Bias Assessment and Reduction for Limited Information Estimation in General Dynamic Simultaneous Equations Models

Dandan Wang

Supervisor: Prof. Garry Phillips

Department of Economics

Cardiff University

This dissertation is submitted for the degree of

Doctor of Philosophy

Business School January 2017

I would like to dedicate this thesis to my loving parents.

ANNEX 1:
Specimen layout for Declaration/Statements page to be included in a thesis.

DECLARATION

This work has not been submitted in substance for any other degree or award at this or any other university or place of learning, nor is being submitted concurrently in candidature for any degree or other award.

Signed ……………………………………………………… (candidate) Date …………………..…………….………

STATEMENT 1

This thesis is being submitted in partial fulfillment of the requirements for the degree of ………(insert MCh, MD, MPhil, PhD etc, as appropriate)

Signed ……………………………………….…………… (candidate) Date ……………………………..……………

STATEMENT 2

This thesis is the result of my own independent work/investigation, except where otherwise stated, and the thesis has not been edited by a third party beyond what is permitted by Cardiff University's Policy on the Use of Third Party Editors by Research Degree Students. Other sources are acknowledged by explicit references. The views expressed are my own.

Signed ……………………………………….……….…… (candidate) Date …………………….…………………

STATEMENT 3

I hereby give consent for my thesis, if accepted, to be available online in the University's Open Access repository and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed ……………………………………………..…..….. (candidate) Date …………………………………………

STATEMENT 4: PREVIOUSLY APPROVED BAR ON ACCESS

I hereby give consent for my thesis, if accepted, to be available online in the University's Open Access repository and for inter-library loans **after expiry of a bar on access previously approved by the Academic Standards & Quality Committee.**

Signed ……………………………………………..……… (candidate) Date ………………………………….………

Acknowledgements

My gratitude goes to all the people who helped me gain this invaluable experience.

It has been a privilege to be supervised by Professor Garry Phillips, I would like to express my special appreciation and thanks to him. It has been an honour to be his Ph.D. student. I appreciate all his contributions of time, ideas, and support to make my Ph.D. experience productive and stimulating. The joy and enthusiasm he has for his research was contagious and motivational for me, even during tough times in the Ph.D. pursuit. I am also thankful for the excellent example he has provided as a successful econometrician and professor.Your advice on both research as well as on my career have been invaluable.

Thanks to the Julian Hodge bank for funding my Ph.D. research. Thanks to Professor Patrick Minford who was always so kind to give me the advise in both research and career.

Thanks also to all the lecturers who have taught me at Cardiff Business School, to Alicante University for providing me the visiting opportunities, and to family and friends for being a force for good in this world.

Abstract

Most of the literature which has considered the small sample bias of limited information estimators in simultaneous equation models has done so in the context of the static rather than the dynamic simultaneous equations model (DSEM). Therefore, an analysis of the performance of estimators in the general dynamic simultaneous equations case is timely and this is what is provided in this paper. By introducing an asymptotic expansion for the estimation errors of estimators, we are able to obtain bias approximations to order T^{-1}. Following this we constructed bias corrected estimators by using the estimated bias approximation to reduce the bias. As an alternative, the use of the non-parametric bootstrap as a bias correction procedure was also examined.

In Chapter 2, we analyse the Two Stage Least Squares (2SLS) Estimator in the general DSEM. Based on the result in Chapter 2, Chapter 3 compared the Fuller modification of the limited information maximum likelihood estimator (FLIML) with the 2SLS estimator. The bias approximation and reduction in the p^{th}-order dynamic reduced form are analysed in Chapter 4.

The results indicate that FLIML gives much less biased estimates than the 2SLS estimation in the general DSEM. We have also observed that the bias correction method based on the estimated bias approximation to order T^{-1} provides almost unbiased estimates and it does not lead to an inflation of the mean squared errors compared with the associated uncorrected estimators. We suggest that the corrected estimators, based upon the $O(T^{-1})$, should be used to reduce the bias of the original estimators in small samples. Alternatively, the numerical results show that the bootstrap method leads to an effective reduction of the bias and an inflation of MSE, however this reduction is not as effective as the first one.

Keywords :General Dynamic simultaneous equations model; Asymptotic approximations; Bias correction; Bootstrap; Monte Carlo simulations; 2SLS; FLIML; OLS; C2SLS; CFLIML; COLS.

JEL classification: C13; C32

Contents

List of Tables

LIST OF ABBREVIATIONS

Acronyms / Abbreviations

2SLS Two Stage Least Square

AR Autoregressive

ARMA Autoregressive–Moving-Average

BOLS Bootstrap Ordinary Least Square

C2SLS Corrected Two Stage Least Square

CFLIML Corrected Fuller Limited Information Maximum Likelihood

CLRM Classical Linear Regression Models

COLS Corrected Ordinary Least Square

DSEM Dynamic Simultaneous Equations Models

FIE Full Information Estimators

FLIML Fuller Limited Information Maximum Likelihood

GLS Generalised Least Squares

ILS Indirect Least Squares

LIE Limited Information Estimators

LIML Limited Information Maximum Likelihood

MC Monte Carlo

MSE Mean Squared Errors

OLS Ordinary Least Square

SEM Simultaneous Equations Models

Chapter 1

Introduction

Economic modelling provides the relationships between economic variables which are useful in making predictions and conducting policy evaluations. Well known examples of econometric models include the classical linear regression models (CLRM), where the regressors are assumed to be non-stochastic, the model is linear and the errors independent and identically distributed normal random variables. However, these assumptions are rather far from economic reality, it has long been realized that a relaxation of these assumptions is necessary in empirical work. The non-classical i.e statistical models for which these assumptions are violated and the inference procedures of CLRM are not applicable, include the simultaneous equations model (SEM) which was introduced by Haavelmo (1943). Since then the SEM has been used extensively to analyse economic phenomena which gives a more realistic representation of an economic process. However this relaxation of the assumptions of the CLRM usually makes it impractical to derive the exact distributional properties of estimators and test statistics in finite samples. Thus a simpler approach is required and instead of basing analysis on the exact properties of estimators reliance is placed on their asymptotic properties which can be obtained over a wide range of models and which are generally assumed to provide a reasonable approximation to finite sample properties.

The asymptotic theory of estimators and test statistics plays an important role in econometric inference for analysing economic phenomena. However, asymptotic properties hinge upon a crucial condition that the number of observations be infinitely large; this condition is generally not met in the practice though and the quantity and quality of economic data are not controllable. In many cases, only small sample sizes of data or poor quality data are observed. How large the sample of observations should be to achieve the asymptotic properties remains unanswered. Fisher (1921) originally pointed out that the asymptotic theory which requires the number of observations to be infinitely large for the asymptotic results to hold true may not imply the finite sample behaviour of economic estimators and the test statistics in many practical applications. As well they may give misleading results for small or even moderately large samples. Even if a large number of observations is available, such as with an increasing number of data sets in finance, development economics, and labour economics etc., it may not desirable to use them because of the non-compliance with the other conditions required for the asymptotic theory to hold. Moreover, Leamer (1978) pointed out that sequential application of asymptotically equivalent procedures, might lead to results that are not asymptotically equivalent under the data instigated models. It was shown that asymptotically equivalent estimators may have very different finite sample properties.

For these and the other important reasons, it seems that the information about the small sample behaviour of estimators and test statistics is of great value in econometrics. This thesis is devoted to obtaining such information in a form that is easy to interpret and practical to use.

There are three main tools of analysis for obtaining more information for the small sample behaviour of the econometric estimators and test statistics: exact finite sample theory, the use of asymptotic expansions or large sample approximation, and more recently, the application of the bootstrapping (resampling) technique.

Fisher (1921), Fisher (1922), and then the work of Cramér (1945) laid the foundation of statistical finite sample theory on the exact distributions and moments which are valid for any sample size. This exact theory on distributions and moments was introduced into econometrics by Haavelmo (1947) and Anderson and Rubin (1950) on the exact confidence region of structural coefficient, Hurwicz (1950) on the exact least square bias in an autoregressive model, Basmann (1963), Richardson (1968), Sawa (1969), Sawa (1972), Anderson and Sawa (1973), Ullah and Nagar (1974), Phillips (1983), and Ullah (2004) on the exact density and moments of the estimators in the simultaneous equations model. All these important contributions were related to obtaining exact results, which hold for any sample size, small, moderately large, or very large. However, the density functions of estimators have a complicated mathematical structure which makes it difficult to draw the meaningful inference from them.

The second method uses asymptotic approximations, with errors of smaller order of magnitude than the first order asymptotic approximation, to obtain more information on small sample behaviour. It provides results which will tend to lie between the exact and asymptotic results. Thus it can tell us how much we lose by using asymptotic theory and how far is it from the exact results. It includes the saddlepoint approximation, the large-T approximation and the small σ approximation.

The saddlepoint approximation can be obtained for any statistic which admits a cumulative generating function. It is based on the Fourier inversion formulae for the density, and applies the steepest decent method to the integration to derive an approximation for the density function. However, it is not widely used in econometrics, see Daniels (1954) and Daniels (1956).

The second approximation method is the large-T approximation. A significant growth in the literature took place following the work of the Sargan school, Nagar school, Basmann school, Anderson school, and P.C B. Phillips school. Most of the contributions of these schools, however, were confined to the analytical derivation of the moments and distribution in the static simultaneous equations model and the dynamic first order autoregressive model, with *i.i.d* normal disturbances. First, we should clarify the difference between the large-T approximation and large sample asymptotic theory. The inferences from the asymptotic theory are simply based on the limiting distribution when T goes to infinity, while, the large-T approximation uses an asymptotic expansion to approximate the exact distribution or moments of the statistic, and then it provides inferences based on some leading terms of the expansion. The accuracy of the large-T approximation increases as the sample size increases. There are two different popular ways to derive this large-T approximation in econometrics; using the Edgeworth approximation and Nagar's approximation. It is useful to distinguish between Sargan school (who used the Edgeworth approximation) and the Nagar school(who focused on finding approximations to estimator moments) to facilitate our discussion. The Sargan school which is exemplified by Sargan (1975), Sargan (1976), Mariano (1972), and Phillips (1980), rigorously developed the theory and applications of the Edgeworth (1896) expansions to derive the approximate distribution function of econometric estimators. The Nagar school, Nagar (1959), Kadane (1971), obtained the approximate moments of the k-class estimators in simultaneous equations. This thesis can be classified as belonging to the Nagar school since the approximation employed is essentially based on a Taylor series expansion to approximate the sampling error (the difference between the statistic and the parameter), so that the successive terms are in the descending order of the sample size T, in probability. He found expressions of bias to the order of T^{-1}, and for the second moment to the order of T^{-2} for general k-class estimators.

To illustrate, suppose a sample size T and an estimate $\hat{\alpha}$ of a coefficient vector α, the large-T approach in Nagar (1959) starts by expanding the estimation error as follows:

$$\sqrt{T}(\hat{\alpha}-\alpha)=\sum_{s=1}^{p}\frac{e_s}{T^{(1/2)(s-1)}}+\frac{r_p}{T^{(1/2)p}}$$

where e_s, for $s = 1, ..., p$, and r_p are all $O_p(1)$ as $T \to \infty$. Here r_p is the remainder term in an expansion up to order p. A bias approximation is then obtained by taking expectations of the terms in the summation.

An alternative is the small-σ approach due to Kadane (1971). This method uses a Taylor series expansion to expand the expression for the estimation error, so that the successive terms are in increasing powers of σ in probability, in contrast to the large-T asymptotic expansion which orders these terms in descending order of the sample size, T, in probability. The general expansion is

$$\frac{1}{\sigma}(\hat{\alpha}-\alpha)=\sum_{s=1}^{p}\sigma^{s-1}\dot{e_s}+\sigma^{p}\dot{r_p}$$

and where $\dot{e_s}$, for $s = 1, ..., p$, and $\dot{r_p}$ are also bounded in probability, this time as σ, the standard deviation of the equation disturbance, tends to zero. The bias is then approximated to an appropriate order by taking expectations of the terms in the summation.

It has been shown that the two approaches give essentially the same bias approximations in the static SEM case. However, as shown by Kiviet and Phillips (1989) the two approximations are not the same in dynamic models and the large-T approach is superior. In addition, the small-σ approximation requires that the disturbance be suitably small and approaches zero in limit. The rationale is that when σ gets progressively smaller the econometric model gives a progressively better explanation of the data. Hence both large-T and small-σ are idealisations.

The third approach and the most recent tool for gaining the information of small sample behaviour was introduced by Efron (1979), Hall (1997). Both Monte Carlo and Bootstrap methods are based on the use of simulation techniques to generate some specific numerical approximations to the sample distribution in selected case. In this thesis, we will explore the bootstrap method in reducing the bias of the estimates based on the approach of Freedman (1984) and Ip (1991), which is the non-parametric residual bootstrap method.

From the literature, it is apparent that most of work done to explore the small sample properties in simultaneous equation models covers only the static case, whilst any work in the dynamic case considers only first order dynamics. This is true, for example, in the work of Phillips and Liu-Evans (2015) which considered the properties of the two stage least squares (2SLS) in a DSEM with just one lagged endogenous variable. We are not aware of any work in the literature which explores the properties of estimators in the general dynamic simultaneous equations model. In the case of the Fuller limited information maximum likelihood (FLIML) estimator, there are no reported results at all for the DSEM even for the one lagged case. In Kiviet, Phillips, and Schipp (1999) a bias approximation was presented for the maximum likelihood estimation of the reduced form parameters of a first order DSEM while again the general case was not considered. However Kiviet and Phillips (1994) had previously presented a bias approximation for the least squares estimator in a single equation dynamic regression of general order. This thesis is interested in extending Kiviet, Phillips, and Schipp (1999), and Kiviet and Phillips (1993) of a one period lagged-dependent variables, and Kiviet and Phillips (1994) of high-order dynamic single regression and that of Phillips and Liu-Evans (2015), to high order dynamic simultaneous equations models.

Two alternative approaches could be used to estimate the simultaneous equations model: limited information estimators(LIE) (single equation estimation) which estimate a system of simultaneous equations by estimating each equation (provided it is identified) separately, and Full information estimators(FIE) (system estimation) which estimate all the (identified) equations in the system simultaneously. FIE is more efficient, but it puts more constraints on the model compared with LIE. FIE incorporates knowledge of all the restrictions in the system when estimating each parameter, while LIE only utilizes knowledge of the restrictions in the particular equation being estimated. Hence, in the FIE approach the misspecification of one behavioural equation affects the other behavioural equations, while the LIE is not as prone to misspecification (due to fewer assumptions). This thesis focuses on the behaviour of LIE in the context of the general dynamic simultaneous equations models in finite sample cases. The commonly used limited information estimators are the least squares (ordinary least squares (OLS), indirect least squares (ILS), generalised least squares (GLS) and two stage least squares (2SLS), and the limited information maximum likelihood (LIML)). When the model has endogenous variables, the OLS estimator could be seriously biased and inconsistent. Hence in this thesis we merely analyse the behaviour of OLS in the reduced form when the sample size is small. As we mentioned before, in truly large samples, asymptotically equivalent estimates should not be very different, however, if the sample size is small or moderate, various asymptotically justified estimators of the same coefficients can assume quite different numerical values, and they exhibit different properties. Such as, in the case of the 2SLS estimator which only possesses finite moments up to the degree of over-identification and the limited information maximum likelihood estimator which does not possess moments of any integral values.

For static models, 2SLS has historically been the most commonly used limited information estimation method. Another popular method is LIML. The fact that LIML does not possess finite moments of any order is well known, see for example Roberto S. Mariano (1972). More recently Chao et al. (2013a), Chao et al. (2013b) provide proofs that LIML does not have any finite moments , which leads to particularly dispersed estimates, see Hahn, Hausman, and Kuersteiner (2004). To solve the problem of LIML estimation Hahn, Hausman, and Kuersteiner (2004), Hausman et al. (2009) suggested the estimator proposed by Fuller (1977) which we refer to as FLIML and which has a k-class representation based on $k = \lambda_{Fuller}$, where $\lambda_{Fuller} = \lambda^* - \frac{\alpha}{T-K} = \min \frac{\beta_*' Y_*' \bar{P}_{Z_1} Y_* \beta_*}{\beta_*' Y_*' \bar{P}_{Z_{Y_* \beta_*}}} - \frac{\alpha}{T-K}$, has all necessary moments and which has a small bias property. In the static case, the FLIML estimator yields estimates which are unbiased to order T^{-1}, and order σ^2 respectively, see Anderson, Kunitomo, and Morimune (1986). As noted earlier,while some literature has explored the small sample properties of 2SLS in a one lagged dependent variable DSEM, no literature has explored the general DSEM case. In reference to the FLIML estimator, I am not aware of any literature that has presented its properties in the DSEM when the sample size is small. Hence in this thesis, we will explore the small sample properties of 2SLS and FLIML estimators in the general DSEM, assess the bias approximation, and subtract the bias from the initial estimator to develop bias corrected estimators.

Chapter 2 examines estimation in over-identified equations for the general DSEM with p lagged dependent variables, strong exogenous variables and innovation errors; moreover, this system is stable. In particular,we show how the new corrected 2SLS estimator gives almost unbiased estimation based on the bias approximation which is obtained by first taking Taylor expansions to order T^{-1}. The bias approximation is then decomposed into the simultaneity bias and dynamic bias components. The dynamic bias follows the results presented in Kiviet, Phillips, and Schipp (1999), Kiviet and Phillips (1993) and Phillips(2011). Interestingly, in our results we observe that the dynamic bias and simultaneity bias have opposite signs which indicates that some correction method which is suitable for the static case may not be so in the case of dynamic models. In this Chapter we also employ the non-parametric bootstrap to correct the bias and the results point out that it could be an alternative way to correct the bias when the computer cost is a consideration. The new bias correction procedure does not lead to an inflation in the mean squared errors. While, the non-parametric bootstrap method leads to an increase in mean squared errors, but this increase is not likely to be substantial.

Chapter 3 contributes to the literature in three ways. First, under regularity conditions, we derive the bias approximation to order $1/T$ for the Fuller limited information maximum likelihood(FLIML) estimator in the general DSEM which has not been reported in previous literature. We do this by using Nagar type expansions which makes the results comparable. Secondly, we compare the analytical bias approximation for FLIML with the corresponding bias approximation for 2SLS and we observe that the FLIML estimator, which removes the $O(T^{-1})$ simultaneity bias completely and dynamic bias partially, gives much less biased estimates. Third, we have constructed a bias corrected procedure for the original FLIML by employing the estimated bias approximation. We observe that this corrected FLIML (CFLIML) gives almost unbiased estimates numerically and analytically. Moreover, the MSE of endogenous and exogenous variables' coefficients in FLIML are much smaller than that of 2SLS. However, the MSEs for some dynamic coefficient estimates are found to increase in FLIML. The corrected estimators do not lead to an inflation in the MSE compared with the non-corrected methods in general.

Chapter 2 and Chapter 3, both analyse the bias approximation in the same general DSEM, with the same assumptions. These two chapters show that the bias corrected estimators, based upon $O(T^{-1})$ can be recommended as a bias reduction technique in the general DSEM.

Independent of the analysis of the structural form model estimator properties, Chapter 4 focuses on exploring the classical estimator OLS in the p^{th} order dynamic reduced form model which comes from the general DSEM. In this chapter, we extend the analysis in Kiviet, Phillips, and Schipp (1999) to the general p^{th}-order dynamic reduced form case. Without losing generality, we focus on the bias in the OLS estimates for the first equation of a multi-equations system, where p lagged endogenous variables are included. The bias approximation to order T^{-1} is derived by using the Nagar expansion, and the bias corrected estimators are constructed by employing the estimated bias approximation. We set up a set of Monte Carlo experiments to examine the performance of COLS and the residual bootstrap OLS in this general reduced form model. The simulations and numerical results suggest that the OLS bias can be substantial which was also observed in Kiviet, Phillips, and Schipp (1999). The COLS estimator gives almost unbiased estimation, and the residual bootstrap method is also well behaved as a bias reduction procedure. From the results, it is obvious that bias correction using the $O(T^{-1})$ bias expansion is more effective compared to the bootstrap method. In addition bias correction with either method does not result in an increase of the MSE. Hence, the bias corrected estimator COLS, based upon the $O(T^{-1})$ bias approximation can be recommended as a bias reduction technique in the p^{th} order dynamic reduced form. Alternatively, the non-parametric bootstrap is also an appropriate way to reduce the bias if the computer cost is a consideration.

In Chapter 5 of this thesis, we provide the summary of conclusions of our research and we suggest the future direction for further research studies in relation to relevant chapters of this thesis.

Chapter 2

The bias of 2SLS estimator in general dynamic simultaneous equations models

2.1 Introduction

To explore the finite sample properties of estimators in the static simultaneous equations model(SEM), Nagar (1959) found the bias approximation for k-class estimators to the order of T^{-1}, and also derived an approximation for the second moment to order T^{-2}, by using asymptotic expansions essentially based on employing Taylor expansions. Later work in this area included Phillips (2000), Mikhail (1972), Hahn and Hausman (2002), and Bun and Windmeijer (2011) examined bias approximation and reduction in the static simultaneous equation models.

In the dynamic regression models, a number of researchers show that least squares estimators can be seriously biased in small samples. They include Grubb and Symons (1987), Hoque and Peters (1986), and Peters (1989). Kiviet, Phillips, and Schipp (1999), Kiviet and Phillips (1993), Kiviet and Phillips (1995) while Phillips and Liu-Evans

(2015) show that the bias in 2SLS in a dynamic simultaneous equation model (DSEM) can be expressed in two parts, a part which derives from simultaneity and a part which is due to the dynamics. However, this latter paper only focuses on the first order DSEM rather than the general DSEM(p lagged dependent variables). In the high order dynamic case, Kiviet and Phillips (1994) present the small sample bias of OLS for the standard ARMAX $(p, 0, k)$ model; however, this is a single equation regression model rather than a DSEM.

In this paper, we are interested in extending the Phillips and Liu-Evans (2015), and Kiviet, Phillips, and Schipp (1999) analysis for the first order DSEM to the general order DSEM assuming that the structural disturbances are normally and independently distributed with mean vector $0'$ and fixed covariance matrix $\Sigma = (\sigma_{ij})$. This general dynamic simultaneous equations model includes the endogenous variables which are lagged p time periods, and strongly exogenous $I(0)$ regressors lagged q time period.

With this model, we analyse the behaviour of 2SLS when sample size is small. Analytically, we derive the bias approximation of 2SLS to order T^{-1}, and confirm the evidence which has been observed in Kiviet, Phillips, and Schipp (1999), Kiviet and Phillips (1993) and Phillips and Liu-Evans (2015), i.e the bias comes from the simultaneity and dynamics respectively. Interestingly, the numerical results show that these two parts actually have opposite signs. In this case, bias correction methods which effectively reduce the bias in the static case (Kiviet and Phillips (1989), Sawa (1973) and Iglesias and Phillips (2012), etc.) may not be suitable for our dynamic models. However, if we subtract the observed bias approximation in estimation from the corresponding estimator, the bias corrected estimator may be unbiased to order T^{-1} theoretically. Kiviet and Phillips (2005) show that $O(\sigma^2)$ bias approximation can be used for corrected 2SLS (C2SLS) estimation of dynamic models. Kiviet, Phillips, and Schipp (1999) and Liu-Evans and Phillips (2012) use the $O(T^{-1})$ bias approximation in

COLS estimation of autoregressive models, and it presents almost unbiased estimators. Phillips and Liu-Evans (2015) show in Monte carlo simulations that by using the C2SLS in the first order DSEM, the new C2SLS method gives almost unbiased estimation. Hence, we develop the bias corrected estimator by employing the estimated bias approximation applied to the traditional 2SLS estimator. Ideally, using the large-T approximation in this paper directly for a reduced-bias estimator may tend to yield more accurate numerical results than any existing approximation. Hence we would expect the $O(T^{-1})$ bias approximation in our paper to yield a substantial improvement over the uncorrected 2SLS estimator.

Our numerical results show that the bias approximation may tend to overstate the magnitude of the "true" bias as given by the Monte Carlo estimates in 2SLS. However, importantly, the bias corrected estimator, based upon $O(T^{-1})$ approximation, very substantially reduces the Monte Carlo 2SLS bias. Moreover, in most cases, it does not inflate the MSE. Hence, the bias corrected estimator, based upon $O(T^{-1})$ bias approximation, can be recommended as a bias reduction technique for practical use. The other alternative bias reduction method is also considered in this chapter. Freedman (1984) pointed out that the residual bootstrap method could be useful in bias reduction in 2SLS estimation, since it may have some effect in eliminating the bias that comes from the dynamic part. Ip (1991) provides strong support that the bootstrap 2SLS can correct bias for both static and dynamic parts to order T^{-1}. In our experiments, the bootstrap method is not as good as C2SLS, but it may still effectively reduce the bias in the 2SLS. When L , the order of over-identification is large, the estimates of endogenous and exogenous coefficients may have small MSE when using the bootstrap 2SLS.

The next section will introduce the general model. Section 2.3 evaluates the bias approximation for the first equation in the structural form. Section 2.4 introduces the new bias correction method C2SLS. The numerical experiments and the associated results are present in section 2.5, and 2.6. In these two sections we also employ the non-parametric residual bootstrap 2SLS estimator. The last section is our conclusion part.

2.2 The Model

The complete system we are interested in:

$$YB + \sum_{i=1}^{p} Y_{-i}A^{(i)} + \sum_{j=0}^{q} X_{-j}C^{(j)} = \tilde{U}, \tag{2.1}$$

where Y is a $T \times G$ matrix of T observations on G endogenous variables, X is a $T \times K$ matrix of observations on K stationary (we will relax this assumption in our further work) and strongly exogenous variables, Y_{-i} is a $T \times G$ matrix of observations on the endogenous variables lagged i time periods (G lagged endogenous explanatory variables) and we assume that the initial values $(Y_{1-p,}, ..., Y_0)$ are non-stochastic. The model also involves K current strictly exogenous variables in the matrix X which is assumed to be of full rank K, and has q lags X_{-j}. They are assumed to be the $I(0)$ process. $\tilde{U}$ is a $T \times G$ matrix of structural disturbances. The matrices $B, A^{(i)}$ and $C^{(j)}$ are of dimension $G \times G, G \times G$ and $K \times G$, respectively, and B is assumed to be non-singular. The rows of $\tilde{U}$ are assumed to be normally and independently distributed with zero mean and fixed covariance matrix $\tilde{\Sigma} = (\tilde{\sigma}_{mn})$.

Furthermore, we assume that the eigenvalues (real or complex values) of the system of difference equations are inside the unit circle which ensures the stability for our system. Thus the roots (real or complex values) of the determinantal equation
$det|B\varpi^p + A^{(1)}\varpi^{p-1} + A^{(2)}\varpi^{p-2} + ... + A^{(p)}| = 0$
are smaller than unity in absolute value: $|\varpi|^h < 1, h = 1, 2, ..., p$. This statement of

the system essentially follows that of Dhrymes (1970), Chapter 12; Davidson (2000), Section 4.3.2.

The reduced form of the model is:

$$\begin{aligned} Y &= -\sum_{i=1}^{p} Y_{-i}A^{(i)}B^{-1} - \sum_{j=0}^{q} X_{-j}C^{(j)}B^{-1} + \tilde{U}B^{-1} \\ &= \sum_{i=1}^{p} L^{i}Y\Gamma^{(i)} + \sum_{j=0}^{q} L^{j}X\Pi^{(j)} + \tilde{V} \\ &= ZA^{*} + \tilde{V}, \end{aligned} \tag{2.2}$$

where $\Gamma^{(i)} = -A^{(i)}B^{-1}$, $\Pi^{(j)} = -C^{(j)}B^{-1}$ and $\tilde{V} = \tilde{U}B^{-1}$. The rows of $\tilde{V}$ are normally and independently distributed with zero mean and covariance matrix $\tilde{\Omega} = (\tilde{\omega}_{mn}) = E(\tilde{V}^{'}\tilde{V})/T$. Also $Z = [R : S]$ is a $T \times (P+Q)$ matrix where $P = \sum_{m=1}^{G} p(m)$ and $Q = \sum_{n=1}^{K} q(n)$. Here the $T \times P$ matrix R includes all the observations for the (stochastic) lagged endogenous variables, and the $T \times Q$ matrix S includes the observations for all the other regressors. A^{*} is the $(P+Q) \times G$ coefficients matrix.

The stochastic part $\tilde{W}$ of $Y = \bar{Y} + \tilde{W}$ from equation (2.2) has rows $\tilde{w}_t^{'}$, $t = 1, 2...T$, which can be written as follows.

$$\begin{aligned}
\tilde{w_1}' &= \tilde{v_1}', \\
\tilde{w_2}' &= \tilde{v_2}' + \tilde{w_1}'\Gamma^{(1)}, \\
\tilde{w_3}' &= \tilde{v_3}' + \tilde{w_1}'\Gamma^{(2)} + \tilde{w_2}'\Gamma^{(1)}, \\
\tilde{w_4}' &= \tilde{v_4}' + \tilde{w_1}'\Gamma^{(3)} + \tilde{w_2}'\Gamma^{(2)} + \tilde{w_3}'\Gamma^{(1)}, \\
&\vdots \\
\tilde{w_p}' &= \tilde{v_p}' + \tilde{w_1}'\Gamma^{(p-1)} + \tilde{w_2}'\Gamma^{(p-2)} + \tilde{w_3}'\Gamma^{(p-3)} + ... + \tilde{w}_{p-1}'\Gamma^{(1)}, \\
\tilde{w_{p+1}}' &= \tilde{v_{p+1}}' + \tilde{w_1}'\Gamma^{(p)} + \tilde{w_2}'\Gamma^{(p-1)} + \tilde{w_3}'\Gamma^{(p-2)} + ... + \tilde{w}_p'\Gamma^{(1)}, \\
&\vdots
\end{aligned}$$

$$\tilde{w_T}^{'} = \tilde{v_T}^{'} + \tilde{w}_{T-p}^{'}\Gamma^{(p)} + ... + \tilde{w}_{T-2}^{'}\Gamma^{(2)} + \tilde{w}_{T-1}^{'}\Gamma^{(1)}.$$

Let the $T \times T$ matrix D be such that,

$$D = \begin{bmatrix} 0 & 0 & . & . & . & . & 0 \\ 1 & 0 & & & & & 0 \\ 0 & 1 & & & & & 0 \\ . & . & . & . & & & . \\ . & . & . & . & & & . \\ 0 & 0 & . & . & 0 & 1 & 0 \end{bmatrix}, \qquad D^2 = \begin{bmatrix} 0 & 0 & . & . & . & . & 0 \\ 0 & 0 & & & & & 0 \\ 1 & 0 & & & & & 0 \\ . & . & . & . & & & . \\ . & . & . & . & & & . \\ 0 & 0 & . & . & 1 & 0 & 0 \end{bmatrix},$$

where $D^{T-1}.D = D^T = 0$ and D^0 is I_T. Also we define a $TG \times G$ matrix J formed by stacking the matrices J_t, $t = 0, 1, ..., T-1$, as follows

$$\begin{aligned}
&J_0 = I_G, \\
&J_1 = \Gamma^{(1)}, \\
&J_2 = \Gamma^{(2)} + \Gamma^{(1)} J_1, \\
&J_3 = \Gamma^{(3)} + \Gamma^{(2)} J_1 + \Gamma^{(1)} J_2, \\
&J_4 = \Gamma^{(4)} + \Gamma^{(3)} J_1 + \Gamma^{(2)} J_2 + \Gamma^{(1)} J_3, \\
&\vdots \\
&J_p = \Gamma^{(p)} + \Gamma^{(p-1)} J^{(1)} + ... + \Gamma^{(1)} J_{p-1}, \\
&J_{p+1} = \Gamma^{(p)} J_1 + ... + \Gamma^{(1)} J_p, \\
&J_{p+2} = \Gamma^{(p)} J_2 + ... + \Gamma^{(1)} J_{p+1}, \\
&\vdots \\
&J_{T-1} = \Gamma^{(p)} J_{T-p} + ... + \Gamma^{(1)} J_{T-2}.
\end{aligned}$$

The matrix J can be written as:

$$J = \begin{bmatrix} J_0 \\ J_1 \\ J_2 \\ \vdots \\ J_p \\ J_{p+1} \\ J_{p+2} \\ \vdots \\ \vdots \\ J_{T-1} \end{bmatrix} = \begin{bmatrix} I & 0 & 0 & 0 & 0 & ... & ... & 0 & 0 & 0 \\ \Gamma^{(1)} & 0 & 0 & 0 & 0 & ... & ... & 0 & 0 & 0 \\ \Gamma^{(2)} & \Gamma^{(1)} & 0 & 0 & 0 & ... & ... & 0 & 0 & 0 \\ \vdots & \vdots & . & . & & & & & \cdots & \vdots \\ \Gamma^{(p)} & \Gamma^{(p-1)} & ... & \Gamma^{(1)} & 0 & ... & ... & & & . \\ 0 & \Gamma^{(p)} & \Gamma^{(p-1)} & ... & \Gamma^{(1)} & 0 & ... & ... & & . \\ 0 & 0 & \Gamma^{(p)} & \Gamma^{(p-1)} & ... & \Gamma^{(1)} & 0 & ... & ... & . \\ \vdots & \vdots & . & . & & & & & \cdots & \vdots \\ \vdots & \vdots & . & . & & & & & \cdots & \vdots \\ 0 & .... & . & \cdots & \Gamma^{(p)} & \Gamma^{(p-1)} & ... & & \cdots & \Gamma^{(1)} \end{bmatrix} \begin{bmatrix} I \\ J_1 \\ J_2 \\ \vdots \\ J_p \\ J_{p+1} \\ J_{p+2} \\ \vdots \\ J_{T-2} \end{bmatrix}.$$

With these definitions $\tilde{W}$ can be represented in terms of $\tilde{V}$ as follows:

$$\tilde{W} = \sum_{t=1}^{T-1} D^t \tilde{V} J_t + \tilde{V} = \sum_{t=0}^{T-1} D^t \tilde{V} J_t. \tag{2.3}$$

In equation (2.3), $\sum_{i=0}^{T-1} D^t \tilde{V} J_t$ is the stochastic part of Y. In equation (2.2) $Z = [R : S]$, and in accordance with our notation, Z may be decomposed as:

$$Z = \bar{Z} + \tilde{W}^*. \tag{2.4}$$

Here $\bar{Z} = [\bar{R} : X]$ is taken to be the non-stochastic part of Z, whose component matrix $\bar{R}$ is the non-stochastic part of R. The stochastic part of Z is $\tilde{W}^* = \omega W^*$, and $\tilde{W}^*$ can be expressed as:

$$\tilde{W}^* = [\tilde{R} : 0] = [L\tilde{W} : L^2\tilde{W} : .... : L^p\tilde{W} : 0]$$

$$= [\sum_{t=1}^{T-1} D^t \tilde{V} J_{t-1} : \sum_{t=2}^{T-1} D^t \tilde{V} J_{t-2} : ... : \sum_{t=p}^{T-1} D^t \tilde{V} J_{t-p} : 0].$$

The standardized form of $\tilde{W}^*$ can be presented as:

$$\tilde{W^*} = \left[\sum_{i=1}^{p} \sum_{t=i}^{T-1} D^t \tilde{V} J_{t-i} \Psi_i^{'} : 0 \right], \tag{2.5}$$

where $\Psi_i^{'} = e_i^{'} \otimes I_G$ is $G \times P$ matrix with $P = \sum_{m=1}^{G} p(m)$ and where all component $G \times G$ matrices are zero except the i^{th} which is an identity matrix. e_i is the $p \times 1$ unit vector with all elements equal to zero except the i^{th} which is unity.

2.3 Structural Form Estimation–Two Stage Least Square Estimation

In this section we derive the large T approximations to the bias of 2SLS estimators when estimating the structural coefficients of the first equation which forms part of the complete system equation (2.1), and we shall write the equation as:

$$y_1 = Y_2 \beta_1 + \sum_{i=1}^{p} L^i Y_1 a_1^{(i)} + \sum_{j=0}^{q} L^j X_1 c_1^{(j)} + \tilde{u}_1 = \Upsilon \delta_1 + \tilde{u}_1, \tag{2.6}$$

where

$$\Upsilon = [Y_2 : R_1 : S_1] \quad and \quad \delta_1^{'} = (\beta_1, a_1^{(1)}, ..., a_1^{(p)}, c_1^{(1)}, ..., c_1^{(q)}).$$

Here $Y_1 = [y_1 : Y_2]$ is a $T \times (g+1)$ matrix of observations on $g+1$ included endogenous variables. $L^i Y_1$ is the i period lagged version of Y_1, X_1 is a $T \times k$ matrix of observations on k stationary exogenous variables. Υ is a $T \times (g + P^* + Q^*)$ matrix which includes the $T \times g$ matrix Y_2, the $T \times P^*$ matrix R_1 contains the lagged endogenous regressor values and the $T \times Q^*$ matrix S_1 contains the exogenous regressor values which are taken as fixed. $P^* = \sum_{m=1}^{g+1} p(m)$ and $Q^* = \sum_{n=1}^{k} q(n)$ which allows for the equations to contain different numbers of lagged endogenous and exogenous regressors respectively. δ_1 is a $(g+P^*+Q^*)\times 1$ vector which contains all the structural form parameters. We shall denote:

$$\bar{\Upsilon} = [\bar{Y}_2 : \bar{R}_1 : \bar{S}_1] \quad and \quad \tilde{F} = [\tilde{W}_2 : \tilde{R}_1 : 0] \tag{2.7}$$

as, respectively, the non-stochastic and stochastic parts of Υ which will be used in later analysis. Notice that the non-stochastic part of Y contains $\bar{Y}_2$ and $\bar{R}_1$ which are the unconditional expectations of Y_2 and R_1 respectively. Note also that in $\tilde{F}$, $\tilde{W}_2$ is the relevant stochastic part of $\tilde{W}$ for Y_2 as given in equation (2.3).

The standard 2SLS estimator of δ_1 can be written as:

$$\begin{aligned}\hat{\delta_1} &= (\hat{\Upsilon}'\hat{\Upsilon})^{-1}\hat{\Upsilon}'y_1 \\ &= \delta_1 + (\hat{\Upsilon}'\hat{\Upsilon})^{-1}\hat{\Upsilon}'\tilde{u}_1\end{aligned} \tag{2.8}$$

where

$$\hat{\Upsilon} = [\hat{Y}_2 : R_1 : S_1] \quad and \quad \hat{Y}_2 = \sum_{i=1}^{p} L^i Y \hat{\Gamma}_2^{(i)} + \sum_{j=0}^{q} L^j X \hat{\Pi}_2^{(j)}$$

and $\hat{Y}_2$ is obtained when the reduced form equation (2.2) is estimated by OLS. The matrix R_1 which refers to $L^i Y_1$ is $T\times P^*$, where $P^* = \sum_{m=1}^{g+1} p(m)$. $\hat{\Gamma}_2^{(i)}$, $i = 1,2,...,p$ and $\hat{\Pi}_2^{(j)}$, $j = 1,2,...,q$ are respectively, $G\times g$ and $K\times g$ matrices of estimated reduced form coefficients in equation (2.2). $\hat{\Upsilon}$ can be also decomposed into non-stochastic part $\bar{\Upsilon}$ and stochastic part $(\hat{\Upsilon}-\bar{\Upsilon})$, hence:

$$\hat{\Upsilon} = \bar{\Upsilon} + (\hat{\Upsilon}-\bar{\Upsilon}).$$

Using equation (2.7) and (2.8), and $\bar{Y}_2 = \sum_{i=1}^{p} L^i \bar{Y} \Gamma_2^{(i)} + \sum_{j=0}^{q} L^j X \Pi_2^{(j)}$, the stochastic part of $\hat{\Upsilon}$ can be written as:

$$
\begin{aligned}
\hat{\Upsilon} - \bar{\Upsilon} &= \left[\sum_{i=1}^{p} L^i \bar{Y} (\hat{\Gamma}_2^i - \Gamma_2^i) + \sum_{j=1}^{q} L^j X (\hat{\Pi}_2^j - \Pi_2^j) + \sum_{i=1}^{p} L^i \tilde{W} \hat{\Gamma}_2^i : \sum_{i=1}^{p} L^i \tilde{W}_1 : 0 \right] \qquad (2.9) \\
&= \left[\sum_{i=1}^{p} L^i \bar{Y} (\hat{\Gamma}_2^{(i)} - \Gamma_2^{(i)}) + \sum_{j=1}^{q} L^j X (\hat{\Pi}_2^j - \Pi_2^j) + \sum_{i=1}^{p} L^i \tilde{W} (\hat{\Gamma}_2^{(i)} - \Gamma_2^{(i)}) : 0 : 0 \right] \\
&\quad + [\sum_{i=1}^{p} L^i \tilde{W} \Gamma_2^{(i)} : \tilde{R}_1 : 0].
\end{aligned}
$$

We define:

$$
\begin{aligned}
\Delta_1 &= \left[\sum_{i=1}^{p} L^i \bar{Y} (\hat{\Gamma}_2^{(i)} - \Gamma_2^{(i)}) + \sum_{j=1}^{q} L^j X (\hat{\Pi}_2^j - \Pi_2^j) + \sum_{i=1}^{p} L^i \tilde{W} (\hat{\Gamma}_2^{(i)} - \Gamma_2^{(i)}) : 0 : 0 \right], \qquad (2.10) \\
\Delta_2 &= [\sum_{i=1}^{p} L^i \tilde{W} \Gamma_2^{(i)} : \tilde{R}_1 : 0].
\end{aligned}
$$

Then $\hat{\Upsilon} = \bar{\Upsilon} + \Delta_1 + \Delta_2$ and it is possible to write

$$
\begin{aligned}
\hat{\Upsilon}' \hat{\Upsilon} &= \bar{\Upsilon}' \bar{\Upsilon} + \Delta_1' \Delta_1 + \Delta_2' \Delta_2 + \bar{\Upsilon}' \Delta_1 + \bar{\Upsilon}' \Delta_2 + \Delta_1' \bar{\Upsilon} + \Delta_2' \bar{\Upsilon} + \Delta_1' \Delta_2 + \Delta_2' \Delta_1 \qquad (2.11) \\
&= \bar{\Upsilon}' \bar{\Upsilon} + E(\Delta_2' \Delta_2) + \Delta_1' \Delta_1 + (\Delta_2' \Delta_2 - E(\Delta_2' \Delta_2)) + (\bar{\Upsilon}' \Delta_1 + \Delta_1' \bar{\Upsilon}) \\
&\quad + (\bar{\Upsilon}' \Delta_2 + \Delta_2' \bar{\Upsilon}) + (\Delta_1' \Delta_2 + \Delta_2' \Delta_1).
\end{aligned}
$$

Let $H^{-1} = \bar{\Upsilon}' \bar{\Upsilon} + \mathbb{E}(\Delta_2' \Delta_2)$ which is $O(T)$, then put the $O_p(T^{1/2})$ component of $\hat{\Upsilon}' \hat{\Upsilon}$ as J_1^* and the $O_p(1)$ component as J_2^*. Where, J_1^* includes $(\Delta_2' \Delta_2 - \mathbb{E}(\Delta_2' \Delta_2))$, $((\bar{\Upsilon}' \Delta_1 + \Delta_1' \bar{\Upsilon})$, and $(\bar{\Upsilon}' \Delta_2 + \Delta_2' \bar{\Upsilon})$ $(\Delta_1' \Delta_2 + \Delta_2' \Delta_1)$, and the component of J_2^* is $\Delta_1' \Delta_1$. We can then express $(\hat{\Upsilon}' \hat{\Upsilon})^{-1}$ from equation (2.11)as follows:

$$
\begin{aligned}
(\hat{\Upsilon}' \hat{\Upsilon})^{-1} &= (H^{-1} + J_1^* + J_2^*)^{-1} = H(I + J_1^* H + J_2^* H)^{-1} \qquad (2.12) \\
&= H - H J_1^* H + o_p(T^{-3/2})
\end{aligned}
$$

and noting that $\hat{\Upsilon} = \bar{\Upsilon} + \Delta_1 + \Delta_2$, we have

$$\hat{\Upsilon}'\tilde{u}_1 = \bar{\Upsilon}'\tilde{u}_1 + \Delta_1'\tilde{u}_1 + \Delta_2'\tilde{u}_1. \tag{2.13}$$

Here $\Delta_1'\tilde{u}_1$ is $O_p(1)$, $\hat{\Upsilon}'\tilde{u}_1$ and $\Delta_2'\tilde{u}_1$ are $O_p(T^{1/2})$. Combining equation (2.12) with (2.13) gives:

$$\begin{aligned}\hat{\delta}_1 - \delta_1 = (\hat{\Upsilon}'\hat{\Upsilon})^{-1}\hat{\Upsilon}'\tilde{u}_1 = H\bar{\Upsilon}\tilde{u}_1 + H\Delta_1'\tilde{u}_1 + H\Delta_2'\tilde{u}_1 - HJ_1^*H\bar{\Upsilon}'\tilde{u}_1 \\ - HJ_1^*H\Delta_2\tilde{u}_1 + o_p(T^{-1}).\end{aligned} \tag{2.14}$$

Taking expectations term by term yields the 2SLS bias, and this is given in *Theorem 1* below. Defining $H^{*-1} = \mathbb{E}(Z'Z)$, where recall that $Z = [\tilde{R} : S]$ which includes all the lagged endogenous variables and all the exogenous variables and let $I_2 = \begin{bmatrix} I_P \\ 0 \end{bmatrix}$ which is $(P+Q) \times P$ selection matrix, then $I_2' H^* I_2 = H^{**}$, a sub-matrix of H^*. We also define the matrix $C^* = \begin{bmatrix} \Gamma_2^{(*)} & : & I_1 & : & 0 \end{bmatrix}$ which is $P \times (g + P^* + Q^*)$ matrices. It contains the $P \times g$ matrix $\Gamma_2^{(*)} = (\Gamma_2^{(1)}, \Gamma_2^{(2)}, ..., \Gamma_2^{(p)})'$, the $P \times P^*$ selection matrix I_1, and the $P \times Q^*$ matrix $(0_{P \times Q^*})$. It then follows that we may write $\Delta_2 = [\sum_{i=1}^{p} L^i \tilde{W} \Gamma_2^{(i)} : \tilde{R}_1 : 0] = \tilde{R}C^*$, where $\tilde{R} = [LW : L^2W : ... : L^pW]$ includes all the stochastic part of lagged dependent variables. We will use this expression for further calculations in the appendix. Define $\tau = \sigma^2\phi$ and $\vartheta = \Lambda^{**'}\tau$, ϕ is defined by using, Nagar (1959), the decomposition for $\tilde{V}$, $\tilde{V} = S^* + \tilde{u}_1\phi'$, where S^* and $\tilde{u}_1$ are normally and independently distributed. Then $\phi\sigma^2 = \mathbb{E}\left(\frac{1}{T}\tilde{V}'\tilde{u}_1\right)$. We define $\Lambda^{**} = \begin{bmatrix} I_g : 0 \\ 0 \end{bmatrix}$ which is a $G \times (g + P^* + Q^*)$ dimension selection matrix, then $\tilde{V}_2$ is the $T \times g$ sub-matrix of matrix $\tilde{V}$, which can be expressed as $[\tilde{V}_2 : 0 : 0] = \tilde{V} \begin{bmatrix} I_g : 0 \\ 0 \end{bmatrix}$. With these and earlier definitions of terms we may state the following:

Theorem 1 . The bias of the 2SLS estimator of the first structural equation parameters to order T^{-1} is given by:

$$
\begin{aligned}
\mathbb{E}(\hat{\delta}_1-\delta_1) =& Htr\{\bar{Z}H^*\bar{Z}^{'}\}\vartheta - H\bar{\Upsilon}^{'}\bar{Z}H^*\bar{Z}^{'}\bar{\Upsilon}H\vartheta \qquad (2.15)\\
&- Htr\{\bar{Z}H^*\bar{Z}^{'}\bar{\Upsilon}H\bar{\Upsilon}^{'}\}\vartheta\\
&+ H\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i,j}^{T-1}(T-t)tr\{\Omega J_{t-i}\Psi_i^{'}H^{**}\Psi_j J_{t-j}^{'}\}\vartheta\\
&- H\sum_{l=1}^{p}\sum_{j=1}^{p}\sum_{r=l,j}^{T-1}(T-r)\left(tr\left\{\bar{Z}H^*I_2\Psi_l J_{r-l}^{'}\Omega J_{s-j}\Psi_j^{'}C^*H\bar{\Upsilon}^{'}\right\}\right)\vartheta\\
&- H\sum_{i=1}^{p}\sum_{l=1}^{p}\sum_{t=i,j}^{T-1}(T-t)C^{*'}\Psi_i J_{t-i}^{'}\Omega J_{t-j}\Psi_j^{'}I_2^{'}H^*\bar{Z}^{'}\bar{\Upsilon}H\vartheta\\
&- H\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i,j}^{T-1}(T-t)(tr\{\Omega J_{t-i}\Psi_i^{'}I_2^{'}H^*\bar{Z}^{'}\bar{\Upsilon}HC^{*'}\Psi_l J_{t-l}^{'}\})\vartheta\\
&- H\bar{\Upsilon}^{'}\bar{Z}H^*\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i,j}^{T-1}I_2(T-t)\Psi_i J_{t-i}^{'}\Omega J_{t-j}\Psi_j^{'}C^{*'}H\vartheta\\
&- H\sum_{i=1}^{p}\sum_{l=1}^{p}\sum_{t=i,l}^{T-1}(T-t)\left(tr\left\{\Omega J_{t-i}\Psi_i^{'}H^{**}\left[\sum_{j=1}^{p}\sum_{m=1}^{p}\sum_{s=j,m}^{T-1}(T-s)\Psi_j\right.\right.\right.\\
&\qquad\left.\left.\left.\times J_{s-j}\Omega J_{s-m}\Psi_m^{'}C^*\right]HC^{*'}\Psi_l J_{s-l}^{'}\right\}\right)\vartheta\\
&- H\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i,j}^{T-1}(T-t)C^{*'}\Psi_i J_{t-i}^{'}\Omega J_{t-j}\Psi_j^{'}H^{**}\\
&\quad\times\sum_{l=1}^{p}\sum_{m=1}^{p}\sum_{r=l,m}^{T-1}(T-l)\Psi_l J_{r-l}^{'}\Omega J_{r-m}\Psi_m^{'}C^*H\vartheta\\
&- H\bar{\Upsilon}^{'}\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1}D^tD^{s'}\bar{\Upsilon}H(tr\{\Omega J_{s-j}\Psi_j^{'}H^{**}\Psi_i J_{t-i}^{'}\})\vartheta\\
&- H\bar{\Upsilon}^{'}\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1}D^tD^s\bar{\Upsilon}H\Lambda^{**'}\Omega J_{t-i}\Psi_i^{'}H^{**}\Psi_j J_{s-j}\tau\\
&- H\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1}(tr\{\Psi_j J_{s-j}^{'}\Omega J_{t-i}\Psi_i^{'}H^{**}\})(tr\{D^tD^{s'}\bar{\Upsilon}H\bar{\Upsilon}^{'}\})\vartheta\\
&- H\Lambda^{**'}\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1}\Omega J_{s-j}\Psi_j^{'}H^{**}\Psi_i J_{t-i}^{'}(tr\{D^t\bar{\Upsilon}H\bar{\Upsilon}^{'}D^s\})\tau
\end{aligned}
$$

$$
\begin{aligned}
&- H\bar{\Upsilon}^{'}\sum_{i=1}^{p}\sum_{t=i}^{T-1} D^{t}\bar{\Upsilon}H^{'}C^{*'}\Psi_{i}J_{t-i}^{'}\tau \\
&- H\sum_{i=1}^{p}\sum_{t=i}^{T-1} C^{*'}\Psi_{i}J_{t-i}^{'}(tr\{\bar{\Upsilon}^{'}D^{t'}\bar{\Upsilon}H\})\tau \\
&- H\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1}(tr\{D^{t}D^{s'}\bar{Z}H^{*}I_{2}\Psi_{i}J_{t-i}^{'}\Omega J_{s-j}\Psi_{j}^{'}C^{*}H\bar{\Upsilon}^{'}\})\vartheta \\
&- H\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1}\Lambda^{**'}\Omega J_{s-j}\Psi_{j}^{'}C^{*}H\bar{\Upsilon}^{'}D^{t'}D^{s'}\bar{Z}H^{*}I_{2}\Psi_{i}J_{t-i}^{'}\tau \\
&- H\Lambda^{**}\sum_{l=1}^{p}\sum_{r=l}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1}\Omega J_{s-j}\Psi_{j}^{'}C^{*}H\bar{\Upsilon}^{'}D^{r'}D^{s}\bar{Z}H^{*}I_{2}\Psi_{l}J_{r-l}^{'}\tau \\
&- H\Lambda^{**}\sum_{l=1}^{p}\sum_{r=l}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1}\Omega J_{r-l}\Psi_{l}^{\prime}I_{2}^{\prime}H^{*}\bar{Z}^{\prime}D^{r'}D^{s}\bar{\Upsilon}HC^{*'}\Psi_{j}J_{s-j}^{\prime}\tau \\
&- H\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1} C^{*'}\Psi_{i}J_{t-i}^{'}\Omega J_{s-j}\Psi_{j}^{'}I_{2}^{'}H^{*}\bar{Z}^{'}D^{t}D^{s'}\bar{\Upsilon}H\vartheta \\
&- H\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1} C^{*'}\Psi_{i}J_{t-i}^{'}\Omega\Lambda^{**}H\bar{\Upsilon}^{'}D^{r'}D^{t'}\bar{Z}H^{*}I_{2}\Psi_{j}J_{s-j}^{'}\tau \\
&- H\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1} C^{*'}\Psi_{i}J_{t-i}^{'}(tr\{D^{t'}D^{s}\bar{Z}H^{*}I_{2}\Psi_{j}J_{s-j}^{'}\Omega\Lambda^{**}H\bar{\Upsilon}^{'}\})\tau \\
&- H\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1} C^{*'}\Psi_{i}J_{t-i}^{'}\Omega\Lambda^{**}H\bar{\Upsilon}^{'}D^{s'}D^{t}\bar{Z}H^{*}I_{2}\Psi_{j}J_{s-j}^{'}\tau \\
&- H\bar{\Upsilon}^{'}\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1} D^{t}D^{s}tr\left\{\Omega J_{t-i}\Psi_{i}^{\prime}C^{*}H\Lambda^{**'}\right\}I_{2}\bar{Z}H^{*}I_{2}\Psi_{j}J_{s-j}^{\prime}\tau \\
&- H\bar{\Upsilon}^{'}\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1} D^{t}D^{s'}\bar{Z}H^{*}I_{2}\Psi_{j}J_{s-j}^{\prime}\Omega J_{t-i}\Psi_{i}^{\prime}C^{*}H\vartheta \\
&- H\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1}\Lambda^{**'}tr\left\{\bar{Z}H^{*}I_{2}\Psi_{i}J_{t-i}^{\prime}\Omega J_{s-j}\Psi_{j}^{\prime}C^{*}H\bar{\Upsilon}^{\prime}D^{t}D^{s'}\right\}\tau \\
&- H\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1}\Lambda^{**'}\Omega J_{s-j}\Psi_{j}^{\prime}C^{*}H\bar{\Upsilon}^{\prime}D^{t}D^{s}\bar{Z}H^{*}I_{2}\Psi_{i}J_{t-i}^{\prime}\tau \\
&- H\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1} C^{*'}\Psi_{i}J_{t-i}^{'}(tr\{\Omega\Lambda^{**}HC^{*'}\Psi_{j}J_{s-j}^{'}\}) \\
&\qquad\qquad \times (tr\{D^{t'}\bar{Z}H^{*}\bar{Z}^{'}D^{s'}\}.I)\tau
\end{aligned}
$$

$$-H\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1}(tr\{D^{t'}\bar{Z}H^{*}\bar{Z}^{'}D^{s'}\})C^{*'}\Psi_{i}J_{t-i}^{'}\Omega HC^{*'}\Psi_{j}J_{s-j}^{'}\vartheta$$

$$-H\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1}(tr\{\Omega J_{t-i}\Psi_{i}^{'}C^{*}HC^{*'}\Psi_{j}J_{s-j}^{'}\})(tr\{\bar{Z}H^{*}\bar{Z}^{'}D^{t}D^{s'}\})\vartheta$$

$$-H\Lambda^{**'}\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1}\Omega J_{s-j}\Psi_{j}^{'}C^{*}HC^{*'}\Psi_{i}J_{t-i}^{'}(tr\{\bar{Z}H^{*}\bar{Z}^{'}D^{t}D^{s}\})\tau$$

$$+o(T^{-1}).$$

A proof of this result is given by Appendix A.1 and it is obtained by evaluating the expectations of each term.

From the result, we note that the bias of 2SLS to order T^{-1} of the first structural form equation has two distinct parts: a part is due to the simultaneity of the system which is represented by the first three terms in the above, and a part which is due to the dynamic nature of the structural equation which is represented by all the remaining terms. Here, $H = (\bar{\Upsilon}'\bar{\Upsilon} + \mathbb{E}(\Delta_2'\Delta_2))^{-1}$, $H^* = (\bar{Z}'\bar{Z} + \mathbb{E}(\tilde{W}^{*'}\tilde{W}^{*})^{-1})$ and $H^{**} = I_2'H^*I_2 = I_2^{'}(\bar{Z}'\bar{Z} + \mathbb{E}(\tilde{W}^{*'}\tilde{W}^{*})^{-1})I_2$, from which, we observe that the bias that comes solely from the simultaneity terms should not include the expected stochastic parts in the first three terms. In fact, the expression in Theorem 1 should reduce to the Nagar (1959) bias approximation in static models when any terms that result from the inclusion of lagged endogenous regressors are removed. This means that a reduction of the above result to that for the static case will obtain with the removal of any terms involving the "D" matrix and the expected stochastic parts in the first three terms, and this may be shown to be the case. Note that the first ten items without D terms, will be removed by using the FLIML which will be analysed in chapter 3. The numerical results will be discussed in section 2.6.

2.4 Bias corrected 2SLS Estimator

Biased corrected 2SLS estimator for structural form equations parameters can be obtained by estimating the approximating bias and then subtracting this bias estimate from the corresponding estimator. As we showed in section 2.3, the bias approximations depend upon the reduced form coefficient matrices $\Gamma^{(1)}, \Gamma^{(2)} ..., \Gamma^{(p)}, \Pi^{(1)}, \Pi^{(2)} ..., \Pi^{(q)}$, the non-stochastic matrices $X, LX, ...L^q X$, the starting values $y_0^{'}, Ly_0^{'}, L^2 y_0^{'}$ and $L^3 y_0^{'}$ vectors. The $G \times 1$ column vector τ, is equal to $\sigma^2 \phi = \mathbb{E}(\frac{1}{T} \tilde{V}' \tilde{u}_1)$.

To obtain the estimated bias terms, the reduced form parameter matrices $\Gamma^{(1)}, \Gamma^{(2)} ...,$ $\Gamma^{(p)}, \Pi^{(1)}, \Pi^{(2)} ..., \Pi^{(q)}$ are replaced by their OLS estimates. The $G \times 1$ column vector τ is estimated from $[Y - \sum_{i=1}^{p} L^i Y \hat{\Gamma}^{(i)} - \sum_{j=0}^{q} L^{(j)} X \Pi^{(j)}]'(y_1 - \Upsilon \hat{\delta}_1)/T$, the inner product of the G reduced form residuals vectors and the first equation of structural form residuals vector, which is obtained when equation 2.6 is estimated by 2SLS. Then ϑ is replaced by estimated ϑ , where $\hat{\vartheta} = \Lambda^{**'} \hat{\tau}$, where $\Lambda^{**'} = \begin{bmatrix} I_g : 0 \\ 0 \end{bmatrix}$ which is $G \times (g + P^* + Q^*)$ dimension selection matrix.

Definition 1. ***Given*** $\hat{\delta}_{1,b(2SLS)}$ ***is estimated 2SLS bias approximations for the coefficient bias*** $\delta_{1,b(2SLS)} = \left(\beta_{1,b(2SLS)}, \alpha^{(1)}_{1,b(2SLS)}, ..., \alpha^{(p)}_{1,b(2SLS)}, c^{(0)}_{1,b(2SLS)}, c^{(1)}_{1,b(2SLS)}, ..., c^{(q)}_{1,b(2SLS)}\right)$, ***and given that*** $\hat{\delta}_{1,2SLS}$ ***is the 2SLS estimator of*** δ_1***, the C2SLS bias corrected estimator*** $\hat{\delta}_{1,C2SLS}$ ***is as following:***

$$\hat{\delta}_{1,C2SLS} = \hat{\delta}_{1,2SLS} - \hat{\delta}_{1,b(2SLS)}. \tag{2.16}$$

To examine how well the C2SLS works for practical bias correction, a set of Monte Carlo experiments were conducted and the results are discussed in section 2.6.

2.5 Numerical Experiments Design

2.5.1 Numerical Model

The experiments were conducted using a three equation dynamic simultaneous equation model with four lagged endogenous variables based on sample sizes 50 and 100. Hence the matrix of endogenous variables is $Y = (y_1, y_2, y_3)$. Under the condition for the existence of the moments for the 2SLS estimator [1], in our experiment the degree of over-identification L is greater or equal to 2, so that 2SLS estimates possess a finite mean and variance. In our experiments, we chose $L = 2$, 4 and 6 . To commence, we generated two exogenous variables in each equation respectively. L is varied by augmenting the exogenous variables in both second and third equations. Hence, when $L = 2$, the exogenous variable matrix $X = (x_1, x_2, x_3, x_4, x_5, x_6)$; when $L = 4$, $X = (x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8)$; and when $L = 6$, $X = (x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10})$. Following the $I(0)$ strictly exogenous assumption in section 2.2, each exogenous variable is generated as Gaussian autoregressive process with mean zero and with an autoregressive coefficient of 0.9, and they are independent of each other.

$$x_{jt} = 0.9 x_{j(t-1)} + \varsigma_{jt} \qquad \varsigma_{jt} \sim \mathcal{N}(0, 1).$$

The coefficient matrices are as follows (The coefficients are chosen arbitrarily following with the stability assumption in section 2.2):

$$B = \begin{bmatrix} 1 & -1.11 & -3 \\ -\beta_{21} & 1 & -4.6 \\ -\beta_{31} & -8 & 1 \end{bmatrix}, \qquad A^{(1)} = \begin{bmatrix} \alpha_{11}^{(1)} & 0.56 & -0.45 \\ -\alpha_{21}^{(1)} & -0.62 & 0.28 \\ -\alpha_{31}^{(1)} & -0.90 & -0.32 \end{bmatrix},$$

[1] Sargan (1974) showed that the moments of the 2SLS exist up to the order of over-identification in the static SEM. We shall assume the result is valid for the DSEM also.

$$A^{(2)} = \begin{bmatrix} \alpha_{11}^{(2)} & -0.80 & -0.82 \\ -\alpha_{21}^{(2)} & 0.72 & -0.90 \\ -\alpha_{31}^{(2)} & -0.50 & 0.78 \end{bmatrix}, \qquad A^{(3)} = \begin{bmatrix} -\alpha_{11}^{(3)} & -0.46 & -0.80 \\ -\alpha_{21}^{(3)} & -0.72 & 0.31 \\ -\alpha_{31}^{(3)} & -0.31 & 0.74 \end{bmatrix},$$

$$A^{(4)} = \begin{bmatrix} \alpha_{11}^{(4)} & -0.36 & -0.2 \\ -\alpha_{21}^{(4)} & -0.46 & 0.58 \\ -\alpha_{31}^{(4)} & 0.58 & 0.70 \end{bmatrix}.$$

$$L = 2 \quad C^{'} = \begin{bmatrix} c_{11} & c_{21} & c_{31} & 0.00 & 0.00 & 0.00 & 0.00 \\ -1.00 & 0.00 & 0.00 & 0.75 & -0.24 & 0.00 & 0.00 \\ 1.00 & 0.00 & 0.00 & 0.00 & 0.00 & -0.15 & 0.86 \end{bmatrix};$$

$$L = 4 \quad C' = \begin{bmatrix} c_{11} & c_{21} & c_{31} & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ -1.00 & 0.00 & 0.00 & 0.75 & -0.24 & 0.35 & 0.00 & 0.00 & 0.00 \\ 1.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & -0.15 & 0.86 & -0.58 \end{bmatrix};$$

$$L = 6 \quad C' = \begin{bmatrix} c_{11} & c_{21} & c_{31} & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ -1.00 & 0.00 & 0.00 & 0.75 & -0.24 & 0.35 & 0.68 & 0.00 & 0.00 & 0.00 & 0.00 \\ 1.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & -0.15 & 0.86 & -0.58 & 0.33 \end{bmatrix}.$$

There are 17 coefficients in the first equation to be estimated and they are given below:

$$\begin{array}{llllll} \beta_{21} = 2.00, & \beta_{31} = 5.00, & \alpha_{11}^{(1)} = 0.50, & \alpha_{21}^{(1)} = 0.36, & \alpha_{31}^{(1)} = 0.40, & \alpha_{11}^{(2)} = 1.20, \\ \alpha_{21}^{(2)} = 0.60, & \alpha_{31}^{(2)} = -0.38, & \alpha_{11}^{(3)} = 0.65, & \alpha_{21}^{(3)} = 1.20, & \alpha_{31}^{(3)} = 0.38, & \alpha_{11}^{(4)} = 0.50, \\ \alpha_{21}^{(4)} = 0.60, & \alpha_{31}^{(4)} = -0.20, & c_{11} = 1.00, & c_{21} = 0.60, & c_{31} = -0.50. & \end{array}$$

The model disturbances are generated as standard normal random variables. The reduced form of the model is:

$$Y = LY\Gamma^{(1)} + L^2 Y\Gamma^{(2)} + L^3 Y\Gamma^{(3)} + L^4 Y\Gamma^{(4)} + X\Pi + \tilde{V},$$

where $\tilde{V} = (\tilde{v}_1, \tilde{v}_2, \tilde{v}_3)$ is a $T \times 3$ matrix of reduced form disturbances. We use a matrix P from a Choleski factorisation of Ω to generate the reduced form errors. Hence each row of $\tilde{V}$ is obtained from the transpose of

$$\begin{bmatrix} \tilde{v}_{1,t} \\ \tilde{v}_{2,t} \\ \tilde{v}_{3,t} \end{bmatrix} = P \begin{bmatrix} \tilde{e}_{1,t} \\ \tilde{e}_{2,t} \\ \tilde{e}_{3,t} \end{bmatrix},$$

where $\tilde{e}_{1,t}$, $\tilde{e}_{2,t}$ and $\tilde{e}_{3,t}$ denote the standardised disturbances. Each row of $\tilde{U}$ has mean $0'$ and covariance matrix Σ, and is i.i.d.. $\tilde{U}$ is the structural form disturbances. Hence the distribution of the structural disturbances can be evaluated from

$$B'\tilde{v}_t = \tilde{u}_t \Rightarrow \tilde{u}_t \sim \mathcal{N}(0, \Sigma) \qquad where \quad \Sigma = B'\Omega B.$$

We set the structural covariance matrix is as follows:

$$\Sigma = \begin{bmatrix} 0.3524 & 0.3448 & 0.3112 \\ 0.3448 & 0.3668 & 0.2984 \\ 0.3112 & 0.2984 & 0.4064 \end{bmatrix},$$

from which the reduced form covariance is:

$$\Omega = \begin{bmatrix} 0.0055 & 0.0054 & 0.0030 \\ 0.0054 & 0.0844 & 0.0085 \\ 0.0030 & 0.0085 & 0.0069 \end{bmatrix}.$$

Based on the above parameters, the relevant eigenvalues of the reduced form equations which determine the stationarity condition can be calculated from the following determinantal equation.

$$det \mid \Gamma^{(4)} + \varpi\Gamma^{(3)} + \varpi^2\Gamma^{(2)} + \varpi^3\Gamma^{(1)} - \varpi^4 I_3 \mid = 0.$$

All the roots ϖ are complex , but they are inside the unit circle Holmgren (2000), which ensures the stability of this system.

$$\varpi_1 = 0.6947 + 0.4789i, \quad \varpi_2 = 0.6947 - 0.4789i, \quad \varpi_3 = -0.0561 + 0.5955i,$$
$$\varpi_4 = -0.0561 - 0.5955i, \quad \varpi_5 = 0.0996 + 0.7235i, \quad \varpi_6 = 0.0996 - 0.7235i,$$
$$\varpi_7 = -0.5847 + 0.0909i, \quad \varpi_8 = -0.5847 - 0.0909i \quad \varpi_9 = -0.2039 + 0.4099i,$$
$$\varpi_{10} = -0.2039 - 0.4099i, \quad \varpi_{11} = 0.4688 + 0.0000i, \quad \varpi_{12} = -0.2651 + 0.0000i.$$

This system above is slightly different from the general model equation (2.1). We have normalized with respect to $\beta_{11} = 1 = \beta_{22} = \beta_{33}$. To achieve the general case (with high lag order), we choose 4 lags (most macroeconomic data are quarterly data). While many of the simulations conducted in the literature focus on two equation models, we decided to simulate a three equations model in this paper.

The initial values, $y_0^{'}, Ly_0^{'}, L^2y_0^{'}, L^3y_0^{'}$ are generated by averaging the simulated reduced form 1000 times. We first take the expectation of the reduced form, where

$\mathbb{E}(y^{'}) = \mathbb{E}(y^{'})_{-1} = \mathbb{E}(y^{'})_{-2} = \mathbb{E}(y^{'})_{-3} = \mathbb{E}(y^{'})_{-4}$, and $\bar{x}' = \mathbb{E}(X)$, which is as follows:

$$\mathbb{E}(y^{'}) = \mathbb{E}(y^{'})\Gamma^{(1)} + \mathbb{E}(y^{'})\Gamma^{(2)} + \mathbb{E}(y^{'})\Gamma^{(3)} + \mathbb{E}(y^{'})\Gamma^{(4)} + \bar{x}'\Pi.$$

From it we obtain $\mathbb{E}(y^{'})$. Then using this 1×3 vector $\mathbb{E}(y^{'})$ as the starting value in the reduced form to generate $T \times G$ matrix $(Y_0)_1$ which is $\Big((y^{'}_{0_1})_1, (y^{'}_{0_2})_1 ... (y^{'}_{0_{T-1}})_1, (y^{'}_{0_T})_1\Big)^{'}$. Following this procedure, we generate the $M = 1,000$ sets of $T \times G$ matrices Y_0 which is $(Y_0)_1, (Y_0)_2 ... (Y_0)_{M-1}, (Y_0)_M$. Then the pool of initial value Y_0 is $Y_0 = \sum_{m=1}^{M}(Y_0)_m$ which is $T \times G$ matrix. Hence, the initial value in this four lagged dependent variables model is given by $y^{'}_0 = \sum_{m=1}^{M}(y^{'}_{0_T})_m/M$, $Ly^{'}_0 = \sum_{m=1}^{M}(y^{'}_{0_{T-1}})_m/M$, $L^2y^{'}_0 = \sum_{m=1}^{M}(y^{'}_{0_{T-2}})_m/M$ and $L^3y^{'}_0 = \sum_{1}^{M}(y^{'}_{0_{T-3}})_m/M$.

2.5.2 The Simulation model

The number of Monte Carlo replications is 20,000, while 199 bootstrap replicates are used when constructing the bias corrected bootstrap.

Bootstrap

Based on Freedman (1984), Ip (1991) provides support for the asymptotic validity of the 2SLS bootstrap in static and dynamic models where errors are normal, and MacKinnon (2002) conducted hypothesis testing in static model which also supports the asymptotic validity of the 2SLS bootstrap.

The residual bootstrap 2SLS is simulated by first estimating the equation of interest using 2SLS. Then by using the estimates and resampling the estimated residuals, pseudo-data (B sets) are generated. Bootstrap replicates are obtained by implementing 2SLS on each of B sets. The bias corrected bootstrap estimate of δ_1 can be calculated as $2\hat{\delta}_1 - \hat{\delta}_{1,\bar{b}}$, where $\hat{\delta}_1$ is the original estimate, and $\hat{\delta}_{1,\bar{b}}$ is the mean of the bootstrap replicates.

Freedman's bootstrap remains the same steps as the usual residual bootstrap, except the generation of the pseudo data.

Our target is to estimate

$$y_1 = Y_2\beta_1 + LY_1\alpha_1^{(1)} + L^2Y_1\alpha_1^{(2)} + L^3Y\alpha_1^{(3)} + L^4Y_1\alpha_1^{(4)} + X_1c_1 + \tilde{u}_1 \tag{2.17}$$

and we would like to generate the pseudo data y_1^* , LY_1^* , $L^2Y_1^*$, $L^3Y_1^*$, $L^4Y_1^*$ and Y_2^* from equation (2.17) by resampling the residuals $\hat{u}_{1,2SLS}$. However, the first element y_1^* cannot be obtained without knowing the first element of Y_2^*. Hence, we use the reduced form of Y_2, which is estimated by OLS as,

$$Y_2 = LY\hat{\Gamma}_2^{(1)} + L^2Y\hat{\Gamma}_2^{(2)} + L^3Y\hat{\Gamma}_2^{(3)} + L^4Y\hat{\Gamma}_2^{(4)} + X\hat{\Pi}_2 + \hat{\tilde{V}}_2. \tag{2.18}$$

Equation (2.18) is used in conjunction with the 2SLS estimate of equation (2.17), which will become,

$$y_1 = Y_2\hat{\beta}_1 + LY_1\hat{\alpha}_1^{(1)} + L^2Y_1\hat{\alpha}_1^{(2)} + L^3Y_1\hat{\alpha}_1^{(3)} + L^4Y_1\hat{\alpha}_1^{(4)} + X_1\hat{c}_1 + \hat{\tilde{u}}_1. \tag{2.19}$$

Then, we can resample the $\hat{\tilde{u}}_1$ in equation (2.19) to generate $\hat{\tilde{u}}_1^*$ and then resample the $\hat{\tilde{V}}_2$ in equation (2.18) to give $\hat{\tilde{V}}_2^*$. Note that the disturbances are resampled from the rows of $(\hat{\tilde{u}}_1, \hat{\tilde{V}}_2)$, so that the elements in the resampled residuals $\hat{\tilde{u}}_1^*$ and $\hat{\tilde{V}}_2^*$ correspond to each other.

Based on the resampled residuals, we can generate the pseudodata which we need. Here we use the same procedure as we defined in section 2.5 to get the initial values of $y_0^{'*}, Ly_0^{'*}, L^2y_0^{'*}, L^3y_0^{'*}$, but the parameters now are replaced by the estimated value in this 2SLS-bootstrap method. Then, it is possible for us to generate $y_{2_1}^{'*}$ from equation (2.18) by using $\hat{\tilde{v}}_{2_1}^{'*}$. This can be then substituted into equation (2.19) and used with

$Ly^{'*}_{1_1}$, $L^2y^{'*}_{1_1}$,$L^3y^{'*}_{1_1}$, $L^4y^{'*}_{1_1}$ and $\hat{\hat{u}}^*_{1_1}$ to generate $y^*_{1_1}$. Then $y^*_{1_1}$ can be put into equation (2.18) to generate the second vector element of (Y^*_2) which can be used in (2.19) to give the next element (y^*_1) to put in equation (2.18). Continuing this iteration gives the full vectors of y^*_1 , LY^*_1 , $L^2Y^*_1$, $L^3Y^*_1$, $L^4Y^*_1$ and Y^*_2.

Finally the actual data are replaced by pseudodata to estimate the equation of interest by using the traditional 2SLS estimation method. Thus, Y^*_2 is regressed on $(LY^* : L^2Y^* : L^3Y^* : L^4Y^* : X)$ in order to generate the fitted values $\hat{Y}^*_2$, the y^*_1 is regressed on $(\hat{Y}^*_2 : LY^* : L^2Y^* : L^3Y^* : L^4Y^* : X_1)$ to give the bootstrap 2SLS replicates $\hat{\beta}^*_{1,b}$, $\hat{\alpha}^{(1)*}_{1,b}$, $\hat{\alpha}^{(2)*}_{1,b}$, $\hat{\alpha}^{(3)*}_{1,b}$, $\hat{\alpha}^{(4)*}_{1,b}$ and $\hat{c}^*_{1,b}$. For each $\delta_1 \in (\beta_1, \alpha^{(1)}, \alpha^{(2)}, \alpha^{(3)}, \alpha^{(4)}, c_1)$, the bias corrected bootstrap 2SLS estimate is given by $\hat{\hat{\delta}}_{1,b} = 2\hat{\delta}_{1,2SLS} - \hat{\delta}^*_{1,\bar{b}}$,where $\hat{\delta}^*_{1,\bar{b}} = \frac{1}{B}\sum_{b=1}^{B}\hat{\delta}^*_{1,b}$.

The bias corrected bootstrap based on our numerical design is as follows:

Definition 1. Given $\hat{\delta}_{1,\bar{b}}$ as the mean of the bootstrap 2SLS replicates for the coefficient $\delta_1 = (\beta_1, \alpha^{(1)}, \alpha^{(2)}, \alpha^{(3)}, \alpha^{(4)}, c_1)$, and given $\hat{\delta}_{1,2SLS}$ as the 2SLS estimator of δ_1, the bootstrap bias corrected estimator $\hat{\hat{\delta}}_{1,b}$ is as follows:

$$\hat{\hat{\delta}}_{1,b} = 2\hat{\delta}_{1,2SLS} - \hat{\delta}^*_{1,\bar{b}}.$$

2.6 Numerical Results

The numerical results show a comparison of the performance of 2SLS, and the residual bootstrap 2SLS and C2SLS, which is summarized in Table A.1 to Table A.3. Table A.1 reports the overall bias approximation, simultaneity bias, and dynamic biases, respectively. Table A.2 presents the bias of 2SLS, the bias of Bootstrap 2SLS, and the bias of C2SLS respectively. Table A.3 presents the MSE of 2SLS, Bootstrap 2SLS, and C2SLS respectively. β_{21}, β_{31} are the coefficients of endogenous variables of the first structural form equation. α^1_{11} to α^4_{31} are the coefficients of the lagged endogenous variables (4 lagged endogenous variables). c_{11} is the constant, and c_{21}, c_{31} are the parameters of exogenous regressors.

Table A.1 shows that the bias approximation may tend to overstate the magnitude of the "true" bias as given by the Monte Carlo estimates in 2SLS. For example, when $\beta_{21} = 2.00$, $L = 2$, and $T = 50$, the 2SLS bias is -0.3042, whilst the bias approximation slightly higher than the 2SLS bias of -0.3229. Moreover, when we numerically evaluate the dynamic bias and the simultaneity bias separately, the results show that they have opposite signs. If we still look at the coefficient above, the approximated bias is -0.3229 where -0.5322 comes from simultaneity, and 0.2093 comes from the dynamics. It implies that if the bias correction method can only eliminate either the simultaneity bias or the dynamic bias but not both, then instead of decreasing the overall bias, the bias correction method could possibly provide more biased estimates. Hence, a bias correction method which effectively reduces the bias in the static case may not do so in the dynamic case.

When the sample size increases, both approximated bias and the bias of 2SLS decreases. At the same time, when the order of over-identification L increases, this is followed by an increase in the 2SLS bias and a corresponding increase in the approximation.

The results for the corrected 2SLS(C2SLS) estimator which was constructed by subtracting the bias estimate are presented in Table A.2. This bias corrected estimator, based upon $O(T^{-1})$ approximation, significantly reduces the 2SLS bias. For α_{31}^2, the coefficient of $L^2 y_3$ in the first equation, in fact when $L = 2$ and sample size is $T = 50$, by using the new C2SLS estimator, the bias reduced from +61% to +9%. Generally, C2SLS gives almost unbiased estimators in both sample size 50 and 100, when over-identification level is $L = 2, 4$ and 6. The alternative approach based on the non-parametric residual bootstrap applied to 2SLS also reduces the bias effectively; in most cases the bootstrap 2SLS gives almost unbiased estimates when sample size is 100. However, in general, compared with C2SLS, the performance of bootstrap 2SLS is weaker in reducing the bias. As we have shown for α_{31}^2, when the bootstrap method is used the bias is reduced to +30%, and when the sample size increases to $T = 100$ and over-identification level is still $L = 2$, both these two bias correction methods yield almost unbiased results eliminating around a 15% bias from the 2SLS estimator. It is clear that, generally, these two bias corrected estimators have a substantially smaller bias than their uncorrected counterparts.

Table A.3 reports the MSE of 2SLS, C2SLS, and Bootstrap 2SLS respectively . Generally, the MSE for C2SLS is smaller than the corresponding MSE for the 2SLS while both are smaller than the bootstrap 2SLS MSE. Interestingly, the MSE of the bootstrap 2SLS is lower than that of the 2SLS for the coefficient of endogenous variables and exogenous variables when $L = 4, 6$, in both sample size sets. In few cases, the MSE of C2SLS is slightly larger than that of 2SLS because of the almost unbiasedness estimates of 2SLS itself when sample size is large. However, this increasing is trivial. For α_{21}^2, which is the coefficient of $L^2 y_2$ in the first equation of the structural form, when sample size is 100 and the over-identification level is $L = 2, 4$ and 6, the percentage of bias for 2SLS is −2% , −3%, −2% and the MSE is 0.0375, 0.0294, 0.0250, while for C2SLS, the MSE is 0.0398, 0.0294, 0.0272. It is clear that the C2SLS has the smallest MSE, and the bootstrap 2SLS has the largest MSE . However,the MSE of the bootstrap 2SLS is not far from the results for 2SLS, and when L increases, the difference becomes smaller.

2.7 Conclusion

The $O(T^{-1})$ bias in 2SLS estimation of a general DSEM can be decomposed into two parts, which come from the simultaneity and dynamics respectively. These two bias components may be of opposite signs which indicates that the bias correction used should be able to reduce the bias that comes from both components; otherwise the overall bias could become absolutely larger. Notice that the bias approximation tends to overstate the magnitude of the "true" bias as given by the Monte Carlo estimates in 2SLS. Even so, the bias corrected estimator, based upon the $O(T^{-1})$ approximation, very substantially reduces the 2SLS bias. In addition, it was found to be better overall in terms of MSE, as there is no inflation of the 2SLS MSE. Hence, from the theoretical and analytic analysis, the bias corrected estimator, based upon $O(T^{-1})$ can be recommended as a bias reduction technique.

The bootstrap simulation results in this paper provide evidence in support of the alternative bias correction technique based on the bootstrap. It performs particularly well in bias correction. While the bias correction is not as effective as C2SLS, the computer cost is less which may be a consideration. The bootstrap also reduces the MSE in 2SLS for both endogenous and exogenous variables when L is large.

Chapter 3

A Comparison of Limited Information Estimators in Dynamic Simultaneous Equations Models

3.1 Introduction

An important distinction of this paper is that it is the first paper that investigates the performance of the modified limited information maximum likelihood (Fuller (1977)(FLIML)) estimator in the general dynamic simultaneous equation model (DSEM) when the sample size is small.

This chapter compares the bias approximations of 2SLS and FLIML in the general DSEM without any restrictions on the order of the dynamics, whereby there may be any number of lagged regressor variables, both lagged endogenous and exogenous, provided that the required identification conditions are met. This chapter represents a considerable extension of earlier work cited below. The corresponding bias approximation in 2SLS is given in Chapter 2. I then evaluate the bias approximation for FLIML and represent this result in the context of the approximated bias in 2SLS to make these two bias approximations comparable. An improved estimator which is unbiased to order T^{-1} is proposed by incorporating estimates of the bias approximation into the corresponding estimators, i.e. 2SLS and FLIML.

Analytically, the FLIML estimator gives much less biased estimates than the 2SLS; it removes the simultaneity bias completely and dynamic bias partially to order T^{-1}. This is a result which has not been observed in the previous literature. The bias corrected estimators, CFLIML and C2SLS, give almost unbiased results to order T^{-1}. Numerically, the Monte Carlo experiments also support the analytical analysis. In addition, in general the mean squared errors of CFLIML and C2SLS are less than the associated uncorrected estimators.

The linear simultaneous equations model is one of the most important models in economics and there are several estimators for estimating its unknown parameters. For static models, two stage least squares (2SLS) has, historically, been the the most commonly used limited information estimation method. Another popular method is that of limited information maximum likelihood, LIML. The fact that LIML does not possess finite moments of any order is well known, see for example Roberto S. Mariano (1972). More recently Chao et al. (2013a), Chao et al. (2013b) also provide proofs that LIML does not have any finite moments , which leads to particularly dispersed estimates, see Hahn, Hausman, and Kuersteiner (2004). To solve the problem of limited information maximum likelihood estimation Hahn, Hausman, and Kuersteiner (2004), Hausman et al. (2009) suggested the estimator proposed by Fuller (1977) which we refer to as FLIML. This estimator has a k-class representation based on $k = \lambda_{Fuller}$, where $\lambda_{Fuller} = \lambda^{*} - \frac{\alpha}{T-K} = \min \frac{\beta_*' Y_*' \bar{P}_{Z_1} Y_* \beta_*}{\beta_*' Y_*' \bar{P}_{Z_{Y_* \beta_*}}} - \frac{\alpha}{T-K}$. It has all necessary moments and has a small bias property. In the static case, the FLIML estimator yields estimates which are unbiased to order T^{-1}, and order σ^2 respectively, see Anderson, Kunitomo, and Morimune (1986).

However, I am not aware of any work in the literature which explores the properties of FLIML in the dynamic simultaneous equations model; in particular, it is not known if the approximate unbiasedness property which obtains in static models carries over to the dynamic case. My paper is interested to explore this area.

The seminal paper is Nagar (1959) where the approximation employed is essentially based on a Taylor series expansion to approximate the sampling error (the difference between the statistic and the parameter), so that the successive terms are in the descending order of the sample size T, in probability. He found expressions for the bias to the order of T^{-1}, and for the second moment to the order of T^{-2}; for general k-class estimators.

To illustrate, suppose a sample size T and an estimate $\hat{\alpha}$ of a coefficient vector α, the large-T approach in Nagar (1959) starts by expanding the estimation error as follows:

$$\sqrt{T}(\hat{\alpha}-\alpha)=\sum_{s=1}^{p}\frac{e_s}{T^{(1/2)(s-1)}}+\frac{r_p}{T^{(1/2)p}},$$

where e_s, for $s=1,...,p$, and r_p are all $O_p(1)$ as $T\to\infty$. Here r_p is the remainder term in an expansion up to order p. A bias approximation is then obtained by taking expectations of the terms in the summation.

An alternative is the small-σ approach due to Kadane (1971). This method uses a Taylor series expansion to expand the expression for the estimation error, so that the successive terms are in a power series of σ in probability, in contrast to the large-T asymptotic expansion which orders these terms in descending order of the sample size, T, in probability. The general expansion is

$$\frac{1}{\sigma}(\hat{\alpha}-\alpha)=\sum_{s=1}^{p}\sigma^{s-1}\dot{e_s}+\sigma^p\dot{r_p}$$

and where $\dot{e_s}$, for $s = 1, ..., p$, and $\dot{r_p}$ are also bounded in probability, this time as σ, the standard deviation of the equation disturbance, tends to zero. The bias is then approximated to an appropriate order by taking expectations of the terms in the summation.

It has been shown that the two approaches give essentially the same bias approximations in the static SEM case. However, as shown by Kiviet and Phillips (1989) the two approximations are not the same in dynamic models and the large-T approach is superior. In addition, the small-σ approximation requires that the disturbance be suitably small and approaches zero in limit. The rationale is that when σ gets progressively smaller the econometric models gives a progressively better explanation of the data. Hence both large-T and small-σ are idealisations.

Hahn and Hausman (2002), and Bun and Windmeijer (2011) investigate the use of bias approximation and reduction in static simultaneous equations models. These papers also use the asymptotic expansion to obtain the bias approximation to order T^{-1}, then investigate the performance of estimated approximated bias in bias reduction.

There has been some research which has explored the properties of estimators in the DSEM. The earliest known work is Kiviet and Phillips (1989). Phillips and Liu-Evans (2015) shows in Monte carlo simulations that by using the C2SLS in the first order DSEM, the new C2SLS method gives almost unbiased estimation. This work however, considered only the restricted case of one lag for the lagged endogenous regressor variables, yet it is recognised that often dynamic models will have higher order lags especially when quarterly models are specified. Chapter 2 extends this paper into p high order lags. Moreover in this chapter a bias approximation was used to obtain a bias corrected estimator which was shown to be, at least, comparable to the bootstrap in reducing bias. In this chapter, based on Chapter 2 and the extended general model, I analytically and numerically compare the corresponding results of 2SLS in Chapter 2 with the FLIML estimator results. The bias corrected 2SLS (Chapter 2), C2SLS, and corrected FLIML, CFLIML, are also presented in this paper.

The structure of the chapter is as follows. The next section outlines the general model. Section 3.3 evaluates and compares the bias approximation of FLIML and 2SLS estimators. Section 3.4 investigates how to apply the approximated bias to correct the original estimators and construct new estimators (C2SLS/CFLIML). Numerical experiments and results are present in section 3.5. The conclusion is presented in the section 3.6.

3.2 Model

In this Chapter, without losing generality, we are still focusing on analysing the first equation of a general dynamic simultaneous equations model, with innovation errors and p lagged-endogenous and q strongly exogenous, explanatory variables, which has been discussed in Chapter 2:

$$y_1 = Y_2\beta_1 + \sum_{i=1}^{p} L^i Y_1 a_1^{(i)} + \sum_{j=0}^{q} L^j X_1 c_1^{(j)} + \tilde{u}_1 = \Upsilon\delta_1 + \tilde{u}_1, \tag{3.1}$$

where

$$\Upsilon = [Y_2 : R_1 : S_1] \quad and \quad \delta_1^{'} = (\beta_1', a_1^{(1)'}, ..., a_1^{(p)'}, c_1(1)', ..., c_1(q)').$$

Here $Y_1 = [y_1 : Y_2]$ is a $T \times (g+1)$ matrix of observations on $g+1$ included endogenous variables. $L^i Y_1$ is the i period lagged version of Y_1, X_1 is a $T \times k$ matrix of observations on k stationary exogenous variables.Υ is a $T \times (g + P^* + Q^*)$ matrix which includes the $T \times g$ matrix Y_2, the $T \times P^*$ matrix R_1 which contains the lagged endogenous regressor values and the $T \times Q^*$ matrix S_1 which contains the exogenous regressor values which are taken as fixed. $P^* = \sum_{m=1}^{g+1} p(m)$ and $Q^* = \sum_{n=1}^{k} q(n)$ which allows for the equations to contain different numbers of lagged endogenous and exogenous regressors respectively. δ_1 is a $(g+P^*+Q^*)\times 1$ vector which contains all the structural form parameters. We shall denote:

$$\bar{\Upsilon} = [\bar{Y}_2 : \bar{R}_1 : \bar{S}_1] \quad and \quad \tilde{F} = [\tilde{W}_2 : \tilde{R}_1 : 0] \tag{3.2}$$

as the non-stochastic and stochastic parts of Υ which will be used in later analysis. Notice that the non-stochastic part of Y contains $\bar{Y}_2$ and $\bar{R}_1$ which are the unconditional expectations of Y_2 and R_1 respectively. Note also that in $\tilde{F}$, $\tilde{W}_2$ is the relevant stochastic part of $\tilde{W}$ for Y_2 as given in equation (2.3).

3.3 Fuller Limited information maximum likelihood Estimators

The Fuller Limited Information Maximum Likelihood (FLIML) estimator, see Fuller (1977), is based on a modification of the LIML estimator and can be written in the form of k-class estimator, where $k = \lambda - \dfrac{a}{T-K}$, λ is the smallest root of the determinantal equation,

$$det|Y_1(I-P_{X_1})Y_1 - \lambda Y_1(I-P_X)Y_1| = 0,$$

$P_{X_1} = X_1(X_1'X_1)^{-1}X_1'$, $P_X = X(X'X)^{-1}X'$, and X_1 and X are the sets of exogenous variables in the first equation and the whole system respectively. K is the total number of exogenous variables in the system, and a is a positive integer in the range 1 to 4, Fuller found that in the static case when $a=1$, the estimator has a small bias, and when $a=4$ the estimator has smallest MSE but the bias is typically larger than when $a=1$. We will explore the behaviour of FLIML when $a=1$ by comparing the FLIML bias approximation with the 2SLS bias approximation. The FLIML estimate we shall write as:

$$\hat{\delta}_{FLIML} = \begin{bmatrix} Y_2{}'Y_2 - \left(\lambda - \dfrac{1}{T-P-Q}\right)\hat{V_2}{}'\hat{V_2} & Y_2'\left[R_1 : S_1\right] \\ \left[R_1 : S_1\right]' Y_2 & \left[R_1 : S_1\right]'\left[R_1 : S_1\right] \end{bmatrix}^{-1} \tag{3.3}$$

$$\times \begin{bmatrix} Y_2' - \left(\lambda - \dfrac{1}{T-P-Q}\right)\hat{V_2}{}' \\ \left[R_1 : S_1\right]' \end{bmatrix} y_1$$

$$= (\hat{\Upsilon}_F'\Upsilon_F)^{-1}\hat{\Upsilon}_F' y_1$$

$$= \delta_1 + (\hat{\Upsilon}_F'\Upsilon_F)^{-1}\hat{\Upsilon}_F' u_1,$$

where $\hat{\Upsilon}_F = \left[\hat{Y}_2 + (1-(\lambda - \frac{1}{T-P-Q}))\hat{V}_2 : R_1 : S_1\right]$ and $\Upsilon_F = \Upsilon = [Y_2 : R_1 : S_1]$. $\hat{Y}_2$ is obtained as the predicted value of Y_2 when the reduced form equation (2.2) is estimated by OLS. Υ_F, and $\hat{\Upsilon}_F$ can be decomposed into their non-stochastic part $\bar{\Upsilon}_F$ and stochastic part $\Upsilon_F - \bar{\Upsilon}_F$, and $(\hat{\Upsilon}_F - \bar{\Upsilon}_F)$ respectively, hence:

$$\hat{\Upsilon}_F = \bar{\Upsilon}_F + (\hat{\Upsilon}_F - \bar{\Upsilon}_F),$$

$$\Upsilon_F = \bar{\Upsilon}_F + (\Upsilon_F - \bar{\Upsilon}),$$

using equation (3.3), and noting that $\Upsilon_F = \Upsilon$, $\bar{\Upsilon}_F = \bar{\Upsilon}$ and $\bar{Y_2} = \sum_{i=1}^{p} L^i \bar{Y}\Gamma_2^{(i)} + \sum_{j=0}^{q} L^j X \Pi_2^{(j)}$.

So that the above stochastic parts can be written as:

$$\hat{\Upsilon}_F - \bar{\Upsilon}_F = \left[\hat{Y}_2 - \bar{Y}_2 + (1-(\lambda - \frac{1}{T-P-Q}))\hat{V}_2 : (R_1 - \bar{R}_1) : 0\right] \tag{3.4}$$

$$= \left[\hat{Y}_2 - \bar{Y}_2 : R_1 - \bar{R}_1 : 0\right] + \left[(1-(\lambda - \frac{1}{T-P-Q}))\hat{V}_2 : 0 : 0)\right].$$

Also,

$$\Upsilon_F - \bar{\Upsilon}_F = \left[Y_2 - \bar{Y}_2 : R_1 - \bar{R}_1 : 0\right] \tag{3.5}$$
$$= \left[\hat{Y}_2 - \bar{Y}_2 : R_1 - \bar{R}_1 : 0\right] + \left[\hat{V}_2 : 0 : 0\right].$$

Then,

$$\hat{\Upsilon}_F' \Upsilon_F = \hat{\Upsilon}'\hat{\Upsilon} + (1 - (\lambda - \frac{1}{T-P-Q})) \begin{bmatrix} \hat{V}_2'\hat{V}_2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

where $\hat{\Upsilon}'\hat{\Upsilon}$ is the counterpart expression in 2SLS. Recall that from equation 2.12:

$$(\hat{\Upsilon}'\hat{\Upsilon})^{-1} = (H^{-1} + J_1^* + J_2^*)^{-1} = H - HJ_1^*H + o_p(T^{-3/2}). \tag{3.6}$$

The extra term in FLIML is $(1-(\lambda - \frac{1}{T-P-Q}))\hat{V}_2'\hat{V}_2 = (1-(\lambda - \frac{1}{T-P-Q}))\tilde{V}_2'(I - P_z)\tilde{V}_2$ which is $O_p(1)$. Hence ,

$$(\hat{\Upsilon}_F'\Upsilon_F)^{-1} = H - HJ_1^*H + o_p(T^{-3/2}), \tag{3.7}$$

so that $(\hat{\Upsilon}_F'\Upsilon_F)^{-1}$ and $(\hat{\Upsilon}'\hat{\Upsilon})^{-1}$ are the same to order $T^{-3/2}$. The remaining part of FLIML compared with 2SLS in Chapter 2 is:

$$\hat{\Upsilon}_F'\tilde{u}_1 = \hat{\Upsilon}\tilde{u}_1 + \left[(1 - (\lambda - \frac{1}{T-P-Q}))\hat{V}_2 : 0 : 0\right]' \tilde{u}_1 \tag{3.8}$$

where, $\hat{\Upsilon}'\tilde{u}_1$ is the relevant part in 2SLS. Recall from equation 2.13 and 2.10 the definitions of Δ_1 and Δ_2 :

$$\hat{\Upsilon}\tilde{u}_1 = \bar{\Upsilon}'\tilde{u}_1 + \Delta_1'\tilde{u}_1 + \Delta_2'\tilde{u}_2.$$

Then combining equations 3.7 and 3.8 gives the approximate estimation error as,

$$\begin{aligned}\hat{\delta}_{FLIML} - \delta &= (\hat{\Upsilon}'_{FLIMl}\Upsilon_{FLIML})^{-1}\hat{\Upsilon}'_{FLIML}\tilde{u}_1 \\ &= \underbrace{H\bar{\Upsilon}\tilde{u}_1 + H\Delta_1'\tilde{u}_1 + H\Delta_2'\tilde{u}_1 - HJ_1^*H\bar{\Upsilon}'\tilde{u}_1 - HJ_1^*H\Delta_2\tilde{u}_1}_{\text{The same bias as in 2SLS}} \\ &\quad \underbrace{+H\left[(1-(\lambda - \frac{1}{T-P-Q}))\hat{V}_2 : 0 : 0\right]'\tilde{u}_1}_{\text{The extra term compared with 2SLS}} + o_p(T^{-1}).\end{aligned} \tag{3.9}$$

Notice that it is the last term in the above which gives an extra bias approximation term to the FLIML expansion compared with 2SLS and taking expectations term by term will yield the FLIML bias approximation to order T^{-1}. As I have mentioned in the Theorem 1(Chapter 2), FLIML eliminates the first ten terms which are without matrix D terms in the Theorem 1. This result is shown in *Theorem 2.*

Theorem2: FLIML gives less biased estimates compared to 2SLS, it removes completely that part of the $O(T^{-1})$ bias approximation which is associated with the simultaneity bias. Moreover FLIML also partially removes the bias comes from the dynamic part. The bias of the FLIML estimator of the first structural form equation to order T^{-1} is given by:

$$\begin{aligned}\mathbb{E}\left(\hat{\delta}_{FLIML} - \delta\right) = &- H\bar{\Upsilon}'\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1} D^t D^s \bar{\Upsilon} H \Lambda^{**'}\Omega J_{t-i}\Psi_i' H^{**}\Psi_j J_{s-j}\tau \\ &- H\bar{\Upsilon}'\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1} D^t D^{s'} \bar{\Upsilon} H(tr\{\Omega J_{s-j}\Psi_j' H^{**}\Psi_i J_{t-i}'\}.I)\vartheta\end{aligned} \tag{3.10}$$

$$-H\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1}(tr\{\Psi_j J'_{s-j}\Omega J_{t-i}\Psi'_i H^{**}\}.I)(tr\{D^t D^{s'}\bar{\Upsilon} H\bar{\Upsilon}'\}.I)\vartheta$$

$$-H\Lambda^{**'}\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1}\Omega J_{s-j}\Psi'_j H^{**}\Psi_i J'_{t-i}(tr\{D^t\bar{\Upsilon} H\bar{\Upsilon}' D^s\}.I)\tau$$

$$-H\bar{\Upsilon}'\sum_{i=1}^{p}\sum_{t=i}^{T-1} D^t\bar{\Upsilon} H' C^{*'}\Psi_i J'_{t-i}\tau$$

$$-H\sum_{i=1}^{p}\sum_{t=i}^{T-1} C^{*'}\Psi_i J'_{t-i}(tr\{\bar{\Upsilon}' D^{t'}\bar{\Upsilon} H\}.I)\tau$$

$$-H\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1}(tr\{D^t D^{s'}\bar{Z} H^* I_2\Psi_i J'_{t-i}\Omega J_{s-j}\Psi'_j C^* H\bar{\Upsilon}'\}.I)\vartheta$$

$$-H\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1}\Lambda^{**'}\Omega J_{s-j}\Psi'_j C^* H\bar{\Upsilon}' D^{t'} D^{s'}\bar{Z} H^* I_2\Psi_i J'_{t-i}\tau$$

$$-H\Lambda^{**}\sum_{l=1}^{p}\sum_{r=l}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1}\Omega J_{s-j}\Psi'_j C^* H\bar{\Upsilon}' D^{r'} D^s\bar{Z} H^* I_2\Psi_l J'_{r-l}\tau$$

$$-H\Lambda^{**}\sum_{l=1}^{p}\sum_{r=l}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1}\Omega J_{r-l}\Psi'_l I'_2 H^*\bar{Z}' D^{r'} D^s\bar{\Upsilon} H C^{*'}\Psi_j J'_{s-j}\tau$$

$$-H\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1} C^{*'}\Psi_i J'_{t-i}\Omega J_{s-j}\Psi'_j I'_2 H^*\bar{Z}' D^t D^{s'}\bar{\Upsilon} H\vartheta$$

$$-H\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1} C^{*'}\Psi_i J'_{t-i}\Omega\Lambda^{**} H\bar{\Upsilon}' D^{r'} D^{t'}\bar{Z} H^* I_2\Psi_j J'_{s-j}\tau$$

$$-H\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1} C^{*'}\Psi_i J'_{t-i}(tr\{D^{t'} D^s\bar{Z} H^* I_2\Psi_j J'_{s-j}\Omega\Lambda^{**} H\bar{\Upsilon}'\}.I)\tau$$

$$-H\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1} C^{*'}\Psi_i J'_{t-i}\Omega\Lambda^{**} H\bar{\Upsilon}' D^{s'} D^t\bar{Z} H^* I_2\Psi_j J'_{s-j}\tau$$

$$-H\bar{\Upsilon}'\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1} D^t D^s tr\left\{\Omega J_{t-i}\Psi'_i C^* H\Lambda^{**'}\right\} I\bar{Z} H^* I_2\Psi_j J'_{s-j}\tau$$

$$-H\bar{\Upsilon}'\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1} D^t D^{s'}\bar{Z} H^* I_2\Psi_j J'_{s-j}\Omega J_{t-i}\Psi'_i C^* H\vartheta$$

$$-H\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1}\Lambda^{**'} tr\left\{\bar{Z} H^* I_2\Psi_i J'_{t-i}\Omega J_{s-j}\Psi'_j C^* H\bar{\Upsilon}' D^t D^{s'}\right\}\tau$$

$$-H\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1}\Lambda^{**'}\Omega J_{s-j}\Psi'_j C^* H\bar{\Upsilon}' D^t D^s\bar{Z} H^* I_2\Psi_i J'_{t-i}\tau$$

$$
\begin{aligned}
&- H\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1} C^{*'}\Psi_i J_{t-i}^{'}(tr\{\Omega\Lambda^{**}HC^{*'}\Psi_j J_{s-j}^{'}\}.I)\\
&\qquad\qquad \times (tr\{D^{t'}\bar{Z}H^{*}\bar{Z}^{'}D^{s'}\}.I)\tau\\
&- H\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1}(tr\{D^{t'}\bar{Z}H^{*}\bar{Z}^{'}D^{s'}\}.I)C^{*'}\Psi_i J_{t-i}^{'}\Omega HC^{*'}\Psi_j J_{s-j}^{'}\vartheta\\
&- H\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1}(tr\{\Omega J_{t-i}\Psi_i^{'}C^{*}HC^{*'}\Psi_j J_{s-j}^{'}\}.I)\\
&\qquad\qquad \times (tr\{\bar{Z}H^{*}\bar{Z}^{'}D^{t}D^{s'}\}.I)\vartheta\\
&- H\Lambda^{**'}\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1}\Omega J_{s-j}\Psi_j^{'}C^{*}HC^{*'}\Psi_i J_{t-i}^{'}(tr\{\bar{Z}H^{*}\bar{Z}^{'}D^{t}D^{s}\}.I)\tau\\
&+ o(T^{-1}).
\end{aligned}
$$

A proof of this result is given in Appendix B.1 by evaluating the expectation of each term. The two stage least squares result is given in Chapter 2.

3.4 Bias corrected 2SLS/FLIML Estimators

The procedure is the same as shown in section 2.4 for the bias corrected 2SLS estimator. Bias corrected 2SLS/FLIML estimators for structural form equations can be obtained by estimating the approximating bias and then subtracting this bias estimate from the corresponding estimators.

Definition 1. Given $\hat{\delta}_{1,b(2SLS/FLIML)}$ ***as estimated 2SLS/FLIML bias approximations replicates for the coefficient bias*** $\delta_{1,b(2SLS/FLIML)} = \big(\beta_{1,b(2SLS/FLIML)}, \alpha^{(1)}_{1,b(2SLS/FLIML)}, \ldots, \alpha^{(p)}_{1,b(2SLS/FLIML)},$ $c^{(0)}_{1,b(2SLS/FLIML)}, c^{(1)}_{1,b(2SLS/FLIML)}, \ldots, c^{(q)}_{1,b(2SLS/FLIML)}\big)$, ***and given*** $\hat{\delta}_{1,2SLS/FLIML}$ ***as the 2SLS/FLIML estimator of*** δ_1***, the C2SLS***

and CFLIML bias corrected estimator $\hat{\hat{\delta}}_{1,C2SLS/CFLIML}$ ***is as follows:***

$$\hat{\hat{\delta}}_{1,C2SLS/CFLIML} = \hat{\hat{\delta}}_{1,2SLS/FLIML} - \hat{\hat{\delta}}_{1,b(2SLS/FLIML)}. \tag{3.11}$$

To examine how well the C2SLS/CFLIML works for practical bias correction, a set of Monte Carlo experiments were conducted and the results are discussed in section 3.5.

3.5 Numerical Results

To be consistent, the numerical model and coefficient set up of this chapter is exactly the same as in Chapter 2, three dependent variables simultaneous equations models, with four lagged dependent variables. See the Chapter 2 numerical design numerical model part. The number of Monte Carlo replications is $20,000$.

The numerical results of 2SLS, FLIM, C2SLS, and CFLIML are summarized in Appendix. B.2 from Table B.1 to Table B.3. Table B.1 shows the bias approximation in both cases. Table B.2 presents the simulated estimation bias of 2SLS, FLIML, C2SLS, and CFLIML respectively. Table B.3 presents the MSE of these estimators. β_{21} and β_{31} are the coefficients of endogenous variables of the first structural form equation. α_{11}^{1} to α_{31}^{4} are the coefficients of the lagged endogenous variables. c_{11} is the constant term, and c_{21}, c_{31} are the parameters of exogenous variables.

In Table B.1 most of the bias approximations actually provide an overstated indication of the magnitude of the "true" bias as given by the Monte Carlo estimates in 2SLS and FLIML. The bias approximations of FLIML are generally smaller than those of 2SLS, and many of them have opposite signs. This latter characteristic arises because the bias that results from simultaneity and dynamics have opposite signs in

2SLS which was observed in Chapter 2, while FLIML removes the simultaneity bias and part of the dynamic bias.

Both Table B.1 and Table B.2 show that when the sample size increases, the bias of estimators decreases, and when the level of over-identification L increases, the bias of estimators also increases. Comparing the two uncorrected estimators, FLIML has superior performance to 2SLS in reducing the estimated bias for the endogenous and exogenous variable's coefficients. The bias that is left is relatively small and many estimated coefficients are almost unbiased. As was discussed in section 3.3, the bias approximation of FLIML concerns only the dynamic part (compared to 2SLS, FLIML eliminates the bias from simultaneity and some parts of the dynamics that are present in 2SLS). The estimated values of parameters based on FLIML are much less biased than the 2SLS in most cases, however, when sample size is 50, the bias of α_{21}^{1}, for the FLIML estimator is larger than for 2SLS. When the sample size is 50, with the over-identification level L =2, the bias of FLIML estimator for α_{31}^{4} is 10% more than 2SLS. The bias corrected estimators in the Monte Carlo simulations, i.e.C2SLS and CFLIML, are obtained by replacing the unknown coefficients in the bias approximation with estimates, which has been discussed in section 3.4.

Comparing the Monte Carlo bias estimates for C2SLS, CFLIML, with those for 2SLS, FLIML, it is clear that, generally, the bias corrected estimators have a substantially smaller bias than their uncorrected counterparts. In general, C2SLS and CFLIML yield almost unbiased estimators in both sample size 50 and 100, when over-identification level is $L = 2, 4$ and 6.

While the uncorrected FLIML estimates are much less biased than the uncorrected 2SLS, the mean squared errors of the parameter estimates of the endogenous and exogenous variables for the FLIML estimator are smaller in the simulations than those of 2SLS. However, the MSEs of dynamic coefficient estimators in FLIML are slightly larger than in 2SLS, though the percentage increase is not large. The MSEs of CFLIML and C2SLS exhibit similar characteristics. The MSE decreases when sample size increases; the MSE of the corrected estimators are smaller than the associated uncorrected estimators in general. In few cases, when sample size is large and the bias of uncorrected estimators is small, the MSE of corrected estimators is slightly larger than the associated uncorrected estimators. However, this increase is not likely to be substantial. For α_{11}^1, which is the coefficient of $L^1 y_1$ in the first equation of the structural form, when sample size is 100 and the over-identification level is $L = 2, 4$ and 6, the percentage of bias for 2SLS is +2% , −0%, +3% and the MSE is 0.0152, 0.0104, 0.0084, while for C2SLS, the MSE is 0.0160, 0.0105, 0.0097. The percentage of bias for FLIML is +1% , −1%, +1% and the MSE is 0.0165, 0.0118, 0.0102, while for the CFLIML, the MSE is 0.0168, 0.0121, 0.0101.

3.6 Conclusion

The $O(T^{-1})$ bias in 2SLS estimation of a general DSEM can be decomposed into two parts, which can be related to the simultaneity and dynamics respectively. However, the bias of FLIML effectively comes from the dynamics. It removes the simultaneity bias to order T^{-1} which is in 2SLS and some part of the dynamic bias. It gives less biased estimates compared to 2SLS. Notice that the bias approximation provides an overstated indication of magnitude of the "true" bias as given by the Monte Carlo estimates in 2SLS/FLIML. The mean squared errors of endogenous and exogenous variables' coefficients in FLIML are smaller than in 2SLS. However, the MSEs for some dynamic coefficient estimates are found to increase in FLIML. The bias corrected estimator, based upon the $O(T^{-1})$ approximation, very substantially reduces the 2SLS/FLIML bias. In addition, it was found to be better overall in terms of MSE, as there is no inflation of the MSE of uncorrected estimators. Hence, the bias corrected estimator, based upon $O(T^{-1})$ can be recommended as a bias reduction technique for either estimator from our analytic and numerical analysis.

Chapter 4

Bias Approximation and Reduction in the pth -order Dynamic Reduced Form

4.1 Introduction

Many papers have examined the properties of least squares estimators in dynamic regression models with white noise innovation disturbances. It is consistent, however, it exhibits serious bias when sample size is small. In this paper, we extend the analysis in Kiviet, Phillips, and Schipp (1999) to the general *pth*-order dynamic reduced form case. Two most popular approximation expansions are used to derive the properties of the ordinary least square estimators in dynamic regression models; Nagar's large-T approximation method and Kadane's small-σ approximation. The latter one was first employed by Kadane (1971) for k-class estimators to analyse the coefficients of a single equation of linear simultaneous stochastic equations with normal disturbances. This method uses a Taylor series expansion to expand the expression for the sample error, so that the successive terms are in descending order of σ in probability, in contrast to the large-T asymptotic expansion which orders these terms in increasing order of the sample size, T, in probability. The general expansion is

$$\frac{1}{\sigma}(\hat{\alpha}-\alpha)=\sum_{s=1}^{p}\sigma^{s-1}\dot{e_s}+\sigma^{p}\dot{r_p}$$

and where $\dot{e_s}$, for $s = 1,...,p$, and $\dot{r_p}$ are also bounded in probability, this time as σ, the standard deviation of the equation disturbance, tends to zero. The bias is then approximated to an appropriate order by taking expectations of the terms in the summation. Kiviet and Phillips (1993), Kiviet and Phillips (1994) applied the small-σ approximation in the context of ARX models.

The large-T approximation in OLS, was first used by Grubb and Symons (1987) for the OLS estimator of the lagged dependent variable parameter in a first order stable autoregressive model with exogenous variables (ARX(1)). Then Kiviet and Phillips (1993), Kiviet and Phillips (1994) provided a series of extensions of Grubb and Symons (1987). They extended the analysis to the estimator of the full coefficient vector and the high order dynamic regression model, ARX(p). All the papers above are in a stable model context. Kiviet and Phillips (2005) extended the Nagar approximation to examine the bias, variance and mean square error of the OLS estimator for the coefficient vector in a linear dynamic regression model with a unit root. Kiviet, Phillips, and Schipp (1999) also explored both small- σ and larger-T approximations for the OLS estimator in the context of a first order dynamic reduced form model with normally distributed white noise disturbances and an arbitrary number of exogenous regressors.

Then the bias approximation was used to construct the corrected OLS (COLS), respectively, unbiased to order σ^2 and order T^{-1}, Kiviet and Phillips (1993), Kiviet and Phillips (1994), Kiviet, Phillips, and Schipp (1999).

It has been shown that the two approaches give essentially the same bias approximations in the static case, and have almost the same effect in the bias correction procedures. However, as shown by Kiviet and Phillips (1989), Kiviet and Phillips (1994) and Kiviet, Phillips, and Schipp (1999), the two approximations are not the same in dynamic models, the results that come from small-σ approximation are proved rather poor and the performance of biased corrected estimator is poor as well. The large-T approach is superior, it leads to the bias corrected estimators continuing to perform satisfactorily and it would be much preferred.

The numerical method, the bootstrap, is a most popular tool in econometrics for improving estimation and inference. In the normal distributed errors ARX(1) model, Ip (1991) has proven that the residual bootstrap bias correction removes the $O(T^{-1})$ part of the bias. Surprisingly, Inoue and Kilian (2002) shows that even under the case where the process is integrated to order unity, as long as the number of lags $p > 1$, the residual bootstrapping in the high order AR(p) processes is valid. Hence, we are interested in exploring the performance of this residual bootstrap method in the *pth* lagged dependent variables reduced form model.

In this chapter, without losing generality, we focus on the bias in the OLS estimates for the first equation of a multi-equations system, where p lagged dependent variables are included. The bias approximation to order T^{-1} is derived by using the Nagar expansion, and the bias corrected estimators are constructed by employing the estimated bias approximation. We set up a series of Monte Carlo experiments to show the performance of COLS, and the residual bootstrap OLS in this general reduced form model. The simulations and numerical results suggest that the OLS bias can be substantial which was also observed in Kiviet, Phillips, and Schipp (1999). The COLS estimator gives almost unbiased estimation, and the residual bootstrap method is also well behaved in the bias reduction procedure. However comparing these two methods, the $O(T^{-1})$ expansion is more effective. Moreover, these two bias correction methods do not lead to an increase in the MSE.

The structure of this chapter is as follows. Section 4.2 introduces the general reduced form model, and section 4.3 presents the bias approximation to order T^{-1} . Section 4.4 constructs the COLS by introducing the bias approximation. The numerical design which includes the bootstrap methodology is presented in section 4.5. The numerical results and our conclusion are in section 4.6 and section 4.7 respectively.

4.2 The Model

The complete system is presented in Chapter 2.2, with innovation errors and p lagged-dependent and q lagged strongly exogenous explanatory variables. The exogenous variables are with $I(0)$ process. The first equation of the reduced form system in equation (2.2) is given by:

$$y_1 = \sum_{i=1}^{p} L^i Y \gamma_1^{(i)} + \sum_{j=0}^{q} L^j X \pi_1^{(j)} + \omega_1 v_1, \tag{4.1}$$

where v_1 is the first column of V, $v_1 \sim N(0,I)$. y_1 is a $T \times 1$ vector and the observations on the regressors are contained in a $T \times (P+Q)$ matrix $Z = [R : S]$. $P = \sum_{m=1}^{G} p(m)$ and $Q = \sum_{n=1}^{K} q(n)$. Here the $T \times P$ matrix R includes all the stochastic lagged dependent variables, and $T \times Q$ matrix S includes all the other regressors.

In what follows we shall rewrite equation (4.1) as:

$$y_1 = Z\alpha_1 + \omega_1 v_1, \tag{4.2}$$

where $\alpha_1' = (\gamma_1^{(i)'}, \pi_1^{(j)'}) = (\gamma_{1,}^{(1)} ..., \gamma_{1,}^{(p)}, \pi_1^{(0)}, ..., \pi_1^{(q)})$ are the reduced form parameters.

4.3 Reduced Form Estimation: OLS bias

The OLS estimator of α_1 in equation (4.2) is given by:

$$\hat{\alpha_1} = (Z^{'}Z)^{-1}Z^{'}y_1 = \alpha_1 + (Z^{'}Z)^{-1}Z^{'}\omega_1 v_1 \tag{4.3}$$
$$= \alpha_1 + (Z^{'}Z)^{-1}Z^{'}\tilde{v_1}.$$

Corollary 1. X is $I(0)$ which implies that $(i)X^{'}X = O(T)$; $(ii)\bar{Z}^{'}\bar{Z} = O(T)$; $(iii)Z^{'}Z = O_p(T)$; $(iv)E(\bar{Z}^{'}\tilde{v_1}) = 0$ and $\bar{Z}^{'}\tilde{v_1} = O_p(T^{1/2})$; $(v)E(Z^{'}\tilde{v_1}) = 0$ and $Z^{'}\tilde{v_1} = O_p(T^{1/2})$.

To find a bias approximation to order T^{-1}, upon substituting equation (2.4), we start from:

$$(Z^{'}Z)^{-1} = \left[\mathbb{E}(Z^{'}Z) + Z^{'}Z - \mathbb{E}(Z^{'}Z)\right]^{-1} \tag{4.4}$$
$$= [H^{*-1} + \bar{Z}'\tilde{W^*} + \tilde{W^*}'\bar{Z} + \tilde{W^*}'\tilde{W^*} - \mathbb{E}(\tilde{W^*}'\tilde{W^*})]^{-1}$$
$$= H^*\left[I + \left(\bar{Z}'\tilde{W^*} + \tilde{W^*}'\bar{Z}\right)H^* + \left(\tilde{W^*}'\tilde{W^*} - \mathbb{E}(\tilde{W^*}'\tilde{W^*})\right)H^*\right]^{-1},$$

where

$$H^{*-1} = \mathbb{E}(Z^{'}Z) = \bar{Z}^{'}\bar{Z} + \mathbb{E}(\tilde{W^*}'\tilde{W^*}).$$

The stochastic term inside the inverse term of equation (4.4) are of stochastic order of $T^{-1/2}$, then

$$Z'\tilde{v_1} = \bar{Z}^{'}\tilde{v_1} + \tilde{W^*}'\tilde{v_1} \tag{4.5}$$

and each term on the right hand side of equation (4.5) is $O_p(T^{1/2})$. Expanding equation (4.4) and keeping terms up to $O_p(T^{-1})$, then substituting the expanding form and equation (4.5) into equation (4.3), yields the following expansion of the estimator of

α_1.

$$\begin{aligned}\hat{\alpha_1} - \alpha_1 &= H^* \left[I + \left(\bar{Z}' \tilde{W^*} + \tilde{W^*}' \bar{Z} \right) H^* + \left(\tilde{W^*}' \tilde{W^*} - \mathbb{E}(\tilde{W^*}' \tilde{W^*}) \right) H^* \right]^{-1} \\ &\quad \times \left(\bar{Z}' \tilde{v_1} + \tilde{W^*}' \tilde{v_1} \right) \\ &= H^* \bar{Z}' \tilde{v}_1 + H^* \tilde{W}^{*'} \tilde{v}_1 - H^* \bar{Z}' \tilde{W}^* H^* \bar{Z}' \tilde{v}_1 \\ &\quad - H^* \tilde{W}^{*'} \bar{Z} H^* \bar{Z}' \tilde{v}_1 - H^* \bar{Z}' \tilde{W}^* H^* \tilde{W}^{*'} \tilde{v}_1 \\ &\quad - H^* \tilde{W}^{*'} \bar{Z} H^* \tilde{W}^{*'} \tilde{v}_1 - H^* \left(\tilde{W^*}' \tilde{W^*} - \mathbb{E}(\tilde{W^*}' \tilde{W^*}) \right) H^* \bar{Z}' \tilde{v}_1 \\ &\quad - H^* \left(\tilde{W^*}' \tilde{W^*} - \mathbb{E}(\tilde{W^*}' \tilde{W^*}) \right) H^* \tilde{W}^{*'} \tilde{v}_1 + o_p(T^{-1}).\end{aligned} \tag{4.6}$$

Taking expectations of each term above leads to the following:

Theorem 3. The bias of the OLS estimator of the first reduced form equation to order T^{-1} is given by

$$\mathbb{E}(\hat{\alpha}_1 - \alpha_1) = -H^* \Xi^* \tilde{\Omega}_{.1} - H^* \begin{bmatrix} \Theta \\ 0 \end{bmatrix} \tilde{\Omega}_{.1} + o(T^{-1}), \tag{4.7}$$

where

$$\begin{aligned} H^* &= \left\{ \bar{Z}' \bar{Z} + \sum_{i=1}^{p} \sum_{j=1}^{p} \sum_{t=i,j}^{T-1} (T-t) \begin{bmatrix} \Psi_i J_{t-i}' \tilde{\Omega} J_{t-i} \Psi_i' & 0 \\ 0 & 0 \end{bmatrix} \right\}^{-1}, \\ \Xi^* &= \sum_{i=1}^{p} \sum_{t=i}^{T-1} \Lambda_t^* \left(J_{t-i} \Psi_i' : 0 \right)', \\ \Lambda_t^* &= \bar{Z}' D^t \bar{Z} H^* + tr \left(\bar{Z}' D^t \bar{Z} H^* \right) I_{P+Q}, \qquad \text{where} \quad P = \sum_{m=1}^{G} p(m), Q = \sum_{n=1}^{K} q(n) \end{aligned}$$

and Ψ is a $P \times P$ matrix which equals to:

$$\Theta = \sum_{i=1}^{p} \sum_{t=i}^{T-1} \sum_{l=1}^{p} \sum_{r=l}^{T-1} \sum_{b=i+l}^{p} \sum_{s=t+r}^{T-1} (T-s) \left[(\Psi_l J_{s-l}^{'} \tilde{\Omega} J_{r-b} \Psi_b^{'} H^{**} + tr \left\{ (\Psi_l J_{s-l}^{'} \tilde{\Omega} J_{r-b} \Psi_b^{'} H^{**} \right\} I \right]$$
$$\times \Psi_i J_{t-i}^{'},$$

where recall that in section 2.3, we defined $H^{**} = I_2^{'} H^* I_2$ as the $P \times P$ leading submatrix of matrix H^*, with $I_2 = \begin{bmatrix} I_P \\ 0 \end{bmatrix}$ which is a $(P+Q) \times P$ selection matrix, and $\tilde{\Omega}_{.1}$ is the first column of

$$\tilde{\Omega} = \mathbb{E}\left[\frac{1}{T}\tilde{V}'\tilde{V}\right].$$

A proof of this result is given in Appendix C.1 by evaluating the expectations of each terms.

Then, the bias approximation can be evaluated straightforwardly once the structure is known.

4.4 Bias corrected OLS Estimator in Reduced Form

Bias corrected OLS is similar to the C2SLS, and CFLIML in the last two chapters, via, employing the simulated bias approximation into the OLS estimator.

Definition 1. Given $\hat{\alpha}_{1,b(OLS)}$ ***as estimated OLS bias approximations replicates for the coefficient bias*** $\alpha_{1,b(OLS)} = \left(\gamma_{1,b(OLS)}^{(1)}, ..., \gamma_{1,b(OLS)}^{(p)}, \pi_{1,b(OLS)}^{(0)}, \pi_{1,b(OLS)}^{(1)}, ..., \pi_{1,b(OLS)}^{(q)}\right)$, ***and given*** $\hat{\alpha}_{1,OLS}$ ***as the OLS estimator of*** α_1***, the COLS bias corrected estimator*** $\hat{\alpha}_{1,COLS}$ ***is as follows:***

$$\hat{\alpha}_{1,COLS} = \hat{\alpha}_{1,OLS} - \hat{\alpha}_{1,b(OLS)}. \tag{4.8}$$

To examine how well the COLS works for practical bias correction, a set of Monte Carlo experiments were conducted and the results are discussed in section 4.6.

4.5 Numerical Experiments Design

4.5.1 Numerical Model

The system was still constructed as in Chapter 2.5 and Chapter 3.2 using a three equations dynamic simultaneous equations model with four lagged endogenous variables based on sample sizes 50 and 100 and it's over-identified. Hence the first equation of reduced form is the following:

$$y_1 = \sum_{i=1}^{4} L^4 Y \gamma_1^{(4)} + \sum_{j=0}^{q} L^j X \pi_1^{(j)} + \omega v_1, \tag{4.9}$$

where $Y = (y_1, y_2, y_3)$. We analyse the reduced form properties of the general dynamic simultaneous equations models and in this chapter we only focus on the case when $L = 2$, $X = (x_1, x_2, x_3, x_4, x_5, x_6)$, since the over-identification level will not influence the properties of the reduced form. Each exogenous variable is generated as Gaussian autoregressive process with mean zero and with an autoregressive coefficient of 0.9, and they are independent of each other.

$$x_{jt} = 0.9 x_{j(t-1)} + \varsigma_{jt}, \qquad \varsigma_{jt} \sim \mathcal{N}(0, 1).$$

The coefficient matrices of reduced form based on the structural form coefficients set up in Chapter 2.5 are as follows:

$$\Gamma^{(1)} = -A^{(1)}B^{-1} = \begin{bmatrix} \gamma_{11}^{(1)} & -0.2093 & 0.0192 \\ \gamma_{21}^{(1)} & 0.0271 & -0.0777 \\ \gamma_{31}^{(1)} & -0.0141 & -0.0979 \end{bmatrix},$$

$$\Gamma^{(2)} = -A^{(2)}B^{-1} = \begin{bmatrix} \gamma_{11}^{(2)} & -0.2374 & -0.1360 \\ \gamma_{21}^{(2)} & -0.2951 & 0.0306 \\ \gamma_{31}^{(2)} & 0.2142 & -0.0251 \end{bmatrix},$$

$$\Gamma^{(3)} = -A^{(3)}B^{-1} = \begin{bmatrix} \gamma_{11}^{(3)} & -0.1637 & -0.0725 \\ \gamma_{21}^{(3)} & -0.1253 & -0.1402 \\ \gamma_{31}^{(3)} & -0.0581 & 0.0154 \end{bmatrix},$$

$$\Gamma^{(4)} = -A^{(4)}B^{-1} = \begin{bmatrix} \gamma_{11}^{(4)} & -0.0808 & -0.0603 \\ \gamma_{21}^{(4)} & -0.0061 & -0.0826 \\ \gamma_{31}^{(4)} & 0.0593 & 0.0479 \end{bmatrix},$$

and for,

$$L = 2, \quad \Pi^{'} = \begin{bmatrix} \pi_{11} & \pi_{21} & \pi_{31} & \pi_{41} & \pi_{51} & \pi_{61} & \pi_{71} \\ -0.1949 & -0.1191 & 0.0992 & -0.0830 & 0.0266 & -0.0161 & 0.0921 \\ -0.1382 & -0.0384 & 0.0320 & 0.0628 & -0.0201 & 0.0014 & -0.0083 \end{bmatrix}.$$

The coefficients we are interested in from the first equation are:

$$\gamma_{11}^{(1)} = 0.1774, \quad \gamma_{21}^{(1)} = 0.0258, \quad \gamma_{31}^{(1)} = -0.1177, \quad \gamma_{11}^{(2)} = 0.0454, \quad \gamma_{21}^{(2)} = 0.1626,$$
$$\gamma_{31}^{(2)} = -0.0768, \quad \gamma_{11}^{(3)} = -0.0397, \quad \gamma_{21}^{(3)} = 0.2487, \quad \gamma_{31}^{(3)} = 0.3409, \quad \gamma_{11}^{(4)} = 0.0371,$$
$$\gamma_{21}^{(4)} = 0.1751, \quad \gamma_{31}^{(4)} = 0.1584.$$

When $L = 2$,

$$\pi_{11} = -0.0806, \quad \pi_{21} = 0.1697, \quad \pi_{31} = -0.1414 \quad \pi_{41} = 0.1482, \quad \pi_{51} = -0.0474,$$
$$\pi_{61} = -0.0249, \quad \pi_{71} = 0.142.$$

We use a matrix P from a Choleski factorosation of the reduced form covariance Ω to generate the reduced form errors. Hence,

$$\begin{bmatrix} \tilde{v}_{1,t} \\ \tilde{v}_{2,t} \\ \tilde{v}_{3,t} \end{bmatrix} = P \begin{bmatrix} \tilde{e}_{1,t} \\ \tilde{e}_{2,t} \\ \tilde{e}_{3,t} \end{bmatrix}$$

where $\tilde{e}_{1,t}$, $\tilde{e}_{2,t}$ and $\tilde{e}_{3,t}$ denote the standardised disturbances. The distribution of the rows of $\tilde{U}$ always have mean 0 and covariance matrix Σ, and they are i.i.d. $\tilde{U}$ is the structural form disturbances. Then, the distribution of the structural disturbances can be evaluated from

$$B^{'}\tilde{v}_t = \tilde{u}_t \Rightarrow \tilde{u}_t \sim \mathcal{N}(0, \Sigma), \qquad \textit{where} \quad \Sigma = B^{'}\Omega B.$$

We arbitrarily set the structural covariance matrix is as follows:

$$\Sigma = \begin{bmatrix} 0.3524 & 0.3448 & 0.3112 \\ 0.3448 & 0.3668 & 0.2984 \\ 0.3112 & 0.2984 & 0.4064 \end{bmatrix} \tag{4.10}$$

from which the reduced form covariance is:

$$\Omega = \begin{bmatrix} 0.0055 & 0.0054 & 0.0030 \\ 0.0054 & 0.0844 & 0.0085 \\ 0.0030 & 0.0085 & 0.0069 \end{bmatrix}. \tag{4.11}$$

The setting up of initial values is exactly the same as in our Chapter 2 and 3.

4.5.2 The Simulation Model

The number of Monte Carlo replications is $20,000$, and 199 bootstrap replicates are used when achieving the bias corrected bootstrap.

Bootstrap

Based on Freedman (1984), Ip (1991) provides support for the asymptotic validity of the 2SLS bootstrap in static and dynamic models where errors are normal, and MacKinnon (2002) conducted hypothesis testing in a static model which also supports the asymptotic validity of the 2SLS bootstrap.

With the standard residual bootstrap, the bias corrected bootstrap estimators are calculated by first estimating the equation of interest using the original estimation method, and then by using this to generate pseudo-date (B sets) by resampling the residual from the initial estimated equation. Bootstrap replicates are obtained by implementing the original estimation method on each of B sets. The bias corrected bootstrap estimate of α_1 can be calculated as $2\hat{\alpha}_1 - \hat{\alpha}_{1,\bar{b}}$, where $\hat{\alpha_1}$ is the original estimate, and $\hat{\alpha}_{1,\bar{b}}$ is the mean of the bootstrap replicates.

Freedman's bootstrap contains the same steps as the usual residual bootstrap , except for the generation of the pseudo data.

The target equation is :

$$y_1 = \sum_{i=1}^{4} L^i Y \gamma_1^{(i)} + \sum_{j=0}^{j} L^j X \pi_1^{(j)} + \tilde{v}_1. \tag{4.12}$$

To set up bootstrap procedure for the above reduced form, we generate the $y_1 = (y_{11}, y_{12}...y_{1T})'$, and the initial value y_0' , Ly_0', L^2y_0', L^3y_0' which are obtained in Chapter 2, then by using OLS estimator, obtain $\hat{\alpha}_1$ and $\hat{v}_1 = (\hat{v}_1, ..., \hat{v}_T)'$. Then the re-sampling data $y_1^* = (y_{11}^*, y_{12}^*...y_{1T}^*)'$ and Y_0^* , LY_0^*, $L^2Y_0^*$, $L^3Y_0^*$ are generated recursively as $y_1^* = \sum_{i=1}^{4} L^i Y^* \hat{\gamma}_1^{(i)} + \sum_{j=0}^{j} L^j X \hat{\pi}_1^{(j)} + \hat{\tilde{v}}_1^*$, where we resample the $\hat{\tilde{v}}_1$ in equation (4.12) to generate $\hat{\tilde{v}}_1^*$.

In each bootstrap replication, y_1^* is regressed on $[LY^* : L^2Y^* : L^3Y^* : L^4Y^* : X]$ to get the bootstrapped estimates $\hat{\alpha}_1^*$, then the bias corrected bootstrap estimates are given by the definition 2:

Definition 2. Given $\hat{\alpha}_{1,\bar{b}}$ as the mean of the bootstrap OLS replicates for the coefficient $\alpha_1 \in (\gamma^{(1)}, \gamma^{(2)}, \gamma^{(3)}, \gamma^{(4)}, \pi_1)$, and given $\hat{\alpha}_{1,OLS}$ as the OLS estimator of α_1, the bootstrap bias corrected estimator $\hat{\alpha}_{1,b}$ is as follows:

$$\hat{\alpha}_{1,b} = 2\hat{\alpha}_{1,OLS} - \hat{\alpha}_{1,\bar{b}}, \quad where \quad \hat{\alpha}_{1,\bar{b}} = \frac{1}{B}\sum_{b=1}^{B} \hat{\alpha}_{1,b}.$$

4.6 Numerical Results

The numerical results of OLS, COLS, and residual bootstrap OLS are summarized in Appendix. C.2 from Table C.1 to Table C.2. Table C.1 shows the bias approximation, the OLS bias, the COLS bias, and the bootstrap bias when the sample size is 50 and 100 respectively. Table C.2 presents the MSE of these estimators. γ_{11}^1 to γ_{31}^4 are the coefficients of lagged endogenous variables of the first reduced form equation. π_{11} is the constant team, and π_{21} to c_{71} are the parameters of exogenous variables in the reduced form.

In Table C.1 most of the bias approximations actually provide an overstated indication of the magnitude of the "true" bias as given by the Monte Carlo estimates in OLS, but this over-stating is not likely to be substantial. When the sample size increases, the bias approximation and OLS bias drop sharply. Both corrected OLS, and residual bootstrap OLS effectively reduce the bias in the relevant uncorrected method, especially for the lagged dependent variables. For instance, the first order dependent variable, $\gamma_{21}^{1} = 0.0258$, while the OLS bias and bias approximation is $-0.0375(-145\%)$, and $-0.0387(150\%)$ respectively. When we apply the bias corrected method on the OLS, the bias drop sharply; in the COLS case, the bias reduces from 145% to 16%, and in the bootstrap case, the the bias decreases from 145% to 47%. In Table C.1, the COLS presents less biased estimates compared with the bootstrap method; however, this advantage is not that obvious in some estimated coefficients. For example, for γ_{21}^{3}, these two correction methods reduce the bias from -4% to -2%. OLS itself gives an almost unbiased estimates for the exogenous variables, hence in this case, the bias correction is not that necessary to employ into the correction. For example, $\pi_{51} = -0.0474$, the OLS bias is $0.0005(+1\%)$ which gives an almost unbiased estimate; certainly the bias approximation in this case is also close to zero $(0.0009(+2\%))$. When sample size increases to 100, many coefficients of lagged endogenous variables actually present almost unbiased properties, so in these cases, the bias advantage of the correction method is not so obvious; however, nearly half of the estimated coefficients are far away from the actual values. The properties of the constant term among these three estimators are similar to those for the lagged endogenous variables.

Table C.2 presents the mean squared errors for these three estimators where it is seen that when sample size increases, MSE decreases. Surprisingly, the results are unlike the results in Chapter 2 and Chapter 3, since the three estimators have almost the same level of MSEs. Meanwhile, the MSE in the case of COLS is slightly smaller than in the other two cases which is similar as in Chapter 2 and Chapter 3. As reported in Table C.1, OLS gives almost unbiased estimates for the exogenous variable coefficients, while the MSE of them are close to zero. However, it is not the case for the constant term, which is 0.0053 for OLS , 0.0051 for COLS, 0.0049 for bootstrap OLS when sample size is 50, and 0.0016 for OLS , 0.0015 for COLS, 0.0015 for bootstrap OLS when sample size is 100.

4.7 Conclusion

The $O(T^{-1})$ bias approximation in the OLS estimation of a *pth* order dynamic reduced form is presented in our analytic analysis part. The ordinary least squares bias can be substantial in dynamic reduced form equations based on both simulations and numerical results, which has also been observed in Kiviet, Phillips, and Schipp (1999).

Analytically, the bias corrected estimator, based upon an $O(T^{-1})$ approximation, very substantially reduces the OLS bias. The residual bootstrap procedure in OLS also effectively reduces the bias. However, from the results, it is obvious that using the $O(T^{-1})$ bias approximation is more effective in bias reduction compared to the bootstrap method. The MSEs, in these three cases, are almost at the same level; in other words, these two bias correction methods do not lead to an increasing of the MSEs. Hence, the bias corrected estimator, COLS, based upon the $O(T^{-1})$ bias correction can be recommended as a bias reduction technique for the p lagged dependent variable reduced form. Alternatively, the non-parametric bootstrap is also a way to reduce the bias and may be considered especially if the computer cost is of importance.

Chapter 5

Summary of the Conclusions

The asymptotic distributions of estimates and test statistics play more and more important roles in the development of econometric theory. However, knowledge of their finite sample properties is limited in many cases. Based on the asymptotic properties of estimators and test statistic, inference may not be reliable for small samples or even moderately large samples. Hence, it is worthwhile to explore the relevant properties when the sample size is small; furthermore, it is important to derive the analytical results at the most general level possible, which can help us understand the quality of inference in practice. This thesis analyses the limited information estimators in general dynamic simultaneous equation models which will extend our knowledge of the small sample properties of estimators in this area of econometrics and which will be of benefit to economists in estimating economic models under the linear DSEM when sample size is small.

In this thesis, a standard system is introduced which contains normally and independently distributed structural disturbances with mean vector 0 and fixed covariance matrix $\Sigma = (\sigma_{ij})$, and strictly exogenous $I(0)$ regressors . This general dynamic simultaneous equations model includes endogenous variables which are lagged p time periods, and exogenous variables which are lagged q time period. Based on this general

DSEM, we explore the properties of the two most popular estimators in the linear SEM, that is the 2SLS and FLIML estimators, in the small sample environment. In the *pth* order reduced form model which comes from the general DSEM, we analyse the behaviour of the classical estimator, the ordinary least squares estimator, when the sample size is small. The bias corrected estimators are constructed by subtracting the estimated bias approximations from the corresponding estimators. We also conducted Monte Carlo experiments to compare the performance of corrected and uncorrected estimators. The three equations model which was employed includes four lagged dependent variables, normally distributed innovation errors, and $I(0)$ strong regressors. The over-identification level of this model is set up as 2, 4, and 6, and the model is stable. An alternative estimation procedure, a numerical bias correction method known as the residual bootstrap, is also introduced and applied to 2SLS and FLIML estimators.

Chapter 2 examines the small sample properties of 2SLS in the general DSEM. We analytically derived the bias approximation of 2SLS to order T^{-1} by using the Nagar expansion method, and showed that the bias approximation has a simultaneity part and a dynamic part (Kiviet, Phillips, and Schipp (1999), Kiviet and Phillips (1993) and Phillips(2011)). Then, the bias corrected method is constructed by estimating the approximating bias and subtracting this bias estimate from the corresponding estimators. Theoretically, it could reduce bias to order T^{-1}.

Numerical results show that the bias that comes from the dynamic part has an opposite sign compared to the bias that comes from the simultaneity part which indicates that the bias correction method which effectively reduces the bias in the static case may not do the same in the dynamic case. The bias corrected estimator, based upon the $O(T^{-1})$ approximation, very substantially reduces the 2SLS bias. It does not inflate the MSE in the most cases. Hence, the bias corrected estimator, based upon the $O(T^{-1})$ bias approximation can be recommended as a bias reduction technique.

The bootstrap simulation results in this paper provide evidence for an alternative bias correction technique; it is shown to perform particularly well at bias correction. The bias correction is not as effective as with C2SLS, but the computing cost is less. The Bootstrap also reduced both standard error and MSE in 2SLS for both endogenous and exogenous variable coefficients when L is large.

A comparison between 2SLS and FLIML estimators is made in Chapter 3. We derive the bias approximation for the FLIML estimator to order T^{-1}, and separate it into two parts, one part representing the bias as in 2SLS, and the other part which we call it the extra term is an additional term compared to the 2SLS bias approximation. Comparing these two parts analytically, the results show that the FLIML estimator gives much less biased estimates than 2SLS ; hence this extra term has a sign opposite to the part which represents the 2SLS bias approximation. As a result the $O(T^{-1})$ simultaneity bias is removed completely and the dynamic bias partially. To remove the $O(T^{-1})$ dynamic part bias completely from the FLIML, the corrected FLIML, CFLIML, is conducted by subtracting this bias estimate from the corresponding estimators.

Numerical results show that the bias in FLIML is smller than the 2SLS bias. The mean squared errors of endogenous and exogenous variables' coefficients in FLIML are also smaller than in 2SLS. However, the MSEs for some dynamic coefficient estimates are found to increase in FLIML.

The bias corrected estimator, based upon the $O(T^{-1})$ approximation, very substantially reduces the 2SLS/FLIML bias. Moreover, it does not inflate the MSE. Hence, the bias corrected estimator, based upon the $O(T^{-1})$ can be recommended as a bias reduction technique for either estimator.

In Chapter 4, we move to the *pth* order reduced form which is transformed from the general DSEM. To estimate the single equation model, OLS is the classical estimator. Analytically, we present the bias approximation of OLS using a large-T approximation and our corrected OLS gives unbiased estimation to order T^{-1} by subtracting this bias estimate from the corresponding estimators. The ordinary least squares bias can be substantial in dynamic reduced form equations as indicated by both simulations and numerical results which has also been shown in Kiviet, Phillips, and Schipp (1999).

Numerically, the bias corrected estimator, based upon the $O(T^{-1})$ approximation, very substantially reduces the Monte Carlo OLS bias. From the results however, it is shown that OLS itself may give almost unbiased estimates of the coefficients of the exogenous regressors which implies that the bias correction is not necessary for these estimated coefficients. The residual bootstrap procedure in OLS also effectively reduces the bias. However, from the results it is obvious that employing the $O(T^{-1})$ bias approximation is more effective in bias reduction compared to the bootstrap method. Surprisingly, the MSEs in these three cases, are almost at the same level; in other words, these two bias correction methods do not lead to an increase of the MSE. Hence, the $O(T^{-1})$ bias corrected estimator, COLS, can be recommended as a bias reduction technique in the pth lagged dependent variables reduced form. Alternatively, the non-parametric bootstrap is also a way to reduce the bias if the computer cost is the consideration.

Notice that the bias approximation provides an overstated indication of magnitude of the "true" bias as given by the Monte Carlo estimates in all these three estimators. In these three chapters, we have not considered the moments existence problem in the dynamic models. We all know these three estimators do not have a moments problem in the static case from our discussion in each chapter and in this thesis we assume there is no moments problem in the dynamic case either.

In this thesis, we numerically and analytically derive some results for limited information estimators in the general DSEM. However, there are still many ways in which further research could extend our current findings. As shown above, exploring the conditions for the existence of moments in the FLIML would be particularly interesting; however investigating the properties of inference procedures based on these three estimators of the general DSEM would be an obvious next step.

Our model is based on normally distributed innovation errors. If this assumption is relaxed to include other distributions including asymmetry, how might the results compare with the current results? Is the bias correction method based on the bias approximation still reliable? What the effect would be if the instruments are weak instruments is also interesting to explore. The small sample properties of estimators in panel data models is the another interesting direction to explore. We are also interested in applying our results in applications of economic and financial models to test to what extent our methods can improve estimation in practice.

References

[1] T. W. Anderson, Naoto Kunitomo, and Kimio Morimune. “Comparing Single-Equation Estimators in a Simultaneous Equation System”. In: *Econometric Theory* 2.1 (1986), pp. 1–32.

[2] T. W. Anderson and Herman Rubin. “The Asymptotic Properties of Estimates of the Parameters of a Single Equation in a Complete System of Stochastic Equations”. In: *The Annals of Mathematical Statistics* 21.4 (1950), pp. 570–582.

[3] T. W. Anderson and Takamitsu Sawa. “Distributions of Estimates of Coefficients of a Single Equation in a Simultaneous System and Their Asymptotic Expansions”. In: *Econometrica* 41.4 (1973), pp. 683–714.

[4] R. L. Basmann. “A Note on the Exact Finite Sample Frequency Functions of Generalized Classical Linear Estimators in a Leading Three-Equation Case”. English. In: *Journal of the American Statistical Association* 58.301 (1963), pp. 161–171.

[5] Maurice J.G. Bun and Frank Windmeijer. “A comparison of bias approximations for the two-stage least squares (2SLS) estimator”. In: *Economics Letters* 113.1 (2011), pp. 76 –79.

[6] John Chao et al. *An Expository Note on the Existence of Moments of Fuller and HFUL Estimators*. Departmental Working Papers 201311. Rutgers University, Department of Economics, July 2013.

[7] John Chao et al. *Combining Two Consistent Estimators*. Departmental Working Papers 201310. Rutgers University, Department of Economics, July 2013.

[8] H. Cramér. *Mathematical Methods of Statistics*. Princeton landmarks in mathematics and physics. Princeton University Press, 1945.

[9] H. E. Daniels. “Saddlepoint Approximations in Statistics”. In: *The Annals of Mathematical Statistics* 25.4 (1954), pp. 631–650.

[10] H. E. Daniels. “The Approximate Distribution of Serial Correlation Coefficients”. In: *Biometrika* 43.1/2 (1956), pp. 169–185.

[11] J. Davidson. *Econometric Theory*. Wiley, 2000.

[12] P.J. Dhrymes. *Econometrics: statistical foundations and applications*. Springer study edition. Springer-Verlag, 1970.

[13] F.Y. Edgeworth. *The Asymmetrical Probability-curve*. Philosophical magazine. Philosophical magazine, 1896.

[14] B. Efron. “Bootstrap Methods: Another Look at the Jackknife”. In: *Ann. Statist.* 7.1 (Jan. 1979), pp. 1–26.

[15] R. A. Fisher. “On the probable error of a coefficient of correlation deduced from a small sample”. In: *Metron* 1 (1921), pp. 3–32.

[16] R. A. Fisher. “The Goodness of Fit of Regression Formulae, and the Distribution of Regression Coefficients”. In: *Journal of the Royal Statistical Society* 85.4 (1922), pp. 597–612.

[17] D. Freedman. “On Bootstrapping Two-Stage Least-Squares Estimates in Stationary Linear Models”. English. In: *The Annals of Statistics* 12.3 (1984), pp. 827–842.

[18] Wayne A. Fuller. “Some Properties of a Modification of the Limited Information Estimator”. In: *Econometrica* 45.4 (1977), pp. 939–953.

[19] David Grubb and James Symons. “Bias in Regressions With a Lagged Dependent Variable”. In: *Econometric Theory* 3.03 (1987), pp. 371–386.

[20] Trygve Haavelmo. “Methods of Measuring the Marginal Propensity to Consume”. In: *Journal of the American Statistical Association* 42.237 (1947), pp. 105–122.

[21] Trygve Haavelmo. “The Statistical Implications of a System of Simultaneous Equations”. In: *Econometrica* 11.1 (1943), pp. 1–12.

[22] Jinyong Hahn and Jerry Hausman. “Notes on bias in estimators for simultaneous equation models”. In: *Economics Letters* 75.2 (2002), pp. 237–241.

[23] Jinyong Hahn, Jerry Hausman, and Guido Kuersteiner. “Estimation with weak instruments: Accuracy of higher-order bias and MSE approximations”. In: *Econometrics Journal* 7.1 (2004), pp. 272–306.

[24] P. Hall. *The Bootstrap and Edgeworth Expansion*. Springer Series in Statistics. Springer New York, 1997.

[25] Hausman et al. *Instrumental Variable Estimation with Heteroskedasticity and Many Instruments*. Economics Working Paper Archive 566. The Johns Hopkins University,Department of Economics, Sept. 2009.

[26] R. Holmgren. *A First Course in Discrete Dynamical Systems*. Universitext. Springer New York, 2000.

[27] Asraul Hoque and Tom A. Peters. “Finite sample analysis of the ARMAX models.” In: *Sankhyā. Series B. Methodological* 48.1-2 (1986), pp. 266–283.

[28] L. Hurwicz. *Least Squares Bias in Time Series*. Statistical Inference in Dynamic Econimic Models. New York, Wiley, 1950.

[29] Emma M. Iglesias and Garry D. A. Phillips. “Almost Unbiased Estimation in Simultaneous Equation Models With Strong and/or Weak Instruments”. In: *Journal of Business & Economic Statistics* 30.4 (2012), pp. 505–520.

[30] Atsushi Inoue and Lutz Kilian. "Bootstrapping Autoregressive Processes with Possible Unit Roots". In: *Econometrica* 70.1 (2002), pp. 377–391.

[31] W.C. Ip. *Bootstrap Methods in Econometrics.* Ph.D. Thesis University of Leeds, 1991.

[32] Joseph B. Kadane. "Comparison of k-Class Estimators When the Disturbances Are Small". In: *Econometrica* 39.5 (1971), pp. 723–737.

[33] Joseph B. Kadane. "Testing Overidentifying Restrictions when the Disturbances are Small". In: *Journal of the American Statistical Association* 65.329 (1970), pp. 182–185.

[34] Joseph B Kadane. "Testing a Subset of the Overidentifying Restrictions". In: *Econometrica* 42.5 (1974), pp. 853–67.

[35] Jan F. Kiviet and Garry D. A. Phillips. "Almost Unbiased Estimation in Dynamic Simultaneous Equations Through a Small Disturbance Correction". In: *Proceedings if the Osaka Econometrics Conferences, Japan* (1995).

[36] Jan F. Kiviet and Garry D. A. Phillips. "Bias assessment and reduction in linear error-correction models". In: *Journal of Econometrics* 63.1 (1994), pp. 215–243.

[37] Jan F. Kiviet and Garry D. A. Phillips. "Comparison of Large T and Small Sigma Asymptotics in Static and Dynamic Simultaneous Equation Models". In: *Working paper* (1989).

[38] Jan F. Kiviet and Garry D. A. Phillips. "Moment approximation for least-squares estimators in dynamic regression models with a unit root;" in: *Econometrics Journal* 8.2 (July 2005), pp. 115–142.

[39] Jan F. Kiviet, Garry D. A. Phillips, and Bernhard Schipp. "Alternative bias approximations in first-order dynamic reduced form models". In: *Journal of Economic Dynamics and Control* 23.7 (1999), pp. 909–928.

[40] Jan F. Kiviet and Garry D.A. Phillips. "Alternative Bias Approximations in Regressions with a Lagged-Dependent Variable". In: *Econometric Theory* 9.01 (1993), pp. 62–80.

[41] E.E. Leamer. *Specification Searches: Ad Hoc Inference ith non-experimental Data.* New York, Wiley, 1978.

[42] Gareth Liu-Evans and Garry D. A. Phillips. "Bootstrap, Jackknife and COLS: Bias and Mean Squared Error in Estimation of Autoregressive Models". In: *Journal of Time Series Econometrics* 4.2 (2012), pp. 1–35.

[43] James G. MacKinnon. "Bootstrap inference in econometrics". In: *Canadian Journal of Economics* 35.4 (2002), pp. 615–645.

[44] Roberto S. Mariano. "The Existence of Moments of the Ordinary Least Squares and Two-Stage Least Squares Estimators". In: *Econometrica* 40.4 (1972), pp. 643–652.

[45] W. M. Mikhail. “The Bias of the Two-Stage Least Squares Estimator”. In: *Journal of the American Statistical Association* 67.339 (1972), pp. 625–627.

[46] A. L. Nagar. “The Bias and Moment Matrix of the General k-Class Estimators of the Parameters in Simultaneous Equations”. In: *Econometrica* 27.4 (1959), pp. 575–595.

[47] Thomas A. Peters. “The exact moments of ols in dynamic regression models with non-normal errors”. In: *Journal of Econometrics* 40.2 (1989), pp. 279–305.

[48] Garry D. A. Phillips. “An alternative approach to obtaining Nagar-type moment approximations in simultaneous equation models”. In: *Journal of Econometrics* 97.2 (2000), pp. 345–364.

[49] Garry D. A. Phillips and Gareth Liu-Evans. “Approximating and Reducing Bias in 2SLS Estimation of Dynamic Simultaneous Equation Models”. In: *Computational Statistics and Data Analysis* (2015).

[50] P. C. B. Phillips. “Finite Sample Theory and the Distributions of Alternative Estimators of the Marginal Propensity to Consume”. In: *The Review of Economic Studies* 47.1 (1980), pp. 183–224.

[51] P.C.B. Phillips. “Exact small sample theory in the simultaneous equations model”. In: *Handbook of Econometrics*. Ed. by Z. Griliches† and M. D. Intriligator. Vol. 1. Handbook of Econometrics. Elsevier, 1983. Chap. 8, pp. 449–516.

[52] David H. Richardson. “The Exact Distribution of a Structural Coefficient Estimator”. In: *Journal of the American Statistical Association* 63.324 (1968), pp. 1214–1226.

[53] Takamitsu Sawa Roberto S. Mariano. “The Exact Finite-Sample Distribution of the Limited-Information Maximum Likelihood Estimator in the Case of Two Included Endogenous Variables”. In: *Journal of the American Statistical Association* 67.337 (1972), pp. 159–163.

[54] J. D. Sargan. “Econometric Estimators and the Edgeworth Approximation”. In: *Econometrica* 44.3 (1976), pp. 421–448.

[55] J. D. Sargan. “Gram-Charlier Approximations Applied to t Ratios of k-Class Estimators”. In: *Econometrica* 43.2 (1975), pp. 327–346.

[56] J D Sargan. “The Validity of Nagar’s Expansion for the Moments of Econometric Estimators”. In: *Econometrica* 42.1 (1974), pp. 169–76.

[57] Takamitsu Sawa. *Almost unbiased estimator in simultaneous equations systems.* mit Bibliogr. Hoboken, NJ, 1973.

[58] Takamitsu Sawa. “Finite-Sample Properties of the k-Class Estimators”. In: *Econometrica* 40.4 (1972), pp. 653–680.

[59] Takamitsu Sawa. “The Exact Sampling Distribution of Ordinary Least Squares and Two-Stage Least Squares Estimators”. In: *Journal of the American Statistical Association* 64.327 (1969), pp. 923–937.

[60] A. Ullah. *Finite Sample Econometrics*. Advanced Texts in Econometrics. OUP Oxford, 2004.

[61] Aman Ullah and A L Nagar. “The Exact Mean of the Two-Stage Least Squares Estimator of the Structural Parameters in an Equation Having Three Endogenous Variables”. In: *Econometrica* 42.4 (1974), pp. 749–58.

[62] Dandan Wang and Garry D.A Phillips. “Bias Assessment and Reduction for the 2SLS Estimator in General Dynamic Simultaneous Equations Models”. In: *Working paper* (2015).

Appendix A

Appendix for Chapter 2

A.1 The Evaluation for Theorem 1

A.1.1 Lemmas

The following lemmas will be used in later evaluations for Theorem 1 in section 2.6.

Lemma 1: **The expectation of a product of three normal (means of zero) random variables is zero.** i.e

$$\mathbb{E}(\Xi A \Psi B \Phi) = 0$$

where Ξ, Ψ, and Φ are three normal (means of zero) random variables.

Lemma 2: $(Z'Z)^{-1} = [\mathbb{E}(Z'Z)]^{-1} + O_p(T^{\frac{-3}{2}})$,
where $Z = [R : S] = \left[LY, L^2Y...L^pY : X, LX, L^2X...L^qX\right]$.

Proof:

$$\begin{aligned}
Z'Z &= \mathbb{E}(Z'Z) + (Z'Z - \mathbb{E}(Z'Z)) \\
&= \mathbb{E}(Z'Z)\left[I + \left[\mathbb{E}(Z'Z)\right]^{-1}\left(Z'Z - \mathbb{E}(Z'Z)\right)\right]. \\
(Z'Z)^{-1} &= \left[I + \left[\mathbb{E}(Z'Z)\right]^{-1}(Z'Z - \mathbb{E}(Z'Z))\right]^{-1}\left(\mathbb{E}(Z'Z)\right)^{-1}
\end{aligned}$$

where,

$$\left[I + \left[\mathbb{E}(Z'Z)\right]^{-1}(Z'Z - \mathbb{E}(Z'Z))\right]^{-1} = \left[I - \left[\mathbb{E}(Z'Z)\right]^{-1}(Z'Z - \mathbb{E}(Z'Z))\right] + o_p(T^{-1/2}).$$

Hence,

$$\begin{aligned}
(Z'Z)^{-1} &= \left[I - \left[\mathbb{E}(Z'Z)\right]^{-1}(Z'Z - \mathbb{E}(Z'Z))\right]\left(\mathbb{E}(Z'Z)\right)^{-1} + o_p(T^{-3/2}) \\
&= \left[\mathbb{E}(Z'Z)\right]^{-1} - \left[\mathbb{E}(Z'Z)\right]^{-1}\left(Z'Z - \mathbb{E}(Z'Z)\right)\left[\mathbb{E}(Z'Z)\right]^{-1} + o_p(T^{-3/2}) \\
&= \left[\mathbb{E}(Z'Z)\right]^{-1} + O_p(T^{-3/2}).
\end{aligned}$$

Lemma 3: **Based on Nagar (1959)'s decomposition, the reduced form disturbances can be decomposed as $\tilde{V} = S^* + \tilde{u}_1\phi^{'}$, where $\tilde{u}_1$ and S^* are normally distributed but independent, $\phi\sigma^2 = \mathbb{E}\left(\frac{1}{T}\tilde{V}^{'}\tilde{u}_1\right)$.**

$$\begin{aligned}
&\mathbb{E}(S^*AS^{*'}) = tr(C_2^*A).I, \\
&\mathbb{E}(S^{*'}AS) = tr(A).IC_2^*, \\
&\mathbb{E}(S^*AS^*) = A^{'}C_2^*, \\
&\mathbb{E}(S^{*'}AS^{*'}) = C_2^*A,
\end{aligned}$$

where A is a corresponding and constant matrix, $C_2^* = \Omega - \sigma^2\phi\phi^{'}$, Ω is the covariance matrix of $\tilde{V}$.

Lemma 4: **Mikhail (1972) Suppose also that U, V, W and X are matrices, with the same number of rows, whose elements are normally distributed random variables with the properties that if ϕ_{ri} and Ψ_{sj} are elements of any of these matrices**

$$\mathbb{E}(\phi_{ri}\Psi_{sj}) = 0, \qquad r \neq s$$
$$= \omega_{\phi}\Psi_{ij}, \qquad r = s$$

and denote the matrix whose elements are $\omega_{\phi}\Psi_{ij}$ by $\Omega_{\phi\Psi}$ for ϕ, $\Psi = U$, V, W and X.

Suppose also that A, B and C are constant matrices of such dimensions that the various products considered below exist, then:

1. $\mathbb{E}(UAVBWCX) = A'\Omega_{uv}BC'\Omega_{wx} + B'\Omega_{vx}\,\mathrm{tr}(\Omega_{uw}CA') + C'\Omega_{wv}BA'\Omega_{ux}$,
2. $\mathbb{E}(U'AVBWCX) = \Omega_{uv}BC'\Omega_{wx}\,\mathrm{tr}(A) + \Omega_{uw}CA'B'\Omega_{vx} + \Omega_{ux}\,\mathrm{tr}(AB'\Omega_{vw}C)$,
3. $\mathbb{E}(UAV'BWCX) = BC'\Omega_{wx}tr(\Omega_{uv}A') + B'C'\Omega_{wu}A\Omega_{vx} + C'\Omega_{wv}A'\Omega_{ux}tr(B)$,
4. $\mathbb{E}(UAVBW'CX) = A'\Omega_{uv}B\Omega_{wx}tr(C) + CA'\Omega_{uw}B'\Omega_{vx} + C'A'\Omega_{ux}tr(B\Omega_{wv})$,
5. $\mathbb{E}(U'AV'BWCX) = \Omega_{uv}A'BC'\Omega_{wx} + \Omega_{uw}CBA\Omega_{vx} + \Omega_{ux}tr(A\Omega_{vw}C)tr(B)$,
6. $\mathbb{E}(U'AVBW'CX) = \Omega_{uv}B\Omega_{wx}tr(C)tr(A) + \Omega_{uw}B'\Omega_{vx}tr(AC') + \Omega_{ux}tr(AC)tr(B'\Omega_{vw})$,
7. $\mathbb{E}(UAV'BW'CX) = B\Omega_{wx}tr(A'\Omega_{uv})trC + CB\Omega_{wu}A\Omega_{vx} + C'B\Omega_{wv}A'\Omega_{ux}$,
8. $\mathbb{E}(U'AV'BW'CX) = \Omega_{uv}A'B\Omega_{wx}tr(C) + \Omega_{uw}B'C'A\Omega_{vx} + \Omega_{ux}tr(A\Omega_{vw}B'C)$.

A.1.2 Evaluating the Expectations

From equation 2.14

$$\mathbb{E}(\hat{\delta}_1 - \delta_1) = \mathbb{E}\left\{H\bar{\Upsilon}\tilde{u}_1 + H\Delta_1^{'}\tilde{u}_1 + H\Delta_2^{'}\tilde{u}_1 - HJ_1^{*}H\bar{\Upsilon}^{'}\tilde{u}_1 - HJ_1^{*}H\Delta_2\tilde{u}_1\right\} + o(T^{-1}), \tag{A.1}$$

evaluating the expectation for each term.

The first term,

$$(i) \qquad \mathbb{E}\{H\bar{\Upsilon}\tilde{u}_1\} = H\bar{\Upsilon}\mathbb{E}\{\tilde{u}_1\} = 0. \tag{A.2}$$

The second term,

$$(ii) \qquad \mathbb{E}\{H\Delta_1^{'}\tilde{u}_1\} = H\mathbb{E}\{\Delta_1^{'}\tilde{u}_1\}.$$

Recalling equation 2.10 for the definition of Δ_1, we have:

$$\begin{aligned}
H\Delta_1^{'}\tilde{u}_1 &= H\left(\bar{Z}(Z^{'}Z)^{-1}Z^{'}[\tilde{V}_2 : 0 : 0] + \left[\tilde{R}I_2^{'}(Z^{'}Z)^{-1}Z^{'}\tilde{V}_2 : 0 : 0)\right]\right)^{'}\tilde{u}_1 \\
&= H\Lambda^{**'}\tilde{V}^{'}\bar{Z}(\mathbb{E}(Z^{'}Z))^{-1}\bar{Z}^{'}\tilde{u}_1 + H\Lambda^{**'}\tilde{V}^{'}\bar{Z}(\mathbb{E}(Z^{'}Z))^{-1}I_2\tilde{R}^{'}\tilde{u}_1 \\
&\quad + H\Lambda^{**'}\bar{V}^{'}\tilde{W}^{*'}(\mathbb{E}(Z^{'}Z))^{-1}\bar{Z}^{'}\tilde{u}_1 + H\Lambda^{**'}\tilde{V}^{'}\tilde{W}^{*}(\mathbb{E}(Z^{'}Z))^{-1}I_2\tilde{R}'u_1 + o_p(T^{-1})
\end{aligned}$$

where, $\tilde{V}_2$ is the $T \times g$ submatrix of matrix $\tilde{V}$, which can be expressed as $[\tilde{V}_2 : 0 : 0] = \tilde{V}\begin{bmatrix} I_g : 0 \\ 0 \end{bmatrix}$, and we define $\Lambda^{**} = \begin{bmatrix} I_g : 0 \\ 0 \end{bmatrix}$ which is with $G \times (g + P^* + Q^*)$ dimension selection matrix. Also by using Lemma 2,

$$\begin{bmatrix} \hat{\Gamma}_2^{(1)} - \Gamma_2^{(1)} \\ \vdots \\ \hat{\Gamma}_2^{(p)} - \Gamma_2^{(p)} \\ \hat{\Pi}_2^{(1)} - \Pi_2^{(1)} \\ \vdots \\ \hat{\Pi}_2^{(q)} - \Pi_2^{(q)} \end{bmatrix} = (Z^{'}Z)^{-1}Z^{'}\tilde{V}_2 = (\mathbb{E}\{Z^{'}Z\})^{-1}\bar{Z}^{'}\tilde{V}_2 + (\mathbb{E}\{Z^{'}Z\})^{-1}\tilde{W}^{*'}\tilde{V}_2 + o_p(T^{-1/2})$$

and

$$\begin{bmatrix} \hat{\Gamma}_2^{(1)} - \Gamma_2^{(1)} \\ \vdots \\ \hat{\Gamma}_2^{(p)} - \Gamma_2^{(p)} \end{bmatrix} = I_2^{'}(Z^{'}Z)^{-1}Z^{'}\tilde{V}_2 = I_2'(\mathbb{E}\{Z^{'}Z\})^{-1}\bar{Z}^{'}\tilde{V}_2 + I_2'(\mathbb{E}\{Z^{'}Z\})^{-1}\tilde{W}^{*'}\tilde{V}_2 + o_p(T^{-1/2}),$$

and

$$[\tilde{V}_2 : 0 : 0] = \tilde{V}\Lambda^{**}.$$

Taking expectation, the last two terms are zero, then this gives

$$\begin{aligned} \mathbb{E}\{H\Delta_1^{'}\tilde{u}_1\} = H\Lambda^{**'}\mathbb{E}\left\{\tilde{V}^{'}\bar{Z}(\mathbb{E}(Z^{'}Z))^{-1}\bar{Z}^{'}\tilde{u}_1\right\} \\ + H\Lambda^{**'}\mathbb{E}\left\{\tilde{V}^{'}\tilde{W}^{*}(\mathbb{E}(Z^{'}Z))^{-1}I_2\tilde{R}'u_1\right\} + o(T^{-1}). \end{aligned} \tag{A.3}$$

The first term can be expressed as:

(1)

$$\begin{aligned} H\Lambda^{**'}\mathbb{E}\left\{\tilde{V}^{'}\bar{Z}(\mathbb{E}(Z^{'}Z))^{-1}\bar{Z}^{'}\tilde{u}_1\right\} &= H\Lambda^{**'}\mathbb{E}\left\{(S^{*} + \tilde{u}_1\phi^{'})\bar{Z}(\mathbb{E}(Z^{'}Z))^{-1}\bar{Z}^{'}\tilde{u}_1\right\} \\ &= H(tr\{\bar{Z}(\mathbb{E}(Z^{'}Z))^{-1}\bar{Z}^{'}\Lambda^{**'}\}.I)(\sigma_1^2\phi). \end{aligned}$$

The second term can be evaluated as:

(2)

$$H\Lambda^{**'}\mathbb{E}\left\{\tilde{V}^{'}\tilde{W}^{*}(\mathbb{E}(Z^{'}Z))^{-1}I_2\tilde{R}'u_1\right\}$$

Note:

1. Recalling equation 2.5 in Chapter 2, $\tilde{W}^* = \left[\sum_{i=1}^{p}\sum_{t=i}^{T-1} D^t\tilde{V}J_{t-i}\Psi_i^{'} : 0\right]$, and $\tilde{R} = \sum_{i=1}^{p}\sum_{t=i}^{T-1} D^t\tilde{V}J_{t-i}\Psi_i^{'}$.

$$
\begin{aligned}
&H\Lambda^{**'}\mathbb{E}\left\{\tilde{V}^{'}\tilde{W}^*(\mathbb{E}(Z^{'}Z))^{-1}I_2\tilde{R}'u_1\right\}\\
&= H\Lambda^{**'}\mathbb{E}\left\{(S^*+\tilde{u}_1\phi')'\left[\sum_{i=1}^{p}\sum_{t=i}^{T-1} D^t\tilde{V}J_{t-i}\Psi_i^{'} : 0\right][\mathbb{E}(Z'Z)]^{-1}I_2\sum_{j=1}^{p}\sum_{s=j}^{T-1}\Psi_jJ_{s-j}'\tilde{V}'D^{s'}\tilde{u}_1\right\}\\
&= H\Lambda^{**'}\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1}\mathbb{E}\left\{\phi\tilde{u}_1'D^t\tilde{u}_1\phi'J_{t-i}\Psi_i'I_2'[\mathbb{E}(Z'Z)]^{-1}I_2\Psi_jJ_{s-j}'\phi\tilde{u}_1'D^{s'}\tilde{u}_1\right\}\\
&+H\Lambda^{**'}\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1}\mathbb{E}\left\{\phi\tilde{u}_1'D^tS^*J_{t-i}\Psi_i'I_2'[E(Z'Z)]^{-1}I_2\Psi_jJ_{s-j}'S^{*'}D^{s'}\tilde{u}_1\right\}\\
&+H\Lambda^{**'}\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1}\mathbb{E}\left\{S^{*'}D^tS^*J_{t-i}\Psi_i'I_2'[\mathbb{E}(Z'Z)]^{-1}I_2\Psi_jJ_{s-j}'\phi\tilde{u}_1'D^{s'}\tilde{u}_1\right\}\\
&+H\Lambda^{**'}\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1}\mathbb{E}\left\{S^{*'}D^t\tilde{u}_1\phi'J_{t-i}\Psi_i'I_2'[\mathbb{E}(Z'Z)]^{-1}I_2\Psi_jJ_{s-j}'S^{*'}D^{s'}\tilde{u}_1\right\}\\
&= (T-t)H\Lambda^{**'}\phi\sigma^2tr\left\{\Omega J_{t-i}\Psi_i'I_2'[\mathbb{E}(Z'Z)]^{-1}I_2\Psi_jJ_{s-j}'\right\}
\end{aligned}
$$

Note:

1. The last step is obtained by using Lemma 4.

Combining these two terms together, the result for equation (A.3) is:

$$
\begin{aligned}
&\mathbb{E}\{H\Delta_1^{'}\tilde{u}_1\} \\
&= H(tr\{\bar{Z}(\mathbb{E}(Z^{'}Z))^{-1}\bar{Z}^{'}\Lambda^{**'}\}.I)(\sigma_1^2\phi)\\
&\quad +(T-t)H\Lambda^{**'}\phi\sigma^2tr\left\{\Omega J_{t-i}\Psi_i'I_2'[\mathbb{E}(Z'Z)]^{-1}I_2\Psi_jJ_{s-j}'\right\}+o(T^{-1}).
\end{aligned}
\tag{A.4}
$$

The third term in equation A.1 is:

$$
(iii) \qquad \mathbb{E}\{H\Delta_2^{'}\tilde{u}_1\} = H\mathbb{E}(\Delta_2^{'}\tilde{u}_1) = 0. \tag{A.5}
$$

Recalling equation 2.10 for the definition of Δ_2, then clearly $\mathbb{E}\{\tilde{R}'\tilde{u}_1\} = 0$.

The fourth term of equation A.1 is:

$$
\begin{aligned}
(iv) \qquad -\mathbb{E}\{HJ_1^*H\bar{\Upsilon}'\tilde{u}_1\} &= -\mathbb{E}\{H[(\Delta_2'\Delta_2 - \mathbb{E}(\Delta_2'\Delta_2)) + ((\bar{\Upsilon}'\Delta_1 + \Delta_1'\bar{\Upsilon}) + (\bar{\Upsilon}'\Delta_2 \\
&\quad + \Delta_2'\bar{\Upsilon}) + (\Delta_1'\Delta_2 + \Delta_2'\Delta_1)]H\bar{\Upsilon}'\tilde{u}_1\} \\
&= -\mathbb{E}\left\{H\bar{\Upsilon}'\left[\sum_{i=1}^{p} L^i\bar{Y}(\hat{\Gamma}_2^{(i)} - \Gamma_2^{(i)}) + \sum_{j=1}^{q} L^j X(\hat{\Pi}_2^{(i)} - \Pi_2^{(i)})\right.\right. \\
&\quad \left.\left. + \sum_{i=1}^{p} L^i\tilde{W}(\hat{\Gamma}_2^{(i)} - \Gamma_2^{(i)}) : 0 : 0\right] H\bar{\Upsilon}'\tilde{u}_1\right\} \\
&- \mathbb{E}\left\{H\left[\sum_{i=1}^{p} L^i\bar{Y}(\hat{\Gamma}_2^{(i)} - \Gamma_2^{(i)}) + \sum_{j=1}^{q} L^j X(\hat{\Pi}_2^{(i)} - \Pi_2^{(i)})\right.\right. \\
&\quad \left.\left. + \sum_{i=1}^{p}\right] L^i\tilde{W}(\hat{\Gamma}_2^{(i)} - \Gamma_2^{(i)}) : 0 : 0]'\bar{\Upsilon}H\bar{\Upsilon}'\tilde{u}_1\right\} \\
&- \mathbb{E}\left\{H\left[\bar{\Upsilon}'\sum_{i=1}^{p} L^i\tilde{W}C^* + C^{*'}\left(\sum_{i=1}^{p} L^i\tilde{W}\right)'\bar{\Upsilon}\right] H\bar{\Upsilon}'\tilde{u}_1\right\} \\
&- \mathbb{E}\left\{H\left[\sum_{i=1}^{p} L^i\bar{Y}(\hat{\Gamma}_2^{(i)} - \Gamma_2^{(i)}) + \sum_{j=1}^{q} L^j X(\hat{\Pi}_2^{i} - \Pi_2^{i})\right.\right. \\
&\quad \left.\left. + \sum_{i=1}^{p} L^i\tilde{W}(\hat{\Gamma}_2^{(i)} - \Gamma_2^{(i)}) : 0 : 0\right]' \left[\sum_{i=1}^{p} L^i\tilde{W}\Gamma_2^{(i)} : \tilde{R}_1 : 0\right] H\bar{\Upsilon}'\tilde{u}_1\right\} \\
&- \mathbb{E}\left\{H\left[\sum_{i=1}^{p} L^i\tilde{W}\Gamma_2^{(i)} : \tilde{R}_1 : 0\right]' \left[\sum_{i=1}^{p} L^i\bar{Y}(\hat{\Gamma}_2^{(i)} - \Gamma_2^{(i)})\right.\right. \\
&\quad \left.\left. + \sum_{j=1}^{q} L^j X(\hat{\Pi}_2^{i} - \Pi_2^{i}) + \sum_{i=1}^{p} L^i\tilde{W}(\hat{\Gamma}_2^{(i)} - \Gamma_2^{(i)}) : 0 : 0\right] H\bar{\Upsilon}'\tilde{u}_1\right\} \\
&+ o(T^{-1}),
\end{aligned}
$$

where the definition of Δ_1 and Δ_2 is from equation 2.10, and the expression of J_1^* is in the footnote 1 in section 2.3.

By using the expression

$$\begin{bmatrix} \hat{\Gamma}_2^{(1)} - \Gamma_2^{(1)} \\ \vdots \\ \hat{\Gamma}_2^{(p)} - \Gamma_2^{(p)} \\ \hat{\Pi}_2^{(1)} - \Pi_2^{(1)} \\ \vdots \\ \hat{\Pi}_2^{(q)} - \Pi_2^{(q)} \end{bmatrix} = (Z^{'}Z)^{-1}Z^{'}\tilde{V}_2 = (\mathbb{E}\{Z^{'}Z\})^{-1}\bar{Z}^{'}\tilde{V}_2 + (\mathbb{E}\{Z^{'}Z\})^{-1}\tilde{W}^{*'}\tilde{V}_2 + o_p(T^{-1/2})$$

and

$$\begin{bmatrix} \hat{\Gamma}_2^{(1)} - \Gamma_2^{(1)} \\ \vdots \\ \hat{\Gamma}_2^{(p)} - \Gamma_2^{(p)} \end{bmatrix} = I_2^{'}(Z^{'}Z)^{-1}Z^{'}\tilde{V}_2 = I_2^{\prime}(\mathbb{E}\{Z^{'}Z\})^{-1}\bar{Z}^{'}\tilde{V}_2 + I_2^{\prime}(\mathbb{E}\{Z^{'}Z\})^{-1}\tilde{W}^{*'}\tilde{V}_2$$
$$+ o_p(T^{-1/2})$$

and

$$[\tilde{V}_2 : 0 : 0] = \tilde{V}\Lambda^{**},$$

the above (iv) expression can be written as:

$$-\mathbb{E}\{H\bar{\Upsilon}^{'}\bar{Z}[\mathbb{E}(Z^{'}Z)]^{-1}\bar{Z}^{'}\tilde{V}\Lambda^{**}H\bar{\Upsilon}^{'}\tilde{u}_1\} - \mathbb{E}\{H\bar{\Upsilon}^{'}\tilde{R}I_2^{'}[\mathbb{E}(Z^{\prime}Z)]^{-1}\tilde{W}^{*'}\tilde{V}\Lambda^{**}H\bar{\Upsilon}^{'}\tilde{u}_1\}$$
$$-\mathbb{E}\left\{H\Lambda^{**'}\tilde{V}^{'}\bar{Z}[\mathbb{E}(Z^{\prime}Z)]^{-1}\bar{Z}^{'}\bar{\Upsilon}H\bar{\Upsilon}^{'}\tilde{u}_1\right\} - \mathbb{E}\{H\Lambda^{**'}\tilde{V}^{'}\tilde{W}^{*}[\mathbb{E}(Z^{\prime}Z)]^{-1}I_2\tilde{R}^{'}\bar{\Upsilon}H\bar{\Upsilon}^{'}\tilde{u}_1\}$$
$$-\mathbb{E}\{H\bar{\Upsilon}^{'}\tilde{R}CH\bar{\Upsilon}^{'}\tilde{u}_1\} - \mathbb{E}\{HC^{*'}\tilde{R}^{'}\bar{\Upsilon}H\bar{\Upsilon}^{'}\tilde{u}_1\}$$
$$-\mathbb{E}\{H\Lambda^{**'}\tilde{V}^{'}\tilde{W}^{*}(\mathbb{E}(Z^{'}Z))^{-1}\bar{Z}^{'}\tilde{R}CH\bar{\Upsilon}^{'}\tilde{u}_1\} - \mathbb{E}\{H\Lambda^{**'}\tilde{V}^{'}\bar{Z}(\mathbb{E}(Z^{'}Z))^{-1}I_2\tilde{R}^{'}\tilde{R}CH\bar{\Upsilon}^{'}\tilde{u}_1\}$$
$$-\mathbb{E}\{HC^{'}\tilde{R}^{'}\bar{Z}(\mathbb{E}(Z^{'}Z))^{-1}\tilde{W}^{*'}\tilde{V}\Lambda^{**}H\bar{\Upsilon}^{'}\tilde{u}_1\} - \mathbb{E}\{HC^{'}\tilde{R}^{'}\tilde{R}I_2^{\prime}(\mathbb{E}(Z^{'}Z))^{-1}\bar{Z}^{'}\tilde{V}\Lambda^{**}H\bar{\Upsilon}^{'}\tilde{u}_1\}.$$

Using $\tilde{R}$ ($\tilde{R} = \sum_{i=1}^{p}\sum_{t=i}^{T-1} D^t\tilde{V}J_t\Psi_i^{'}$, and $\tilde{W}^* = \left[\sum_{i=1}^{p}\sum_{t=i}^{T-1} D^t\tilde{V}J_t\Psi_i^{'} : 0\right]$ from equation 2.5, and the decomposition of $\tilde{V}$, $\tilde{V} = S^* + \tilde{u}_1\phi^{'}$, then (iv) can be obtained from the sum of $(1)-(8)$ below:

(1)

$$-\mathbb{E}\{H\bar{\Upsilon}^{'}\bar{Z}[\mathbb{E}(Z^{'}Z)]^{-1}\bar{Z}^{'}\tilde{V}\Lambda^{**}H\bar{\Upsilon}^{'}\tilde{u}_1\} = -\mathbb{E}\{H\bar{\Upsilon}^{'}\bar{Z}[\mathbb{E}(Z^{'}Z)]^{-1}\bar{Z}^{'}\tilde{u}_1\phi^{'}\Lambda^{**}H\bar{\Upsilon}^{'}\tilde{u}_1\} \tag{A.6}$$

$$= -H\bar{\Upsilon}^{'}\bar{Z}[\mathbb{E}(Z^{'}Z)]^{-1}\bar{Z}^{'}\bar{\Upsilon}H^{'}\Lambda^{**'}(\sigma_1^2\phi).$$

(2)

$$-\mathbb{E}\{H\bar{\Upsilon}^{'}\tilde{R}I_2^{'}[\mathbb{E}(Z'Z)]^{-1}\tilde{F}^{*'}\tilde{V}\Lambda^{**}H\bar{\Upsilon}^{'}\tilde{u}_1\} \tag{A.7}$$

$$= -\mathbb{E}\left\{H\bar{\Upsilon}^{'}\sum_{i=1}^{p}\sum_{t=i}^{T-1} D^t\tilde{V}J_{t-i}\Psi_i^{'}I_2^{'}[\mathbb{E}(Z'Z)]^{-1}\begin{bmatrix}\sum_{j=1}^{p}\sum_{s=j}^{T-1}\Psi_jJ_{s-j}^{'}\tilde{V}^{'}D^{t'}\\ 0\end{bmatrix}\tilde{V}\Lambda^{**}H\bar{\Upsilon}^{'}\tilde{u}_1\right\}.$$

For the moment, we shall focus on the the following equation (Moving the summations and first three fixed terms H, $\bar{\Upsilon}^{'}$, and D^t outside of expectation symbol):

$$\mathbb{E}\left\{\tilde{V}J_{t-i}\Psi_i^{'}I_2^{'}[\mathbb{E}(Z'Z)]^{-1}\begin{bmatrix}\Psi_jJ_{s-j}^{'}\tilde{V}^{'}D^{s'}\\ 0\end{bmatrix}\tilde{V}\Lambda^{**}H\bar{\Upsilon}^{'}\tilde{u}_1\right\} \tag{A.8}$$

$$\begin{aligned}
&= \mathbb{E}\{\tilde{u}_1\phi^{'}J_{t-i}\Psi_i^{'}I_2^{'}[\mathbb{E}(Z'Z)]^{-1}I_2\Psi_jJ_{s-j}^{'}\phi\tilde{u}_1^{'}D^{s'}\tilde{u}_1\phi^{'}\Lambda^{**}H\bar{\Upsilon}^{'}\tilde{u}_1\}\\
&\quad+\mathbb{E}\{S^*J_{t-i}\Psi_i^{'}I_2^{'}[\mathbb{E}(Z'Z)]^{-1}I_2\Psi_jJ_{s-j}^{'}S^{*'}D^{s'}\tilde{u}_1\phi^{'}\Lambda^{**}H\bar{\Upsilon}^{'}\tilde{u}_1\}\\
&\quad+\mathbb{E}\{\tilde{u}_1\phi^{'}J_{t-i}\Psi_i^{'}I_2^{'}[\mathbb{E}(Z'Z)]^{-1}I_2\Psi_jJ_{s-j}^{'}S^{*'}D^{s'}S^*\Lambda^{**}H\bar{\Upsilon}^{'}\tilde{u}_1\}\\
&\quad+\mathbb{E}\{S^*J_{t-i}\Psi_i^{'}I_2^{'}[\mathbb{E}(Z'Z)]^{-1}I_2\Psi_jJ_{s-j}^{'}\phi\tilde{u}_1^{'}D^{s'}S^*\Lambda^{**}H\bar{\Upsilon}^{'}\tilde{u}_1\}\}.
\end{aligned}$$

Then equation (A.8) will be calculated from $(a)-(d)$ below:

(a)

$$\begin{aligned}
&\mathbb{E}\{\tilde{u}_1\phi^{'}J_{t-i}\Psi_i^{'}I_2^{'}[\mathbb{E}(Z'Z)]^{-1}I_2\Psi_jJ_{s-j}^{'}\phi\tilde{u}_1^{'}D^{s^{'}}\tilde{u}_1\phi^{'}\Lambda^{**}H\bar{\Upsilon}^{'}\tilde{u}_1\} \\
&\quad=\mathbb{E}\left\{\tilde{u}_1\tilde{u}_1^{'}\bar{\Upsilon}H^{'}\Lambda^{**^{'}}\phi\phi^{'}J_{t-i}\Psi_i^{'}I_2^{'}[\mathbb{E}(Z'Z)]^{-1}\begin{bmatrix}\Psi_jJ_{s-j}^{'}\\0\end{bmatrix}\phi\tilde{u}_1^{'}D^{s^{'}}\tilde{u}_1\right\} \\
&\quad=\sigma_1^4(tr(D^{s^{'}})I+D^s+D^{s^{'}})\bar{\Upsilon}H^{'}\Lambda^{**^{'}}\phi\phi^{'}J_{t-i}\Psi_i^{'}I_2^{'}[\mathbb{E}(Z'Z)]^{-1}I_2\Psi_jJ_{s-j}^{'}\phi \\
&\quad=\sigma_1^4(D^s+D^{s^{'}})\bar{\Upsilon}H^{'}\Lambda^{**^{'}}\phi\phi^{'}J_{t-i}\Psi_i^{'}I_2^{'}[\mathbb{E}(Z'Z)]^{-1}I_2\Psi_jJ_{s-j}^{'}\phi.
\end{aligned}$$

Using Lemma 4.

(b)

$$\begin{aligned}
&\mathbb{E}\{S^*J_{t-i}\Psi_i^{'}I_2^{'}[\mathbb{E}(Z'Z)]^{-1}I_2\Psi_jJ_{s-j}^{'}S^{*^{'}}D^{s^{'}}\tilde{u}_1\phi^{'}\Lambda^{**}H\bar{\Upsilon}^{'}\tilde{u}_1\} \\
&\quad=\mathbb{E}\left\{S^*J_{t-i}\Psi_i^{'}I_2^{'}[\mathbb{E}(Z'Z)]^{-1}\begin{bmatrix}\Psi_jJ_{s-j}^{'}\\0\end{bmatrix}S^{*^{'}}D^{s^{'}}\tilde{u}_1\tilde{u}_1^{'}\bar{\Upsilon}H^{'}\Lambda^{**^{'}}\phi\right\} \\
&\quad=\sigma^2tr\{C_2^*J_{t-i}\Psi_i^{'}I_2^{'}[\mathbb{E}(Z'Z)]^{-1}I_2\Psi_jJ_{s-j}^{'}\}D^{s^{'}}\bar{\Upsilon}H^{'}\Lambda^{**^{'}}\phi \\
&\quad=\sigma^2tr\{(\Omega-\phi\phi^{'}\sigma^2)J_{t-i}\Psi_i^{'}I_2^{'}[\mathbb{E}(Z'Z)]^{-1}I_2\Psi_jJ_{s-j}^{'}\}D^{s^{'}}\bar{\Upsilon}H^{'}\Lambda^{**^{'}}\phi.
\end{aligned}$$

Using Lemma 3, $\mathbb{E}\{S^*AS^{*^{'}}\}=tr\{C_2^*A\}I$.

(c)

$$\begin{aligned}
&\mathbb{E}\{\tilde{u}_1\phi^{'}J_{t-i}\Psi_i^{'}I_2^{'}[\mathbb{E}(Z'Z)]^{-1}I_2\Psi_jJ_{s-j}^{'}S^{*^{'}}D^{s^{'}}S^*\Lambda^{**}H\bar{\Upsilon}^{'}\tilde{u}_1\} \\
&\quad=\mathbb{E}\left\{\tilde{u}_1\tilde{u}_1^{'}\bar{\Upsilon}H^{'}\Lambda^{**^{'}}S^{*^{'}}D^sS^*\begin{bmatrix}\Psi_jJ_{s-j}^{'}\\0\end{bmatrix}^{'}I_2^{'}[\mathbb{E}(Z'Z)]^{-1}I_2\Psi_iJ_{t-i}^{'}\phi\right\} \\
&\quad=\sigma^2\bar{\Upsilon}H\Lambda^{**^{'}}\mathbb{E}(S^{*^{'}}D^{s^{'}}S^*)J_{s-j}\Psi_j^{'}I_2^{'}[\mathbb{E}(Z'Z)]^{-1}I_2\Psi_iJ_{t-i}^{'}\phi \\
&\quad=\sigma^2\bar{\Upsilon}H\Lambda^{**^{'}}tr\{D^s\}C_2^*J_{s-j}\Psi_j^{'}I_2^{'}[\mathbb{E}(Z'Z)]^{-1}I_2J_{t-i}^{'}\Psi_i\phi=0.
\end{aligned}$$

Using Lemma 3, $\mathbb{E}\{S^{*'}AS^*\} = tr\{A\}C_2^*$.

(d)

$$\begin{aligned}
&\mathbb{E}\{S^* J_{t-i}\Psi_i^{'} I_2^{'}[\mathbb{E}(Z'Z)]^{-1} I_2 \Psi_j J_{s-j}^{'} \phi \tilde{u}_1^{'} D^{s'} S^* \Lambda^{**} H \bar{\Upsilon}^{'} \tilde{u}_1\} \\
&\quad = \mathbb{E}\{D^s \tilde{u}_1 \phi^{'} J_{s-j} \Psi_j^{'} I_2^{'}[\mathbb{E}(Z'Z)]^{-1} I_2 \Psi_i J_{t-i}^{'} C_2^* \Lambda^{**} H \bar{\Upsilon}^{'} \tilde{u}_1\}.
\end{aligned}$$

Using the definition of S^* and $\tilde{u}_1$: S^* and $\tilde{u}_1$ are independent, and Lemma 3 that $\mathbb{E}\{S^* A S^*\} = A^{'} C_2^*$.

Then, we have:

$$\begin{aligned}
&\mathbb{E}\{D^s \tilde{u}_1 \phi^{'} J_{s-j} \Psi_j^{'} I_2^{'}[\mathbb{E}(Z'Z)]^{-1} I_2 \Psi_i J_{t-i}^{'} C_2^* \Lambda^{**} H \bar{\Upsilon}^{'} \tilde{u}_1\} \\
&\quad = \mathbb{E}\{D^s \tilde{u}_1 \tilde{u}_1^{'} \bar{\Upsilon} H^{'} \Lambda^{**'} C_2^* J_{t-i} \Psi_i^{'} I_2^{'}[\mathbb{E}(Z'Z)]^{-1} I_2 \Psi_j J_{s-j}^{'} \phi\} \\
&\quad = \sigma^2 D^s \bar{\Upsilon} H \Lambda^{**'} C_2^* J_{t-i} \Psi_i^{'} I_2^{'}[\mathbb{E}(Z'Z)]^{-1} I_2 \Psi_j J_{s-j}^{'} \phi \\
&\quad = \sigma^2 D^s \bar{\Upsilon} H \Lambda^{**'} \Omega J_{t-i} \Psi_i^{'} I_2^{'}[\mathbb{E}(Z'Z)]^{-1} I_2 \Psi_j J_{s-j}^{'} \phi \\
&\qquad - \sigma^4 2 D^s \bar{\Upsilon} H \Lambda^{**'} \phi \phi^{'} J_{t-i} \Psi_i^{'} I_2^{'}[\mathbb{E}(Z'Z)]^{-1} I_2 \Psi_i J_{s-j}^{'} \phi.
\end{aligned}$$

Putting $(a)-(d)$ together, we have:

$$\begin{aligned}
&\mathbb{E}\left\{ \tilde{V} J_{t-i} \Psi_i^{'} I_2^{'}[\mathbb{E}(Z'Z)]^{-1} \begin{bmatrix} \Psi_j J_{s-j}^{'} \tilde{V}^{'} D^{s'} \\ 0 \end{bmatrix} \tilde{V} \Lambda^{**} H \bar{\Upsilon}^{'} \tilde{u}_1 \right\} \\
&\quad = \sigma^2 D^{s'} \bar{\Upsilon} H \Lambda^{**'} \phi tr\{\Omega J_{t-i} \Psi_i^{'} I_2^{'}[\mathbb{E}(Z'Z)]^{-1} I_2 \Psi_i J_{s-j}^{'}\} \\
&\qquad + \sigma^2 D^s \bar{\Upsilon} H \Lambda^{**'} \Omega J_{t-i} \Psi_i^{'} I_2^{'}[\mathbb{E}(Z'Z)]^{-1} I_2 \Psi_j J_{s-j}^{'} \phi.
\end{aligned}$$

Then, equation A.7 becomes:

$$-\mathbb{E}\{H \bar{\Upsilon}^{'} \tilde{R} I_2^{'}[\mathbb{E}(Z'Z)]^{-1} \tilde{F}^{*'} \tilde{V} \Lambda^{**} H \bar{\Upsilon}^{'} \tilde{u}_1\} \tag{A.9}$$

$$= -H\bar{\Upsilon}'\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1} D^t D^{s'}\bar{\Upsilon}Htr\left\{\Omega\left[J_{s-j}\Psi_j^{'}:0\right][\mathbb{E}(Z'Z)]^{-1}I_2\Psi_i J_{t-i}^{'}\right\}\Lambda^{**'}(\sigma^2\phi)$$
$$-H\bar{\Upsilon}'\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1} D^t D^s\bar{\Upsilon}H\Lambda^{**'}\Omega J_{t-i}\Psi_i^{'}I_2^{'}[\mathbb{E}(Z'Z)]^{-1}I_2\Psi_j J_{s-j}^{'}(\sigma^2\phi).$$

(3)

$$\begin{aligned} &-\mathbb{E}\{H\Lambda^{**'}\tilde{V}'\bar{Z}[\mathbb{E}(Z'Z)]^{-1}\bar{Z}'\bar{\Upsilon}H\bar{\Upsilon}'\tilde{u}_1\} \\ &= -\mathbb{E}\{H\Lambda^{**'}\phi\tilde{u}_1^{'}\bar{Z}[\mathbb{E}(Z'Z)]^{-1}\bar{\Upsilon}H\bar{\Upsilon}'\tilde{u}_1\} \\ &= -H\Lambda^{**'}\phi\sigma^2 tr\{\bar{Z}[\mathbb{E}(Z'Z)]^{-1}\bar{\Upsilon}H\bar{\Upsilon}'\} \\ &= -Htr\{\bar{Z}[\mathbb{E}(Z'Z)]^{-1}\bar{\Upsilon}H\bar{\Upsilon}'\}\Lambda^{**'}(\sigma^2\phi). \end{aligned} \tag{A.10}$$

(4)

$$\begin{aligned} &-\mathbb{E}\{H\Lambda^{**'}\tilde{V}'\tilde{W}^*[\mathbb{E}(Z'Z)]^{-1}I_2\tilde{R}'\bar{\Upsilon}H\bar{\Upsilon}'\tilde{u}_1\} \\ &= -\mathbb{E}\left\{H\Lambda^{**'}\tilde{V}'\left[\sum_{i=1}^{p}\sum_{t=i}^{T-1} D^t\tilde{V}J_{t-i}\Psi_i^{'}:0\right][\mathbb{E}(Z'Z)]^{-1}I_2\sum_{j=1}^{p}\sum_{s=j}^{T-1}\Psi_j J_{s-j}^{'}\tilde{V}'D^{s'}\bar{\Upsilon}H\bar{\Upsilon}'\tilde{u}_1\right\}. \end{aligned} \tag{A.11}$$

Here:

$$\begin{aligned} &\mathbb{E}\left\{\tilde{V}'D^t\tilde{V}J_{t-i}\Psi_i^{'}I_2^{'}[\mathbb{E}(Z'Z)]^{-1}I_2\Psi_j J_{s-j}^{'}\tilde{V}'D^{s'}\bar{\Upsilon}H\bar{\Upsilon}'\tilde{u}_1\right\} \\ &= \mathbb{E}\{\phi\tilde{u}_1^{'}D^t\tilde{u}_1\phi' J_{t-i}\Psi_i^{'}I_2^{'}[E(Z'Z)]^{-1}I_2\Psi_j J_{s-j}^{'}\phi\tilde{u}_1^{'}D^{s'}\bar{\Upsilon}H\bar{\Upsilon}'\tilde{u}_1\} \\ &\quad +\mathbb{E}\{\phi\tilde{u}_1^{'}D^t S J_{t-i}\Psi_i^{'}I_2^{'}[\mathbb{E}(Z'Z)]^{-1}I_2 v_j J_{s-j}^{'}S' D^{s'}\bar{\Upsilon}H\bar{\Upsilon}'\tilde{u}_1\} \\ &\quad +\mathbb{E}\{S' D^t\tilde{u}_1\phi' J_{t-i}\Psi_i^{'}I_2^{'}[\mathbb{E}(Z'Z)]^{-1}I_2\Psi_j J_{s-j}^{'}S' D^{s'}\bar{\Upsilon}H\bar{\Upsilon}'\tilde{u}_1\} \\ &\quad +\mathbb{E}\{S' D^t S J_{t-i}\Psi_i^{'}I_2^{'}[\mathbb{E}(Z'Z)]^{-1}I_2\Psi_j J_{s-j}^{'}\phi\tilde{u}_1^{'}D^{s'}\bar{\Upsilon}H\bar{\Upsilon}'\tilde{u}_1\} \\ &= \sigma^4\phi\phi' J_{t-i}\Psi_i^{'}I_2^{'}[\mathbb{E}(Z'Z)]^{-1}I_2\Psi_j J_{s-j}^{'}\phi tr\{(D^t+D^{t'})D^{s'}\bar{\Upsilon}H\bar{\Upsilon}'\} \end{aligned}$$

$$
\begin{aligned}
&+\sigma^2 tr\{\Omega J_{t-i}\Psi_i^{'}I_2^{'}[\mathbb{E}(Z'Z)]^{-1}I_2\Psi_j J_{s-j}^{'}\}tr\{D^t D^{s'}\bar{\Upsilon}H\bar{\Upsilon}^{'}\}\phi \\
&-\sigma^4 tr\{\phi J_{t-i}\Psi_{\phi}^{'}i^{'}I_2^{'}[\mathbb{E}(Z'Z)]^{-1}I_2\Psi_j J_{s-j}^{'}\}tr\{D^t D^{s'}\bar{\Upsilon}H\bar{\Upsilon}^{'}\}\phi \\
&+\sigma^2 tr\{D^t\bar{\Upsilon}H\bar{\Upsilon}^{'}D^s\}\Omega J_{s-j}\Psi_j^{'}I_2^{'}[\mathbb{E}(Z'Z)]^{-1}I_2 v_i J_{t-i}^{'}\phi \\
&-\sigma^4 tr\{D^t\bar{\Upsilon}H\bar{\Upsilon}^{'}D^s\}\phi\phi^{'}J_{s-j}\Psi_j^{'}I_2^{'}[\mathbb{E}(Z'Z)]^{-1}I_2\Psi_i J_{t-i}^{'}\phi \\
&+0 \\
=\;&\sigma^2 tr\{\Omega J_{t-i}\Psi_i^{'}I_2^{'}[\mathbb{E}(Z'Z)]^{-1}I_2\Psi_j J_{s-j}^{'}\}tr\{D^t D^{s'}\bar{\Upsilon}H\bar{\Upsilon}^{'}\}\phi \\
&+\sigma^2 tr\{D^t\bar{\Upsilon}H\bar{\Upsilon}^{'}D^s\}\Omega J_{s-j}\Psi_j^{'}I_2^{'}[\mathbb{E}(Z'Z)]^{-1}I_2\Psi_i J_{t-i}^{'}\phi.
\end{aligned}
$$

Using Lemma 3.

The final expression for equation (A.11) is :

$$
\begin{aligned}
&-\mathbb{E}\{H\Lambda^{**'}\tilde{V}^{'}\tilde{W}^{*}[\mathbb{E}(Z'Z)]^{-1}I_2\tilde{R}^{'}\bar{\Upsilon}H\bar{\Upsilon}^{'}\tilde{u}_1\} \qquad (A.12)\\
&\quad=-H\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1}(tr\{\Omega J_{t-i}\Psi_i^{'}I_2^{'}[\mathbb{E}(Z'Z)]^{-1}I_2\Psi_j J_{s-j}^{'}\}.I)\\
&\qquad\qquad\times(tr\{D^t D^{s'}\bar{\Upsilon}H\bar{\Upsilon}^{'}\}.I)\Lambda^{**'}(\sigma^2\phi)\\
&\qquad-H\Lambda^{**'}\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1}\Omega J_{s-j}\Psi_j^{'}I_2^{'}[\mathbb{E}(Z'Z)]^{-1}I_2\Psi_i J_{t-i}^{'}(tr\{D^t\bar{\Upsilon}H\bar{\Upsilon}^{'}D^s\}.I)(\sigma^2\phi).
\end{aligned}
$$

(5)

$$
\begin{aligned}
&-\mathbb{E}\left\{H\bar{\Upsilon}^{'}\tilde{R}C^{*}H\bar{\Upsilon}^{'}\tilde{u}_1\right\} \qquad (A.13)\\
&\quad=-\mathbb{E}\left\{H\bar{\Upsilon}^{'}\sum_{i=1}^{p}\sum_{t=i}^{T-1}D^t S J_{t-i}\Psi_i^{'}C^{*}H\bar{\Upsilon}^{'}\tilde{u}_1\right\}-\mathbb{E}\left\{H\bar{\Upsilon}^{'}\sum_{i=1}^{p}\sum_{t=i}^{T-1}D^t\tilde{u}_1\phi^{'}J_{t-i}\Psi_i^{'}C^{*}H\bar{\Upsilon}^{'}\tilde{u}_1\right\}\\
&\quad=-\mathbb{E}\left\{H\bar{\Upsilon}^{'}\sum_{i=1}^{p}\sum_{t=i}^{T-1}D^t\tilde{u}_1\tilde{u}_1^{'}\bar{\Upsilon}H^{'}C^{*'}\Psi_i J_{t-i}^{'}\phi\right\}\\
&\quad=-H\bar{\Upsilon}^{'}\sum_{i=1}^{p}\sum_{t=i}^{T-1}D^t\bar{\Upsilon}H^{'}C^{*'}\Psi_i J_{t-i}^{'}(\sigma^2\phi).
\end{aligned}
$$

(6)

$$-\mathbb{E}\left\{HC^{*'}\tilde{R}'\bar{\Upsilon}H\bar{\Upsilon}'\tilde{u}_1\right\} = -\mathbb{E}\left\{HC^{*'}\sum_{i=1}^{p}\sum_{t=i}^{T-1}\Psi_i J_{t-i}'\tilde{V}'D^{t'}\bar{\Upsilon}H\bar{\Upsilon}'\tilde{u}_1\right\} \tag{A.14}$$

$$= -\mathbb{E}\left\{H\sum_{i=1}^{p}\sum_{t=i}^{T-1}C^{*'}\Psi_i J_{t-i}'\phi\tilde{u}_1'D^{t'}\bar{\Upsilon}H\bar{\Upsilon}'\tilde{u}_1\right\}$$

$$= -H\sum_{i=1}^{p}\sum_{t=i}^{T-1}C^{*'}\Psi_i J_{t-i}'(tr\{\bar{\Upsilon}'D^{t'}\bar{\Upsilon}H\}.I)(\sigma^2\phi).$$

(7)

$$-\mathbb{E}\{H\Lambda^{**'}\tilde{V}'\tilde{W}^*(\mathbb{E}(Z'Z))^{-1}\bar{Z}'\tilde{R}C^*H\bar{\Upsilon}'\tilde{u}_1\} - \mathbb{E}\{H\Lambda^{**'}\tilde{V}'\bar{Z}(\mathbb{E}(Z'Z))^{-1}I_2\tilde{R}'\tilde{R}C^*H\bar{\Upsilon}'\tilde{u}_1\} \tag{A.15}$$

$$= -\mathbb{E}\left\{H\Lambda^{**'}\tilde{V}'\left[\sum_{i=1}^{p}\sum_{t=i}^{T-1}D^t\tilde{V}J_{t-i}\Psi_i' : 0\right](\mathbb{E}(Z'Z))^{-1}\bar{Z}'\sum_{j=1}^{p}\sum_{s=j}^{T-1}D^s\tilde{V}J_{s-j}\Psi_j'C^*H\bar{\Upsilon}'\tilde{u}_1\right\}$$

$$-\mathbb{E}\left\{H\Lambda^{**'}\tilde{V}'\bar{Z}(\mathbb{E}(Z'Z))^{-1}I_2\left(\sum_{l=1}^{p}\sum_{r=l}^{T-1}D^r\tilde{V}J_{r-l}\Psi_l'\right)'\sum_{j=1}^{p}\sum_{s=j}^{T-1}D^s\tilde{V}J_{s-j}\Psi_j'C^*H\bar{\Upsilon}'\tilde{u}_1\right\}.$$

This is calculated in two parts *(7a)* and *(7b)*:

(7a)

$$\mathbb{E}\{\tilde{V}'[D^t\tilde{V}J_{t-i}\Psi_i' : 0](\mathbb{E}(Z'Z))^{-1}\bar{Z}'D^s\tilde{V}J_{t-j}\Psi_j'C^*H\bar{\Upsilon}'\tilde{u}_1\}$$

$$= \mathbb{E}\{\phi\tilde{u}_1'D^t\tilde{u}_1\phi'J_{t-i}\Psi_i'I_2'(\mathbb{E}(Z'Z))^{-1}\bar{Z}'D^s\tilde{u}_1\phi'J_{s-j}'\Psi_jC^*H\bar{\Upsilon}'\tilde{u}_1\}$$

$$+\mathbb{E}\{\phi\tilde{u}_1'D^tSJ_{t-i}\Psi_i'I_2'(\mathbb{E}(Z'Z))^{-1}\bar{Z}'D^sSJ_{s-j}\Psi_j'C^*H\bar{\Upsilon}'\tilde{u}_1\}$$

$$+\mathbb{E}\{S'D^{t'}SJ_{t-i}\Psi_i'I_2'(\mathbb{E}(Z'Z))^{-1}\bar{Z}'D^s\tilde{u}_1\phi'J_{s-j}\Psi_j'C^*H\bar{\Upsilon}'\tilde{u}_1\}$$

$$+\mathbb{E}\{S'D^t\tilde{u}_1\phi'J_{t-i}\Psi_i'I_2'(\mathbb{E}(Z'Z))^{-1}\bar{Z}'D^sSJ_{s-j}\Psi_j'C^*H\bar{\Upsilon}'\tilde{u}_1\}$$

$$= \phi\sigma^4 tr\{(D^t + D^{t'})D^{s'}\bar{Z}(\mathbb{E}(Z'Z))^{-1}I_2\Psi_iJ_{t-i}'\phi\phi'J_{s-j}\Psi_j'C^*H\bar{\Upsilon}'\}$$

$$+\sigma^2\phi tr\{D^tD^{s'}\bar{Z}(\mathbb{E}(Z'Z))^{-1}I_2\Psi_iJ_{t-i}'\Omega J_{s-j}\Psi_j'C^*H\bar{\Upsilon}'\}$$

$$
\begin{aligned}
&-\sigma^4\phi tr\{D^t D^{s'}\bar{Z}(\mathbb{E}(Z^{'}Z))^{-1}I_2\Psi_i J^{'}_{t-i}\phi\phi^{'}J_{s-j}\Psi^{'}_j C^* H\bar{\Upsilon}^{'}\}\\
&+0\\
&+\sigma^2\Omega(D^t\bar{\Upsilon}HC^{*'}\Psi_j J^{'}_{s-j})^{'}D^{s'}\bar{Z}(\mathbb{E}(Z^{'}Z))^{-1}I_2\Psi_i J^{'}_{t-i}\phi\\
&-\sigma^4\phi\phi^{'}(D^t\bar{\Upsilon}HC^{*'}\Psi_j J^{'}_{s-j})^{'}D^{s'}\bar{Z}(\mathbb{E}(Z^{'}Z))^{-1}I_2\Psi_i J^{'}_{t-i}\phi\\
=&\ (tr\{D^t D^{s'}\bar{Z}(\mathbb{E}(Z^{'}Z))^{-1}I_2\Psi_i J^{'}_{t-i}\Omega J_{s-j}\Psi^{'}_j cH\bar{\Upsilon}^{'}\}.I)(\sigma^2\phi)\\
&+\Omega J_{s-j}\Psi^{'}_j C^* H\bar{\Upsilon}^{'}D^{t'}D^{s'}\bar{Z}(\mathbb{E}(Z^{'}Z))^{-1}I_2\Psi_i J^{'}_{t-i}(\sigma^2\phi).
\end{aligned}
$$

Using Lemma 3.

The final expression of the first part of equation (A.15) can be written as:

$$
\begin{aligned}
&-\mathbb{E}\{H\Lambda^{**'}\tilde{V}^{'}\tilde{W}^*(\mathbb{E}(Z^{'}Z))^{-1}\bar{Z}^{'}\tilde{R}C^*H\bar{\Upsilon}^{'}\tilde{u}_1\} \qquad \text{(A.16)}\\
&\quad=-\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1}H(tr\{D^t D^{s'}\bar{Z}(\mathbb{E}(Z^{'}Z))^{-1}I_2\Psi_i J^{'}_{t-i}\Omega J_{s-j}\Psi^{'}_j C^* H\bar{\Upsilon}^{'}\}.I)\Lambda^{**'}(\sigma^2\phi)\\
&\quad\quad-\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1}H\Lambda^{**'}\Omega J_{s-j}\Psi^{'}_j C^* H\bar{\Upsilon}^{'}D^{t'}D^{s'}\bar{Z}(\mathbb{E}(Z^{'}Z))^{-1}I_2\Psi_i J^{'}_{t-i}(\sigma^2\phi).
\end{aligned}
$$

(7b)

$$
\begin{aligned}
&-\mathbb{E}\left\{\tilde{V}^{'}\bar{Z}(E(Z^{'}Z))^{-1}I_2\Psi_l J^{'}_{r-l}\tilde{V}'D^{r'}D^s\tilde{V}J_{s-j}\Psi^{'}_j C^* H\bar{\Upsilon}^{'}\tilde{u}_1\right\}\\
&\quad=-\mathbb{E}\left\{\phi u^{'}_1\bar{Z}(E(Z^{'}Z))^{-1}I_2\Psi_l J^{'}_{r-l}\phi u^{'}_1 D^{r'}D^s u_1\phi' J_{s-j}\Psi^{'}_j C^* H\bar{\Upsilon}^{'}\tilde{u}_1\right\}\\
&\quad=-E\left\{\phi u^{'}_1\bar{Z}(\mathbb{E}(Z^{'}Z))^{-1}I_2\Psi_l J^{'}_{r-l}S'D^{r'}D^s S J_{s-j}\Psi^{'}_j C^* H\bar{\Upsilon}^{'}\tilde{u}_1\right\}\\
&\quad=-E\left\{S'\bar{Z}(\mathbb{E}(Z^{'}Z))^{-1}I_2\Psi_l J^{'}_{r-l}\phi u^{'}_1 D^{r'}D^s S J_{s-j}\Psi^{'}_j C^* H\bar{\Upsilon}^{'}\tilde{u}_1\right\}\\
&\quad=-\mathbb{E}\left\{S'\bar{Z}(\mathbb{E}(Z^{'}Z))^{-1}I_2\Psi_l J^{'}_{r-l}S'D^{r'}D^s u_1\phi' J_{s-j}\Psi^{'}_j C^* H\bar{\Upsilon}^{'}\tilde{u}_1\right\}\\
&\quad=-\phi\sigma^2 tr\left\{\bar{Z}(\mathbb{E}(Z^{'}Z))^{-1}I_2\Psi_l J^{'}_{r-l}\Omega J_{s-j}\Psi^{'}_j C^* H\bar{\Upsilon}^{'}\right\}tr\left\{D^{r'}D^s\right\}\\
&\quad\quad-\Omega J_{s-j}\Psi^{'}_j C^* H\bar{\Upsilon}^{'}D^{r'}D^s\bar{Z}(\mathbb{E}(Z^{'}Z))^{-1}I_2\Psi_l J^{'}_{r-l}\phi\sigma^2\\
&\quad\quad-\Omega J_{r-l}\Psi'_l I'_2(\mathbb{E}(Z^{'}Z))^{-1}\bar{Z}'D^{r'}D^s\bar{\Upsilon}HC^{*'}\Psi_j J'_{s-j}\phi\sigma^2.
\end{aligned}
$$

The final expression of the second part of equation (A.15) can be written as:

$$
\begin{aligned}
&-\mathbb{E}\{H\Lambda^{**'}\tilde{V}'\bar{Z}(\mathbb{E}(Z'Z))^{-1}I_2\tilde{R}'\tilde{R}C^*H\bar{\Upsilon}'\tilde{u}_1\} \qquad \text{(A.17)}\\
&= -H\sum_{l=1}^{p}\sum_{j=1}^{p}\sum_{r=l,j}^{T-1}(T-r)\left(tr\left\{\bar{Z}(\mathbb{E}(Z'Z))^{-1}I_2\Psi_l J_{r-l}'\Omega J_{s-j}\Psi_j'C^*H\bar{\Upsilon}'\right\}.I\right)\vartheta\\
&- H\Lambda^{**}\sum_{l=1}^{p}\sum_{r=l}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1}\Omega J_{s-j}\Psi_j'C^*H\bar{\Upsilon}'D^{r'}D^s\bar{Z}(\mathbb{E}(Z'Z))^{-1}I_2\Psi_l J_{r-l}'\phi\sigma^2\\
&- H\Lambda^{**}\sum_{l=1}^{p}\sum_{r=l}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1}\Omega J_{r-l}\Psi_l' I_2'(\mathbb{E}(Z'Z))^{-1}\bar{Z}'D^{r'}D^s\bar{\Upsilon}HC^{*'}\Psi_j J_{s-j}'\phi\sigma^2.
\end{aligned}
$$

(8)

$$
-\mathbb{E}\{HC^{*'}\tilde{R}'\bar{Z}(\mathbb{E}(Z'Z))^{-1}\tilde{W}^{*'}\tilde{V}\Lambda^{**}H\bar{\Upsilon}'\tilde{u}_1\} - \mathbb{E}\{HC^{*'}\tilde{R}'\tilde{R}I_2'(\mathbb{E}(Z'Z))^{-1}\bar{Z}\tilde{V}\Lambda^{**}H\bar{\Upsilon}'\tilde{u}_1\}. \qquad \text{(A.18)}
$$

Equation (A.18) can be written as the sum of two parts ($8a$) and ($8b$):

$$
\begin{aligned}
&(8a) \qquad \text{(A.19)}\\
&-\mathbb{E}\{HC^{*'}\tilde{R}'\bar{Z}(\mathbb{E}(Z'Z))^{-1}\tilde{W}^{*'}\tilde{V}\Lambda^{**}H\bar{\Upsilon}'\tilde{u}_1\}\\
&= -\mathbb{E}\left\{H\sum_{i=1}^{p}\sum_{t=i}^{T-1}C^{*'}\Psi_i J_{t-i}'\tilde{V}'D^{t'}\bar{Z}(\mathbb{E}(Z'Z))^{-1}\begin{bmatrix}\sum_{i=1}^{p}\sum_{s=j}^{T-1}\Psi_j J_{s-j}'\tilde{V}'D^{s'}\\ 0'\end{bmatrix}\right.\\
&\qquad\qquad \left.\times\tilde{V}\Lambda^{**}H\bar{\Upsilon}'\tilde{u}_1\right\}.
\end{aligned}
$$

Here,

$$
\begin{aligned}
&\mathbb{E}\left\{\tilde{V}'D^{t'}\bar{Z}(\mathbb{E}(Z'Z))^{-1}\begin{bmatrix}\Psi_j J_{s-j}'\\ 0\end{bmatrix}\tilde{V}'D^{s'}\tilde{V}\Lambda^{**}H\bar{\Upsilon}'\tilde{u}_1\right\}\\
&= \mathbb{E}\left\{\phi\tilde{u}_1'D^{t'}\bar{Z}(\mathbb{E}(Z'Z))^{-1}I_2\Psi_j J_{s-j}'\phi\tilde{u}_1'D^{s'}\tilde{u}_1\phi'\Lambda^{**}H\bar{\Upsilon}'\tilde{u}_1\right\}
\end{aligned}
$$

$$
\begin{aligned}
&+\mathbb{E}\left\{\phi\tilde{u}_1^{'}D^{t'}\bar{Z}(\mathbb{E}(Z^{'}Z))^{-1}I_2\Psi_j J_{s-j}^{'}S^{'}D^{s'}S\Lambda^{**}H\bar{\Upsilon}^{'}\tilde{u}_1\right\}\\
&+\mathbb{E}\left\{S^{\prime}D^{t'}\bar{Z}(\mathbb{E}(Z^{'}Z))^{-1}I_2\Psi_j J_{s-j}^{'}S^{'}D^{s'}\tilde{u}_1\phi^{'}\Lambda^{**}H\bar{\Upsilon}^{'}\tilde{u}_1\right\}\\
&+\mathbb{E}\left\{S^{'}D^{t'}\bar{Z}(\mathbb{E}(Z^{'}Z))^{-1}I_2\Psi_j J_{s-j}^{'}\phi\tilde{u}_1^{'}D^{s'}S\Lambda^{**}H\bar{\Upsilon}^{'}\tilde{u}_1\right\}\\
=\;&\sigma^4\phi\phi^{'}J_{s-j}\Psi_j^{'}I_2^{'}(\mathbb{E}(Z^{'}Z))^{-1}\bar{Z}^{'}D^tD^{s'}\bar{\Upsilon}H\Lambda^{**'}\phi\\
&+\sigma^4\phi\phi^{'}\Lambda^{**}H\bar{\Upsilon}D^{s'}D^{t'}\bar{Z}(\mathbb{E}(Z^{'}Z))^{-1}I_2\Psi_j J_{s-j}^{'}\phi\\
&+0\\
&+\sigma^2\Omega J_{s-j}\Psi_j^{'}I_2^{'}(\mathbb{E}(Z^{'}Z))^{-1}\bar{Z}^{'}D^tD^{s'}\bar{\Upsilon}H\Lambda^{**'}\phi\\
&-\sigma^4\phi\phi^{'}J_{s-j}\Psi_j^{'}I_2^{'}(\mathbb{E}(Z^{'}Z))^{-1}\bar{Z}^{'}D^tD^{s'}\bar{\Upsilon}H\Lambda^{**'}\phi\\
&+\sigma^2\Omega\Lambda^{**}H\bar{\Upsilon}D^{s'}D^{t'}\bar{Z}(\mathbb{E}(Z^{'}Z))^{-1}I_2\Psi_j J_{s-j}^{'}\phi\\
&-\sigma^4\phi\phi^{'}\Lambda^{**}H\bar{\Upsilon}D^{s'}D^{t'}\bar{Z}(\mathbb{E}(Z^{'}Z))^{-1}I_2\Psi_j J_{s-j}^{'}\phi\\
=\;&\Omega J_{s-j}\Psi_j^{'}I_2^{'}(\mathbb{E}(Z^{'}Z))^{-1}\bar{Z}^{'}D^tD^{s'}\bar{\Upsilon}H\Lambda^{**'}(\sigma^2\phi)\\
&+\Omega\Lambda^{**}H\bar{\Upsilon}^{'}D^{s'}D^{t'}\bar{Z}(\mathbb{E}(Z^{'}Z))^{-1}I_2\Psi_j J_{s-j}^{'}(\sigma^2\phi).
\end{aligned}
$$

Using Lemma 3 and 4.

The final result for equation (A.19) is:

$$
\begin{aligned}
&-\mathbb{E}\{HC^{*'}\tilde{R}^{'}\bar{Z}(\mathbb{E}(Z^{'}Z))^{-1}\tilde{W}^{*'}\tilde{V}\Lambda^{**}H\bar{\Upsilon}^{'}\tilde{u}_1\} \qquad (A.20)\\
&=-H\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1}C^{*'}\Psi_iJ_{t-i}^{'}\Omega J_{s-j}\Psi_j^{'}I_2^{'}(\mathbb{E}(Z^{'}Z))^{-1}\bar{Z}^{'}D^tD^{s'}\bar{\Upsilon}H\Lambda^{**'}(\sigma^2\phi)\\
&\quad-H\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1}C^{*'}\Psi_iJ_{t-i}^{'}\Omega\Lambda^{**}H\bar{\Upsilon}D^{r'}D^{t'}\bar{Z}(\mathbb{E}(Z^{'}Z))^{-1}I_2\Psi_jJ_{s-j}^{'}(\sigma^2\phi).
\end{aligned}
$$

(8b)

$$
\begin{aligned}
&-\mathbb{E}\{HC^{*'}\tilde{R}^{'}\tilde{R}I_2^{'}(\mathbb{E}(Z^{'}Z))^{-1}\bar{Z}^{'}\tilde{V}\Lambda^{**}H\bar{\Upsilon}^{'}\tilde{u}_1\} \qquad (A.21)\\
&=-\mathbb{E}\left\{H\sum_{i=1}^{p}\sum_{t=i}^{T-1}C^{*'}\Psi_iJ_{t-i}^{'}\tilde{V}^{'}D^{t'}\sum_{j=1}^{p}\sum_{s=j}^{T-1}D^s\tilde{V}J_{s-j}\Psi_j^{'}I_2^{'}(\mathbb{E}(Z^{'}Z))^{-1}\bar{Z}^{'}\tilde{V}\Lambda^{**}H\bar{\Upsilon}^{'}\tilde{u}_1\right\}.
\end{aligned}
$$

Here,

$$
\begin{aligned}
&\mathbb{E}\Big\{\tilde{V}^{'}D^{t^{'}}D^{s}\tilde{V}J_{s-j}\Psi_{j}^{'}I_{2}^{'}(\mathbb{E}(Z^{'}Z))^{-1}\bar{Z}^{'}\tilde{V}\Lambda^{**}H\bar{\Upsilon}^{'}\tilde{u}_{1}\Big\}\\
&\quad=\mathbb{E}\{\phi\tilde{u}_{1}^{'}D^{t^{'}}D^{s}\tilde{u}_{1}\phi^{'}J_{s-j}\Psi_{j}^{'}I_{2}^{'}(\mathbb{E}(Z^{'}Z))^{-1}\bar{Z}^{'}\tilde{u}_{1}\phi^{'}\Lambda^{**}H\bar{\Upsilon}^{'}\tilde{u}_{1}\}\\
&\qquad+\mathbb{E}\{\phi\tilde{u}_{1}^{'}D^{t^{'}}D^{s}SJ_{s-j}\Psi_{j}^{'}I_{2}^{'}(\mathbb{E}(Z^{'}Z))^{-1}\bar{Z}^{'}S\Lambda^{**}H\bar{\Upsilon}^{'}\tilde{u}_{1}\}\\
&\qquad+\mathbb{E}\{S^{'}D^{t^{'}}D^{s}SJ_{s-j}\Psi_{j}^{'}I_{2}^{'}(\mathbb{E}(Z^{'}Z))^{-1}\bar{Z}^{'}\tilde{u}_{1}\phi^{'}\Lambda^{**}H\bar{\Upsilon}^{'}\tilde{u}_{1}\}\\
&\qquad+\mathbb{E}\{S^{'}D^{t^{'}}D^{s}\tilde{u}_{1}\phi^{'}J_{s-j}\Psi_{j}^{'}I_{2}^{'}(\mathbb{E}(Z^{'}Z))^{-1}\bar{Z}^{'}S\Lambda^{**}H\bar{\Upsilon}^{'}\tilde{u}_{1}\}\\
&\quad=\sigma^{4}\phi tr\{\frac{1}{2}(D^{t^{'}}D^{s}+D^{s^{'}}D^{t})\}tr\{\bar{\Upsilon}H\Lambda^{**^{'}}\phi\phi^{'}J_{s-j}\Psi_{j}^{'}I_{2}^{'}(\mathbb{E}(Z^{'}Z))^{-1}\bar{Z}^{'}\}\\
&\qquad+2\sigma^{4}\phi tr\{\frac{1}{2}(D^{t^{'}}D^{s}+D^{s^{'}}D^{t})\}\bar{\Upsilon}H\Lambda^{**^{'}}\phi\phi^{'}J_{s-j}\Psi_{j}^{'}I_{2}^{'}(\mathbb{E}(Z^{'}Z))^{-1}\bar{Z}^{'}\\
&\qquad+\sigma^{2}\phi tr\{D^{t^{'}}D^{s}\bar{Z}(\mathbb{E}(Z^{'}Z))^{-1}I_{2}\Psi_{j}J_{s-j}^{'}\Omega\Lambda^{**}H\bar{\Upsilon}^{'}\}\\
&\qquad-\sigma^{4}\phi tr\{D^{t^{'}}D^{s}\bar{Z}(\mathbb{E}(Z^{'}Z))^{-1}I_{2}\Psi_{j}J_{s-j}^{'}\phi\phi^{'}\Lambda^{**}H\bar{\Upsilon}^{'}\}\\
&\qquad+\sigma^{2}tr\{D^{t^{'}}D^{r}\}\Omega J_{s-j}\Psi_{j}^{'}I_{2}^{'}(\mathbb{E}(Z^{'}Z))^{-1}\bar{Z}^{'}\bar{\Upsilon}H\Lambda^{**^{'}}\phi\\
&\qquad-\sigma^{4}tr\{D^{t^{'}}D^{r}\}\phi\phi^{'}J_{s-j}\Psi_{j}^{'}I_{2}^{'}(\mathbb{E}(Z^{'}Z))^{-1}\bar{Z}^{'}\bar{\Upsilon}H\Lambda^{**^{'}}\phi\\
&\qquad+\sigma^{2}\Omega\Lambda^{**}H\bar{\Upsilon}^{'}D^{s^{'}}D^{t}\bar{Z}(\mathbb{E}(Z^{'}Z))^{-1}I_{2}\Psi_{j}J_{s-j}^{'}\phi\\
&\qquad-\sigma^{4}\phi\phi^{'}\Lambda^{**}H\bar{\Upsilon}^{'}D^{s^{'}}D^{t}\bar{Z}(\mathbb{E}(Z^{'}Z))^{-1}I_{2}\Psi_{j}J_{s-j}^{'}\phi\\
&\quad=(tr\{D^{t^{'}}D^{s}\bar{Z}(\mathbb{E}(Z^{'}Z))^{-1}I_{2}\Psi_{j}J_{s-j}^{'}\Omega\Lambda^{**}H\bar{\Upsilon}^{'}\}.I)(\sigma^{2}\phi)\\
&\quad+\Omega J_{s-j}\Psi_{j}^{'}I_{2}^{'}(\mathbb{E}(Z^{'}Z))^{-1}\bar{Z}^{'}\bar{\Upsilon}H(tr\{D^{t^{'}}D^{r}\}.I)\Lambda^{**^{'}}(\sigma^{2}\phi)\\
&\quad+\Omega\Lambda^{**}H\bar{\Upsilon}^{'}D^{s^{'}}D^{t}\bar{Z}(\mathbb{E}(Z^{'}Z))^{-1}I_{2}\Psi_{j}J_{s-j}^{'}(\sigma^{2}\phi).
\end{aligned}
$$

Using Lemma 4.

The final result for equation (A.21) is

$$
\begin{aligned}
&-\mathbb{E}\{HC^{*^{'}}\tilde{R}^{'}\tilde{R}I_{2}^{'}(\mathbb{E}(Z^{'}Z))^{-1}\bar{Z}^{'}\tilde{V}\Lambda^{**}H\bar{\Upsilon}^{'}\tilde{u}_{1}\} \qquad \text{(A.22)}\\
&\quad=-H\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1}C^{*^{'}}\Psi_{i}J_{t-i}^{'}(tr\{D^{t^{'}}D^{s}\bar{Z}(\mathbb{E}(Z^{'}Z))^{-1}I_{2}e_{j}J_{s-j}^{'}\Omega\Lambda^{**}H\bar{\Upsilon}^{'}\}.I)(\sigma^{2}\phi)
\end{aligned}
$$

$$-H\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1}C^{*'}\Psi_i J_{t-i}^{'}\Omega J_{s-j}\Psi_j^{'}I_2^{'}(\mathbb{E}(Z^{'}Z))^{-1}\bar{Z}^{'}\bar{\Upsilon}H(tr\{D^{t'}D^r\}.I)\Lambda^{**'}(\sigma^2\phi)$$
$$-H\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1}C^{*'}\Psi_i J_{t-i}^{'}\Omega\Lambda^{**}H\bar{\Upsilon}^{'}D^{s'}D^t\bar{Z}(\mathbb{E}(Z^{'}Z))^{-1}I_2\Psi_j J_{s-j}^{'}(\sigma^2\phi).$$

Therefore, by combining equations (A.6), (A.9), (A.10), (A.12), (A.13), (A.14), (A.16), (A.17),(A.20), (A.22), we can get the final expression for (iv).

$$
\begin{aligned}
(v)\qquad -\mathbb{E}\{HJ_1^*H\Delta_2^{'}\tilde{u}_1\} = & -\mathbb{E}\{H\bar{\Upsilon}^{'}\bar{Z}[\mathbb{E}(Z'Z)]^{-1}\tilde{W}^{*'}\tilde{V}\Lambda^{**}HC^{*'}\tilde{R}^{'}\tilde{u}_1\}\\
& -\mathbb{E}\{H\bar{\Upsilon}^{'}\tilde{R}I_2^{'}[\mathbb{E}(Z'Z)]^{-1}\bar{Z}'\tilde{V}\Lambda^{**}HC^{*'}\tilde{R}^{'}\tilde{u}_1\}\\
& -\mathbb{E}\{H\Lambda^{**'}\tilde{V}^{'}\tilde{W}^{*}[\mathbb{E}(Z'Z)]^{-1}\bar{Z}^{'}\bar{\Upsilon}HC^{*'}\tilde{R}^{'}\tilde{u}_1\}\\
& -\mathbb{E}\{H\Lambda^{**'}\tilde{V}^{'}\bar{Z}[\mathbb{E}(Z'Z)]^{-1}I_2\tilde{R}'\bar{\Upsilon}HC^{*'}\tilde{R}^{'}\tilde{u}_1\}\\
& -\mathbb{E}\{HC^{*'}\tilde{R}^{'}\tilde{R}I_2^{'}[\mathbb{E}(Z'Z)]^{-1}\tilde{W}^{*'}\tilde{V}\Lambda^{**}HC^{*'}\tilde{R}^{'}\tilde{u}_1\}\\
& -\mathbb{E}\{HC^{*'}\tilde{R}^{'}\bar{Z}[\mathbb{E}(Z'Z)]^{-1}\bar{Z}^{'}\tilde{V}\Lambda^{**}HC^{*'}\tilde{R}^{'}\tilde{u}_1\}\\
& -\mathbb{E}\{H\Lambda^{**'}\tilde{V}^{'}\tilde{W}^{*}(\mathbb{E}(Z^{'}Z))^{-1}I_2\tilde{R}^{'}\tilde{R}C^{*}HC^{*'}\tilde{R}^{'}\tilde{u}_1\}\\
& -\mathbb{E}\{H\Lambda^{**'}\tilde{V}^{'}\bar{Z}(\mathbb{E}(Z^{'}Z))^{-1}\bar{Z}^{'}\tilde{R}C^{*}HC^{*'}\tilde{R}^{'}\tilde{u}_1\}\\
& -\mathbb{E}\{HC^{*'}\tilde{R}'\tilde{R}C^{*}HC^{*'}\tilde{R}^{'}\tilde{u}_1\},
\end{aligned}
$$

where the definition of Δ_1 and Δ_2 is from equation 2.10.

Then v can be obtained from the sum of $(1') - (9')$ below:

$(1')$

$$
\begin{aligned}
& -\mathbb{E}\{H\bar{\Upsilon}^{'}\bar{Z}[\mathbb{E}(Z'Z)]^{-1}\tilde{W}^{*'}\tilde{V}\Lambda^{**}HC^{*'}\tilde{R}^{'}\tilde{u}_1\} \qquad \text{(A.23)}\\
& = -\mathbb{E}\left\{H\bar{\Upsilon}^{'}\bar{Z}[\mathbb{E}(Z'Z)]^{-1}\begin{bmatrix}\sum_{i=1}^{p}\sum_{t=i}^{T-1}\Psi_i J_{t-i}^{'}\\ 0\end{bmatrix}\tilde{V}^{'}D^{t'}\tilde{V}\Lambda^{**}\right.
\end{aligned}
$$

$$\left.\times H \sum_{j=1}^{p} \sum_{s=j}^{T-1} C^{*'} \Psi_j J_{s-j}^{'} \tilde{V}^{'} D^{s'} \tilde{u}_1 \right\}.$$

Then, equation (A.23) can be calculated from:

$$\begin{aligned}
&\mathbb{E}\left\{\tilde{V}^{'} D^{t'} \tilde{V} \Lambda^{**} H C^{*'} \Psi_j J_{s-j}^{'} \tilde{V}^{'} D^{s'} \tilde{u}_1\right\} \\
&\quad = \mathbb{E}\{\phi \tilde{u}_1^{'} D^{t'} \tilde{u}_1 \phi^{'} \Lambda^{**} H C^{*'} \Psi_j J_{s-j}^{'} \phi \tilde{u}_1^{'} D^{s'} \tilde{u}_1\} \\
&\qquad + \mathbb{E}\{\phi \tilde{u}_1^{'} D^{t'} S \Lambda^{**} H C^{*'} \Psi_j J_{s-j}^{'} S^{'} D^{s'} \tilde{u}_1\} \\
&\qquad + \mathbb{E}\{S^{'} D^{t'} S \Lambda^{**} H C^{*'} \Psi_j J_{s-j}^{'} \phi \tilde{u}_1^{'} D^{s'} \tilde{u}_1\} \\
&\qquad + \mathbb{E}\{S^{'} D^{t'} \tilde{u}_1 \phi^{'} \Lambda^{**} H C^{*'} \Psi_j J_{s-j}^{'} S^{'} D^{s'} \tilde{u}_1\} \\
&\quad = \sigma^4 \phi \phi^{'} \Lambda^{**} H C^{*'} \Psi_j J_{s-j}^{'} \phi tr\{(D^t + D^{t'}) D^{s'}\} \\
&\qquad + 0 \\
&\qquad + 0 \\
&\qquad + \sigma^2 tr\{D^{t'} D^s\} \Omega J_{s-j} \Psi_j^{'} C^* H \Lambda^{**'} \phi \\
&\qquad - \sigma^4 tr\{D^{t'} D^s\} \phi \phi^{'} J_{s-j} \Psi_j^{'} C^* H \Lambda^{**'} \phi \\
&\quad = (tr\{D^{t'} D^s\}.I) \Omega J_{s-j} \Psi_j^{'} C^* H \Lambda^{**'} (\phi \sigma^2).
\end{aligned}$$

Then, the final result for equation (A.23) is:

$$\begin{aligned}
&-\mathbb{E}\{H \bar{\Upsilon}^{'} \bar{Z} [\mathbb{E}(Z'Z)]^{-1} \tilde{W}^{*'} \tilde{V} \Lambda^{**} H C^{*'} \tilde{R}^{'} \tilde{u}_1\} \\
&\quad = -H \bar{\Upsilon}^{'} \bar{Z} I_2 [\mathbb{E}(Z'Z)]^{-1} \sum_{i=1}^{p} \sum_{t=i}^{T-1} \sum_{j=1}^{p} \sum_{s=j}^{T-1} \Psi_i J_{t-i}^{'} (tr\{D^{t'} D^s\}.I) \Omega J_{s-j} \Psi_j^{'} C^* H \Lambda^{**'} (\phi \sigma^2).
\end{aligned} \tag{A.24}$$

(2′)

$$-\mathbb{E}\{H \bar{\Upsilon}^{'} \tilde{R} I_2^{\prime} [\mathbb{E}(Z'Z)]^{-1} \bar{Z}' \tilde{V} \Lambda^{**} H C^{*'} \tilde{R}^{'} \tilde{u}_1\} \tag{A.25}$$

$$= -\mathbb{E}\{H\bar{\Upsilon}' \sum_{i=1}^{p} \sum_{t=i}^{T-1} D^t \tilde{V} J_{t-i} \Psi_i' I_2' [\mathbb{E}(Z'Z)]^{-1} \bar{Z}' \tilde{V} \Lambda^{**} HC^{*'} \sum_{j=1}^{p} \sum_{s=j}^{T-1} \Psi_j J_{s-j} \tilde{V}' D^{t'} \tilde{u}_1\}.$$

Then, equation (A.25) can be calculated from:

$$\begin{aligned}
&-\mathbb{E}\{\tilde{V} J_{t-i} \Psi_i' I_2' [\mathbb{E}(Z'Z)]^{-1} \bar{Z}' \tilde{V} \Lambda^{**} HC^{*'} \Psi_j J_{s-j} \tilde{V}' D^{s'} \tilde{u}_1\} \\
&\quad = -\mathbb{E}\{u_1 \phi' J_{t-i} \Psi_i' I_2' [\mathbb{E}(Z'Z)]^{-1} \bar{Z}' u_1 \phi' \Lambda^{**} HC^{*'} \Psi_j J_{s-j} \phi u_1' D^{s'} \tilde{u}_1\} \\
&\qquad - \mathbb{E}\{S J_{t-i} \Psi_i' I_2' [\mathbb{E}(Z'Z)]^{-1} \bar{Z}' S \Lambda^{**} HC^{*'} \Psi_j J_{s-j} \phi u_1' D^{s'} \tilde{u}_1\} \\
&\qquad - \mathbb{E}\{S \phi' J_{t-i} \Psi_i' I_2' [\mathbb{E}(Z'Z)]^{-1} \bar{Z}' u_1 \phi' \Lambda^{**} HC^{*'} \Psi_j J_{s-j} S' D^{s'} \tilde{u}_1\} \\
&\qquad - \mathbb{E}\{u_1 \phi' J_{t-i} \Psi_i' I_2' [\mathbb{E}(Z'Z)]^{-1} \bar{Z}' S \Lambda^{**} HC^{*'} \Psi_j J_{s-j} S' D^{s'} \tilde{u}_1\} \\
&\quad = -D^s tr \left\{\Omega J_{t-i} \Psi_i' C^* H \Lambda^{**'}\right\} I \bar{Z} [\mathbb{E}(Z'Z)]^{-1} I_2 \Psi_j J_{s-j}' \phi \sigma^2 \\
&\qquad - D^{s'} \bar{Z} [\mathbb{E}(Z'Z)]^{-1} I_2 \Psi_j J_{s-j}' \Omega J_{t-i} \Psi_i' C^* H \Lambda^{**'} \phi \sigma^2.
\end{aligned}$$

Then, the final expression of equation (A.25) is:

$$\begin{aligned}
&-\mathbb{E}\{H\bar{\Upsilon}' \tilde{R} I_2' [\mathbb{E}(Z'Z)]^{-1} \bar{Z}' \tilde{V} \Lambda^{**} HC^{*'} \tilde{R}' \tilde{u}_1\} \qquad\qquad (A.26) \\
&\quad = -H\bar{\Upsilon}' \sum_{i=1}^{p} \sum_{t=i}^{T-1} \sum_{j=1}^{p} \sum_{s=j}^{T-1} D^t D^s tr \left\{\Omega J_{t-i} \Psi_i' C^* H \Lambda^{**'}\right\} I \bar{Z} [\mathbb{E}(Z'Z)]^{-1} I_2 \Psi_j J_{s-j}' \phi \sigma^2 \\
&\qquad - H\bar{\Upsilon}' \sum_{i=1}^{p} \sum_{t=i}^{T-1} \sum_{j=1}^{p} \sum_{s=j}^{T-1} D^t D^{s'} \bar{Z} [\mathbb{E}(Z'Z)]^{-1} I_2 \Psi_j J_{s-j}' \Omega J_{t-i} \Psi_i' C^* H \Lambda^{**'} \phi \sigma^2.
\end{aligned}$$

(3′)

$$\begin{aligned}
&-\mathbb{E}\{H\Lambda^{**'} \tilde{V}' \tilde{W}^* [\mathbb{E}(Z'Z)]^{-1} \bar{Z}' \bar{\Upsilon} HC^{*'} \tilde{R}' \tilde{u}_1\} \qquad\qquad (A.27) \\
&\quad = -\mathbb{E}\{H\Lambda^{**'} \tilde{V}' \left[\sum_{i=1}^{p} \sum_{t=i}^{T-1} D^t \tilde{V} J_{t-i} \Psi_i' : 0\right] [\mathbb{E}(Z'Z)]^{-1} \bar{Z}' \bar{\Upsilon} H \sum_{j=1}^{p} \sum_{s=j}^{T-1} C^{*'} \Psi_j J_{s-j}' \tilde{V}' D^{s'} \tilde{u}_1\}.
\end{aligned}$$

Here,

$$\begin{aligned}
&\mathbb{E}\{\tilde{V}^{'}D^{t}\tilde{V}J_{t-i}\Psi_{i}^{'}I_{2}^{'}[\mathbb{E}(Z'Z)]^{-1}\bar{Z}^{'}\bar{\Upsilon}HC^{*^{'}}\Psi_{j}J_{s-j}^{'}\tilde{V}^{'}D^{s^{'}}\tilde{u}_{1}\} \\
&\quad = \mathbb{E}\{\phi\tilde{u}_{1}^{'}D^{t}\tilde{u}_{1}\phi^{'}J_{t-i}\Psi_{i}^{'}I_{2}^{'}[\mathbb{E}(Z'Z)]^{-1}\bar{Z}^{'}\bar{\Upsilon}HC^{*^{'}}\Psi_{j}J_{s-j}^{'}\phi\tilde{u}_{1}^{'}D^{s^{'}}\tilde{u}_{1}\} \\
&\qquad + \mathbb{E}\{\phi\tilde{u}_{1}^{'}D^{t}SJ_{t-i}\Psi_{i}^{'}I_{2}^{'}[\mathbb{E}(Z'Z)]^{-1}\bar{Z}^{'}\bar{\Upsilon}HC^{*^{'}}\Psi_{j}J_{s-j}^{'}S^{'}D^{s^{'}}\tilde{u}_{1}\} \\
&\qquad + \mathbb{E}\{S^{'}D^{t}SJ_{t-i}\Psi_{i}^{'}I_{2}^{'}[\mathbb{E}(Z'Z)]^{-1}\bar{Z}^{'}\bar{\Upsilon}HC^{*^{'}}\Psi_{j}J_{s-j}^{'}\phi\tilde{u}_{1}^{'}D^{s^{'}}\tilde{u}_{1}\} \\
&\qquad + \mathbb{E}\{S^{'}D^{t}\tilde{u}_{1}\phi^{'}J_{t-i}\Psi_{i}^{'}I_{2}^{'}[\mathbb{E}(Z'Z)]^{-1}\bar{Z}^{'}\bar{\Upsilon}HC^{*^{'}}\Psi_{j}J_{s-j}^{'}S^{'}D^{s^{'}}\tilde{u}_{1}\} \\
&\quad = \sigma^{4}\phi\phi^{'}J_{t-i}\Psi_{i}^{'}I_{2}^{'}[\mathbb{E}(Z'Z)]^{-1}\bar{Z}^{'}\bar{\Upsilon}HC^{*^{'}}\Psi_{j}J_{s-j}^{'}\phi tr\{\frac{1}{2}(D^{t}+D^{t^{'}})D^{r^{'}}\}+ \\
&\qquad + 0 \\
&\qquad + 0 \\
&\qquad + \sigma^{2}\phi tr\{D^{t}D^{s^{'}}\}tr\{\Omega J_{t-i}\Psi_{i}^{'}I_{2}^{'}[\mathbb{E}(Z'Z)]^{-1}\bar{Z}^{'}\bar{\Upsilon}HC^{*^{'}}\Psi_{j}J_{s-j}^{'}\} \\
&\qquad - \sigma^{4}\phi tr\{D^{t}D^{s^{'}}\}tr\{\phi\phi^{'}J_{t-i}\Psi_{i}^{'}I_{2}^{'}[\mathbb{E}(Z'Z)]^{-1}\bar{Z}^{'}\bar{\Upsilon}HC^{*^{'}}\Psi_{j}J_{s-j}^{'}\} \\
&\quad = tr\{D^{t}D^{s^{'}}\}tr\{\Omega J_{t-i}\Psi_{i}^{'}I_{2}^{'}[\mathbb{E}(Z'Z)]^{-1}\bar{Z}^{'}\bar{\Upsilon}HC^{*^{'}}\Psi_{j}J_{s-j}^{'}\}(\sigma^{2}\phi).
\end{aligned}$$

Then, the final result for equation (A.27) is:

$$\begin{aligned}
&-\mathbb{E}\{H\Lambda^{**^{'}}\tilde{V}^{'}\tilde{W}^{*}[\mathbb{E}(Z'Z)]^{-1}\bar{Z}^{'}\bar{\Upsilon}HC^{*^{'}}\tilde{R}^{'}\tilde{u}_{1}\} \\
&\quad = -H\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1}tr\{D^{t}D^{r^{'}}\}(tr\{\Omega J_{t-i}\Psi_{i}^{'}I_{2}^{'}[\mathbb{E}(Z'Z)]^{-1}\bar{Z}^{'}\bar{\Upsilon}HC^{*^{'}}\Psi_{j}J_{s-j}^{'}\}.I) \\
&\qquad\qquad \times\Lambda^{**^{'}}(\sigma^{2}\phi).
\end{aligned} \tag{A.28}$$

(4′)

$$\begin{aligned}
&-\mathbb{E}\{H\Lambda^{**^{'}}\tilde{V}^{'}\bar{Z}[\mathbb{E}(Z'Z)]^{-1}I_{2}\tilde{R}'\bar{\Upsilon}HC^{*^{'}}\tilde{R}^{'}\tilde{u}_{1}\} \\
&\quad = -\mathbb{E}\{H\Lambda^{**^{'}}\tilde{V}^{'}\bar{Z}[\mathbb{E}(Z'Z)]^{-1}I_{2}\sum_{i=1}^{p}\sum_{t=i}^{T-1}\Psi_{i}J_{t-i}'\tilde{V}'D^{t^{'}}\bar{\Upsilon}HC^{*^{'}}\sum_{j=1}^{p}\sum_{s=j}^{T-1}\Psi_{j}J_{s-j}'\tilde{V}'D^{s^{'}}\tilde{u}_{1}\}.
\end{aligned} \tag{A.29}$$

Then, equation (A.29) can be calculated from:

$$
\begin{aligned}
&-\mathbb{E}\{\tilde{V}^{'}\bar{Z}[\mathbb{E}(Z'Z)]^{-1}I_2\Psi_iJ_{t-i}'\tilde{V}'D^{t'}\bar{\Upsilon}HC^{*'}\Psi_jJ_{s-j}'\tilde{V}'D^{s'}\tilde{u}_1\}\\
&\quad=-\mathbb{E}\{\phi\tilde{u}_1^{'}\bar{Z}[\mathbb{E}(Z'Z)]^{-1}I_2\Psi_iJ_{t-i}'\phi\tilde{u}_1^{'}D^{t'}\bar{\Upsilon}HC^{*'}\Psi_jJ_{s-j}'\phi\tilde{u}_1^{'}D^{s'}\tilde{u}_1\}\\
&\qquad-\mathbb{E}\{S^{'}\bar{Z}[\mathbb{E}(Z'Z)]^{-1}I_2\Psi_iJ_{t-i}'S'D^{t'}\bar{\Upsilon}HC^{*'}\Psi_jJ_{s-j}'\phi\tilde{u}_1^{'}D^{s'}\tilde{u}_1\}\\
&\qquad-\mathbb{E}\{S^{'}\bar{Z}[\mathbb{E}(Z'Z)]^{-1}I_2\Psi_iJ_{t-i}'\phi\tilde{u}_1^{'}D^{t'}\bar{\Upsilon}HC^{*'}\Psi_jJ_{s-j}'S'D^{s'}\tilde{u}_1\}\\
&\qquad-\mathbb{E}\{\phi\tilde{u}_1^{'}\bar{Z}[\mathbb{E}(Z'Z)]^{-1}I_2\Psi_iJ_{t-i}'S'D^{t'}\bar{\Upsilon}HC^{*'}\Psi_jJ_{s-j}'S'D^{s'}\tilde{u}_1\}\\
&\quad=-\sigma^4\phi\phi'J_{s-j}\Psi_j'C^*H\bar{\Upsilon}'D^tD^s\bar{Z}[\mathbb{E}(Z'Z)]^{-1}I_2\Psi_iJ_{t-i}'\phi\\
&\qquad-\phi\sigma^2\,\mathrm{tr}\left\{\bar{Z}[\mathbb{E}(Z'Z)]^{-1}I_2\Psi_iJ_{t-i}'\sigma^2\phi\phi'J_{s-j}\Psi_j'C^*H\bar{\Upsilon}'D^tD^{s'}\right\}\\
&\qquad-0\\
&\qquad-(\Omega-\sigma^2\phi\phi')J_{s-j}\Psi_j'C^*H\bar{\Upsilon}'D^tD^s\bar{Z}[\mathbb{E}(Z'Z)]^{-1}I_2\Psi_iJ_{t-i}'\phi\sigma^2\\
&\qquad-tr\left\{\bar{Z}[\mathbb{E}(Z'Z)]^{-1}I_2\Psi_iJ_{t-i}'(\Omega-\sigma^2\phi\phi')J_{s-j}\Psi_j'C^*H\bar{\Upsilon}'D^tD^{s'}\right\}\phi\sigma^2\\
&\quad=-tr\left\{\bar{Z}[\mathbb{E}(Z'Z)]^{-1}I_2\Psi_iJ_{t-i}'\Omega J_{s-j}\Psi_j'C^*H\bar{\Upsilon}'D^tD^{s'}\right\}\phi\sigma^2\\
&\qquad-\Omega J_{s-j}\Psi_j'C^*H\bar{\Upsilon}'D^tD^s\bar{Z}[\mathbb{E}(Z'Z)]^{-1}I_2\Psi_iJ_{t-i}'\phi\sigma^2.
\end{aligned}
$$

Using Lemma 3 and 4.

The final expression for equation (A.29) is:

$$
\begin{aligned}
&-\mathbb{E}\{H\Lambda^{**'}\tilde{V}^{'}\bar{Z}[\mathbb{E}(Z'Z)]^{-1}I_2\tilde{R}'\bar{\Upsilon}HC^{*'}\tilde{R}^{'}\tilde{u}_1\} \qquad \text{(A.30)}\\
&=-\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1}H\Lambda^{**'}tr\left\{\bar{Z}[\mathbb{E}(Z'Z)]^{-1}I_2\Psi_iJ_{t-i}'\Omega J_{s-j}\Psi_j'C^*H\bar{\Upsilon}'D^tD^{s'}\right\}\phi\sigma^2\\
&\quad-\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1}H\Lambda^{**'}\Omega J_{s-j}\Psi_j'C^*H\bar{\Upsilon}'D^tD^s\bar{Z}[\mathbb{E}(Z'Z)]^{-1}I_2\Psi_iJ_{t-i}'\phi\sigma^2.
\end{aligned}
$$

(5′)

$$-\mathbb{E}\{HC^{*'}\tilde{R}^{'}\tilde{R}I_2^{'}[\mathbb{E}(Z'Z)]^{-1}\tilde{W}^{*'}\tilde{V}\Lambda^{**}HC^{*'}\tilde{R}^{'}\tilde{u}_1\} \qquad \text{(A.31)}$$

$$= -\mathbb{E}\left\{ H\mathbb{E}(\tilde{R}'\tilde{R})I_2^{'}[\mathbb{E}(Z'Z)]^{-1\ 1} \right.$$

$$\left. \times \begin{bmatrix} \sum_{l=1}^{p}\sum_{r=l}^{T-1}\Psi_l J_{r-l}^{'}\tilde{V}^{'}D^{r'} \\ 0 \end{bmatrix} \tilde{V}\Lambda^{**}H\sum_{b=1}^{p}\sum_{h=b}^{T-1}C^{*'}\Psi_b J_{h-b}^{'}\tilde{V}^{'}D^{t'}\tilde{u}_1 \right\}$$

Note:

1. $\tilde{R}'\tilde{R} = \mathbb{E}(\tilde{R}'\tilde{R}) + (\tilde{R}'\tilde{R} - \mathbb{E}(\tilde{R}'\tilde{R})) \equiv \mathbb{E}(\tilde{R}'\tilde{R}) + O_p(T^{1/2})$, where

$$\mathbb{E}(\tilde{R}'\tilde{R}) = \sum_{i=1}^{p}\sum_{t=i}^{T-1}\Psi_i J_{t-i}^{'}\tilde{V}^{'}D^{t'}\sum_{j=1}^{p}\sum_{s=j}^{T-1}D^s\tilde{V}J_{s-j}\Psi_j = \sum_{i=1}^{p}\sum_{t=i}^{T-1}(T-t)\Psi_i J_{t-i}^{'}\Omega J_{t-i}\Psi_i^{'})$$

. In the following calculation I can replace $\tilde{R}'\tilde{R}$ with $\mathbb{E}(\tilde{R}'\tilde{R})$ to the order of the approximation.

Hence, equation A.31 can be expressed as:

$$-\mathbb{E}\{HC^{*'}\tilde{R}^{'}\tilde{R}I_2^{'}[\mathbb{E}(Z'Z)]^{-1}\tilde{W}^{*'}\tilde{V}\Lambda^{**}HC^{*'}\tilde{R}^{'}\tilde{u}_1\} \tag{A.32}$$

$$= -\mathbb{E}\left\{ \left(H\sum_{i=1}^{p}\sum_{t=i}^{T-1}(T-t)C^{*'}\Psi_i J_{t-i}^{'}\Omega J_{t-i}\Psi_i^{'} \right) I_2^{'}[\mathbb{E}(Z'Z)]^{-1} \begin{bmatrix} \sum_{l=1}^{p}\sum_{r=l}^{T-1}\Psi_l J_{r-l}^{'}\tilde{V}^{'}D^{r'} \\ 0 \end{bmatrix} \right.$$

$$\left. \times \tilde{V}\Lambda^{**}H\sum_{b=1}^{p}\sum_{h=b}^{T-1}C^{*'}\Psi_b J_{h-b}^{'}\tilde{V}^{'}D^{h'}\tilde{u}_1 \right\} + o(T^{-1}).$$

Here,

$$\mathbb{E}\{\tilde{V}^{'}D^{r'}\tilde{V}\Lambda^{**}HC^{*'}\Psi_b J_{h-b}^{'}\tilde{V}^{'}D^{h'}\tilde{u}_1\}$$

$$= \mathbb{E}\{\phi\tilde{u}_1^{'}D^{r'}\tilde{u}_1\phi^{'}\Lambda^{**}HC^{*'}\Psi_b J_{h-b}^{'}\phi\tilde{u}_1^{'}D^{h'}\tilde{u}_1\}$$

$$+ \mathbb{E}\{\phi\tilde{u}_1^{'}D^{r'}S^{*}\Lambda^{**}HC^{*'}\Psi_b J_{h-b}^{'}S^{*'}D^{h'}\tilde{u}_1\}$$

$$+ \mathbb{E}\{S^{*'}D^{r'}S^{*}\Lambda^{**}HC^{*'}\Psi_b J_{h-b}^{'}\phi\tilde{u}_1^{'}D^{h'}\tilde{u}_1\}$$

$$
\begin{aligned}
&+\mathbb{E}\{S^{*'}D^{r'}\tilde{u}_1\phi'\Lambda^{**}HC^{*'}\Psi_bJ'_{h-b}S^{*'}D^{h'}\tilde{u}_1\}\\
=\;&\sigma^4\phi\phi'\Lambda^{**}HC^{*'}\Psi_bJ'_{h-b}\phi tr\{(D^t+D^{t'})D^{r'}\}\\
&+0\\
&+\sigma^2 tr\{D^{t'}D^r\}\Omega J_{h-b}\Psi'_bC^*H\Lambda^{**'}\phi-\sigma^4 tr\{D^{t'}D^r\}\phi\phi'J_{h-b}\Psi'_bC^*H\Lambda^{**'}\phi\\
&+0\\
=\;&\Omega J_{h-b}\Psi'_bC^*H(tr\{D^{t'}D^r\}.I)\Lambda^{**'}(\sigma^2\phi).
\end{aligned}
$$

Therefore, the final result for equation (A.32) is :

$$
\begin{aligned}
&-\mathbb{E}\{HC^{*'}\tilde{R}'\tilde{R}I'_2[\mathbb{E}(Z'Z)]^{-1}\tilde{W}^{*'}\tilde{V}\Lambda^{**}HC^{*'}\tilde{R}'\tilde{u}_1\} \qquad (A.33)\\
&=-H\sum_{i=1}^{p}\sum_{t=i}^{T-1}(T-t)C^{*'}\Psi_iJ'_{t-i}\Omega J_{t-i}\Psi_i)I'_2[\mathbb{E}(Z'Z)]^{-1}\\
&\sum_{l=1}^{p}\sum_{r=l}^{T-1}\sum_{b=1}^{p}\sum_{h=b}^{T-1}I_2\Psi_lJ'_{r-l}\Omega J_{h-b}\Psi'_bC^*H(tr\{D^{t'}D^r\}.I)\Lambda^{**'}(\sigma^2\phi)+o(T^{-1}).
\end{aligned}
$$

(6′)

$$
\begin{aligned}
&-\mathbb{E}\{HC^{*'}\tilde{R}'\bar{Z}[\mathbb{E}(Z'Z)]^{-1}\bar{Z}'\tilde{V}\Lambda^{**}HC^{*'}\tilde{R}'\tilde{u}_1\} \qquad (A.34)\\
&=-\mathbb{E}\{H\sum_{i=1}^{p}\sum_{t=i}^{T-1}C^{*'}\Psi_iJ'_{t-i}\tilde{V}'D^{t'}\bar{Z}[\mathbb{E}(Z'Z)]^{-1}\bar{Z}'\tilde{V}\Lambda^{**}H\sum_{j=1}^{p}\sum_{s=j}^{T-1}C^{*'}\Psi_jJ'_{s-j}\tilde{V}'D^{s'}\tilde{u}_1\}.
\end{aligned}
$$

Here,

$$
\begin{aligned}
&\mathbb{E}\{\tilde{V}'D^{t'}\bar{Z}[\mathbb{E}(Z'Z)]^{-1}\bar{Z}'\tilde{V}\Lambda^{**}HC^{*'}\Psi_jJ'_{s-j}\tilde{V}'D^{s'}\tilde{u}_1\}\\
&=\mathbb{E}\{\phi\tilde{u}'_1D^{t'}\bar{Z}[\mathbb{E}(Z'Z)]^{-1}\bar{Z}'\tilde{u}_1\phi'\Lambda^{**}HC^{*'}\Psi_jJ'_{s-j}\phi\tilde{u}'_1D^{s'}\tilde{u}_1\}\\
&\quad+\mathbb{E}\{\phi\tilde{u}'_1D^{t'}\bar{Z}[\mathbb{E}(Z'Z)]^{-1}\bar{Z}'S\Lambda^{**}HC^{*'}\Psi_jJ'_{s-j}S'D^{s'}\tilde{u}_1\}\\
&\quad+\mathbb{E}\{S'D^{t'}\bar{Z}[\mathbb{E}(Z'Z)]^{-1}\bar{Z}'S\Lambda^{**}HC^{*'}\Psi_jJ'_{s-j}\phi\tilde{u}'_1D^{s'}\tilde{u}_1\}\\
&\quad+\mathbb{E}\{S'D^{t'}\bar{Z}[\mathbb{E}(Z'Z)]^{-1}\bar{Z}'\tilde{u}_1\phi'\Lambda^{**}HC^{*'}\Psi_jJ'_{s-j}S'D^{s'}\tilde{u}_1\}
\end{aligned}
$$

$$
\begin{aligned}
&= \sigma^4 \phi\phi^{'} \Lambda^{**} HC^{*'} \Psi_j J^{'}_{s-j} \phi tr\{(D^s + D^{s'}) D^{t'}\} \bar{Z}[\mathbb{E}(Z'Z)]^{-1} \bar{Z}^{'} \\
&\quad + \sigma^2 \phi tr\{\Omega \Lambda^{**} HC^{*'} \Psi_j J^{'}_{s-j}\} tr\{D^{t'} \bar{Z}[\mathbb{E}(Z'Z)]^{-1} \bar{Z}^{'} D^{s'}\} \\
&\quad - \sigma^4 \phi tr\{\phi\phi^{'} \Lambda^{**} HC^{*'} \Psi_j J^{'}_{s-j}\} tr\{D^{t'} \bar{Z}[\mathbb{E}(Z'Z)]^{-1} \bar{Z}^{'} D^{s'}\} \\
&\quad + \sigma^2 tr\{D^{t'} \bar{Z}[\mathbb{E}(Z'Z)]^{-1} \bar{Z}^{'} D^{s'}\} \Omega J_{s-j} \Psi^{'}_j C^* H \Lambda^{**'} \phi \\
&\quad - \sigma^4 tr\{D^{t'} \bar{Z}[\mathbb{E}(Z'Z)]^{-1} \bar{Z}^{'} D^{s'}\} \phi\phi^{'} J_{s-j} \Psi^{'}_j C^* H \Lambda^{**'} \phi \\
&= tr\{\Omega \Lambda^{**} HC^{*'} \Psi_j J^{'}_{s-j}\} (tr\{D^{t'} \bar{Z}[\mathbb{E}(Z'Z)]^{-1} \bar{Z}^{'} D^{s'}.I)(\sigma^2 \phi) \\
&\quad + (tr\{D^{t'} \bar{Z}[\mathbb{E}(Z'Z)]^{-1} \bar{Z}^{'} D^{s'}\}.I) \Omega J_{s-j} \Psi^{'}_j C^* H \Lambda^{**'} (\sigma^2 \phi).
\end{aligned}
$$

Then, the final result for equation (A.34) is:

$$
\begin{aligned}
&- \mathbb{E}\{HC^{*'} \tilde{R}^{'} \bar{Z}[\mathbb{E}(Z'Z)]^{-1} \bar{Z}^{'} \tilde{V} \Lambda^{**} HC^{*'} \tilde{R}^{'} \tilde{u}_1\} \qquad (A.35) \\
&= -H \sum_{i=1}^{p} \sum_{t=i}^{T-1} \sum_{j=1}^{p} \sum_{s=j}^{T-1} \Psi_i J^{'}_{t-i} (tr\{\Omega \Lambda^{**} HC^{*'} \Psi_j J^{'}_{s-j}\}.I) \\
&\qquad \times (tr\{D^{t'} \bar{Z}[\mathbb{E}(Z'Z)]^{-1} \bar{Z}^{'} D^{s'}.I)(\sigma^2 \phi) \\
&\quad - H \sum_{i=1}^{p} \sum_{t=i}^{T-1} \sum_{j=1}^{p} \sum_{s=j}^{T-1} (tr\{D^{t'} \bar{Z} H^* \bar{Z}^{'} D^{s'}\}.I) C^{*'} \Psi_i J^{'}_{t-i} \Omega HC^{*'} \Psi_j J^{'}_{s-j} H \Lambda^{**'} \sigma^2 \phi.
\end{aligned}
$$

(7′)

$$
\begin{aligned}
&- \mathbb{E}\{H \Lambda^{**'} \tilde{V}^{'} \tilde{W}^* (\mathbb{E}(Z^{'} Z))^{-1} I^{'}_2 \tilde{R}^{'} \tilde{R} C^* HC^{*'} \tilde{R}^{'} \tilde{u}_1\} \qquad (A.36) \\
&= -\mathbb{E}\left\{ H \Lambda^{**'} \tilde{V}^{'} \left[\sum_{i=1}^{p} \sum_{t=i}^{T-1} D^t \tilde{V} J_{t-i} \Psi^{'}_i : 0 \right] (\mathbb{E}(Z^{'} Z))^{-1} I^{'}_2 \sum_{j=1}^{p} \sum_{s=j}^{T-1} \Psi_j J_{s-j} \tilde{V}^{'} D^{s'} \right. \\
&\qquad \left. \times \sum_{l=1}^{p} \sum_{r=l}^{T-1} D^r \tilde{V} J_{r-l} \Psi^{'}_l C^* H \sum_{b=1}^{p} \sum_{h=b}^{T-1} C^{*'} D^h \tilde{V} J_{h-b} \Psi^{'}_b \tilde{u}_1 \right\} \\
&= -\mathbb{E}\left\{ H \Lambda^{**'} \tilde{V}^{'} \left[\sum_{i=1}^{p} \sum_{t=i}^{T-1} D^t \tilde{V} J_{t-i} \Psi^{'}_i : 0 \right] (\mathbb{E}(Z^{'} Z))^{-1} I^{'}_2 \right. \\
&\qquad \left. \times \sum_{j=1}^{p} \sum_{s=j}^{T-1} (T-s) \Psi_j J_{s-j} \Omega J_{s-j} \Psi^{'}_j C^* H \sum_{b=1}^{p} \sum_{h=b}^{T-1} C^{*'} \Psi_b J^{'}_{h-b} \tilde{V}^{'} D^{h'} \tilde{u}_1 \right\}.
\end{aligned}
$$

Here,

$$
\begin{aligned}
&\mathbb{E}\left\{\tilde{V}^{'}D^{t}\tilde{V}J_{t-i}\Psi_i^{'}I_2^{'}(\mathbb{E}(Z^{'}Z))^{-1}I_2^{'}\left[\sum_{j=1}^{p}\sum_{s=j}^{T-1}(T-s)\Psi_jJ_{s-j}\Omega J_{s-j}\Psi_j^{'}C^{*}\right]\right.\\
&\qquad\left.\times HC^{*'}\Psi_bJ_{h-b}^{'}\tilde{V}^{'}D^{h'}\tilde{u}_1\right\}\\
&=\mathbb{E}\left\{\phi\tilde{u}_1^{'}D^{t}\tilde{u}_1\phi^{'}J_{t-i}\Psi_i^{'}I_2^{'}(\mathbb{E}(Z^{'}Z))^{-1}I_2^{'}\left[\sum_{j=1}^{p}\sum_{s=j}^{T-1}(T-s)\Psi_jJ_{s-j}\Omega J_{s-j}\Psi_j^{'}C^{*}\right]\right.\\
&\qquad\left.\times HC^{*'}\Psi_bJ_{h-b}^{'}\phi\tilde{u}_1^{'}D^{h'}\tilde{u}_1\right\}\\
&\quad+\mathbb{E}\left\{\phi\tilde{u}_1^{'}D^{t}SJ_{t-i}e_i^{'}I_2^{'}(\mathbb{E}(Z^{'}Z))^{-1}I_2^{'}\left[\sum_{j=1}^{p}\sum_{s=j}^{T-1}(T-s)\Psi_jJ_{s-j}\Omega J_{s-j}\Psi_j^{'}C^{*}\right]\right.\\
&\qquad\left.\times HC^{*'}\Psi_bJ_{h-b}^{'}S^{'}D^{h'}\tilde{u}_1\right\}\\
&\quad+\mathbb{E}\left\{S^{'}D^{t}SJ_{t-i}\Psi_i^{'}I_2^{'}(\mathbb{E}(Z^{'}Z))^{-1}I_2^{'}\left[\sum_{j=1}^{p}\sum_{s=j}^{T-1}(T-s)e_jJ_{s-j}\Omega J_{s-j}\Psi_j^{'}C^{*}\right]\right.\\
&\qquad\left.\times HC^{*'}\Psi_bJ_{h-b}^{'}\phi\tilde{u}_1^{'}D^{h'}\tilde{u}_1\right\}\\
&\quad+\mathbb{E}\left\{S^{'}D^{t}\tilde{u}_1\phi^{'}J_{t-i}\Psi_i^{'}I_2^{'}(\mathbb{E}(Z^{'}Z))^{-1}I_2^{'}\left[\sum_{j=1}^{p}\sum_{s=j}^{T-1}(T-s)\Psi_jJ_{s-j}\Omega J_{s-j}\Psi_j^{'}C^{*}\right]\right.\\
&\qquad\left.\times HC^{*'}\Psi_bJ_{h-b}^{'}S^{'}D^{h'}\tilde{u}_1\right\}\\
&=\sigma^4\phi\phi^{'}J_{t-i}\Psi_i^{'}I_2^{'}(\mathbb{E}(Z^{'}Z))^{-1}I_2^{'}\left[\sum_{j=1}^{p}\sum_{s=j}^{T-1}(T-s)\Psi_jJ_{s-j}\Omega J_{s-j}\Psi_j^{'}C^{*}\right]\\
&\qquad\times HC^{*'}\Psi_bJ_{h-b}^{'}\phi tr\left\{(D^{t}+D^{t'})D^{h'}\right\}\\
&\quad+\sigma^2\phi tr\left\{\Omega J_{t-i}\Psi_i^{'}I_2^{'}(\mathbb{E}(Z^{'}Z))^{-1}I_2^{'}\left[\sum_{j=1}^{p}\sum_{s=j}^{T-1}(T-s)\Psi_jJ_{s-j}\Omega J_{s-j}\Psi_j^{'}C^{*}\right]\right.\\
&\qquad\left.\times HC^{*'}\Psi_bJ_{h-b}^{'}\right\}tr\left\{D^{t}D^{h'}\right\}\\
&\quad-\sigma^4\phi tr\left\{\phi\phi^{'}J_{t-i}\Psi_i^{'}I_2^{'}(\mathbb{E}(Z^{'}Z))^{-1}I_2^{'}\left[\sum_{j=1}^{p}\sum_{s=j}^{T-1}(T-s)\Psi_jJ_{s-j}\Omega J_{s-j}\Psi_j^{'}C^{*}\right]\right.
\end{aligned}
$$

$$\left. \times HC^{*'}\Psi_b J_{h-b}^{'} \right\} tr\left\{D^t D^{h'}\right\}$$

$$+0$$

$$+0$$

$$= tr\left\{ \Omega J_{t-i}\Psi_i^{'} I_2^{'}(\mathbb{E}(Z^{'}Z))^{-1} I_2^{\prime} \left[\sum_{j=1}^{p}\sum_{s=j}^{T-1}(T-s)\Psi_j J_{s-j}\Omega J_{s-j}\Psi_j^{'} C^* \right] \right.$$

$$\left. \times HC^{*'}\Psi_b J_{h-b}^{'} \right\} tr\left\{D^t D^{h'}\right\}(\sigma^2\phi).$$

Therefore, the final result for equation (A.36) is:

$$-\mathbb{E}\left\{H\Lambda^{**'}\tilde{V}^{'}\tilde{W}^{*}(\mathbb{E}(Z^{'}Z))^{-1}I_2^{'}\tilde{R}^{'}\tilde{R}C^{*}HC^{*'}\tilde{R}^{'}\tilde{u}_1\right\} \tag{A.37}$$

$$= -H\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{b=1}^{p}\sum_{h=b}^{T-1}\left(tr\{\Omega J_{t-i}\Psi_i^{'}I_2^{'}(\mathbb{E}(Z^{'}Z))^{-1}I_2^{\prime}\left[\sum_{j=1}^{p}\sum_{s=j}^{T-1}(T-s)\Psi_j J_{s-j}\Omega J_{s-j}\Psi_j^{'}C^{*}\right]\right.$$

$$\left. \times HC^{*'}\Psi_b J_{h-b}^{'}\}.I \right)\left(tr\{D^t D^{h'}.I\}\right)\Lambda^{**'}(\sigma^2\phi).$$

(8′)

$$-\mathbb{E}\left\{H\Lambda^{**'}\tilde{V}^{'}\bar{Z}(\mathbb{E}(Z^{'}Z))^{-1}\bar{Z}^{'}\tilde{R}C^{*}HC^{*'}\tilde{R}^{'}\tilde{u}_1\right\} \tag{A.38}$$

$$= -\mathbb{E}\{H\Lambda^{**'}\tilde{V}^{'}\bar{Z}(\mathbb{E}(Z^{'}Z))^{-1}\bar{Z}^{'}\sum_{i=1}^{p}\sum_{t=i}^{T-1}D^t\tilde{V}J_{t-i}\Psi_i^{'}C^{*}H\sum_{j=1}^{p}\sum_{s=j}^{T-1}C^{*'}\Psi_j J_{s-j}^{'}\tilde{V}^{'}D^{s'}\tilde{u}_1\}.$$

Here,

$$\mathbb{E}\{\tilde{V}^{'}\bar{Z}(\mathbb{E}(Z^{'}Z))^{-1}\bar{Z}^{'}D^t\tilde{V}J_{t-i}v_i^{'}C^{*}HC^{*'}\Psi_j J_{s-j}^{'}\tilde{V}^{'}D^{s'}\tilde{u}_1\}$$

$$= \mathbb{E}\{\phi\tilde{u}_1^{'}\bar{Z}(\mathbb{E}(Z^{'}Z))^{-1}\bar{Z}^{'}D^t\tilde{u}_1\phi^{'}J_{t-i}\Psi_i^{'}C^{*}HC^{*'}v_j J_{s-j}^{'}\phi\tilde{u}_1^{'}D^{s'}\tilde{u}_1\}$$

$$+\mathbb{E}\{\phi\tilde{u}_1^{'}\bar{Z}(\mathbb{E}(Z^{'}Z))^{-1}\bar{Z}^{'}D^t S J_{t-i}v_i^{'}C^{*}HC^{*'}\Psi_j J_{s-j}^{'}S^{'}D^{s'}\tilde{u}_1\}$$

$$+\mathbb{E}\{S^{'}\bar{Z}(\mathbb{E}(Z^{'}Z))^{-1}\bar{Z}^{'}D^t S J_{t-i}\Psi_i^{'}C^{*}HC^{*'}\Psi_j J_{s-j}^{'}\phi\tilde{u}_1^{'}D^{s'}\tilde{u}_1\}$$

$$
\begin{aligned}
&+\mathbb{E}\{S^{'}\bar{Z}(\mathbb{E}(Z^{'}Z))^{-1}\bar{Z}^{'}D^{t}\tilde{u}_1\phi^{'}J_{t-i}\Psi_i^{'}C^{*}HC^{*^{'}}\Psi_j J_{s-j}^{'}S^{'}D^{s^{'}}\tilde{u}_1\}\\
=&\,\sigma^4\phi\phi^{'}J_{t-i}\Psi_i^{'}cHc^{'}\Psi_j J_{s-j}^{'}\phi tr\{(D^{s}+D^{s^{'}})\bar{Z}(\mathbb{E}(Z^{'}Z))^{-1}\bar{Z}^{'}D^{t}\}\\
&+\sigma^2\phi tr\{\Omega J_{t-i}\Psi_i^{'}C^{*}HC^{*^{'}}\Psi_j J_{s-j}^{'}\}tr\{\bar{Z}(\mathbb{E}(Z^{'}Z))^{-1}\bar{Z}^{'}D^{t}D^{s^{'}}\}\\
&-\sigma^4\phi tr\{\phi\phi^{'}J_{t-i}\Psi_i^{'}C^{*}HC^{*^{'}}\Psi_j J_{s}^{'}\}tr\{\bar{Z}(\mathbb{E}(Z^{'}Z))^{-1}\bar{Z}^{'}D^{t}D^{s^{'}}\}\\
&+\sigma^2 tr\{\bar{Z}(\mathbb{E}(Z^{'}Z))^{-1}\bar{Z}^{'}D^{t}D^{s}\}\Omega J_{s-j}\Psi_j^{'}C^{*}HC^{*^{'}}\Psi_i J_{t-i}^{'}\phi\\
&-\sigma^4 tr\{\bar{Z}(\mathbb{E}(Z^{'}Z))^{-1}\bar{Z}^{'}D^{t}D^{s}\}\phi\phi^{'}J_{s-j}\Psi_j^{'}C^{*}HC^{*^{'}}\Psi_i J_{t-i}^{'}\phi\\
&+0\\
=&\,(tr\{\Omega J_{t-i}\Psi_i^{'}C^{*}HC^{*^{'}}\Psi_j J_{s-j}^{'}\}.I)(tr\{\bar{Z}(\mathbb{E}(Z^{'}Z))^{-1}\bar{Z}^{'}D^{t}D^{s^{'}}\}.I)(\sigma^2\phi)\\
&+\Omega J_{s-j}\Psi_j^{'}C^{*}HC^{*^{'}}\Psi_i J_{t-i}^{'}(tr\{\bar{Z}(\mathbb{E}(Z^{'}Z))^{-1}\bar{Z}^{'}D^{t}D^{s}\}.I)(\sigma^2\phi).
\end{aligned}
$$

Therefore, the final result of equation (A.38) is:

$$
\begin{aligned}
&-\mathbb{E}\{H\Lambda^{**^{'}}\tilde{V}^{'}\bar{Z}(\mathbb{E}(Z^{'}Z))^{-1}\bar{Z}^{'}\tilde{R}C^{*}HC^{*^{'}}\tilde{R}^{'}\tilde{u}_1\} \qquad (A.39)\\
&=-H\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1}(tr\{\Omega J_{t-i}\Psi_i^{'}C^{*}HC^{*^{'}}\Psi_j J_{s-j}^{'}\}.I)\\
&\qquad\qquad\times(tr\{\bar{Z}(\mathbb{E}(Z^{'}Z))^{-1}\bar{Z}^{'}D^{t}D^{s^{'}}\}.I)\Lambda^{**^{'}}(\sigma^2\phi)\\
&\quad-H\Lambda^{**^{'}}\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1}\Omega J_{s-j}\Psi_j^{'}C^{*}HC^{*^{'}}\Psi_i J_{t-i}^{'}(tr\{\bar{Z}(\mathbb{E}(Z^{'}Z))^{-1}\bar{Z}^{'}D^{t}D^{s}\}.I)(\sigma^2\phi).
\end{aligned}
$$

(9′)

$$
\begin{aligned}
&-\mathbb{E}\{HC^{*^{'}}\tilde{R}'\tilde{R}C^{*}HC^{*^{'}}\tilde{R}^{'}\tilde{u}_1\} \qquad (A.40)\\
&=-\mathbb{E}\left\{H\sum_{i=1}^{p}\sum_{t=i}^{T-1}C^{*^{'}}\Psi_i J_{t-i}^{'}\tilde{V}^{'}D^{t^{'}}\sum_{j=1}^{p}\sum_{s=j}^{T-1}D^{s}\tilde{V}J_{s-j}\Psi_j^{'}C^{*}H\sum_{l=1}^{p}\sum_{r=l}^{T-1}C^{*^{'}}\Psi_l J_{r-l}^{'}\tilde{V}^{'}D^{r^{'}}\tilde{u}_1\right\}\\
&=-\mathbb{E}\left\{H\left[\sum_{i=1}^{p}\sum_{t=i}^{T-1}(T-t)C^{*^{'}}\Psi_i J_{t-i}^{'}\Omega J_{t-i}\Psi_i^{'}C^{*}\right]H\sum_{l=1}^{p}\sum_{r=l}^{T-1}C^{*^{'}}\Psi_l J_{r-l}^{'}\tilde{V}^{'}D^{r^{'}}\tilde{u}_1\right\}
\end{aligned}
$$

$$= -\mathbb{E}\left\{H\left[\sum_{i=1}^{p}\sum_{t=i}^{T-1}(T-t)C^{*'}\Psi_i J_{t-i}^{'}\Omega J_{t-i}\Psi_i^{'}C^{*}\right]H\sum_{l=1}^{p}\sum_{r=l}^{T-1}C^{*'}\Psi_l J_{r-l}^{'}\phi\tilde{u}_1^{'}D^{r'}\tilde{u}_1\right\}$$

$$= 0.$$

Therefore, by combining equation (A.24),(A.26), (A.28), (A.30), (A.33), (A.35), (A.37), (A.39), (A.40), we can get the final expression for (v).

Rearranging for the final expression

Recall $H^* = [E(Z^{'}Z)]^{-1}$, set $H^{**} = I_2^{'}H^*I_2$ and assume $\tau = \sigma^2\phi$ and $\vartheta = \Lambda^{**'}\tau$. We will add all the expectations from $((i) - (v)$ which refer to equation (A.4, (A.6),(A.9), (A.10), (A.12), (A.13), (A.14), (A.16),(A.19) (A.20), (A.22), (A.24),(A.26), (A.28),(A.30), (A.33), (A.35), (A.37), (A.39), (A.40)) we can get the final expression which is our Theorem 1 equation (2.15).

A.2 Numerical Results

Table A.1 Approximation bias and MC 2SLS bias, when L=2, 4, 6; T=50, 100

		$T = 50$			$T = 100$		
		$L = 2$	$L = 4$	$L = 6$	$L = 2$	$L = 4$	$L = 6$
$\beta_{21} = 2.00$	MC 2SLS bias	-0.3042	-0.6434	-0.6338	-0.1253	-0.1600	-0.2250
	Approximation bias	-0.3229	-0.7150	-0.8931	-0.1597	-0.1799	-0.3233
	Simultaneity part	-0.5322	-0.8123	-0.9305	-0.2831	-0.3641	-0.5643
	Dynamic Part	0.2093	0.0973	0.0374	0.1234	0.1842	0.2410
$\beta_{31} = 5.00$	MC 2SLS bias	-0.6466	-1.0910	-1.0130	-0.2604	-0.3115	-0.3958
	Approximation bias	-0.6439	-1.1902	-1.4003	-0.3015	-0.2159	-0.4608
	Simultaneity part	-0.9908	-1.4162	-2.7651	-0.6001	-0.4621	-0.7661
	Dynamic Part	0.3469	0.2260	1.3648	0.2986	0.2462	0.3053
$\alpha_{11}^{1} = 0.50$	MC 2SLS bias	0.0241	-0.0919	-0.0127	0.0082	0.0078	0.0120
	Approximation bias	0.0365	-0.0784	-0.0241	0.0120	0.0106	0.0198
	Simultaneity part	0.1815	-0.0926	-0.1079	0.0310	0.1028	0.0603
	Dynamic Part	-0.1450	0.0142	0.0838	-0.0190	-0.0922	-0.0405
$\alpha_{21}^{1} = 0.36$	MC 2SLS bias	0.0291	-0.0110	0.0158	0.0134	0.0206	0.0243
	Approximation bias	0.0513	-0.0216	0.0251	0.0252	0.0196	0.0351
	Simultaneity part	0.0501	-0.0285	-0.0732	0.1096	0.0561	0.0571
	Dynamic Part	0.0012	0.0069	0.0481	-0.0844	-0.0365	-0.0220
$\alpha_{31}^{1} = 0.40$	MC 2SLS bias	0.0297	0.1265	-0.2573	0.0346	0.0343	-0.0264
	Approximation bias	0.0337	0.1593	-0.2149	0.0283	0.0525	-0.0407
	Simultaneity part	0.1247	0.3770	-0.5128	0.0403	0.1698	-0.0700
	Dynamic Part	-0.091	-0.2177	0.2979	-0.012	-0.1173	0.0293
$\alpha_{11}^{2} = 1.20$	MC 2SLS bias	-0.1651	-0.2898	-0.2569	-0.0636	-0.0688	-0.1028
	Approximation bias	-0.1324	-0.3514	-0.3281	-0.0804	-0.1095	-0.1502
	Simultaneity part	-0.1889	-0.7067	-1.4103	-0.1007	-0.2154	-0.2771
	Dynamic Part	0.0565	0.3553	1.0822	0.0203	0.1095	0.1269

Continued on next page

Table A.1 – continued from previous page

		$T = 50$			$T = 100$		
		$L = 2$	$L = 4$	$L = 6$	$L = 2$	$L = 4$	$L = 6$
$\alpha^2_{21} = 0.60$	MC 2SLS bias	-0.0613	-0.0793	-0.0952	-0.0152	-0.0197	-0.0114
	Approximation bias	-0.0580	-0.0803	-0.0811	-0.0191	-0.0217	-0.0247
	Simultaneity part	-0.0590	-0.0972	-0.1033	-0.0679	-0.1000	-0.0189
	Dynamic Part	0.0010	0.0169	0.0222	0.0488	0.0783	-0.0058
$\alpha^2_{31} = -0.38$	MC 2SLS bias	0.2232	0.0391	0.1675	0.0800	0.0688	0.0994
	Approximation bias	0.3746	0.0407	0.1803	0.1018	0.0825	0.1098
	Simultaneity part	0.7055	0.0967	0.4849	0.1164	0.0755	0.3245
	Dynamic Part	-0.3309	-0.0056	-0.3046	-0.0146	0.007	-0.2147
$\alpha^3_{11} = 0.65$	MC 2SLS bias	-0.0639	-0.0962	-0.2440	-0.0231	-0.0297	-0.0596
	Approximation bias	-0.0702	-0.1208	-0.2921	-0.0259	-0.0540	-0.0732
	Simultaneity part	-0.1840	-0.5007	-0.3786	-0.1027	-0.0708	-0.1102
	Dynamic Part	0.1138	0.3799	0.0865	0.0768	0.0168	0.0370
$\alpha^3_{21} = 1.20$	MC 2SLS bias	-0.1081	-0.2849	-0.2184	-0.053	-0.0465	-0.0876
	Approximation bias	-0.1399	-0.2087	-0.2034	-0.0507	-0.0603	-0.1280
	Simultaneity part	-0.1539	-0.5886	-0.2733	-0.1497	-0.1010	-0.3024
	Dynamic Part	0.0140	0.3799	0.0699	0.0990	0.0407	0.1744
$\alpha^3_{31} = 0.38$	MC 2SLS bias	-0.0874	-0.1323	-0.1399	-0.0386	-0.0318	-0.0435
	Approximation bias	-0.1064	-0.2073	-0.1601	-0.0411	-0.0535	-0.0739
	Simultaneity part	-0.1559	-0.3960	-0.2609	-0.0533	-0.0720	-0.1032
	Dynamic Part	0.0495	0.1887	0.1008	0.0122	-0.0185	0.0293
$\alpha^4_{11} = 0.50$	MC 2SLS bias	-0.0006	-0.1251	-0.0851	0.0017	0.0062	-0.0023
	Approximation bias	-0.0011	-0.0987	-0.1003	0.0020	0.0110	-0.0004
	Simultaneity part	-0.0096	-0.1703	-0.0673	-0.0170	-0.0413	-0.0107
	Dynamic Part	0.0085	0.0716	-0.033	0.0150	0.0303	0.0103

Continued on next page

Table A.1 – continued from previous page

		$T=50$			$T=100$		
		$L=2$	$L=4$	$L=6$	$L=2$	$L=4$	$L=6$
$\alpha_{21}^4=0.60$	MC 2SLS bias	-0.0261	-0.0246	-0.1820	-0.0055	-0.0187	-0.0450
	Approximation bias	-0.0258	-0.0208	-0.1921	-0.0068	-0.0335	-0.0410
	Simultaneity part	-0.1456	-0.0736	-0.8031	-0.0108	-0.1024	-0.1599
	Dynamic Part	0.1198	0.0528	0.6110	0.0040	0.0689	0.1189
$\alpha_{31}^4=-0.20$	MC 2SLS bias	0.0204	0.2545	-0.0040	0.0076	0.0181	0.0100
	Approximation bias	0.0340	0.3612	-0.0091	0.0030	0.0211	0.0263
	Simultaneity part	0.1723	0.4532	-0.0195	0.0010	0.0760	0.0781
	Dynamic Part	-0.1383	-0.092	0.0104	0.0020	-0.0549	-0.0518
$c_{11}=1.00$	MC 2SLS bias	-0.0944	-0.2015	-0.2674	-0.0373	-0.0437	-0.0721
	Approximation bias	-0.0821	-0.2872	-0.3813	-0.0374	-0.0386	-0.1071
	Simultaneity part	-0.1966	-0.5648	-0.5241	-0.1067	-0.0977	-0.2654
	Dynamic Part	0.1145	0.2776	0.1428	0.0693	0.0591	0.1583
$c_{21}=0.60$	MC 2SLS bias	-0.0570	-0.1148	-0.1111	-0.0231	-0.0275	-0.0410
	Approximation bias	-0.0846	-0.1590	-0.1846	-0.0252	-0.0290	-0.0572
	Simultaneity part	-0.1129	-0.3222	-0.4027	-0.1016	-0.0713	-0.1404
	Dynamic Part	0.0283	0.1632	0.2181	0.0764	0.0423	0.0832
$c_{31}=-0.50$	MC 2SLS bias	0.0471	0.1004	0.0973	0.0188	0.0247	0.0325
	Approximation bias	0.0778	0.1264	0.0703	0.0221	0.0240	0.0488
	Simultaneity part	0.1543	0.2651	0.1176	0.0731	0.0381	0.0529
	Dynamic Part	-0.0765	-0.1387	-0.0473	-0.0510	-0.0141	-0.0041

Table A.1 presents the bias approximation of the 17 first structural form coefficients in two stage least square estimators and the bias of the Monte Carlo two stage least square estimator. The bias approximation comes from dynamic part and simultaneity part are also reported separately in Table A.1.The sample size is 50 and 100 respectively, and for the over-identification level we choose three different cases ($L=2$, $L=4$ and $L=6$).

* Both the Monte Carlo bias and the bias approximation increase when the sample size increases from 50 to 100 in the coefficients $\alpha_{21}^1=0.36$, when $L=4,6$; $\alpha_{31}^2=-0.38$, when $L=4$; $\alpha_{31}^4=-0.2$, when $L=6$). It seems abnormal, however,that as in my other experiments, the bias increases when the sample size increases from 50 to 70, then decreases again when the sample size increases. Thus, the trend of the bias of these coefficients decreases when sample size increases. Please see the Note table D.1 .

Table A.2 Bootstrap and C2SLS bias, when L=2, 4, 6; T=50, 100

		$T=50$			$T=100$		
		$L=2$	$L=4$	$L=6$	$L=2$	$L=4$	$L=6$
$\beta_{21}=2.00$	MC 2SLS bias	-0.3032(-15%)	-0.6444(-32%)	-0.6318(-32%)	-0.1260(-6%)	-0.1600(-8%)	-0.2243(-11%)
	Bootstrap bias	-0.1262(-6%)	-0.4187(-21%)	-0.4490(-22%)	-0.0289(-1%)	-0.0518(-3%)	-0.0851(-4%)
	C2SLS bias	-0.0200(-1%)	0.0303(-2%)	0.0396(+2%)	0.00678(+0%)	0.0188(+0%)	-0.0278(-1%)
$\beta_{31}=5.00$	MC 2SLS bias	-0.6473(-13%)	-1.0954(-22%)	-1.0068(-20%)	-0.2539(-5%)	-0.3109(-6%)	-0.3947(-8%)
	Bootstrap bias	-0.2788(-6%)	-0.6910(-14%)	-0.7109(-14%)	-0.0505(-1%)	-0.0938(-3%)	-0.14339-3%
	C2SLS bias	0.0973(+2%)	0.0140(+6%)	-0.0016(-0%)	0.0830(+2%)	0.0315(+0%)	-0.0828(-2%)
$\alpha^1_{11}=0.50$	MC 2SLS bias	0.0259(+5%)	-0.0915(-18%)	-0.0136(-3%)	0.0078(+2%)	-0.0001(-0%)	0.0144(+3%)
	Bootstrap bias	0.0208(+4%)	-0.051(-10%)	-0.0056(-1%)	0.0021(+0%)	-0.0014(-0%)	0.0063(+1%)
	C2SLS bias	0.0171(+2%)	0.0057(+3%)	0.0015(+0%)	0.0068(+1%)	0.0004(+0%)	0.0030(-1%)
$\alpha^1_{21}=0.36$	MC 2SLS bias	0.0327(+9%)	-0.0107(-3%)	0.0140(+4%)	0.0151(+4%)	0.0195+5%	0.0282(+8%)
	Bootstrap bias	0.0216(+6%)	0.0013(+0%)	0.0137(+4%)	0.0050(+1%)	0.0048(+1%)	0.01449(+3%)
	C2SLS bias	0.0120(+2%)	0.0033(+0%)	0.0066(+1%)	-0.0050(-1%)	0.0016(+0%)	0.0041(+1%)
$\alpha^1_{31}=0.40$	MC 2SLS bias	0.0287(+7%)	0.1280(+32%)	-0.2583(-65%)	0.0310(+8%)	0.0346(+9%)	-0.0312(-8%)
	Bootstrap bias	0.0297(+7%)	0.0745(+19%)	-0.1949(-49%)	0.0042(+1%)	0.0061(+2%)	-0.0167(-4%)
	C2SLS bias	0.0284(+1%)	0.0331(+6%)	-0.0197(-0%)	-0.0011(-0%)	-0.0030(-1%)	0.0077(+0%)
$\alpha^2_{11}=1.20$	MC 2SLS bias	-0.1680(-14%)	-0.2939(-24%)	-0.2552(-21%)	-0.0615(-5%)	-0.0664(-6%)	-0.1025(-9%)
	Bootstrap bias	-0.0696(-6%)	-0.1845(-15%)	-0.1735(-14%)	-0.0119(-1%)	-0.0195(-2%)	-0.0371(-3%)
	C2SLS bias	0.0292(+2%)	0.0826(+7%)	0.0470(+4%)	0.0024(+0%)	0.0097(+0%)	0.0059(+0%)

Continued on next page

Table A.2 – continued from previous page

		$T = 50$			$T = 100$		
		$L = 2$	$L = 4$	$L = 6$	$L = 2$	$L = 4$	$L = 6$
$\alpha^2_{21} = 0.60$	MC 2SLS bias	-0.0646(-11%)	-0.0778(-13%)	-0.0923(-15%)	-0.0147(-2%)	-0.0186(-3%)	-0.0131(-2%)
	Bootstrap bias	-0.0202(-3%)	-0.0514(-9%)	-0.0659(-11%)	-0.0024(-0%)	-0.0051(-1%)	-0.0055(-1%)
	C2SLS bias	0.0168(+0%)	-0.0079(-0%)	0.0204(+4%)	0.0013(+0%)	0.0042(+0%)	-0.0041(-1%)
$\alpha^2_{31} = -0.38$	MC 2SLS bias	0.2324(+61%)	0.0433(+11%)	0.1645(+43%)	0.0793(+21%)	0.0726(+19%)	0.1056(+28%)
	Bootstrap bias	0.1138(+30%)	0.0385(+10%)	0.1209(+32%)	0.0207(+5%)	0.0258(+7%)	0.0414(+11%)
	C2SLS bias	0.0345(+9%)	0.0200/(+5%)	0.0241(+6%)	0.0148(4%)	0.0043(+1%)	-0.0186(-5%)
$\alpha^3_{11} = 0.65$	MC 2SLS bias	-0.0601(-9%)	-0.0919(-14%)	-0.2445(-38%)	-0.0237(-4%)	-0.0304(-5%)	-0.0568(-9%)
	Bootstrap bias	-0.0201(-3%)	-0.0592(+9%)	-0.1830(-28%)	-0.0052(-1%)	-0.0101(-2%)	-0.0224(-3%)
	C2SLS bias	-0.0129(-2%)	-0.0437(-7%)	0.0182(+3%)	0.0020(+0%)	0.0042(+0%)	0.0073(+1%)
$\alpha^3_{21} = 1.20$	MC 2SLS bias	-0.1116(-9%)	-0.2871(-24%)	-0.2161(-18%)	-0.0518(-4%)	-0.0466(-4%)	-0.0834(-7%)
	Bootstrap bias	-0.0476(-4%)	-0.1772(-15%)	-0.1522(-13%)	-0.0109(-1%)	-0.0139(-1%)	-0.0298(-2%)
	C2SLS bias	-0.0192(-1%)	0.0382(+3%)	0.0063(+1%)	0.0024(+0%)	0.0097(+0%)	0.0060(+0%)
$\alpha^3_{31} = 0.38$	MC 2SLS bias	-0.0941(-25%)	-0.1343(-35%)	-0.1425(-38%)	-0.0383(-10%)	-0.0361(-10%)	-0.0520(-14%)
	Bootstrap bias	-0.0467(-12%)	-0.0788(-21%)	-0.0878(-23%)	-0.0102(-3%)	-0.0105(-3%)	-0.0196(-5%)
	C2SLS bias	-0.0060(-1%)	-0.0301(-8%)	-0.0208(-6%)	-0.0027(-0%)	0.0048(+1%)	-0.0268(-7%)
$\alpha^4_{11} = 0.50$	MC 2SLS bias	-0.0005(-0%)	-0.1267(-25%)	-0.0836(-17%)	0.0015(+0%)	0.0053(+1%)	-0.0022(-0%)
	Bootstrap bias	0.0116(+2%)	-0.0740(-15%)	-0.0522(-10%)	0.0007(+0%)	0.0019(+0%)	-0.0006(-0%)
	C2SLS bias	0.0003(+0%)	0.0250(+5%)	0.0084(+0%)	0.0001(+0%)	0.0010(+0%)	0.0010(+0%)

Continued on next page

Table A.2 – continued from previous page

		$T=50$			$T=100$		
		$L=2$	$L=4$	$L=6$	$L=2$	$L=4$	$L=6$
$\alpha_{21}^4=0.60$	MC 2SLS bias	-0.0241(-4%)	-0.0254(-4%)	-0.1816(-30%)	-0.0059(-1%)	-0.0194(-3%)	-0.0454(-8%)
	Bootstrap bias	-0.0064(-1%)	-0.01999(-3%)	-0.1306(-22%)	-0.0012(-0%)	-0.0062(-1%)	-0.0185(-3%)
	C2SLS bias	0.0011(+0%)	-0.0157(-2%)	-0.0135(-2%)	0.0013(-0%)	0.0023(+0%)	0.0038(+0%)
$\alpha_{31}^4=-0.20$	MC 2SLS bias	0.0201(+10%)	0.2538(+127%)	-0.0084(-4%)	0.0037(+2%)	0.0171(+9%)	0.0134(+7%)
	Bootstrap bias	0.0074(+4%)	0.1500(+75%)	-0.0156(-8%)	-0.0006(-0%)	0.0033(+2%)	0.0031(+2%)
	C2SLS bias	0.0036(+2%)	0.0201(+10%)	-0.0024(-0%)	-0.0001(-0%)	-0.0104(-1%)	0.0031(+2%)
$c_{11}=1.00$	MC 2SLS bias	-0.0947(-9%)	-0.2024(-20%)	-0.2683(-27%)	-0.0365(-4%)	-0.0440(-4%)	-0.0707(-7%)
	Bootstrap bias	-0.0316(-3%)	-0.1237(-12%)	-0.1898(-19%)	-0.0070(-1%)	-0.0138(-1%)	-0.0267(-3%)
	C2SLS bias	-0.0252(+2%)	-0.0080(-1%)	-0.0091(-1%)	0.0024(+0%)	0.0072(+1%)	0.0081(+1%)
$c_{21}=0.60$	MC 2SLS bias	-0.0572(-10%)	-0.1154(-19%)	-0.1108(-18%)	-0.0231(-4%)	-0.0273(-5%)	-0.0404(-7%)
	Bootstrap bias	-0.0233(-4%)	-0.0747(-12%)	-0.0799(-13%)	-0.0051(-1%)	-0.0088(-1%)	-0.0154(-3%)
	C2SLS bias	-0.0127(-2%)	-0.0291(-5%)	-0.0076(-1%)	0.0014(+0%)	0.0031(+0%)	0.0040(+1%)
$c_{31}=-0.50$	MC 2SLS bias	0.0471(+9%)	0.1006(+20%)	0.0969(+19%)	0.0189(+4%)	0.0247(+5%)	0.0322(+6%)
	Bootstrap bias	0.0197(+4%)	-0.0424(+13%)	0.0694(+14%)	0.0041(+1%)	0.0079(+2%)	0.0123(+2%)
	C2SLS bias	-0.0087(-2%)	-0.0232(-5%)	-0.0097(-2%)	-0.0008(-0%)	-0.0030(-1%)	-0.0023(-0%)

Table A.2presents the bias of the 17 first structural form coefficients in the corrected estimators, i.e.residual bootstrap two stage least square estimator, and corrected two stage least square estimator by subtracting the estimated bias approximation. The sample size is 50 and 100 respectively, and for the over-identification level we choose three different cases ($L=2$, $L=4$ and $L=6$).

The Monte Carlo bias increase when the sample size increases from 50 to 100 in the coefficients $\alpha_{21}^1=0.36$, when $L=4,6$; $\alpha_{31}^2=-0.38$, when $L=4$; $\alpha_{31}^4=-0.2$, when $L=6$). It seems abnormal, however,that as in my other experiments, the bias increases when the sample size increases from 50 to 70, then decreases again when the sample size increases. Thus, the trend of the bias of these coefficients decreases when sample size increases. Please see the Note table D.1 .

Table A.3 The MSE of Bootstrap and C2SLS, when L=2, 4, 6; T=50, 100

		$T=50$			$T=100$		
		$L=2$	$L=4$	$L=6$	$L=2$	$L=4$	$L=6$
$\beta_{21}=2.00$	MSE of MC 2SLS	0.5033	0.6161	0.5510	0.1580	0.0956	0.1169
	MSE of Bootstrap	0.9837	0.5892	0.4768	0.1970	0.0992	0.1077
	MSE of C2SLS	0.4173	0.4360	0.4275	0.1334	0.0829	0.0991
$\beta_{31}=5.00$	MSE of MC 2SLS	2.0151	1.8829	1.4067	0.6353	0.3847	0.3905
	MSE of Bootstrap	3.6960	1.8449	1.2041	0.7848	0.4005	0.3647
	MSE of C2SLS	1.5251	1.3289	1.0154	0.6003	0.3419	0.3509
$\alpha_{11}^{1}=0.50$	MSE of MC 2SLS	0.0399	0.0326	0.0175	0.0152	0.0104	0.0084
	MSE of Bootstrap	0.0544	0.0395	0.0224	0.0172	0.0119	0.0098
	MSE of C2SLS	0.0368	0.0322	0.0180	0.0160	0.0105	0.0097
$\alpha_{21}^{1}=0.36$	MSE of MC 2SLS	0.0703	0.0514	0.0438	0.0315	0.0267	0.0254
	MSE of Bootstrap	0.1024	0.0738	0.0561	0.0359	0.0306	0.0294
	MSE of C2SLS	0.0464	0.0477	0.0451	0.0332	0.0290	0.0277
$\alpha_{31}^{1}=0.40$	MSE of MC 2SLS	0.3922	0.2388	0.3113	0.1783	0.1446	0.1257
	MSE of Bootstrap	0.6026	0.3167	0.3708	0.2034	0.1654	0.1501
	MSE of C2SLS	0.3392	0.2297	0.3106	0.1542	0.1445	0.1302
$\alpha_{11}^{2}=1.20$	MSE of MC 2SLS	0.1794	0.1567	0.1102	0.0585	0.0332	0.0362
	MSE of Bootstrap	0.3237	0.1647	0.1012	0.0713	0.0367	0.0366
	MSE of C2SLS	0.1774	0.1581	0.1012	0.0546	0.0337	0.0308
$\alpha_{21}^{2}=0.60$	MSE of MC 2SLS	0.1081	0.0625	0.0577	0.0375	0.0294	0.0250
	MSE of Bootstrap	0.1897	0.0868	0.0713	0.0442	0.0346	0.0298
	MSE of C2SLS	0.1056	0.0610	0.0569	0.0398	0.0294	0.0272
$\alpha_{31}^{2}=-0.38$	MSE of MC 2SLS	0.4005	0.1734	0.1813	0.1604	0.1238	0.1123
	MSE of Bootstrap	0.6020	0.2404	0.2166	0.1863	0.1413	0.1270
	MSE of C2SLS	0.4003	0.1641	0.1802	0.1409	0.1238	0.1107
$\alpha_{11}^{3}=0.65$	MSE of MC 2SLS	0.0940	0.0415	0.1167	0.0331	0.0214	0.0245
	MSE of Bootstrap	0.1662	0.0550	0.1194	0.0394	0.0246	0.0277
	MSE of C2SLS	0.0923	0.0404	0.1069	0.0328	0.0213	0.0270

Continued on next page

Table A.3 – continued from previous page

		$T=50$			$T=100$		
		$L=2$	$L=4$	$L=6$	$L=2$	$L=4$	$L=6$
$\alpha_{21}^{3}=1.20$	MSE of MC 2SLS	0.1317	0.1758	0.1044	0.055	0.0348	0.0401
	MSE of Bootstrap	0.2106	0.1893	0.1070	0.0653	0.0392	0.0434
	MSE of C2SLS	0.1300	0.1752	0.0833	0.0545	0.0342	0.0417
$\alpha_{31}^{3}=0.38$	MSE of MC 2SLS	0.2642	0.1577	0.1408	0.1273	0.1057	0.0923
	MSE of Bootstrap	0.3739	0.2131	0.1668	0.1469	0.1234	0.1086
	MSE of C2SLS	0.2138	0.1483	0.1357	0.1264	0.1068	0.0920
$\alpha_{11}^{4}=0.50$	MSE of MC 2SLS	0.0592	0.0523	0.0434	0.0196	0.0124	0.0188
	MSE of Bootstrap	0.0925	0.0631	0.0524	0.0224	0.0141	0.0138
	MSE of C2SLS	0.0584	0.0522	0.0434	0.0203	0.0124	0.0136
$\alpha_{21}^{4}=0.60$	MSE of MC 2SLS	0.0631	0.0317	0.0791	0.0252	0.0203	0.0216
	MSE of Bootstrap	0.0990	0.0438	0.0842	0.0293	0.0234	0.0246
	MSE of C2SLS	0.0635	0.0301	0.0677	0.0260	0.0199	0.0234
$\alpha_{31}^{4}=-0.20$	MSE of MC 2SLS	0.2089	0.1930	0.1031	0.0879	0.0663	0.0602
	MSE of Bootstrap	0.2914	0.2207	0.1326	0.1009	0.0754	0.706
	MSE of C2SLS	0.2075	0.1852	0.1031	0.0820	0.0661	0.0585
$c_{11}=1.00$	MSE of MC 2SLS	0.1111	0.0892	0.1125	0.0277	0.0138	0.0169
	MSE of Bootstrap	0.2136	0.100	0.1042	0.0344	0.0153	0.0171
	MSE of C2SLS	0.0982	0.0901	0.1042	0.0254	0.0135	0.0137
$c_{21}=0.60$	MSE of MC 2SLS	0.0214	0.0209	0.0177	0.007	0.0036	0.0046
	MSE of Bootstrap	0.0425	0.0210	0.0160	0.0088	0.0040	0.0045
	MSE of C2SLS	0.0150	0.0203	0.0086	0.0069	0.0034	0.0038
$c_{31}=-0.50$	MSE of MC 2SLS	0.0139	0.0160	0.01380	0.0049	0.0030	0.0030
	MSE of Bootstrap	0.0272	0.0115	0.01250	0.0062	0.0032	0.0029
	MSE of C2SLS	0.0097	0.0164	0.0115	0.0040	0.0026	0.0031

Table A.3 presents the mean squared errors of the 17 first structural form coefficients in the Monte Carlo two stage least squares, corrected two stage least squares and bootstrap two stage least squares respectively. The sample size is 50 and 100 respectively, and for the over-identification level we choose three different cases ($L=2$, $L=4$ and $L=6$) .

Appendix B

Appendix for Chapter 3

B.1 The Evaluation for Theorem 2

Taking the expectation of equation 3.9

$$
\begin{aligned}
\mathbb{E}\left(\hat{\delta}_{FLIML}-\delta\right) &= \mathbb{E}\left((\hat{\Upsilon}'_{FLIMl}\Upsilon_{FLIML})^{-1}\hat{\Upsilon}'_{FLIML}\tilde{u}_1\right) \\
&= \underbrace{\mathbb{E}\left(H\bar{\Upsilon}\tilde{u}_1 + H\Delta_1'\tilde{u}_1 + H\Delta_2'\tilde{u}_1 - HJ_1^*H\bar{\Upsilon}'\tilde{u}_1 - HJ_1^*H\Delta_2\tilde{u}_1\right)}_{\text{The same as 2SLS bias}} \\
&\quad + \underbrace{\mathbb{E}\left(H\left[(1-(\lambda-\frac{1}{T-P-Q}))\hat{V}_2 : 0 : 0\right]'\tilde{u}_1\right)}_{\text{Extra term in FLIML compared with 2SLS}} + o(T^{-1}).
\end{aligned} \tag{B.1}
$$

In the expansion, Δ_1 and Δ_2 are defined in equation 2.10.

$\mathbb{E}\left(H\bar{\Upsilon}\tilde{u}_1 + H\Delta_1'\tilde{u}_1 + H\Delta_2'\tilde{u}_1 - HJ_1^*H\bar{\Upsilon}'\tilde{u}_1 - HJ_1^*H\Delta_2\tilde{u}_1\right)$ is the bias approximation in 2SLS, and $\mathbb{E}\left(H\left[(1-(\lambda-\frac{1}{T-P-Q}))\hat{V}_2 : 0 : 0\right]'\tilde{u}_1\right)$ is the extra term of bias approximation in the Fuller limited information maximum likelihood estimation compared with 2SLS. Hence, we start to calculate from this extra term, then compare the results with the 2SLS results obtained in section 2.3.

$$
\begin{aligned}
&\mathbb{E}\left(H\left[(1-(\lambda-\frac{1}{T-P-Q}))\hat{V}_2:0:0\right]^{'}\tilde{u}_1\right) \qquad \text{(B.2)}\\
&=H\mathbb{E}\left((1-\lambda+\frac{1}{T-P-Q})(\tilde{V}_2:0:0)'\bar{P}_Z\tilde{u}_1\right)\\
&=H\Lambda^{**'}\mathbb{E}\left((1-\lambda+\frac{1}{T-P-Q})\tilde{V}'\bar{P}_Z\tilde{u}_1\right),
\end{aligned}
$$

where $\hat{V}_2=\bar{P}_Z\tilde{V}_2$, $\bar{P}_Z=I-Z(Z'Z)^{-1}Z'$, and recall that $[\tilde{V}_2:0:0]=\tilde{V}\begin{bmatrix} I_g:0\\ 0\end{bmatrix}=\tilde{V}\Lambda^{**}$, and $Z=[R:S]$ is $T\times(P+Q)$ dimension matrix.

To proceed, using Kadane (1974) and Kadane (1970), we have

$$
1-\lambda=\frac{-\tilde{u}_1'(\bar{P}_{\tilde{\Upsilon}}-\bar{P}_Z)'\tilde{u}_1}{\tilde{u}_1'\bar{P}_Z\tilde{u}_1}+o_p(T^{-1})
$$

where $\bar{P}_{\tilde{\Upsilon}}=I-\tilde{\Upsilon}(\tilde{\Upsilon}'\tilde{\Upsilon})^{-1}\tilde{\Upsilon}'$, and recall that $\tilde{\Upsilon}=[\tilde{Y}_2:R_1:S_1]$ is a $T\times(g+P^*+Q^*)$ dimension matrix, and $\tilde{Y}_2=Y_2-V_2=\sum_{i=1}^{p}L^{(i)}Y\Gamma_2^i+\sum_{j=0}^{q}L^jX\Pi_2^{(j)}$. Then, we can rewrite equation B.2 as:

$$
\begin{aligned}
&H\Lambda^{**'}\mathbb{E}\left((1-\lambda+\frac{1}{T-P-Q})\tilde{V}'\bar{P}_Z\tilde{u}_1\right) \qquad \text{(B.3)}\\
&=H\Lambda^{**'}\mathbb{E}\left(\frac{-\tilde{u}_1'(\bar{P}_{\tilde{\Upsilon}}-\bar{P}_Z)\tilde{u}_1}{\tilde{u}_1'\bar{P}_Z\tilde{u}_1}\tilde{V}'\bar{P}_Z\tilde{u}_1\right)+\frac{\mathbb{E}(\tilde{V}'\bar{P}_Z\tilde{u}_1)}{T-P-Q}
\end{aligned}
$$

Notes:

1.Recalling $\tilde{V}=S^*+\tilde{u}_1\phi^{'}$, where $\tilde{u}_1$ and S^* are normally distributed but independent, then,

$$
\begin{aligned}
\tilde{V}'\bar{P}_Z\tilde{u}_1&=\mathbb{E}\left(\tilde{V}'\bar{P}_Z\tilde{u}_1\right)+\left(\tilde{V}'\bar{P}_Z\tilde{u}_1-\mathbb{E}\left(\tilde{V}'\bar{P}_Z\tilde{u}_1\right)\right)\\
&\equiv\mathbb{E}\left(\tilde{V}'\bar{P}_Z\tilde{u}_1\right)+o_p(T)
\end{aligned}
$$

$$= \mathbb{E}\left(S^{*'}\bar{P}_Z\tilde{u}_1\right) + \mathbb{E}\left(\phi\tilde{u}_1'\bar{P}_Z\tilde{u}_1\right) + o_p(T).$$

2. $\mathbb{E}(S^{*'}\bar{P}_Z\tilde{u}_1) = 0$

Proof:

$\mathbb{E}(S^{*'}\bar{P}_Z\tilde{u}_1) = \mathbb{E}(S^{*'}\left(I - Z(Z'Z)^{-1}Z'\right)\tilde{u}_1) = -\mathbb{E}(S^{*'}\left(Z(Z'Z)^{-1}Z'\right)\tilde{u}_1)$($\tilde{u}_1$ and S^* are independent) $\Rightarrow -\mathbb{E}(S^{*'}\left((\bar{Z}+\tilde{W}^*)H^*(\bar{Z}'+\tilde{W}^{*'})\right)\tilde{u}_1)$($\tilde{W}^*$ is lagged stochastic part and $\tilde{W}^*$, $\tilde{u}_1$ and S^* are independent between each other)$\Rightarrow -\mathbb{E}(S^{*'}\left(\tilde{W}^*H^*\tilde{W}^{*'}\right)\tilde{u}_1) = 0$.

Hence, we can express equation (B.3) as follows:

$$\begin{aligned} &H\Lambda^{**'}\mathbb{E}\left((1-\lambda+\frac{1}{T-P-Q})\tilde{V}'\bar{P}_Z\tilde{u}_1\right) \\ &= H\Lambda^{**'}\phi\mathbb{E}\left(-\tilde{u}_1'(\bar{P}_{\tilde{\Upsilon}}-\bar{P}_Z)\tilde{u}_1\frac{\mathbb{E}(\tilde{u}_1'\bar{P}_Z\tilde{u}_1)+\mathbb{E}(S'\bar{P}_Z\tilde{u}_1)}{\tilde{u}_1'\bar{P}_Z u_1}\right) + H\Lambda^{**'}\phi\frac{\mathbb{E}(\tilde{u}_1'\bar{P}_Z\tilde{u}_1)}{T-P-Q} \end{aligned} \tag{B.4}$$

Note:

1. $\dfrac{1}{\tilde{u}_1'\bar{P}_Z\tilde{u}_1} \equiv \dfrac{1}{\mathbb{E}(\tilde{u}_1'\bar{P}_Z\tilde{u}_1)} + o_p(T^{-3/2})$

Proof:

$$\tilde{u}_1'\bar{P}_Z\tilde{u}_1 = \mathbb{E}(\tilde{u}_1'\bar{P}_Z\tilde{u}_1) + \tilde{u}_1'\bar{P}_Z\tilde{u}_1 - \mathbb{E}(\tilde{u}_1'\bar{P}_Z\tilde{u}_1) = \mathbb{E}(\tilde{u}_1'\bar{P}_Z\tilde{u}_1)\left[1+\frac{\tilde{u}_1'\bar{P}_Z\tilde{u}_1-\mathbb{E}(\tilde{u}_1'\bar{P}_Z\tilde{u}_1)}{\mathbb{E}(\tilde{u}_1'\bar{P}_Z\tilde{u}_1)}\right].$$

$$\Rightarrow \frac{1}{\tilde{u}_1'\bar{P}_Z\tilde{u}_1} = \frac{1}{\mathbb{E}(\tilde{u}_1'\bar{P}_Z\tilde{u}_1)}\left[1+\frac{\tilde{u}_1'\bar{P}_Z\tilde{u}_1-\mathbb{E}(\tilde{u}_1'\bar{P}_Z\tilde{u}_1)}{\mathbb{E}(\tilde{u}_1'\bar{P}_Z\tilde{u}_1)}\right]^{-1} \equiv \frac{1}{\mathbb{E}(\tilde{u}_1'\bar{P}_Z\tilde{u}_1)} + O_p(T^{-3/2}).$$

Hence, we can express equation (B.3) as follows:

$$\begin{aligned} &H\Lambda^{**'}\mathbb{E}\left((1-\lambda+\frac{1}{T-P-Q})\tilde{V}'\bar{P}_Z\tilde{u}_1\right) \\ &= H\Lambda^{**'}\phi\mathbb{E}\left(-\tilde{u}_1'((I-\tilde{\Upsilon}(\tilde{\Upsilon}'\tilde{\Upsilon})^{-1}\tilde{\Upsilon}')-(I-Z(Z'Z)^{-1}Z'))\tilde{u}_1\frac{\mathbb{E}(\tilde{u}_1'\bar{P}_Z\tilde{u}_1)}{\mathbb{E}(\tilde{u}_1'\bar{P}_Z u_1)}\right) \end{aligned} \tag{B.5}$$

$$
\begin{aligned}
&+H\Lambda^{**'}\phi\frac{\mathbb{E}(\tilde{u}_1'\bar{P}_Z\tilde{u}_1)}{T-P-Q}\\
&=H\Lambda^{**'}\phi\mathbb{E}\left(-\tilde{u}_1'(Z(Z'Z)^{-1}Z'-\tilde{\Upsilon}(\tilde{\Upsilon}'\tilde{\Upsilon})^{-1}\tilde{\Upsilon}')\tilde{u}_1\frac{\mathbb{E}(\tilde{u}_1'\bar{P}_Z\tilde{u}_1)}{\mathbb{E}(\tilde{u}_1'\bar{P}_Z u_1)}\right)+H\Lambda^{**'}\phi\frac{\mathbb{E}(\tilde{u}_1'\bar{P}_Z\tilde{u}_1)}{T-P-Q}\\
&=-H\Lambda^{**'}\phi\mathbb{E}\left(\tilde{u}_1'Z(Z'Z)^{-1}Z'\tilde{u}_1\right)+H\Lambda^{**'}\phi\mathbb{E}\left(\tilde{u}_1'\tilde{\Upsilon}(\tilde{\Upsilon}'\tilde{\Upsilon})^{-1}\tilde{\Upsilon}'\tilde{u}_1\right)+H\Lambda^{**'}\phi\frac{\mathbb{E}(\tilde{u}_1'\bar{P}_Z\tilde{u}_1)}{T-P-Q}.
\end{aligned}
$$

We shall shortly evaluate these terms but first we note the following Lemmas.

B.1.1 Lemmas

The following lemma may be used in the late evaluations.

Lemma 1: **The expectation of a product of three normal random variables is zero.** i.e

$$\mathbb{E}(\Xi A\Psi B\Phi)=0$$

, where Ξ,Ψ, and Φ are three normal (means of zero) random variables.

Lemma 2: $(Z'Z)^{-1}=[\mathbb{E}(Z'Z)]^{-1}+O_p(T^{\frac{-3}{2}})$ **and** $(\tilde{\Upsilon}'\tilde{\Upsilon})^{-1}=[\mathbb{E}(\Upsilon'\Upsilon)]^{-1}+O_p(T^{\frac{-3}{2}})$.
where $Z=[R:S]=\left[LY,L^2Y...L^pY:X,LX,L^2X...L^qX\right]$,
$\Upsilon=[Y_2:R_1:S_1]=\left[Y_2:LY_1,L^2Y_1...L^pY_1:X,LX_1,L^2X_1...L^qX_1\right]$.

Proof:
$(Z'Z)^{-1}=[\mathbb{E}(Z'Z)]^{-1}+O_p(T^{\frac{-3}{2}})$, see A.1, Lemma 2.
One may follow the same procedure to derive $(\tilde{\Upsilon}'\tilde{\Upsilon})^{-1}=[\mathbb{E}(\Upsilon'\Upsilon)]^{-1}+O_p(T^{\frac{-3}{2}})$.

In what follows frequent use will be made of this lemma when replacing $(Z'Z)^{-1}$ and $(\tilde{\Upsilon}'\tilde{\Upsilon})^{-1}$ with $[\mathbb{E}(Z'Z)]^{-1}$ and $[\mathbb{E}(\Upsilon'\Upsilon)]^{-1}$ respectively, while retaining the order of the approximation at T^{-1}.

B.1.2 The extra term expression

As has been noted the bias approximation to order T^{-1} for FLIML is equal to the order T^{-1} bias approximation for 2SLS plus the expected value to order T^{-1} of the additional terms presented in equation B.2 and equation B.5. Hence we shall evaluate the three terms in equation B.5. Recall that $H = (\mathbb{E}(\Upsilon'\Upsilon))^{-1} = (\bar{\Upsilon}'\bar{\Upsilon} + \mathbb{E}(\Delta_2'\Delta_2))^{-1}$ and $H^* = (\mathbb{E}(Z'Z))^{-1} = (\bar{Z}'\bar{Z} + \mathbb{E}(W^{*'}W^*))^{-1}$.

1. The first term in equation B.5 can be evaluated as:

$$\begin{aligned} &-H\Lambda^{**'}\phi\mathbb{E}\left(\tilde{u}_1'(Z(Z'Z)^{-1}Z'\tilde{u}_1)\right) \\ &\quad = -H\Lambda^{**'}\phi\mathbb{E}\left(\tilde{u}_1'(\bar{Z}+\tilde{W}^*)\left[\mathbb{E}(Z'Z)\right]^{-1}(\bar{Z}'+\tilde{W}^{*'})\tilde{u}_1\right) \end{aligned} \tag{B.6}$$

Using Lemma 2

$$\begin{aligned} &= -H\Lambda^{**'}\phi\mathbb{E}\left(\tilde{u}_1'\bar{Z}H^*\bar{Z}'\tilde{u}_1\right) - H\Lambda^{**'}\phi\mathbb{E}\left(\tilde{u}_1'\tilde{W}^*H^*\tilde{W}^{*'}\tilde{u}_1\right) \\ &= -H\Lambda^{**'}\phi\sigma^2\left(tr\left\{\bar{Z}H^*\bar{Z}'\right\}.I\right) \\ &\quad - H\Lambda^{**'}\phi\mathbb{E}\left(\tilde{u}_1'\left[\sum_{i=1}^{p}\sum_{t=i}^{T-1}D^t\tilde{V}J_{t-i}\Psi_i' : 0\right]H^*\left[\sum_{j=1}^{p}\sum_{s=j}^{T-1}\Psi_j J_{s-j}'\tilde{V}'D^{s'} : 0\right]\tilde{u}_1\right) \\ &= -H\Lambda^{**'}\phi\sigma^2\left(tr\left\{\bar{Z}H^*\bar{Z}'\right\}.I\right) \\ &\quad - H\Lambda^{**'}\phi\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1}\mathbb{E}\left(\tilde{u}_1'D^t\tilde{V}J_{t-i}\Psi_i'I_2'H^*I_2\Psi_j J_{s-j}'\tilde{V}'D^{s'}\tilde{u}_1\right) \\ &= -H\Lambda^{**'}\phi\sigma^2\left(tr\left\{\bar{Z}H^*\bar{Z}'\right\}.I\right) \\ &\quad - H\Lambda^{**'}\phi\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1}\mathbb{E}\left(\tilde{u}_1'D^t(S^*+\tilde{u}_1\phi')J_{t-i}\Psi_i'H^{**}\Psi_j J_{s-j}'(S^{*'}+\phi\tilde{u}_1')D^{s'}\tilde{u}_1\right) \\ &= -H\Lambda^{**'}\phi\sigma^2\left(tr\left\{\bar{Z}H^*\bar{Z}'\right\}.I\right) \\ &\quad - H\Lambda^{**'}\phi\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1}\mathbb{E}\left(\tilde{u}_1'D^t\tilde{u}_1\phi'J_{t-i}\Psi_i'H^{**}\Psi_j J_{s-j}'\phi\tilde{u}_1'D^{s'}\tilde{u}_1\right) \\ &\quad - H\Lambda^{**'}\phi\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1}\mathbb{E}\left(\tilde{u}_1'D^tS^*J_{t-i}\Psi_i'H^{**}\Psi_j J_{s-j}'S^{*'}D^{s'}\tilde{u}_1\right) \end{aligned} \tag{B.7}$$

$$= -H\left(tr\left\{\bar{Z}H^{*}\bar{Z}'\right\}.I\right)\vartheta - H\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i,j}^{T-1}(T-t)tr\left(\Omega J_{t-i}\Psi_i' H^{**}\Psi_j J_{t-j}'\right)\vartheta.$$

By using Lemma 2 and Lemma 4 in Appendix.A.

2. The second term in equation B.5 :

$$H\Lambda^{**'}\phi\mathbb{E}\left(\tilde{u}_1'\tilde{\Upsilon}(\tilde{\Upsilon}'\tilde{\Upsilon})^{-1}\tilde{\Upsilon}'\tilde{u}_1\right) \tag{B.8}$$

$$\begin{aligned}
&= H\Lambda^{**'}\phi\mathbb{E}\left(\tilde{u}_1'(\bar{\Upsilon}+\Delta_2)\left[\mathbb{E}(\Upsilon'\Upsilon)\right]^{-1}(\bar{\Upsilon}'+\Delta_2')\tilde{u}_1\right)\\
&= H\Lambda^{**'}\phi\mathbb{E}\left(\tilde{u}_1'\bar{\Upsilon}H\bar{\Upsilon}'\tilde{u}_1\right)+H\Lambda^{**'}\phi\mathbb{E}\left(\tilde{u}_1'\Delta_2 H\Delta_2'\tilde{u}_1\right)\\
&= H\Lambda^{**'}\phi\sigma^2\left(\mathrm{tr}\left\{\bar{\Upsilon}H\bar{\Upsilon}'\right\}.I\right)\\
&\quad + H\Lambda^{**'}\phi\mathbb{E}\left(\tilde{u}_1'\sum_{i=1}^{p}\sum_{t=i}^{T-1}D^t\tilde{V}J_{t-i}\Psi_i'C^{*}H\sum_{j=1}^{p}\sum_{s=j}^{T-1}C^{*'}\Psi_j J_{s-j}'\tilde{V}'D^{s'}\tilde{u}_1\right)\\
&= H\Lambda^{**'}\phi\sigma^2\left(\mathrm{tr}\left\{\bar{\Upsilon}H\bar{\Upsilon}'\right\}.I\right)\\
&\quad + H\Lambda^{**'}\phi\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1}\mathbb{E}\left(\tilde{u}_1'D^t(S^{*}+\tilde{u}_1\phi')J_{t-i}\Psi_i'C^{*}HC^{*'}\Psi_j J_{s-j}'(S^{*'}+\phi\tilde{u}_1')D^{s'}\tilde{u}_1\right)\\
&= H\Lambda^{**'}\phi\sigma^2\left(\mathrm{tr}\left\{\bar{\Upsilon}H\bar{\Upsilon}'\right\}.I\right)\\
&\quad + H\Lambda^{**'}\phi\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1}\mathbb{E}\left(\tilde{u}_1'D^t\tilde{u}_1\phi'J_{t-i}\Psi_i'C^{*}HC^{*'}\Psi_j J_{s-j}'\phi\tilde{u}_1'D^{s'}\tilde{u}_1\right)\\
&\quad + H\Lambda^{**'}\phi\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{j=1}^{p}\sum_{s=j}^{T-1}\mathbb{E}\left(\tilde{u}_1'D^tS^{*}J_{t-i}\Psi_i'C^{*}HC^{*'}\Psi_j J_{s-j}'S^{*'}D^{s'}\tilde{u}_1\right)\\
&= H\left(\mathrm{tr}\left\{\bar{\Upsilon}H\bar{\Upsilon}'\right\}.I\right)\vartheta + H\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i,j}^{T-1}(T-t)\,\mathrm{tr}\left(\Omega J_{t-i}\Psi_i'C^{*}HC^{*'}\Psi_j J_{t-j}'\right)\vartheta.
\end{aligned}$$

By using Lemma 2 and Lemma 4 in Appendix.A.

3. The third term in equation B.5 :

$$H\Lambda^{**'}\phi\frac{\mathbb{E}(\tilde{u}_1'\bar{P}_Z\tilde{u}_1)}{T-P-Q} \tag{B.9}$$

$$
\begin{aligned}
&= H\Lambda^{**'}\phi\frac{\mathbb{E}(\tilde{u}_1'(I-P_Z)\tilde{u}_1)}{T-P-Q} \\
&= H\Lambda^{**'}\phi\frac{T\sigma^2-\mathbb{E}(\tilde{u}_1'P_Z\tilde{u}_1)}{T-P-Q} \\
&= H\Lambda^{**'}\phi\frac{T\sigma^2-\mathbb{E}\left(\tilde{u}_1'(Z(Z'Z)^{-1}Z'\tilde{u}_1)\right)}{T-P-Q} \\
&= H\Lambda^{**'}\phi\frac{T\sigma^2-\mathbb{E}\left(\tilde{u}_1'(\bar{Z}+\tilde{W}^*)\left[\mathbb{E}(Z'Z)\right]^{-1}(\bar{Z}'+\tilde{W}^{*'})\tilde{u}_1\right)}{T-P-Q} \\
&= H\Lambda^{**'}\phi\frac{T\sigma^2-\mathbb{E}\left(\tilde{u}_1'\bar{Z}H^*\bar{Z}'\tilde{u}_1\right)-\mathbb{E}\left(\tilde{u}_1'\tilde{W}^*H^*\tilde{W}^{*'}\tilde{u}_1\right)}{T-P-Q}.
\end{aligned}
$$

Notes:

1. Using the result in equation B.6 $\mathbb{E}\left(\tilde{u}_1'\bar{Z}\bar{H}^*\bar{Z}'\tilde{u}_1\right)=\text{tr}\{\bar{Z}H^*\bar{Z}'\}$, then $\dfrac{H\Lambda^{**'}\phi\,\text{tr}\{\bar{Z}H^*\bar{Z}'\}}{T-P-Q}$ is $O(T^{-2})$ which is of lower order than order T^{-1} .

2. $\mathbb{E}\left(\tilde{u}_1'\tilde{W}^*H^*\tilde{W}^{*'}\tilde{u}_1\right)$ is $O(1)$, then $\dfrac{H\Lambda^{**'}\phi\mathbb{E}\left(\tilde{u}_1'\tilde{W}^*H^*\tilde{W}^{*'}\tilde{u}_1\right)}{T-P-Q}$ is of lower order than T^{-1} .

Hence, we can express equation (B.9) as follows:

$$
\begin{aligned}
&H\Lambda^{**'}\phi\frac{\mathbb{E}(\tilde{u}_1'\bar{P}_Z\tilde{u}_1)}{T-P-Q} \\
&= H\vartheta+o(T^{-1}).
\end{aligned}
\tag{B.10}
$$

Then combining these three terms, the extra bias term in FLIML given by B.5 can be written as:

$$
\begin{aligned}
&\mathbb{E}\left(H\left[(1-(\lambda-\frac{1}{T-P-Q}))\hat{V}_2:0:0\right]'\tilde{u}_1\right) \\
&= -H\left(tr\left\{\bar{Z}H^*\bar{Z}'\right\}.I\right)\vartheta-H\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i,j}^{T-1}(T-t)tr\left(\Omega J_{t-i}\Psi_i'H^{**}\Psi_jJ_{t-j}'\right)\vartheta \\
&\quad +H\left(\text{tr}\left\{\bar{\Upsilon}H\bar{\Upsilon}'\right\}.I\right)\vartheta+H\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i,j}^{T-1}(T-t)\,\text{tr}\left(\Omega J_{t-i}\Psi_i'C^*HC^{*'}\Psi_jJ_{t-j}'\right)\vartheta \\
&\quad +H\vartheta+o(T^{-1}).
\end{aligned}
\tag{B.11}
$$

B.1.3 The Related terms in 2SLS compared with B.1.2

The extra terms in equation B.11 which have been presented in Theorem 1, section 2.3. We shall show that the first ten terms in Theorem 1 are cancelled out by the terms in equation B.11 thus demonstrating that the FLIML bias approximation contains only those components which explicitly involve the D matrices. To show this we first note the first ten terms in Theorem 1:

$$H\left(tr\{\bar{Z}H^*\bar{Z}'\}.I\right)\vartheta \tag{B.12}$$

$$H\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i,j}^{T-1}(T-t)tr\left(\Omega J_{t-i}\Psi_i'H^{**}\Psi_j J_{t-j}'\right)\vartheta \tag{B.13}$$

$$-H\left(tr\{\bar{Z}H^*\bar{Z}'\bar{\Upsilon}H\bar{\Upsilon}'\}.I\right)\vartheta \tag{B.14}$$

$$-H\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i,j}^{T-1}(T-t)(tr\{\Omega J_{t-i}\Psi_i' I_2' H^*\bar{Z}'\bar{\Upsilon}HC^{*'}\Psi_j J_{t-j}'\}.I)\vartheta \tag{B.15}$$

$$-H\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i,j}^{T-1}(T-t)tr\left\{\bar{Z}(\mathbb{E}(Z'Z))^{-1}I_2\Psi_i J_{t-i}'\Omega J_{t-j}\Psi_j' C^* H\bar{\Upsilon}'\right\}\vartheta \tag{B.16}$$

$$\begin{aligned}-H\sum_{i=1}^{p}\sum_{l=1}^{p}\sum_{t=i,l}^{T-1}(T-t)\Bigg(tr\Bigg\{\Omega J_{t-i}\Psi_i' H^{**}\Bigg[&\sum_{j=1}^{p}\sum_{m=1}^{p}\sum_{s=j,m}^{T-1}(T-s)\Psi_j J_{s-j} \\ &\times\Omega J_{s-m}\Psi_m' C^*\Bigg]Hc'\Psi_l J_{s-l}'\Bigg\}.I\Bigg)\vartheta\end{aligned} \tag{B.17}$$

$$-H\bar{\Upsilon}'\bar{Z}H^*\bar{Z}'\bar{\Upsilon}H\vartheta \tag{B.18}$$

$$-H\sum_{i=1}^{p}\sum_{l=1}^{p}\sum_{t=i,j}^{T-1}(T-t)C^{*'}\Psi_i J_{t-i}'\Omega J_{t-j}\Psi_j' I_2' H^*\bar{Z}'\bar{\Upsilon}H\vartheta \tag{B.19}$$

$$-H\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i,j}^{T-1}(T-t)C^{*'}\Psi_i J_{t-i}'\Omega J_{t-j}\Psi_j' H^{**}\sum_{l=1}^{p}\sum_{m=1}^{p}\sum_{r=l,m}^{T-1}(T-l)\Psi_l J_{r-l}'\Omega J_{r-m}\Psi_m' C^* H\vartheta \tag{B.20}$$

$$-H\bar{\Upsilon}'\bar{Z}H^*\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i,j}^{T-1}I_2(T-t)\Psi_i J_{t-i}'\Omega J_{t-j}\Psi_j' C^{*'}H\vartheta. \tag{B.21}$$

B.1.4 Comparing the terms in Appendix B.1.2 and Appendix B.1.3

Here we show how the the FLIML estimator eliminates the bias terms in B.12 to B.21.

1. Terms B.12 and B.13 are eliminated by the first two terms in equation B.11.

$$\begin{aligned}
&H\left(tr\{\bar{Z}H^{*}\bar{Z}'\}.I\right)\vartheta - H\left(tr\{\bar{Z}H^{*}\bar{Z}'\}.I\right)\vartheta \\
&+H\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i,j}^{T-1}(T-t)tr\left(\Omega J_{t-i}\Psi_i' H^{**}\Psi_j J_{t-j}'\right)\vartheta \\
&-H\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i,j}^{T-1}(T-t)tr\left(\Omega J_{t-i}\Psi_i' H^{**}\Psi_j J_{t-j}'\right)\vartheta \\
&=0.
\end{aligned} \tag{B.22}$$

2. Comparing the terms B.14, B.15 with the third term in equation B.11:

Recalling $\tilde{W}^{*} = \left[\tilde{R}:0\right] = \tilde{R}I_2^{'}$(Chapter 2), where I_2 is defined as $I_2 = \begin{bmatrix} I_P \\ 0 \end{bmatrix}$ in Chapter 2

$$\begin{aligned}
&H\left(tr\{\bar{\Upsilon}H\bar{\Upsilon}'\}.I\right)\vartheta - H\left(tr\{\bar{Z}H^{*}\bar{Z}'\bar{\Upsilon}H\bar{\Upsilon}'\}.I\right)\vartheta \\
&-H\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i,j}^{T-1}(T-t)(tr\{\Omega J_{t-i}\Psi_i^{'}I_2^{'}H^{*}\bar{Z}^{'}\bar{\Upsilon}HC^{*'}\Psi_j J_{t-j}^{'}\}.I)\vartheta \\
&=H\left(tr\left\{\bar{\Upsilon}H\bar{\Upsilon}'\bar{Z}\left(\bar{Z}'\bar{Z}\right)^{-1}\bar{Z}'\right\}.I\right)\vartheta - H\left(tr\left\{\bar{\Upsilon}H\bar{\Upsilon}'\bar{Z}\left(\bar{Z}'\bar{Z}+\mathbb{E}\left\{\tilde{W}^{*'}\tilde{W}\right\}\right)^{-1}\bar{Z}'\right\}.I\right)\vartheta \\
&\quad -H\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i,j}^{T-1}(T-t)(tr\{\Omega J_{t-i}\Psi_i^{'}I_2^{'}H^{*}\bar{Z}^{'}\bar{\Upsilon}HC^{*'}\Psi_j J_{t-j}^{'}\}.I)\vartheta \\
&=H\left(tr\left\{\bar{\Upsilon}H\bar{\Upsilon}'\bar{Z}\left(\left(\bar{Z}'\bar{Z}\right)^{-1}-\left(\bar{Z}'\bar{Z}+\mathbb{E}\left\{\tilde{W}^{*'}\tilde{W}\right\}\right)^{-1}\right)\bar{Z}'\right\}.I\right)\vartheta \\
&\quad -H\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i,j}^{T-1}(T-t)(tr\{\Omega J_{t-i}\Psi_i^{'}I_2^{'}H^{*}\bar{Z}^{'}\bar{\Upsilon}HC^{*'}\Psi_j J_{t-j}^{'}\}.I)\vartheta \\
&=H\left(tr\left\{\bar{\Upsilon}H\bar{\Upsilon}'\bar{Z}\left(\bar{Z}'\bar{Z}\right)^{-1}\left(\bar{Z}'\bar{Z}+\mathbb{E}\left\{\tilde{W}^{*'}\tilde{W}^{*}\right\}\right)^{-1}\mathbb{E}\left\{\tilde{W}^{*'}\tilde{W}^{*}\right\}\bar{Z}'\right\}.I\right)\vartheta
\end{aligned} \tag{B.23}$$

$$
\begin{aligned}
&-H\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i,j}^{T-1}(T-t)(tr\{\Omega J_{t-i}\Psi_i^{'}I_2^{'}H^*\bar{Z}^{'}\bar{\Upsilon}HC^{*'}\Psi_jJ_{t-j}^{'}\}.I)\vartheta\\
&=H\left(tr\left\{\bar{\Upsilon}H\bar{\Upsilon}^{'}\bar{Z}\left(\bar{Z}^{\prime}\bar{Z}\right)^{-1}H^*I_2\mathbb{E}\left\{\tilde{R}^{\prime}\tilde{R}\right\}I_2^{\prime}\bar{Z}^{\prime}\right\}.I\right)\vartheta\\
&-H\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i,j}^{T-1}(T-t)(tr\{\Omega J_{t-i}\Psi_i^{'}I_2^{'}H^*\bar{Z}^{'}\bar{\Upsilon}HC^{*'}\Psi_jJ_{t-j}^{'}\}.I)\vartheta\\
&=\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i}^{T-1}\sum_{s=j}^{T-1}H\left(tr\left\{\bar{\Upsilon}H\bar{\Upsilon}^{'}\bar{Z}\left(\bar{Z}^{\prime}\bar{Z}\right)^{-1}H^*I_2\mathbb{E}\left\{\Psi_iJ_{t-i}^{\prime}V^{\prime}D^{t^{\prime}}D^sVJ_{s-j}\Psi_j^{\prime}\right\}I_2^{\prime}\bar{Z}^{\prime}\right\}.I\right)\vartheta\\
&-H\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i,j}^{T-1}(T-t)(tr\{\Omega J_{t-i}\Psi_i^{'}I_2^{'}H^*\bar{Z}^{'}\bar{\Upsilon}HC^{*'}\Psi_jJ_{t-j}^{'}\}.I)\vartheta\\
&=\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i}^{T-1}\sum_{s=j}^{T-1}H\left(tr\left\{\bar{\Upsilon}H\bar{\Upsilon}^{'}\bar{Z}\left(\bar{Z}^{\prime}\bar{Z}\right)^{-1}H^*I_2\mathbb{E}\left\{\Psi_iJ_{t-i}^{\prime}S^{\prime}D^{t^{\prime}}D^sSJ_{s-j}\Psi_j^{\prime}\right\}I_2^{\prime}\bar{Z}^{\prime}\right\}.I\right)\vartheta\\
&+\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i}^{T-1}\sum_{s=j}^{T-1}H\left(tr\left\{\bar{\Upsilon}H\bar{\Upsilon}^{'}\bar{Z}\left(\bar{Z}^{\prime}\bar{Z}\right)^{-1}H^*I_2\mathbb{E}\left\{\Psi_iJ_{t-i}^{\prime}\phi u_1^{\prime}D^{t^{\prime}}D^su_1\phi^{\prime}J_{s-j}\Psi_j^{\prime}\right\}I_2^{\prime}\bar{Z}^{\prime}\right\}.I\right)\vartheta\\
&-H\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i,j}^{T-1}(T-t)(tr\{\Omega J_{t-i}\Psi_i^{'}I_2^{'}H^*\bar{Z}^{'}\bar{\Upsilon}HC^{*'}\Psi_jJ_{t-j}^{'}\}.I)\vartheta\\
&=H\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i,j}^{T-1}(T-t)\left(tr\left\{\bar{\Upsilon}H\bar{\Upsilon}^{'}\bar{Z}\left(\bar{Z}^{\prime}\bar{Z}\right)^{-1}H^*I_2\Psi_iJ_{t-i}^{\prime}\Omega J_{s-j}\Psi_j^{\prime}I_2^{'}\bar{Z}^{\prime}\right\}.I\right)\vartheta\\
&-H\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i,j}^{T-1}(T-t)(tr\{\Omega J_{t-i}\Psi_i^{'}I_2^{'}H^*\bar{Z}^{'}\bar{\Upsilon}HC^{*'}\Psi_jJ_{t-j}^{'}\}.I)\vartheta.
\end{aligned}
$$

Note:

1. $\bar{\Upsilon}=\left[A_2^{(*)}:I_3:I_4\right]\bar{Z}$. Where, $\left[A_2^{(*)}:I_3:I_4\right]$ is a $(P+Q)\times(g+P^*+Q^*)$ matrix. $A_2^{(*)}=\left[\Gamma_2^{(*)}:\Pi_2^{(*)}\right]^{\prime}$ is a $(P+Q)\times g$ matrix and recalling that $\Gamma_2^{(*)}=(\Gamma_2^{(1)},\Gamma_2^{(2)},...,\Gamma_2^{(p)})'$ and $\Gamma_2^{(*)}=(\Pi_2^{(1)},\Pi_2^{(2)},...,\Pi_2^{(q)})'$. I_3 is a $(P+Q)\times P^*$ selection matrix and I_4 is $(P+Q)\times Q^*$ selection matrix.

Hence, we can express equation (B.23) as follows:

$$
\begin{aligned}
&H\left(tr\{\bar{\Upsilon}H\bar{\Upsilon}'\}.I\right)\vartheta-H\left(tr\{\bar{Z}H^*\bar{Z}'\bar{\Upsilon}H\bar{\Upsilon}'\}.I\right)\vartheta \qquad \text{(B.24)}\\
&-H\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i,j}^{T-1}(T-t)(tr\{\Omega J_{t-i}\Psi_i^{'}I_2^{'}H^*\bar{Z}^{'}\bar{\Upsilon}HC^{*'}\Psi_jJ_{t-j}^{'}\}.I)\vartheta
\end{aligned}
$$

$$= H\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i,j}^{T-1}(T-t)\left(tr\left\{\bar{\Upsilon}H\left(\left[A_2^{(*)}:I_3:I_4\right]\right)^{'}\bar{Z}'\bar{Z}\left(\bar{Z}'\bar{Z}\right)^{-1}H^{*}I_2\Psi_i J'_{t-i}\Omega J_{s-j}\Psi'_j I_2^{'}\right\}.I\right)\vartheta$$
$$- H\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i,j}^{T-1}(T-t)(tr\{\Omega J_{t-i}\Psi_i^{'}I_2^{'}H^{*}\bar{Z}^{'}\bar{\Upsilon}HC^{*'}\Psi_j J_{t-j}^{'}\}.I)\vartheta$$
$$= H\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i,j}^{T-1}(T-t)(tr\{\Omega J_{t-i}\Psi_i^{'}I_2^{'}H^{*}\bar{Z}^{'}\bar{\Upsilon}H\left(C^{*'}-\left(\left[A_2^{(*)}:I_3:I_4\right]\right)^{'}I_2\right)\Psi_j J_{t-j}^{'}\}.I)\vartheta$$
$$= 0.$$

Note:

1. $C^{*'}-\left[A_2^{(*)}:I_3:I_4\right]^{'}I_2=0$, where I_2 is defined as $I_2=\begin{bmatrix}I_P\\0\end{bmatrix}$ in Chapter 2.

3. Comparing the term B.16, B.17 with the fourth term in equation B.11:

$$H\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i,j}^{T-1}(T-t)\,\mathrm{tr}\left(\Omega J_{t-i}\Psi_i' C^{*}HC^{*'}\Psi_j J_{t-j}'\right)\vartheta \tag{B.25}$$
$$- H\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i,j}^{T-1}(T-t)tr\left\{\bar{Z}(\mathbb{E}(Z^{'}Z))^{-1}I_2\Psi_i J_{t-i}^{'}\Omega J_{t-j}\Psi_j^{'}C^{*}H\bar{\Upsilon}^{'}\right\}\vartheta$$
$$- H\sum_{i=1}^{p}\sum_{l=1}^{p}\sum_{t=i,l}^{T-1}(T-t)\left(tr\left\{\Omega J_{t-i}\Psi_i^{'}H^{**}\left[\sum_{j=1}^{p}\sum_{m=1}^{p}\sum_{s=j,m}^{T-1}(T-s)\Psi_j J_{s-j}\times\Omega J_{s-m}\Psi_m^{'}C^{*}\right]HC^{*'}\Psi_l J_{s-l}^{'}\right\}.I\right)\vartheta$$
$$= H\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i,j}^{T-1}(T-t)\,\mathrm{tr}\left(\Omega J_{t-i}\Psi_i' C^{*}HC^{*'}\Psi_j J_{t-j}'\right)\vartheta$$
$$- H\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i,j}^{T-1}(T-t)tr\left\{\bar{\Upsilon}^{'}\bar{Z}(\mathbb{E}(Z^{'}Z))^{-1}I_2\Psi_i J_{t-i}^{'}\Omega J_{t-j}\Psi_j^{'}C^{*}H\right\}\vartheta$$
$$- H\sum_{i=1}^{p}\sum_{l=1}^{p}\sum_{t=i,l}^{T-1}(T-t)\left(tr\left\{\Omega J_{t-i}\Psi_i^{'}H^{**}\left[\sum_{j=1}^{p}\sum_{m=1}^{p}\sum_{s=j,m}^{T-1}(T-s)\Psi_j J_{s-j}\times\Omega J_{s-m}\Psi_m^{'}C^{*}\right]HC^{*'}\Psi_l J_{s-l}^{'}\right\}.I\right)\vartheta$$
$$= H\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i,j}^{T-1}(T-t)\,\mathrm{tr}\left(\Omega J_{t-i}\Psi_i' C^{*}HC^{*'}\Psi_j J_{t-j}'\right)\vartheta$$

$$
\begin{aligned}
&- H\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i,j}^{T-1}(T-t)tr\left\{\left(\left[A_2^{(*)}:I_3:I_4\right]\right)^{'}\bar{Z}^{'}\bar{Z}\left(\bar{Z}^{'}\bar{Z}+\mathbb{E}\left\{\tilde{W}^{*'}\tilde{W}\right\}\right)^{-1}\right.\\
&\qquad\qquad\left.\times I_2\Psi_i J_{t-i}^{'}\Omega J_{t-j}\Psi_j^{'}C^{*}H\right\}\vartheta\\
&- H\sum_{i=1}^{p}\sum_{l=1}^{p}\sum_{t=i,l}^{T-1}(T-t)\left(tr\left\{\Omega J_{t-i}\Psi_i^{'}H^{**}\left[\sum_{j=1}^{p}\sum_{m=1}^{p}\sum_{s=j,m}^{T-1}(T-s)\Psi_j J_{s-j}\right.\right.\right.\\
&\qquad\qquad\left.\left.\left.\times\Omega J_{s-m}\Psi_m^{'}C^{*}\right]HC^{*'}\Psi_l J_{s-l}^{'}\right\}.I\right)\vartheta\\
&= H\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i,j}^{T-1}(T-t)\,\mathrm{tr}\left(\Omega J_{t-i}\Psi_i' C^{*}HC^{*'}\Psi_j J_{t-j}'\right)\vartheta\\
&- H\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i,j}^{T-1}(T-t)tr\left\{\left(\left[A_2^{(*)}:I_3:I_4\right]\right)^{'}\bar{Z}^{'}\bar{Z}\right.\\
&\qquad\times\left(\left(\bar{Z}'\bar{Z}\right)^{-1}-\left(\bar{Z}'\bar{Z}\right)^{-1}\left(\bar{Z}^{'}\bar{Z}+\mathbb{E}\left\{\tilde{W}^{*'}\tilde{W}^{*}\right\}\right)^{-1}\mathbb{E}\left\{\tilde{W}^{*'}\tilde{W}\right\}\right)\\
&\qquad\left.\times I_2\Psi_i J_{t-i}^{'}\Omega J_{t-j}\Psi_j^{'}C^{*}H\right\}\vartheta\\
&- H\sum_{i=1}^{p}\sum_{l=1}^{p}\sum_{t=i,l}^{T-1}(T-t)\left(tr\left\{\Omega J_{t-i}\Psi_i^{'}H^{**}\left[\sum_{j=1}^{p}\sum_{m=1}^{p}\sum_{s=j,m}^{T-1}(T-s)\Psi_j J_{s-j}\right.\right.\right.\\
&\qquad\qquad\left.\left.\left.\times\Omega J_{s-m}\Psi_m^{'}C^{*}\right]HC^{*'}\Psi_l J_{s-l}^{'}\right\}.I\right)\vartheta
\end{aligned}
$$

Using the result in Appendix.A.1, $\mathbb{E}\{\tilde{W}^{*'}\tilde{W}^{*}\}=\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i,j}^{T-1}(T-t)tr\left(\Omega I_2 J_{t-i}\Psi_i^{'}\Psi_i J_{t-j}^{'}I_2'\right)$,

$$
\begin{aligned}
&= H\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i,j}^{T-1}(T-t)\,\mathrm{tr}\left(\Omega J_{t-i}\Psi_i' C^{*}HC^{*'}\Psi_j J_{t-j}'\right)\vartheta\\
&- H\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i,j}^{T-1}(T-t)\,\mathrm{tr}\left(\Omega J_{t-i}\Psi_i' C^{*}HC^{*'}\Psi_j J_{t-j}'\right)\vartheta\\
&+ H\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i,j}^{T-1}(T-t)tr\left\{\left(\left[A_2^{(*)}:I_3:I_4\right]\right)^{'}H^{*}I_2\right.\\
&\qquad\left.\times\sum_{l=1}^{p}\sum_{m=1}^{p}\sum_{s=l,m}^{T-1}(T-s)\Psi_l J_{t-l}^{'}\Omega J_{t-m}\Psi_m^{'}I_2^{'}I_2\Psi_i J_{t-i}^{'}\Omega J_{t-j}\Psi_j^{'}C^{*}H\right\}\vartheta\\
&- H\sum_{i=1}^{p}\sum_{l=1}^{p}\sum_{t=i,l}^{T-1}(T-t)\left(tr\left\{\Omega J_{t-i}\Psi_i^{'}H^{**}\left[\sum_{j=1}^{p}\sum_{m=1}^{p}\sum_{s=j,m}^{T-1}(T-s)\Psi_j J_{s-j}\right.\right.\right.\\
&\qquad\qquad\left.\left.\left.\times\Omega J_{s-m}\Psi_m^{'}C^{*}\right]HC^{*'}\Psi_l J_{s-l}^{'}\right\}.I\right)\vartheta
\end{aligned}
$$

*Note: $\left[A_2^{(*)} : I_3 : I_4\right]$ is a $(P+Q)\times(g+P^*+Q^*)$ matrix, $C^{*'} = \left[A_2^{(*)} : I_3 : I_4\right]' I_2$

$$
\begin{aligned}
&= H\sum_{i=1}^{p}\sum_{l=1}^{p}\sum_{t=i,l}^{T-1}(T-t)\left(tr\left\{\Omega J_{t-i}\Psi_i' H^{**}\left[\sum_{j=1}^{p}\sum_{m=1}^{p}\sum_{s=j,m}^{T-1}(T-s)\Psi_j J_{s-j}\right.\right.\right.\\
&\qquad\qquad\left.\left.\left.\times\,\Omega J_{s-m}\Psi_m' C^*\right] HC^{*'}\Psi_l J_{s-l}'\right\}.I\right)\vartheta\\
&\quad - H\sum_{i=1}^{p}\sum_{l=1}^{p}\sum_{t=i,l}^{T-1}(T-t)\left(tr\left\{\Omega J_{t-i}\Psi_i' H^{**}\left[\sum_{j=1}^{p}\sum_{m=1}^{p}\sum_{s=j,m}^{T-1}(T-s)\Psi_j J_{s-j}\right.\right.\right.\\
&\qquad\qquad\left.\left.\left.\times\,\Omega J_{s-m}\Psi_m' C^*\right] HC^{*'}\Psi_l J_{s-l}'\right\}.I\right)\vartheta\\
&= 0.
\end{aligned}
$$

4. Comparing the term B.18, B.20, B.19, B.21 with the last term in equation B.11:

$$
\begin{aligned}
&H\vartheta - H\bar{\Upsilon}'\bar{Z}H^*\bar{Z}'\bar{\Upsilon}H\vartheta \qquad\qquad \text{(B.26)}\\
&- H\sum_{i=1}^{p}\sum_{l=1}^{p}\sum_{t=i,j}^{T-1}(T-t)C^{*'}\Psi_i J_{t-i}'\Omega J_{t-j}\Psi_j' I_2' H^*\bar{Z}'\bar{\Upsilon}H\vartheta\\
&- H\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i,j}^{T-1}(T-t)C^{*'}\Psi_i J_{t-i}'\Omega J_{t-j}\Psi_j' H^{**}\sum_{l=1}^{p}\sum_{m=1}^{p}\sum_{r=l,m}^{T-1}(T-l)\Psi_l J_{r-l}'\Omega J_{r-m}\Psi_m' C^* H\vartheta\\
&- H\bar{\Upsilon}'\bar{Z}H^*\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i,j}^{T-1} I_2(T-t)\Psi_i J_{t-i}'\Omega J_{t-j}\Psi_j' C^{*'}H\vartheta\\
&= H\vartheta - H\bar{\Upsilon}'\bar{Z}H^*\bar{Z}'\left[\bar{Z}A_2^{(*)} \quad \bar{Z}\left[I_3 : I_4\right]\right]H\vartheta\\
&\quad - H\sum_{i=1}^{p}\sum_{l=1}^{p}\sum_{t=i,j}^{T-1}(T-t)C^{*'}\Psi_i J_{t-i}'\Omega J_{t-j}\Psi_j' I_2' H^*\bar{Z}'\left[\bar{Z}A_2^{(*)} \quad \bar{Z}\left[I_3 : I_4\right]\right]H\vartheta\\
&\quad - H\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i,j}^{T-1}(T-t)C^{*'}\Psi_i J_{t-i}'\Omega J_{t-j}\Psi_j' H^{**}\sum_{l=1}^{p}\sum_{m=1}^{p}\sum_{r=l,m}^{T-1}(T-l)\Psi_l J_{r-l}'\Omega J_{r-m}\Psi_m' C^* H\vartheta\\
&\quad - H\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i,j}^{T-1}(T-t)\bar{\Upsilon}'(\bar{Z})H^* I_2\Psi_i J_{t-i}'\Omega J_{t-j}\Psi_j' C^* H\vartheta\\
&= H\vartheta - H\bar{\Upsilon}'\bar{Z}\left(\bar{Z}'\bar{Z}+\mathbb{E}\left\{\tilde{W}^{*'}\tilde{W}^*\right\}\right)^{-1}\bar{Z}'\bar{Z}\left[A_2^{(*)} \quad \left[I_3 : I_4\right]\right]H\vartheta\\
&\quad - H\sum_{i=1}^{p}\sum_{l=1}^{p}\sum_{t=i,j}^{T-1}(T-t)C^{*'}\Psi_i J_{t-i}'\Omega J_{t-j}\Psi_j' I_2'\left(\bar{Z}'\bar{Z}+\mathbb{E}\left\{\tilde{W}^{*'}\tilde{W}^*\right\}\right)^{-1}\bar{Z}'
\end{aligned}
$$

$$\times \left[\bar{Z}A_2^{(*)} \quad \bar{Z}\left[I_3 : I_4\right]\right] H\vartheta$$

$$-H\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i,j}^{T-1}(T-t)C^{*'}\Psi_i J_{t-i}^{'}\Omega J_{t-j}\Psi_j' H^{**}\sum_{l=1}^{p}\sum_{m=1}^{p}\sum_{r=l,m}^{T-1}(T-l)\Psi_l J_{r-l}^{'}\Omega J_{r-m}\Psi_m^{'}C^{*}H\vartheta$$

$$-H\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i,j}^{T-1}(T-t)\bar{\Upsilon}^{'}(\bar{Z})H^{*}I_2\Psi_i J_{t-i}^{'}\Omega J_{t-j}\Psi_j^{'}C^{*}H\vartheta$$

$$= H\vartheta - H\bar{\Upsilon}'\bar{Z}\left((\bar{Z}'\bar{Z})^{-1} - H^{*}\mathbb{E}\{\tilde{W}^{*'}\tilde{W}^{*}\}(\bar{Z}'\bar{Z})^{-1}\right)\bar{Z}'\bar{Z}\left[A_2^{(*)} \quad \left[I_3 : I_4\right]\right] H\vartheta$$

$$-H\sum_{i=1}^{p}\sum_{l=1}^{p}\sum_{t=i,j}^{T-1}(T-t)C^{*'}\Psi_i J_{t-i}^{'}\Omega J_{t-j}\Psi_j^{'}I_2^{'}\left((\bar{Z}'\bar{Z})^{-1} - H^{*}\mathbb{E}\{\tilde{W}^{*'}\tilde{W}^{*}\}(\bar{Z}'\bar{Z})^{-1}\right)$$

$$\bar{Z}^{'}\left[\bar{Z}A_2^{(*)} \quad \bar{Z}\left[I_3 : I_4\right]\right] H\vartheta$$

$$-H\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i,j}^{T-1}(T-t)C^{*'}\Psi_i J_{t-i}^{'}\Omega J_{t-j}\Psi_j' H^{**}\sum_{l=1}^{p}\sum_{m=1}^{p}\sum_{r=l,m}^{T-1}(T-l)\Psi_l J_{r-l}^{'}\Omega J_{r-m}\Psi_m^{'}C^{*}H\vartheta$$

$$-H\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i,j}^{T-1}(T-t)\bar{\Upsilon}^{'}(\bar{Z})H^{*}I_2\Psi_i J_{t-i}^{'}\Omega J_{t-j}\Psi_j^{'}C^{*}H\vartheta$$

*Note: $\bar{\Upsilon} = \bar{Z}\left[A_2^{(*)} : I_3 : I_4\right]$

$$= H\vartheta - H\bar{\Upsilon}'\bar{\Upsilon}H\vartheta + H\bar{\Upsilon}'\bar{Z}H^{*}\mathbb{E}\{\tilde{W}^{*'}\tilde{W}^{*}\}(\bar{Z}'\bar{Z})^{-1}\bar{Z}'\bar{\Upsilon}H\vartheta$$

$$-H\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i,j}^{T-1}(T-t)C^{*'}\Psi_i J_{t-i}^{'}\Omega J_{t-j}\Psi_j^{'}C^{*}H\vartheta$$

$$+H\sum_{i=1}^{p}\sum_{l=1}^{p}\sum_{t=i,j}^{T-1}(T-t)C^{*'}\Psi_i J_{t-i}^{'}\Omega J_{t-j}\Psi_j^{'}I_2^{'}H^{*}I_2\sum_{l=1}^{p}\sum_{m=1}^{p}\sum_{r=l,m}^{T-1}(T-l)\Psi_l J_{r-l}^{'}$$

$$\Omega J_{r-m}\Psi_m^{'}I_2^{'}\left[A_2^{(*)} : I_3 : I_4\right] H\vartheta$$

$$-H\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i,j}^{T-1}(T-t)C^{*'}\Psi_i J_{t-i}^{'}\Omega J_{t-j}\Psi_j' H^{**}\sum_{l=1}^{p}\sum_{m=1}^{p}\sum_{r=l,m}^{T-1}(T-l)\Psi_l J_{r-l}^{'}\Omega J_{r-m}\Psi_m^{'}C^{*}H\vartheta$$

$$-H\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i,j}^{T-1}(T-t)\bar{\Upsilon}^{'}(\bar{Z})H^{*}I_2\Psi_i J_{t-i}^{'}\Omega J_{t-j}\Psi_j^{'}C^{*}H\vartheta$$

*Note: $\left[A_2^{(*)} : I_3 : I_4\right]$ is a $(P+Q)\times(g+P^{*}+Q^{*})$ matrix, $C^{*'} = \left[A_2^{(*)} : I_3 : I_4\right]^{'} I_2$

$$= H\vartheta - H\bar{\Upsilon}'\bar{\Upsilon}\left(\left(\bar{\Upsilon}'\bar{\Upsilon}\right)^{-1} - \left(\bar{\Upsilon}'\bar{\Upsilon}\right)^{-1}H\mathbb{E}\{\Delta_2^{'}\Delta_2\}\right)\vartheta$$

$$+H\bar{\Upsilon}'\bar{Z}H^{*}\mathbb{E}\{\tilde{W}^{*'}\tilde{W}^{*}\}(\bar{Z}'\bar{Z})^{-1}\bar{Z}'\bar{\Upsilon}H\vartheta$$

$$- H\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i,j}^{T-1}(T-t)C^{*'}\Psi_i J_{t-i}^{'}\Omega J_{t-j}\Psi_j^{'}C^{*}H\vartheta$$

$$- H\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i,j}^{T-1}(T-t)\bar{\Upsilon}^{'}(\bar{Z})H^{*}I_2\Psi_i J_{t-i}^{'}\Omega J_{t-j}\Psi_j^{'}C^{*}H\vartheta$$

Using the result in Appendix.A.1, $\mathbb{E}\{\tilde{W}^{*'}\tilde{W}^{*}\} = \sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i,j}^{T-1}(T-t)tr\left(\Omega I_2 J_{t-i}\Psi_i^{'}\Psi_i J_{t-j}^{'}I_2^{\prime}\right)$,

$$= H\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i,j}^{T-1}(T-t)C^{*'}\Psi_i J_{t-i}^{'}\Omega J_{t-j}\Psi_j^{'}C^{*}H\vartheta$$

$$+ H\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i,j}^{T-1}(T-t)\bar{\Upsilon}^{'}(\bar{Z})H^{*}I_2\Psi_i J_{t-i}^{'}\Omega J_{t-j}\Psi_j^{'}I_2^{'}(\bar{Z})^{-1}\bar{\Upsilon}H\vartheta$$

$$- H\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i,j}^{T-1}(T-t)C^{*'}\Psi_i J_{t-i}^{'}\Omega J_{t-j}\Psi_j^{'}C^{*}H\vartheta$$

$$- H\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{t=i,j}^{T-1}(T-t)\bar{\Upsilon}^{'}(\bar{Z})H^{*}I_2\Psi_i J_{t-i}^{'}\Omega J_{t-j}\Psi_j^{'}C^{*}H\vartheta$$

$$= 0.$$

Then the bias approximation to $O(T^{-1})$ for FLIML estimator is the remaining part in 2SLS which is summarized in Theorem 2 in section 3.3.

B.2 Numerical Results

Table B.1 **Bias Approximation of 2SLS and FLIML to $O(T^{-1})$, when L=2, 4, 6; T=50, 100**

		$T=50$			$T=100$		
		$L=2$	$L=4$	$L=6$	$L=2$	$L=4$	$L=6$
$\beta_{21}=2.00$	2SLS bias	-0.3229	-0.7150	-0.8931	-0.1597	-0.1799	-0.3233
	FLIML bias	-0.1359	-0.2886	-0.0517	-0.0739	-0.0629	-0.0567
$\beta_{31}=5.00$	2SLS bias	-0.6439	-1.1902	-1.4003	-0.3015	-0.2159	-0.4608
	FLIML bias	-0.2487	-0.3905	-0.1410	-0.1531	-0.2017	-0.0712
$\alpha^1_{11}=0.50$	2SLS bias	0.0365	-0.0784	-0.0241	0.0120	0.0106	0.0198
	FLIML bias	0.0212	-0.0395	0.0261	0.0073	-0.0070	0.0016
$\alpha^1_{21}=0.36$	2SLS bias	0.0513	-0.0216	0.0251	0.0252	0.0196	0.0351
	FLIML bias	0.0613	0.0324	0.0715	0.0208	0.0189	0.0398
$\alpha^1_{31}=0.40$	2SLS bias	0.0337	0.1593	-0.2149	0.0283	0.0525	-0.0407
	FLIML bias	0.0784	0.1396	0.0701	0.0716	0.0478	0.0300
$\alpha^2_{11}=1.20$	2SLS bias	-0.1324	-0.3514	-0.3281	-0.0804	-0.1095	-0.1502
	FLIML bias	-0.0765	-0.1032	-0.0378	-0.0297	-0.0469	-0.0210
$\alpha^2_{21}=0.60$	2SLS bias	-0.0580	-0.0803	-0.0811	-0.0191	-0.0217	-0.0247
	FLIML bias	-0.0183	-0.0108	0.0319	0.0175	0.0011	0.0089
$\alpha^2_{31}=-0.38$	2SLS bias	0.3746	0.0407	0.1803	0.1018	0.0825	0.1098
	FLIML bias	0.1692	0.0610	0.0996	0.0667	0.0578	0.0501
$\alpha^3_{11}=0.65$	2SLS bias	-0.0702	-0.1208	-0.2921	-0.0259	-0.0540	-0.0732
	FLIML bias	-0.0321	-0.0261	0.0157	-0.0183	-0.0072	0.0028
$\alpha^3_{21}=1.20$	2SLS bias	-0.1399	-0.2087	-0.2034	-0.0507	-0.0603	-0.1280
	FLIML bias	-0.0702	-0.1305	-0.0201	-0.0214	-0.0161	-0.0211
$\alpha^3_{31}=0.38$	2SLS bias	-0.1064	-0.2073	-0.1601	-0.0411	-0.0535	-0.0739
	FLIML bias	-0.0805	-0.0388	-0.1057	-0.0309	-0.0200	-0.0401
$\alpha^4_{11}=0.50$	2SLS bias	-0.0011	-0.0987	-0.1003	0.0020	0.0110	-0.0004
	FLIML bias	0.0340	-0.0270	0.0397	0.0209	0.0132	0.0147

Continued on next page

Table B.1 – continued from previous page

		$T = 50$			$T = 100$		
		$L = 2$	$L = 4$	$L = 6$	$L = 2$	$L = 4$	$L = 6$
$\alpha^4_{21} = 0.60$	2SLS bias	-0.0258	-0.0208	-0.1921	-0.0068	-0.0335	-0.0410
	FLIML bias	0.0067	-0.0190	0.0701	0.0081	0.0091	0.0070
$\alpha^4_{31} = -0.20$	2SLS bias	0.0340	0.3612	-0.0091	0.0030	0.0211	0.0263
	FLIML bias	0.0421	0.0800	0.0201	0.0049	0.0189	0.0310
$c_{11} = 1.00$	2SLS bias	-0.0821	-0.2872	-0.3813	-0.0374	-0.0386	-0.1071
	FLIML bias	0.0351	-0.1926	0.0386	0.0006	-0.0077	-0.0007
c_{21}=0.60	2SLS bias	-0.0846	-0.1590	-0.1846	-0.0252	-0.0290	-0.0572
	FLIML bias	-0.0241	-0.0491	0.0015	-0.0064	-0.0070	-0.0051
$c_{31} = -0.50$	2SLS bias	0.0778	0.1264	0.0703	0.0221	0.0240	0.0488
	FLIML bias	0.0217	0.0781	-0.0103	0.0042	0.0062	0.0002

Table B.1 presents the bias approximation of the 17 first structural form coefficients in two stage least squares and Fuller limited information maximum likelihood. The sample size is 50 and 100 respectively, and for the over-identification we choose three different cases ($L = 2$, $L = 4$ and $L = 6$).

The bias approximation of 2SLS increase when the sample size increases from 50 to 100 in the coefficients $\alpha^1_{21} = 0.36$, when $L = 4,6$; $\alpha^2_{31} = -0.38$, when $L = 4$; $\alpha^4_{31} = -0.2$, when $L = 6$). It seems abnormal, however,that as in my other experiments, the bias increases when the sample size increases from 50 to 70, then decreases again when the sample size increases. Thus, the trend of the bias of these coefficients decreases when sample size increases. Please see the Note table D.1 .

Table B.2 **Monte Carlo 2SLS vs Monte Carlo FLIML vs C2SLS vs CFLIML, when L=2, 4, 6; T=50, 100**

		$T=50$			$T=100$		
		$L=2$	$L=4$	$L=6$	$L=2$	$L=4$	$L=6$
β_{21}=2.00	MC 2SLS bias	-0.3032(-15%)	-0.6444(-32%)	-0.6318(-32%)	-0.1260(-6%)	-0.1600(-8%)	-0.2243(-11%)
	MC FLIML bias	-0.1099(-5%)	-0.1528(-8%)	-0.0573(-3%)	-0.0544(-3%)	-0.0476(-2%)	-0.0355(-2%)
	C2SLS bias	-0.0200(-1%)	0.0452(-2%)	0.0396(+2%)	0.0292(+2%)	0.0188(+1%)	-0.0278(-1%)
	CFLIML bias	-0.0220(-1%)	0.0424(-2%)	0.0118(+1%)	-0.0125(-1%)	-0.0178(-1%)	-0.0149(-1%)
$\beta_{31}=5.00$	MC 2SLS bias	-0.6473(-13%)	-1.0954(-22%)	-1.0068(-20%)	-0.2539(-5%)	-0.3109(-6%)	-0.3947(-8%)
	MC FLIML bias	-0.2665(-5%)	-0.2039(-4%)	-0.1170(-2%)	-0.1199(-2%)	-0.0878(-2%)	-0.0502 (-1%)
	C2SLS bias	0.0973(+2%)	0.1875(+3%)	-0.0016(-0%)	0.0830(+2%)	0.0315(+0%)	-0.0828(-2%)
	CFLIML bias	0.0902(-2%)	-0.1222(-2%)	-0.00178(-0%)	0.0182(+0%)	0.0330(+0%)	-0.0206(-0%)
$\alpha^1_{11}=0.50$	MC 2SLS bias	0.0259(+5%)	-0.0915(-18%)	-0.0136(-3%)	0.0078(+2%)	-0.0001(-0%)	0.0144(+3%)
	MC FLIML bias	0.0296(+6%)	-0.0256(-5%)	+0.0173(+3%)	0.0070(+1%)	-0.0048(-1%)	0.0030(+1%)
	C2SLS bias	0.0171(+3%)	0.0057(+3%)	0.0015(+0%)	0.0068(+1%)	0.0004(+0%)	0.0030(-1%)
	CFLIML bias	0.0166(+3%)	-0.0111(-2%)	0.0010(+0%)	-0.0039(-1%)	0.0006(+0%)	0.0033(+1%)
$\alpha^1_{21}=0.36$	MC 2SLS bias	0.0327(+9%)	-0.0107(-3%)	0.0140(+4%)	0.0151(+4%)	0.0195+5%	0.0282(+8%)
	MC FLIML bias	0.0419(+12%)	0.0271(+8%)	0.0524(+15%)	0.0179(+5%)	0.0130(+4%)	0.0229(+6%)
	C2SLS bias	0.0121(+2%)	0.0033(+1%)	0.0109(+3%)	-0.0050(-1%)	0.0031(+1%)	0.0041(+1%)
	CFLIML bias	0.0130(+3%)	- 0.0032(-1%)	0.0066(+2%)	-0.0052(-1%)	0.0033(+1%)	0.0068(+2%)

Continued on next page

Table B.2 – continued from previous page

		$T = 50$			$T = 100$		
		$L = 2$	$L = 4$	$L = 6$	$L = 2$	$L = 4$	$L = 6$
$\alpha_{31}^{1} = 0.40$	MC 2SLS bias	0.0287(+7%)	0.1280(+32%)	-0.2583(-65%)	0.0310(+8%)	0.0346(+9%)	-0.0312(-8%)
	MC FLIML bias	0.0660(+16%)	0.1178(+22%)	0.0451(+11%)	0.0398(+10%)	0.0350(+9%)	0.0218(+5%)
	C2SLS bias	0.0284(+1%)	0.0331(+6%)	0.0235(+6%)	-0.0039(-1%)	-0.0030(-1%)	0.0077(+0%)
	CFLIML bias	0.0008(+1%)	0.0198(+5%)	-0.0198(-5%)	+0.0042(+1%)	0.0045(+1%)	0.0003(+0%)
$\alpha_{11}^{2} = 1.20$	MC 2SLS bias	-0.1680(-14%)	-0.2939(-24%)	-0.2552(-21%)	-0.0615(-5%)	-0.0664(-6%)	-0.1025(-9%)
	MC FLIML bias	-0.0642(-5%)	-0.0834(-7%)	-0.0300(-2%)	-0.0281(-2%)	-0.0202(-2%)	-0.0122(-1%)
	C2SLS bias	0.0292(+2%)	0.0827(+7%)	0.0470(+4%)	0.0024(+0%)	0.0097(+0%)	0.0059(+0%)
	CFLIML bias	-0.0297(-2%)	0.0443(-4%)	-0.0026(-0%)	0.0041(+0%)	-0.0052(-0%)	-0.0057(-0%)
$\alpha_{21}^{2} = 0.60$	MC 2SLS bias	-0.0646(-11%)	-0.0778(-13%)	-0.0923(-15%)	-0.0147(-2%)	-0.0186(-3%)	-0.0131(-2%)
	MC FLIML bias	-0.0105(-2%)	-0.0122(-2%)	0.0276(+5%)	0.0012(+0%)	0.0027(+0%)	0.0093(+2%)
	C2SLS bias	0.0168(+0%)	-0.0079(-0%)	0.0204(+4%)	0.0013(+0%)	0.0042(+0%)	-0.0041(-1%)
	CFLIML bias	-0.0018(-0%)	0.0076(+0%)	0.0131(+2%)	0.0002(+0%)	-0.0017(-0%)	0.0013(+0%)
$\alpha_{31}^{2} = -0.38$	MC 2SLS bias	0.2324(+61%)	0.0433(+11%)	0.1645(+43%)	0.0793(+21%)	0.0726(+19%)	0.1056(+28%)
	MC FLIML bias	0.1570(+41%)	0.0505(+13%)	0.0927(+24%)	0.0517(+14%)	0.0411(+11%)	0.0390(+10%)
	C2SLS bias	0.0345(+9%)	0.0200/(+5%)	0.0241(+6%)	0.0148(4%)	0.0043(+1%)	-0.0186(-5%)
	CFLIML bias	0.0299(+8%)	0.0115/(+3%)	0.0131(+3%)	0.0081(+2%)	0.0023(+1%)	0.0149(+4%)

Continued on next page

Table B.2 – continued from previous page

		$T = 50$			$T = 100$		
		$L = 2$	$L = 4$	$L = 6$	$L = 2$	$L = 4$	$L = 6$
$\alpha_{11}^3 = 0.65$	MC 2SLS bias	-0.0601(-9%)	-0.0919(-14%)	-0.2445(-38%)	-0.0237(-4%)	-0.0304(-5%)	-0.0568(-9%)
	MC FLIML bias	-0.0238(-4%)	-0.0133(-2%)	0.0182(+3%)	-0.0229(-2%)	-0.0030(-0%)	0.0032(+0%)
	C2SLS bias	-0.0129(-2%)	-0.0437(-7%)	0.0182(+3%)	0.0040(+1%)	0.0042(+0%)	0.0073(+1%)
	CFLIML bias	-0.0018(+0%)	0.0100(+2%)	0.0104(+2%)	-0.0012(-0%)	-0.0024(-0%)	-0.0015(-0%)
$\alpha_{21}^3 = 1.20$	MC 2SLS bias	-0.1116(-9%)	-0.2871(-24%)	-0.2161(-18%)	-0.0518(-4%)	-0.0466(-4%)	-0.0834(-7%)
	MC FLIML bias	-0.0606(-5%)	-0.0939(-8%)	-0.0183(-2%)	-0.0220(-2%)	-0.0134(-1%)	-0.0121(-1%)
	C2SLS bias	-0.0192(-2%)	-0.0382(-3%)	0.0063(+1%)	0.0024(+0%)	0.0097(+0%)	0.0060(+0%)
	CFLIML bias	-0.0184(-2%)	-0.0316(-3%)	0.0041(+1%)	-0.0045(-0%)	-0.0040(-0%)	-0.0042(-0%)
$\alpha_{31}^3 = 0.38$	MC 2SLS bias	-0.0941(-25%)	-0.1343(-35%)	-0.1425(-38%)	-0.0383(-10%)	-0.0361(-10%)	-0.0520(-14%)
	MC FLIML bias	-0.0962(-25%)	-0.0924(-24%)	-0.0639(-17%)	-0.0346(-9%)	-0.0283(-7%)	-0.0317(-8%)
	C2SLS bias	-0.0060(-1%)	-0.0301(-8%)	-0.0208(-6%)	-0.0038(-1%)	0.0048(+1%)	-0.0268(-7%)
	CFLIML bias	+0.0091(-2%)	-0.0292(-8%)	-0.0184(-5%)	0.0044(+1%)	0.0046(+1%)	-0.0093(-2%)
$\alpha_{11}^4 = 0.50$	MC 2SLS bias	-0.0005(-0%)	-0.1267(-25%)	-0.0836(-17%)	0.0015(+0%)	0.0053(+1%)	-0.0022(-0%)
	MC FLIML bias	0.0231(+5%)	-0.0183(-4%)	0.0470(9%)	0.0126(+3%)	0.0123(+2%)	0.0126(+3%)
	C2SLS bias	-0.0031(-1%)	0.0250(+5%)	0.0117(-2%)	0.0007(+0%)	0.0010(+0%)	0.0010(+0%)
	CFLIML bias	-0.0046(-1%)	0.0047(+1%)	-0.0175(-4%)	0.0019(+0%)	0.0019(+0%)	0.0033(+1%)

Continued on next page

Table B.2 – continued from previous page

		$T = 50$			$T = 100$		
		$L = 2$	$L = 4$	$L = 6$	$L = 2$	$L = 4$	$L = 6$
$\alpha_{21}^4 = 0.60$	MC 2SLS bias	-0.0241(-4%)	-0.0254(-4%)	-0.1816(-30%)	-0.0059(-1%)	-0.0194(-3%)	-0.0454(-8%)
	MC FLIML bias	0.0093(+2%)	0.0144(+2%)	0.0404(7%)	0.0075(+0%)	0.0072(+1%)	0.0052(+1%)
	C2SLS bias	0.0011(+0%)	-0.0157(-2%)	-0.0135(-2%)	0.0070(-0%)	0.0023(+0%)	0.0038(+0%)
	CFLIML bias	0.0074(-1%)	0.0073(1%)	0.0149(2%)	0.0066(+0%)	0.0001(+0%)	-0.0015(-0%)
$\alpha_{31}^4 = -0.20$	MC 2SLS bias	0.0201(+10%)	0.2538(+127%)	-0.0084(-4%)	0.0037(+2%)	0.0171(+9%)	0.0134(+7%)
	MC FLIML bias	0.0390(+20%)	0.0844(+42%)	0.012(+6%)	0.0061(+3%)	0.0134(+7%)	0.0134(+7%)
	C2SLS bias	0.0036(+2%)	0.0201(+10%)	-0.0024(-0%)	-0.0012(-1%)	-0.01040(-1%)	0.0031(+2%)
	CFLIML bias	+0.0026(+1%)	-0.0071(-4%)	0.0025(+1%)	-0.0014(-1%)	0.0013(+1%)	0.0026(+1%)
$c_{11=1.00}$	MC 2SLS bias	-0.0947(-9%)	-0.2024(-20%)	-0.2683(-27%)	-0.0365(-4%)	-0.0440(-4%)	-0.0707(-7%)
	MC FLIML bias	-0.0380(-4%)	-0.1237(-12%)	0.0338(+3%)	-0.0089(-1%)	-0.0045(-0%)	-0.0014(-0%)
	C2SLS bias	-0.0252(+2%)	-0.0080(-1%)	-0.0091(-1%)	0.0024(+0%)	0.0072(+1%)	0.0081(+1%)
	CFLIML bias	-0.009(-0%)	0.0001(+0%)	0.0038(+0%)	0.0013(+0%)	-0.0028(-0%)	-0.0023(-0%)
$c_{21} = 0.60$	MC 2SLS bias	-0.0572(-10%)	-0.1154(-19%)	-0.1108(-18%)	-0.0231(-4%)	-0.0273(-5%)	-0.0404(-7%)
	MC FLIML bias	-0.0317(-5%)	-0.0376(-6%)	0.0072(+1%)	-0.0073(-1%)	-0.0056(-1%)	-0.0037(-1%)
	C2SLS bias	-0.0127(-2%)	-0.0291(-5%)	-0.0076(-1%)	0.0014(+0%)	0.0031(+0%)	0.0040(+1%)
	CFLIML bias	-0.0082(-1%)	-0.0208(-3%)	-0.0051(-1%)	0.0013(+0%)	-0.0026(-0%)	-0.0021(-0%)

Continued on next page

Table B.2 – continued from previous page

		$T = 50$			$T = 100$		
		$L = 2$	$L = 4$	$L = 6$	$L = 2$	$L = 4$	$L = 6$
$c_{31} = -0.50$	MC 2SLS bias	0.0471(+9%)	0.1006(+20%)	0.0969(+19%)	0.0189(+4%)	0.0247(+5%)	0.0322(+6%)
	MC FLIML bias	0.0269(+5%)	0.0322(+6%)	-0.0087(-2%)	0.0057(+1%)	0.0048(+1%)	0.0033(+1%)
	C2SLS bias	-0.0087(-2%)	-0.0232(-5%)	-0.0097(-2%)	0.0041(-0%)	-0.0030(-1%)	-0.0023(-0%)
	CFLIML bias	0.0074(+1%)	0.0173(+3%)	0.0029(+1%)	-0.0020(-0%)	0.0023(+0%)	0.0022(+0%)

Table B.2 presents the bias and the bias proportion of four different estimators (the uncorrected two stage least squares, the uncorrected Fuller limited information maximum likelihood, the corrected two stage lease squares and the corrected Fuller limited information maximum likelihood estimators) compare with the actual value of the 17 coefficients we are interested in . The sample size is 50 and 100 respectively, and for the over-identification we choose three different cases ($L = 2$, $L = 4$ and $L = 6$)

** The Monte Carlo bias of 2SLS increases when the sample size increases from 50 to 100 in the coefficients $\alpha_{21}^{1} = 0.36$, when $L = 4, 6$; $\alpha_{31}^{2} = -0.38$, when $L = 4$; $\alpha_{31}^{4} = -0.2$, when $L = 6$). It seems abnormal (we suppose when sample size increases, the bias would decrease), however, that as in my other experiments, the bias increases when the sample size increases from 50 to 70, then decreases again when the sample size increases. Thus, the trend of the bias of these coefficients decreases when sample size increases. Please see the Note table D.1.

Table B.3 **The MSE of 2SLS, C2SLS, FLIML, and CFLIML when L=2, 4, 6; T=50, 100**

		$T = 50$			$T = 100$		
		$L = 2$	$L = 4$	$L = 6$	$L = 2$	$L = 4$	$L = 6$
$\beta_{21} = 2.00$	MSE of MC 2SLS	0.5033	0.6161	0.5510	0.1580	0.0956	0.1169
	MSE of MC FLIML	0.3549	0.4109	0.4297	0.1554	0.0868	0.0952
	MSE of C2SLS	0.4173	0.4360	0.4275	0.1334	0.0829	0.0991
	MSE of CFLIML	0.2704	0.3209	0.4305	0.1565	0.0844	0.0897
$\beta_{31} = 5.00$	MSE of MC 2SLS	2.0151	1.8829	1.4067	0.6353	0.3847	0.3905
	MSE of MC FLIML	1.4421	1.3333	1.0301	0.6237	0.3515	0.3280
	MSE of C2SLS	1.5251	1.3289	1.0154	0.6003	0.3419	0.3509
	MSE of CFLIML	1.1372	1.5247	1.0397	0.5971	0.3521	0.2962
$\alpha^1_{11} = 0.50$	MSE of MC 2SLS	0.0399	0.0326	0.0175	0.0152	0.0104	0.0084
	MSE of MC FLIML	0.0486	0.0460	0.0359	0.0165	0.0118	0.0102
	MSE of C2SLS	0.0368	0.0322	0.0180	0.0160	0.0105	0.0097
	MSE of CFLIML	0.0382	0.0394	0.0227	0.0168	0.0121	0.0101
$\alpha^1_{21} = 0.36$	MSE of MC 2SLS	0.0703	0.0514	0.0438	0.0315	0.0267	0.0254
	MSE of MC FLIML	0.0820	0.0932	0.0890	0.0342	0.0302	0.0308
	MSE of C2SLS	0.0464	0.0477	0.0451	0.0332	0.0290	0.0277
	MSE of CFLIML	0.0550	0.0520	0.0438	0.0338	0.0303	0.0247
$\alpha^1_{31} = 0.40$	MSE of MC 2SLS	0.3922	0.2388	0.3113	0.1783	0.1446	0.1257
	MSE of MC FLIML	0.4544	0.4071	0.5253	0.1949	0.1640	0.1553
	MSE of C2SLS	0.3392	0.2297	0.3106	0.1542	0.1445	0.1302
	MSE of CFLIML	0.3877	0.2178	0.3112	0.1652	0.1441	0.1597
$\alpha^2_{11} = 1.20$	MSE of MC 2SLS	0.1794	0.1567	0.1102	0.0585	0.0332	0.0362
	MSE of MC FLIML	0.1452	0.1347	0.1032	0.0591	0.0341	0.0346
	MSE of C2SLS	0.1774	0.1581	0.1012	0.0546	0.0337	0.0308
	MSE of CFLIML	0.1422	0.1348	0.1034	0.0551	0.0353	0.0331
$\alpha^2_{21} = 0.60$	MSE of MC 2SLS	0.1081	0.0625	0.0577	0.0375	0.0294	0.0250
	MSE of MC FLIML	0.1055	0.1055	0.1085	0.0406	0.0336	0.0310
	MSE of C2SLS	0.1056	0.0610	0.0569	0.0398	0.0294	0.0272
	MSE of CFLIML	0.1056	0.0609	0.1063	0.0410	0.0323	0.0313

Continued on next page

Table B.3 – continued from previous page

		$T = 50$			$T = 100$		
		$L = 2$	$L = 4$	$L = 6$	$L = 2$	$L = 4$	$L = 6$
$\alpha_{31}^{2} = -0.38$	MSE of MC 2SLS	0.4005	0.1734	0.1813	0.1604	0.1238	0.1123
	MSE of MC FLIML	0.4043	0.2999	0.3119	0.1695	0.1379	0.1287
	MSE of C2SLS	0.4003	0.1641	0.1802	0.1409	0.1238	0.1107
	MSE of CFLIML	0.3935	0.1707	0.1811	0.1500	0.1238	0.1121
$\alpha_{11}^{3} = 0.65$	MSE of MC 2SLS	0.0940	0.0415	0.1167	0.0331	0.0214	0.0245
	MSE of MC FLIML	0.0928	0.0612	0.1369	0.0355	0.0235	0.0274
	MSE of C2SLS	0.0923	0.0404	0.1069	0.0328	0.0213	0.0270
	MSE of CFLIML	0.0939	0.0553	0.1165	0.0363	0.0209	0.0299
$\alpha_{21}^{3} = 1.20$	MSE of MC 2SLS	0.1317	0.1758	0.1044	0.0550	0.0348	0.0401
	MSE of MC FLIML	0.1306	0.1725	0.1277	0.0570	0.0379	0.0426
	MSE of C2SLS	0.1300	0.1752	0.0833	0.0545	0.0342	0.0417
	MSE of CFLIML	0.1307	0.1732	0.1041	0.0546	0.0361	0.0381
$\alpha_{31}^{3} = 0.38$	MSE of MC 2SLS	0.2642	0.1577	0.1408	0.1273	0.1057	0.0923
	MSE of MC FLIML	0.3152	0.2556	0.2438	0.1384	0.1205	0.1119
	MSE of C2SLS	0.2138	0.1483	0.1357	0.1264	0.1068	0.0920
	MSE of CFLIML	0.2453	0.1444	0.1400	0.1272	0.1132	0.1114
$\alpha_{11}^{4} = 0.50$	MSE of MC 2SLS	0.0592	0.0523	0.0434	0.0196	0.0124	0.0188
	MSE of MC FLIML	0.0680	0.0647	0.0856	0.0213	0.0141	0.0146
	MSE of C2SLS	0.0584	0.0522	0.0434	0.0203	0.0124	0.0136
	MSE of CFLIML	0.0592	0.00639	0.0448	0.0180	0.0120	0.0129
$\alpha_{21}^{4} = 0.60$	MSE of MC 2SLS	0.0631	0.0317	0.0791	0.0252	0.0203	0.0216
	MSE of MC FLIML	0.0690	0.0552	0.1099	0.0274	0.0228	0.0249
	MSE of C2SLS	0.0635	0.0301	0.0677	0.0260	0.0199	0.0234
	MSE of CFLIML	0.0701	0.0438	0.0850	0.0264	0.0203	0.0216
$\alpha_{31}^{4} = -0.20$	MSE of MC 2SLS	0.2089	0.1930	0.1031	0.0879	0.0663	0.0602
	MSE of MC FLIML	0.2472	0.2344	0.2041	0.0951	0.0746	0.0734
	MSE of C2SLS	0.2075	0.1852	0.1031	0.0820	0.0661	0.0585
	MSE of CFLIML	0.2003	0.1892	0.1500	0.0764	0.0663	0.0602

Continued on next page

Table B.3 – continued from previous page

		$T=50$			$T=100$		
		$L=2$	$L=4$	$L=6$	$L=2$	$L=4$	$L=6$
$c_{11}=1.00$	MSE of MC 2SLS	0.1111	0.0892	0.1125	0.0277	0.0138	0.0169
	MSE of MC FLIML	0.0932	0.0890	0.1076	0.0282	0.0140	0.0159
	MSE of C2SLS	0.0982	0.0901	0.1042	0.0254	0.0135	0.0137
	MSE of CFLIML	0.0980	0.0890	0.1038	0.0211	0.0145	0.0149
$c_{21}=0.60$	MSE of MC 2SLS	0.0214	0.0209	0.0177	0.0070	0.0036	0.0046
	MSE of MC FLIML	0.0154	0.0147	0.0148	0.0069	0.0035	0.0041
	MSE of C2SLS	0.0150	0.0203	0.0086	0.0069	0.0034	0.0038
	MSE of CFLIML	0.0161	0.0167	0.0152	0.0070	0.0037	0.0048
$c_{31}=-0.50$	MSE of MC 2SLS	0.0139	0.0160	0.0138	0.0049	0.0030	0.0030
	MSE of MC FLIML	0.0103	0.0114	0.0122	0.0050	0.0029	0.0027
	MSE of C2SLS	0.0097	0.0164	0.0115	0.0040	0.0026	0.0031
	MSE of CFLIML	0.0100	0.0131	0.0124	0.0046	0.0029	0.0023

Table B.3 presents the mean squared errors of the 17 target coefficients in four different estimators (the uncorrected two stage least square, the uncorrected Fuller limited information maximum likelihood, the corrected two stage lease square and the corrected Fuller limited information maximum likelihood estimators). The sample size is 50 and 100 respectively, and for the over-identification we choose three different cases ($L=2$, $L=4$ and $L=6$).

Appendix C

Appendix for Chapter 4

C.1 The evaluation of Theorem 3

$$\begin{aligned}\mathbb{E}(\hat{\alpha}_1 - \alpha_1) = \mathbb{E}\Big\{ & H^*\bar{Z}^{'}\tilde{v}_1 + H^*\tilde{W}^{*'}\tilde{v}_1 - H^*\bar{Z}^{'}\tilde{W}^*H^*\bar{Z}^{'}\tilde{v}_1 - H^*\tilde{W}^{*'}\bar{Z}H^*\bar{Z}^{'}\tilde{v}_1 \\ & - H^*\bar{Z}^{'}\tilde{W}^*H^*\tilde{W}^{*'}\tilde{v}_1 \\ & - H^*\tilde{W}^{*'}\bar{Z}H^*\tilde{W}^{*'}\tilde{v}_1 - H^*\left(\tilde{W^{*'}}\tilde{W^*} - \mathbb{E}(\tilde{W^{*'}}\tilde{W^*})\right)H^*\bar{Z}^{'}\tilde{v}_1 \\ & - H^*\left(\tilde{W^{*'}}\tilde{W^*} - \mathbb{E}(\tilde{W^{*'}}\tilde{W^*})\right)H^*\tilde{W}^{*'}\tilde{v}_1\Big\} + o(T^{-1}).\end{aligned} \tag{C.1}$$

Clearly, terms involving a product of an odd number of normal random variables have zero expectation and these items can be eliminated. Hence, we should evaluate the first four terms and the last term in equation (C.1). In section 4.3 we define $H^* = \left[\mathbb{E}(Z^{'}Z)\right]^{-1}$ and $\tilde{W}^* = \left[\sum_{i=1}^{p}\sum_{t=i}^{T-1} D^t\tilde{V}J_{t-i}\Psi_i^{'} : 0\right]$. Hence we can write the expectation for each term:

$$(i) \quad \mathbb{E}(H^*\bar{Z}^{'}\tilde{v}_1) = H^*\bar{Z}\mathbb{E}(\tilde{v}_1) = 0,$$

$$(ii) \quad \mathbb{E}(H^*\tilde{W}^{*'}\tilde{v}_1) = H^*\mathbb{E}(\tilde{W}^{*'}\tilde{v}_1).$$

The definition of $\tilde{W}^{*'}$ is from section 2.2, and following Nagar (1959) we decompose the reduced form disturbance matrix $\tilde{V}$ into two independent parts as:

$$\tilde{V} = \tilde{S} + \tilde{v}_1\tilde{\Omega}_{.1}^{'},$$

where $\tilde{\Omega}_{.1}$ is the first column of the reduced form disturbance's covariance matrix, $\tilde{S}$, and $\tilde{v}_1$ are independent, then we can rewrite $\tilde{W}^*$ as:

$$\tilde{W}^* = \left(\sum_{i=1}^{p}\sum_{t=i}^{T-1} D^t\tilde{S}J_{t-i}\Psi_i^{'} : 0\right) + \left(\sum_{i=1}^{p}\sum_{t=i}^{T-1} D^t\tilde{v}_1\tilde{\Omega}_{.1}^{'}J_{t-i}\Psi_i^{'} : 0\right). \tag{C.2}$$

Therefore for the second term (ii),the expectation becomes:

$$\begin{aligned}\mathbb{E}(\tilde{W}^{*'}\tilde{v}_1) &= \mathbb{E}\left\{\left(\sum_{i=1}^{p}\sum_{t=i}^{T-1} D^t\tilde{v}_1\tilde{\Omega}_{.1}^{'}J_{t-i}\Psi_i^{'} : 0\right)^{'}\tilde{v}_1\right\} \\ &= \mathbb{E}\left\{\sum_{i=1}^{p}\sum_{t=i}^{T-1}\Psi_i J_{t-i}^{'}\tilde{\Omega}_{.1}\tilde{v}_1^{'}D^{t'}\tilde{v}_1\right\} \\ &= \mathbb{E}\left\{\sum_{i=1}^{p}\sum_{t=i}^{T-1}\Psi_i J_{t-i}^{'}\tilde{\Omega}_{.1}E\left(\tilde{v}_1^{'}D^{t'}\tilde{v}_1\right)\right\},\end{aligned}$$

where $\mathbb{E}\left(\tilde{v}_1^{'}D^{t'}\tilde{v}_1\right) = 0, \quad t = 1,...,T-1$. Notice that the first part of equation (C.2) is independent of $\tilde{v}_1$, hence this element is eliminated in evaluation. It is obvious that our final result of (ii) is:

$$\mathbb{E}(H^*\tilde{W}^{*'}\tilde{v}_1) = 0$$

$$\begin{aligned}(iii) \quad \mathbb{E}(-H^*\bar{Z}^{'}\tilde{W}^*H^*\bar{Z}^{'}\tilde{v}_1) &= -H^*\bar{Z}^{'}\mathbb{E}(\tilde{W}^*H^*\bar{Z}^{'}\tilde{v}_1) \\ &= -H^*\bar{Z}^{'}\mathbb{E}\left\{\left(\sum_{i=1}^{p}\sum_{t=i}^{T-1} D^t\tilde{v}_1\tilde{\Omega}_{.1}^{'}J_{t-i}\Psi_i^{'} : 0\right)H^*\bar{Z}'\tilde{v}_1\right\}\end{aligned}$$

$$= -H^* \bar{Z}^{'} \sum_{i=1}^{p} \sum_{t=i}^{T-1} D^t \tilde{v}_1 \tilde{v}_1^{'} \bar{Z} H^* \left(I_P : 0\right)^{'} \Psi_i J_{t-i}^{\prime} \tilde{\Omega}_{.1},$$

where $P = \sum_{m=1}^{G} p(m)$ which is clarified in equation (4.2).

$$(iv) \quad \mathbb{E}(-H^* \tilde{W}^{*^{'}} \bar{Z} H^* \bar{Z}^{'} \tilde{v}_1) = -H^* \mathbb{E}(\tilde{W}^{*^{'}} \bar{Z} H^* \bar{Z}^{'} \tilde{v}_1)$$
$$= -H^* \mathbb{E}\left\{ \left(\sum_{i=1}^{p} \sum_{t=i}^{T-1} D^t \tilde{v}_1 \tilde{\Omega}_{.1}^{'} J_{t-i} \Psi_i^{'} : 0 \right)^{'} \bar{Z} H^* \bar{Z}^{\prime} \tilde{v}_1 \right\}$$
$$= -H^* tr \left(\sum_{i=1}^{p} \sum_{t=i}^{T-1} \bar{Z}^{'} D^t \bar{Z} H^* \right) \left(J_{t-i} \Psi_i^{'} : 0 \right)^{'} \tilde{\Omega}_{.1}.$$

Then, we add the first four items together which yields:

$$-H^* \sum_{i=1}^{p} \sum_{t=i}^{T-1} \Lambda_t^* \left(J_{t-i} \Psi_i^{'} : 0 \right)^{'} \tilde{\Omega}_{.1} \tag{C.3}$$

where $\Lambda_t^* = \bar{Z}^{'} D^t \bar{Z} H^* + tr \left(\bar{Z}^{'} D^t \bar{Z} H^* \right) I_{P+Q}$, and here $P = \sum_{m=1}^{G} p(m)$ and $Q = \sum_{n=1}^{K} q(n)$ which is clarified in equation (4.2).

The remaining term is:

$$(v) \quad \mathbb{E}\left\{ -H^* \left(\tilde{W^{*\prime}} \tilde{W^*} - \mathbb{E}(\tilde{W^{*\prime}} \tilde{W^*}) \right) H^* \tilde{W}^{*^{'}} \tilde{v}_1 \right\}$$
$$= -H^* \mathbb{E}\left\{ \left(\tilde{W^{*\prime}} \tilde{W^*} \right) H^* \tilde{W}^{*^{'}} \tilde{v}_1 \right\} + H^* \mathbb{E}(\tilde{W^{*\prime}} \tilde{W^*}) H^* \mathbb{E}(\tilde{W}^{*^{'}} \tilde{v}_1)$$
$$= -H^* \mathbb{E}\left\{ \left(\tilde{W^{*\prime}} \tilde{W^*} \right) H^* \tilde{W}^{*^{'}} \tilde{v}_1 \right\} \qquad since \quad \mathbb{E}(\tilde{W}^{*^{'}} \tilde{v}_1) = 0$$
$$= -H^* \mathbb{E}\left\{ \begin{bmatrix} \sum_{i=1}^{p} \sum_{t=i}^{T-1} \Psi_i J_{t-i}^{'} \tilde{V}^{\prime} D^{t^{'}} \sum_{l=1}^{p} \sum_{s=l}^{T-1} D^s \tilde{V} J_{s-l} \Psi_l^{'} & 0 \\ 0 & 0 \end{bmatrix} H^* \right.$$
$$\left. \times \begin{bmatrix} \sum_{b=1}^{p} \sum_{r=b}^{T-1} \Psi_b J_{r-b}^{'} \tilde{V}^{\prime} D^{r^{'}} \\ 0 \end{bmatrix} \tilde{v}_1 \right\}$$

$$= -H^*\mathbb{E}\left\{\begin{bmatrix} \sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{l=1}^{p}\sum_{s=l}^{T-1}\sum_{b=1}^{p}\sum_{r=b}^{T-1}\Psi_i J_{t-i}^{'}\tilde{V}'D^{t'}D^s\tilde{V}J_{s-l}\Psi_l^{'}H^{**}\Psi_b J_{r-b}^{'}\tilde{V}'D^{r'} \\ 0 \end{bmatrix} \times \tilde{v}_1\right\},$$

recalling that in section 2.3, we defined $H^{**} = I_2^{'}H^*I_2$ is the $P \times P$ leading submatrix of matrix H^*, where $I_2 = \begin{bmatrix} I_P \\ 0 \end{bmatrix}$ which is $(P+Q) \times P$ selection matrix.

Follow with Mikhail (1972) Lemma** 6**:

Suppose A, B and C are constant matrices of such dimensions that the various products exists:

$$\mathbb{E}\{\tilde{V}^{'}A\tilde{V}B\tilde{V}^{'}C\tilde{v}_1\} = \tilde{\Omega}B\tilde{\Omega}_{.1}trCtrA + \tilde{\Omega}B^{'}\tilde{\Omega}_{1.}tr(AC^{'}) + \tilde{\Omega}_{.1}tr(AC)tr(B'\tilde{\Omega}). \quad \text{(C.4)}$$

Hence, we can evaluate the expectation by using this lemma:

$$\mathbb{E}\left\{\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{l=1}^{p}\sum_{s=l}^{T-1}\sum_{b=1}^{p}\sum_{r=b}^{T-1}\Psi_i J_{t-i}^{'}\tilde{V}'D^{t'}D^s\tilde{V}J_{s-l}\Psi_l^{'}H^{**}\Psi_b J_{r-b}^{'}\tilde{V}'D^{r'}\tilde{v}_1\right\}$$
$$= \sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{l=1}^{p}\sum_{s=l}^{T-1}\sum_{b=1}^{p}\sum_{r=b}^{T-1}\Psi_i J_{t-i}^{'}\tilde{\Omega}J_{s-l}\Psi_l H^{**}\Psi_b J_{r-b}^{'}\tilde{\Omega}_{.1}tr\left\{D^{r'}\right\}tr\left\{D^{t'}D^s\right\}$$
$$+\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{l=1}^{p}\sum_{s=l}^{T-1}\sum_{b=1}^{p}\sum_{r=b}^{T-1}\Psi_i J_{t-i}^{'}\tilde{\Omega}J_{r-l}\Psi_l^{'}H^{**}\Psi_b J_{s-b}^{'}\tilde{\Omega}_{1.}tr\left\{D^{t'}D^sD^r\right\}$$
$$+\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{l=1}^{p}\sum_{s=l}^{T-1}\sum_{b=1}^{p}\sum_{r=b}^{T-1}\Psi_i J_{t-i}^{'}\tilde{\Omega}_{.1}tr\left\{D^{t'}D^sD^{r'}\right\}tr\left\{J_{r-l}\Psi_l^{'}H^{**}\Psi_b J_{s-b}^{'}\tilde{\Omega}\right\}.$$

The first term equals to zero, since $tr\left\{D^{r'}\right\} = tr\left\{D^{r}\right\} = 0$. The trace in the second term would be:

$$tr\left\{D^{t'}D^{s}D^{r}\right\} = \begin{cases} 0, & t \neq s+r \\ T-t, & t = s+r \end{cases}.$$

The trace in the third term becomes:

$$tr\left\{D^{t'}D^{s}D^{r'}\right\} = \begin{cases} 0, & s \neq t+r \\ T-s, & s = t+r \end{cases}.$$

Hence, finally, it has been shown that:

$$\begin{aligned} & -H^{*}\mathbb{E}\left\{\left(\tilde{W^{*'}}\tilde{W^{*}}\right)H^{*}\tilde{W}^{*'}\tilde{v}_1\right\} \qquad \text{(C.5)} \\ &= -H^{*}\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{l=1}^{p}\sum_{s=l}^{T-1}\sum_{b=1}^{p}\sum_{r=b}^{T-1}(T-s)\Psi_l J_{s-l}^{'}\tilde{\Omega}J_{r-b}\Psi_b^{'}H^{**}\Psi_i J_{t-i}^{'}\tilde{\Omega}_{.1} \\ &\quad -H^{*}\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{l=1}^{p}\sum_{s=l}^{T-1}\sum_{b=1}^{p}\sum_{r=b}^{T-1}(T-s)tr\left\{\Psi_l J_{s-l}^{'}\tilde{\Omega}J_{r-b}e_b^{'}H^{**}\right\}\Psi_i J_{t-i}^{'}\tilde{\Omega}_{.1} \\ &= -H^{*}\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{l=1}^{p}\sum_{r=l}^{T-1}\sum_{b=i+l}^{p}\sum_{s=t+r}^{T-1}(T-s)(J_{s-l}\Psi_l^{'}:0)^{'}\tilde{\Omega}J_{r-b}\Psi_b^{'}H^{**}\Psi_i J_{t-i}^{'}\tilde{\Omega}_{.1} \\ &\quad -H^{*}\sum_{i=1}^{p}\sum_{t=i}^{T-1}\sum_{l=1}^{p}\sum_{r=l}^{T-1}\sum_{b=i+l}^{p}\sum_{s=t+r}^{T-1}(T-s)tr\left\{(\Psi_l J_{s-l}^{'}\tilde{\Omega}J_{r-b}\Psi_b^{'}H^{**}\right\}(J_{t-i}\Psi_i^{'}:0)^{'}\tilde{\Omega}_{.1}. \end{aligned}$$

Collecting equation (C.3) and (C.5), Theorem 3 is proved. Notice that in section 4.3 we define $H^{*} = \left[E(Z^{'}Z)\right]^{-1}$. It can be evaluated as follows. We start from

$\bar{Z}'\bar{Z}+E(\tilde{W^{*\prime}}\tilde{W^*})$, then:

$$\mathbb{E}(\tilde{W^{*\prime}}\tilde{W^*}) = \mathbb{E}\left\{\begin{bmatrix} \sum_{i=1}^{p}\sum_{t=i}^{T-1}\Psi_i J_{t-i}^{'}\tilde{V}'D^{t'}\sum_{l=1}^{p}\sum_{s=l}^{T-1}D^s\tilde{V}J_{s-l}\Psi_l^{'} & 0 \\ 0 & 0 \end{bmatrix}\right\}.$$

Using Mikhail (1972) lemma 6,we can write:

$$\mathbb{E}\left(\tilde{V}'D^{t'}D^s\tilde{V}\right) = \tilde{\Omega}tr\left\{D^{t'}D^s\right\} = \begin{cases} 0, & t \neq s \\ (T-t)\tilde{\Omega}, & t = s. \end{cases}$$

Hence,

$$H^* = \left\{\bar{Z}'\bar{Z} + \sum_{i=1}^{p}\sum_{l=1}^{p}\sum_{t=i,l}^{T-1}(T-t)\begin{bmatrix} \Psi_i J_{t-i}^{'}\tilde{\Omega}J_{t-l}\Psi_l^{'} & 0 \\ 0 & 0 \end{bmatrix}\right\}^{-1}.$$

C.2 Numerical Results

Table C.1 Bias approximation, MC OLS bias, MC COLS bias and BOLS; T=50, 100

		$T=50$	$T=100$
	Bias Approximation	-0.0341(-19%)	-0.0142 (8%)
$\gamma_{11}^{1}=0.1774$	MC OLS	-0.0330 (-19%)	-0.0146 (-8%)
	MC COLS	-0.0071(-4%)	-0.0020 (-1%)
	BOLS	-0.0107(-6%)	-0.0020 (-1%)
	Bias Approximation	-0.0387(150 %)	-0.0211 (82%)
$\gamma_{21}^{1}=0.0258$	MC OLS	-0.0375(-145%)	-0.0198 (-77%)
	MC COLS	-0.0041(-16%)	-0.0010 (-4%)
	BOLS	-0.0122(-47%)	-0.0023 (-9%)
	Bias Approximation	-0.0301(-25 %)	0.0001 (0%)
$\gamma_{31}^{1}=-0.1177$	MC OLS	-0.0254(-22%)	0.0004 (0%)
	MC COLS	-0.0032(-3%)	-0.0005 (-1%)
	BOLS	-0.0109(-9%)	-0.0007 (-1%)
	Bias Approximation	-0.0109(-24%)	-0.0055 (-12%)
$\gamma_{11}^{2}=0.0454$	MC OLS	-0.0100(-22%)	-0.0053(-12%)
	MC COLS	-0.0005(-1%)	-0.0007(-2%)
	BOLS	-0.0050(-11%)	-0.0007(-2%)
	Bias Approximation	-0.0119(-7 %)	-0.0041 (-3%)
$\gamma_{21}^{2}=0.1626$	MC OLS	-0.0132 (-8%)	-0.0056 (-3%)
	MC COLS	-0.0027(-2%)	-0.0009 (-0%)
	BOLS	-0.0062(-4%)	-0.0015 (-1%)
	Bias Approximation	-0.0518(-67 %)	-0.0251 (-33%)
$\gamma_{31}^{2}=-0.0768$	MC OLS	-0.0480 (-62%)	-0.0232 (-30%)
	MC COLS	-0.0097(-13%)	-0.0010 (-1%)
	BOLS	-0.0193(-25%)	-0.0041 (-5%)
	Bias Approximation	-0.0061(-15 %)	-0.0015 (-4%)
$\gamma_{11}^{3}=-0.0397$	MC OLS	-0.0056 (-14%)	-0.0012 (-3%)
	MC COLS	-0.0010(3%)	-0.0004 (-1%)
	BOLS	-0.0014(-4%)	-0.0002 (-1%)

Continued on next page

Table C.1 – continued from previous page

		$T=50$	$T=100$
	Bias Approximation	-0.0135(-5%)	-0.0043 (-2%)
$\gamma_{21}^3 = 0.2487$	MC OLS	-0.0100 (-4%)	-0.0039 (-2%)
	MC COLS	-0.0042(-2%)	-0.0011 (-0%)
	BOLS	-0.0038(-2%)	-0.0004 (-0%)
	Bias Approximation	0.0172(+5%)	0.0023 (+1%)
$\gamma_{31}^3 = 0.3409$	MC OLS	0.0132 (+4%)	0.0021 (+1%)
	MC COLS	0.0031(+1%)	-0.0002 (-0%)
	BOLS	0.0037(+1%)	-0.0002 (-0%)
	Bias Approximation	-0.0293(-77 %)	-0.0135 (-36%)
$\gamma_{11}^4 = 0.0371$	MC OLS	-0.0235 (-63%)	-0.0126 (-34%)
	MC COLS	-0.0012(-3%)	-0.0020 (-6%)
	BOLS	-0.0093(-25%)	-0.0024 (-6%)
	Bias Approximation	0.0041(+2 %)	-0.0023 (+1%)
$\gamma_{21}^4 = 0.1751$	MC OLS	0.0033 (+1%)	0.0022 (+1%)
	MC COLS	0.0017(+1%)	-0.0000 (+0%)
	BOLS	0.0010(+1%)	-0.0000 (+0%)
	Bias Approximation	0.0206(+13 %)	0.0088 (+6%)
$\gamma_{31}^4 = 0.1584$	MC OLS	0.0122 (+8%)	0.0088 (+6%)
	MC COLS	0.0029(+2%)	0.0017 (+1%)
	BOLS	0.0060(+4%)	0.0016 (+1%)
	Bias Approximation	-0.0.0301(-37 %)	-0.0139 (-17%)
$\pi_{11} = -0.0806$	MC OLS	-0.0233 (-29%)	-0.0101 (-13%)
	MC COLS	-0.0007(-1%)	-0.0010 (-1%)
	BOLS	-0.0091(-11%)	-0.0019 (-2%)
	Bias Approximation	-0.0041(-2 %)	-0.0020 (-1%)
$\pi_{21} = 0.1697$	MC OLS	-0.0034 (-2%)	-0.0014 (-1%)
	MC COLS	-0.0009(-0%)	-0.0001 (-0%)
	BOLS	-0.0013(-1%)	-0.0003(-0%)

Continued on next page

Table C.1 – continued from previous page

		$T = 50$	$T = 100$
	Bias Approximation	0.0030(+2 %)	0.0006 (+0%)
$\pi_{31} = -0.1414$	MC OLS	0.0023 (2%)	0.0006 (+0%)
	MC COLS	0.0015(+1%)	-0.0001 (-0%)
	BOLS	0.0010(+1%)	0.0001(+0%)
	Bias Approximation	0.0000(+0 %)	0.0001 (0%)
$\pi_{41} = 0.1482$	MC OLS	-0.0000 (-0%)	0.0012 (+1%)
	MC COLS	0.0003(+0%)	-0.0003 (-0%)
	BOLS	0.0001(+0%)	0.0002 (+0%)
	Bias Approximation	0.0009(+2 %)	-0.0004 (-1%)
$\pi_{51} = -0.0474$	MC OLS	0.0005 (+1%)	-0.0004 (-1%)
	MC COLS	0.0001(+0%)	0.0000 (+0%)
	BOLS	0.0003(+1%)	-0.0001 (-0%)
	Bias Approximation	-0.0029(-12 %)	-0.0010 (-4%)
$\pi_{61} = -0.0249$	MC OLS	-0.0032 (-14%)	-0.0006 (-2%)
	MC COLS	-0.0007(-3%)	+0.0003 (+1%)
	BOLS	-0.0011(-4%)	-0.0001 (-0%)
	Bias Approximation	0.0102(+7 %)	0.0059 (+4%)
$\pi_{71} = 0.1427$	MC OLS	0.0086 (+6%)	0.0043 (+3%)
	MC COLS	-0.0013(-1%)	0.0010 (+0%)
	BOLS	0.0031(+2%)	0.0007 (+0%)

Table C.1 presents the bias approximation of the 19 first reduced form coefficients (the over-identification level of the structural form is $L = 2$) in the ordinary least square estimator. It also reports the bias of the Monte Carlo ordinary least square estimator, the bias of the corrected ordinary least square estimator, and the residual bootstrap ordinary least square estimator. The sample size is 50 and 100 respectively.

Table C.2 The MSE of OLS, COLS, BOLS; T=50, 100

		$T = 50$	$T = 100$
$\gamma_{11}^{1} = 0.1774$	MSE of MC OLS	0.0072	0.0026
	MSE of MC COLS	0.0060	0.0020
	MSE of BOLS	0.0068	0.0025
$\gamma_{21}^{1} = 0.0258$	MSE of MC OLS	0.0129	0.0054
	MSE of MC COLS	0.0120	0.0049
	MSE of BOLS	0.0125	0.0052
$\gamma_{31}^{1} = -0.1177$	MSE of MC OLS	0.0750	0.0286
	MSE of MC COLS	0.0759	0.0279
	MSE of BOLS	0.0771	0.0294
$\gamma_{11}^{2} = 0.0454$	MSE of MC OLS	0.0049	0.0024
	MSE of MC COLS	0.0049	0.0022
	MSE of BOLS	0.0051	0.0025
$\gamma_{21}^{2} = 0.1626$	MSE of MC OLS	0.0122	0.0047
	MSE of MC COLS	0.0117	0.0050
	MSE of BOLS	0.0129	0.0049
$\gamma_{31}^{2} = -0.0768$	MSE of MC OLS	0.0604	0.0223
	MSE of MC COLS	0.0521	0.0223
	MSE of BOLS	0.0614	0.025
$\gamma_{11}^{3} = -0.0397$	MSE of MC OLS	0.0067	0.0026
	MSE of MC COLS	0.0068	0.0021
	MSE of BOLS	0.0072	0.0027
$\gamma_{21}^{3} = 0.2487$	MSE of MC OLS	0.0110	0.0048
	MSE of MC COLS	0.0100	0.0043
	MSE of BOLS	0.0119	0.0051
$\gamma_{31}^{3} = 0.3409$	MSE of MC OLS	0.0484	0.0188
	MSE of MC COLS	0.0472	0.0190
	MSE of BOLS	0.0497	0.0194
$\gamma_{11}^{4} = 0.0371$	MSE of MC OLS	0.0057	0.0023
	MSE of MC COLS	0.0055	0.0020
	MSE of BOLS	0.0055	0.0022

Continued on next page

Table C.2 – continued from previous page

		$T = 50$	$T = 100$
$\gamma_{21}^{4} = 0.1751$	MSE of MC OLS	0.0091	0.0039
	MSE of MC COLS	0.0093	0.0031
	MSE of BOLS	0.0095	0.0040
$\gamma_{31}^{4} = 0.1584$	MSE of MC OLS	0.0542	0.0144
	MSE of MC COLS	0.0493	0.0137
	MSE of BOLS	0.0545	0.0144
$\pi_{11} = -0.0806$	MSE of MC OLS	0.0053	0.0016
	MSE of MC COLS	0.0051	0.0015
	MSE of BOLS	0.0049	0.0015
$\pi_{21} = 0.1697$	MSE of MC OLS	0.0003	0.0001
	MSE of MC COLS	0.0002	0.0002
	MSE of BOLS	0.0003	0.0001
$\pi_{31} = -0.1414$	MSE of MC OLS	0.0002	0.0001
	MSE of MC COLS	0.0002	0.0001
	MSE of BOLS	0.0002	0.0001
$\pi_{41} = 0.1482$	MSE of MC OLS	0.0005	0.0002
	MSE of MC COLS	0.0003	0.0002
	MSE of BOLS	0.0005	0.0002
$\pi_{51} = -0.0474$	MSE of MC OLS	0.0003	0.0001
	MSE of MC COLS	0.0003	0.0001
	MSE of BOLS	0.0003	0.0001
$\pi_{61} = -0.0249$	MSE of MC OLS	0.0005	0.0001
	MSE of MC COLS	0.0005	0.0001
	MSE of BOLS	0.0005	0.0001
$\pi_{71} = 0.1427$	MSE of MC OLS	0.0006	0.0001
	MSE of MC COLS	0.0006	0.0001
	MSE of BOLS	0.0006	0.0001

Table C.2 presents the mean squared errors of the 19 target coefficients in least squares on three occasions (the uncorrected ordinary least square estimator, the corrected ordinary least square estimator, and the residual bootstrap ordinary least square estimator) in the *pth* order reduced form. The sample size is 50 and 100 respectively.

Appendix D

Appendix for the Note table

D.1 Other Experiments for 2SLS

Table D.1 **Percentages of the bias of 2SLS estimation, when L= 4, 6; T=50, 70, 90, 100**

	$L=4$				$L=6$			
	$T=50$	$T=70$	$T=90$	$T=100$	$T=50$	$T=70$	$T=90$	$T=100$
$\alpha_{21}^{1}=0.36$	-3%	8%	6%	5%	4%	8%	8%	8%
$\alpha_{31}^{2}=-0.38$	-	-	-	-	11%	25%	21%	19%
$\alpha_{31}^{4}=-0.20$	-4%	51%	18%	7%	-	-	-	-

Table D.1 presents the trend of some related coefficients in certain cases. The over-identification level is $L=4,6$. The sample size is 50, 70, 90 and 100 respectively. Note that when $L=4$, the bias of α_{31}^{2} decreases when sample size increases and so does α_{31}^{4}, when $L=6$. Hence these two are not in this abnormal case.

Appendix E

Appendix for the Programming

Main Script

```
%   Main


%MAX_SAMPLE_SIZE = 1000;
%matlabpool close force
%matlabpool open
clear all;
load 'exgenous.mat'

config.model(1).Bi      = [  1.000   2.000   5.000 ;
                             1.110   1.000   8.000;
                             3.000    4.600  1.000]';

config.model(1).Ai1      = [   0.500    0.360   0.400 ;
                              0.560    0.620   0.900 ;
                             0.450   0.280   0.320 ]';
```

```
config.model(1).Ai2       =  [ 1.200  0.600 0.380;
                               0.800  0.720  0.500;
                               0.820  0.900 0.780]';

config.model(1).Ai3       =  [ 0.65  1.200  0.380;
                               0.46   0.720 0.560;
                               0.800 0.310 0.740]';

config.model(1).Ai4       =  [ 0.500  0.600 0.200;
                               0.360   0.460 0.500;
                               0.200 0.580 0.700]';

config.model(1).Ci       = [  1.000   0.600   0.5   0        0      0
0      0.00    0.00   0.00  0.00;
                            1.000    0       0    0.750   0.24    0.35
0.68   0       0.00   0.00  0.00;
                            1.000   0       0      0       0       0.00
0.00  0.15   0.86    0.58 0.33]';


% config.model(1).Ci       = [  1.000   0.600   0.5   0        0      0
0.00    0.00  0.00   ;
%                             1.000    0       0    0.750   0.24    0.35
0.00   0.00   0.00   ;
%                             1.000   0        0      0       0       0.00
 0.15   0.86   0.58 ]';

%  config.model(1).Ci       = [  1.000   0.600   0.5   0        0
0.00  0.00   ;
%                              1.000    0       0    0.750   0.24
0.00  0.00   ;
```

```
%                               1.000    0         0       0        0
 0.15    0.86    ]';
config.model(1).Sigma    =  [   0.3524    0.3448   0.3112;
                             0.3448    0.3668   0.2984;
                             0.3112    0.2984   0.4064];



config.eq                = 1;
config.replications      = 20000;
config.bst               = 199;
config.X                 = X_N_M0_V1_B09;
%Exgenous, Normal Distribution, Mean=0,  Variance=1, AR(1) Beta=0.9

config.constT            = 1;
config.dList             = [1];            % 1=Normal; 2=Uniform
config.tList      = [100];
%config.tlist      = [50]

config.eList=[...
              10001 ...
              10002 ...
              10003 ...
              30011 ...
              30021 ...
              30012 ...
               ];
doMonteCarlosSimulation(config);

%matlabpool close
```

Monte Carlo Experiment

```
%   domontecarlosimulation

function doMonteCarlosSimulation (config)

    replications    = config.replications;  % number of replications

    eq              = config.eq;            % esitmated equation

    bst             = config.bst;           % number of bootstrapping

    constT          = config.constT;        % with/without constant term


    mListLen = length(config.model);

    tListLen = length(config.tList);

    dListLen = length(config.dList);

    eListLen = length(config.eList);
```

```
eList     = config.eList;

O = zeros(mListLen,eListLen,tListLen,dListLen,17,14);



%R1 = zeros(eListLen,1,r);



%** LAYER 1: Model*********************************

for mptr = 1:mListLen

    Bi     = config.model(mptr).Bi;

    Ci        = config.model(mptr).Ci;

    Ai1   = config.model(mptr).Ai1;

    Ai2   = config.model(mptr).Ai2;

    Ai3   = config.model(mptr).Ai3;

    Ai4   = config.model(mptr).Ai4;



    Sigma    = config.model(mptr).Sigma;

    Gamma1    =  1 * Ai1/Bi;
```

```
Gamma2   =  1 * Ai2/Bi;

Gamma3   =  1 * Ai3/Bi;

Gamma4   =  1 * Ai4/Bi;

Pi      =  1 * Ci/Bi;

Omega   = Bi'\Sigma/Bi;



EigVal  = polyeig(Bi, Ai1, Ai2, Ai3, Ai4);

rho     = chol(Omega);

% Get included (endogenous/lagged endogenous/exgenous) index

indX        = find(Ci(:,eq) ~= 0);





indY        = find(Bi(:,eq) ~= 0);

indY(eq,:)  = [];



indLY       = find(Ai1(:,eq) ~= 0);
```

```
indL2Y      = find(Ai2(:,eq) ~= 0);



indL3Y      = find(Ai3(:,eq) ~= 0);



indL4Y      = find(Ai4(:,eq) ~= 0);

% Ignore the index of (Ci) with row elements all equal to zero

inxX0   = find(sum(abs(Ci),2) ~= 0);

exdX    = find(Ci(:,eq) == 0);

exdX    = intersect(exdX, inxX0);






% Ignore the index of (Bi) with row elements all equal to zero

inxYO   = find(sum(abs(Bi),2) ~= 0);

exdY    = find(Bi(:,eq) == 0);

exdY    = intersect(exdY, inxYO);
```

```
% Ignore the index of (Ai) with row elements all equal to zero

inxLYO  = find(sum(abs(Ai1),2) ~= 0);

exdLY   = find(Ai1(:,eq) == 0);

exdLY   = intersect(exdLY, inxLYO);



inxL2YO  = find(sum(abs(Ai2),2) ~= 0);

exdL2Y   = find(Ai2(:,eq) == 0);

exdL2Y   = intersect(exdL2Y, inxL2YO);



inxL3YO  = find(sum(abs(Ai3),2) ~= 0);

exdL3Y   = find(Ai3(:,eq) == 0);

exdL3Y   = intersect(exdL3Y, inxL3YO);



inxL4YO  = find(sum(abs(Ai4),2) ~= 0);

exdL4Y   = find(Ai4(:,eq) == 0);
```

```
exdL4Y   = intersect(exdL4Y, inxL4YO);



% Coefficients(True value) of estimated equation

delta1      =  1*[Bi(indY, eq); Ai1(indLY, eq);Ai2(indL2Y,
 eq);Ai3(indL3Y, eq);Ai4(indL4Y, eq); Ci(indX, eq)];

delta1Len   = length(delta1);



K = length([indX; exdX ]);

G = length([indY; exdY]) + 1;

J = length([indLY; exdLY; indL2Y; exdL2Y; indL3Y; exdL3Y; indL4Y;
 exdL4Y]);

k = length(indX );

g = length(indY);

j = length([indLY; indL2Y; indL3Y; indL4Y]  );

L = (K + J)   (k + j)   g;
```

```
%** LAYER 2: Sample Size ******************************************

for tptr = 1:tListLen

    T = config.tList(tptr);



    rw = 200;             % runway length







    if constT == 1

        cX = config.X(rw+1:T+rw,1:K);

        % below is for calculation of expected inital value of Y

        mX=mean(cX);

        ftX = config.X(rw T+1:rw,1:K);

    else

        cX = config.X(rw+1:T+rw,2:K+1);
```

```
    % below is for calculation of expected inital value of Y

     mX=mean(cX);

     ftX = config.X(rw T+1:rw,2:K+1);

end




X = cX;

X1 = cX(:,indX);

X2 = cX(:,exdX);










%** LAYER 3: Distributions ************************************

for dptr = 1:dListLen

    dist = config.dList(dptr);
```

```
% Below is for calculation of expected starting value of Y
  % instead of setting as zero or other arbitrary choice.
 Iy=eye(3);


 mY=(mX*Pi)/(Iy Gamma1 Gamma2 Gamma3 Gamma4);

 l3my0 =mY;

 l2my0 =mY;

 lmy0 =mY;

 my0 =mY;

 y01=zeros(T,G);

 for i=1:10000

[temp11, ~, ~, ~ , ~ ] = genY(l3my0,l2my0,lmy0,my0, ftX, rho,
 Gamma1, Gamma2, Gamma3, Gamma4,  Pi, T, G, dist) ;


y01 =y01+temp11;

 end
```

```
y0=y01(T,:)/10000;

ly0 =y01(T1,:)/10000;

l2y0 =y01(T2,:)/10000;

l3y0 =y01(T3,:)/10000;

%** LAYER 4: Simulation ***************************************
```

```
V = zeros(delta1Len,eListLen,replications);

R = zeros(2,eListLen,replications);

tic

h=waitbar(0, 'please wait...');

for rptr = 1:replications

   waitbar(rptr/replications)

    [Y, LY, L2Y,  L3Y,  L4Y ] = genY(l3y0,l2y0,ly0,y0, X,
     rho, Gamma1, Gamma2, Gamma3, Gamma4,  Pi, T, G, dist);

    y            = Y(:,eq);

    Y2           = Y;

    Y2(:,eq)     = [];

    LY1          = LY(:,indLY);

    L2Y1         = L2Y(:,indL2Y);

    L3Y1         = L3Y(:,indL3Y);

    L4Y1         = L4Y(:,indL4Y);
```

```
LY1X1       = [LY1 L2Y1 L3Y1 L4Y1 X1];

LYX         = [LY(:,[indLY exdLY]) L2Y(:,[indL2Y
 exdL2Y]) L3Y(:,[indL3Y exdL3Y]) L4Y(:,[indL4Y exdL4Y])
  X];



for eptr = 1:eListLen



    switch eList(eptr)



        case 10001;   [V(:,eptr,rptr), R(:,eptr,rptr) ]
         =c2SLS( y, Y2,LY1X1, LYX);



       case 10002;   [V(:,eptr,rptr), R(:,eptr,rptr) ]
       =Cc2SLS( y,Y, Y2,LY1X1, LYX,l3y0, l2y0, ly0 ,y0,
        X, X1, G, g, K, k, j, indY,
         indLY,indL2Y,indL3Y,indL4Y, T)



      case 10003;  [V(:,eptr,rptr), R(:,eptr,rptr)] =
       c2SLS_boot( y0,ly0,l2y0,l3y0,y,Y, Y2,LY1X1, LYX,
        ftX, X,mX,Iy, X1, indLY, exdLY, indL2Y, exdL2Y,
         indL3Y, exdL3Y, indL4Y, exdL4Y, T,g,G, bst );
```

```
                    case 30011;  [V(:,eptr,rptr), R(:,eptr,rptr)]
                     =fuller(y, Y2, LYX, LY1X1, T, 1/(T K J));

                    case 30021;  [V(:,eptr,rptr), R(:,eptr,rptr)]
                     =fuller(y, Y2, LYX, LY1X1, T, 4/(T K J));

                    case 30012;   [V(:,eptr,rptr), R(:,eptr,rptr)]
                     =cfuller(cfuller(y,Y, Y2,LY1X1, LYX,l3y0, l2y0,
                      ly0 ,y0, X, X1, G, g, K, k, j, indY,
                       indLY,indL2Y,indL3Y,indL4Y, T, adj), 1/(T K));




                      otherwise;

                  end

              end

          end

          close(h)

          toc



      a  = getApproximation(l3y0, l2y0, ly0 ,y0,ftX, X, X1,
       Gamma1,Gamma2,Gamma3,Gamma4,Bi, Omega,Sigma,
```

```
  Pi, G, g, K, k,
  J,j, indY, indLY,indL2Y,indL3Y,indL4Y,eList, T,
   eq);



O(mptr,:,tptr,dptr,:,:) = getInfo(V,R,delta1);



O(mptr,:,tptr,dptr,:,2) = a;



ahat = reshape(O(mptr,:,tptr,dptr,:,4),
eListLen,delta1Len);




O(mptr,:,tptr,dptr,:,3) = ((a   ahat)./abs(ahat))*100;



B = reshape(O(mptr,:,tptr,dptr,:,:),eListLen,delta1Len,14);



printResult(mptr, eq, T, replications, bst, constT,
dist, L, Bi, Ai1, Ai2, Ai3,Ai4, Ci, Gamma1,
Gamma2,Gamma3, Gamma4, Pi, Sigma,
```

```
                  Omega, EigVal, B, eList)



         end

        end

    end



    date_string = datestr(now(), 'yyyymmdd_HHMMSS');

    savefile = ['output' ' (' date_string ').mat'];

    save(savefile, 'O');





end



function A = getInfo(V, R, delta1)

    eListLen = size(V(:,:,1),2);

    cListLen = size(V(:,:,1),1);
```

```
% Total 13 things to be reported.

A = zeros(eListLen,cListLen, 14);



for eptr = 1:eListLen

    for cptr = 1:cListLen

        tmp = V(cptr,eptr,:);

        A(eptr,cptr,1)    = delta1(cptr);

        A(eptr,cptr,4)    = mean(tmp)   delta1(cptr);

        A(eptr,cptr,5)    = ((mean(tmp)
         delta1(cptr))/abs(delta1(cptr)))*100;

        A(eptr,cptr,6)    = std(tmp);

        A(eptr,cptr,7)    = max(tmp);

        A(eptr,cptr,8)    = min(tmp);

        A(eptr,cptr,9)    = median(tmp)   delta1(cptr);

        A(eptr,cptr,10)    = iqr(tmp);

        A(eptr,cptr,11)   = mean((tmp delta1(cptr)).^2);
```

```
            A(eptr,cptr,12)    = mean(R(1,eptr,:));

            A(eptr,cptr,13)    = mean(R(2,eptr,:));

            A(eptr,cptr,14)    = var(tmp);

        end

    end

end



function [b, I] = c2SLS( y, Y2,LY1X1, LYX)



       %Y = [y Y2];





       %GP_hat = LYX\Y;





%         Gamma_hat = GP_hat(1:G,:);
```

```
%         Pi_hat = GP_hat(G+1:end,:)



       V2 = Y2   LYX*(LYX\Y2);

%         V = Y   LYX*GP_hat;

%         V2 = V(:,2:end);



       UL = Y2'*Y2 V2'*V2;

       UR = Y2'*LY1X1;

       LL = UR';

       LR = LY1X1'*LY1X1;



       b = ([UL UR; LL LR])\([(Y2  V2)'*y; LY1X1'*y]);
```

```
       I (1)= 0;

       %I = 1   (V2'*V2)/(Y2'*M*Y2);


end


function [b, I] = Cc2SLS( y,Y, Y2,LY1X1, LYX,l3y0, l2y0, ly0 ,y0, X, X1, G, g,
 K, k, j, indY, indLY,indL2Y,indL3Y,indL4Y, T)


         [b, I] = c2SLS( y, Y2,LY1X1, LYX);

         GP_hat = LYX\Y;

         Gamma1_hat=GP_hat(1:G,:);

         Gamma2_hat=GP_hat(G+1:G+G,:);

         Gamma3_hat=GP_hat(G+G+1:G+G+G,:);

         Gamma4_hat=GP_hat(G+G+G+1:j,:);

         Pi_hat=GP_hat(j+1:end,:);

         V = Y   LYX*GP_hat;
```

```
R = [Y2 LY1X1];

phi = V'*(y   R*b)/T;

Omega_hat = V'*V;



a = Dynamic_2SLS_LT_bias_approximation(l3y0, l2y0, ly0 ,y0, X, X1,
 Gamma1_hat,Gamma2_hat,Gamma3_hat,Gamma4_hat, Omega_hat,phi, Pi_hat, G,
  g, K, k,j, indY, indLY,indL2Y,indL3Y,indL4Y, T);



 b = b a;



I (1)= 0;

%I = 1   (V2'*V2)/(Y2'*M*Y2);


end
```

```
%% Fuller

function [b, I] = fuller(y, Y2, LYX, LY1X1, T, adj)

    Yd = [y Y2];

    YdtYd = Yd'*Yd;

    %Wsdd = Yd' * Yd   Yd' * X1 * inv(X1' * X1) * X1' * Yd; %#ok<MINV>

    Wsdd = YdtYd   ( Yd' *LY1X1/(LY1X1' * LY1X1)*LY1X1' * Yd );



    %Wdd = Yd' * Yd   Yd' * X * inv(X' * X) * X' * Yd;

    Wdd = YdtYd   ( Yd' *LYX/(LYX' * LYX) * LYX' * Yd );



    %lambda = min(eig(inv(Wdd) * Wsdd));

    lambda = min(eig( Wsdd/Wdd));



    lambda = lambda   adj;
```

```
    [b, I] = kClass(y, Y2, LYX, LY1X1, lambda, T);

end



function [b, I] = cfuller(y,Y, Y2,LY1X1, LYX,l3y0, l2y0, ly0 ,y0, X, X1, G, g,
 K, k, j, indY, indLY,indL2Y,indL3Y,indL4Y, T, adj)

    [b, I] = cfuller(y, Y2, LYX, LY1X1, T, adj);

    GP_hat = LYX\Y;

         Gamma1_hat=GP_hat(1:G,:);

         Gamma2_hat=GP_hat(G+1:G+G,:);

         Gamma3_hat=GP_hat(G+G+1:G+G+G,:);

         Gamma4_hat=GP_hat(G+G+G+1:j,:);

         Pi_hat=GP_hat(j+1:end,:);

         V = Y   LYX*GP_hat;

         R = [Y2 LY1X1];

         phi = V'*(y   R*b)/T;

         Omega_hat = V'*V;

         af = Dynamic_Full_LT_bias_approximation(l3y0, l2y0, ly0 ,y0, X, X1,
```

```
        Gamma1_hat,Gamma2_hat,Gamma3_hat,Gamma4_hat, Omega_hat,phi, Pi_hat, G,
         g, K, k,j, indY, indLY,indL2Y,indL3Y,indL4Y, T);




         b = b a;




       I (1)= 0;



end



function [b, I] = c2SLS_boot( y0,ly0,l2y0,l3y0,y,Y, Y2,LY1X1, LYX,ftX,
 X,mX,Iy,X1, indLY, exdLY, indL2Y, exdL2Y, indL3Y, exdL3Y, indL4Y, exdL4Y,
  T,g,G, bst )

    [b1, I1] = c2SLS( y, Y2,LY1X1, LYX);

     GP_hat=LYX\Y;

     Gamma1_hat= GP_hat(1:G,:);

     Gamma2_hat= GP_hat(G+1:G+G,:);
```

```
Gamma3_hat= GP_hat(G+G+1:G+G+G,:);

Gamma4_hat= GP_hat(G+G+G+1:G+G+G+G,:);

Pi_hat=GP_hat(G+G+G+G+1:end,:);


mYS=(mX*Pi_hat)/(Iy Gamma1_hat Gamma2_hat Gamma3_hat Gamma4_hat);


l3my0s =mYS;

l2my0s =mYS;

lmy0s=mYS;

my0s=mYS;


y01_star=zeros(T,G);

y011_star=zeros(T,G);

VI = Y   LYX*GP_hat;

for i=1:1000
```

```
        y01_star(1,:)=[l3my0s l2my0s lmy0s
        my0s ftX(1,:)]*GP_hat+VI(1,:);

        y01_star(2,:)=[l2my0s lmy0s my0s y01_star(1,:)
        ftX(2,:)]*GP_hat+VI(2,:);

        y01_star(3,:)=[lmy0s  my0s y01_star(1,:) y01_star(2,:)
         ftX(3,:)]*GP_hat+VI(3,:);

        y01_star(4,:)=[my0s   y01_star(1,:)
         y01_star(2,:) y01_star(3,:) ftX(4,:)]*GP_hat+VI(4,:);

        for t=5:T

            y01_star(t,:)=[ y01_star(t 1,:) y01_star(t 2,:)
            y01_star(t 3,:) y01_star(t 4,:) ftX(t,:)]*GP_hat+VI(t,:);

        end

        y011_star=y011_star+y01_star;

    end

        y0_star=y011_star(T,:)/1000;



        ly0_star=y011_star(T1,:)/1000;
```

```
        l2y0_star=y011_star(T2,:)/1000;

        l3y0_star=y011_star(T3,:)/1000;

y0_star=y0;

ly0_star=ly0;

l2y0_star=l2y0;

l3y0_star=l3y0;

    GP2_hat =LYX\Y2;
```

```
    V2 = Y2   LYX*GP2_hat;

    %Y2_hat = Y2   V2;

    U1 = y   [Y2 LY1X1]*b1;

%     y_hat=[Y2_hat LY1X1]*b1 U1;



    b_star = zeros(bst, size(b1,1));

    I_star = zeros(bst, 2);









    LY_star = zeros(T, g+1);

    y_star = zeros(T, 1);

    Y2_star = zeros(T, g);
```

```
for p = 1:bst

    inx = randi(T,T,1);

    V2_star = V2(inx,:);

    U1_star = U1(inx,:);

    Y2_star(1,:) =[y0_star(1, indLY) ly0_star(1, indL2Y) l2y0_star(1,
     indL3Y) l3y0_star(1, indL4Y) X(1,:)]*GP2_hat+ V2_star(1,:);

    y_star(1,:) = [Y2_star(1,:) y0_star(1, indLY) ly0_star(1, indL2Y)
     l2y0_star(1, indL3Y) l3y0_star(1, indL4Y) X1(1,:)]*b1 + U1_star(1,:);
```

```
LY_star(1, :) = [y_star(1,:) Y2_star(1,:)];



Y2_star(2,:) = [LY_star(1,[indLY exdLY]) y0_star(1, [indL2Y exdL2Y])
 ly0_star(1, [indL3Y exdL3Y]) l2y0_star(1, [indL4Y exdL4Y])
  X(2,:)]*GP2_hat + V2_star(2,:);

 y_star(2,:) = [Y2_star(2,:) LY_star(1,indLY) y0_star(1, indL2Y)
  ly0_star(1, indL3Y) l2y0_star(1, indL4Y) X1(2,:)]*b1 + U1_star(2,:);

LY_star(2, :) = [y_star(2,:) Y2_star(2,:)];



Y2_star(3,:) = [LY_star(2, [indLY exdLY]) LY_star(1, [indL2Y exdL2Y])
 y0_star(1, [indL3Y exdL3Y]) ly0_star(1, [indL4Y exdL4Y])
  X(3,:)]*GP2_hat + V2_star(3,:);

 y_star(3,:) = [Y2_star(3,:) LY_star(2, indLY) LY_star(1,indL2Y)
  y0_star(1, indL3Y) ly0_star(1, indL4Y) X1(3,:)]*b1 + U1_star(3,:);

LY_star(3, :) = [y_star(3,:) Y2_star(3,:)];



Y2_star(4,:) = [LY_star(3, [indLY exdLY]) LY_star(2, [indL2Y exdL2Y])
 LY_star(1, [indL3Y exdL3Y]) y0_star(1, [indL4Y exdL4Y]) X(4,:)]*GP2_hat
  + V2_star(4,:);

y_star(4,:) = [Y2_star(4,:) LY_star(3, indLY) LY_star(2,indL2Y)
 LY_star(1, indL3Y) y0_star(1, indL4Y) X1(4,:)]*b1 + U1_star(4,:);
```

```
LY_star(4, :) = [y_star(4,:) Y2_star(4,:)];

for t = 5:T

    Y2_star(t, :) = [LY_star(t 1 , :) LY_star(t 2 , :) LY_star(t 3 , :)
     LY_star(t 4 , :) X(t,:)]*GP2_hat + V2_star(t,:);

    y_star(t,:) = [Y2_star(t,:) LY_star(t 1 , indLY) LY_star(t 2 , indL2Y)
     LY_star(t 3 , indL3Y) LY_star(t 4 , indL4Y) X1(t,:)]*b1 +
      U1_star(t,:);

    LY_star(t, :) = [y_star(t,:) Y2_star(t,:)];

end

LYS = [y0_star(indLY); LY_star(1:T1,:)];

L2YS= [ly0_star; y0_star; LY_star(1:T2,:)];
```

```
        L3YS=[l2y0_star; ly0_star; y0_star; LY_star(1:T3,:)];

        L4YS=[l3y0_star; l2y0_star; ly0_star; y0_star;
        LY_star(1:T4,:)];

       % LYY_star = [LYS L2YS L3YS L4YS];

        LYX_star = [LYS(:,[indLY exdLY]),L2YS(:,[indL2Y exdL2Y]),
        L3YS(:,[indL3Y exdL3Y]),L4YS(:,[indL4Y exdL4Y]), X];

        LY1X1_star = [LYS(:,indLY),L2YS(:,indL2Y),
        L3YS(:,indL3Y),L4YS(:,indL4Y) X1];



        [b_star(p,:), I_star(p,:) ] = c2SLS( y_star, Y2_star,LY1X1_star,
         LYX_star);

    end

    b = 2*b1'   mean(b_star);

    I(1) = mean(I_star(1,:));

       %I(2) = mean(I_star(2,:));

end
```

```
function printResult(m, eq, T, replications, bst, const, dist, L, Bi, Ai1, Ai2,
 Ai3,Ai4, Ci, Gamma1,Gamma2,Gamma3, Gamma4, Pi, Sigma, Omega, EigVal, B, eList)

   eListLen = length(B(:,1,1));

    cptrLen = length(B(1,:,1));



    date_string = datestr(now(), 'yyyy mm dd HHMMSS');



    fname = ['M',num2str(m), '_L', num2str(L), '_N',num2str(T),
     '_B',num2str(bst), '_',num2str(dist), ' (',date_string, ')', '.txt'];

    fileID = fopen(fname,'w');

    %fileID = 1;



    fprintf(fileID, [date_string, '\r\n']);

    fprintf(fileID,'\r\n');



    fprintf(fileID, 'Elapsed time is %.4f seconds. \r\n',    toc );

    fprintf(fileID,'\r\n');
```

```
fprintf(fileID, 'Equation(%d)  N(%d)  R(%d)  Boot(%d)  ConstT(%d)  Dist(%d)
  L(%d) \r\n', eq, T, replications, bst, const, dist, L);

fprintf(fileID, '%s\r\n', '');



fprintf(fileID,'endogenous (structural) = \r\n');

ftmp = [repmat('%+2.4f   ', 1, size(Bi',2)), '\r\n'];

fprintf(fileID, ftmp, transpose(Bi'));

fprintf(fileID,'\r\n');



fprintf(fileID,'lagged endogenous|exgenous (structural) = \r\n');

ftmp = [repmat('%+2.4f   ', 1, size([Ai1; Ai2; Ai3; Ai4; Ci]',2)),
 '\r\n'];

fprintf(fileID, ftmp, transpose([Ai1;Ai2;Ai3;Ai4; Ci]'));

fprintf(fileID,'\r\n');



fprintf(fileID,'lagged endogenous|exgenous (reduced) = \r\n');
```

```
ftmp = [repmat('%+2.4f    ', 1, size([Gamma1; Gamma2; Gamma3; Gamma4;
 Pi]',2)), '\r\n'];

fprintf(fileID, ftmp, transpose([Gamma1; Gamma2; Gamma3;
 Gamma4; Pi]'));

fprintf(fileID,'\r\n');



fprintf(fileID,'Sigma = \r\n');

ftmp = [repmat('%+2.4f    ', 1, size(Sigma,2)), '\r\n'];

fprintf(fileID, ftmp, transpose(Sigma));

fprintf(fileID,'\r\n');



fprintf(fileID,'Omega = \r\n');

ftmp = [repmat('%+10.4f    ', 1, size(Omega,2)), '\r\n'];

fprintf(fileID, ftmp, transpose(Omega));

fprintf(fileID,'\r\n');



fprintf(fileID,'Eign roots = \r\n');
```

```
ftmp = [repmat('%+2.4f    ', 1, size(EigVal,2)), '\r\n'];

fprintf(fileID, ftmp, transpose(EigVal));

fprintf(fileID,'\r\n');



for p = 1:cptrLen

    fprintf(fileID,'%11s', 'Coefficient');

    fprintf(fileID,'(%1.f)', p);

    fprintf(fileID,'%10s', '');

    %fprintf(fileID,'%14s', '');

    fprintf(fileID,'%8s', 'True');

    fprintf(fileID,'%17s', 'Approx.');

    fprintf(fileID,'%20s', 'Bias');

    %fprintf(fileID,'%5s', '(R)');

    fprintf(fileID,'%19s', 'Std');

    fprintf(fileID,'%12s', 'Max');

    fprintf(fileID,'%12s', 'Min');
```

```
fprintf(fileID,'%13s', 'Median');

fprintf(fileID,'%11s', 'IQR');

fprintf(fileID,'%14s', 'Mse');

%fprintf(fileID,'%5s', '(R)');

fprintf(fileID,'%13s', 'R1');

fprintf(fileID,'%11s', 'R2');

fprintf(fileID,'%13s', 'Var');

fprintf(fileID,'\r\n');

fprintf(fileID,'=====================================================
===');

fprintf(fileID,'===================================================
=====');

fprintf(fileID,'=================================================
========\r\n');



%

meanInx = abs(B(:,p,4));

[dummy, meanInx] = sort(meanInx);
```

```
[dummy, meanInx] = sort(meanInx);



mseInx = abs(B(:,p,11));

[dummy, mseInx] = sort(mseInx);

[dummy, mseInx] = sort(mseInx);

%}



 for q = 1:eListLen

    fprintf(fileID,'%25s', getName(eList(q)));

    fprintf(fileID, '%8.4f', B(q,p,1));

    fprintf(fileID, '%12.4f',B(q,p,2));

    fprintf(fileID, ' (%+4.0f%%)',B(q,p,3));

    fprintf(fileID, '     ');

    fprintf(fileID, '(%2.0f) %6.4f', meanInx(q), B(q,p,4));

    %fprintf(fileID, '%12.4f (%2.0f)', B(q,p,4), meanInx(q));

    %fprintf(fileID, '%12.4f',B(q,p,4));
```

```
    fprintf(fileID, ' (%+4.0f%%)',B(q,p,5));



    fprintf(fileID, '%12.4f', B(q,p,6));

    fprintf(fileID, '%12.4f', B(q,p,7));

    fprintf(fileID, '%12.4f', B(q,p,8));

    fprintf(fileID, '%12.4f', B(q,p,9));

    fprintf(fileID, '%12.4f', B(q,p,10));

    %fprintf(fileID, '%12.4f', B(q,p,11));

    fprintf(fileID, '     ');

    fprintf(fileID, '(%2.0f) %6.4f', mseInx(q), B(q,p,11));

    fprintf(fileID, '%12.4f', B(q,p,12));

    fprintf(fileID, '%12.4f', B(q,p,13));

    fprintf(fileID, '%12.4f', B(q,p,14));

    fprintf(fileID, '\r\n');

end
```

```
        fprintf(fileID, '\r\n');
    end
    fclose(fileID);
end
```

```
function estName = getName(est)
    switch est
         case 10001; estName='2SLS       : ';
       case 10002; estName='C2SLS      : ';
        case 10003; estName = '2SLS_bt   : ';

        case 30011; estName='FLIML(1)      : ';
        case 30021; estName='FLIML(4)      : ';
```

```
        case 30012; estName = 'CFLIML   : ';










         otherwise;

    end

end




 function X = genX(T, k, constT, dist, beta)

     switch dist

         case 1; X = randn(T ,k);

         case 2; X =  sqrt(12)/2 + sqrt(12).*rand(T, k);

         otherwise;

     end
```

```
     for i = 2:T
         X(i,:) = beta * X(i1,:) + X(i,:);
     end

     if constT == 1
         X(:,1) = 1;
     end
 end

function [Y, LY, L2Y,  L3Y,  L4Y ] = genY(l3y0,l2y0,ly0,y0, X, rho, Gamma1,
 Gamma2, Gamma3, Gamma4,  Pi, T, G, dist)
    switch dist
        case 1; e = randn(T, G);
```

```
 case 2; e =  sqrt(12)/2 + sqrt(12).*rand(T, G);

otherwise;

end



Vb = e*rho';



Y  = zeros(T, G);






XV = X*Pi + Vb;

Y(1,:) = l3y0*Gamma4+l2y0*Gamma3+ly0*Gamma2+y0*Gamma1 + XV(1,:);

Y(2,:) = l2y0*Gamma4+ly0*Gamma3+y0*Gamma2+Y(1,:)*Gamma1 + XV(2,:);

Y(3,:) = ly0*Gamma4+y0*Gamma3+Y(1,:)*Gamma2+Y(2,:)*Gamma1 + XV(3,:);

Y(4,:) = y0*Gamma4+Y(1,:)*Gamma3+Y(2,:)*Gamma2 +
Y(3,:)*Gamma1+ XV(4,:);

for t = 5:T
```

```
        %Y(t,:) = Y(t1,:)*Gamma + X(t,:)*Pi + Vb(t,:);

        Y(t,:) = Y(t1,:)*Gamma1+Y(t2,:)*Gamma2+Y(t3,:)*Gamma3+
         Y(t4,:)*Gamma4+XV(t,:);

    end

    %LY = [y0; Y(1:T1,:)];

    LY = [y0; Y(1:T1,:)];

    L2Y= [ly0; y0; Y(1:T2,:)];

    L3Y= [l2y0; ly0;  y0; Y(1:T3,:)];

    L4Y= [l3y0; l2y0; ly0; y0; Y(1:T4,:)];


end


function [Yb, LYb, L2Yb, L3Yb, L4Yb] = genYb(l3y0,l2y0,ly0,y0, X, Gamma1,
 Gamma2,Gamma3,Gamma4,Pi, T, G)

    Yb       = zeros(T, G);

    Yb(1,:) =l3y0*Gamma4+l2y0*Gamma3+ly0*Gamma2+y0*Gamma1 + X(1,:)*Pi;

     Yb(2,:) = l2y0*Gamma4+ly0*Gamma3+y0*Gamma2+Yb(1,:)*Gamma1
```

```
     + X(2,:)*Pi;

        Yb(3,:) = ly0*Gamma4+y0*Gamma3+Yb(1,:)*Gamma2+Yb(2,:)*Gamma1
        + X(3,:)*Pi;

        Yb(4,:) = y0*Gamma4+Yb(1,:)*Gamma3+Yb(2,:)*Gamma2 +
        Yb(3,:)*Gamma1+ X(4,:)*Pi;

    for t = 5:T

        Yb(t,:) = Yb(t 1,:)*Gamma1+Yb(t 2,:)*Gamma2+Yb(t 3,:)*Gamma3+
         Yb(t 4,:)*Gamma4 + X(t,:)*Pi;

    end

    LYb = [y0; Yb(1:T1,:)];

    L2Yb= [ly0; y0; Yb(1:T2,:)];

    L3Yb= [l2y0; ly0;  y0; Yb(1:T3,:)];

    L4Yb= [l3y0; l2y0; ly0; y0; Yb(1:T4,:)];

end



function a = Dynamic_2SLS_LT_bias_approximation(l3y0, l2y0, ly0 ,y0, X, X1,
 Gamma1,Gamma2,Gamma3,Gamma4, Omega,phi, Pi, G, g, K, k,j, indY,
  indLY,indL2Y,indL3Y,indL4Y, T)
```

```
JJ =cell(T,1);

JJ{1,1}=eye(G,G);

JJ{2,1}=Gamma1;

JJ{3,1}=Gamma2+Gamma1*JJ{2,1};

JJ{4,1}=Gamma3+Gamma2*JJ{2,1}+Gamma1*JJ{3,1};

JJ{5,1}=Gamma4+Gamma3*JJ{2,1}+Gamma2*JJ{3,1}+Gamma1*JJ{4,1};

JJ{6,1}=Gamma4*JJ{2,1}+Gamma3*JJ{3,1}+Gamma2*JJ{4,1}+Gamma1*JJ{5,1};

for i=7:T

    JJ{i,1}=Gamma4*JJ{(i 4),1}+Gamma3*JJ{(i 3),1}
    +Gamma2*JJ{(i 2),1}+Gamma1*JJ{(i 1),1};

end



e=cell(4,1);

[e{1:4, 1}] = deal(zeros(G));

el=cell(4,1);

[el{1:4, 1}] = deal(zeros(G));
```

```
 D= diag(ones(1,T1) , 1);

  d_hat_t=cell(T1,1);

d_hat_r=cell(T1,1);

for t = 1:T1

  d_hat_t{t,1} = D^t;

end

for r = 1:T1

  d_hat_r{r,1} = D^r;

end



 Qw = zeros(j,j);

 for l=1:4;

     e{l,1}=eye(G);

      e=cell2mat(e);
```

```
for h=1:4

     el{h,1}=eye(G);

     el=cell2mat(el);

  for  t = l:T 1

   Qw2 =e*(JJ{t l+1,:})';

     for r=h:T 1



  Qw1 =trace((d_hat_t{t,1} )'*(d_hat_r{r,1}));



  Qw = Qw + Qw2*Omega*Qw1*(JJ{r h+1,:})*el';



     end

  end

el= mat2cell(el, [G, G, G,G]);

[el{1:4, 1}] = deal(zeros(G));

end
```

```
      e= mat2cell(e, [G, G, G,G]);

    [e{1:4, 1}] = deal(zeros(G));

end

[Yb, LYb, L2Yb, L3Yb, L4Yb] = genYb(l3y0,l2y0,ly0,y0, X, Gamma1,
 Gamma2,Gamma3,Gamma4,Pi, T, G);

 Y2b = Yb(:,indY);

    LY1b = LYb(:,indLY);

    L2Y1b = L2Yb(:,indL2Y);

    L3Y1b = L3Yb(:,indL3Y);

    L4Y1b = L4Yb(:,indL4Y);

    Rb = [Y2b LY1b L2Y1b L3Y1b L4Y1b X1];

    Zb = [LYb L2Yb L3Yb L4Yb X];
```

```
        I2s = [eye(4*G); zeros(K,4*G)];

        Qz = Zb'*Zb + [Qw zeros(G+G+G+G,K);
        zeros(K, G+G+G+G) zeros(K, K)];

        Qz = Qz\eye(size(Qz));

        Qzs = I2s'*Qz*I2s;

%         Gamma12=Gamma1(:,indY);

%         Gamma22=Gamma2(:,indY);

%         Gamma32=Gamma3(:,indY);

%         Gamma42=Gamma4(:,indY);

        Gammas2=[Gamma1(:,indY);Gamma2(:,indY);Gamma3(:,indY);Gamma4(:,indY)];

    Ed2d2 =[ Gammas2'*Qw*Gammas2 Gammas2'*Qw zeros(g,k); ...

             Qw*Gammas2 Qw zeros(4*G,k); ...

             zeros(k,g) zeros(k,j) zeros(k,k)];
```

```
Qs = Rb'*Rb + Ed2d2;

Qs =Qs\eye(size(Qs));

I4 = eye(G);

B = [I4(:,indY) zeros(G,j+k)];

psi=B'*phi;

I1 = eye(j);

A = [Gammas2 I1 zeros(G+G+G+G,k)];

I = eye(g+j+k);
```

```
T1 = zeros(g+j+k,1);

 T2 = T1; T3 = T1; T4 = T1; T5 = T1; T6 = T1; T7 = T1; T8 = T1;
 T9 = T1; T10 = T1; T11 = T1; T12 = T1; T14=T1;T13 = T1; T15 = T1;
  T16 = T1; T17 = T1; T18 = T1; T19=T1;T20=T1;

T21=T1;T22=T1;

T1 = (Rb'*Zb*Qz*Zb'*Rb*Qs + (trace(Zb*Qz*Zb'*Rb*Qs*Rb')*I))*psi;



T5 = ( (trace(Qw*I2s'*Qz*Zb'*Rb*Qs*A')*I) +
Rb'*Zb*Qz*I2s*Qw*A*Qs + A'*Qw*I2s'*Qz*Zb'*Rb*Qs )*psi;



T8 = (A'*Qw*Qzs*Qw*A*Qs + (trace(Qw*Qzs*Qw*A*Qs*A')*I))*psi;



T14=(trace(Qw*Qzs)*I)*psi;




el=cell(4,1);

er=cell(4,1);
```

```
ek=cell(4,1);

[el{1:4, 1}] = deal(zeros(G));

[er{1:4, 1}] = deal(zeros(G));



[ek{1:4, 1}] = deal(zeros(G));




AQs = A*Qs;

AQsA_ = AQs*A';

AQsB_ = AQs*B';

AQsRb_ = AQs*Rb';

RbQs_A_ =AQsRb_';

BQs = B*Qs;

BQsA_ = BQs*A';

BQsRb_ = BQs*Rb';

RbQs = Rb*Qs;
```

```
RbQsB_ = RbQs*B';

RbQsRb_ = RbQs*Rb';

ZbQz = Zb*Qz;

ZbQzI2s = ZbQz*I2s;

I2s_QzZb_ = ZbQzI2s';

ZbQzZb_ = ZbQz*Zb';

B_Omega = B'*Omega;





for ll=1:4

     el{ll,1}=eye(G);

     el=cell2mat(el);

     for t = ll:T 1
```

```
T2a = Rb'*d_hat_t{t,1}*RbQs;

T2b = trace(Rb'*d_hat_t{t,1}*RbQs)*I;



T2 = T2 + (T2a + T2b)*A'*el*(JJ{t ll+1,:})'* phi;

T3c = Qzs*el*(JJ{t ll+1,:})';

T4b = ZbQzI2s*el*(JJ{t ll+1,:})'*Omega;

T6b = Omega*JJ{t ll+1,:}*el'*AQsA_;

T6c = A'*el*(JJ{t ll+1,:})'*Omega;

T7a = T6c;

T9b = RbQs*B'*Omega*JJ{t ll+1,:}*el'*Qzs;

T10a =  A'*el*(JJ{t ll+1,:})';

T11a = T10a;

T12a = Qzs*el*(JJ{t ll+1,:})';

T12c = AQsA_*el*(JJ{t ll+1,:})';

T12f = ZbQzI2s*el*(JJ{t ll+1,:})';

T20a = B_Omega*JJ{t ll+1,:}*el'*AQsRb_;
```

```
    T20c = B_Omega*JJ{t ll+1,:}*el'*I2s_QzZb_;



    T19a =T12f;

    T17a = trace(Omega*JJ{t ll+1,:}*el'*AQsB_);

 for rr=1:4



        er{rr,1}=eye(G);

         er=cell2mat(er);

   for r = rr:T 1








        T3a  = Rb'*(d_hat_t{t,1})*(d_hat_r{r,1})'*RbQs;

        T3b  = trace((d_hat_t{t,1})*(d_hat_r{r,1})'*RbQsRb_)*I;



        T4a  = (d_hat_t{t,1})*(d_hat_r{r,1})';
```

```
T4c = AQsRb_;

T21a = (d_hat_t{t,1})'*(d_hat_r{r,1});



T6a = trace(ZbQzZb_*(d_hat_t{t,1})*(d_hat_r{r,1})')*I;

T7b = I2s_QzZb_*(d_hat_t{t,1})*(d_hat_r{r,1})'*RbQs*psi;



T9a = Rb'*(d_hat_t{t,1})*(d_hat_r{r,1});

T10b = (d_hat_r{r,1})'*(d_hat_t{t,1})'
+(d_hat_r{r,1})'*(d_hat_r{r,1});

T10c = (d_hat_t{t,1})'*(d_hat_r{r,1})*ZbQzI2s;

T11b = (d_hat_t{t,1})'*ZbQzZb_*(d_hat_r{r,1})';



T12b = (d_hat_t{t,1})*RbQsRb_*(d_hat_r{r,1});

T12d = ZbQzZb_*(d_hat_t{t,1})*(d_hat_r{r,1});

T12e = AQsRb_*(d_hat_t{t,1})'*(d_hat_r{r,1})';



T18a = trace((d_hat_t{t,1})*(d_hat_r{r,1})');
```

```
T19b =AQsRb_*T4a;

T20b = ((d_hat_t{t,1})*(d_hat_r{r,1})+
(d_hat_t{t,1})'*(d_hat_r{r,1}))*ZbQzI2s;

T20d = ((d_hat_t{t,1})'*(d_hat_r{r,1}))*RbQs_A_;

T16a =  Rb'*(d_hat_t{t,1})*(d_hat_r{r,1})';



T3  = T3  + (T3a + T3b)*(trace(Omega*JJ{r rr+1,:}*er'*T3c)*I)*psi;



T4  = T4  + (trace(T4a*T4b*JJ{r rr+1,:}*er'*T4c)*I)*psi;



T6  = T6  + T6a * ((trace(T6b*er*(JJ{r rr+1,:})')*I) +
 T6c*JJ{r rr+1,:}*er'*AQs)*psi;



T7  = T7  + (T7a*JJ{r rr+1,:}*er'*T7b);

T9  = T9  + ( T9a*T9b*er*(JJ{r rr+1,:})'*phi);

%T9  = T9  +
 ((D^t)*(D^r)*Rb*Qs*B'*Omega*(Gamma^(t 1))*
 Qzs*(Gamma^(r 1))'*vphi);
```

```
T10 =T10+ ( T10a* (Omega*BQsRb_*(T10b)*ZbQzI2s*er
*(JJ{r rr+1,:})'
+ ...
    (trace(T10c*er*(JJ{r rr+1,:})'*Omega*BQsRb_)
    *eye(G)))*phi);

T11 = T11 + ( T11a*
 (trace(Omega*BQsA_*er*(JJ{r rr+1,:})')*
 trace(T11b)*eye(G)) * phi);

T12 = T12 + ( B'*Omega*JJ{r rr+1,:}*er'*
(T12a*(trace(T12b)*eye(G))
 + ...

T12c*(trace(T12d)*eye(G)) + T12e*T12f)*phi);



T16 =T16+T16a*T4b*JJ{r rr+1,:}*er'*AQs*psi;

T17 =T17+T9a*(T17a*eye(T))*ZbQzI2s*er*(JJ{r rr+1,:})'*phi;

T18 =T18+B'*trace(T19a*Omega*JJ{r rr+1,:}*er'*AQsRb_)*
T18a*eye(G)*phi;



T19 = T19+B'*trace(T19a*Omega*JJ{r rr+1,:}*er'*T19b)*phi;
```

```
        T20 = T20+(T20a*T20b+T20c*T20d)*er*(JJ{r rr+1,:})'*phi;

        T21b = AQsA_*er*(JJ{r rr+1,:})';



for k=1:4

        ek{k,1}=eye(G);

        ek=cell2mat(ek);

    for s=k:T

           if t == r+s

               T21 = T21 + ( T6c*JJ{t r k+1,:}*ek'*T21b
                *trace(T21a*D^(t r)))*eye(G)*phi;

           elseif r == t+s

               T22 = T22 + (T10a* trace(JJ{r t k+1,:}*ek'*T21b *Omega)
                *trace(T21a*(D^(r t))')*eye(G))*phi;

           else

           end

    end
```

```
 ek= mat2cell(ek, [G G G G]);

[ek{1:4, 1}] = deal(zeros(G));

end

  end

  er= mat2cell(er, [G G G G]);

[er{1:4, 1}] = deal(zeros(G));

end

   end

 el= mat2cell(el, [G G G G]);
```

```
 [el{1:4, 1}] = deal(zeros(G));

end

 T1 =    Qs*T1;

 T2 =  Qs*(trace(Zb*Qz*Zb')*I)*psi Qs*T2 ;

 T3 =    Qs*T3;

 T4  =   Qs*T4;

 T5  =   Qs*T5;

 T6  =   Qs*T6;

 T7  =   Qs*T7;

 T8  =   Qs*T8;

 T9  =   Qs*T9;

 T10 =   Qs*T10;

 T11 =   Qs*T11;

 T12 =   Qs*T12;
```

```
    T14= Qs*T14;

    T16 =   Qs*T16;

    T17 =   Qs*T17;

    T18 =   Qs*T18;

    T19 =   Qs*T19;

    T20 =   Qs*T20;

    T21 =   Qs*T21;

    T22 =   Qs*T22;



   a = T1 + T2 + T3 + T4 + T5 + T6 + T7 + T8 + T9 + T10 + T11 +
    T12+T21+T22+T14+T16+T17+T18+T19+T20

end



function af = Dynamic_Full_LT_bias_approximation(l3y0, l2y0, ly0 ,y0, X, X1,
 Gamma1,Gamma2,Gamma3,Gamma4, Omega,phi, Pi, G, g, K, k,j, indY,
  indLY,indL2Y,indL3Y,indL4Y, T)



   JJ =cell(T,1);
```

```
JJ{1,1}=eye(G,G);

JJ{2,1}=Gamma1;

JJ{3,1}=Gamma2+Gamma1*JJ{2,1};

JJ{4,1}=Gamma3+Gamma2*JJ{2,1}+Gamma1*JJ{3,1};

JJ{5,1}=Gamma4+Gamma3*JJ{2,1}+Gamma2*JJ{3,1}+Gamma1*JJ{4,1};

JJ{6,1}=Gamma4*JJ{2,1}+Gamma3*JJ{3,1}+Gamma2*JJ{4,1}+Gamma1*JJ{5,1};

for i=7:T

    JJ{i,1}=Gamma4*JJ{(i 4),1}+Gamma3*JJ{(i 3),1}+Gamma2*JJ{(i 2),1}
    +Gamma1*JJ{(i 1),1};

end



e=cell(4,1);

[e{1:4, 1}] = deal(zeros(G));

el=cell(4,1);

[el{1:4, 1}] = deal(zeros(G));
```

```
 D= diag(ones(1,T 1) , 1);

  d_hat_t=cell(T 1,1);

d_hat_r=cell(T 1,1);

for t = 1:T 1

  d_hat_t{t,1} = D^t;

end

for r = 1:T 1

  d_hat_r{r,1} = D^r;

end



 Qw = zeros(j,j);

 for l=1:4;

     e{l,1}=eye(G);

      e=cell2mat(e);

     for h=1:4
```

```
      el{h,1}=eye(G);

      el=cell2mat(el);

   for  t = 1:T 1

    Qw2 =e*(JJ{t l+1,:})';

      for r=h:T 1



   Qw1 =trace((d_hat_t{t,1} )'*(d_hat_r{r,1}));



   Qw = Qw + Qw2*Omega*Qw1*(JJ{r h+1,:})*el';



      end

   end

 el= mat2cell(el, [G, G, G,G]);

 [el{1:4, 1}] = deal(zeros(G));

 end

   e= mat2cell(e, [G, G, G,G]);
```

```
    [e{1:4, 1}] = deal(zeros(G));

end

[Yb, LYb, L2Yb, L3Yb, L4Yb] = genYb(l3y0,l2y0,ly0,y0, X, Gamma1,
 Gamma2,Gamma3,Gamma4,Pi, T, G);

 Y2b = Yb(:,indY);

    LY1b = LYb(:,indLY);

    L2Y1b = L2Yb(:,indL2Y);

    L3Y1b = L3Yb(:,indL3Y);

    L4Y1b = L4Yb(:,indL4Y);

    Rb = [Y2b LY1b L2Y1b L3Y1b L4Y1b X1];

    Zb = [LYb L2Yb L3Yb L4Yb X];

    I2s = [eye(4*G); zeros(K,4*G)];
```

```
    Qz = Zb'*Zb + [Qw zeros(G+G+G+G,K); zeros(K, G+G+G+G) zeros(K, K)];

    Qz = Qz\eye(size(Qz));

    Qzs = I2s'*Qz*I2s;



    Gammas2=[Gamma1(:,indY);Gamma2(:,indY);Gamma3(:,indY);Gamma4(:,indY)];



Ed2d2 =[ Gammas2'*Qw*Gammas2 Gammas2'*Qw zeros(g,k); ...

         Qw*Gammas2 Qw zeros(4*G,k); ...

         zeros(k,g) zeros(k,j) zeros(k,k)];



Qs = Rb'*Rb + Ed2d2;

Qs =Qs\eye(size(Qs));
```

```
I4 = eye(G);

B = [I4(:,indY) zeros(G,j+k)];



psi=B'*phi;



I1 = eye(j);

A = [Gammas2 I1 zeros(G+G+G+G,k)];



I = eye(g+j+k);



T1 = zeros(g+j+k,1);

 T2 = T1;  T5 = T1;

T1 = (Rb'*Zb*Qz*Zb'*Rb*Qs + (trace(Zb*Qz*Zb'*Rb*Qs*Rb')*I))*psi;



T5 = ( (trace(Qw*I2s'*Qz*Zb'*Rb*Qs*A')*I) + Rb'*Zb*Qz*I2s*Qw*A*Qs
```

```
 + A'*Qw*I2s'*Qz*Zb'*Rb*Qs )*psi;

 el=cell(4,1);

 [el{1:4, 1}] = deal(zeros(G));

RbQs = Rb*Qs;

for ll=1:4
     el{ll,1}=eye(G);
     el=cell2mat(el);
     for t = ll:T 1

     T2a = Rb'*d_hat_t{t,1}*RbQs;
     T2b = trace(Rb'*d_hat_t{t,1}*RbQs)*I;
```

```
        T2 = T2 + (T2a + T2b)*A'*el*(JJ{t ll+1,:})'* phi;

    end

        end

    T1 =    Qs*T1;

    T2 =  Qs*(trace(Zb*Qz*Zb')*I)*psi Qs*T2 ;

    T5  =   Qs*T5;

    af = a  (T1+T2 T5 )

end

function b = iqr(X)

    XS  = sort(X);
```

```
    N    = length(X);



    q1       = (N+1) / 4;

    q1L      = floor(q1);

    q1R      = q1L + 1;

    qDiff    = q1   q1L;

    Q1       = XS(q1L) + qDiff * ( XS(q1R)   XS(q1L) );



    q3       = 3*(N+1) / 4;

    q3L      = floor(q3);

    q3R      = q3L + 1;

    qDiff    = q3   q3L;

    Q3       = XS(q3L) + qDiff * ( XS(q3R)   XS(q3L) );



    b = Q3   Q1;

end
```

```
function b = mse(X, trueValue)

    tmp = squeeze(X);

    n = length(tmp);

    b = ((tmp   trueValue)' * (tmp   trueValue))/n;

end
```