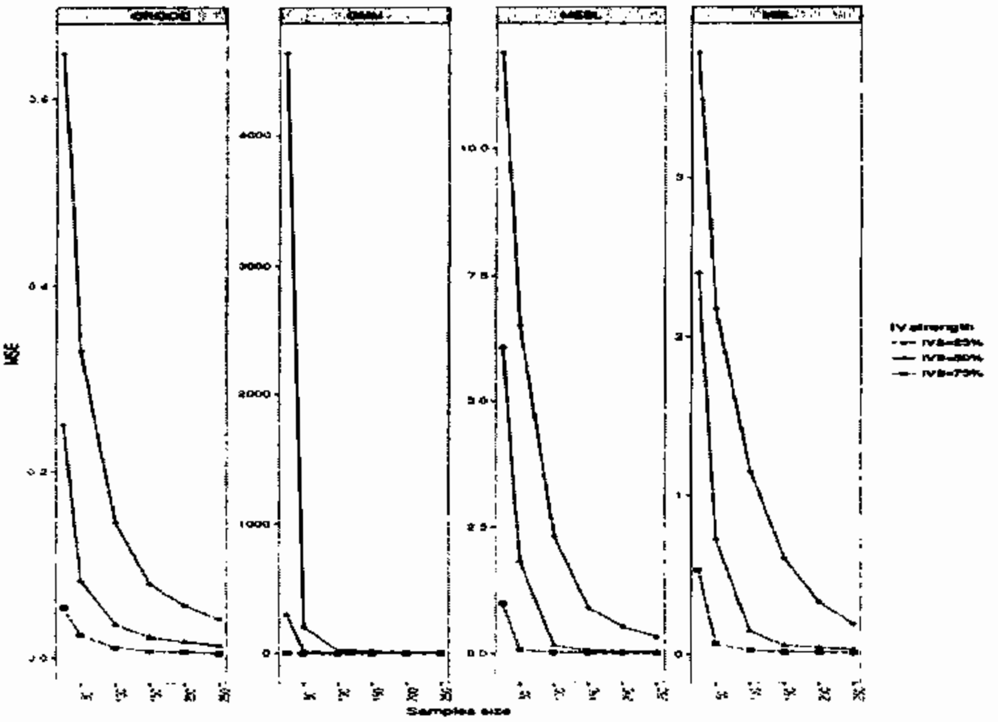
Figure 4-14 further explores the results of MSE for a high degree of endogeneity (60%) and varied IVS and sample sizes. GMM estimator documented high MSE than other estimators, while CROCC showed the least MSE. Especially for weak IVS (25%), the GMM showed very high MSE. However, as the IVS increases, the MSE of all estimators reduced. Overall, the MSE of all the estimators for a high degree of endogeneity can be concluded as CROCC < MEL < MEEL < GMM.

Figure 4-14: MSE with High Degree of Endogeneity in Exactly, Determined Model



At last, Figure 4-15 analyses the efficiency of the estimators when the degree of endogeneity is very high, and IVS and sample size vary. Figure 4-15 is showing that for weak IVS (25%), the MSE of all the estimators is very high. However, as the IVS increases the MSE reduced for all the estimators. In general, CROCC showed the least, while GMM depicted the highest MSE comparatively for a very high degree of endogeneity.

Figure 4-15: MSE with Very High Degree of Endogeneity in Exactly, Determined Model



4.3.3 Summary of Results for MSE in Exactly, Determined Model

The above discussion provides some useful general conclusions about MSE in an exactly determined model for all the estimators. In general, MSE of all the estimators reduced as the IVS increases. Similarly, the MSE of the estimators decreased with the increase in sample size. Table 4-5 further compares the efficiency of all the estimators in all the scenarios. The table is providing two useful implications regarding the efficiency of the estimators. First, CROCC is found as the most efficient estimator in all the combinations. The reason for its superiority can be due to its extraction of information from MEL and MEEL. CROCC uses the minimum quadratic risk estimation rule that ultimately decreases its MSE lower than the individual MSE of MEL and MEEL. Second, GMM is the worst estimator as compared to information-theoretic estimators irrespective of the degree of endogeneity, instrumental variable strength, and sample sizes. Hence, future studies are recommended to use CROCC in estimating the regression model having the problem of endogeneity.

Table 4-5: MSE and Implications for Exactly Determined Model

	Weak Instruments	Moderate Instruments	Strong Instruments
Low Degree	[CROCC < MEL < MEEL < GMM]^{25} [CROCC < MEL < MEEL < GMM]^{50} [CROCC < MEL < MEEL < GMM]^{100} [CROCC < MEL < MEEL < GMM]^{150} [CROCC < MEL < MEEL < GMM]^{200} [CROCC < MEL < MEEL < GMM]^{250}	[CROCC < MEL < MEEL < GMM]^{25} [CROCC < MEL < MEEL < GMM]^{50} [CROCC < MEL < MEEL < GMM]^{100} [CROCC < MEL < MEEL < GMM]^{150} [CROCC < MEL < MEEL < GMM]^{200} [CROCC < MEL < MEEL < GMM]^{250}	[CROCC < MEL < MEEL < GMM]^{25} [CROCC < MEL < MEEL < GMM]^{50} [CROCC < MEL < MEEL < GMM]^{100} [CROCC < MEL < MEEL < GMM]^{150} [CROCC < MEL < MEEL < GMM]^{200} [CROCC < MEL < MEEL < GMM]^{250}
Moderate Degree	[CROCC < MEL < MEEL < GMM]^{25} [CROCC < MEL < MEEL < GMM]^{50} [CROCC < MEL < MEEL < GMM]^{100} [CROCC < MEL < MEEL < GMM]^{150} [CROCC < MEL < MEEL < GMM]^{200} [CROCC < MEL < MEEL < GMM]^{250}	[CROCC < MEL < MEEL < GMM]^{25} [CROCC < MEL < MEEL < GMM]^{50} [CROCC < MEL < MEEL < GMM]^{100} [CROCC < MEL < MEEL < GMM]^{150} [CROCC < MEL < MEEL < GMM]^{200} [CROCC < MEL < MEEL < GMM]^{250}	[CROCC < MEL < MEEL < GMM]^{25} [CROCC < MEL < MEEL < GMM]^{50} [CROCC < MEL < MEEL < GMM]^{100} [CROCC < MEL < MEEL < GMM]^{150} [CROCC < MEL < MEEL < GMM]^{200} [CROCC < MEL < MEEL < GMM]^{250}
High Degree	[CROCC < MEL < MEEL < GMM]^{25} [CROCC < MEL < MEEL < GMM]^{50} [CROCC < MEL < MEEL < GMM]^{100} [CROCC < MEL < MEEL < GMM]^{150} [CROCC < MEL < MEEL < GMM]^{200} [CROCC < MEL < MEEL < GMM]^{250}	[CROCC < MEL < MEEL < GMM]^{25} [CROCC < MEL < MEEL < GMM]^{50} [CROCC < MEL < MEEL < GMM]^{100} [CROCC < MEL < MEEL < GMM]^{150} [CROCC < MEL < MEEL < GMM]^{200} [CROCC < MEL < MEEL < GMM]^{250}	[CROCC < MEL < MEEL < GMM]^{25} [CROCC < MEL < MEEL < GMM]^{50} [CROCC < MEL < MEEL < GMM]^{100} [CROCC < MEL < MEEL < GMM]^{150} [CROCC < MEL < MEEL < GMM]^{200} [CROCC < MEL < MEEL < GMM]^{250}
Very High Degree	[CROCC < MEL < MEEL < GMM]^{25} [CROCC < MEL < MEEL < GMM]^{50} [CROCC < MEL < MEEL < GMM]^{100} [CROCC < MEL < MEEL < GMM]^{150} [CROCC < MEL < MEEL < GMM]^{200} [CROCC < MEL < MEEL < GMM]^{250}	[CROCC < MEL < MEEL < GMM]^{25} [CROCC < MEL < MEEL < GMM]^{50} [CROCC < MEL < MEEL < GMM]^{100} [CROCC < MEL < MEEL < GMM]^{150} [CROCC < MEL < MEEL < GMM]^{200} [CROCC < MEL < MEEL < GMM]^{250}	[CROCC < MEL < MEEL < GMM]^{25} [CROCC < MEL < MEEL < GMM]^{50} [CROCC < MEL < MEEL < GMM]^{100} [CROCC < MEL < MEEL < GMM]^{150} [CROCC < MEL < MEEL < GMM]^{200} [CROCC < MEL < MEEL < GMM]^{250}

The table is summarising the results of MCS (5000 simulations) for all the possible combinations of the degree of endogeneity, strength of instrumental variables, and finite sample sizes. The most left column is representing the different degrees of endogeneity while the first row is showing different strengths of instrumental variables. In the cells, all the estimators are ranked according to their MSE for each sample size (as given in superscript).

4.4 Biasedness of MCS for Overdetermined Model

This section will explore the results of biasedness in the case of the over-determined model. The overdetermined model represents a model where the number of moment conditions exceeds the number of parameters. In such a condition, the GMM estimator is compared with information-theoretic estimators within different combinations of sample sizes, IVS, and degree of endogeneity, as defined in the previous section.

Table 4-6: Small Samples Bias in Overdetermined Model

Samples size	Instrumental variables strength (IVS)	Degree of Endogeneity (D.E)	Estimators			
			MEL	MEEL	GMM	CROCC
25	IVS= 25%	D.E.= 20%	0.131721	0.126771	0.147980	0.137196
		D.E.= 40%	0.255778	0.265153	0.308088	0.285296
		D.E.= 60%	0.419872	0.416329	0.446539	0.439316
		D.E.= 80%	0.557805	0.540770	0.583616	0.588621
	IVS= 50%	D.E.= 20%	0.063521	0.061779	0.092644	0.080708
		D.E.= 40%	0.160483	0.147456	0.180026	0.178066
		D.E.= 60%	0.231653	0.233323	0.275849	0.280108
		D.E.= 80%	0.323395	0.334926	0.376499	0.389285
	IVS= 75%	D.E.= 20%	0.013393	0.017021	0.034329	0.008604
		D.E.= 40%	0.051946	0.027429	0.073958	0.057690
		D.E.= 60%	0.085858	0.062809	0.107013	0.084953
		D.E.= 80%	0.122039	0.100465	0.145094	0.131407
50	IVS= 25%	D.E.= 20%	0.084847	0.086975	0.109530	0.099169
		D.E.= 40%	0.194070	0.188890	0.214384	0.219998
		D.E.= 60%	0.305693	0.309972	0.339093	0.351371
		D.E.= 80%	0.420732	0.430935	0.454369	0.480281
	IVS= 50%	D.E.= 20%	0.032428	0.033764	0.055731	0.048250
		D.E.= 40%	0.077674	0.073730	0.093648	0.108783
		D.E.= 60%	0.129355	0.139178	0.159221	0.178634
		D.E.= 80%	0.177741	0.185994	0.206100	0.237675
	IVS= 75%	D.E.= 20%	-0.002060	-0.009244	0.020730	0.001869
		D.E.= 40%	0.006056	0.009584	0.037914	0.022720
		D.E.= 60%	0.031088	0.028756	0.056226	0.046752
		D.E.= 80%	0.043472	0.043857	0.071486	0.061580
100	IVS= 25%	D.E.= 20%	0.045836	0.049923	0.073881	0.065476
		D.E.= 40%	0.113530	0.112850	0.136220	0.149210
		D.E.= 60%	0.190590	0.203440	0.224390	0.247360
		D.E.= 80%	0.262040	0.271130	0.289670	0.330620
	IVS= 50%	D.E.= 20%	0.006301	0.006367	0.025032	0.017840
		D.E.= 40%	0.030365	0.036634	0.053491	0.057349
		D.E.= 60%	0.054735	0.060672	0.077781	0.092048
		D.E.= 80%	0.081789	0.084509	0.099419	0.124040
	IVS= 75%	D.E.= 20%	-0.017333	-0.011853	0.007158	-0.009614
		D.E.= 40%	-0.007903	-0.003907	0.016408	0.003349
		D.E.= 60%	0.004249	0.007787	0.027684	0.018149
		D.E.= 80%	0.011084	0.016503	0.034947	0.029071

Samples size	Instrumental variables strength (IVS)	Degree of Endogeneity (D.E)	Estimators			
			MEL	MEEL	GMM	CROCC
150	IVS= 25%	D.E.= 20%	0.036437	0.025326	0.047482	0.045859
		D.E.= 40%	0.080001	0.080126	0.100643	0.114459
		D.E.= 60%	0.132863	0.131192	0.151017	0.177000
		D.E.= 80%	0.177526	0.178516	0.196481	0.242562
	IVS= 50%	D.E.= 20%	-0.002232	-0.005472	0.013285	0.004050
		D.E.= 40%	0.015236	0.016666	0.034308	0.031352
		D.E.= 60%	0.032886	0.033054	0.049785	0.055758
		D.E.= 80%	0.050313	0.052517	0.067710	0.081191
	IVS= 75%	D.E.= 20%	-0.017929	-0.010914	0.005726	-0.010443
		D.E.= 40%	-0.014269	-0.004795	0.011653	-0.002789
		D.E.= 60%	-0.007115	0.000206	0.016805	0.005607
		D.E.= 80%	-0.003435	0.005501	0.022361	0.012708
200	IVS= 25%	D.E.= 20%	0.024242	0.009787	0.030911	0.027196
		D.E.= 40%	0.057735	0.054168	0.073706	0.085949
		D.E.= 60%	0.097268	0.099014	0.118139	0.141739
		D.E.= 80%	0.134736	0.131185	0.148410	0.185813
	IVS= 50%	D.E.= 20%	-0.009964	-0.005453	0.013446	-0.000536
		D.E.= 40%	0.007572	0.010855	0.028760	0.022307
		D.E.= 60%	0.018122	0.017151	0.034493	0.035654
		D.E.= 80%	0.033634	0.031535	0.047156	0.055484
	IVS= 75%	D.E.= 20%	-0.023375	-0.009541	0.005307	-0.011780
		D.E.= 40%	-0.018830	-0.007538	0.007453	-0.007048
		D.E.= 60%	-0.013016	0.000585	0.015349	0.001963
		D.E.= 80%	-0.010187	0.000460	0.016356	0.004819
250	IVS= 25%	D.E.= 20%	0.006895	0.007784	0.028306	0.024592
		D.E.= 40%	0.044298	0.042601	0.061166	0.070489
		D.E.= 60%	0.078955	0.068639	0.087549	0.110249
		D.E.= 80%	0.105269	0.100516	0.117594	0.153508
	IVS= 50%	D.E.= 20%	-0.009586	-0.009058	0.010180	-0.003788
		D.E.= 40%	0.001191	0.001824	0.019751	0.011339
		D.E.= 60%	0.012996	0.014494	0.032058	0.028657
		D.E.= 80%	0.022240	0.019940	0.036377	0.040216
	IVS= 75%	D.E.= 20%	-0.021903	-0.008891	0.003629	-0.011515
		D.E.= 40%	-0.019799	-0.006214	0.006700	-0.007259
		D.E.= 60%	-0.014519	-0.005079	0.008738	-0.003586
		D.E.= 80%	-0.012958	-0.000181	0.013685	0.001848

4.4.1 Varying Endogeneity and Samples Against Different IVS

This section will provide the results of biasedness for varied endogeneity and sample size in case of an over-determined model. Figure 4-16 exhibits the biasedness of the estimators for weak IVS and varying degree of endogeneity and sample. On the vertical axis, biasedness is measured while the horizontal axis represents sample sizes. As we increased the endogeneity, the biasedness of all estimators increased. Similarly, when the sample size increased, the biasedness of all estimators decreased and converged to its true parameter value. Figure 4-16 also revealed that MEL and MEEL have low biasedness than GMM and CROCC in all settings. GMM produced the highest biasedness, especially for small sample sizes and low degree of endogeneity.

Figure 4-16: Biasedness with Weak IVS in Overdetermined Model

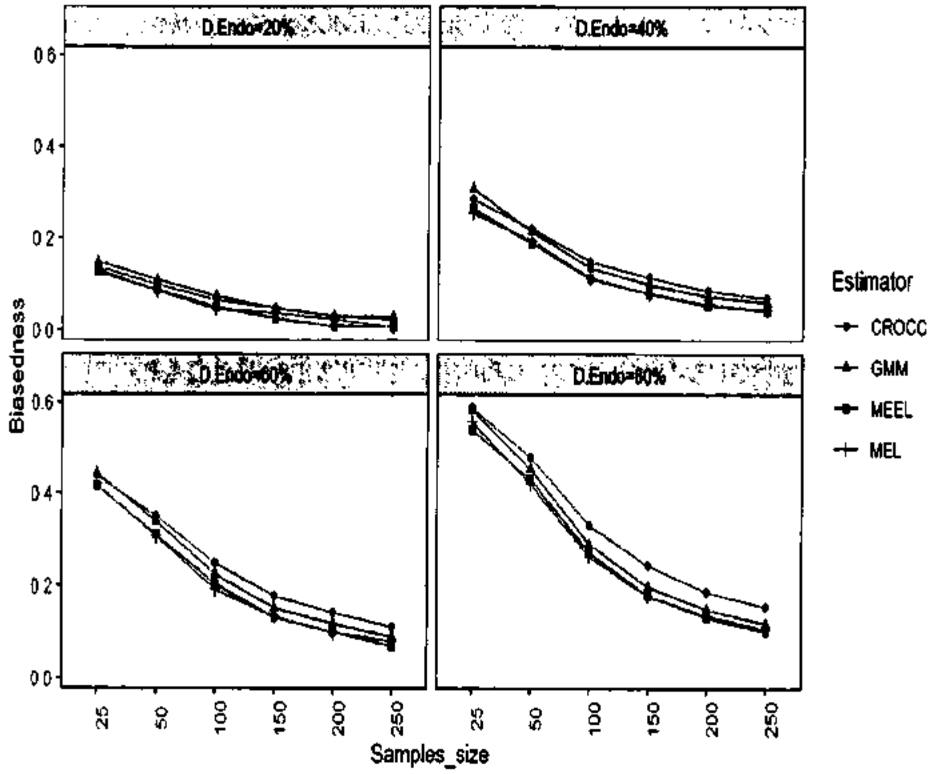


Figure 4-17 reports the biasedness of the estimators when IVS is moderate, and the degree of endogeneity and sample size varies. The increment of the degree of endogeneity from low to very high degree intensifies the biasedness of all estimators, however with the increase in the sample sizes, they presented better performances. Overall, MEL and MEEL estimators are showing lower bias than GMM and CROCC.

Figure 4-17: Biasedness with Moderate IVS in Overdetermined Model

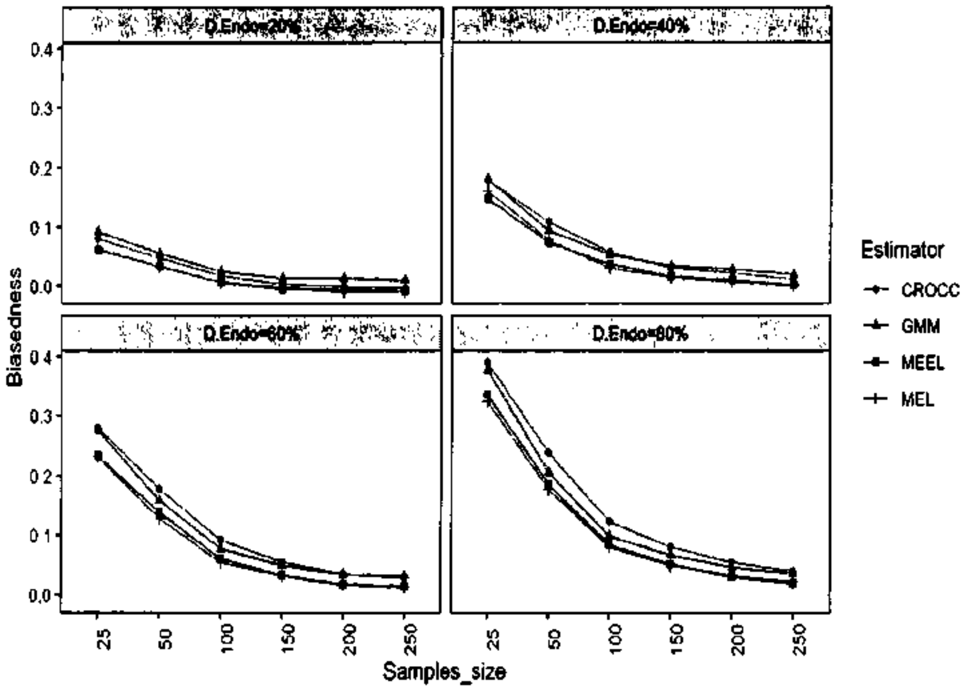
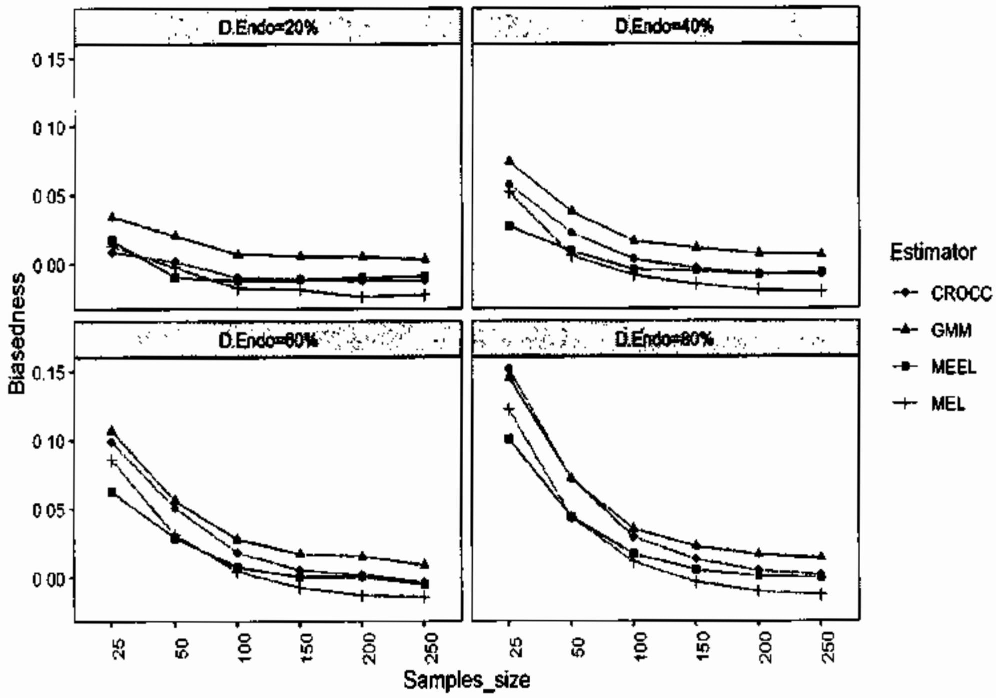


Figure 4-18 explores the biasedness of estimators for strong IVS and varied endogeneity and sample sizes. Results showed that with the increase in the degree of endogeneity, the biasedness of all the estimators also increase. On the other side, all the estimators improved their performance as we increased the sample size. However, the GMM has the highest biasedness than all three estimators in all combinations. This analysis also confirmed the superiority of MEL and MEEL over GMM and CROCC.

Figure 4-18: Biasedness with Strong IVS in Overdetermined Model



4.4.2 Varying IVS and Samples against Different Degrees of Endogeneity

Figure 4-19 analyses the biasedness from MCS for a low degree of endogeneity and varied IVS and sample sizes. In this analysis, the biasedness of GMM is found higher than all other information-theoretic estimators, especially for small sample size and weak IVS. With the increase of IVS, biasedness of all the estimators is reduced. Similarly, the biasedness of all estimators decreases for large sample sizes.

Figure 4-19: Biasedness with Low Degree of Endogeneity in Overdetermined Model

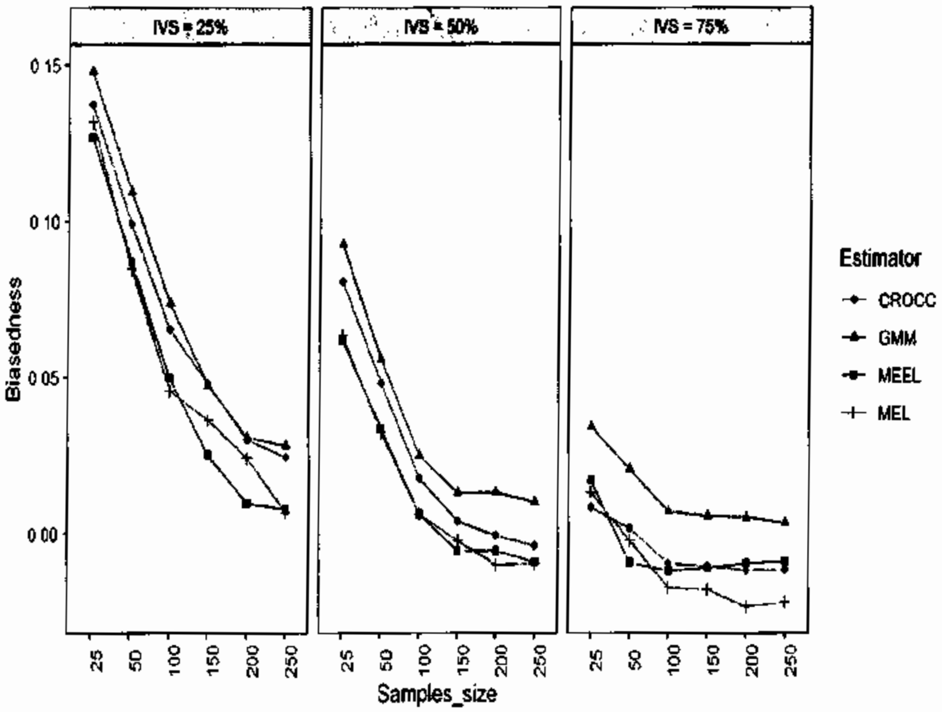
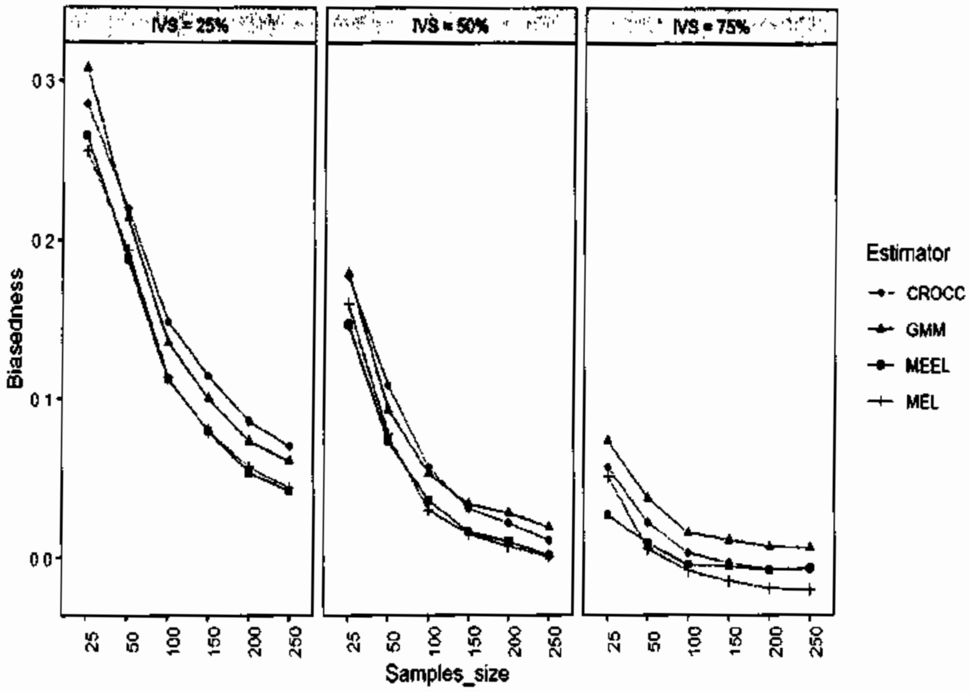


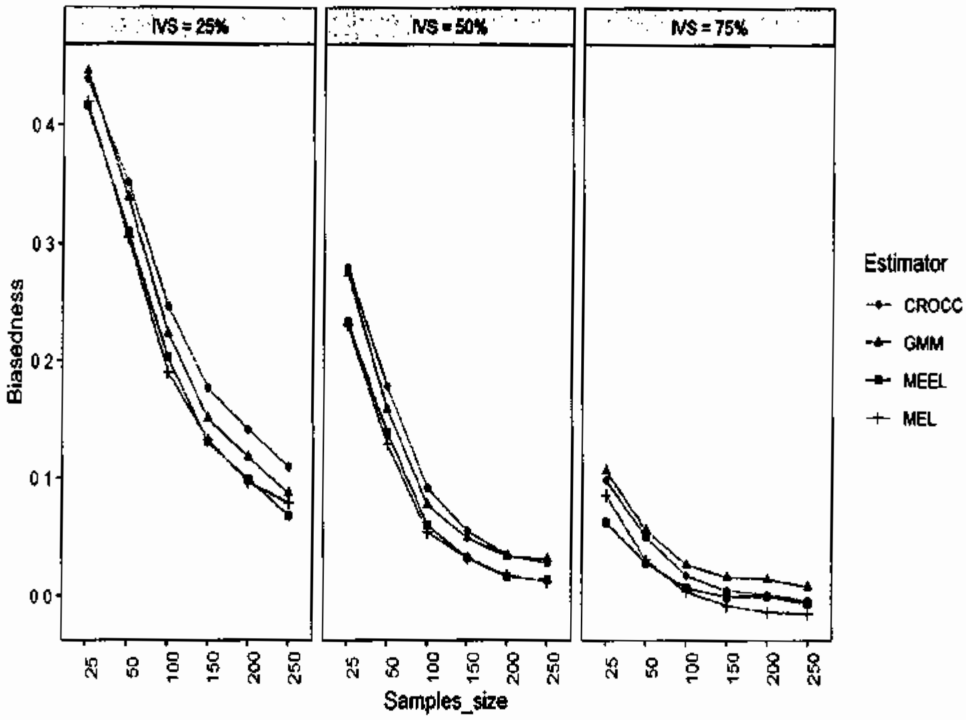
Figure 4-20 provides the detail of biasedness of MCS analysis for varied IVS and samples when the degree of endogeneity is low. In this investigation, the GMM estimator produced more biasedness comparatively for the sample of "25" in all the combinations. Conversely, MEL and MEEL generated the lowest biasedness. Similar to previous results, all the biasedness of all the estimators reduced for large samples and strong IVS.

Figure 4-20: Biasedness with Moderate Degree of Endogeneity in Overdetermined Model



Similarly, Figure 4-21 evaluates the biasedness of estimators when the degree of endogeneity is high. The results showed improved performances of all the estimators for the large sample. Likewise, the biasedness of the estimators reduced when IVS is strong, especially in the case of MEEL. Results further revealed that MEL and MEEL produce less biasedness than GMM and CROCC in all the combinations while GMM performed worst in the case of the small sample and weak IVS.

Figure 4-21: Biasedness with High Degree of Endogeneity in Overdetermined Model



At last, the graphs in Figure 4-22 are provided to analyse the biasedness of the estimators for a very high degree of endogeneity. Results again concluded that the biasedness of the estimators reduces as the sample size increases. Similar to previous results, the performance of all the estimators increase for strong IVS as compared to weak IVS. However, overall results again confirmed the superiority of MEL and MEEL over GMM and CROCC.

Figure 4-22: Biasedness with Very High Degree of Endogeneity in Overdetermined Model

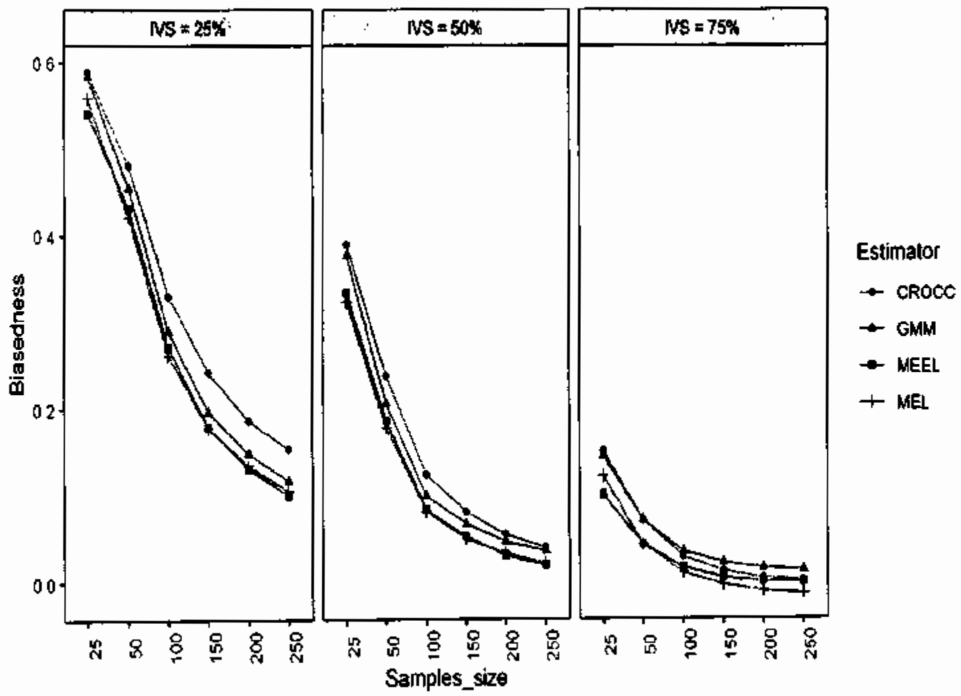
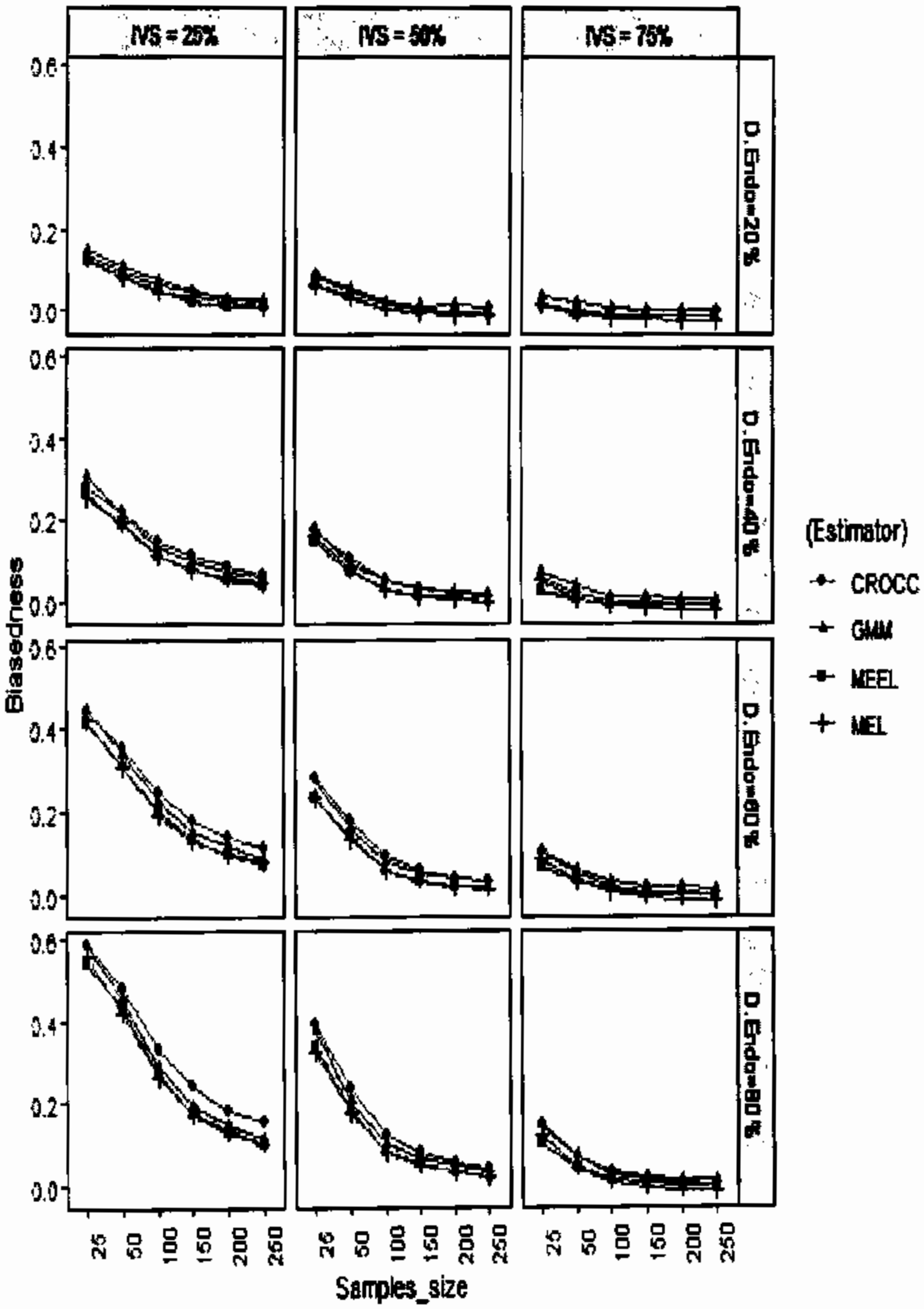


Figure 4-23: Summary of Biasedness in Overdetermined Model



4.4.3 Summary of Results for Biasedness in Overdetermined Model

Figure 4-23 provides the summary results of the biasedness of all the estimators across different samples, degrees of endogeneity, and IVS in the case of the over-determined model. The figure provides some useful general conclusions about estimators' performances. As a whole, the performances of all the estimators increase for a large sample. It indicates that all the estimators are consistent. Similarly, all the estimators produced less biasedness when a strong instrumental variable is used as compared to weak IVS. Conversely, as the degree of endogeneity increases, the biasedness of the estimators also increases comparatively. Table 4-7 is comparing the four estimators for each scenario. The table is providing two useful implications for the results.

First, MEEL is found better estimator in the case of a large sample (200 & 250) irrespective of degrees of endogeneity and instrumental variable strengths. Second, MEL and MEEL are better estimators than CROCC and GMM for weak and moderate instrumental variable strengths. It is notable that in the case of the exactly determined model, CROCC was found better estimators than MEL and MEEL for a very high degree of endogeneity. However, in the case of the overdetermined model, MEL and MEEL are performing better. We noticed that in exactly determined models, MEL and MEEL produced negative biasedness while their optimal convex combination (CROCC) produced positive but close to zero biasedness. Conversely, in the case of an overdetermined model where four instrumental variables are used to estimate one endogenous variable, MEL and MEEL produced positive biasedness. As a result, their optimal convex combination went away from their individual positive biasedness that increases the biasedness of CROCC comparatively. Hence, it is concluded that MEL and MEEL produce less bias relatively in the presence of endogeneity.

Table 4-7: Biasedness and Practical Implications for Overdetermined Model

	Weak Instruments	Moderate Instruments	Strong Instruments
Low Degree	[MEEL < MEL < CROCC < GMM]^{25} [MEL < MEEL < CROCC < GMM]^{50} [MEL < MEEL < CROCC < GMM]^{100} [MEEL < MEL < CROCC < GMM]^{150} [MEEL < MEL < CROCC < GMM]^{200} [MEL < MEEL < CROCC < GMM]^{250}	[MEEL < MEL < CROCC < GMM]^{25} [MEL < MEEL < CROCC < GMM]^{50} [MEEL < MEL < CROCC < GMM]^{100} [MEEL < MEL < CROCC < GMM]^{150} [MEL < MEEL < CROCC < GMM]^{200} [MEL < MEEL < CROCC < GMM]^{250}	[CROCC < MEL < MEEL < GMM]^{25} [CROCC < MEL < MEEL < GMM]^{50} [GMM < CROCC < MEEL < MEL]^{100} [GMM < CROCC < MEEL < MEL]^{150} [GMM < MEEL < CROCC < MEL]^{200} [GMM < MEEL < CROCC < MEL]^{250}
Moderate Degree	[MEL < MEEL < CROCC < GMM]^{25} [MEL < MEEL < GMM < CROCC]^{50} [MEEL < MEL < GMM < CROCC]^{100} [MEEL < MEL < GMM < CROCC]^{150} [MEEL < MEL < GMM < CROCC]^{200} [MEEL < MEL < GMM < CROCC]^{250}	[MEEL < MEL < CROCC < GMM]^{25} [MEEL < MEL < GMM < CROCC]^{50} [MEL < MEEL < GMM < CROCC]^{100} [MEL < MEEL < CROCC < GMM]^{150} [MEL < MEEL < CROCC < GMM]^{200} [MEL < MEEL < CROCC < GMM]^{250}	[MEEL < MEL < CROCC < GMM]^{25} [MEL < MEEL < CROCC < GMM]^{50} [MEL < MEEL < CROCC < GMM]^{100} [CROCC < MEEL < MEL < GMM]^{150} [CROCC < MEEL < GMM < MEL]^{200} [MEEL < CROCC < GMM < MEL]^{250}
High Degree	[MEEL < MEL < CROCC < GMM]^{25} [MEL < MEEL < GMM < CROCC]^{50} [MEL < MEEL < GMM < CROCC]^{100} [MEEL < MEL < GMM < CROCC]^{150} [MEEL < MEL < GMM < CROCC]^{200} [MEEL < MEL < GMM < CROCC]^{250}	[MEL < MEEL < GMM < CROCC]^{25} [MEL < MEEL < GMM < CROCC]^{50} [MEL < MEEL < GMM < CROCC]^{100} [MEL < MEEL < GMM < CROCC]^{150} [MEEL < MEL < GMM < CROCC]^{200} [MEL < MEEL < CROCC < GMM]^{250}	[MEEL < CROCC < MEL < GMM]^{25} [MEEL < MEL < CROCC < GMM]^{50} [MEL < MEEL < CROCC < GMM]^{100} [MEEL < MEL < CROCC < GMM]^{150} [MEEL < MEL < CROCC < GMM]^{200} [MEEL < CROCC < GMM < MEL]^{250}
Very High Degree	[MEEL < MEL < CROCC < GMM]]^{25} [MEL < MEEL < GMM < CROCC]^{50} [MEL < MEEL < GMM < CROCC]^{100} [MEL < MEEL < GMM < CROCC]^{150} [MEEL < MEL < GMM < CROCC]^{200} [MEEL < MEL < GMM < CROCC]^{250}	[MEL < MEEL < GMM < CROCC]^{25} [MEL < MEEL < GMM < CROCC]^{50} [MEL < MEEL < GMM < CROCC]^{100} [MEL < MEEL < GMM < CROCC]^{150} [MEEL < MEL < GMM < CROCC]^{200} [MEEL < MEL < GMM < CROCC]^{250}	[MEEL < CROCC < GMM < MEL]^{25} [MEEL < CROCC < GMM < MEL]^{50} [MEL < MEEL < CROCC < GMM]^{100} [MEL < MEEL < CROCC < GMM]^{150} [MEEL < CROCC < MEL < GMM]^{200} [MEEL < CROCC < MEL < GMM]^{250}

The table is summarising the results of MCS (5000 simulations) for all the possible combinations of the degree of endogeneity, strength of instrumental variables, and finite sample sizes. The most left column is representing the different degrees of endogeneity while the first row is showing different strengths of instrumental variables. In the cells, all the estimators are ranked according to their biasedness for each sample size (as given in superscript).

4.5 MSE of MCS for Overdetermined Model

This section will explore the results for MSE of MCS analysis in the case of an overdetermined model. Similar to previous sections, results are presented with respect to varying samples, IVS, and degree of endogeneity.

Table 4-8: Mean Square Error in Overdetermined Model

Samples size	Instrumental variables strength (IVS)	Degree of Endogeneity (D.E)	Estimators			
			MEL	MEEL	GMM	CROCC
25	IVS= 25%	D.E.= 20%	0.384619	0.565216	0.396791	0.142165
		D.E.= 40%	0.420759	0.708238	0.393339	0.195545
		D.E.= 60%	0.481954	0.885732	0.444742	0.287124
		D.E.= 80%	0.545038	0.538630	0.519353	0.414447
	IVS= 50%	D.E.= 20%	0.279600	0.250216	0.219084	0.093487
		D.E.= 40%	0.262562	0.295217	0.243078	0.135700
		D.E.= 60%	0.252804	0.296491	0.257719	0.151019
		D.E.= 80%	0.295358	0.354622	0.302199	0.226938
	IVS= 75%	D.E.= 20%	0.100485	0.140272	0.090139	0.048587
		D.E.= 40%	0.101984	0.185907	0.097544	0.052662
		D.E.= 60%	0.119218	0.186806	0.099543	0.055584
		D.E.= 80%	0.103223	0.184966	0.098453	0.069119
50	IVS= 25%	D.E.= 20%	0.285730	0.298636	0.268605	0.100122
		D.E.= 40%	0.292070	0.285157	0.295041	0.132308
		D.E.= 60%	0.321175	0.338128	0.337904	0.197253
		D.E.= 80%	0.349902	0.373502	0.374043	0.290284
	IVS= 50%	D.E.= 20%	0.117578	0.129706	0.130162	0.055653
		D.E.= 40%	0.126849	0.140433	0.138831	0.062960
		D.E.= 60%	0.124191	0.137880	0.136634	0.076562
		D.E.= 80%	0.134618	0.129555	0.126966	0.097222
	IVS= 75%	D.E.= 20%	0.044570	0.047386	0.043746	0.022831
		D.E.= 40%	0.043417	0.046137	0.043572	0.023322
		D.E.= 60%	0.041735	0.048471	0.045646	0.025534
		D.E.= 80%	0.044841	0.045488	0.044463	0.025519
100	IVS= 25%	D.E.= 20%	0.173550	0.177550	0.181350	0.067122
		D.E.= 40%	0.173680	0.165970	0.171090	0.081123
		D.E.= 60%	0.176090	0.186210	0.193750	0.113820
		D.E.= 80%	0.193930	0.200370	0.203970	0.155950
	IVS= 50%	D.E.= 20%	0.063308	0.063903	0.063975	0.030024
		D.E.= 40%	0.062511	0.065364	0.065144	0.032317
		D.E.= 60%	0.060711	0.066518	0.067020	0.036067
		D.E.= 80%	0.059598	0.066392	0.066304	0.040417
	IVS= 75%	D.E.= 20%	0.020956	0.021839	0.021053	0.011646
		D.E.= 40%	0.021989	0.022472	0.021808	0.012315
		D.E.= 60%	0.021311	0.021142	0.020850	0.011333
		D.E.= 80%	0.020368	0.021209	0.020592	0.011940

Samples size	Instrumental variables strength (IVS)	Degree of Endogeneity (D.E)	Estimators			
			MEL	MEEL	GMM	CROCC
150	IVS= 25%	D.E.= 20%	0.120202	0.116633	0.120753	0.049170
		D.E.= 40%	0.113844	0.123333	0.126852	0.057323
		D.E.= 60%	0.118662	0.119804	0.123769	0.072846
		D.E.= 80%	0.117712	0.130283	0.133029	0.094864
	IVS= 50%	D.E.= 20%	0.039606	0.042295	0.042621	0.020605
		D.E.= 40%	0.039906	0.041364	0.041822	0.021531
		D.E.= 60%	0.039657	0.041460	0.041644	0.022505
		D.E.= 80%	0.038199	0.042712	0.042864	0.024914
	IVS= 75%	D.E.= 20%	0.014087	0.014187	0.013763	0.007789
		D.E.= 40%	0.013310	0.014556	0.014301	0.007627
		D.E.= 60%	0.013522	0.014159	0.013961	0.007562
		D.E.= 80%	0.013300	0.014559	0.014657	0.007570
200	IVS= 25%	D.E.= 20%	0.090858	0.087821	0.090059	0.039088
		D.E.= 40%	0.093598	0.094068	0.097319	0.044515
		D.E.= 60%	0.094557	0.089119	0.092539	0.054335
		D.E.= 80%	0.087001	0.092732	0.096485	0.063819
	IVS= 50%	D.E.= 20%	0.031710	0.030247	0.030654	0.015889
		D.E.= 40%	0.029372	0.030617	0.031037	0.015863
		D.E.= 60%	0.029497	0.030958	0.030976	0.016255
		D.E.= 80%	0.028220	0.033044	0.033058	0.018259
	IVS= 75%	D.E.= 20%	0.010414	0.010794	0.010554	0.005790
		D.E.= 40%	0.010625	0.010793	0.010566	0.005750
		D.E.= 60%	0.010535	0.010944	0.010531	0.005957
		D.E.= 80%	0.010220	0.010696	0.010532	0.005580
250	IVS= 25%	D.E.= 20%	0.074604	0.074094	0.076236	0.032360
		D.E.= 40%	0.080767	0.073081	0.075773	0.036704
		D.E.= 60%	0.070883	0.072105	0.075433	0.040152
		D.E.= 80%	0.075599	0.076733	0.079593	0.048716
	IVS= 50%	D.E.= 20%	0.023952	0.025753	0.025833	0.013589
		D.E.= 40%	0.023861	0.023784	0.023779	0.012581
		D.E.= 60%	0.023663	0.024933	0.025135	0.013141
		D.E.= 80%	0.023232	0.025222	0.025413	0.013391
	IVS= 75%	D.E.= 20%	0.008309	0.008761	0.008372	0.004967
		D.E.= 40%	0.008704	0.008928	0.008516	0.005144
		D.E.= 60%	0.007929	0.008789	0.008400	0.004771
		D.E.= 80%	0.008475	0.008778	0.008470	0.004769

4.5.1 Varying Endogeneity and Samples Against Different IVS

Figure 4-24 presents the MSE for varied endogeneity and samples when IVS is weak. Results are showing that as the sample size increases, the MSE of all the estimators decreases. However, the graphs are clearly showing that CROCC produced least MSE regardless of the sample size, IVS, or degree of endogeneity. For a small sample size, MEEL performed worst and produced the highest MSE.

Figure 4-24: MSE for Weak IVS in Overdetermined Model

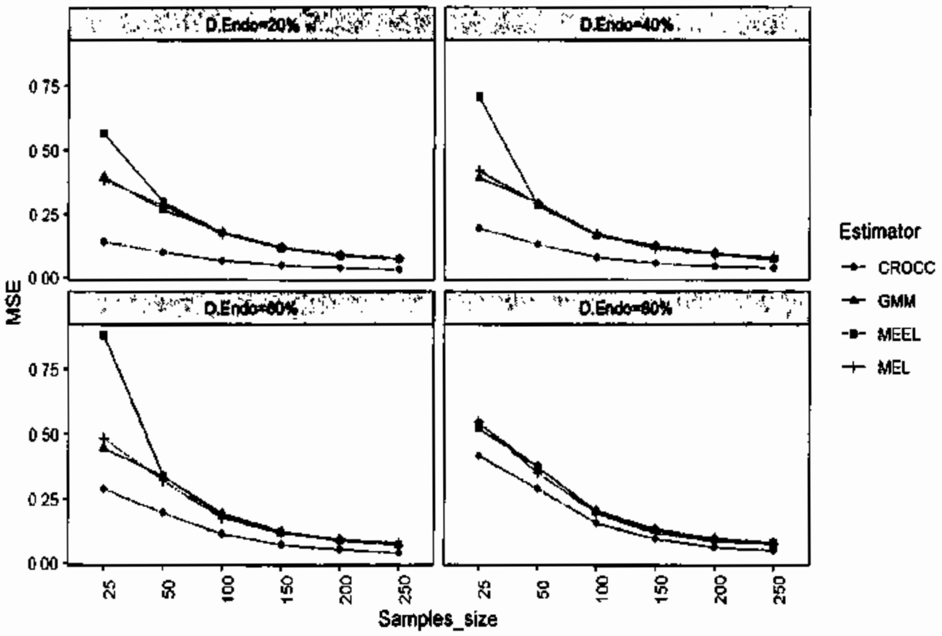
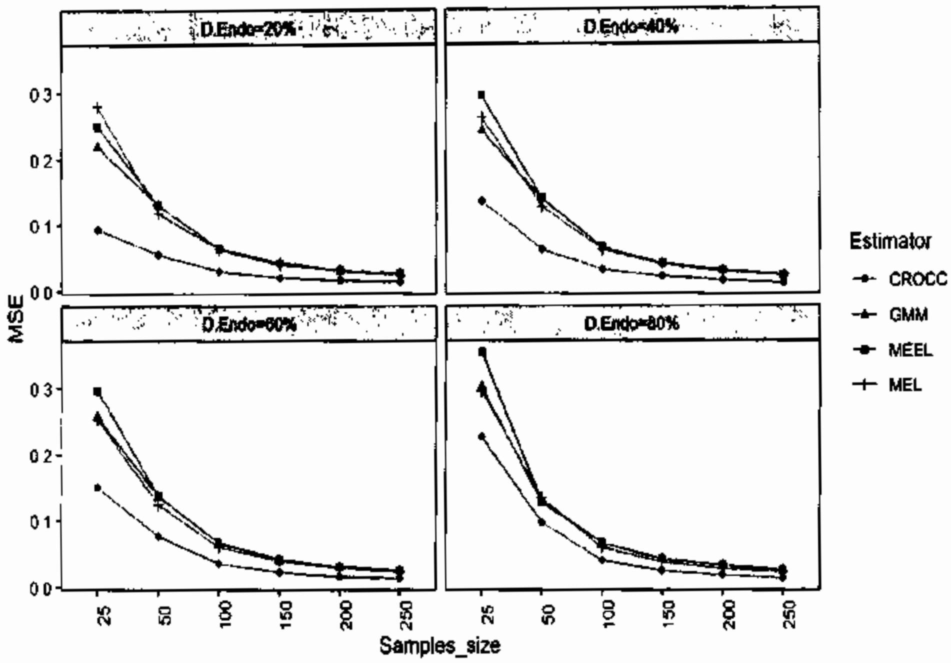


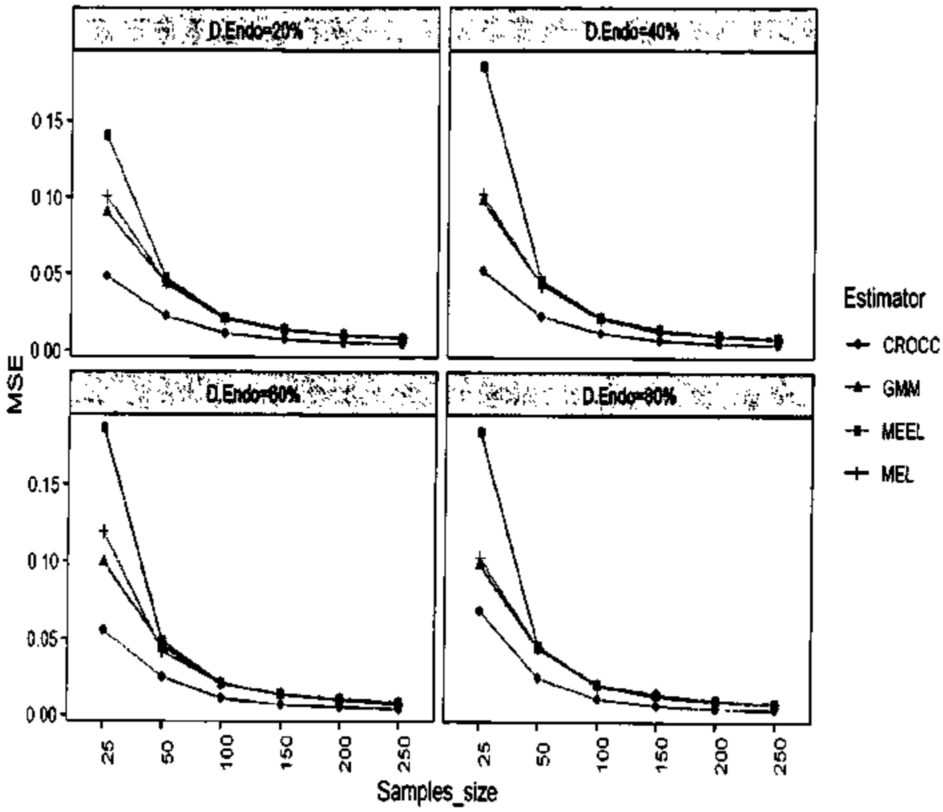
Figure 4-25 represents the graphs of MSE of estimators for moderate IVS in the overdetermined model. Results are showing that CROCC is performing best in all the combinations. However, the MSE of all the estimators reduces with the increase in sample size. MEEL also performed worst when the degree of endogeneity is 40% or above.

Figure 4-25: MSE for Moderate IVS in Overdetermined Model



Similarly, Figure 4-26 explores the results of MSE for varied endogeneity and sample sizes when IVS is strong. This analysis exhibited that the MSE of CROCC is less than the MEL, MEEL, and GMM estimators. As the degree of endogeneity increases, the MSE of all estimators increased when the sample is small. Conversely, the MSE decreased as the sample size increases. Overall, CROCC performed best while MEEL produced more MSE, especially for a small sample.

Figure 4-26: MSE for Strong IVS in Overdetermined Model



4.5.2 Varying IVS and Sample Sizes Against Different Degrees of Endogeneity

Figure 4-27 exhibited the graphs about MSE when the degree of endogeneity is low in an overdetermined model. The graphs are clearly showing that CROCC is the best estimator comparatively regardless of the variations in IVS, degree of endogeneity, and sample sizes. However, the MSE of all the estimators decreases as the sample size increases, and IVS becomes strong.

Figure 4-27: MSE for Low Degree of Endogeneity in Overdetermined Model

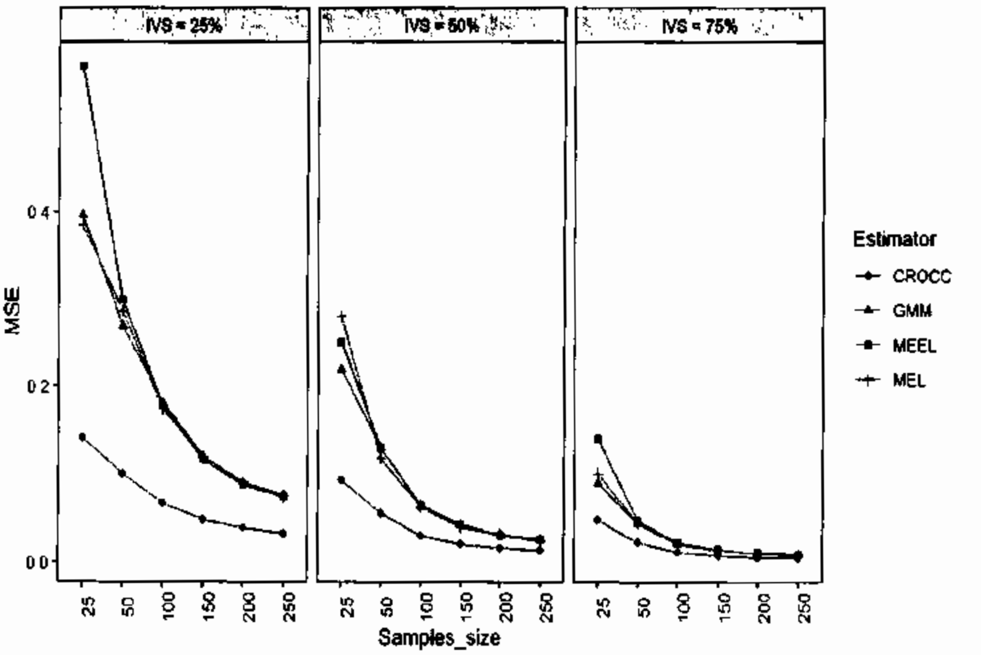
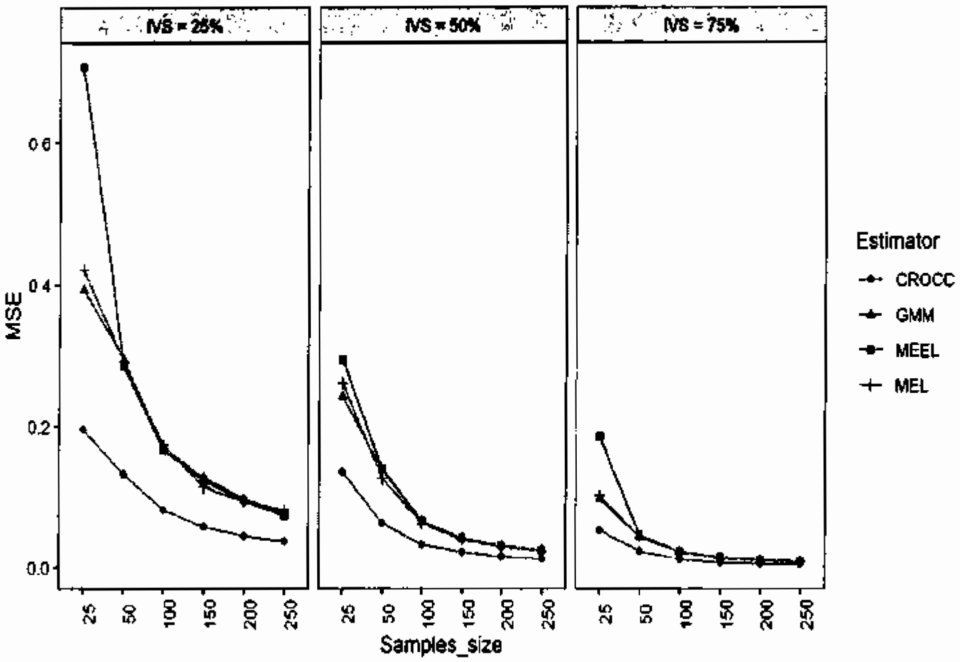


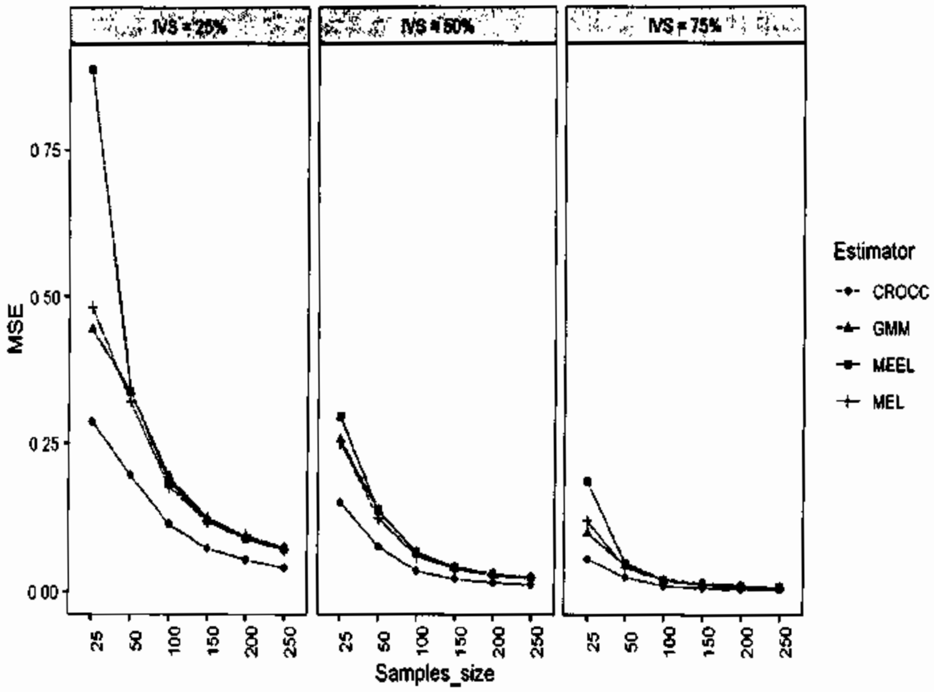
Figure 4-28 fix the degree of endogeneity at a moderate level and show MSE for varied IVS and samples. Results showed that as the IVS changes from weak to strong, the MSE decreased of all the parameters. Similar to previous results, MSE also reduced with the increase in sample size in this context. Similarly, for all the combinations, CROCC is found the best option comparatively.

Figure 4-28: MSE for Moderate Degree of Endogeneity in Overdetermined Model



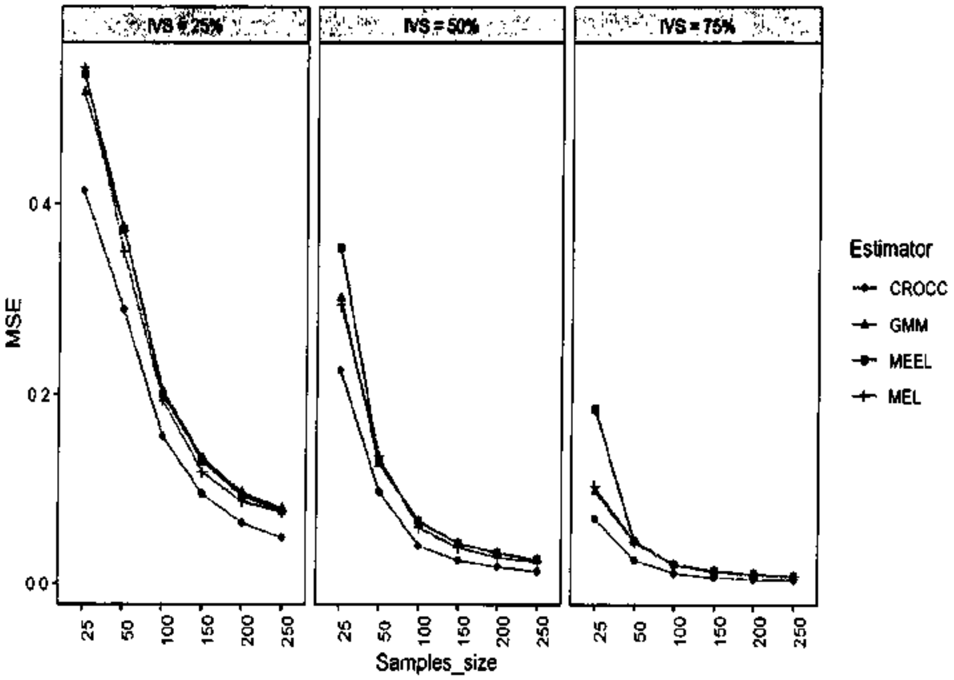
In Figure 4-29, MSE of estimators at a high degree of endogeneity is presented. The graphs are depicting CROCC as the best estimator comparatively in terms of MSE. The MEEL estimator produced higher MSE for the small sample, but for a larger sample size, its MSE documented similar results of MEL and GMM. On the other side, the MSE of all estimators decreased as we have increased the strength of instruments. Similarly, with the increase in sample size, all the estimators improved their MSE.

Figure 4-29: MSE for High Degree of Endogeneity in Overdetermined Model



At last, Figure 4-30 is showing the MSE of all estimators at a very high degree of endogeneity. This investigation also reveals that the MSE of CROCC is less than the MEL, MEEL, and GMM. In the samples of "25," the MEEL estimator has a higher MSE than the MEL and GMM estimator. On the other hand, the MSE of all estimators has decreased as we increased the sample size.

Figure 4-30: MSE for Very High Degree of Endogeneity in Overdetermined Model



4.5.3 Summary of Results for MSE in Overdetermined Model

In Figure 4-31, the MSE of all the estimators is summarized across different sample sizes, IVS, and endogeneity. All the estimators improved their MSE as the strength of the instrumental variable increases. Similarly, as the sample size increases, the MSE of MEL, MEEL, and GMM estimators showed performances close to each other. In most of the cases, the MSE of all the estimators increased with an increase in the degree of endogeneity from low to very high.

Table 4-9 further compares the efficiency of four estimators. Results showed that CROCC is found the most efficient estimator as compared to other estimators in all the combinations. Conversely, GMM is found the least efficient estimator in case of low and moderate strength of instrument variables in most cases. However, for strong instrument variable strength, MEEL is the least efficient estimator. Therefore, it is concluded that CROCC is the most efficient estimator in both exactly determined and overdetermined models having the problem of endogeneity.

Figure 4-31: Summary of MSE in Overdetermined Model

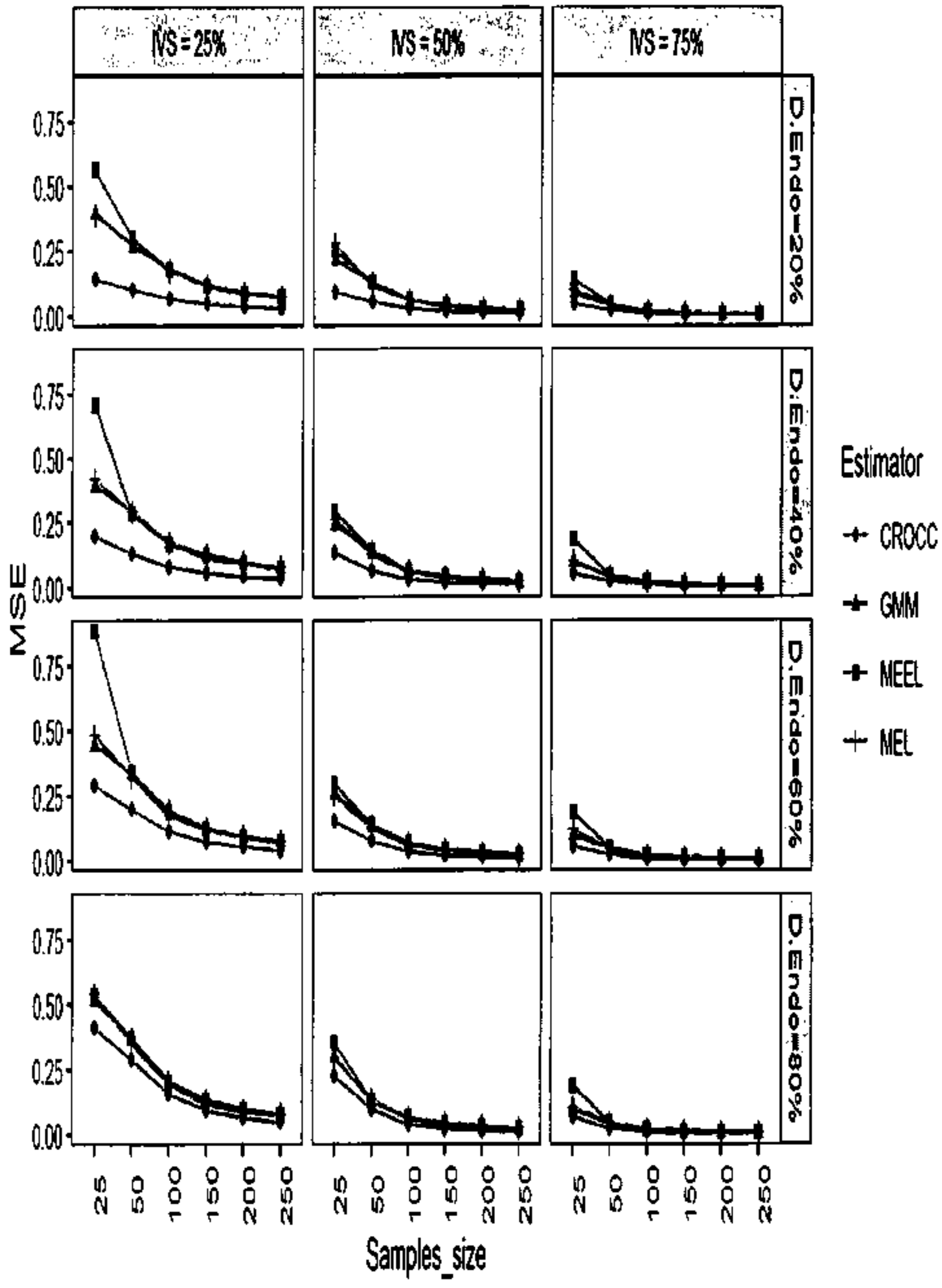


Table 4-9: MSE and Practical Implications for Overdetermined Model

	Weak Instruments	Moderate Instruments	Strong Instruments
Low Degree	[CROCC < MEL < GMM < MEEL]^{25} [CROCC < GMM < MEL < MEEL]^{50} [CROCC < MEL < MEEL < GMM]^{100} [CROCC < MEL < MEEL < GMM]^{150} [CROCC < MEEL < GMM < MEL]^{200} [CROCC < MEEL < GMM < MEL]^{250}	[CROCC < GMM < MEEL < MEL]^{25} [CROCC < MEL < MEEL < GMM]^{50} [CROCC < MEL < MEEL < GMM]^{100} [CROCC < MEL < MEEL < GMM]^{150} [CROCC < MEL < MEEL < GMM]^{200} [CROCC < MEL < MEEL < GMM]^{150}	[CROCC < GMM < MEEL < MEL]^{25} [CROCC < GMM < MEEL < MEL]^{50} [CROCC < MEL < GMM < MEEL]^{100} [CROCC < MEL < GMM < MEEL]^{150} [CROCC < MEL < GMM < MEEL]^{200} [CROCC < MEL < GMM < MEEL]^{250}
Moderate Degree	[CROCC< GMM < MEL < MEEL]^{25} [CROCC < MEEL < MEL < GMM]^{50} [CROCC < MEEL < MEL < GMM]^{100} [CROCC < MEL < MEEL < GMM]^{150} [CROCC < MEL < MEEL < GMM]^{200} [CROCC < MEEL < MEL < GMM]^{250}	[CROCC< GMM < MEL < MEEL]^{25} [CROCC < MEL < MEEL < GMM]^{50} [CROCC < MEL < MEEL < GMM]^{100} [CROCC < MEL < MEEL < GMM]^{150} [CROCC < MEL < MEEL < GMM]^{200} [CROCC < MEL < MEEL < GMM]^{250}	[CROCC< GMM < MEL < MEEL]^{25} [CROCC < MEL < GMM < MEEL]^{50} [CROCC < MEL < GMM < MEEL]^{100} [CROCC < MEL < GMM < MEEL]^{150} [CROCC < MEL < GMM < MEEL]^{200} [CROCC < MEL < GMM < MEEL]^{250}
High Degree	[CROCC < GMM < MEL < MEEL]^{25} [CROCC < MEL < GMM < MEEL]^{50} [CROCC < MEL < MEEL < GMM]^{100} [CROCC < MEL < MEEL < GMM]^{150} [CROCC < MEEL < GMM < MEL]^{200} [CROCC < MEEL < MEL < GMM]^{250}	[CROCC < MEL < GMM < MEEL]^{25} [CROCC < MEL < MEEL < GMM]^{50} [CROCC < MEL < MEEL < GMM]^{100} [CROCC < MEL < MEEL < GMM]^{150} [CROCC < MEL < MEEL < GMM]^{200} [CROCC < MEL < MEEL < GMM]^{250}	[CROCC < GMM < MEL < MEEL]^{25} [CROCC < MEL < GMM < MEEL]^{50} [CROCC < MEL < GMM < MEEL]^{100} [CROCC < MEL < GMM < MEEL]^{150} [CROCC < MEL < GMM < MEEL]^{200} [CROCC < MEL < GMM < MEEL]^{250}
Very High Degree	[CROCC < MEL < MEEL < GMM]^{25} [CROCC < MEL < MEEL < GMM]^{50} [CROCC < MEL < MEEL < GMM]^{100} [CROCC < MEL < MEEL < GMM]^{150} [CROCC < MEL < MEEL < GMM]^{200} [CROCC < MEL < MEEL < GMM]^{250}	[CROCC < MEL < GMM < MEEL]^{25} [CROCC < GMM < MEEL < MEL]^{50} [CROCC < MEL < MEEL < GMM]^{100} [CROCC < MEL < MEEL < GMM]^{150} [CROCC < MEL < MEEL < GMM]^{200} [CROCC < MEL < MEEL < GMM]^{250}	[CROCC < GMM < MEL < MEEL]^{25} [CROCC < GMM < MEL < MEEL]^{50} [CROCC < MEL < MEEL < GMM]^{100} [CROCC < MEL < MEEL < GMM]^{150} [CROCC < MEL < MEEL < GMM]^{200} [CROCC < MEL < MEEL < GMM]^{250}

The table is summarising the results of MCS (5000 simulations) for all the possible combinations of the degree of endogeneity, strength of instrumental variables, and finite sample sizes. The most left column is representing the different degrees of endogeneity while the first row is showing different strengths of instrumental variables. In the cells, all the estimators are ranked according to their MSE for each sample size (as given in superscript).

4.6 Comparative Analysis of Convex Combinations of Cressie and Read Information Theoretic Approach

Previous results have explored that CROCC is better than GMM and even other individual information-theoretic estimators in many cases. CROCC is based on Cressie and Read (CR) information-theoretic approach that optimizes the convex combination, as suggested by (Judge & Mittelhammer, 2011). However, the CROCC estimator has some limitations, as we have discussed in the previous section that may influence the performance of the estimator. For this purpose, we introduced another estimator that estimates Cressie and Read Arbitrary Convex Combinations (CRACC) from MEL and MEEL. This dissertation compares the finite sample properties of convex combinations of CROCC with CRACC. The descriptions of CROCC and CRACC are given as below.

The CROCC uses the following estimation method

$$\bar{\boldsymbol{\beta}}(\hat{\boldsymbol{\alpha}}) = \hat{\boldsymbol{\alpha}}\,\hat{\boldsymbol{\beta}}\,(\mathrm{MEL}) + (1-\hat{\boldsymbol{\alpha}})\,\hat{\boldsymbol{\beta}}\,(\mathrm{MEEL})$$

Where '$\hat{\alpha}$' is estimated as follow

$$\hat{\boldsymbol{\alpha}} = \frac{\mathrm{tr}\,(\mathrm{cov}\left(\hat{\boldsymbol{\beta}}(\mathrm{MEL})\right)}{\mathrm{tr}\,(\mathrm{cov}\left(\hat{\boldsymbol{\beta}}(\mathrm{MEL})\right) + \mathrm{tr}\,(\mathrm{cov}\left(\hat{\boldsymbol{\beta}}(\mathrm{MEEL})\right)}$$

While CRACC uses the following estimation method

$$\bar{\boldsymbol{\beta}}(\mathbf{a}) = \mathbf{a}\,\hat{\boldsymbol{\beta}}\,(\mathrm{MEL}) + (1-\mathbf{a})\,\hat{\boldsymbol{\beta}}\,(\mathrm{MEEL})$$

Where 'a' is an arbitrary value, as shown in Table 4-10, this dissertation proposed three arbitrary convex combinations (i.e., 25%, 50%, and 75%) for the CR method. To better understand the results, three arbitrary combinations are labeled as CRACC1 (for the value of 25%), CRACC2 (for the value of 50%), and CRACC3 (for the value of 75%).

The subsequent part will provide the comparison of biasedness and MSE on different sample sizes using the "5000" Monte Carlo simulation (MCS) analysis for CROCC and three proposed arbitrary combinations of CRACC. Similar to previous sections, the performances of CROCC and CRACC is analysed for different IVS and degree of endogeneity.

Table 4-10: Arbitrary Values Applied in CRACC

Arbitrary value	Arbitrary Convex Combination
a = 25%	$\text{CRACC1} = 0.25\ \widehat{\boldsymbol{\beta}}\,(\text{MEL}) + (1 - 0.25)\,\widehat{\boldsymbol{\beta}}\,(\text{MEEL})$
a = 50%	$\text{CRACC2} = 0.50\ \widehat{\boldsymbol{\beta}}\,(\text{MEL}) + (1 - 0.50)\,\widehat{\boldsymbol{\beta}}\,(\text{MEEL})$
a = 75%	$\text{CRACC3} = 0.75\ \widehat{\boldsymbol{\beta}}\,(\text{MEL}) + (1 - 0.75)\,\widehat{\boldsymbol{\beta}}\,(\text{MEEL})$

4.6.1 Biasedness of CROCC and CRACC in Exactly Determined Model

Table 4-11 is comparing the biasedness of CROCC and three arbitrary CRACC estimators having a different level of instrumental variable strengths and degree of endogeneity of different sample sizes. CROCC is found better estimators in small samples (25 & 50), especially for the high degree of endogeneity, moderate instrumental variable, and strong instrumental variable. Conversely, for large samples, CRACC performed better in most cases. Therefore, the convex combination of CROCC needs to improve in terms of biasedness to estimate the regression model having an endogeneity problem with large samples.

Table 4-11: Biasedness of CROCC and CRACC in Exactly Determined Model

	Weak Instruments	Moderate Instruments	Strong Instruments
Low Degree	[CROCC< CRACC1,2,3]^{25} [CRACC1< CROCC< CRACC2,3]^{50} [CRACC1< CROCC< CRACC2,3]^{100} [CRACC1< CROCC< CRACC2,3]^{150} [CRACC1< CROCC< CRACC2,3]^{200} [CROCC< CRACC1,2,3]^{250}	[CROCC< CRACC1,2,3]^{25} [CROCC< CRACC1,2,3]^{50} [CRACC1< CROCC< CRACC2,3]^{100} [CRACC1< CROCC< CRACC2,3]^{150} [CROCC< CRACC1,2,3]^{200} [CRACC2< CROCC< CRACC1,3]^{250}	[CROCC< CRACC1,2,3]^{25} [CROCC< CRACC1,2,3]^{50} [CROCC< CRACC1,2,3]^{100} [CRACC2< CROCC< CRACC1,3]^{150} [CRACC2< CROCC< CRACC1,3]^{200} [CRACC2< CROCC< CRACC1,3]^{250}
Moderate Degree	[CROCC< CRACC1,2,3]^{25} [CRACC1< CROCC< CRACC2,3]^{50} [CROCC< CRACC1,2,3]^{100} [CROCC< CRACC1,2,3]^{150} [CRACC1< CROCC< CRACC2,3]^{200} [CRACC1<CRACC2<CROCC< CRACC3]^{250}	[CROCC< CRACC1,2,3]^{25} [CROCC< CRACC1,2,3]^{50} [CROCC< CRACC1,2,3]^{100} [CRACC1< CROCC< CRACC2,3]^{150} [CRACC1<CRACC2<CROCC< CRACC3]^{200} [CRACC1<CRACC2<CROCC< CRACC3]^{250}	[CROCC< CRACC1,2,3]^{25} [CROCC< CRACC1,2,3]^{50} [CRACC1< CROCC< CRACC2,3]^{100} [CRACC1<CRACC2<CROCC< CRACC3]^{150} [CRACC1<CRACC2<CROCC< CRACC3]^{200} [CROCC< CRACC1,2,3]^{250}
High Degree	[CROCC< CRACC1,2,3]^{25} [CROCC< CRACC1,2,3]^{50} [CROCC< CRACC1,2,3]^{100} [CROCC< CRACC1,2,3]^{150} [CROCC< CRACC1,2,3]^{200} [CRACC1<CRACC2<CROCC< CRACC3]^{250}	[CROCC< CRACC1,2,3]^{25} [CROCC< CRACC1,2,3]^{50} [CROCC< CRACC1,2,3]^{100} [CROCC< CRACC1,2,3]^{150} [CRACC1< CROCC< CRACC2,3]^{200} [CRACC1< CROCC< CRACC2,3]^{250}	[CROCC< CRACC1,2,3]^{25} [CROCC< CRACC1,2,3]^{50} [CROCC< CRACC1,2,3]^{100} [CRACC1< CROCC< CRACC2,3]^{150} [CRACC1< CROCC< CRACC2,3]^{200} [CRACC1<CRACC2<CROCC< CRACC3]^{250}
Very High Degree	[CROCC< CRACC1,2,3]^{25} [CROCC< CRACC1,2,3]^{50} [CROCC< CRACC1,2,3]^{100} [CROCC< CRACC1,2,3]^{150} [CROCC< CRACC1,2,3]^{200} [CROCC< CRACC1,2,3]^{250}	[CROCC< CRACC1,2,3]^{25} [CROCC< CRACC1,2,3]^{50} [CROCC< CRACC1,2,3]^{100} [CROCC< CRACC1,2,3]^{150} [CROCC< CRACC1,2,3]^{200} [CRACC1< CROCC< CRACC2,3]^{250}	[CROCC< CRACC1,2,3]^{25} [CROCC< CRACC1,2,3]^{50} [CROCC< CRACC1,2,3]^{100} [CROCC< CRACC1,2,3]^{150} [CROCC< CRACC1,2,3]^{200} [CRACC1< CROCC< CRACC2,3]^{250}

The table is summarising the results of MCS (5000 simulations) for all the possible combinations of the degree of endogeneity, strength of instrumental variables, and finite sample sizes. The most left column is representing the different degrees of endogeneity while the first row is showing different strengths of instrumental variables. In the cells, CROCC and three arbitrary CRACC estimators are ranked according to their Biasedness for each sample size (as given in superscript).

4.6.1 MSE of CROCC and CRACC in Exactly Determined Model

Table 4-12 is comparing the MSE of CROCC and CRACC in the exactly determined model. Results are clearly showing the least MSE of CROCC in all the combinations of instrumental variable strengths, degrees of endogeneity, and sample sizes. Therefore, it can be concluded that CROCC is an efficient estimator in an exactly determined model.

Table 4-12: MSE of CROCC and CRACC in Exactly Determined Model

	Weak Instruments	Moderate Instruments	Strong Instruments
Low Degree	$[CROCC< CRACC_{1,2,3}]^{25}$ $[CROCC< CRACC_{1,2,3}]^{50}$ $[CROCC< CRACC_{1,2,3}]^{100}$ $[CROCC< CRACC_{1,2,3}]^{150}$ $[CROCC< CRACC_{1,2,3}]^{200}$ $[CROCC< CRACC_{1,2,3}]^{250}$	$[CROCC< CRACC_{1,2,3}]^{25}$ $[CROCC< CRACC_{1,2,3}]^{50}$ $[CROCC< CRACC_{1,2,3}]^{100}$ $[CROCC< CRACC_{1,2,3}]^{150}$ $[CROCC< CRACC_{1,2,3}]^{200}$ $[CROCC< CRACC_{1,2,3}]^{250}$	$[CROCC< CRACC_{1,2,3}]^{25}$ $[CROCC< CRACC_{1,2,3}]^{50}$ $[CROCC< CRACC_{1,2,3}]^{100}$ $[CROCC< CRACC_{1,2,3}]^{150}$ $[CROCC< CRACC_{1,2,3}]^{200}$ $[CROCC< CRACC_{1,2,3}]^{250}$
Moderate Degree	$[CROCC< CRACC_{1,2,3}]^{25}$ $[CROCC< CRACC_{1,2,3}]^{50}$ $[CROCC< CRACC_{1,2,3}]^{100}$ $[CROCC< CRACC_{1,2,3}]^{150}$ $[CROCC< CRACC_{1,2,3}]^{200}$ $[CROCC< CRACC_{1,2,3}]^{250}$	$[CROCC< CRACC_{1,2,3}]^{25}$ $[CROCC< CRACC_{1,2,3}]^{50}$ $[CROCC< CRACC_{1,2,3}]^{100}$ $[CROCC< CRACC_{1,2,3}]^{150}$ $[CROCC< CRACC_{1,2,3}]^{200}$ $[CROCC< CRACC_{1,2,3}]^{250}$	$[CROCC< CRACC_{1,2,3}]^{25}$ $[CROCC< CRACC_{1,2,3}]^{50}$ $[CROCC< CRACC_{1,2,3}]^{100}$ $[CROCC< CRACC_{1,2,3}]^{150}$ $[CROCC< CRACC_{1,2,3}]^{200}$ $[CROCC< CRACC_{1,2,3}]^{250}$
High Degree	$[CROCC< CRACC_{1,2,3}]^{25}$ $[CROCC< CRACC_{1,2,3}]^{50}$ $[CROCC< CRACC_{1,2,3}]^{100}$ $[CROCC< CRACC_{1,2,3}]^{150}$ $[CROCC< CRACC_{1,2,3}]^{200}$ $[CROCC< CRACC_{1,2,3}]^{250}$	$[CROCC< CRACC_{1,2,3}]^{25}$ $[CROCC< CRACC_{1,2,3}]^{50}$ $[CROCC< CRACC_{1,2,3}]^{100}$ $[CROCC< CRACC_{1,2,3}]^{150}$ $[CROCC< CRACC_{1,2,3}]^{200}$ $[CROCC< CRACC_{1,2,3}]^{250}$	$[CROCC< CRACC_{1,2,3}]^{25}$ $[CROCC< CRACC_{1,2,3}]^{50}$ $[CROCC< CRACC_{1,2,3}]^{100}$ $[CROCC< CRACC_{1,2,3}]^{150}$ $[CROCC< CRACC_{1,2,3}]^{200}$ $[CROCC< CRACC_{1,2,3}]^{250}$
Very High Degree	$[CROCC< CRACC_{1,2,3}]^{25}$ $[CROCC< CRACC_{1,2,3}]^{50}$ $[CROCC< CRACC_{1,2,3}]^{100}$ $[CROCC< CRACC_{1,2,3}]^{150}$ $[CROCC< CRACC_{1,2,3}]^{200}$ $[CROCC< CRACC_{1,2,3}]^{250}$	$[CROCC< CRACC_{1,2,3}]^{25}$ $[CROCC< CRACC_{1,2,3}]^{50}$ $[CROCC< CRACC_{1,2,3}]^{100}$ $[CROCC< CRACC_{1,2,3}]^{150}$ $[CROCC< CRACC_{1,2,3}]^{200}$ $[CROCC< CRACC_{1,2,3}]^{250}$	$[[CROCC< CRACC_{1,2,3}]^{25}$ $[CROCC< CRACC_{1,2,3}]^{50}$ $[CROCC< CRACC_{1,2,3}]^{100}$ $[CROCC< CRACC_{1,2,3}]^{150}$ $[CROCC< CRACC_{1,2,3}]^{200}$ $[CROCC< CRACC_{1,2,3}]^{250}$

The table is summarising the results of MCS (5000 simulations) for all the possible combinations of the degree of endogeneity, strength of instrumental variables, and finite sample sizes. The most left column is representing the different degrees of endogeneity while the first row is showing different strengths of instrumental variables. In the cells, CROCC and three arbitrary CRACC estimators are ranked according to their MSE for each sample size (as given in superscript).

4.6.2 Biasedness of CROCC and CRACC in Overdetermined Model

Therefore, it is concluded that, in general, CROCC produces more biasedness than CRACC. Table 4-13 compares the biasedness of CROCC and CRACC for all the combinations of IVS, endogeneity, and sample sizes in an overdetermined model having four instruments. Results are showing that CRACC exhibited the least biasedness than CROCC when the weak or moderate instrumental variable is used. However, for a strong instrumental variable, the CROCC performed better in some cases. Therefore, it is concluded that, in general, CROCC produces more biasedness than CRACC.

Table 4-13: Biasedness of CROCC and CRACC in Overdetermined Model

	Weak Instruments	Moderate Instruments	Strong Instruments
Low Degree	$[CRACC_{1,2,3} < CROCC]^{25}$ $[CRACC_{1,2,3} < CROCC]^{50}$ $[CRACC_{1,2,3} < CROCC]^{100}$ $[CRACC_{1,2,3} < CROCC]^{150}$ $[CRACC_{1,2,3} < CROCC]^{200}$ $[CRACC_{1,2,3} < CROCC]^{250}$	$[CRACC_{1,2,3} < CROCC]^{25}$ $[CRACC_{1,2,3} < CROCC]^{50}$ $[CRACC_{1,2,3} < CROCC]^{100}$ $[CRACC_3 < CRACC_2 < CROCC < CRACC_1]$ $[CROCC < CRACC_{1,2,3}]^{200}$ $[CROCC < CRACC_{1,2,3}]^{250}$	$[CROCC < CRACC_{1,2,3}]^{25}$ $[CROCC < CRACC_{1,2,3}]^{50}$ $[CROCC < CRACC_{1,2,3}]^{100}$ $[CROCC < CRACC_{1,2,3}]^{150}$ $[CROCC < CRACC_{1,2,3}]^{200}$ $[CROCC < CRACC_{1,2,3}]^{250}$
Moderate Degree	$[CRACC_{1,2,3} < CROCC]^{25}$ $[CRACC_{1,2,3} < CROCC]^{50}$ $[CRACC_{1,2,3} < CROCC]^{100}$ $[CRACC_{1,2,3} < CROCC]^{150}$ $[CRACC_{1,2,3} < CROCC]^{200}$ $[CRACC_{1,2,3} < CROCC]^{250}$	$[CRACC_{1,2,3} < CROCC]^{25}$ $[CRACC_{1,2,3} < CROCC]^{50}$ $[CRACC_{1,2,3} < CROCC]^{100}$ $[CRACC_{1,2,3} < CROCC]^{150}$ $[CRACC_{1,2,3} < CROCC]^{200}$ $[CRACC_{1,2,3} < CROCC]^{250}$	$[CRACC_{1,2,3} < CROCC]^{25}$ $[CRACC_{1,2,3} < CROCC]^{50}$ $[CROCC < CRACC_{1,2,3}]^{100}$ $[CROCC < CRACC_{1,2,3}]^{150}$ $[CROCC < CRACC_{1,2,3}]^{200}$ $[CROCC < CRACC_{1,2,3}]^{250}$
High Degree	$[CRACC_{1,2,3} < CROCC]^{25}$ $[CRACC_{1,2,3} < CROCC]^{50}$ $[CRACC_{1,2,3} < CROCC]^{100}$ $[CRACC_{1,2,3} < CROCC]^{150}$ $[CRACC_{1,2,3} < CROCC]^{200}$ $[CRACC_{1,2,3} < CROCC]^{250}$	$[CRACC_{1,2,3} < CROCC]^{25}$ $[CRACC_{1,2,3} < CROCC]^{50}$ $[CRACC_{1,2,3} < CROCC]^{100}$ $[CRACC_{1,2,3} < CROCC]^{150}$ $[CRACC_{1,2,3} < CROCC]^{200}$ $[CRACC_{1,2,3} < CROCC]^{250}$	$[CRACC_{1,2,3} < CROCC]^{25}$ $[CRACC_{1,2,3} < CROCC]^{50}$ $[CRACC_{1,2,3} < CROCC]^{100}$ $[CRACC_{1,2,3} < CROCC]^{150}$ $[CROCC < CRACC_{1,2,3}]^{200}$ $[CROCC < CRACC_{1,2,3}]^{250}$
Very High Degree	$[CRACC_{1,2,3} < CROCC]^{25}$ $[CRACC_{1,2,3} < CROCC]^{50}$ $[CRACC_{1,2,3} < CROCC]^{100}$ $[CRACC_{1,2,3} < CROCC]^{150}$ $[CRACC_{1,2,3} < CROCC]^{200}$ $[CRACC_{1,2,3} < CROCC]^{250}$	$[CRACC_{1,2,3} < CROCC]^{25}$ $[CRACC_{1,2,3} < CROCC]^{50}$ $[CRACC_{1,2,3} < CROCC]^{100}$ $[CRACC_{1,2,3} < CROCC]^{150}$ $[CRACC_{1,2,3} < CROCC]^{200}$ $[CRACC_{1,2,3} < CROCC]^{250}$	$[CRACC_{1,2,3} < CROCC]^{25}$ $[CRACC_{1,2,3} < CROCC]^{50}$ $[CRACC_{1,2,3} < CROCC]^{100}$ $[CRACC_{1,2,3} < CROCC]^{150}$ $[CRACC_1 < CROCC < CRACC_{2,3}]^{200}$ $[CROCC < CRACC_{1,2,3}]^{250}$

The table is summarising the results of MCS (5000 simulations) for all the possible combinations of the degree of endogeneity, strength of instrumental variables, and finite sample sizes. The most left column is representing the different degrees of endogeneity while the first row is showing different strengths of instrumental variables. In the cells, all the estimators are ranked according to their Biasedness for each sample size (as given in superscript).

MSE of CROCC and CRACC in Overdetermined Model in Table 4-14 compares the MSE of CROCC and CRACC in the overdetermined model. It is found that CROCC is an efficient estimator for very small samples (25) in all the combinations. Similarly, CROCC is an efficient estimator than CROCC when the instrument variable is weak, and the degree of endogeneity is low or moderate. However, in general, CRACC performed better than CROCC for a moderate degree of endogeneity, a high degree of endogeneity, and strong instrumental variables.

Table 4-14: MSE of CROCC and CRACC in Overdetermined Model

	Weak Instruments	Moderate Instruments	Strong Instruments
Low Degree	$[CROCC < CRACC_{1,2,3}]^{25}$ $[CROCC < CRACC_{1,2,3}]^{50}$ $[CROCC < CRACC_{1,2,3}]^{100}$ $[CROCC < CRACC_{1,2,3}]^{150}$ $[CROCC < CRACC_{1,2,3}]^{200}$ $[CROCC < CRACC_{1,2,3}]^{250}$	$[CROCC < CRACC_{1,2,3}]^{25}$ $[CROCC < CRACC_{1,2,3}]^{50}$ $[CROCC < CRACC_{1,2,3}]^{100}$ $[CRACC_2 < CROCC < CRACC_{1,3}]^{150}$ $[CRACC_2 < CROCC < CRACC_{1,3}]^{200}$ $[CRACC_2 < CROCC < CRACC_{1,3}]^{250}$	$[CROCC < CRACC_{1,2,3}]^{25}$ $[CRACC_2 < CROCC < CRACC_{1,3}]^{50}$ $[CRACC_2 < CROCC < CRACC_{1,3}]^{100}$ $[CRACC_2 < CROCC < CRACC_{1,3}]^{150}$ $[CRACC_2 < CROCC < CRACC_{1,3}]^{200}$ $[CRACC_2 < CROCC < CRACC_{1,3}]^{250}$
Moderate Degree	$[CROCC < CRACC_{1,2,3}]^{25}$ $[CROCC < CRACC_{1,2,3}]^{50}$ $[CROCC < CRACC_{1,2,3}]^{100}$ $[CROCC < CRACC_{1,2,3}]^{150}$ $[CROCC < CRACC_{1,2,3}]^{200}$ $[CROCC < CRACC_{1,2,3}]^{250}$	$[CROCC < CRACC_{1,2,3}]^{25}$ $[CROCC < CRACC_{1,2,3}]^{50}$ $[CROCC < CRACC_{1,2,3}]^{100}$ $[CRACC_2 < CROCC < CRACC_{1,3}]^{150}$ $[CRACC_2 < CROCC < CRACC_{1,3}]^{200}$ $[CRACC_2 < CROCC < CRACC_{1,3}]^{250}$	$[CROCC < CRACC_{1,2,3}]^{25}$ $[CRACC_2 < CROCC < CRACC_{1,3}]^{50}$ $[CRACC_2 < CROCC < CRACC_{1,3}]^{100}$ $[CRACC_2 < CROCC < CRACC_{1,3}]^{150}$ $[CRACC_2 < CROCC < CRACC_{1,3}]^{200}$ $[CRACC_2 < CROCC < CRACC_{1,3}]^{250}$
High Degree	$[CROCC < CRACC_{1,2,3}]^{25}$ $[CROCC < CRACC_{1,2,3}]^{50}$ $[CRACC_2 < CROCC < CRACC_{1,3}]^{100}$ $[CRACC_2 < CROCC < CRACC_{1,3}]^{150}$ $[CRACC_2 < CROCC < CRACC_{1,3}]^{200}$ $[CRACC_2 < CROCC < CRACC_{1,3}]^{250}$	$[CROCC < CRACC_{1,2,3}]^{25}$ $[CRACC_2 < CROCC < CRACC_{1,3}]^{50}$ $[CRACC_2 < CROCC < CRACC_{1,3}]^{100}$ $[CRACC_2 < CROCC < CRACC_{1,3}]^{150}$ $[CRACC_2 < CROCC < CRACC_{1,3}]^{200}$ $[CRACC_2 < CROCC < CRACC_{1,3}]^{250}$	$[CROCC < CRACC_{1,2,3}]^{25}$ $[CRACC_2 < CROCC < CRACC_{1,3}]^{50}$ $[CRACC_2 < CROCC < CRACC_{1,3}]^{100}$ $[CRACC_2 < CROCC < CRACC_{1,3}]^{150}$ $[CRACC_2 < CROCC < CRACC_{1,3}]^{200}$ $[CRACC_2 < CROCC < CRACC_{1,3}]^{250}$
Very High Degree	$[CROCC < CRACC_{1,2,3}]^{25}$ $[CRACC_2 < CRACC_3 < CROCC < CRACC_1]^{50}$ $[CRACC_{1,2,3} < CROCC]^{100}$ $[CRACC_{1,2,3} < CROCC]^{150}$ $[CRACC_{1,2,3} < CROCC]^{200}$ $[CRACC_{1,2,3} < CROCC]^{250}$	$[CRACC_2 < CROCC < CRACC_{1,3}]^{25}$ $[CRACC_{1,2,3} < CROCC]^{50}$ $[CRACC_2 < CROCC < CRACC_{1,3}]^{100}$ $[CRACC_2 < CROCC < CRACC_{1,3}]^{150}$ $[CRACC_2 < CROCC < CRACC_{1,3}]^{200}$ $[CRACC_2 < CROCC < CRACC_{1,3}]^{250}$	$[CROCC < CRACC_{1,2,3}]^{25}$ $[CRACC_2 < CROCC < CRACC_{1,3}]^{50}$ $[CRACC_2 < CROCC < CRACC_{1,3}]^{100}$ $[CRACC_2 < CROCC < CRACC_{1,3}]^{150}$ $[CRACC_2 < CROCC < CRACC_{1,3}]^{200}$ $[CRACC_2 < CROCC < CRACC_{1,3}]^{250}$

The table is summarising the results of MCS (5000 simulations) for all the possible combinations of the degree of endogeneity, strength of instrumental variables, and finite sample sizes. The most left column is representing the different degrees of endogeneity while the first row is showing different strengths of instrumental variables. In the cells, all the estimators are ranked according to their MSE for each sample size (as given in superscript).

4.6.3 Summary of Results

The comparative analysis of CROCC and CRACC provided two useful implications. First, CROCC is found an efficient estimator than CRACC in case of an exactly determined model. It is also noteworthy that CROCC was an efficient estimator than GMM, MEL, and MEEL in an exactly determined model, as explored in the previous section. Therefore, future studies are recommended to use CROCC in estimating a regression model having the problem of endogeneity in an exactly determined model. Second, CROCC performed less efficient than CRACC in most cases in an overdetermined model that alludes to the deficiency of CROCC. Therefore, it is suggested that a convex combination of CROCC can be improved by taking arbitrary values to produce efficient and less biased estimations.

CHAPTER 5. EMPIRICAL ANALYSIS

This dissertation also applied the MEL, MEEL, GMM, and CROCC to estimate economic models on real data set. Since the focus of this research is to investigate the estimators providing less biased and efficient estimations to resolve the endogeneity problem. Therefore economic models having endogeneity in their system will be suitable for analysing the application of studied estimators. Consumption function and money demand function are two famous models that can help to investigate the implication of studied estimators for real data set. Previously, various studies have confirmed the existence of an endogeneity problem for these two economic models (Charemza & Deadman, 1997; Thomas, 1993). To deal with such an endogeneity problem, GMM is often used.

In practice, economic series for such models have limited or small samples. While the GMM estimator produces biased and less efficient results for small samples as explored in the previous section. Similarly, this dissertation proposed that applying information-theoretic approaches can produce less biased and efficient estimations, even for small samples comparatively. Therefore, it can be useful to apply an information-theoretic approach to estimate the models having endogeneity, especially with small samples. To conclude, these propositions in the case of real data set, MEL, MEEL, GMM, CROCC, and CRACC2 (CRACC2 is used because it was found efficient in most cases) will be used to estimate consumption function and money demand function. If the results confirm the superiority of information-theoretic models, then it is a new direction for researchers to make the inference better, especially for a small sample. The subsequent part will compare the performances of estimators for both consumption function and money demand function separately.

5.1 Consumption Function and Information-Theoretic Approach

This section will estimate the consumption function as proposed by Keynes (1936) and also referred to as an absolute income hypothesis (AIH). The consumption function is as follow:

$$C_t = a + bY_t + \varepsilon_t \qquad 5\text{-}1$$

$$Y_t = C_t + I_t \qquad 5\text{-}2$$

In the above-stated system of equations, "C_t" is household consumption, "Y_t" is per capita income, and "I_t" is the investment. The equation "5-1" has a disturbance term. "ε_t" is known as the Keynesian consumption function. On the other hand, we do not need to estimate the equation "5-2", because this is aggregate demand equation and the coefficient of the consumption and investment are equal to unity as stated in (Thomas, 1993). The reduced form of the above consumption system in (section: 3) has shown that "C_t" and "Y_t" are endogenous variables, and "I_t" is the only exogenous variable. To solve the equation "5-1", investment "I_t" and government expenditures "G" are used as instrumental variables.

The Instrumental variables matrix is as below:

$$Z = [\text{ constant Investment Government expenditures}] = [\text{ones I G}]$$

This system of equations has simultaneity bias. Ordinary least square (OLS) estimator is not suitable to estimate this model because of the simultaneity bias. As an alternative, the GMM estimator can estimate such models. However, GMM does not hold good finite samples properties reported by numerous authors[9] in literature. This research proposed that the information-theoretic approach holds such good finite sample properties that can estimate the model has simultaneity bias more accurately. Therefore, GMM and information-theoretic estimators are applied to the consumption function for the purpose of comparison. The data set regarding the above equations of consumption function is collected from the International Financial Statistics (IFS) database with a sample size of less than "30".

In consumption function, the slope coefficient of income is known as marginal propensity to consume (MPC), whose range is between "0 < MPC < 1" in absolute measures. However, we have used the log transformation (semi-log model) where the coefficient of income is known as elasticity of consumption with respect to income. The consumption function (semi-log model) is estimated, and the coefficients of endogenous regressor (Elasticity of consumption) are plotted in Figure 5-1.

The vertical axis of Figure 5-1 represents the elasticity, while the horizontal axis is denoted with countries. Results showed that the estimated elasticity obtained from the GMM estimator is greater than the estimates of information-theoretic estimators. For instance, in the case of Australia, the GMM estimator showed that the elasticity of unity, while all the information-theoretic estimators documented the elasticity around "0.9" for the same country. Similarly, for Austria, the GMM showed the elasticity of greater than unity. Conversely, the elasticity measured by all the information-theoretic estimators is less than unity (around 0.85). In the case of New Zealand, the analysis

---

[9] The GMM estimator have poor finite samples properties, in this regard various authors (Altonji & Segal, 1996; Hall & Horowitz, 1996; Imbens et al., 1998; Imbens, 1997; Kitamura & Stutzer, 1997; Qin & Lawless, 1994) have reported the number of problems of GMM estimator.

showed more differences between the estimate of GMM (Elasticity equal 0.98) and information-theoretic estimators (Elasticity around 0.6). Overall results are alluding that the estimates of information-theoretic estimators are very close to each other while the GMM estimates are far away from them.

Figure 5-1: Estimates of the Coefficients of Endogenous Regressor Consumption Function

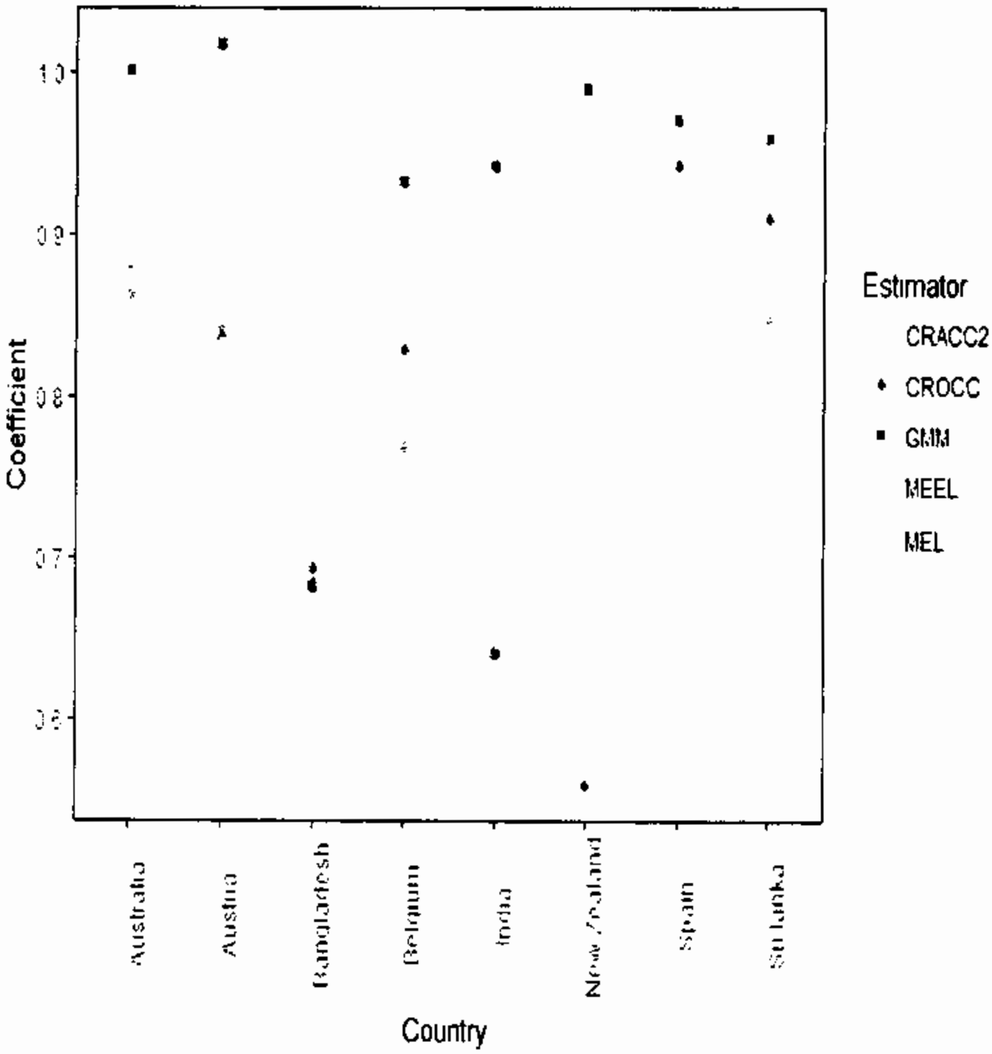


Table 5-1: Coefficients of Endogenous Regressor of Consumption Function

Countries	MEL	MEEL	GMM	CROCC	CRACC2
Australia	0.9039	0.86243	1.0011	0.88217	0.883166
Austria	0.82243	0.84207	1.01731	0.83641	0.832249
Bangladesh	0.75455	0.68421	0.68085	0.6927	0.719382
Belgium	0.85588	0.76893	0.93221	0.82829	0.812404
India	0.76346	0.6423	0.94176	0.63896	0.702876
New Zealand	0.55905	0.65	0.98931	0.55905	0.604525
Spain	0.9514	0.91881	0.97032	0.94214	0.935105
Sri Lanka	0.91957	0.84658	0.95835	0.91007	0.883077

5.2 Money Demand Function and Information-Theoretic Approach

The money demand function is also estimated using GMM and information-theoretic estimators. It was Keynes (1936) who proposed the idea of liquidity preference theory against the quantity theory of money (QTM). In his model, he explained the role of interest rate with respect to supply and demand for money. According to him, the demand for money can be segregated into three parts, i.e., precautionary demand for money, transaction demand for money, and speculative demand for money. Later on, various authors (Friedman, 1956; Tobin, 1958)[10] presented different ideas about money demand functions. In this concern, (Thomas, 1993) discussed different theories about money demand functions and also stated the problem of simultaneity bias in the estimation of the money demand function. The money demand function and its simultaneity bias also have been discussed in the previous section "3". Due to a special

[10] (Friedman, 1956) presented the reformulation in quantity theory of money.
(Tobin, 1958) discussed the idea of non-human wealth in money demand function and presented the portfolio diversification.

case of simultaneity bias, this study selected the money demand function for empirical analysis. The money demand function can be stated as follow:

$$\mathbf{m}_d = b_o + b_1\mathbf{r} + b_2\mathbf{y} + \varepsilon_1$$

$$\mathbf{m}_s = a_o\lambda + a_1 r + h + \varepsilon_2$$

Equilibrium condition

$$\mathbf{m}_d = \mathbf{m}_s = m$$

In this system, the rate of interest "r" and money supply "M" (money supply M2) are endogenous variables. On the other side, high-powered money "h" and permanent income "Y" are exogenous variables. If we apply the OLS to estimate the money demand function, we will face the biased parameters. One can apply GMM to estimate the above money demand function in the presence of simultaneity bias. However, as noted earlier by Monte Carlo simulation analysis and various studies that the GMM approach has large finite samples bias[11] than alternative estimators (information-theoretic approach)[12], GMM could provide biased results for a small sample.

Therefore, the above money demand function is estimated using MEL, MEEL, CROCC, and GMM approaches to compare their results in the real data set. We have estimated the money demand function for several countries and collect data form the World Bank database with sample size "30". The results are discussed as follows.

---

[11] Various authors (Altonji & Segal, 1996; Hall & Horowitz, 1996; Pagan & Robertson, 1997) discussed number of problems of GMM estimator and its poor finite samples properties.
[12] These authors (Imbens et al., 1998; Guido W Imbens, 1997; Kitamura & Stutzer, 1997; Qin & Lawless, 1994) presented alternative approaches over GMM approach to avoid small samples biasedness.

Table 5-2: Money Demand Function and Information-Theoretic Approach

Country	Estimators & Coefficient of Endogenous Regressor				
	MEL	MEEL	GMM	CROCC	CRACC2
Australia	-0.3658	-0.30779	-0.40061	-0.33667	-0.3368
Bahamas	-1.13306	-0.86625	-0.39599	-0.95958	-0.99966
Barbados	-0.49286	-0.42026	-0.37214	-0.45014	-0.45656
Belize	-1.24919	-1.18611	-1.83796	-1.2167	-1.21765
Canada	-1.10219	-1.12506	-1.15918	-1.11413	-1.11363
Dominica	-0.95018	-0.98642	-1.22315	-0.96792	-0.9683
Fiji	-0.24042	-0.19742	-0.23783	-0.21703	-0.21892
Grenada	-1.19684	-0.75252	-6.93384	-0.94935	-0.97468
Jamaica	-1.13985	-0.99432	-1.2073	-1.05989	-1.06709
Japan	-0.04571	-0.05048	-0.03408	-0.04784	-0.0481
Kenya	-0.04695	-0.0572	-0.13197	-0.05196	-0.05208
Lesotho	-2.22421	-2.40153	-3.04603	-2.37353	-2.31287
Malawi	-0.71861	-0.67078	-1.2893	-0.69352	-0.6947
Papua New Guinea	-1.16703	-1.03652	1.27666	-1.10821	-1.10178
Philippines	-0.01764	-0.00811	-0.00779	-0.01267	-0.01288
Sierra Leone	-1.11315	-1.00566	-1.62164	-1.06123	-1.05941
Singapore	-0.99924	-0.95047	-1.01302	-0.97321	-0.97486
Solomon Islands	-1.85608	-1.78699	-0.55372	-1.82001	-1.82154
South Africa	-1.13882	-1.08208	-0.34012	-1.10637	-1.11045
Sri Lanka	-1.23509	-1.18308	-2.19007	-1.26018	-1.20909
St. Kitts and Nevis	-5.82938	-5.98211	5.323506	-5.89038	-5.90575
St. Lucia	-1.06729	-1.16184	-0.65	-1.11103	-1.11457
St. Vincent and Grenadines	-1.26051	-1.11241	0.079798	-1.18891	-1.18646
Swaziland	-0.92946	-0.99515	-1.39359	-0.96088	-0.96231
Uganda	-0.98117	-0.97035	-1.43159	-0.97573	-0.97576

In Figure 5-2, we have reported the coefficients of the endogenous regressor (interest rate/ coefficient in terms of elasticity) for several countries. According to the theory, the interest rate is inversely related to the money demand (negative elasticity). Still, GMM estimators have shown the positive elasticity of "interest rate" for three countries (Papua New Guinea, St. Kitts & Nevis, and St. Vincent & the Grenadines). Conversely, GMM showed a very high negative elasticity than information-theoretic estimators in the case of "Grenada." In the overall analysis, it is found that the estimates of GMM are away from the information-theoretic estimators.

Figure 5-2: Estimates of the Coefficients of Endogenous Regressor of Money Demand Function

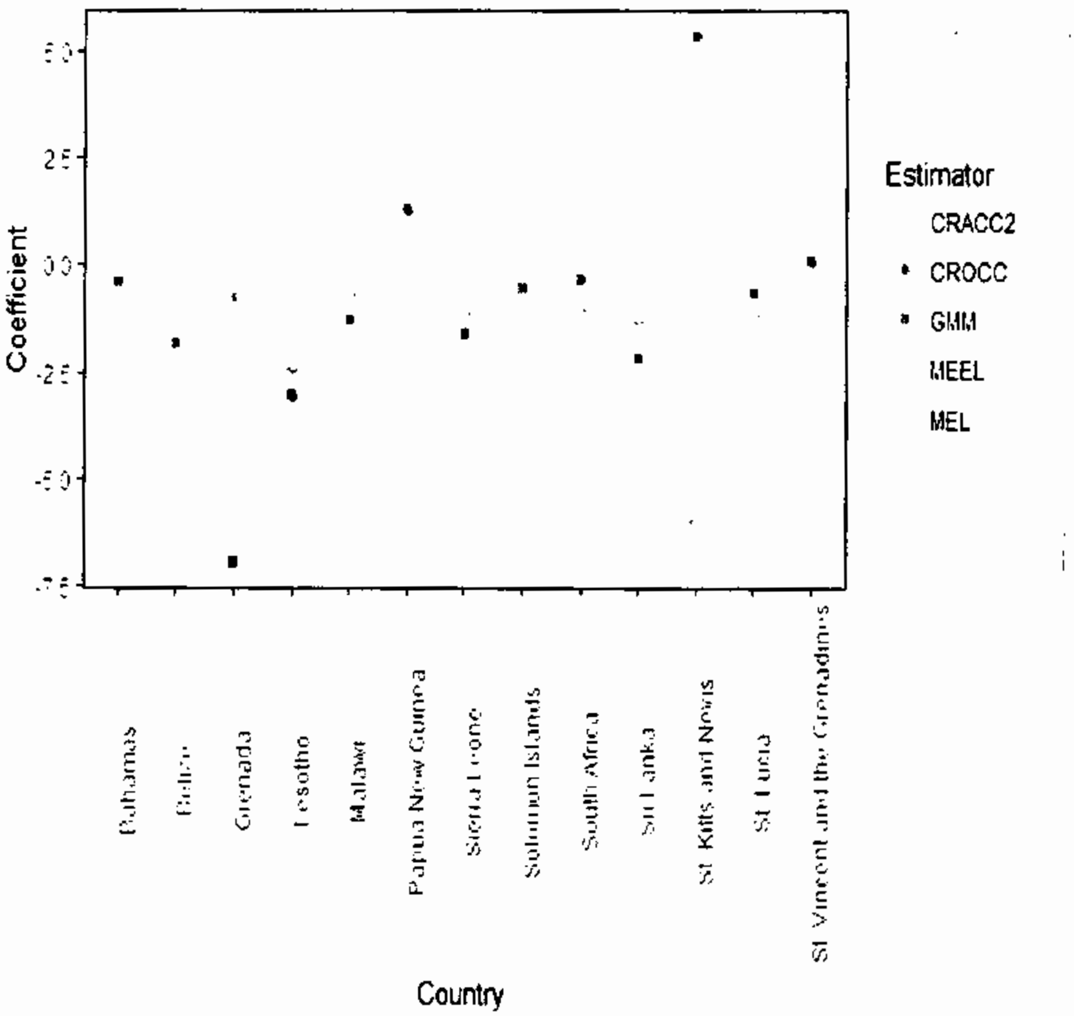
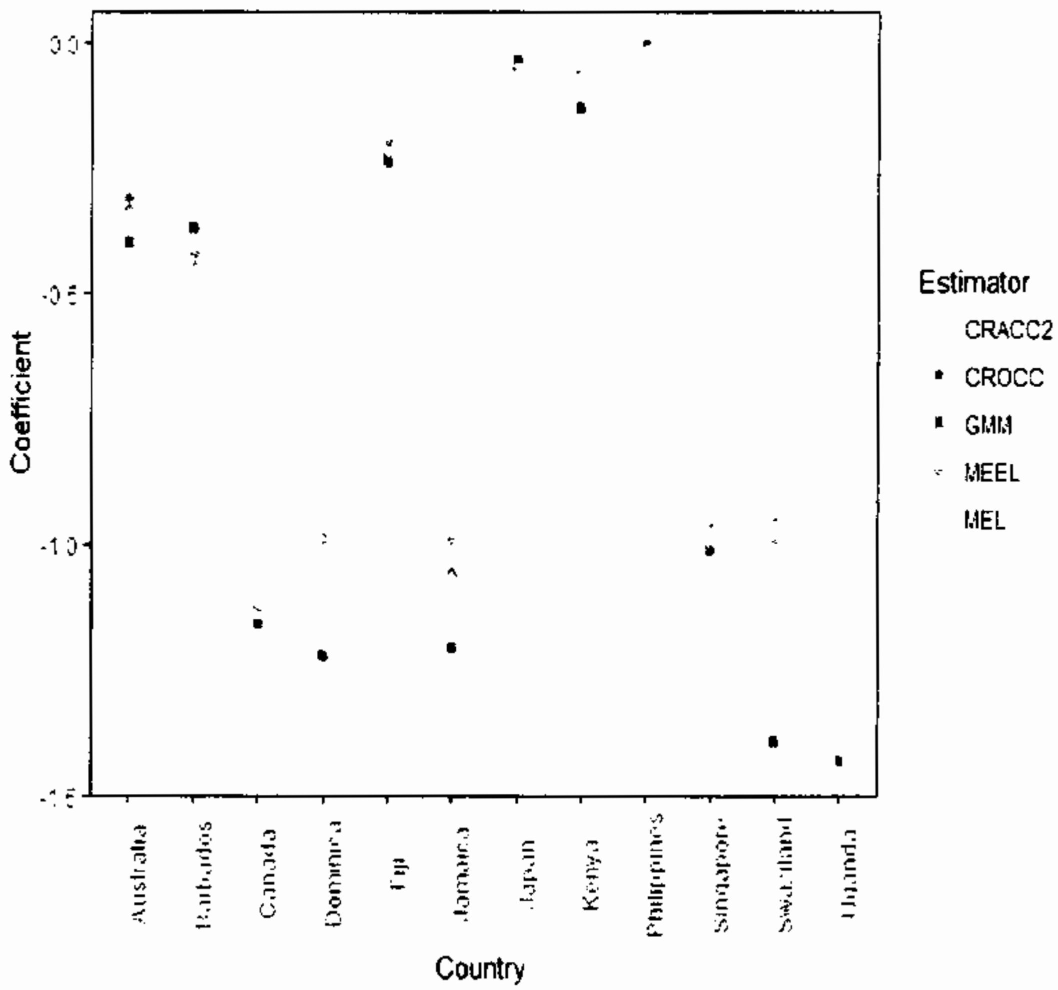


Figure 5-3 also demonstrated that all the estimators have negative elasticity as per the theoretic relationship between "interest rate" and "money demand." In this analysis, GMM still showed elasticity greater than the unity of the countries (Dominica, Swaziland, and Uganda), while for some countries, information-theoretic estimators documented the elasticity less than unity. In some cases, all estimators are also very close to each other. However, in most of the cases, the elasticity calculated by GMM is away from the information-theoretic estimators.

Figure 5-3: Estimates of the Coefficients of Endogenous Regressor of Money Demand Function



5.3 Summary of Empirical Analysis

This chapter applied the GMM and information-theoretic approaches to estimate money demand and consumption function. All the information-theoretic estimators provided similar estimations of both economic models. However, the estimations of GMM were quite far from the estimations of information-theoretic estimators. These results confirm the reliability and efficiencies of information-theoretic estimators than GMM. Hence, future studies should use information-theoretic estimators to estimate the economic models having the problem of endogeneity.

CHAPTER 6. CONCLUSION & RECOMMENDATIONS

The wide range of data sampling experiment (i.e., based on above discussed Monte Carlo Simulation analysis) is used to analyse and compare the finite sample properties of GMM and information-theoretic estimators (MEL, MEEL, and CROCC) in estimating the linear regression model with non-orthogonal condition. The findings of the thesis are entirely based on Monte Carlo simulation Experiment.

The entire analysis is segregated into four parts. This study primarily compared the bias and Mean Square Error (MSE) of GMM and information theoretic estimators including MEL, MEEL and CROCC using different sample sizes with a variety of degrees of endogeneity and strength of instrumental variables. First, the biases and Mean Square Error (MSE) of GMM and information theoretic estimators are compared in exactly determined model. Overall analysis found that CROCC and MEL produced the least biasedness than GMM estimator in exactly determined model. But CROCC produced the least MSE of all other estimators (indicates CROCC is most efficient estimator than other) in all conditions for exactly determined model.

Secondly, The Monte Carlo simulation analysis is conducted for overdetermined models, where the number of moment conditions IV's exceeds from the number of unknown parameters. In this case, GMM is most commonly used in estimation of econometric models, but the performance of GMM estimator is based on the choice of the weight matrix. In contrast, information theoretic estimators give us unique estimates in overdetermined models. Because these estimators handled unknown weight matrix within the estimation procedure. In this case, the information theoretic estimators are superior to GMM. In this part of the thesis, we are also analysed the biasedness and MSE's of GMM and information theoretic estimators within overdetermined model. In general, the MEL and MEEL documented lower bias than the GMM and even CROCC in an overdetermined model. For a small sample of 25 observations, GMM produced the highest bias than other estimators. Conversely, CROCC is found most efficient estimator when MSE of all the estimators are compared. Therefore, it is recommended to use CROCC in future research studies rather than using GMM and other individual estimators of information theoretic approach in order to obtain efficient estimates of regression model having the problem of endogeneity in finite samples.

In third part, we introduced an alternative method (i.e. Cressie and Read Arbitrary Convex Combination - CRACC) based of arbitrary convex combination that is independent from the optimization of the coefficient of convex combination of CROCC. The proposed method (CRACC) is efficient and applicable than CROCC due to its assumption free estimation of convex combination. The results evidenced that CROCC is a more efficient estimator than CRACC in case of exactly determined model. However, CROCC performed less efficient than CRACC in most cases in an overdetermined model that alludes the deficiency of CROCC. Therefore, it is suggested that convex combination of CROCC can be improved by taking arbitrary values to produce efficient and less biased estimations.

Overall simulation analysis, we examine a range of data generating processes and found consider able conclusions. As the effectiveness of the instruments decreases, and the degree of endogeneity increases the performance of all the estimators decreases in term of biasedness, but information-theoretic estimator are still better than GMM in small samples scenario. Similarly, the MSE of all the estimators increases but convex combination still remains superior to all. Conversely, with any type of the degree of endogeneity with the strong instrumental variable strength, Information theoretic estimator are better than GMM[13], especially convex combinations.

In general, all the estimators improve their performance (shows consistency) as the sample size increases. However, the most of the cases, the Information-theoretic type estimators demonstrate less bias than the traditional GMM estimator. In terms of MSE, the Information-theoretic type are usually superior but, convex combinations are more better than individual estimators (MEL and MEEL) and GMM.

Finally in the last part; GMM and information theoretic estimators are applied to estimate money demand and consumption function. It is found that elasticity estimated by information theoretic estimators were close to each other. However, the elasticity produced by GMM are quite far from the estimations of information theoretic estimators. Therefore, it is concluded that GMM is the less reliable and less efficient estimator in finite samples.

This research demonstrates significant implications in the estimation of econometric models. As discussed earlier; the endogeneity problem and finite samples are key issues in econometric models. Our findings would help the researchers to use aproperiate estimators to resolve endogeneity issues in finite samples accoding to their estimating model. The newly proposed estimator: CRACC which has superiorty to CROCC due to its assumption free estimation of two independent samples. It would help the

---

[13] Varioys author reported that the GMM approach has poor finite samples properties of GMM and they suggested that The EL approach is an alternative to GMM (Altonji & Segal, 1996; Hall & Horowitz, 1996; Imbens et al., 1998; Imbens, 1997; Kitamura & Stutzer, 1997; Qin & Lawless, 1994).
The authors (Imbens et al., 1998; Kitamura & Stutzer, 1997) proposed an alternative approach that is exponential tilting (i.e., Kullback-Leibler Information Criterion (KLIC) / Maximum Exponential Empirical Likelihood (MEEL)) as replacement of the GMM approach.

researchers to use best and appropriate estimator. It would open up another research stream in current literature.

6.1 Limitations of the study

There are following limitations of this study.

Our findings are limited to mentioned Monte Carlo simulation anslysis. The used estimators can be evaluated on different type of DGP. Similarly, there are many other nonparametric approaches that can also be evaluated with Information theoretic estimators and GMM. We have just focused to endogeneity issue, but it can be enhance to heteroskedasticity and autocorrelation.

6.2 Policy Recommendation and Future Research

The results provide strong policy implications for the estimation of econometric models having the problem of endogeneity. Since GMM does not hold better finite sample properties, therefore, future studies are recommended to use information-theoretic estimators rather than GMM, especially for small samples. It is suggested that researchers should use CROCC to estimate economic models for the exactly determined system. However, for an overdetermined model, CRACC is recommended to estimate more efficient coefficients. There are some recommendations for future research, as follows:

- The efficiency of CROCC and CRACC can be improved by developing efficient and assumption-free optimization method of convex combination.
- In this study, we used two individual information-theoretic estimators (MEL and MEEL) to construct the convex combinations. Therefore future studies are

recommended to evaluate the efficiency of convex combination based on more than two individual information-theoretic estimators.

- In this study, we used four instrumental variables for one endogenous regressor in the case of the overdetermined model. The future study can evaluate the efficiency of CRACC and CROCC by varying the number of instrumental variables in the overdetermined model.
- The scope of this study was limited to the endogeneity problem. Information-theoretic estimators should also be evaluated in the presence of Heteroskedasticity and Autocorrelation.

References

Altonji, J. G., & Segal, L. M. (1996). Small-Sample Bias in GMM Estimation of Covariance Structures. *Journal of Business & Economic Statistics*, *14*(3), 353–366. https://doi.org/10.1080/07350015.1996.10524661

Baltagi, B. (2007). *Econometrics*. Springer Science & Business Media.

Baum, C. F., Schaffer, M. E., & Stillman, S. (2003). Instrumental Variables and GMM: Estimation and Testing. *The Stata Journal: Promoting Communications on Statistics and Stata*, *3*(1), 1–31. https://doi.org/10.1177/1536867x0300300101

Bera, A. K., & Bilias, Y. (2002). The MM, ME, ML, EL, EF and GMM approaches to estimation: a synthesis. *Journal of Econometrics*, *107*(1–2), 51–86. https://doi.org/10.1016/s0304-4076(01)00113-0

Bond, S. R., Hoeffler, A., & Temple, J. (2001). *GMM Estimation of Empirical Growth Models* (CEPR Discussion No. 3048). Retrieved from http://pascal.iseg.utl.pt/~depeco/summerschool2005/CEPR-DP3048.pdf

Charemza, W., & Deadman, D. F. (1997). *New Directions in Econometric Practice* (Second). Edward Elgar Publishing. Retrieved from https://www.e-elgar.com/shop/new-directions-in-econometric-practice-second-edition

Cameron, A. C., & Trivedi, P. K. (2005). *Microeconometrics: methods and applications*. Cambridge University Press.

Campbell, R. C. (1999). *An Empirical Examination of Maximum Entropy Estimation.*

Chen, S. X., & Keilegom, I. Van. (2009). A review on empirical likelihood methods for regression. *TEST*, *18*(3), 415–447. https://doi.org/10.1007/s11749-009-0159-5

Cressie, N., & Read, T. R. C. (1984). Multinomial Goodness-Of-Fit Tests. *Journal of the Royal Statistical Society: Series B (Methodological)*, *46*(3), 440–464. https://doi.org/10.1111/j.2517-6161.1984.tb01318.x

Feige, E. L. (1967). Expectations and adjustments in the monetary sector. *The American Economic Review*, *57*(2), 462–473.

Ferguson, T. S. (1958). A Method of Generating Best Asymptotically Normal Estimates with Application to the Estimation of Bacterial Densities. *The Annals of Mathematical Statistics*, *29*(4), 1046–1062. https://doi.org/10.1214/aoms/1177706440

Friedman, M. (1956). *Studies in the quantity theory of money*. University of Chicago Press.

Golan, A., Judge, G. G., & Miller, D. (1996). *Maximum entropy econometrics: Robust estimation with limited data.* John Wiley and Sons.

Gorban, A. N., Gorban, P. A., & Judge, G. (2010). *Entropy: the Markov ordering approach.* Entropy, 12(5), 1145-1193.

Hall, P., & Horowitz, J. L. (1996). Bootstrap Critical Values for Tests Based on Generalized-Method-of-Moments Estimators. *Econometrica, 64*(4), 891–916. https://doi.org/10.2307/2171849

Hansen, B. E., & West, K. D. (2002). Generalized Method of Moments and Macroeconomics. *Journal of Business & Economic Statistics, 20*(4), 460–469. https://doi.org/10.1198/073500102288618603

Hansen, L. P. (1982). Large Sample Properties of Generalized Method of Moments Estimators. *Econometrica, 50*(4), 1029–1054. https://doi.org/10.2307/1912775

Hansen, L. P., Heaton, J., & Yaron, A. (1996). Finite-Sample Properties of Some Alternative GMM Estimators. *Journal of Business & Economic Statistics, 14*(3), 262–280. https://doi.org/10.1080/07350015.1996.10524656

Hsing, Y., & Jamal, A. M. M. (2013). A Simultaneous-Equation Model of Money Demand and Money Supply for Canada. *Modern Economy, 4*(1), 32–36. https://doi.org/10.4236/me.2013.41004

Imbens, G. W. (2002). Generalized Method of Moments and Empirical Likelihood. *Journal of Business & Economic Statistics, 20*(4), 493–506. https://doi.org/10.1198/073500102288618630

Imbens, G. W. (1997). One-Step Estimators for Over-Identified Generalized Method of Moments Models. *The Review of Economic Studies, 64*(3), 359–383. https://doi.org/10.2307/2971718

Imbens, G. W., & Spady, R. (2002). Confidence intervals in generalized method of moments models. *Journal of Econometrics, 107*(1–2), 87–98. https://doi.org/10.1016/s0304-4076(01)00114-2

Imbens, G., Johnson, P., & Spady, R. (1998). Information Theoretic Approaches to Inference in Moment Condition Models. *Econometrica, 66*(2), 333–358. https://doi.org/10.3386/t0186

Judge, G. G., & Mittelhammer, R. C. (2011). *An Information Theoretic Approach to Econometrics.* Cambridge University Press. https://doi.org/10.1017/cbo9781139033848

Judge, G. G., & Mittelhammer, R. C. (2004). A Semiparametric Basis for Combining Estimation Problems Under Quadratic Loss. *Journal of the American Statistical Association, 99*(466), 479–487. https://doi.org/10.1198/016214504000000430

Judge, G., & Mittelhammer, R. (2003). Finite Sample Performance of the Empirical Likelihood Estimator under Endogeneity. In *Computer-Aided Econometrics*. CRC Press. https://doi.org/10.1201/9780203911570.ch6

Judge, G., & Mittelhammer, R. (2012). Implications of the Cressie-Read Family of Additive Divergences for Information Recovery. *Entropy*, *14*(12), 2427–2438. https://doi.org/10.3390/e14122427

Kapetanios, G., & Marcellino, M. (2010). Factor-GMM estimation with large sets of possibly weak instruments. *Computational Statistics & Data Analysis*, *54*(11), 2655–2675. https://doi.org/10.1016/j.csda.2010.04.008

Kendix, M., & Walls, W. D. (2010). Oil industry consolidation and refined product prices: Evidence from US wholesale gasoline terminals. *Energy Policy*, *38*(7), 3498–3507.

Keynes, J. M. (1936). *The general theory of money, interest and employment*. London: Macmillan.

Kitamura, Y. (2001). Asymptotic Optimality of Empirical Likelihood for Testing Moment Restrictions. *Econometrica*, *69*(6), 1661–1672. https://doi.org/10.1111/1468-0262.00261

Kitamura, Y., & Stutzer, M. (1997). An Information-Theoretic Alternative to Generalized Method of Moments Estimation. *Econometrica*, *65*(4), 861–874. https://doi.org/10.2307/2171942

Kitamura, Y., Tripathi, G., & Ahn, H. (2004). Empirical Likelihood-Based Inference in Conditional Moment Restriction Models. *Econometrica*, *72*(6), 1667–1714. https://doi.org/10.1111/j.1468-0262.2004.00550.x

Kolaczyk, E. D. (1994). Empirical likelihood for generalized linear models. *Statistica Sinica*, *4*, 199–218.

Latif, R., & Javid, A. Y. (2016). The Determinants of Pakistan Exports of Textile: An Integrated Demand and Supply Approach. *The Pakistan Development Review*, *55*(3), 191–210. https://doi.org/10.30541/v55i3pp.191-210

Ma, A. (2002). GMM estimation of the new Phillips curve. *Economics Letters*, *76*(3), 411–417. https://doi.org/10.1016/s0165-1765(02)00083-6

McCullagh, P., & Nelder, J. A. (1989). An outline of generalized linear models. In *Generalized Linear Models* (pp. 21–47). Springer {US}. https://doi.org/10.1007/978-1-4899-3242-6_2

Mittelhammer, R. C., Judge, G. G., & Miller, D. J. (2000). *Econometric Foundations*. Cambridge University Press.

Newey, W., & Smith, R. J. (2004). Higher Order Properties of Gmm and Generalized Empirical Likelihood Estimators. *Econometrica*, *72*(1), 219–255. https://doi.org/10.1111/j.1468-0262.2004.00482.x

Ngo, P. T. H., & others. (2006). *Endogenous Capital and Profitability in Banking* (No. 464). Retrieved from https://www.cbe.anu.edu.au/researchpapers/econ/wp464.pdf

Owen, A. (1988). Empirical likelihood ratio confidence intervals for a single functional. *Biometrika*, *75*(2), 237–249. https://doi.org/10.1093/biomet/75.2.237

Owen, A. (1990). Empirical Likelihood Ratio Confidence Regions. *The Annals of Statistics*, *18*(1), 90–120. https://doi.org/10.1214/aos/1176347494

Owen, A. (1991). Empirical Likelihood for Linear Models. *The Annals of Statistics*, 19(4), 1725–1747. https://doi.org/10.1214/aos/1176348368

Pagan, A. R., & Robertson, J. (1997). GMM and its Problems. Australia National University.

Qin, J., & Lawless, J. (1994). Empirical Likelihood and General Estimating Equations. *The Annals of Statistics*, *22*(1), 300–325. https://doi.org/10.1214/aos/1176325370

Qin, J., & Lawless, J. (1995). Estimating equations, empirical likelihood and constraints on parameters. *Canadian Journal of Statistics*, *23*(2), 145–159. https://doi.org/10.2307/3315441

Ramalho, J. J. S. (2006). Bootstrap bias-adjusted GMM estimators. *Economics Letters*, *92*(1), 149–155. https://doi.org/10.1016/j.econlet.2006.01.026

Read, T. R. C., & Cressie, N. A. C. (1988). *Goodness-of-Fit Statistics for Discrete Multivariate Data*. New York: Springer. https://doi.org/10.1007/978-1-4612-4578-0

Stock, J. H., & Yogo, M. (2002). Testing for weak instruments in linear IV regression (No. t0284). National Bureau of Economic Research.

Thomas, D. R., & Grunkemeier, G. L. (1975). Confidence Interval Estimation of Survival Probabilities for Censored Data. *Journal of the American Statistical Association*, *70*(352), 865–871. https://doi.org/10.1080/01621459.1975.10480315

Thomas, R. L. (1993). *Introductory econometrics*. Addison-Wesley Longman Limited.

Tobin, J. (1958). Liquidity Preference as Behavior Towards Risk. *The Review of Economic Studies*, *25*(2), 65–86. https://doi.org/10.2307/2296205

Ullah, A. (2002). Uses of entropy and divergence measures for evaluating econometric approximations and inference. *Journal of Econometrics*, *107*(1–2), 313–326. https://doi.org/10.1016/s0304-4076(01)00126-9

Van Akkeren, M., Judge, G., & Mittelhammer, R. (2002). Generalized moment based estimation and inference. *Journal of Econometrics*, *107*(1–2), 127–148. https://doi.org/10.1016/s0304-4076(01)00116-6

Van Beveren, I. (2012). Total factor productivity estimation: A practical review. *Journal of Economic Surveys*, *26*(1), 98–128.

Windmeijer, F. A. G., & Santos Silva, J. M. C. (1996). *Estimation of count data models with endogenous regressors; an application to demand for health care* (Collaborative Research Center No. 20). Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.55.5687&rep=rep1&type=pdf

Wooldridge, J. M. (2001). Applications of Generalized Method of Moments Estimation. *Journal of Economic Perspectives*, *15*(4), 87–100. https://doi.org/10.1257/jep.15.4.87

Appendix A Results of Monte Carlo Simulation Analysis

Appendix A-1 Biasedness of Estimators for Exactly Determined Model

Samples size	Instrumental variables strength	Degree of Endogeneity (D.E)	Biasedness of Estimators			
			CROCC	CRACC1	CRACC2	CRACC3
25	IVS= 25%	D.E= 20%	0.065381	-0.06903	-0.07762	-0.0862
		D.E= 40%	0.157847	-0.03944	-0.0316	-0.02376
		D.E= 60%	0.217529	0.032671	0.045677	0.058684
		D.E= 80%	0.286005	0.01925	0.040218	0.061185
	IVS= 50%	D.E= 20%	0.041374	-0.05808	-0.05443	-0.05079
		D.E= 40%	0.058197	-0.09795	-0.10695	-0.11594
		D.E= 60%	0.075905	-0.14509	-0.12358	-0.10207
		D.E= 80%	0.067181	-0.229	-0.21655	-0.2041
	IVS= 75%	D.E= 20%	0.003091	-0.0383	-0.04308	-0.04787
		D.E= 40%	0.005751	-0.0556	-0.06135	-0.06709
		D.E= 60%	0.011885	-0.08284	-0.08066	-0.07847
		D.E= 80%	0.01391	-0.11211	-0.11664	-0.12118
50	IVS= 25%	D.E= 20%	0.055253	-0.05495	-0.07006	-0.08517
		D.E= 40%	0.09914	-0.09876	-0.10238	-0.106
		D.E= 60%	0.11508	-0.12712	-0.14012	-0.15312
		D.E= 80%	0.117639	-0.15699	-0.14237	-0.12775
	IVS= 50%	D.E= 20%	0.019603	-0.02664	-0.04187	-0.05709
		D.E= 40%	0.027505	-0.06194	-0.07822	-0.0945
		D.E= 60%	0.033036	-0.11406	-0.12283	-0.13159
		D.E= 80%	0.030993	-0.1602	-0.16586	-0.17152
	IVS= 75%	D.E= 20%	0.002416	-0.00453	-0.00906	-0.0136
		D.E= 40%	0.003593	-0.01573	-0.01968	-0.02364
		D.E= 60%	0.003925	-0.02497	-0.03099	-0.03701
		D.E= 80%	0.004241	-0.0371	-0.04298	-0.04886
100	IVS= 25%	D.E= 20%	0.044539	-0.03768	-0.06018	-0.08268
		D.E= 40%	0.058569	-0.07468	-0.10019	-0.1257
		D.E= 60%	0.06706	-0.11702	-0.13926	-0.1615
		D.E= 80%	0.058549	-0.20155	-0.22344	-0.24534
	IVS= 50%	D.E= 20%	0.010184	-0.00177	-0.01318	-0.0246
		D.E= 40%	0.012135	-0.02107	-0.03199	-0.0429
		D.E= 60%	0.017219	-0.04103	-0.05138	-0.06173
		D.E= 80%	0.020696	-0.05312	-0.06044	-0.06775
	IVS= 75%	D.E= 20%	0.000284	-0.00042	-0.00476	-0.0091
		D.E= 40%	0.004384	0.000167	-0.0055	-0.01116
		D.E= 60%	0.004469	-0.0051	-0.01064	-0.01619
		D.E= 80%	0.004554	-0.0106	-0.01477	-0.01895
			Biasedness of Estimators			

Samples size	Instrumental variables strength (IVS)	Degree of Endogeneity (D.E)	CROCC	CRACC1	CRACC2	CRACC3
150	IVS= 25%	D.E= 20%	0.031509	-0.02365	-0.04626	-0.06888
		D.E= 40%	0.043599	-0.0556	-0.08142	-0.10724
		D.E= 60%	0.053063	-0.08748	-0.10532	-0.12316
		D.E= 80%	0.051465	-0.13645	-0.15081	-0.16517
	IVS= 50%	D.E= 20%	0.005392	0.003469	-0.00783	-0.01913
		D.E= 40%	0.011033	-0.0042	-0.01639	-0.02858
		D.E= 60%	0.012651	-0.01697	-0.02639	-0.03581
		D.E= 80%	0.010852	-0.02521	-0.03603	-0.04685
	IVS= 75%	D.E= 20%	0.003268	0.005369	6.44E-06	-0.00536
		D.E= 40%	0.003663	0.001804	-0.00301	-0.00783
		D.E= 60%	0.003796	-0.00094	-0.00613	-0.01131
		D.E= 80%	0.002618	-0.00519	-0.00978	-0.01437
200	IVS= 25%	D.E= 20%	0.029155	0.007843	-0.01659	-0.04103
		D.E= 40%	0.032546	-0.02441	-0.04548	-0.06654
		D.E= 60%	0.043563	-0.0557	-0.07657	-0.09745
		D.E= 80%	0.036544	-0.10209	-0.12319	-0.1443
	IVS= 50%	D.E= 20%	0.004395	0.008441	-0.00471	-0.01785
		D.E= 40%	0.011819	0.004204	-0.00716	-0.01852
		D.E= 60%	0.011312	-0.00325	-0.01494	-0.02664
		D.E= 80%	0.009196	-0.01563	-0.02504	-0.03445
	IVS= 75%	D.E= 20%	0.002988	0.005725	0.000488	-0.00475
		D.E= 40%	0.002851	0.002296	-0.00213	-0.00655
		D.E= 60%	0.001028	-0.00071	-0.00636	-0.01201
		D.E= 80%	-0.00073	-0.00533	-0.0105	-0.01567
250	IVS= 25%	D.E= 20%	0.02941	0.022266	-0.00083	-0.02393
		D.E= 40%	0.036814	-0.00173	-0.02373	-0.04573
		D.E= 60%	0.037394	-0.03095	-0.0509	-0.07084
		D.E= 80%	0.040691	-0.06282	-0.08409	-0.10537
	IVS= 50%	D.E= 20%	0.008856	0.015284	0.00191	-0.01146
		D.E= 40%	0.008097	0.00607	-0.00645	-0.01896
		D.E= 60%	0.008201	-0.00059	-0.01231	-0.02403
		D.E= 80%	0.009935	-0.004	-0.01688	-0.02977
	IVS= 75%	D.E= 20%	0.004753	0.008742	0.002876	-0.00299
		D.E= 40%	0.001847	0.003637	-0.00209	-0.00781
		D.E= 60%	0.003311	0.002761	-0.00239	-0.00754
		D.E= 80%	0.002278	9.45E-05	-0.00524	-0.01057

Appendix A-2 MSE of Estimators for Exactly Determined Model

Samples size	Instrumental variables strength (IVS)	Degree of Endogeneity (D.E)	MSE of Estimators			
			CROCC	CRACC1	CRACC2	CRACC3
25	IVS= 25%	D.E= 20%	0.608279	3.441282	3.185371	4.495913
		D.E= 40%	0.605876	3.263028	3.181471	4.668685
		D.E= 60%	0.846179	3.523326	3.358625	4.860366
		D.E= 80%	0.648154	2.813016	3.841555	6.867023
	IVS= 50%	D.E= 20%	0.223713	1.328453	1.22797	1.743284
		D.E= 40%	0.19268	1.335503	1.337437	1.985697
		D.E= 60%	0.185143	1.484876	1.854176	3.164496
		D.E= 80%	0.250193	1.762291	2.158854	3.59565
	IVS= 75%	D.E= 20%	0.052027	0.21974	0.172165	0.208793
		D.E= 40%	0.056278	0.161666	0.141053	0.186786
		D.E= 60%	0.05875	0.217909	0.226622	0.35247
		D.E= 80%	0.053525	0.363032	0.384228	0.594966
50	IVS= 25%	D.E= 20%	0.371587	2.100016	1.690888	2.097704
		D.E= 40%	0.362297	1.718808	1.843339	2.817162
		D.E= 60%	0.328868	1.888849	2.085729	3.288267
		D.E= 80%	0.329152	1.658933	2.205789	3.818831
	IVS= 50%	D.E= 20%	0.069041	0.35614	0.256796	0.286736
		D.E= 40%	0.070282	0.260713	0.281402	0.434113
		D.E= 60%	0.078162	0.375373	0.367721	0.547447
		D.E= 80%	0.082957	0.534261	0.65655	1.090143
	IVS= 75%	D.E= 20%	0.022068	0.034545	0.026301	0.031412
		D.E= 40%	0.022292	0.03754	0.028418	0.033502
		D.E= 60%	0.022746	0.035759	0.030377	0.039582
		D.E= 80%	0.02418	0.041562	0.035683	0.046571
100	IVS= 25%	D.E= 20%	0.121573	0.581823	0.537227	0.752781
		D.E= 40%	0.123083	0.558316	0.553682	0.823715
		D.E= 60%	0.120467	0.59055	0.754655	1.289644
		D.E= 80%	0.145091	0.806113	0.88733	1.392552
	IVS= 50%	D.E= 20%	0.032313	0.060897	0.046698	0.056617
		D.E= 40%	0.034019	0.067039	0.05333	0.065132
		D.E= 60%	0.033908	0.090485	0.078846	0.104886
		D.E= 80%	0.035828	0.093172	0.07896	0.101072
	IVS= 75%	D.E= 20%	0.011051	0.014634	0.011659	0.014721
		D.E= 40%	0.010994	0.014276	0.011767	0.01499
		D.E= 60%	0.011134	0.014519	0.012244	0.015905
		D.E= 80%	0.010956	0.014567	0.01238	0.016204

Samples size	Instrumental variables strength (IVS)	Degree of Endogeneity (D.E)	MSE of Estimators			
			CROCC	CRACC1	CRACC2	CRACC3
150	IVS= 25%	D.E= 20%	0.071177	0.371964	0.26981	0.302891
		D.E= 40%	0.076082	0.286645	0.292229	0.438329
		D.E= 60%	0.073157	0.307383	0.328738	0.5089
		D.E= 80%	0.079044	0.404871	0.391217	0.560143
	IVS= 50%	D.E= 20%	0.020884	0.032737	0.025179	0.031097
		D.E= 40%	0.021553	0.033257	0.027883	0.036979
		D.E= 60%	0.021699	0.035634	0.028776	0.036617
		D.E= 80%	0.021762	0.034852	0.030119	0.039717
	IVS= 75%	D.E= 20%	0.007111	0.009316	0.007346	0.009208
		D.E= 40%	0.006982	0.008762	0.007223	0.009556
		D.E= 60%	0.007057	0.008851	0.007456	0.009946
		D.E= 80%	0.007149	0.008955	0.007851	0.010462
200	IVS= 25%	D.E= 20%	0.051263	0.165853	0.125626	0.148616
		D.E= 40%	0.053005	0.138027	0.133149	0.193603
		D.E= 60%	0.053489	0.175486	0.195593	0.311827
		D.E= 80%	0.055447	0.220916	0.219978	0.323563
	IVS= 50%	D.E= 20%	0.016429	0.023981	0.018646	0.022917
		D.E= 40%	0.01545	0.021867	0.017408	0.022696
		D.E= 60%	0.016022	0.022814	0.019345	0.025488
		D.E= 80%	0.017044	0.026254	0.021663	0.027113
	IVS= 75%	D.E= 20%	0.005442	0.006695	0.005563	0.007251
		D.E= 40%	0.005234	0.006528	0.005418	0.007182
		D.E= 60%	0.005343	0.006455	0.005583	0.007571
		D.E= 80%	0.00548	0.006707	0.005944	0.008124
250	IVS= 25%	D.E= 20%	0.038891	0.076097	0.057294	0.067673
		D.E= 40%	0.04146	0.081505	0.064404	0.079293
		D.E= 60%	0.042105	0.103859	0.082953	0.101977
		D.E= 80%	0.041137	0.12746	0.130338	0.196009
	IVS= 50%	D.E= 20%	0.013156	0.017792	0.014209	0.018
		D.E= 40%	0.012787	0.017814	0.014182	0.018058
		D.E= 60%	0.012793	0.017942	0.014591	0.018604
		D.E= 80%	0.01259	0.017535	0.016231	0.02293
	IVS= 75%	D.E= 20%	0.004328	0.005586	0.004424	0.005521
		D.E= 40%	0.004244	0.005256	0.004378	0.005784
		D.E= 60%	0.004209	0.004947	0.004356	0.005975
		D.E= 80%	0.004128	0.004888	0.004381	0.006096

Appendix A-3 Biasedness of Estimators for Overdetermined Model

Samples size	Instrumental variables strength (IVS)	Degree of Endogeneity (D.E)	Biasedness of Estimators			
			CROCC	CRACC1	CRACC2	CRACC3
25	IVS= 25%	D.E= 20%	0.137196	0.127931	0.129186	0.130463
		D.E= 40%	0.285296	0.262809	0.260465	0.258121
		D.E= 60%	0.439316	0.417215	0.418101	0.418987
		D.E= 80%	0.588621	0.545017	0.549221	0.553404
	IVS= 50%	D.E= 20%	0.080708	0.062154	0.062658	0.063078
		D.E= 40%	0.178066	0.150712	0.153969	0.157226
		D.E= 60%	0.280108	0.232853	0.232492	0.232033
		D.E= 80%	0.389285	0.332043	0.329161	0.326278
	IVS= 75%	D.E= 20%	0.008604	0.020599	0.044186	0.037768
		D.E= 40%	0.05769	0.033612	0.039797	0.045933
		D.E= 60%	0.098953	0.068763	0.074373	0.080571
		D.E= 80%	0.151407	0.105978	0.111231	0.116772
50	IVS= 25%	D.E= 20%	0.099169	0.086443	0.085911	0.085379
		D.E= 40%	0.219998	0.190185	0.19148	0.192775
		D.E= 60%	0.351371	0.308902	0.307832	0.306762
		D.E= 80%	0.480281	0.428384	0.425833	0.423283
	IVS= 50%	D.E= 20%	0.04825	0.03343	0.033096	0.032762
		D.E= 40%	0.108783	0.074716	0.075702	0.076688
		D.E= 60%	0.178634	0.136722	0.134267	0.131811
		D.E= 80%	0.237675	0.184119	0.181864	0.180051
	IVS= 75%	D.E= 20%	0.001869	-0.00745	-0.00565	-0.00386
		D.E= 40%	0.02272	0.008559	0.007899	0.00702
		D.E= 60%	0.050752	0.029625	0.030245	0.031083
		D.E= 80%	0.07158	0.043787	0.043856	0.043596
100	IVS= 25%	D.E= 20%	0.065476	0.048901	0.047879	0.046858
		D.E= 40%	0.14921	0.11302	0.11319	0.11336
		D.E= 60%	0.24736	0.20023	0.19701	0.1938
		D.E= 80%	0.33062	0.26886	0.26659	0.26431
	IVS= 50%	D.E= 20%	0.01784	0.006351	0.006334	0.006318
		D.E= 40%	0.057349	0.035067	0.0335	0.031932
		D.E= 60%	0.092048	0.059188	0.057704	0.056219
		D.E= 80%	0.12404	0.083829	0.083149	0.082469
	IVS= 75%	D.E= 20%	-0.00961	-0.01307	-0.01455	-0.01588
		D.E= 40%	0.003349	-0.0049	-0.00591	-0.00683
		D.E= 60%	0.018149	0.006985	0.006071	0.005303
		D.E= 80%	0.029071	0.015411	0.014199	0.012977

Samples size	Instrumental variables strength (IVS)	Degree of Endogeneity (D.E)	Biasedness of Estimators			
			CROCC	CRACC1	CRACC2	CRACC3
150	IVS= 25%	D.E= 20%	0.047859	0.028104	0.030881	0.033659
		D.E= 40%	0.114459	0.080095	0.080063	0.080032
		D.E= 60%	0.177	0.13161	0.132027	0.132445
		D.E= 80%	0.242562	0.178268	0.178021	0.177773
	IVS= 50%	D.E= 20%	0.00405	-0.00466	-0.00385	-0.00304
		D.E= 40%	0.031352	0.016308	0.015951	0.015593
		D.E= 60%	0.055758	0.033012	0.03297	0.032928
		D.E= 80%	0.081191	0.051966	0.051415	0.050864
	IVS= 75%	D.E= 20%	-0.01044	-0.01267	-0.01442	-0.01618
		D.E= 40%	-0.00279	-0.00716	-0.00953	-0.0119
		D.E= 60%	0.005607	-0.0018	-0.00339	-0.00509
		D.E= 80%	0.012708	0.003345	0.001146	-0.00113
200	IVS= 25%	D.E= 20%	0.030196	0.013401	0.017015	0.020629
		D.E= 40%	0.085949	0.05506	0.055952	0.056844
		D.E= 60%	0.141739	0.098578	0.098141	0.097705
		D.E= 80%	0.185813	0.132073	0.132961	0.133849
	IVS= 50%	D.E= 20%	-0.00054	-0.00658	-0.00771	-0.00884
		D.E= 40%	0.022307	0.010034	0.009213	0.008393
		D.E= 60%	0.035654	0.017394	0.017636	0.017879
		D.E= 80%	0.055484	0.03206	0.032584	0.033109
	IVS= 75%	D.E= 20%	-0.01178	-0.0128	-0.01627	-0.01983
		D.E= 40%	-0.00705	-0.01036	-0.01318	-0.01601
		D.E= 60%	0.001963	-0.00275	-0.0062	-0.0095
		D.E= 80%	0.004819	-0.0022	-0.00486	-0.00752
250	IVS= 25%	D.E= 20%	0.024592	0.007562	0.007339	0.007117
		D.E= 40%	0.070489	0.043026	0.04345	0.043874
		D.E= 60%	0.110249	0.071218	0.073797	0.076376
		D.E= 80%	0.153508	0.101704	0.102893	0.104081
	IVS= 50%	D.E= 20%	-0.00379	-0.00919	-0.00932	-0.00945
		D.E= 40%	0.011339	0.00159	0.00157	0.001534
		D.E= 60%	0.028657	0.01412	0.013745	0.013371
		D.E= 80%	0.040216	0.020515	0.02109	0.021665
	IVS= 75%	D.E= 20%	-0.01152	-0.01227	-0.0154	-0.01858
		D.E= 40%	-0.00726	-0.00953	-0.01278	-0.01611
		D.E= 60%	-0.00359	-0.00733	-0.00972	-0.01207
		D.E= 80%	0.001848	-0.00338	-0.00657	-0.00976

Appendix A-4 MSE of Estimators for Overdetermined Model

Samples size	Instrumental variables strength (IVS)	Degree of Endogeneity (D.E)	MSE of Estimators			
			CROCC	CRACC1	CRACC2	CRACC3
25	IVS= 25%	D.E= 20%	0.142165	0.346788	0.243925	0.256547
		D.E= 40%	0.195545	0.452534	0.319386	0.308795
		D.E= 60%	0.287124	0.593679	0.429032	0.39179
		D.E= 80%	0.414447	0.452205	0.424468	0.455284
	IVS= 50%	D.E= 20%	0.093487	0.159199	0.133853	0.173954
		D.E= 40%	0.1357	0.193208	0.153763	0.176881
		D.E= 60%	0.151019	0.201417	0.16252	0.179598
		D.E= 80%	0.226938	0.257841	0.215703	0.228209
	IVS= 75%	D.E= 20%	0.048587	0.23433	0.074612	0.029667
		D.E= 40%	0.052662	0.111713	0.073063	0.069851
		D.E= 60%	0.055584	0.114206	0.078419	0.080459
		D.E= 80%	0.069119	0.114947	0.077774	0.074028
50	IVS= 25%	D.E= 20%	0.100122	0.188594	0.149763	0.182141
		D.E= 40%	0.132308	0.190683	0.160344	0.19414
		D.E= 60%	0.197253	0.245258	0.211476	0.236782
		D.E= 80%	0.290284	0.299833	0.271344	0.288034
	IVS= 50%	D.E= 20%	0.055653	0.081615	0.063563	0.075551
		D.E= 40%	0.06296	0.088075	0.068358	0.081283
		D.E= 60%	0.076562	0.092109	0.074571	0.085264
		D.E= 80%	0.097222	0.094641	0.083617	0.097183
	IVS= 75%	D.E= 20%	0.022831	0.029017	0.022425	0.027609
		D.E= 40%	0.023322	0.028537	0.022441	0.027361
		D.E= 60%	0.025534	0.030893	0.023822	0.027741
		D.E= 80%	0.025519	0.029087	0.023667	0.028716
100	IVS= 25%	D.E= 20%	0.067122	0.11074	0.087806	0.10874
		D.E= 40%	0.081123	0.10901	0.091306	0.11286
		D.E= 60%	0.11382	0.13005	0.10964	0.12499
		D.E= 80%	0.15595	0.15188	0.13464	0.14865
	IVS= 50%	D.E= 20%	0.030024	0.039395	0.031126	0.039097
		D.E= 40%	0.032317	0.041335	0.03285	0.039909
		D.E= 60%	0.036067	0.04217	0.033086	0.039266
		D.E= 80%	0.040417	0.043859	0.035215	0.040462
	IVS= 75%	D.E= 20%	0.011646	0.013888	0.010869	0.013303
		D.E= 40%	0.012315	0.014097	0.011207	0.013879
		D.E= 60%	0.011333	0.013208	0.010495	0.013324
		D.E= 80%	0.01194	0.013644	0.011035	0.013479

Samples size	Instrumental variables strength (IVS)	Degree of Endogeneity (D.E)	MSE of Estimators			
			CROCC	CRACC1	CRACC2	CRACC3
150	IVS= 25%	D.E= 20%	0.04917	0.072433	0.058295	0.074218
		D.E= 40%	0.057323	0.078605	0.062114	0.07386
		D.E= 60%	0.072846	0.082061	0.06929	0.081491
		D.E= 80%	0.094864	0.092136	0.077325	0.085851
	IVS= 50%	D.E= 20%	0.020605	0.026119	0.020279	0.024774
		D.E= 40%	0.021531	0.025943	0.02056	0.025214
		D.E= 60%	0.022505	0.026265	0.0209	0.025364
		D.E= 80%	0.024914	0.027473	0.021642	0.025217
	IVS= 75%	D.E= 20%	0.007789	0.008993	0.007245	0.008943
		D.E= 40%	0.007627	0.009012	0.006956	0.008389
		D.E= 60%	0.007562	0.008606	0.00689	0.008595
		D.E= 80%	0.00757	0.009019	0.006981	0.008372
200	IVS= 25%	D.E= 20%	0.039088	0.055197	0.044828	0.056715
		D.E= 40%	0.044515	0.059738	0.048216	0.059503
		D.E= 60%	0.054335	0.059906	0.051075	0.062625
		D.E= 80%	0.063819	0.064176	0.053703	0.061311
	IVS= 50%	D.E= 20%	0.015889	0.018976	0.015462	0.019707
		D.E= 40%	0.015863	0.019027	0.014955	0.018404
		D.E= 60%	0.016255	0.019329	0.015209	0.018598
		D.E= 80%	0.018259	0.02101	0.016195	0.018598
	IVS= 75%	D.E= 20%	0.00579	0.006871	0.005409	0.006517
		D.E= 40%	0.00575	0.006718	0.005331	0.006634
		D.E= 60%	0.005957	0.006948	0.005483	0.006767
		D.E= 80%	0.00558	0.006593	0.005145	0.006354
250	IVS= 25%	D.E= 20%	0.03236	0.045927	0.036624	0.046182
		D.E= 40%	0.036704	0.047213	0.039871	0.051056
		D.E= 60%	0.040152	0.047368	0.038919	0.046757
		D.E= 80%	0.048716	0.051637	0.043082	0.05107
	IVS= 50%	D.E= 20%	0.013589	0.016412	0.012998	0.015511
		D.E= 40%	0.012581	0.014886	0.012085	0.015189
		D.E= 60%	0.013141	0.015523	0.012174	0.014888
		D.E= 80%	0.013391	0.015684	0.012173	0.014689
	IVS= 75%	D.E= 20%	0.004967	0.00546	0.00439	0.005373
		D.E= 40%	0.005144	0.005753	0.004825	0.00587
		D.E= 60%	0.004771	0.005598	0.004319	0.005165
		D.E= 80%	0.004769	0.005531	0.004398	0.00538

Appendix B Estimation of Consumption Function

Countries	Coefficients	MEL	MEEL	GMM	CROCC	CRACC2
Australia	constant	0.040883	0.297131	-0.56099	0.175187	0.169007
	slope	0.903904	0.862429	1.001097	0.882166	0.883167
Austria	constant	0.697428	0.558382	-0.69017	0.598455	0.627905
	slope	0.822427	0.84207	1.01731	0.836409	0.832249
Bangladesh	constant	2.888245	3.275478	3.293675	3.228758	3.081862
	slope	0.754554	0.684209	0.680853	0.692696	0.719382
Belgium	constant	0.653818	3.873487	0.016325	0.883373	2.263653
	slope	0.855879	0.768929	0.932212	0.82829	0.812404
India	constant	0.87257	2.904443	0.102782	2.935136	1.888507
	slope	0.763455	0.642297	0.941763	0.638956	0.702876
New Zealand	constant	5.328512	4	-0.39655	5.328525	4.664256
	slope	0.559051	0.65	0.989305	0.559049	0.604526
Spain	constant	0.017307	0.344233	-0.17278	0.110209	0.18077
	slope	0.951401	0.918809	0.970317	0.94214	0.935105
Sri lanka	constant	0.781741	1.663143	0.250428	0.896459	1.222442
	Slope	0.919569	0.846584	0.958349	0.91007	0.883077

Appendix C Estimation of Money Demand Function

Country	coefficients	Estimators				
		MEL	MEEL	GMM	CROCC	CRACC2
Australia	alpha	13.48706	13.41692	13.5331	13.45184	13.45199
	beta 1	1.332292	1.326763	1.335197	1.329515	1.329528
	Beta 2	-0.3658	-0.30779	-0.40061	-0.33667	-0.3368
Bahamas	alpha	15.82671	13.4607	10.30451	14.28837	14.64371
	beta 1	0.692968	0.912388	1.190769	0.835632	0.802678
	Beta 2	-1.13306	-0.86625	-0.39599	-0.95958	-0.99966
Barbados	alpha	7.405828	5.684565	4.514197	6.392939	6.545197
	beta 1	1.557935	1.718882	1.828512	1.652645	1.638409
	Beta 2	-0.49286	-0.42026	-0.37214	-0.45014	-0.45656
Belize	alpha	12.73991	11.326	26.41333	12.0116	12.03296
	beta 1	1.086851	1.238312	-0.38156	1.164869	1.162582
	Beta 2	-1.24919	-1.18611	-1.83796	-1.2167	-1.21765
Canada	alpha	27.08581	23.49773	25.4025	25.00163	25.29177
	beta 1	0.206776	0.531555	0.381692	0.395428	0.369166
	Beta 2	-1.10219	-0.94985	-1.15918	-1.0137	-1.02602
Dominica	alpha	12.67075	11.6948	12.52629	12.193	12.18278
	beta 1	0.983077	1.094832	1.050444	1.037783	1.038955
	Beta 2	-0.95018	-0.98642	-1.22315	-0.96792	-0.9683
Fiji	alpha	10.75997	10.40242	10.73836	10.56546	10.5812
	beta 1	1.2829	1.321045	1.285203	1.303652	1.301973
	Beta 2	-0.24042	-0.19742	-0.23783	-0.21703	-0.21892
Grenada	alpha	8.659073	7.690368	21.56799	8.119499	8.174721
	beta 1	1.503722	1.519411	1.258571	1.512461	1.511567
	Beta 2	-1.19684	-0.75252	-6.93384	-0.94935	-0.97468
Jamaica	alpha	20.47284	18.80992	20.01009	19.55918	19.64138
	beta 1	0.739674	0.845766	0.795041	0.797964	0.79272
	Beta 2	-1.13985	-0.99432	-1.2073	-1.05989	-1.06709
Japan	alpha	22.11652	25.62112	13.55515	23.68147	23.86882
	beta 1	0.81326	0.58153	1.37935	0.709783	0.697395
	Beta 2	-0.04571	-0.05048	-0.03408	-0.04784	-0.0481
Kenya	alpha	12.76104	12.81586	13.21006	12.78782	12.78845
	beta 1	1.351085	1.34816	1.327386	1.349656	1.349623
	Beta 2	-0.04695	-0.0572	-0.13197	-0.05196	-0.05208
Lesotho	alpha	21.53174	22.30248	24.5633	22.1808	21.91711
	beta 1	0.621614	0.575286	0.481128	0.5826	0.59845
	Beta 2	-2.22421	-2.40153	-3.04603	-2.37353	-2.31287

Country	coefficients	Estimators				
		MEL	MEEL	GMM	CROCC	CRACC2
Malawi	alpha	15.99173	15.85597	17.61425	15.92052	15.92385
	beta 1	1.080856	1.080097	1.08975	1.080458	1.080477
	Beta 2	-0.71861	-0.67078	-1.2893	-0.69352	-0.6947
Papua New Guinea	alpha	16.82337	16.38107	7.181783	16.62403	16.60222
	beta 1	1.020966	1.039991	1.581664	1.02954	1.030479
	Beta 2	-1.16703	-1.03652	1.27666	-1.10821	-1.10178
Philippines	alpha	11.36105	11.23736	11.23241	11.29659	11.29921
	beta 1	1.577005	1.586923	1.587328	1.582174	1.581964
	Beta 2	-0.01764	-0.00811	-0.00779	-0.01267	-0.01288
Sierra Leone	alpha	16.60935	16.83191	19.0395	16.71685	16.72063
	beta 1	1.018737	0.975616	0.942348	0.997909	0.997177
	Beta 2	-1.11315	-1.00566	-1.62164	-1.06123	-1.05941
Singapore	alpha	15.47461	15.11676	15.57828	15.2836	15.29569
	beta 1	1.018082	1.049462	1.008937	1.034831	1.033772
	Beta 2	-0.99924	-0.95047	-1.01302	-0.97321	-0.97486
Solomon Islands	alpha	17.29033	16.77748	11.67634	17.02253	17.03391
	beta 1	0.69039	0.736705	1.090838	0.714575	0.713548
	Beta 2	-1.85608	-1.78699	-0.55372	-1.82001	-1.82154
South Africa	alpha	20.25904	19.93359	15.63602	20.0729	20.09632
	beta 1	0.946602	0.966046	1.226517	0.957723	0.956324
	Beta 2	-1.13882	-1.08208	-0.34012	-1.10637	-1.11045
Sri Lanka	alpha	17.51744	17.42959	20.23614	17.55981	17.47352
	beta 1	1.121903	1.117462	1.095787	1.124045	1.119683
	Beta 2	-1.23509	-1.18308	-2.19007	-1.26018	-1.20909
St. Kitts and Nevis	alpha	25.19735	28.26567	-2.29956	26.42273	26.73151
	beta 1	0.645141	0.36479	1.277738	0.533178	0.504966
	Beta 2	-5.82938	-5.98211	5.323506	-5.89038	-5.90575
St. Lucia	alpha	8.480981	8.628457	7.721107	8.549208	8.554719
	beta 1	1.535306	1.538467	1.533181	1.536769	1.536887
	Beta 2	-1.06729	-1.16184	-0.65	-1.11103	-1.11457
St. Vincent and the Grenadines	alpha	11.84475	10.94959	8.079521	11.41195	11.39717
	beta 1	1.173992	1.243587	1.32085	1.207641	1.20879
	Beta 2	-1.26051	-1.11241	0.079798	-1.18891	-1.18646
Swaziland	alpha	15.17194	15.37623	16.64025	15.26964	15.27409
	beta 1	0.916861	0.910297	0.867772	0.913722	0.913579
	Beta 2	-0.92946	-0.99515	-1.39359	-0.96088	-0.96231
Uganda	alpha	17.75942	17.71008	19.93859	17.73462	15.92385
	beta 1	0.994686	0.996368	0.913917	0.995531	1.080477
	Beta 2	-0.98117	-0.97035	-1.43159	-0.97573	-0.6947