# Shrinkage Estimation and Forecasting in Dynamic Regression Models under Structural Instability

Ali Mehrabani Department of Economics University of Kansas Shahnaz Parsaeian\* Department of Economics University of Kansas Aman Ullah Department of Economics University of California, Riverside

June 10, 2024

#### Abstract

This paper introduces a Stein-like shrinkage method for estimating slope coefficients and forecasting in first order dynamic regression models under structural breaks. The model allows for unit root and non-stationary regressors. The proposed shrinkage estimator is a weighted average of a restricted estimator that ignores the break in the slope coefficients, and an unrestricted estimator that uses the observations within each regime. The restricted estimator is the most efficient estimator but inconsistent when there is a break. However, the unrestricted estimator is consistent but not efficient. Therefore, the proposed shrinkage estimator balances the trade-off between the bias and variance efficiency of the restricted estimator. The averaging weight is proportional to the weighted distance of the restricted estimator, and the unrestricted estimator. We derive the analytical large-sample approximation of the bias, mean squared error, and risk for the shrinkage estimator, the unrestricted estimator, and the restricted estimator. We show that the risk of the shrinkage estimator is lower than the risk of the unrestricted estimator under any break size and break points. Moreover, we extend the results for the model with a unit root and non-stationary regressors. We evaluate the finite sample performance of our proposed method via extensive simulation study, and empirically in forecasting output growth.

*Keywords*: ARX-model; Asymptotic approximation; Dynamic regressions; Forecasting; Moment approximation; Non-stationary regressors; Structural breaks; Unit root.

JEL Classification: C13, C22, C53

<sup>\*</sup>Address correspondence to: Shahnaz Parsaeian, sh.parsaeian@ku.edu.

## 1 Introduction

A sizeable strand of the literature in economics and finance focuses on autoregressive (AR) models. These models are heavily used in forecasting economic and financial variables and are frequently considered as benchmarks in forecast competitions, as they are difficult to beat. Nevertheless, since many time series data in economics and finance are characterized by parameter instability, which is now widely recognized as an important source of forecast failure as recorded by Stock and Watson (1996), Hansen (2001), Giacomini and Rossi (2009), Rossi and Sekhposyan (2010), Inoue and Rossi (2011), Clements and Hendry (2006, 2011), and Rossi (2013) inter alia, there is an increasing evidence that parameters of AR models in many economic and financial time series are unstable and subject to structural breaks. For example, Mankiw and Miron (1986) and Mankiw et al. (1987) considered AR(1) models and find parameter instability in the short-term interest rate. Phillips et al. (2011) find parameter instability for 1990's NASDAQ stock prices. See also Garcia and Perron (1996), and Stock and Watson (1996) who document instability related to the autoregressive terms in a wide variety of economic time series. This suggests a need to study the forecasting performance of AR models when they are subject to structural breaks.

This paper considers an ARX model that contains structural breaks, and proposes a Stein-like shrinkage estimator that exploits observations in neighboring regimes.<sup>1</sup> Our proposed estimator is a weighted average of a restricted estimator (which uses full-sample of observations under the restriction of no breaks in the parameters) with an unrestricted estimator (which uses observations within each regime). The restricted estimator is inconsistent when there is a break while it is the most efficient. On the other hand, the unrestricted estimator is consistent but the consistency comes at the cost of losing efficiency. Hence, the proposed shrinkage estimator trades off an increased bias introduced by the restricted estimator against a reduction in error variance resulting from using a full-sample of observations. The averaging weight depends on the weighted distance of the restricted and unrestricted estimators, which is similar to the James-Stein weight, cf. Stein (1956) and James and Stein (1961).<sup>2</sup> Therefore, it assigns appropriate weights to each of the

<sup>&</sup>lt;sup>1</sup>Over the past decades, a considerable literature on the merits of forecasts averaging obtained from different models has been developed, reviewed by Clemen (1989), Newbold and Harvey (2002), Stock and Watson (2004), and Timmerman (2006). Specifically under structural break models, forecast averaging procedures have been shown to improve forecast performance in the sense of mean squared forecast errors (MSFE), see Pesaran and Timmermann (2002, 2007), Hansen (2009), Clark and McCracken (2010), Pesaran and Pick (2011), Pesaran et al. (2013), Inoue et al. (2017), and Lee et al. (2022a, 2022b, 2022c), among others.

 $<sup>^{2}</sup>$ See also Massoumi (1978), Hansen (2016, 2017), and Mehrabani and Ullah (2020) for similar use of the James-Stein weight in different contexts.

two estimators by measuring the magnitude of the structural break. We derive the analytical large-sample approximation of the bias, mean squared error (MSE) and risk for our proposed shrinkage estimator, the unrestricted estimator, and the restricted estimator.<sup>3</sup> We derive the condition under which the risk of the shrinkage estimator is lower than the risk of the unrestricted estimator under any break size and break points. We also show how the method can be used in out-of-sample forecasting. Furthermore, we extend the results to the model with unit root and non-stationary regressors.

To the best of our knowledge, this is the first paper that analytically derives the asymptotic approximation of the bias, and risk up to order  $T^{-1}$ , and MSE up to order  $T^{-2}$  in ARX models under structural breaks, where T is the total number of observations. Furthermore, we provide new results to the model with unit root and non-stationary regressors under structural breaks, which have not been considered before in the literature. Because of these considerations, we would like to point out that our results differ from the recent work by Lee et al. (2022a, 2022b, 2022c) who consider different types of combined estimators to improve forecasts under structural breaks in a static stationary time series model without lagged dependent variable and using different weights. For example, in Lee et al. (2022a) the combination weight is different and set to be a constant between zero and one, Lee et al. (2022b) consider Stein-like combined estimator, however, the optimal weights in this paper are different due to the presence of lagged dependent variable and theoretical frameworks. Lee et al. (2022c) consider another type of combined estimators which assigns a full weight of one to the post-break sample observations and a weight between zero and one to the pre-break sample observations. In addition, there are three additional main differences that have not been discussed in the previous works by Lee et al. (2022a, 2022b, 2022c). First, this paper considers a dynamic ARX model and allows for unit roots (or integrated of higher order) and non-stationary regressors. Second, in this paper the dominance property of the proposed Stein-like shrinkage estimator holds for any fixed deviations from the restrictions. This complements the "local asymptotic" argument discussed in Lee et al. (2022b). Third, we derive the analytical large sample approximation, using Nagar (1959) method, of the bias and risk up to order  $T^{-1}$ , and MSE up to order  $T^{-2}$  for the proposed shrinkage estimator, the unrestricted estimator, and the restricted estimator under structural breaks. This allows us to theoretically and numerically compare the performance of these estimators. Because of the presence of the lagged dependent variable in the

<sup>&</sup>lt;sup>3</sup>The bias and MSE of the unrestricted estimator are provided in the Supplementary Online Appendix B. The detailed derivations of bias and MSE of the restricted estimator are available upon request.

model considered in this paper, the theoretical results presented here are noticeably different than those in Lee et al. (2022a, 2022b, 2022c). Even in the special case of no lagged variable, our results are different with them because the combined estimator in Lee et al. (2022a) has different weights, and in Lee et al. (2022b, 2022c) they either derive only local asymptotic theory results instead of large sample approximation results and/or their weights in combined estimator are different.

We present an extensive Monte Carlo simulation study to evaluate the finite sample forecasting performance of the proposed shrinkage method. The results support our theoretical findings, and show the outperformance of our shrinkage estimator relative to the unrestricted estimator in finite sample. In particular, our numerical results show the benefits of exploiting the neighboring observations in estimation and forecasting relative to using only the observations within each regime. Furthermore, we undertake an empirical analysis for forecasting the output growth using 131 macroeconomic and financial time series to compare the forecasting performance of the shrinkage estimator with the unrestricted estimator, the restricted estimator, and a range of alternative methods existing in the literature. The empirical results suggest the out-performance of our method relative to the alternative methods, in the sense of mean squared forecast errors (MSFE).

Analysis of AR models subject to structural breaks have been also considered by Clements and Hendry (1998, 1999) and Pesaran and Timmermann (2005), but they focus on the analysis of forecast errors decomposition, while the focus of this paper is developing an estimator to deal with the parameter instability. Specifically, Clements and Hendry (1998, 1999) analyze the forecast errors from AR models subject to structural change by assuming that the parameters of the AR models remain unchanged during the estimation period. Pesaran and Timmermann (2005) consider the small sample properties of forecasts from AR models estimated from windows of different sizes. See also Banerjee and Urga (2006) for a comprehensive review of developments in the fields of modelling structural breaks.

The analysis of approximating the moments of estimators in autoregressive models dates back to Bartlett (1946), who finds a first-order variance approximation in an autoregressive gaussian process (see also Hurwicz (1950)). White (1961), Shenton and Johnson (1965) give approximations of the first two moments in the AR(1) model. Kendall (1954) and Marriott and Pope (1954) consider AR(1) models with intercepts, and find the approximate bias of the least-squares estimator of the lagged dependent variable coefficient. Recently, a number of papers have studied the small sample bias of the ordinary least-squares estimator in single dynamic regression models, see for example Grubb and Symons (1987), and Kiviet and Phillips (1993). Kiviet and Phillips (1994) extend the analysis of Kiviet and Phillips (1993) to the higher-order dynamic regression models. More recently, Kiviet and Phillips (2012) found the higher-order approximate bias of the least-squares estimator of the slope coefficients in a stable ARX(1) model. Kiviet and Phillips (2005) extend these results to non-stable ARX(1) models, and examine the moments of the least-squares estimator in the single normal ARX(1) model with an arbitrary number of exogenous regressors when the true coefficient of the lagged-dependent variable is unity.

The paper is organized as follows. Section 2 describes the model. For simplicity, we discuss the problem under a single break, which simplifies the essential idea without complicating notation. However, the generalization of the method for the multiple breaks is straightforward. In Section 3, we introduce the estimators. We give the bias, MSE, and the risk of the Stein-like shrinkage estimator using large-sample approximations in Section 4, while those of the unrestricted estimator are provided in the Supplementary Online Appendix B. We extend the results of Section 4 to a model where the slope coefficient of the lagged dependent variable is unity in Section 5. Monte Carlo results are given in Section 6. Results from our empirical example are given in Section 7. Conclusions are given is section 8. Proofs and detailed calculations are provided in Appendix A.

Notation: Throughout the paper we adopt the following notation.  $I_p$  and  $0_{p \times q}$  denote the  $p \times p$ identity matrix and  $p \times q$  matrix of zeros, respectively.  $tr(\cdot)$  denotes trace,  $\otimes$  denotes Kronecker products, and  $1(\cdot)$  denotes the indicator function. For an  $m \times n$  real matrix  $A = (a_{ij})$ , we write the transpose A'. We write A = O(b) when the non-zero elements of A are of order b, i.e.,  $a_{ij} = O(b)$ . For a stochastic matrix A, we write  $A = O_p(b)$  when the non-zero elements of A are of order b, i.e.,  $a_{ij} = O_p(b)$ .

# 2 The Model

Consider the following first-order linear dynamic regression model  $(ARX(1) \text{ model})^4$  defined over the period t = 1, 2, ..., T, which is subject to a single structural break at time  $T_1$ 

$$y_t = \begin{cases} \lambda_1 y_{t-1} + x'_t \beta_1 + u_t, & \text{for } t \le T_1 \\ \lambda_2 y_{t-1} + x'_t \beta_2 + u_t, & \text{for } t > T_1, \end{cases}$$
(2.1)

<sup>&</sup>lt;sup>4</sup>Our findings can be generalized—though will not yield qualitatively different results as it seems—for models with higher order lags. For ease of exposition we thus restrict ourselves to the relatively simple ARX(1) model.

where  $y_t$  is the dependent variable,  $x_t = (x_{t,1}, \ldots, x_{t,k})'$  is a  $k \times 1$  vector of exogenous regressors, and  $u_t$  is the unobserved error term.<sup>5</sup> The  $(k+1) \times 1$  vectors of the pre-break and post-break slope coefficients are denoted by  $\alpha_1 = (\lambda_1, \beta'_1)'$ , and  $\alpha_2 = (\lambda_2, \beta'_2)'$ , respectively, which are the parameters of interest. In a matrix form, the model can be expressed as

$$y = Z\alpha + u, \tag{2.2}$$

where  $y = (y^{(1)'}, y^{(2)'})'$  is a  $T \times 1$  vector of the dependent variables,  $y^{(1)} = (y_1, \ldots, y_{T_1})'$ ,  $y^{(2)} = (y_{T_1+1}, \ldots, y_T)'$ . The  $T \times 2(k+1)$  matrix of observations on the regressors is denoted by  $Z = \text{diag}(Z_1, Z_2)$ , where  $Z_i = (y_{-1}^{(i)}, X_i)$  for  $i = 1, 2, y_{-1}^{(1)} = (y_0, \ldots, y_{T_1-1})'$ ,  $y_{-1}^{(2)} = (y_{T_1}, \ldots, y_{T-1})'$ ,  $X_1 = (x_1, \ldots, x_{T_1})'$ , and  $X_2 = (x_{T_1+1}, \ldots, x_T)'$ .  $u = (u_1, \ldots, u_T)'$  is a  $T \times 1$  vector of disturbances, and  $\alpha = (\alpha'_1, \alpha'_2)'$  is a  $2(k+1) \times 1$  vector of the unknown slope coefficients.

Assumption 1 (i)  $|\lambda_i| < 1$ , for i = 1, 2; (ii) the matrix Z is such that  $Z'Z = O_p(T)$ ; (iii) the  $T \times 2(k+1)$  matrix Z has rank 2(k+1) with probability one; (iv) the regressors in X are strongly exogenous; (v) the disturbances follow  $u \sim N(0, \Omega_u)$ , where  $\Omega_u = diag(\sigma_1^2 I_{T_1}, \sigma_2^2 I_{T-T_1})$  with  $0 < \sigma_i^2 < \infty$ ; (vi) the start-up value follows  $y_0 = \bar{y}_0 + du_0 \sim N(\bar{y}_0, d^2 \sigma_1^2)$ , with  $0 \le d < \infty$ ; (vii)  $y_0$  and u are mutually independent.

We note that  $u_0$  is the start-up error term, i.e., the error term at time 0. d = 0 represents the fixed start-up, and if  $d \neq 0$  the start-up is random. Assumption 1(ii) excludes non-stationary regressors including deterministic or stochastic trends, but the presence of such variables will not change the approximation bias, MSE, MSFE, and risk formulas in Section 4. Their inclusion only reduces the order of magnitude of the moments and the order of their remainder terms (for a similar discussion see Kiviet and Phillips (2012)). We demonstrate in Section 5 that our results are still applicable under a unit root (relaxing Assumption 1(i)) and non-stationary regressors (relaxing Assumption 1(ii)). Assumption 1(v) assumes that the error terms are normally distributed and excludes serial correlation in the errors. This assumption can be relaxed at the expense of extra terms to be added in the bias and MSE of the estimators. However, since the model in (2.1) has a dynamic structure, it captures much of the serial correlation.

Remark 1 A main difference between the model considered in (2.1) and those considered in Lee

 $<sup>\</sup>overline{}^{5}$  For simplicity, we consider a single break in (2.1), however, the extension to multiple breaks is straightforward.

et al. (2022a, 2022b, 2022c) is that the model considered here is a dynamic ARX, and allows for a unit root (or higher order integration), and non-stationary regressors.

We analyze the model conditional on the observed matrix X in Section 4. Therefore, in order to distinguish the fixed and zero mean stochastic elements of the regressor matrix  $Z_i$ , i = 1, 2, we decompose  $Z_i = \bar{Z}_i + \tilde{Z}_i$ , where  $\bar{Z}_i$  is defined as the expectation of  $Z_i$  conditional on X and  $\bar{y}_0$ . For notational simplicity, we denote  $\mathbb{E}(\cdot) \equiv \mathbb{E}(\cdot|X, \bar{y}_0)$ . Hence,

$$\bar{Z}_i = \mathbb{E}(Z_i) = \left(\mathbb{E}(y_{-1}^{(i)}), X_i\right) = \left(\bar{y}_{-1}^{(i)}, X_i\right), \quad \text{for } i = 1, 2,$$
(2.3)

and

$$\tilde{Z}_{i} = Z_{i} - \bar{Z}_{i} = \left(\tilde{y}_{-1}{}^{(i)}, 0_{T_{i} \times k}\right) = \tilde{y}_{-1}{}^{(i)}\varsigma_{1}', \quad \text{for } i = 1, 2,$$
(2.4)

where  $\varsigma_1 = (1, 0, \dots, 0)'$  is a vector of  $(k+1) \times 1$ , and  $T_2 = T - T_1$  is the number of post-break sample observations. Define the  $T_i \times T$  matrix  $\Lambda_i = (F_i, \Delta_i)$  for i = 1, 2. Then, we define  $\bar{y}_{-1}^{(i)} = \Lambda_i \bar{y}_{-1}^*$ , where  $\bar{y}_{-1}^* = (\bar{y}_0, x'_1 \beta_1, \dots, x'_{T_1} \beta_1, x'_{T_1+1} \beta_2, \dots, x'_{T_{-1}} \beta_2)'$ ,  $F_1 = (1, \lambda_1, \dots, \lambda_1^{T_1-1})'$  is  $T_1 \times 1$ ,  $F_2 = (\lambda_1^{T_1}, \lambda_1^{T_1} \lambda_2, \dots, \lambda_1^{T_1} \lambda_2^{T_2-1})'$  is  $T_2 \times 1$ , and

are  $T_1 \times (T-1)$  and  $T_2 \times (T-1)$ , respectively. Moreover, define the  $(T+1) \times 1$  random vector  $\nu = (u_0, u')' = (u_0, u_1, \dots, u_T)' \sim N(0, \Omega_{\nu})$ , where  $\Omega_{\nu} = diag(\sigma_1^2 I_{T_1+1}, \sigma_2^2 I_{T_2})$ , and the  $T \times (T+1)$  matrix  $G = (G'_1, G'_2)'$ , with the  $T_i \times (T+1)$  matrix  $G_i = (dF_i, C_i)$ , and  $C_i = (\Delta_i, 0_{T_i \times 1})$ , for i = 1, 2. Then, it can be easily verified that  $\tilde{Z}_i = G_i \nu \varsigma_1'$  for i = 1, 2. Therefore, we have  $Z = \bar{Z} + \tilde{Z}$ , where

$$\begin{split} \bar{Z} &= \operatorname{diag}(\bar{Z}_1, \bar{Z}_2), \text{ and } \tilde{Z} &= \operatorname{diag}(\tilde{Z}_1, \tilde{Z}_2) = \sum_{i=1}^2 L_i G \nu e'_{1,i}, \text{ where the } 2(k+1) \times 1 \text{ vectors } e_{1,1} = \\ (\varsigma'_1, 0_{1 \times (k+1)})' \text{ and } e_{1,2} &= (0_{1 \times (k+1)}, \varsigma'_1)', \text{ and } L_1 = \begin{pmatrix} I_{T_1} & 0_{T_1 \times T_2} \\ 0_{T_2 \times T_1} & 0_{T_2 \times T_2} \end{pmatrix}, L_2 = \begin{pmatrix} 0_{T_1 \times T_1} & 0_{T_1 \times T_2} \\ 0_{T_2 \times T_1} & I_{T_2} \end{pmatrix}, \\ are T \times T \text{ selection matrices.} \end{split}$$

**Remark 2** In our analysis below, we assume that the break point is known. In practice, one has to estimate the break point from data. There are several methods in the literature that can be used to estimate the break point, e.g., Bai and Perron (1998, 2003), who also show that the estimated break fraction converges to its true value at rate T or lower, which is sufficient to establish the consistency of the estimators. In the simulation study and the empirical example in Sections 6 and 7, initially we have estimated the break point, and then applied our method.<sup>6</sup>

### 3 Estimation

Our goal is to estimate the vector of slope parameters,  $\alpha$ , in equation (2.2). We consider three estimators of the slope parameters: (i) an unrestricted least squares estimator that estimates the slope parameters of each regime only using the observations within the same regime, (ii) a restricted estimator that shrinks the unrestricted estimator towards a restricted parameters space which ignores the break in the slope coefficients, and (iii) a Stein-like shrinkage estimator which is a weighted average of the restricted estimator and the unrestricted estimator, and the averaging weight is proportional to a weighted quadratic loss function.

#### 3.1 Unrestricted Estimator

If we allow for the structural break in the slope coefficients, the standard estimator of  $\alpha$  is an unrestricted least-squares estimator. The unrestricted estimator, denoted by  $\hat{\alpha}$ , is defined as

$$\hat{\alpha} = (Z'Z)^{-1}Z'y = \alpha + (Z'Z)^{-1}Z'u, \tag{3.1}$$

<sup>&</sup>lt;sup>6</sup>We note that the literature under structural breaks divides into two strands. The first strand of the literature focuses on break points estimation. There are extensive contributions in this area. The second strand of the literature focuses on forecast improvement under given structural breaks. This paper falls into the second strand of the literature. It would be interesting to include the estimation uncertainty of the estimated break point into the theoretical derivations. This is beyond the scope of this paper due to additional technical challenges that will be introduced to the model, and we defer such analysis to future research.

or equivalently,

$$\hat{\alpha} = \begin{pmatrix} \hat{\alpha}_1 \\ \hat{\alpha}_2 \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} + \begin{pmatrix} (Z_1'Z_1)^{-1}Z_1'u_1 \\ (Z_2'Z_2)^{-1}Z_2'u_2 \end{pmatrix}.$$
(3.2)

#### 3.2 Restricted Estimator

Because of a belief that the break in slope parameters may be small, it can be assumed that the parameter values may be close to a restricted parameter space  $\Xi_0 = \{\alpha \in \mathbb{R}^{2(k+1)} : r(\alpha) = 0\}$ , where  $r(\alpha) = R\alpha : \mathbb{R}^{2(k+1)} \to \mathbb{R}^{k+1}$ . Thus, we shrink  $\hat{\alpha}$  towards the restriction space  $\Xi_0$ . For example, a restriction matrix R that considers no break in all slope parameters is  $R = (I_{(k+1)}, -I_{(k+1)})$ .<sup>7</sup> Thus, the restricted least-squares estimator is obtained as the solution to the following minimization

$$\underset{\text{s.t. }\alpha}{\text{Minimize}} (y - Z\alpha)'(y - Z\alpha) \qquad \text{subject to } r(\alpha) = 0.$$
(3.3)

Therefore, the restricted least-squares estimator, denoted by  $\tilde{\alpha}$ , can be formulated as

$$\tilde{\alpha} = \hat{\alpha} - (Z'Z)^{-1}R' \Big[ R(Z'Z)^{-1}R' \Big]^{-1} R\hat{\alpha}.$$
(3.4)

#### 3.3 Stein-like Shrinkage Estimator

We use the restricted estimator and the unrestricted estimator to construct a Stein-like shrinkage estimator. Then, we show that the proposed Stein-like shrinkage estimator improves efficiency. The improved efficiency is a result of making an appropriate trade-off between the bias due to possible incorrect restrictions and variance efficiency gains from imposing the restrictions.

Our proposed Stein-like shrinkage estimator of the slope coefficients, denoted by  $\check{\alpha}$ , is a weighted

$$\underset{m(k+1)\times(m+1)(k+1)}{R} = \begin{bmatrix} -I_{k+1} & I_{k+1} & 0 & 0 & 0 & 0 \\ 0 & -I_{k+1} & I_{k+1} & 0 & 0 & 0 \\ & & \vdots & \vdots & & \\ 0 & 0 & 0 & & -I_{k+1} & I_{k+1} & 0 \\ 0 & 0 & 0 & 0 & & 0 & -I_{k+1} & I_{k+1} \end{bmatrix}.$$

<sup>&</sup>lt;sup>7</sup>In the case of m breaks in the model of (2.1), a restriction matrix that ignores breaks in all slope parameters takes the following form

average of the unrestricted estimator and the restricted estimator defined as

$$\check{\alpha} = \omega \,\,\hat{\alpha} + (1 - \omega) \,\,\tilde{\alpha},\tag{3.5}$$

where the weight takes the form

$$\omega = \left(1 - \frac{\tau}{D(\hat{\alpha}, \tilde{\alpha})}\right), \qquad D(\hat{\alpha}, \tilde{\alpha}) = (\hat{\alpha} - \tilde{\alpha})' Z' Z(\hat{\alpha} - \tilde{\alpha}), \tag{3.6}$$

where  $D(\hat{\alpha}, \tilde{\alpha})$  measures the weighted distance between  $\hat{\alpha}$  and  $\tilde{\alpha}$ , and  $\tau$  is a positive shrinkage parameter that controls the degree of shrinkage. We will defer describing the optimal choice for shrinkage parameter in Section 4. One may consider the positive part version of the weight  $(\omega)_{+} = \omega \mathbb{1}(\omega \geq 0)$  which ensures the weight is bounded between zero and one. Let  $\check{\alpha}_{+}$  denotes the Stein-like shrinkage estimator with the positive part weight. Then as shown by Hansen (2016), the Risk of  $\check{\alpha}_{+}$  is strictly smaller than that of  $\check{\alpha}$ . Hence, in the Monte Carlo simulations and the empirical study of the paper in Sections 6 and 7, we use the positive part version of the weight.

The shrinkage estimator defined above shrinks the unrestricted estimator towards the restricted estimator by the ratio  $\tau/D(\hat{\alpha}, \tilde{\alpha})$ . When the difference between these two estimators is small  $(D(\hat{\alpha}, \tilde{\alpha})$  is small, and  $(1 - \omega)$  is large), the shrinkage estimator gives a larger weight to the restricted estimator, as it is the most efficient estimator. However, when the difference between the two estimators is substantial or large  $(D(\hat{\alpha}, \tilde{\alpha}) > \tau)$ , the bias of the restricted estimator could be more than its variance efficiency gain. Thus, the shrinkage estimator becomes a weighted average of the restricted estimator and the unrestricted estimator, while giving a larger weight to the unrestricted estimator. Therefore, the shrinkage estimator prevails regardless of the break size.<sup>8</sup> We also show this theoretically in the following section.

## 4 Large-Sample Approximation of the Shrinkage Estimator

We employ the large-sample approximation method developed by Nagar (1959), to analyze the bias, MSE, and risks of the shrinkage estimator (conditional on X) under Assumption 1. To find

<sup>&</sup>lt;sup>8</sup>We note that if there is no structural breaks but a structural break is assumed, the Stein-like shrinkage estimator assigns a large weight to the restricted estimator and a small weight to the unrestricted estimator. At the same time, since  $D(\hat{\alpha}, \tilde{\alpha})$  is very small in this case, the weight is likely to be negative, hence the weight assigned to the unrestricted estimator is zero (using the positive part weight), while a full weight of one is assigned to the restricted estimator. On the other hand, if there is no structural breaks and no structural break is assumed, then the restricted estimator, the unrestricted estimator, and the Stein-like shrinkage estimator are all the same.

moment approximations using the Nagar approach, we begin by expressing the estimation error in term of stochastic components which are of decreasing order of magnitude in terms of the sample size.<sup>9</sup>

**Theorem 1** Under Assumption 1, the bias of the Stein-like shrinkage estimator up to order  $T^{-1}$  is

$$Bias(\check{\alpha}) = \mathbb{E}(\check{\alpha} - \alpha) = \Theta - \frac{\tau}{\phi} \bar{P}\alpha, \tag{4.1}$$

where  $\phi = \alpha' \bar{P}' Q^{-1} \bar{P} \alpha = O(T), \ \bar{P} = QR'(RQR')^{-1}R = O(1), \ Q = [\mathbb{E}(Z'Z)]^{-1} = O(T^{-1}),$ and the first term above is the bias of the unrestricted estimator up to order  $O(T^{-1})$  which is  $\Theta = -Q \sum_{i=1}^{2} \left[ \bar{Z}' L_i C \Omega_u \bar{Z} Q e_{1,i} + e_{1,i} tr(Q \bar{Z}' L_i C \Omega_u \bar{Z}) + 2e_{1,i} e'_{1,i} Q e_{1,i} tr(L_i G \Omega_\nu G' L_i C \Omega_u) \right], \ C = (C'_1, C'_2)',$  and the MSE of the Stein-like shrinkage estimator up to order  $T^{-2}$  is

$$MSE\left(\breve{\alpha}\right) = \mathbb{E}\left[\left(\breve{\alpha} - \alpha\right)(\breve{\alpha} - \alpha)'\right] = MSE\left(\widehat{\alpha}\right) + \frac{\tau^2}{\phi^2}\bar{P}\alpha\alpha'\bar{P}' - \frac{\tau}{\phi}\left[\bar{P}Q\Sigma + \Sigma Q\bar{P}' + \Theta\alpha'P' + \bar{P}\alpha\Theta' + \Psi + \Psi'\right] + \frac{2\tau}{\phi^2}\left[\bar{P}\alpha\alpha'\bar{P}'Q^{-1}\left[\bar{P}Q\Sigma + \Psi\right] + \left[\Sigma Q\bar{P}' + \Psi'\right]Q^{-1}\bar{P}\alpha\alpha'\bar{P}'\right] + \frac{\tau}{\phi^2}\left[\bar{P}\alpha\alpha'\bar{P}'\Phi + \Phi'\bar{P}\alpha\alpha'\bar{P}'\right],$$

$$(4.2)$$

where  $\Psi = (\bar{P} - I)Q\Phi$ ,  $\Phi = \sum_{i=1}^{2} e'_{1,i}\bar{P}\alpha\bar{Z}'L_iC\Omega_u\bar{Z}Q + \sum_{i=1}^{2} e_{1,i}\alpha'\bar{P}'\bar{Z}'L_iC\Omega_u\bar{Z}Q + 2\sum_{i=1}^{2}\sum_{j=1}^{2}\sigma_j^2e_{1,i}e'_{1,j}Q tr(G\Omega_{\nu}G'L_iCL_j)e'_{1,i}\bar{P}\alpha$ , and MSE ( $\hat{\alpha}$ ) is given in equation (B.20) in the Supplementary Online Appendix B.

Further, for any fixed symmetric positive definite weight matrix W of order O(T), the risk of the Stein-like shrinkage estimator up to order  $T^{-1}$  is

$$Risk\ (\breve{\alpha}) = \mathbb{E}\left[(\breve{\alpha} - \alpha)'W(\breve{\alpha} - \alpha)\right] = Risk\ (\widehat{\alpha}) + \frac{\tau^2}{\phi^2}\alpha'\bar{P}'W\bar{P}\alpha - \frac{2\tau}{\phi}\left[tr(W\bar{P}Q\Sigma) + \alpha'\bar{P}'W\Theta + tr(W\Psi)\right] + \frac{2\tau}{\phi^2}\left[2\alpha'\bar{P}'Q^{-1}\bar{P}Q\Sigma W\bar{P}\alpha + 2\alpha'\bar{P}'Q^{-1}\Psi W\bar{P}\alpha + \alpha'\bar{P}'\Phi W\bar{P}\alpha\right],$$

$$(4.3)$$

where  $\Sigma = diag(\sigma_1^2 I_{k+1}, \sigma_2^2 I_{k+1})$ , and Risk  $(\widehat{\alpha}) = tr(MSE(\widehat{\alpha})W)$ .

<sup>&</sup>lt;sup>9</sup>For example, for an estimator  $\hat{a}$  of the unknown parameter a, we determine a positive constant g such that we have  $T^{g}(\hat{a}-a) = \varepsilon_0 + T^{-1/2}\varepsilon_1 + T^{-1}\varepsilon_2 + T^{-3/2}\varepsilon_3 + \cdots + T^{-r/2}\varepsilon_r + O_p(T^{-(r+1)/2})$ , where  $\varepsilon_j$  for  $j = 0, \ldots, r$  are all  $O_p(1)$  as  $T \to \infty$ . The value of g for least-squares estimators in stationary models is 1/2, but it may take other values in non-stationary models (e.g., see Remarks 5–6).

Theorem 1 gives the bias, MSE, and Risk of the Stein-like shrinkage estimator. The bias and MSE of the unrestricted estimator is given in Lemma B.3 in the Supplementary Online Appendix. We note that the risk of the Stein-like shrinkage estimator is generalized to allow for any positive definite weight matrix, W, of order O(T). Two arbitrary choices of W are  $TI_{2(k+1)}$ , and  $W_f$  defined in Section 4.1, where the former provides an unweighted mean squared error, and the latter gives the mean squared forecast error studied in the next section.

**Remark 3** The large sample approximations of the bias and MSE of the unrestricted estimator are provided in the Supplementary Online Appendix B. To the best of our knowledge, this is the first paper that provides these theoretical results under structural breaks.<sup>10</sup>

In the following Corollary we show the dominance conditions of the Stein-like shrinkage estimator relative to the unrestricted estimator. Let  $\phi_W = \alpha' \bar{P}' W \bar{P} \alpha$ ,  $\mu = \text{tr}(W \bar{P} Q \Sigma) + \alpha' \bar{P}' W \Theta + \text{tr}(W\Psi)$ , and  $\eta = 2\alpha' \bar{P}' Q^{-1} \bar{P} Q \Sigma W \bar{P} \alpha + 2\alpha' \bar{P}' Q^{-1} \Psi W \bar{P} \alpha + \alpha' \bar{P}' \Phi W \bar{P} \alpha$ .

**Corollary 1.1** Under Assumption 1, if  $\mu > \frac{1}{\phi}\eta$ , and  $0 < \tau \leq \frac{2\phi}{\phi_W} \left(\mu - \frac{1}{\phi}\eta\right)$ , then the risk of the Stein-like shrinkage estimator up to order  $T^{-1}$  is

 $Risk\ (\breve{\alpha}) < Risk\ (\widehat{\alpha}). \tag{4.4}$ 

In addition, the optimal shrinkage parameter, denoted by  $\tau_{opt}$ , that minimizes the risk of the Stein-like shrinkage estimator up to order  $T^{-1}$ , is

$$\tau_{opt} = \frac{\phi}{\phi_W} \left( \mu - \frac{1}{\phi} \eta \right). \tag{4.5}$$

Therefore, the risk of the optimal Stein-like shrinkage estimator up to order  $T^{-1}$  is<sup>11</sup>

$$Risk\left(\breve{\alpha}_{opt}\right) = Risk\left(\widehat{\alpha}\right) - \frac{1}{\phi_W}\left(\mu - \frac{1}{\phi}\eta\right)^2.$$
(4.6)

Proof: See Appendix A, page 33.

<sup>&</sup>lt;sup>10</sup>The detailed derivations of large sample approximations of bias and MSE of the restricted estimator are also available upon request.

<sup>&</sup>lt;sup>11</sup>The Stein-like shrinkage estimator defined in (3.5) using the optimal shrinkage parameter,  $\tau_{opt}$ .

Corollary 1.1 shows that the proposed shrinkage estimator dominates the unrestricted estimator in terms of having a smaller risk when the shrinkage parameter satisfies the condition  $0 < \tau \leq \frac{2\phi}{\phi_W} \left(\mu - \frac{1}{\phi}\eta\right)$ . In addition, as the choice of the shrinkage parameter is user-specified, its optimal value and our ideal choice that minimizes the risk of the Stein-like shrinkage estimator up to  $O(T^{-1})$ , is  $\tau_{opt}$ . This implies that the Stein-like shrinkage estimator always performs better or performs equal to the unrestricted estimator which is one of the common methods of estimating the coefficients under structural breaks.<sup>12</sup>

Since  $\tau_{opt}$  depends on  $\sigma_1^2$ ,  $\sigma_2^2$ , Q,  $\bar{Z}$ , C, G,  $\omega$ ,  $\alpha$ , and  $\bar{y}_0$ , which are unobserved, we consider the estimated optimal shrinkage parameter, denoted by  $\hat{\tau}_{opt}$ , defined as

$$\hat{\tau}_{opt} = \frac{\hat{\phi}}{\hat{\phi}_W} \left( \hat{\mu} - \frac{1}{\hat{\phi}} \hat{\eta} \right)_+.$$
(4.7)

In equation (4.7),  $\hat{\mu}$ ,  $\hat{\eta}$ ,  $\hat{\phi}$  and  $\hat{\phi}_W$  correspond to  $\mu$ ,  $\eta$ ,  $\phi$  and  $\phi_W$  respectively, after replacing  $(\sigma_i^2, \alpha, Q, \bar{Z}, C, G)$  with their estimates, denoted by  $(\hat{\sigma}_i^2, \hat{\alpha}, \hat{Q}, \hat{Z}, \hat{C}, \hat{G})$ , where  $\hat{\alpha}$  and  $\hat{\sigma}_i^2$  are the unrestricted estimators,  $\hat{Q} = (Z'Z)^{-1}$ ,  $\hat{Z} = diag(\hat{Z}_1, \hat{Z}_2)$ , with  $\hat{Z}_i = (\hat{F}_i y_0 + \hat{C}_i X_i \hat{\beta}_i, X_i)$ ,  $\hat{C}$  and  $\hat{F}$  correspond to  $C = (C'_1, C'_2)'$  and  $F = (F'_1, F'_2)'$  after replacing  $\lambda_1$ , and  $\lambda_2$  by  $\hat{\lambda}_1$ , and  $\hat{\lambda}_2$ . Similarly,  $\hat{\Theta}$ ,  $\hat{\Psi}$ , and  $\hat{P}$  correspond to  $\Theta$ ,  $\Psi$ , and  $\bar{P}$  after replacing the unobserved parameters with their estimates. Since we define the Stein-like shrinkage estimator with positive shrinkage parameter in (3.5), we set  $\hat{\tau}_{opt}$  to zero when the condition  $\hat{\mu} > \frac{1}{\hat{\phi}}\hat{\eta}$  does not hold. In this case, the Stein-like shrinkage estimator and a zero weight to the restricted estimator.

In the following corollary, we show that the estimated optimal shrinkage parameter,  $\hat{\tau}_{opt}$ , is an unbiased estimator of the infeasible optimal shrinkage parameter  $\tau_{opt}$  up to order  $T^{-1}$ . Hence, when the sample size is large enough the risk of the Stein-like shrinkage estimator using the estimated optimal shrinkage parameter is smaller than the risk of the unrestricted estimator.

<sup>&</sup>lt;sup>12</sup>We note that since the model in (2.1) includes dynamic structures, the unrestricted estimator is biased and the terms that contribute to the bias will contribute to the risk of the Stein-like shrinkage estimator as well. Thus, the dominance condition,  $\tau > 0$ , or equivalently  $\mu > \eta/\phi$ , is more mathematically involved compared to the models without the dynamic structures (classical models). If the dynamic structure of the model is excluded, it can be easily verified that the proposed Stein-like shrinkage estimator dominates the unrestricted estimator in terms of having a smaller risk when the number of restrictions is more than 2 (which is the same as the classical Stein dominance condition).

**Corollary 1.2** Under Assumption 1, if  $\mu > \frac{1}{\phi}\eta$ , we have

$$\mathbb{E}(\hat{\tau}_{opt}) = \tau_{opt} + O(T^{-1}). \tag{4.8}$$

Proof: See Appendix A, page 33.

**Remark 4** The risk of the Stein-like shrinkage estimator presented in Theorem 1 is noticeably different from the risk of the proposed estimators derived in Lee et al. (2022a), Lee et al. (2022b), and Lee et al. (2022c). In this paper, we derive the large sample approximation of risk for any break size, while Lee et al. (2022b) derive the asymptotic risk for their proposed combined estimator under local-to-zero asymptotic framework. Also, we consider a dynamic regression model, while the model considered in Lee et al. (2022a) and Lee et al. (2022c) is a static stationary model. Furthermore, we allow for non-stationary regressors and unit roots (discussed in Section 5) which have not been considered in Lee et al. (2022a), Lee et al. (2022b), and Lee et al. (2022c).

#### 4.1 Forecasting under Structural Instability

In this section, we explain how the proposed shrinkage estimator can be used for the out-of-sample forecasting. The true parameters that enter the forecasting period are the coefficients in the most recent regime. Thus, we define a selection matrix  $S = [0_{(k+1)\times(k+1)}, I_{k+1}]$  to select the post-break slope parameters. By pre-multiplying the selection matrix to the shrinkage estimator, we get  $S\breve{\alpha} = \omega S\hat{\alpha} + (1 - \omega) S\tilde{\alpha}$ , where for example  $\breve{\alpha}_2 = S\breve{\alpha}$  denotes the estimated post-break slope parameters of the shrinkage estimator.

We define the one-step-ahead mean squared forecast error (MSFE) of the shrinkage estimator as

$$MSFE(\breve{\alpha}_2) = \mathbb{E}\left[(\breve{\alpha} - \alpha)'W_f(\breve{\alpha} - \alpha)\right],\tag{4.9}$$

where  $W_f = TS' z'_{2,T+1} z_{2,T+1} S$ , where  $z_{2,T+1} = (y_T, x'_{T+1})$ . Since  $W_f$  contains lagged dependent variables and is random, we give the MSFE in the following theorem.

**Theorem 2** Under Assumption 1, the MSFE of the Stein-like shrinkage estimator up to order  $T^{-1}$ 

$$MSFE(\check{\alpha}_{2}) = MSFE(\hat{\alpha}_{2}) + \frac{\tau^{2}}{\phi^{2}} \alpha' \bar{P}' \bar{W}_{f} \bar{P} \alpha - \frac{2\tau}{\phi} \bigg[ tr(\bar{W}_{f} \bar{P} Q \Sigma) + \alpha' \bar{P}' \bar{W}_{f} \Theta + tr(\bar{W}_{f} \Psi) + \alpha' \bar{P}' \Upsilon_{1} + \alpha' \bar{P}' \Upsilon_{3} + tr(\bar{P} \Upsilon_{2}) + tr \left( (\bar{P} - I_{2(k+1)}) Q \Upsilon_{4} \right) \bigg] + \frac{2\tau}{\phi^{2}} \bigg[ 2\alpha' \bar{P}' Q^{-1} \bar{P} Q \Sigma \bar{W}_{f} \bar{P} \alpha + 2\alpha' \bar{P}' Q^{-1} \Psi \bar{W}_{f} \bar{P} \alpha + \alpha' \bar{P}' \Phi \bar{W}_{f} \bar{P} \alpha + \alpha' \bar{P}' \Upsilon_{4} \bar{P} \alpha + 2\alpha' \bar{P}' Q^{-1} \bar{P} \Upsilon_{2} \bar{P} \alpha + 2\alpha' \bar{P}' Q^{-1} (\bar{P} - I_{2(k+1)}) Q \Upsilon_{4} \bar{P} \alpha \bigg],$$

$$(4.10)$$

where  $\overline{W}_f = T \mathbb{E}(S' z'_{2,T+1} z_{2,T+1} S)$ , and the expressions for  $\Upsilon_1 - \Upsilon_4$  are given in equations (A.19) - (A.22).

Proof: See Appendix A, page 34.

The following Corollary shows the conditions under which the MSFE of the Stein-like shrinkage estimator is less than the MSFE of the post-break unrestricted estimator ( $\hat{\alpha}_2 = S\hat{\alpha}$ ). Before stating the Corollary, we define some notations. Let  $\phi_f = \alpha' \bar{P}' \bar{W}_f \bar{P} \alpha$ ,  $\mu_f = \text{tr}(\bar{W}_f \bar{P} Q \Sigma) + \alpha' \bar{P}' \bar{W}_f \Theta + \text{tr}(\bar{W}_f \Psi) + \alpha' \bar{P}' \Upsilon_1 + \alpha' \bar{P}' \Upsilon_3 + \text{tr}(\bar{P} \Upsilon_2) + \text{tr}((\bar{P} - I_{2(k+1)})Q \Upsilon_4)$ , and  $\eta_f = 2\alpha' \bar{P}' Q^{-1} \bar{P} Q \Sigma \bar{W}_f \bar{P} \alpha + 2\alpha' \bar{P}' Q^{-1} \Psi \bar{W}_f \bar{P} \alpha + \alpha' \bar{P}' \Phi \bar{W}_f \bar{P} \alpha + \alpha' \bar{P}' \Upsilon_4 \bar{P} \alpha + 2\alpha' \bar{P}' Q^{-1} \bar{P} \Upsilon_2 \bar{P} \alpha + 2\alpha' \bar{P}' Q^{-1} (\bar{P} - I_{2(k+1)})Q \Upsilon_4 \bar{P} \alpha$ .

**Corollary 2.1** Under Assumption 1, if  $\mu_f > \frac{1}{\phi}\eta_f$ , and  $0 < \tau \leq \frac{2\phi}{\phi_f}(\mu_f - \frac{1}{\phi}\eta_f)$ , then the MSFE of the Stein-like shrinkage estimator up to order  $T^{-1}$  is

$$MSFE(\check{\alpha}_2) < MSFE(\hat{\alpha}_2). \tag{4.11}$$

In addition, the optimal shrinkage parameter, denoted by  $\tau_{opt}^{f}$ , that minimizes the MSFE of the Stein-like shrinkage estimator up to order  $T^{-1}$ , is

$$\tau_{opt}^f = \frac{\phi}{\phi_f} \left( \mu_f - \frac{1}{\phi} \eta_f \right). \tag{4.12}$$

Proof: See Appendix A, page 37.

We note that the optimal shrinkage parameter is data dependent. Therefore, when it is used for forecast evaluations, it needs to be calculated for each estimation window of the sample (rolling or recursive window).

15

#### 4.2 Non-Stationary Regressors

Our analysis can also allow for the unit root and non-stationary regressors. It is possible for the  $y_t$  process to contain a unit root (be integrated of order 1, I(1)) or be integrated of higher order in both of the regimes. In this section, we demonstrate the results of Theorem 1 while we relax Assumption 1(ii) on the stationarity of the exogenous regressors. We extend the results of Theorem 1 by relaxing Assumptions 1(i)-(ii) on the stability, and the stationarity of the exogenous regressors in Section 5.

Following Kiviet and Phillips (2005), we rescale the regressors and the coefficients so that all elements of the estimation error vector are of the same stochastic magnitude. We assume that for j = 1, ..., k + 1, the series of real positive constants  $g_j$  is given, such that  $\dot{Z}'\dot{Z} = O_p(T)$ , where  $\dot{Z} = Z N$ , and  $N = I_2 \otimes diag(T^{-g_1}, ..., T^{-g_{k+1}})$ . In practice, one needs to determine the orders of integration. The Augmented Dickey-Fuller test statistic is a commonly used method to determine the order of integration of a time series. The basic idea is to test for the level of difference at which the series is stationary. For more discussion and alternative approaches, we refer the readers to chapter 14 of Gourieroux and Monfort (1997), and Smeekes and Wijler (2020).

**Remark 5 Non-Stationary Regressors:** In the model of equation (2.1), if Assumption 1(i) holds but the  $l^{th}$  column of exogenous regressors is a linear trend or an I(1) process while others are stationary, then we have  $g_1 = 1$ , and  $g_{l+1} = 1$ , while  $g_j = 0$  for j = 2, ..., l, l+2, ..., k+1.

**Remark 6 Unit Root, and Non-Stationary Regressors:** In the unit root model of equation (2.1) (i.e.  $\lambda_1 = \lambda_2 = 1$ ), if the exogenous regressors are stationary, then we have  $g_1 = 1$ , and  $g_j = 0$ , for j = 2, ..., k + 1. When the  $l^{th}$  column of X is a linear trend or an I(1) process, then  $g_1 = 2$  because of the unit root,  $g_{l+1} = 1$ , and  $g_j = 0$  for j = 2, ..., l, l+2, ..., k+1.

Using the rescaled regressors, we write the model in (2.2) as

$$y = ZN(N^{-1}\alpha) + u, \tag{4.13}$$

where  $N^{-1}\alpha$ , are the rescaled coefficients. Moreover, the rescaled unrestricted estimator can be formulated as

$$N^{-1}(\hat{\alpha} - \alpha) = (\dot{Z}'\dot{Z})^{-1}\dot{Z}'u = N^{-1}(Z'Z)^{-1}Z'u,$$
(4.14)

the rescaled restricted estimator can be formulated as

$$N^{-1}(\tilde{\alpha} - \alpha) = N^{-1}(\hat{\alpha} - \alpha) - (\dot{Z}'\dot{Z})^{-1}R' \Big[ R(\dot{Z}'\dot{Z})^{-1}R' \Big]^{-1}RN^{-1}\hat{\alpha}$$
  
=  $N^{-1}(\hat{\alpha} - \alpha) - N^{-1}(Z'Z)^{-1}R' \Big[ R(Z'Z)^{-1}R' \Big]^{-1}R\hat{\alpha},$  (4.15)

and the rescaled Stein-like shrinkage estimator is

$$N^{-1}(\check{\alpha} - \alpha) = \omega N^{-1}(\hat{\alpha} - \alpha) + (1 - \omega) N^{-1}(\check{\alpha} - \alpha).$$

$$(4.16)$$

Therefore, in order to find the bias and MSE of the Stein-like shrinkage estimator, we need to pre-multiply the bias of the estimator in Theorem 1 by N, and pre- and post- multiply the MSE of the estimator by N.

**Theorem 3** The bias and MSE formulas of the Stein-like shrinkage estimator in Theorem 1 also apply when the exogenous regressors contain non-stationary components. However, the  $j^{th}$  element of bias is approximated up to order  $T^{-1-g_j}$ , and the  $(j,l)^{th}$  element of MSE is approximated up to order  $T^{-2-g_j-g_l}$ .

Proof: The proof follows from proof of Theorem 1, and we omit it for brevity.

## 5 Unit Root

In this section, we extend the results of Theorem 1 by relaxing Assumption 1(i)-(ii) on the stability, and the stationarity of the exogenous regressors. We assume that for j = 1, ..., k + 1, the series of real positive constants  $g_j$  is given, such that  $\dot{Z}'\dot{Z} = O_p(T)$ , where  $\dot{Z} = ZN$ , and  $N = I_2 \otimes$  $diag(T^{-g_1}, ..., T^{-g_{k+1}})$ .

**Theorem 4** Under Assumption 1(iii)-(vii), when the coefficients of the lagged dependent variable is equal to unity, the bias of the rescaled Stein-like shrinkage estimator up to order  $T^{-1}$  is

$$Bias(\check{\alpha}) = \mathbb{E}(\check{\alpha} - \dot{\alpha}) = \dot{\Theta} - \frac{\tau}{\dot{\phi}}\dot{P}\dot{\alpha},\tag{5.1}$$

where  $\dot{\phi} = \dot{\alpha}' \dot{P}' \dot{Q}^{-1} \dot{P} \dot{\alpha} = O(T), \ \dot{P} = \dot{Q}R'(R\dot{Q}R')^{-1}R = O(1), \ \dot{Q} = [\mathbb{E}(\dot{Z}')\mathbb{E}(\dot{Z})]^{-1} = O(T^{-1}),$ and the first term above is the bias of the unrestricted estimator up to order  $T^{-1}$  which is  $\dot{\Theta} = C^{-1}$   $-\dot{Q}\sum_{i=1}^{2} \left[ \dot{Z}'L_i \dot{C}\Omega_u \dot{Z} \dot{Q} N e_{1,i} + N e_{1,i} \operatorname{tr}(Q \dot{Z}'L_i \dot{C}\Omega_u \dot{Z}) \right], \text{ and } \dot{C} \text{ has zeroes on and above its main diagonal and components unity below. Also, the MSE of the rescaled Stein-like shrinkage estimator up to order <math>T^{-2}$  is

$$MSE\left(\check{\alpha}\right) = \mathbb{E}\left[\left(\check{\alpha}-\dot{\alpha}\right)\left(\check{\alpha}-\dot{\alpha}\right)'\right] = MSE\left(\widehat{\alpha}\right) + \frac{\tau^{2}}{\dot{\phi}^{2}}\bar{P}\dot{\alpha}\dot{\alpha}'\bar{P}' - \frac{\tau}{\dot{\phi}}\left[\bar{P}\dot{Q}\Sigma + \Sigma\dot{Q}\bar{P}' + \dot{\Theta}\dot{\alpha}'\dot{P}' + \bar{P}\dot{\alpha}\dot{\Theta}' + \dot{\Psi}\dot{Q}\dot{Q}^{-1}\left[\bar{P}\dot{Q}\Sigma + \dot{\Psi}\right] + \left[\Sigma\dot{Q}\bar{P}' + \dot{\Psi}'\right]\dot{Q}^{-1}\bar{P}\dot{\alpha}\dot{\alpha}'\bar{P}'\right] + \frac{\tau}{\dot{\phi}^{2}}\left[\bar{P}\dot{\alpha}\dot{\alpha}'\bar{P}'\dot{\Phi} + \dot{\Phi}'\bar{P}\dot{\alpha}\dot{\alpha}'\bar{P}'\right],$$

$$(5.2)$$

where  $\dot{\Psi}$  and  $\dot{\Phi}$  are given in equations (A.35)–(A.36), and MSE  $(\hat{\alpha})$  is given in equation (C.5) in the Supplementary Online Appendix C.

Further, for the fixed symmetric positive definite weight matrix W of order O(T), the risk of the Stein-like shrinkage estimator up to order  $T^{-1}$  is

$$\begin{aligned} Risk\ (\check{\alpha}) &= \mathbb{E}\left[(\check{\alpha} - \dot{\alpha})'W(\check{\alpha} - \dot{\alpha})\right] = Risk\ (\widehat{\alpha}) + \frac{\tau^2}{\dot{\phi}^2}\dot{\alpha}'\bar{P}'W\bar{P}\dot{\alpha} - \frac{2\tau}{\dot{\phi}}\left[tr(W\bar{P}\dot{Q}\Sigma) + \dot{\alpha}'\bar{P}'W\dot{\Theta} + tr(W\dot{\Psi})\right] \\ &+ \frac{2\tau}{\dot{\phi}^2}\left[2\dot{\alpha}'\bar{P}'\dot{Q}^{-1}\bar{P}\dot{Q}\Sigma W\bar{P}\dot{\alpha} + 2\dot{\alpha}'\bar{P}'\dot{Q}^{-1}\dot{\Psi}W\bar{P}\dot{\alpha} + \dot{\alpha}'\bar{P}'\dot{\Phi}W\bar{P}\dot{\alpha}\right], \end{aligned}$$

$$(5.3)$$

where  $\Sigma = diag(\sigma_1^2 I_{k+1}, \sigma_2^2 I_{k+1})$ , and Risk  $(\hat{\alpha}) = tr(MSE(\hat{\alpha})W)$ . Proof: See Appendix A, page 37.

**Remark 7** We note that the bias and MSE of the Stein-like shrinkage estimator can be obtained using the bias and MSE of the rescaled Stein-like shrinkage estimator as follows:

 $Bias(\breve{\alpha}) = \mathbb{E}(\breve{\alpha} - \alpha) = NBias(\breve{\dot{\alpha}}),$ 

$$MSE(\breve{\alpha}) = \mathbb{E}\left[(\breve{\alpha} - \alpha)(\breve{\alpha} - \alpha)'\right] = N MSE(\breve{\alpha})N,$$

where the  $j^{th}$  element of bias is approximated up to order  $T^{-1-g_j}$ , and the  $(j,l)^{th}$  element of MSE is approximated up to order  $T^{-2-g_j-g_l}$ .

Similar to Section 4.1, we define the one-step-ahead mean squared forecast error (MSFE) of the

shrinkage estimator as

$$MSFE(\check{\alpha}_2) = \mathbb{E}\left[(\check{\alpha} - \alpha)'W_f(\check{\alpha} - \alpha)\right] = \mathbb{E}\left[(\check{\alpha} - \dot{\alpha})'\dot{W}_f(\check{\alpha} - \dot{\alpha})\right],\tag{5.4}$$

where  $\dot{W}_f = NW_f N = O_p(T)$ . In the following Theorem, we give the MSFE of the Stein-like shrinkage estimator when the model contains a unit root.

**Theorem 5** Under Assumption 1(ii)–(vii), when the coefficients of the lagged dependent variable is equal to unity, the MSFE of the Stein-like shrinkage estimator up to order  $T^{-1}$  is

$$MSFE(\check{\alpha}_{2}) = MSFE(\hat{\alpha}_{2}) + \frac{\tau^{2}}{\phi^{2}} \alpha' \bar{P}' \bar{W}_{f} \bar{P} \alpha - \frac{2\tau}{\phi} \bigg[ tr(\bar{W}_{f} \bar{P} Q \Sigma) + \alpha' \bar{P}' \bar{W}_{f} \Theta + tr(\bar{W}_{f} \Psi) + \alpha' \bar{P}' \bar{\Upsilon} \bigg] + \frac{2\tau}{\phi^{2}} \bigg[ 2\alpha' \bar{P}' Q^{-1} \bar{P} Q \Sigma \bar{W}_{f} \bar{P} \alpha + 2\alpha' \bar{P}' Q^{-1} \Psi \bar{W}_{f} \bar{P} \alpha + \alpha' \bar{P}' \Phi \bar{W}_{f} \bar{P} \alpha \bigg],$$

$$(5.5)$$

where  $\bar{W}_f = T \mathbb{E}(S'z'_{2,T+1}) \mathbb{E}(z_{2,T+1}S)$ , and the expression for  $\bar{\Upsilon} = N^{-1}\dot{\Upsilon}$  with  $\dot{\Upsilon}$  given in equation (A.40).

Proof: See Appendix A, page 41.

## 6 Monte Carlo Simulation

In this section, we report the finite sample performance of our proposed Stein-like shrinkage estimator. We compare the forecasting performance of the shrinkage estimator with those of the unrestricted estimator and the restricted estimator. We consider the following data generating process (DGP)

$$y_{t} = \begin{cases} a_{1} + \lambda_{1} y_{t-1} + x_{t}' \beta_{1} + \sigma_{1} \epsilon_{t}, & \text{for } t \leq T_{1} \\ a_{2} + \lambda_{2} y_{t-1} + x_{t}' \beta_{2} + \sigma_{2} \epsilon_{t}, & \text{for } t > T_{1}, \end{cases}$$
(6.1)

where  $a_1 = (1 - \lambda_1)$  is the intercept in the first regime,  $x_t$  follows a multivariate normal distribution with mean zero and unit variance, and  $\epsilon_t \sim \text{i.i.d. N}(0, 1)$ .<sup>13</sup>  $\beta_1$  is a  $k \times 1$  unit vector of slope parameters of the exogenous regressors. We consider  $T = 100, b_1 \equiv T_1/T \in \{0.2, 0.4, 0.6, 0.8\}$ , and

<sup>&</sup>lt;sup>13</sup>We also consider the case where the error term is distributed from a non-normal distribution. These simulation results are reported in Appendix D of the Supplementary Online Appendix, and are consistent with the results reported here under normal distributions.

 $k \in \{3, 5, 8\}$ , where the results for  $k \in \{3, 8\}$  are available in the Supplementary Online Appendix D. The break size in the slope coefficients of the exogenous regressors,  $\delta \iota_k \equiv \beta_1 - \beta_2$ , takes values of  $\delta \in \{0, 0.25, 0.5, 0.75, 1\}$ , where  $\iota_k$  is a  $k \times 1$  vector of ones, and the intercept in the second regime is  $a_2 = (1 + \delta)(1 - \lambda_2)$ .<sup>14</sup> For the initial observation,  $y_0$ , we consider a fixed start-up by setting d = 0, and  $\bar{y}_0 = a_1/(1 - \lambda_1)$ .

For the break in the autoregressive slope coefficients, we study different experiments summarized in Table 1. Experiment 1 considers no break in the autoregressive parameters. Small breaks, moderate breaks, and large breaks in the autoregressive slope coefficients in either direction are considered in Experiments 2–7. Higher post break volatility ( $\sigma_2 > \sigma_1$ ) and lower post-break volatility ( $\sigma_2 < \sigma_1$ ) are considered in Experiments 8 and 9. Experiments 10–12 consider a unit root process ( $\lambda_1 = \lambda_2 = 1$ ) with no break in the error variance, higher post break volatility, and lower post break volatility, respectively. In all simulations, we estimate the unknown parameters, such as the break point, and the break size in the slope coefficients and in the error variance.

Table 1: Break specifications

Experiments	$\lambda_1$	$\lambda_2$	$\sigma_1$	$\sigma_2$
#1 : No break	0.9	0.9	1	1
#2 : Small break in $\lambda$	0.8	0.9	1	1
#3 : Small break in $\lambda$ (decline)	0.9	0.8	1	1
#4 : Moderate break in $\lambda$	0.3	0.6	1	1
#5 : Moderate break in $\lambda$ (decline)	0.6	0.3	1	1
#6 : Large break in $\lambda$	0.2	0.8	1	1
#7 : Large break in $\lambda$ (decline)	0.8	0.2	1	1
#8: Higher post-break volatility	0.9	0.9	0.5	1
#9: Lower post-break volatility	0.9	0.9	1	0.5
#10: Unit root	1	1	1	1
#11: Unit root with higher post-break volatility	1	1	0.5	1
#12: Unit root with lower post-break volatility	1	1	1	0.5

The results of 5000 Monte Carlo simulations are reported in Table 2. The first column in each table represents the experiment number based on Table 1, the second column shows the break size in the slope coefficients ( $\delta$ ), and the rest report the relative mean squared forecast error (RMSFE) of the post-break slope coefficients of the shrinkage estimator, and the restricted estimator for

<sup>&</sup>lt;sup>14</sup>The intercepts in both regimes are set equal to zero in the presence of unit roots in  $y_t$ , to avoid the possibility of generating linear trends.

different break points. Thus, the RMSFE of the shrinkage estimator, and the restricted estimator are denoted by  $\text{RMSFE}(\check{\alpha}_2) = \text{MSFE}(\check{\alpha}_2)/\text{MSFE}(\hat{\alpha}_2)$ , and  $\text{RMSFE}(\check{\alpha}_2) = \text{MSFE}(\check{\alpha}_2)/\text{MSFE}(\hat{\alpha}_2)$ , respectively.

The Monte Carlo results support our theoretical findings presented in Section 4. The results show that the RMSFE of our proposed shrinkage estimator is uniformly less than or equal to that of the unrestricted estimator over different break sizes and break points. This shows the superiority of the shrinkage estimator relative to the unrestricted estimator. For small break sizes in the slope coefficients (small  $\delta$ ), the restricted estimator performs better than the unrestricted estimator. This is expected, because under small breaks, the bias of the restricted estimator will be dominated by its variance efficiency. In this case, the shrinkage estimator tends to gain more from the efficiency of the restricted estimator by assigning a larger weight to this estimator, and therefore remains one of the best choices. When there is a large break in the slope coefficients, the shrinkage estimator This happens because under a large break, the restricted estimator has a large bias, so the shrinkage estimator assigns more weight to the unrestricted estimator.

When the break occurs towards the end of the sample (e.g.,  $b_1 = 0.8$ ), there are a few observations in the post-break sample, which results in a poor performance of the unrestricted estimator as it uses fewer observations. consequently, in this case, the RMSFE( $\tilde{\alpha}_2$ ) tends to be smaller relative to the case where the break occurs towards the beginning of the sample (e.g.,  $b_1 = 0.2$ ). However, since the shrinkage estimator is a weighted average of the restricted and unrestricted estimators, it prevails.

When the pre-break sample is less volatile compared to the post-break sample ( $\sigma_1 < \sigma_2$ ), the shrinkage estimator gains from exploiting the less volatile pre-break sample observations. Therefore, the shrinkage estimator provides a better estimation of the post-break slope coefficients and outperforms the unrestricted estimator. On the other hand, when  $\sigma_1 > \sigma_2$ , the gain obtained from exploiting the pre-break sample observations decrease as the pre-break sample is more volatile. In this case, the shrinkage estimator gives a larger weight to the unrestricted estimator. We also find that as the number of regressors increases, the shrinkage estimator performs better. A similar results hold when the dependent variable follows a unit root process in Experiments 10–12.

Furthermore, the results show that for a high degree of dependence or persistency (e.g., experiments #2 and #3 in which we also have a small break in the autoregressive slope coefficients

 $\lambda$ ), the gain obtained by using the Stein-like shrinkage estimator over the unrestricted estimator is around 10%–23.5% for a small to moderate break in the slope coefficients of the exogenous regressors  $(\delta \leq 0.5)$ , i.e., the MSFE decreases around 10%–23.5% by using the Stein-like shrinkage estimator relative to the unrestricted estimator. On the other hand, when there is a large break in the slope coefficients of the exogenous regressors ( $\delta > 0.5$ ), this gain is almost less than 5%. However, it is worth mentioning that the Stein-like shrinkage estimator never under-performs the unrestricted estimator. We see a similar pattern when the coefficient in one of the regime is persistent (e.g., experiments #6 and #7 in which we also have a large break in the autoregressive slope coefficients  $\lambda$ ). When we have a unit root (e.g., experiments #10, #11, and #12), the gain obtained by using the Stein-like shrinkage estimator over the unrestricted estimator is around 7%–15% for a small to moderate break in the slope coefficients of the exogenous regressors. Under these experiments, when there is a large break in the slope coefficients of the exogenous regressors, the Stein-like shrinkage estimator performs equivalent to the unrestricted estimator. This is expected since for a large break size, the bias created by the restricted estimator is greater than its variance efficiency. Thus, the Stein-like shrinkage estimator assigns a weight one to the unrestricted estimator and a zero weight to the restricted estimator. In general, we find that the shrinkage estimator performs well under any break sizes and break points.

## 7 Empirical Analysis

In this section, we present an empirical application that highlights the utility of the proposed shrinkage estimator in forecasting. In particular, we provide empirical analysis of forecasting the output growth using 131 macroeconomic and financial time series from the St. Louis Federal Reserve (FRED-MD) database. The 131 series are split into 8 groups: output and income, labor market, consumption and orders, orders and inventories, money and credit, interest rates and exchange rates, prices and stock market.

We use the monthly data from January 1959 up to March 2020. The data are described by McCracken and Ng (2016), who suggest various transformations to render the series stationary and to deal with missing values. Since the number of macroeconomic and financial variables is large, to reduce dimension, we estimate static factors by principal component analysis adapted to allow for missing values, see McCracken and Ng (2016). We select the number of significant factors using a generalization of the Mallow's  $C_p$  criteria for large dimensional panels developed in Bai and Ng

(2002). The criterion finds eight factors which can be interpreted as real activity/employment, term spreads, inflation, housing, interest rate variables, stock market variables, and output and inventories factors.

We adopt the following h-step ahead forecasting equation

$$y_{t+h} = \lambda_h y_t + \hat{f}_t^{\prime} \beta_h + u_{t+h}, \quad t = 1, \dots, T,$$

$$(7.1)$$

where the dependent variable is the output growth over the next h months, that is,  $y_{t+h} = (1200/h) \ln(\text{IP}_{t+h}/\text{IP}_t)$ ,  $IP_t$  denotes index of industrial production in levels, and  $\hat{f}_t$  is the estimated eight factors at time t.

In order to evaluate the performance of our proposed estimator, we compute *h*-step-ahead forecasts using the shrinkage estimator and compare it with those from the existing methods: the unrestricted estimator (labeled "Unrest." in tables), the restricted estimator ("Rest."), the method proposed by Pesaran et al. (2013) ("PPP"), the methods used in Pesaran and Timmermann (2007), namely, "Troff", "Pooled", "WA", "CV", and the average window forecast proposed by Pesaran and Pick (2011) ("AveW"). The forecasts employ a recursive (or expanding) estimation window. Each time that we expand the estimation window, we apply the Schwarz's Bayesian Information Criteria (BIC) to select among the nine predictors except the lag dependent variable (so the model is dynamic). We estimate the break point using Bai and Perron (1998, 2003) method, by setting the significance level at 5% and trimming rate at 0.2.

Table 3 reports the MSFEs for different methods. The first column in this table shows the forecast horizon, h = 1, 6, and 12. For comparison, we report the MSFE for different methods with estimated break date under the label *estimated break dates* in columns 2–9 of Table 3. Besides, as the "CV", "WA" and "Pooled" methods can also be implemented without an estimation of the break date, we report the MSFE of these methods without estimating the break date. The results without estimating break date, treating the break date as unknown, are reported under the label *unknown break dates* in columns 10–13 of Table 3. We consider different out-of-sample forecast periods (Panels A–D in Table 3) running from 1980:01-2020:03 to 2005:01-2020:03 to see the performmance of different methods with this choice.

We test for equal forecast performance of different methods compared to the unrestricted forecast using the Diebold and Mariano (1995) test. Table 3 also shows the results for the Diebold and Mariano test statistic, indicated by asterisks. The 5% and 10% significance levels are denoted by \*\* and \*, respectively. Based on the results, the shrinkage forecasts perform better than the unrestricted forecasts in the sense of MSFE, over different forecast horizons, and over different out-of-sample forecast periods. The improvement is statistically significant at 5% or 10% levels. The forecast improvement ranges from 2 to 5.3 percent. The other alternative methods often outperform the unrestricted forecast, however, their out-performance is not statistically significant which might be a sign of having high variations. Besides, there are some cases that the alternative methods under-perform the unrestricted forecast.<sup>15</sup>

## 8 Conclusion

We introduce a method of estimation and forecasting in ARX(1) models under structural breaks. The theoretical results can be generalized to models with higher order lags, however, it will not yield qualitatively different results. The proposed method has four main advantages relative to the other model averaging and shrinkage estimation methods existing in the literature. First, our method allows for a unit root and non-stationary exogenous regressors. This is important because it can be used in a large class of economic empirical applications. Second, the dominance and optimality of the shrinkage estimator is not limited to MSE and holds for any weighted quadratic loss function where the weight is positive definite and symmetric. This allows to use the proposed Stein-like shrinkage method for out-of-sample forecasting by choosing an appropriate weight. Third, the averaging weight is proportional to the reciprocal of the weighted distance of the restricted estimator and the unrestricted estimator, which provides a shrinkage estimator with a uniformly lower MSE than the unrestricted estimator. Lastly, we provide the large-sample analytical approximations of the bias, MSE and risk of the Stein-like shrinkage estimator, the unrestricted estimator, and the restricted estimator. We also evaluate the performance of our estimator in Monte Carlo simulations, and in an empirical application of forecasting output growth using macroeconomic and financial variables, and show that the Stein-like shrinkage estimator performs well relative to alternative methods.

<sup>&</sup>lt;sup>15</sup>We also calculated the equal forecast performance of the Stein-like shrinkage estimator compared to the unrestricted estimator using the Clark and West (2007) statistic. We found that the statistical significance results are similar to those obtained using the Diebold and Mariano test, i.e., the Stein-like shrinkage estimator significantly out-performs the unrestricted estimator at 5% or 10% significance levels.

## References

- Bai, J. and Ng, S. (2002), "Determining the number of factors in approximate factor models." *Econometrica* 70, 191–221.
- Bai, J. and Perron, P. (1998). "Estimating and testing linear models with multiple structural changes." *Econometrica* 66, 47-78.
- Bai, J. and Perron, P. (2003). "Computation and analysis of multiple structural change models." Journal of Applied Econometrics 18, 1-22.
- Banerjee, A. and Urga, G. (2006). "Modelling structural breaks, long memory and stock market volatility: An overview." *Journal of Econometrics* 129, 1–34.
- Bartlett, M.S. (1946). "On the theoretical specification and sampling properties of autocorrelated time-series." *Journal of the Royal Statistical Society* 8, 27-41.
- Clark, T. E. and McCracken, M. W. (2010). "Averaging forecasts from VARs with uncertain instabilities." *Journal of Applied Econometrics* 25, 5-29.
- Clark, T. E. and West, M. K. D. (2007). "Approximately normal tests for equal predictive accuracy in nested models." *Journal of Econometrics* 138, 291–311.
- Clemen, R. T. (1989). "Combining Forecasts: A review and annotated bibliography." International Journal of Forecasting 5, 559–581.
- Clements, M.P. and D.F. Hendry (1998). "Forecasting economic time series." Cambridge University Press.
- Clements, M.P. and D.F. Hendry (1999). "Forecasting non-stationary economic time series." The MIT Press.
- Clements, M.P. and Hendry, D.F. (2006). "Forecasting with breaks." In Elliott, G., C.W.J. Granger and A. Timmermann (Eds.), *Handbook of Economic Forecasting* vol. 1, 605-658. North-Holland.
- Clements, M. P. and Hendry, D.F. (2011). "The Oxford Handbook of Economic Forecasting." Oxford University Press.
- Diebold, F. X. and Mariano R. S. (1995). "Comparing predictive accuracy." Journal of Business
   & Economic Statistics 13, 253-263.
- Garcia, R. and Perron, P. (1996). "An analysis of the real interest rate under regime shifts." *Review of Economics and Statistics* 78, 111–125.
- Giacomini, R. and Rossi, B. (2009). "Detecting and predicting forecast breakdowns." Review of Economic Studies 76, 669–705.

- Gourieroux, C., Monfort, A. (1997). emphTime Series and Dynamic Models. Cambridge University Press.
- Grubb, D., Symons, J. (1987). "Bias in regressions with a lagged dependent variable." Econometric Theory 3, 371–386.
- Hansen, B.E. (2001). "The new econometrics of structural change: Dating breaks in U.S. labor productivity." *Journal of Economic Perspectives* 15, 117-128.
- Hansen, B.E. (2009). "Averaging estimators for regressions with a possible structural break." *Econometric Theory* 25, 1498-1514.
- Hansen, B.E. (2016). "Efficient shrinkage in parametric models." Journal of Econometrics 190, 115-132.
- Hansen, B.E. (2017). "Stein-like 2SLS estimator." Econometric Reviews 36, 840-852.
- Hurwicz, L., (1950). "Least squares bias in time series." In: Koopmans, T. (Ed.), Statistical Inference in Dynamic Economic Models. Wiley, New York, pp. 365–383.
- Inoue, A. and Rossi, B. (2011). "Identifying the sources of instabilities in macroeconomic fluctuations." *Review of Economics and Statistics* 164, 158–172.
- Inoue, A., Jin, L. and Rossi, B. (2017). "Rolling window selection for out-of-sample forecasting with time-varying parameters." *Journal of Econometrics* 196, 55-67.
- James, W. and Stein, C. M. (1961). "Estimation with quadratic loss." Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Vol. 1, pp. 361-380.
- Kendall, M.G. (1954). "Note on the bias in the estimation of autocorrelation." *Biometrika*, 61, 403–404.
- Kiviet, J. F. and Phillips, G.D.A. (1993). "Alternative bias approximations in regressions with a lagged dependent variable." *Econometric Theory* 9, 62–80.
- Kiviet, J. F. and Phillips, G.D.A. (1994). "Bias assessment and reduction in linear error-correction models." *Journal of Econometrics* 63, 215–243.
- Kiviet, J. F. and Phillips, G.D.A. (2005). "Moment approximation for least-squares estimators in dynamic regression models with a unit root." *Econometrics Journal* 8, 115-142.
- Kiviet, J. F. and Phillips, G.D.A. (2003). "Improved coefficient and variance estimation in stable first-order dynamic regression models." *Working paper*.
- Kiviet, J. F. and Phillips, G.D.A. (2012). "Higher-order asymptotic expansions of the least-squares estimation bias in first-order dynamic regression models." *Computational Statistics and Data Analysis* 56, 3705-3729.

- Lee, T., Parsaeian, S. and Ullah, A. (2022a). "Efficient combined estimation under structural breaks." *Advances in Econometrics*, Vol. 43, ed. by Chudik, A., Hsiao C., and A. Timmermann.
- Lee, T., Parsaeian, S. and Ullah, A. (2022b). "Optimal forecast under structural breaks." Journal of Applied Econometrics 37, 965-987.
- Lee, T., Parsaeian, S. and Ullah, A. (2022c). "Forecasting Under Structural Breaks Using Improved Weighted Estimation." Oxford Bulletin of Economics and Statistics 84, 1485-1501.
- Maasoumi, E. (1978). "A modified Stein-like estimator for the reduced form coefficients of simultaneous equations." *Econometrica* 46, 695–703.
- Marriott, F.H.C., Pope, J.A. (1954). "Bias in the estimation of autocorrelations." *Biometrika*, 61, 393–403.
- Mankiw, N.G. and Miron, J.A. (1986). "The changing behavior of the term structure of interest rates." *The Quarterly Journal of Economics* 101, 211-228.
- Mankiw, N.G., Miron, J.A., and D.N. Weil (1987). "The adjustment of expectations to a change in regime: A study of the founding of the Federal Reserve." *The American Economic Review* 77, 358-374.
- McCracken, M. W. and Ng, S. (2016). "FRED-MD: A monthly database for macroeconomic research." *Journal of Business & Economic Statistics* 34, 574-589.
- Mehrabani, A. and Ullah, A. (2020). "Improved Average Estimation in Seemingly Unrelated Regressions". *Econometrics* 8, 15.
- Nagar, A.L., (1959). "The Bias and Moment Matrix of the General K-Class Estimators of the Parameters in Simultaneous Equations." *Econometrica* 27: 575–95.
- Newbold, P., and Harvey, D. I. (2002). "Forecast combination and encompassing." In M. P. Clements, and D. F. Hendry (Eds.), A companion to economic forecasting. Oxford: Blackwells.
- Pesaran, M.H. and Pick, A. (2011). "Forecast combination across estimation windows." Journal of Business & Economic Statistics 29, 307–318.
- Pesaran, M.H., Pick, A. and Pranovich, M. (2013). "Optimal forecasts in the presence of structural breaks." Journal of Econometrics 177, 134-152.
- Pesaran, M.H. and Timmermann, A. (2007). "Selection of estimation window in the presence of breaks." Journal of Econometrics 137, 134-161.
- Pesaran, M.H. and Timmermann, A. (2002). "Market timing and return prediction under model instability." Journal of Empirical Finance 9, 495–510.

- Pesaran, M.H., Schuermann, T. and Smith, L.Vanessa, (2009). "Forecasting economic and financial variables with global VARs." *International Journal of Forecasting* 25, 642-675.
- Pesaran, M.H. and Timmermann, A. (2005). "Small sample properties of forecasts from auto-regressive models under structural breaks." *Journal of Econometrics* 129, 183-217.
- Phillips, P.C.B., Wu, Y. and Yu J. (2011). "Explosive behavior in the 1990s Nasdaq: When did exuberance escalate asset values?" *International Economic Review* 52, 201–226.
- Rossi, B. (2013). "Advances in forecasting under instability." In G. Elliott, and A. Timmermann (Eds.), *Handbook of Economic Forecasting*, Vol. 2, Part B, 1203 1324. Elsevier.
- Rossi, B. and Sekhposyan, T. (2010). "Have economic models' forecasting performance for US output growth and inflation changed over time, and when?" *International Journal of Forecasting* 26, 808-835.
- Shenton, L.R., Johnson, W.L., (1965). "Moments of a serial correlation coefficient." Journal of the Royal Statistical Society, Series B 27, 308–318.
- Smeekes, S., Wijler, E., (2020). Unit roots and cointegration. In P. Fuleky (Ed.) Macroeconomic Forecasting in the Era of Big Data (Advanced Studies in Theoretical and Applied Econometrics), vol. 52. Springer.
- Stein, Charles M., (1956). "Inadmissibility of the usual estimator for the mean of a multivariate normal distribution." Proceedings of the Third Berkeley Symposium on Math. Statist. and Probability, Vol. 1, 197–206.
- Stock, J.H. and Watson, M.W. (1996). "Evidence on structural instability in macroeconomic time series relations." Journal of Business & Economic Statistics 14, 11–30.
- Stock, J. H. and Watson, M. W. (2003). "Forecasting output and inflation: the role of asset prices." Journal of Economic Literature, 788-829.
- Stock, J. H. and Watson, M. W. (2004). "Combination Forecasts of Output Growth in a Seven-Country Data Set." *Journal of Forecasting* 23, 405-430.
- Timmerman, Allan (2006). "Forecast combinations." In G. Elliott, C. W. Granger, and A. Timmermann (Eds.), Handbook of economic forecasting, 135-196. Elsevier.
- White, J.S., (1961). "Asymptotic expansions for the mean and variance of the serial correlation coefficient." *Biometrika* 48, 95-94.

# A Appendix A

# Proof of Theorem 1:

From the proof of Lemma B.3 in the Supplementary Online Appendix, we have

$$\widehat{\alpha} - \alpha = (Z'Z)^{-1}Z'u = \xi_{-\frac{1}{2}} + \xi_{-1} + \xi_{-\frac{3}{2}} + O_p(T^{-2}),$$
(A.1)

where  $\xi_{-\frac{1}{2}}$ ,  $\xi_{-1}$ , and  $\xi_{-\frac{3}{2}}$  are defined below, and the suffixes show the order of magnitude in probability,

$$\begin{aligned} \xi_{-\frac{1}{2}} &= Q\Big[\bar{Z}'(0,I_T)\nu + \sum_{i=1}^{2}\nu' H_i\nu e_{1,i}\Big] = O_p(T^{-\frac{1}{2}}), \\ \xi_{-1} &= -Q(A+B)\Big[\bar{Z}'(0,I_T)\nu + \sum_{i=1}^{2}\nu' H_i\nu e_{1,i}\Big] = O_p(T^{-1}), \\ \xi_{-\frac{3}{2}} &= Q(A+B)^2\Big[\bar{Z}'(0,I_T)\nu + \sum_{i=1}^{2}\nu' H_i\nu e_{1,i}\Big] = O_p(T^{-\frac{3}{2}}). \end{aligned}$$

Also, let  $W_D \equiv Z'Z = \bar{W}_D + \tilde{W}_D$ , where  $\bar{W}_D = \mathbb{E}(Z'Z) = \bar{Z}'\bar{Z} + \mathbb{E}(\tilde{Z}'\tilde{Z}) = \bar{Z}'\bar{Z} + \sum_{i=1}^2 \operatorname{tr}(G'L_iG\Omega_{\nu})e_{1,i}e'_{1,i} = O_p(T)$ , and  $\tilde{W}_D = \bar{Z}'\tilde{Z} + \tilde{Z}'\bar{Z} + \left(\tilde{Z}'\tilde{Z} - \mathbb{E}(\tilde{Z}'\tilde{Z})\right) = (A+B)Q^{-1} = O_p(T^{\frac{1}{2}})$ , where A and B are given in Lemma B.2 in the Supplementary Online Appendix.

Using equation (A.1) in the inverse of equation (3.6), we have

$$\frac{1}{D(\hat{\alpha}, \tilde{\alpha})} = \left[ (\hat{\alpha} - \tilde{\alpha})' W_D(\hat{\alpha} - \tilde{\alpha}) \right]^{-1} = \left[ \hat{\alpha}' P' W_D P \hat{\alpha} \right]^{-1} \\
= \left[ \left( \alpha + \xi_{-\frac{1}{2}} + O_p(T^{-1}) \right)' \left[ \bar{P} + \tilde{P}_{-\frac{1}{2}} + O_p(T^{-1}) \right]' (\bar{W}_D + \tilde{W}_D) \left[ \bar{P} + \tilde{P}_{-\frac{1}{2}} + O_p(T^{-1}) \right] \left( \alpha + \xi_{-\frac{1}{2}} + O_p(T^{-1}) \right) \right]^{-1} \\
= \left[ \phi + \tilde{\phi}_{\frac{1}{2}} + 2\alpha' \bar{P}' \bar{W}_D \bar{P}_{\xi_{-\frac{1}{2}}} + 2\alpha' \bar{P}' \bar{W}_D \tilde{P}_{-\frac{1}{2}} \alpha + O_p(1) \right]^{-1} \\
= \frac{1}{\phi} \left[ 1 + \frac{1}{\phi} \tilde{\phi}_{\frac{1}{2}} + \frac{2}{\phi} \alpha' \bar{P}' \bar{W}_D \bar{P}_{\xi_{-\frac{1}{2}}} + \frac{2}{\phi} \alpha' \bar{P}' \bar{W}_D \tilde{P}_{-\frac{1}{2}} \alpha + O_p(T^{-1}) \right]^{-1} \\
= \frac{1}{\phi} \left[ 1 - \frac{1}{\phi} \tilde{\phi}_{\frac{1}{2}} - \frac{2}{\phi} \alpha' \bar{P}' \bar{W}_D \bar{P}_{\xi_{-\frac{1}{2}}} - \frac{2}{\phi} \alpha' \bar{P}' \bar{W}_D \tilde{P}_{-\frac{1}{2}} \alpha \right] + O_p(T^{-2}) \\
\equiv \underbrace{\frac{1}{\phi}}_{O_p(T^{-1})} - \underbrace{\frac{1}{\phi^2} D_{\frac{1}{2}}}_{O_p(T^{-2})} + O_p(T^{-2}), \quad (A.2)$$

where  $D_{\frac{1}{2}} = \tilde{\phi}_{\frac{1}{2}} + 2\left[\alpha' \bar{P}' \bar{W}_D \bar{P} \xi_{-\frac{1}{2}} + \alpha' \bar{P}' \bar{W}_D \tilde{P}_{-\frac{1}{2}} \alpha\right] = O_p(T^{\frac{1}{2}}), \ \phi = \alpha' \bar{P}' \bar{W}_D \bar{P} \alpha = O(T),$ 

 $\tilde{\phi}_{\frac{1}{2}} = \alpha' \bar{P}' \tilde{W}_D \bar{P} \alpha = O(T^{1/2})$ , and the last equality above holds by using the standard geometric expansion. The terms with order  $O_p(T^{-2})$  and smaller are dropped, because they will not enter in the calculation of the bias and MSE of the Stein-like shrinkage estimator up to the orders of interest.

Employing equations (A.1)-(A.2) in equation (3.5), we obtain

$$\begin{split} \breve{\alpha} - \alpha &= (\widehat{\alpha} - \alpha) - \tau \Big[ \frac{1}{\phi} - \frac{1}{\phi^2} D_{\frac{1}{2}} + O_p(T^{-2}) \Big] \Big[ \bar{P} + \tilde{P}_{-\frac{1}{2}} + O_p(T^{-1}) \Big] \widehat{\alpha} \\ &= \zeta_{-\frac{1}{2}} + \zeta_{-1} + \zeta_{-\frac{3}{2}} + O_p(T^{-2}), \end{split}$$
(A.3)

where  $\zeta_{-\frac{1}{2}}, \zeta_{-1}$  and  $\zeta_{-\frac{3}{2}}$  are defined below

$$\begin{split} \zeta_{-\frac{1}{2}} &= \xi_{-\frac{1}{2}} = O_p(T^{-\frac{1}{2}}), \\ \zeta_{-1} &= \xi_{-1} - \frac{\tau}{\phi} \bar{P} \alpha = O_p(T^{-1}), \\ \zeta_{-\frac{3}{2}} &= \xi_{-\frac{3}{2}} - \frac{\tau}{\phi} \bar{P} \xi_{-\frac{1}{2}} - \frac{\tau}{\phi} \tilde{P}_{-\frac{1}{2}} \alpha + \frac{\tau}{\phi^2} D_{\frac{1}{2}} \bar{P} \alpha = O_p(T^{-\frac{3}{2}}). \end{split}$$

The bias of the Stein-like shrinkage estimator using the approximations in equation (A.3) up to order  $O(T^{-1})$  is

$$\mathbb{E}(\breve{\alpha} - \alpha) = \mathbb{E}(\zeta_{-\frac{1}{2}} + \zeta_{-1}) = \mathbb{E}(\xi_{-\frac{1}{2}}) + \mathbb{E}(\xi_{-1}) - \frac{\tau}{\phi}\bar{P}\alpha = \Theta - \frac{\tau}{\phi}\bar{P}\alpha,$$
(A.4)

where the last equality holds because of (B.24) and (B.25) in Lemma B.3.

The MSE of the Stein-like shrinkage estimator up to order  ${\cal O}(T^{-2})$  is

$$\mathbb{E}\left[(\breve{\alpha} - \alpha)(\breve{\alpha} - \alpha)'\right] = \mathbb{E}(\Gamma_{-1} + \Gamma_{-\frac{3}{2}} + \Gamma_{-2}),\tag{A.5}$$

where  $\Gamma_{-1}$ ,  $\Gamma_{-\frac{3}{2}}$  and  $\Gamma_{-2}$  are

$$\begin{split} \Gamma_{-1} &= \zeta_{-\frac{1}{2}} \zeta'_{-\frac{1}{2}}, \\ \Gamma_{-\frac{3}{2}} &= \zeta_{-\frac{1}{2}} \zeta'_{-1} + \zeta_{-1} \zeta'_{-\frac{1}{2}}, \\ \Gamma_{-2} &= \zeta_{-\frac{1}{2}} \zeta'_{-\frac{3}{2}} + \zeta_{-\frac{3}{2}} \zeta'_{-\frac{1}{2}} + \zeta_{-1} \zeta'_{-1}, \end{split}$$

and we derive their expectations in the rest of the proof using Lemma B.1 and Lemma B.3.

$$\mathbb{E}(\Gamma_{-1}) = \mathbb{E}(\zeta_{-\frac{1}{2}}\zeta_{-\frac{1}{2}}') = \mathbb{E}(\xi_{-\frac{1}{2}}\xi_{-\frac{1}{2}}') = Q\Sigma,$$
(A.6)

where the last equality holds because of (B.27) in Lemma B.3.

$$\mathbb{E}(\Gamma_{-\frac{3}{2}}) = \mathbb{E}(\zeta_{-\frac{1}{2}}\zeta_{-1}') + \mathbb{E}(\zeta_{-1}\zeta_{-\frac{1}{2}}') = \mathbb{E}(\xi_{-1}\xi_{-\frac{1}{2}}') + \mathbb{E}(\xi_{-\frac{1}{2}}\xi_{-1}'),$$
(A.7)

because

$$\mathbb{E}(\zeta_{-1}\zeta'_{-\frac{1}{2}}) = \mathbb{E}\left[\left(\xi_{-1} - \frac{\tau}{\phi}\bar{P}\alpha\right)\xi'_{-\frac{1}{2}}\right] = \mathbb{E}(\xi_{-1}\xi'_{-\frac{1}{2}}) - \frac{\tau}{\phi}\bar{P}\alpha\,\mathbb{E}(\xi'_{-\frac{1}{2}}) = \mathbb{E}(\xi_{-1}\xi'_{-\frac{1}{2}}),\tag{A.8}$$

where the last equality holds by (B.24), and  $\mathbb{E}(\xi_{-1}\xi'_{-\frac{1}{2}})$  is given in (B.28). Also, we have

$$\begin{split} \mathbb{E}(\Gamma_{-2}) &= \mathbb{E}(\zeta_{-\frac{1}{2}}\zeta_{-\frac{3}{2}}') + \mathbb{E}(\zeta_{-\frac{3}{2}}\zeta_{-\frac{1}{2}}') + \mathbb{E}(\zeta_{-1}\zeta_{-1}') \\ &= \mathbb{E}(\xi_{-\frac{3}{2}}\xi_{-\frac{1}{2}}') + \mathbb{E}(\xi_{-\frac{1}{2}}\xi_{-\frac{3}{2}}') + \mathbb{E}(\xi_{-1}\xi_{-1}') + \frac{\tau^{2}}{\phi^{2}}\bar{P}\alpha\alpha'\bar{P}' - \frac{\tau}{\phi}\Big[\Theta\alpha'\bar{P}' + \bar{P}\alpha\Theta'\Big] \\ &- \frac{\tau}{\phi}\Big[\bar{P}Q\Sigma + \Sigma Q\bar{P}' + \mathbb{E}(\tilde{P}_{-\frac{1}{2}}\alpha\xi_{-\frac{1}{2}}') + \mathbb{E}(\tilde{P}_{-\frac{1}{2}}\alpha\xi_{-\frac{1}{2}}')'\Big] + \frac{2\tau}{\phi^{2}}\Big[\bar{P}\alpha\alpha'\bar{P}'\bar{W}_{D}\bar{P}Q\Sigma \\ &+ \Sigma Q\bar{P}'\bar{W}_{D}\bar{P}\alpha\alpha'\bar{P}' + \bar{P}\alpha\alpha'\bar{P}'\bar{W}_{D}\mathbb{E}(\tilde{P}_{-\frac{1}{2}}\alpha\xi_{-\frac{1}{2}}') + \mathbb{E}(\tilde{P}_{-\frac{1}{2}}\alpha\xi_{-\frac{1}{2}}')'\bar{W}_{D}\bar{P}\alpha\alpha'\bar{P}'\Big] \\ &+ \frac{\tau}{\phi^{2}}\Big[\bar{P}\alpha\alpha'\bar{P}'\mathbb{E}\left(\tilde{W}_{D}\bar{P}\alpha\xi_{-\frac{1}{2}}'\right) + \mathbb{E}\left(\xi_{-\frac{1}{2}}\alpha'\bar{P}'\tilde{W}_{D}\right)\bar{P}\alpha\alpha'\bar{P}'\Big], \end{split}$$
(A.9)

where

$$\mathbb{E}(\zeta_{-1}\zeta_{-1}') = \mathbb{E}(\xi_{-1}\xi_{-1}') - \frac{\tau}{\phi} \mathbb{E}(\xi_{-1}\alpha'\bar{P}') - \frac{\tau}{\phi} \mathbb{E}(\bar{P}\alpha\xi_{-1}') + \frac{\tau^2}{\phi^2}\bar{P}\alpha\alpha'\bar{P}'$$

$$= \mathbb{E}(\xi_{-1}\xi_{-1}') - \frac{\tau}{\phi} \Big[ \mathbb{E}(\xi_{-1})\alpha'\bar{P}' + \bar{P}\alpha \mathbb{E}(\xi_{-1}') \Big] + \frac{\tau^2}{\phi^2}\bar{P}\alpha\alpha'\bar{P}', \tag{A.10}$$

 $\mathbb{E}(\xi_{-1}) = \Theta$ , by (B.25), and  $\mathbb{E}(\xi_{-1}\xi'_{-1})$  is given in (B.29). Also,

$$\mathbb{E}(\zeta_{-\frac{3}{2}}\zeta'_{-\frac{1}{2}}) = \mathbb{E}(\xi_{-\frac{3}{2}}\xi'_{-\frac{1}{2}}) - \frac{\tau}{\phi}\bar{P}\,\mathbb{E}(\xi_{-\frac{1}{2}}\xi'_{-\frac{1}{2}}) - \frac{\tau}{\phi}\mathbb{E}(\tilde{P}_{-\frac{1}{2}}\alpha\xi'_{-\frac{1}{2}}) + \frac{\tau}{\phi^{2}}\mathbb{E}(D_{\frac{1}{2}}\bar{P}\alpha\xi'_{-\frac{1}{2}}) \\
= \mathbb{E}(\xi_{-\frac{3}{2}}\xi'_{-\frac{1}{2}}) - \frac{\tau}{\phi}\Big[\bar{P}Q\Sigma + \mathbb{E}(\tilde{P}_{-\frac{1}{2}}\alpha\xi'_{-\frac{1}{2}})\Big] + \frac{2\tau}{\phi^{2}}\bar{P}\alpha\alpha'\bar{P}'\bar{W}_{D}\Big[\bar{P}Q\Sigma + \mathbb{E}(\tilde{P}_{-\frac{1}{2}}\alpha\xi'_{-\frac{1}{2}})\Big] \\
+ \frac{\tau}{\phi^{2}}\bar{P}\alpha\alpha'\bar{P}'\,\mathbb{E}\left(\tilde{W}_{D}\bar{P}\alpha\xi'_{-\frac{1}{2}}\right), \tag{A.11}$$

where  $\mathbb{E}(\xi_{-\frac{3}{2}}\xi'_{-\frac{1}{2}})$  is given in (B.30), and the last equality above holds by using

$$\mathbb{E}(D_{\frac{1}{2}}\bar{P}\alpha\xi'_{-\frac{1}{2}}) = 2\bar{P}\alpha\alpha'\bar{P}'\bar{W}_D \mathbb{E}\left[\left(\bar{P}\xi_{-\frac{1}{2}} + \tilde{P}_{-\frac{1}{2}}\alpha\right)\xi'_{-\frac{1}{2}}\right] + \bar{P}\alpha\alpha'\bar{P}' \mathbb{E}\left(\tilde{W}_D\bar{P}\alpha\xi'_{-\frac{1}{2}}\right) \\
= 2\bar{P}\alpha\alpha'\bar{P}'\bar{W}_D\bar{P}Q\Sigma + 2\bar{P}\alpha\alpha'\bar{P}'\bar{W}_D \mathbb{E}(\tilde{P}_{-\frac{1}{2}}\alpha\xi'_{-\frac{1}{2}}) + \bar{P}\alpha\alpha'\bar{P}' \mathbb{E}\left(\tilde{W}_D\bar{P}\alpha\xi'_{-\frac{1}{2}}\right),$$
(A.12)

and using equation (B.27). Moreover, we find that

$$\mathbb{E}(\tilde{P}_{-\frac{1}{2}}\alpha\xi'_{-\frac{1}{2}}) = \mathbb{E}\left\{\left(\bar{P}-I\right)Q(A+B)Q^{-1}\bar{P}\alpha\left[\nu'(0,I_{T})'\bar{Z} + \sum_{i=1}^{2}\nu'H_{i}\nu e'_{1,i}\right]Q\right\}$$
$$= \left(\bar{P}-I\right)Q\left\{\sum_{i=1}^{2}e'_{1,i}\bar{P}\alpha\bar{Z}'L_{i}C\Omega_{u}\bar{Z}Q + \sum_{i=1}^{2}e_{1,i}\alpha'\bar{P}'\bar{Z}'L_{i}C\Omega_{u}\bar{Z}Q + 2\sum_{i=1}^{2}\sum_{j=1}^{2}\sigma_{j}^{2}e_{1,i}e'_{1,j}Q\operatorname{tr}(G\Omega_{\nu}G'L_{i}CL_{j})e'_{1,i}\bar{P}\alpha\right\} = \Psi,$$
(A.13)

and

$$\mathbb{E}(\tilde{W}_{D}\bar{P}\alpha\xi'_{-\frac{1}{2}}) = \mathbb{E}\left[\left[\sum_{l=1}^{2} \left(\bar{Z}'_{l}L_{l}G\nu e'_{1,l} + e_{1,l}\nu'G'L_{l}\bar{Z}_{l}\right) + \sum_{l=1}^{2} \left(\nu'G'L_{l}G\nu - \operatorname{tr}(G'L_{l}G\Omega_{\nu})\right)e_{1,l}e'_{1,l}\right]\bar{P}\alpha\right] \\ \left[\nu'(0,I_{T})'\bar{Z} + \sum_{i=1}^{2} \nu'H_{i}\nu e'_{1,i}\right]Q\right] \\ = \sum_{l=1}^{2} \left[\left(e'_{1,l}\bar{P}\alpha\right)\bar{Z}'L_{l}C\Omega_{u}\bar{Z}Q + e_{1,l}\alpha'\bar{P}'\bar{Z}'L_{l}C\Omega_{u}\bar{Z}Q\right] \\ + 2\sum_{i=1}^{2}\sum_{l=1}^{2} \left(e'_{1,l}\bar{P}\alpha\right)e_{1,l}e'_{1,i}Q\operatorname{tr}(C'L_{l}G\Omega_{\nu}G'L_{i}\Omega_{u}) \equiv \Phi.$$
(A.14)

By employing the results of equations (A.6),(A.7) and (A.9), in equation (A.5), we obtain the MSE of the Stein-like shrinkage estimator up to order  $O(T^{-2})$ , as

$$MSE(\breve{\alpha}) = MSE(\widehat{\alpha}) + \frac{\tau^2}{\phi^2} \bar{P} \alpha \alpha' \bar{P}' - \frac{\tau}{\phi} \Big[ \bar{P} Q \Sigma + \Sigma Q \bar{P}' + \Theta \alpha' P' + \bar{P} \alpha \Theta' + \Psi + \Psi' \Big] \\ + \frac{2\tau}{\phi^2} \Big[ \bar{P} \alpha \alpha' \bar{P}' \bar{W}_D \Big[ \bar{P} Q \Sigma + \Psi \Big] + \Big[ \Sigma Q \bar{P}' + \Psi' \Big] \bar{W}_D \bar{P} \alpha \alpha' \bar{P}' \Big] \\ + \frac{\tau}{\phi^2} \Big[ \bar{P} \alpha \alpha' \bar{P}' \Phi + \Phi' \bar{P} \alpha \alpha' \bar{P}' \Big].$$
(A.15)

Further, the risk of the Stein-like shrinkage estimator up to order  $O(T^{-1})$ , can be written as

Risk 
$$(\check{\alpha}) = \mathbb{E}\left[(\check{\alpha} - \alpha)'W(\check{\alpha} - \alpha)\right] = \operatorname{tr}\left[W\mathbb{E}\left[(\check{\alpha} - \alpha)(\check{\alpha} - \alpha)'\right]\right] = \operatorname{tr}\left[W\operatorname{MSE}\left(\check{\alpha}\right)\right]$$
  

$$= \operatorname{Risk}\left(\widehat{\alpha}\right) + \frac{\tau^{2}}{\phi^{2}}\alpha'\bar{P}'W\bar{P}\alpha - \frac{2\tau}{\phi}\left[\operatorname{tr}(W\bar{P}Q\Sigma) + \alpha'\bar{P}'W\Theta + \operatorname{tr}(W\Psi)\right]$$

$$+ \frac{4\tau}{\phi^{2}}\left[\alpha'\bar{P}'\bar{W}_{D}\bar{P}Q\Sigma W\bar{P}\alpha + \alpha'\bar{P}'\bar{W}_{D}\Psi W\bar{P}\alpha\right] + \frac{2\tau}{\phi^{2}}\alpha'\bar{P}'\Phi W\bar{P}\alpha.$$
(A.16)

This completes the proof of Theorem 1.

# Proof of Corollary 1.1:

We note that from Theorem 1, we have

Risk 
$$(\check{\alpha})$$
 = Risk  $(\widehat{\alpha}) + \frac{\phi_W}{\phi^2} \left( \tau^2 - 2\tau \frac{\phi}{\phi_W} (\mu - \frac{1}{\phi} \eta) \right) <$  Risk  $(\widehat{\alpha})$ ,

where the last inequality holds, since  $\mu > \frac{1}{\phi}\eta$ , and  $0 < \tau \le \frac{2\phi}{\phi_W} \left(\mu - \frac{1}{\phi}\eta\right)$ .

In addition, since the shrinkage parameter  $\tau$  in the risk of the Stein-like shrinkage estimator appears as a quadratic expression, there is a unique choice that minimizes the risk which is equal to  $\tau_{opt} = \frac{\phi}{\phi_W} \left(\mu - \frac{1}{\phi}\eta\right)$ .

# Proof of Corollary 1.2:

From Lemma B.2 in the Supplementary Online Appendix and the proof of Theorem 1, we have the followings

$$\hat{\phi} = \phi + \hat{\phi}_{\frac{1}{2}} + O_p(1), \text{ where } \hat{\phi}_{\frac{1}{2}} = 2\alpha' \tilde{P}'_{-\frac{1}{2}} \bar{W}_D \bar{P} \alpha + \alpha' \bar{P}' \tilde{W}_D \bar{P} \alpha + 2\xi'_{-\frac{1}{2}} \bar{P}' \bar{W}_D \bar{P} \alpha,$$

$$\hat{\phi}_{W} = \phi_{W} + \hat{\phi}_{W_{\frac{1}{2}}} + O_{p}(1), \text{ where } \hat{\phi}_{W_{\frac{1}{2}}} = 2\alpha' \tilde{P}'_{-\frac{1}{2}} W \bar{P} \alpha + 2\xi'_{-\frac{1}{2}} \bar{P}' W \bar{P} \alpha.$$

Therefore, we have

$$\hat{\tau}_{opt} = \frac{1}{\phi_W} \left[ \phi + \hat{\phi}_{\frac{1}{2}} - \frac{\phi \hat{\phi}_{W_{\frac{1}{2}}}}{\phi_W} + O_p(1) \right] \left( \hat{\mu} - \frac{1}{\phi} \left[ 1 - \frac{\hat{\phi}_{\frac{1}{2}}}{\phi} + O_p(T^{-1}) \right] \hat{\eta} \right).$$
(A.17)

Using the proof of Lemma B.2 and Lemma B.3 in the Supplementary Online Appendix, it can be easily verified that  $\mathbb{E}(\hat{\phi}_{\frac{1}{2}}) = \mathbb{E}(\hat{\phi}_{W_{\frac{1}{2}}}) = 0$ . To complete the proof, we need to show that (i)  $\hat{\mu} = \mu + \hat{\mu}_{-\frac{1}{2}} + O_p(T^{-1})$  where  $\mathbb{E}(\hat{\mu}_{-\frac{1}{2}}) = 0$  and (ii)  $\hat{\eta} = \eta + \hat{\eta}_{-\frac{1}{2}} + O_p(T^{-1})$  where  $\mathbb{E}(\hat{\eta}_{-\frac{1}{2}}) = 0$ .

Let  $\hat{\mu}_1 = \operatorname{tr}(WP(Z'Z)^{-1}\hat{\Sigma})$  be the first term of  $\hat{\mu}$ . Then, we have

$$\hat{\mu}_1 = \operatorname{tr}(W\bar{P}Q\Sigma) + \operatorname{tr}(W\tilde{P}_{-\frac{1}{2}}Q\Sigma + W\bar{P}Q\Sigma_{-\frac{1}{2}} + W\bar{P}Q_{-\frac{3}{2}}\Sigma) + O_p(T^{-1}),$$

where  $Q_{-\frac{3}{2}} = -Q(A+B)$ , and  $\Sigma_{-\frac{1}{2}} = \hat{\Sigma} - \Sigma$ . It can be easily verified that the expected value of  $\hat{\mu}_1$  up to order  $T^{-1}$  is  $\mathbb{E}(\hat{\mu}_1) = \operatorname{tr}(W\bar{P}Q\Sigma) + O(T^{-1})$ .

Let  $\hat{\mu}_2 = \hat{\alpha}' P' W \hat{\Theta}$  be the second term of  $\hat{\mu}$ . Using Theorem 3.1 in Kiviet and Phillips (2003), we have

$$\hat{\mu}_{2} = \alpha' \bar{P}' W \Theta + \alpha' \tilde{P}'_{\frac{1}{2}} W \Theta + \alpha' \bar{P}' W \Theta_{-\frac{3}{2}} + \xi'_{-\frac{1}{2}} \bar{P}' W \Theta + O_{p}(T^{-1}),$$

where  $\Theta_{-\frac{3}{2}} = \hat{\Theta} - \Theta$ . It can be easily verified that  $\mathbb{E}(\hat{\mu}_2) = \alpha' \bar{P}' W \Theta + O(T^{-1})$ .

Let  $\hat{\mu}_3 = \operatorname{tr}(W\hat{\Psi})$  be the third term of  $\hat{\mu}$ . Using the results in Appendix D of Kiviet and Phillips (2003), it readily follows that  $\mathbb{E}(\hat{\mu}_3) = \operatorname{tr}(W\Psi) + O(T^{-1})$ . This completes the proof of (*i*). Similarly, using the results in Appendix D of Kiviet and Phillips (2003), (*ii*) can be easily verified.

#### **Proof of Theorem 2:**

Note that  $z_{2,T+1}S = (0, y_T, x'_{T+1}) \equiv z_{T+1}$ , and  $z_{T+1} = \bar{z}_{T+1} + \tilde{z}_{T+1}$ , and similar to the arguments in section 2, we write  $\bar{z}_{T+1} = \mathbb{E}(z_{T+1}) = (\mathbf{0}, \bar{y}_{-1,T}, x'_T), \ \bar{y}_{-1,T} = (f_T, c_T)\bar{y}^*_{-1,T}$ , where  $f_T = \lambda_1^{T_1}\lambda_2^{T_2}, c_T = (\lambda_1^{T_1-1}\lambda_2^{T_2}, \dots, \lambda_1\lambda_2^{T_2}, \lambda_2^{T_2-1}, \dots, \lambda_2, 1), \ \bar{y}^*_{-1,T} = (\bar{y}_0, x'_1\beta_1, \dots, x'_{T_1}\beta_1, x'_{T_1+1}\beta_2, \dots, x'_T\beta_2)',$ and  $\tilde{z}_{T+1} = g\nu e'_{1,2}$ , with  $g = (df_T, c_T)$ . Let  $W_f = Tz'_{T+1}z_{T+1} = \bar{W}_f + \tilde{W}_f$ , where  $\bar{W}_f = T\mathbb{E}(z'_{T+1}z_{T+1}) = T\bar{z}'_{T+1}\bar{z}_{T+1} + T\mathbb{E}(\tilde{z}'_{T+1}\tilde{z}_{T+1}) = O_p(T)$ , and  $\tilde{W}_f = T\bar{z}'_{T+1}g\nu e'_{1,2} + Te_{1,2}\nu'g'\bar{z}_{T+1} + T(\nu'g'g\nu - \operatorname{tr}(g'g\Omega_{\nu}))e_{1,2}e'_{1,2} = O_p(T)$ . The one-step ahead MSFE of the Stein-like shrinkage estimator up to order  $O(T^{-1})$  is

$$\begin{split} \mathrm{MSFE}(\check{\alpha}_{2}) &= \mathbb{E}\left[(\check{\alpha}-\alpha)'W_{f}(\check{\alpha}-\alpha)\right] = \mathrm{tr}\left[\bar{W}_{f}\,\mathrm{MSE}\left(\hat{\alpha}\right)\right] + \mathrm{tr}\left[\mathbb{E}(\Gamma_{-1}\tilde{W}_{f}) + \mathbb{E}(\Gamma_{-\frac{3}{2}}\tilde{W}_{f}) + \mathbb{E}(\Gamma_{-2}\tilde{W}_{f})\right] \\ &= \mathrm{MSFE}(\hat{\alpha}_{2}) + \frac{\tau^{2}}{\phi^{2}}\alpha'\bar{P}'\bar{W}_{f}\bar{P}\alpha - \frac{2\tau}{\phi}\left[\mathrm{tr}(\bar{W}_{f}\bar{P}Q\Sigma) + \alpha'\bar{P}'\bar{W}_{f}\Theta + \mathrm{tr}(\bar{W}_{f}\Psi)\right] \\ &+ \frac{4\tau}{\phi^{2}}\left[\alpha'\bar{P}'\bar{W}_{D}\bar{P}Q\Sigma\bar{W}_{f}\bar{P}\alpha + \alpha'\bar{P}'\bar{W}_{D}\Psi\bar{W}_{f}\bar{P}\alpha\right] + \frac{2\tau}{\phi^{2}}\alpha'\bar{P}'\Phi\bar{W}_{f}\bar{P}\alpha \\ &- \frac{2\tau}{\phi}\,\mathrm{tr}\left[\mathbb{E}(\tilde{W}_{f}\xi_{-\frac{1}{2}})\alpha'\bar{P}'\right] - \frac{2\tau}{\phi}\,\mathrm{tr}\left[\mathbb{E}(\tilde{W}_{f}\xi_{-1})\alpha'\bar{P}'\right] - \frac{2\tau}{\phi}\,\mathrm{tr}\left[\bar{P}\,\mathbb{E}(\xi_{-\frac{1}{2}}\xi_{-\frac{1}{2}}'\tilde{W}_{f})\right] \\ &- \frac{2\tau}{\phi}\,\mathrm{tr}\left[\mathbb{E}(\tilde{P}_{-\frac{1}{2}}\alpha\xi_{-\frac{1}{2}}'\tilde{W}_{f})\right] + \frac{2\tau}{\phi^{2}}\,\mathrm{tr}\left[\mathbb{E}(D_{\frac{1}{2}}\bar{P}\alpha\xi_{-\frac{1}{2}}'\tilde{W}_{f})\right] \\ &= \mathrm{MSFE}(\hat{\alpha}_{2}) + \frac{\tau^{2}}{\phi^{2}}\alpha'\bar{P}'\bar{W}_{f}\bar{P}\alpha - \frac{2\tau}{\phi}\left[\mathrm{tr}(\bar{W}_{f}\bar{P}Q\Sigma) + \alpha'\bar{P}'\bar{W}_{f}\Theta + \mathrm{tr}(\bar{W}_{f}\Psi)\right] \\ &+ \frac{4\tau}{\phi^{2}}\left[\alpha'\bar{P}'\bar{W}_{D}\bar{P}Q\Sigma\bar{W}_{f}\bar{P}\alpha + \alpha'\bar{P}'\bar{W}_{D}\Psi\bar{W}_{f}\bar{P}\alpha\right] + \frac{2\tau}{\phi^{2}}\alpha'\bar{P}'\Phi\bar{W}_{f}\bar{P}\alpha \\ &- \frac{2\tau}{\phi}\,\mathrm{tr}\left[\Upsilon_{1}\alpha'\bar{P}'\right] - \frac{2\tau}{\phi}\,\mathrm{tr}\left[\Upsilon_{3}\alpha'\bar{P}'\right] - \frac{2\tau}{\phi}\,\mathrm{tr}\left[\bar{P}\Upsilon_{2}\right] - \frac{2\tau}{\phi}\,\mathrm{tr}\left[(\bar{P}-I_{2(k+1)})Q\Upsilon_{4}\right] \\ &+ \frac{2\tau}{\phi^{2}}\,\mathrm{tr}\left[\bar{P}\alpha\alpha'\bar{P}'\Upsilon_{4} + 2\bar{P}\alpha\alpha'\bar{P}'\bar{W}_{D}\bar{P}\Upsilon_{2} + 2\bar{P}\alpha\alpha'\bar{P}'\bar{W}_{D}(\bar{P}-I_{2(k+1)})Q\Upsilon_{4}\right]. \end{split}$$

$$\tag{A.18}$$

Now, we give the expressions for the expectations of the above equation in the rest of the proof.

$$\mathbb{E}(\tilde{W}_{f}\xi_{-\frac{1}{2}}) = T \mathbb{E}\left[\left[\bar{Z}'_{T+1}g\nu e'_{1,2} + e_{1,2}\nu'g'\bar{Z}_{T+1} + \left(\nu'g'g\nu - \operatorname{tr}(g'g\Omega_{\nu})\right)e_{1,2}e'_{1,2}\right]\right]$$

$$Q\left[\bar{Z}'(0,I_{T})\nu + \sum_{i=1}^{2}\nu'H_{i}\nu e_{1,i}\right]$$

$$= T\left[\bar{Z}'_{T+1}c_{T}\Omega_{u}\bar{Z}Qe_{1,2} + e_{1,2}\operatorname{tr}(Q\bar{Z}'\Omega_{u}c'_{T}\bar{Z}_{T+1}) + 2(e'_{1,2}Qe_{1,2})e_{1,2}\operatorname{tr}(c'_{T}g\Omega_{\nu}G'L_{2}\Omega_{u})\right]$$

$$\equiv \Upsilon_{1}.$$
(A.19)

$$\mathbb{E}(\xi_{-\frac{1}{2}}\xi'_{-\frac{1}{2}}\tilde{W}_{f}) = T \mathbb{E}\left[Q\left[\bar{Z}'(0,I_{T})\nu + \sum_{i=1}^{2}\nu'H_{i}\nu e_{1,i}\right]\left[\nu'(0,I_{T})'\bar{Z} + \sum_{i=1}^{2}\nu'H_{i}\nu e_{1,i}'\right]Q\right]$$
$$\left[\bar{Z}'_{T+1}g\nu e_{1,2}' + e_{1,2}\nu'g'\bar{Z}_{T+1} + \left(\nu'g'g\nu - \operatorname{tr}(g'g\Omega_{\nu})\right)e_{1,2}e_{1,2}'\right]$$

$$= T \left[ 2Q\bar{Z}'\Omega_{u}c'_{T}c_{T}\Omega_{u}\bar{Z}Qe_{1,2}e'_{1,2} + \sum_{i=1}^{2}Q\bar{Z}'[\Omega_{u}C'L_{i}\Omega_{u}c'_{T} + \Omega_{u}L_{i}G\Omega_{\nu}g']\bar{Z}_{T+1}Qe_{1,i}e'_{1,2} \right] \\ + (e'_{1,2}Qe_{1,2})Q\bar{Z}'[\Omega_{u}C'L_{2}\Omega_{u}c'_{T} + \Omega_{u}L_{2}G\Omega_{\nu}g']\bar{Z}_{T+1} \\ + \sum_{i=1}^{2}Qe_{1,i}e'_{1,2}\left[\operatorname{tr}(\bar{Z}Q\bar{Z}'_{T+1}g\Omega_{\nu}G'L_{i}\Omega_{u}) + \operatorname{tr}(\bar{Z}Q\bar{Z}'_{T+1}c_{T}\Omega_{u}L_{i}C\Omega_{u})\right] \\ + \sum_{i=1}^{2}Qe_{1,i}e'_{1,2}Q\bar{Z}'\left[\Omega_{u}C'L_{i}\Omega_{u}c'_{T} + \Omega_{u}L_{i}G\Omega_{\nu}g'\right]\bar{Z}_{T+1} \\ + 2\sum_{i=1}^{2}(e'_{1,2}Qe_{1,2})Qe_{1,i}e'_{1,2}\left[\operatorname{tr}(g\Omega_{\nu}G'L_{2}\Omega_{u}C'L_{i}\Omega_{u}c'_{T}) + \operatorname{tr}(c_{T}\Omega_{u}L_{2}G\Omega_{\nu}G'L_{i}\Omega_{u}c'_{T}) \\ + \operatorname{tr}(g\Omega_{\nu}G'L_{i}\Omega_{u}C'L_{2}\Omega_{u}c'_{T}) + \operatorname{tr}(g'g\Omega_{\nu}G'L_{i}\Omega_{u}L_{2}G\Omega_{\nu})\right] = \Upsilon_{2}.$$

$$(A.20)$$

$$\begin{split} \mathbb{E}(\tilde{W}_{f}\xi_{-1}) &= -T \mathbb{E}\left[\left[\bar{Z}'_{T+1}g\nu e'_{1,2} + e_{1,2}\nu'g'\bar{Z}_{T+1} + \left(\nu'g'g\nu - \operatorname{tr}(g'g\Omega_{\nu})\right)e_{1,2}e'_{1,2}\right] \\ Q(A+B)\left[\bar{Z}'(0,I_{T})\nu + \sum_{i=1}^{2}\nu'H_{i}\nu e_{1,i}\right]\right] \\ &= -T\left[\sum_{i=1}^{2}(e'_{1,i}Qe_{1,i})\bar{Z}'_{T+1}\left[g\Omega_{\nu}G'L_{i}\Omega_{u}C' + c_{T}\Omega_{u}L_{i}G\Omega_{\nu}G'\right]L_{i}\bar{Z}Qe_{1,2} \\ &+ \sum_{i=1}^{2}(e'_{1,i}Qe_{1,i})e_{1,2}\left[\operatorname{tr}(\bar{Z}_{T+1}Q\bar{Z}'L_{i}G\Omega_{\nu}G'L_{i}\Omega_{u}c'_{T}) + \operatorname{tr}(\bar{Z}_{T+1}Q\bar{Z}'L_{i}C\Omega_{u}L_{i}G\Omega_{\nu}g')\right] \\ &+ \sum_{i=1}^{2}(e'_{1,2}Qe_{1,2})\bar{Z}'_{T+1}\left[g\Omega_{\nu}G'L_{i}\Omega_{u}C' + c_{T}\Omega_{u}L_{i}G\Omega_{\nu}G'\right]L_{2}\bar{Z}Qe_{1,i} \\ &+ \sum_{i=1}^{2}\sum_{j=1}^{2}e_{1,2}e'_{1,i}Q\bar{Z}'_{T+1}\left[g\Omega_{\nu}G'L_{j}\Omega_{u}C' + c_{T}\Omega_{u}L_{j}G\Omega_{\nu}G'\right]L_{i}\bar{Z}Qe_{1,j} \\ &+ 2\sum_{i=1}^{2}e_{1,2}e'_{1,2}Q\bar{Z}'L_{i}G\Omega_{\nu}g'c_{T}\Omega_{u}\bar{Z}Qe_{1,i} + 2(e'_{1,2}Qe_{1,2})e_{1,2}\operatorname{tr}(c'_{T}g\Omega_{\nu}G'L_{2}\bar{Z}Q\bar{Z}'\Omega_{u}) \\ &+ 2(e'_{1,2}Qe_{1,2})\bar{Z}'_{T+1}g\Omega_{\nu}G'L_{2}C\Omega_{u}\bar{Z}Qe_{1,2} + 2\sum_{i=1}^{2}e_{1,2}e'_{1,i}Q\bar{Z}'\Omega_{u}C'L_{i}G\Omega_{\nu}g'\bar{Z}_{T+1}Qe_{1,i} \\ &+ 4(e'_{1,2}Qe_{1,2})\bar{Z}'_{T+1}g\Omega_{\nu}G'L_{2}G\Omega_{\nu}G'L_{2}G\Omega_{\nu}G'L_{2}\Omega_{u}) + \operatorname{tr}(g'g\Omega_{\nu}G'L_{2}C\Omega_{u}L_{2}G\Omega_{\nu})\right] \\ &= \Upsilon_{3}. \end{split}$$

$$(A.21)$$
$$\begin{split} \mathbb{E}\Big((A+B)Q^{-1}\bar{P}\alpha\xi'_{-\frac{1}{2}}\bar{W}_{f}\Big) &= T \mathbb{E}\left[(A+B)Q^{-1}\bar{P}\alpha\Big[\nu'(0,I_{T})'\bar{Z} + \sum_{i=1}^{2}\nu'H_{i}\nu e'_{1,i}\Big]Q \\ &\left[\bar{Z}'_{T+1}g\nu e'_{1,2} + e_{1,2}\nu'g'\bar{Z}_{T+1} + \left(\nu'g'g\nu - \operatorname{tr}(g'g\Omega_{\nu})\right)e_{1,2}e'_{1,2}\Big]\right] \\ &= T\left[2\sum_{i=1}^{2}(e'_{1,i}\bar{P}\alpha)\bar{Z}'L_{i}G\Omega_{\nu}g'c_{T}\Omega_{u}\bar{Z}Qe_{1,2}e'_{1,2} + 2\sum_{i=1}^{2}e_{1,i}\alpha'\bar{P}'\bar{Z}'L_{i}G\Omega_{\nu}g'c_{T}\Omega_{u}\bar{Z}Qe_{1,2}e'_{1,2} \\ &+ \sum_{i=1}^{2}\sum_{j=1}^{2}(e'_{1,i}\bar{P}\alpha)\bar{Z}'L_{i}\Big[G\Omega_{\nu}G'L_{j}\Omega_{u}c'_{T} + C\Omega_{u}L_{j}G\Omega_{\nu}g'\Big]\bar{Z}_{T+1}Qe_{1,j}e'_{1,2} \\ &+ \sum_{i=1}^{2}\sum_{j=1}^{2}(e'_{1,i}\bar{P}\alpha)(e'_{1,j}Qe_{1,2})\bar{Z}'L_{i}\Big[G\Omega_{\nu}G'L_{j}\Omega_{u}c'_{T} + C\Omega_{u}L_{j}G\Omega_{\nu}g'\Big]\bar{Z}_{T+1} \\ &+ \sum_{i=1}^{2}\sum_{j=1}^{2}e_{1,i}e'_{1,2}\Big(e'_{1,j}Q\bar{Z}'_{T+1}\Big[g\Omega_{\nu}G'L_{j}\Omega_{u}C' + c_{T}\Omega_{u}L_{j}G\Omega_{\nu}G'\Big]L_{i}\bar{Z}\bar{P}\alpha\Big) \\ &+ \sum_{i=1}^{2}(e'_{1,2}Qe_{1,2})e_{1,i}\alpha'\bar{P}'\bar{Z}'L_{i}\Big[G\Omega_{\nu}G'L_{2}\Omega_{u}c'_{T} + C\Omega_{u}L_{2}G\Omega_{\nu}g'\Big]\bar{Z}_{T+1} \\ &+ 2\sum_{i=1}^{2}(e'_{1,i}\bar{P}\alpha)e_{1,i}e'_{1,2}\operatorname{tr}(G'L_{i}C\Omega_{u}\bar{Z}Q\bar{Z}'_{T+1}g\Omega_{\nu}) + 2\sum_{i=1}^{2}(e'_{1,i}\bar{P}\alpha)e_{1,i}e'_{1,2}Q\bar{Z}'\Omega_{u}C'L_{i}G\Omega_{\nu}g'\bar{Z}_{T+1} \\ &+ 4\sum_{i=1}^{2}(e'_{1,i}\bar{P}\alpha)(e'_{1,2}Qe_{1,2})e_{1,i}e'_{1,2}\Big[\operatorname{tr}(c'_{T}g\Omega_{\nu}G'L_{i}G\Omega_{\nu}G'L_{2}\Omega_{u}) + \operatorname{tr}(g'g\Omega_{\nu}G'L_{i}C\Omega_{u}L_{2}G\Omega_{\nu})\Big]\Big] \equiv \Upsilon_{4} \\ &(A.22) \end{split}$$

This completes the proof of Theorem 2.

### Proof of Corollary 2.1:

The proof follows readily from the proof of Corollary 1.1, so it is omitted.

### **Proof of Theorem 4:**

From the proof of Lemma C.2 in the Supplementary Online Appendix, we have

$$\hat{\dot{\alpha}} - \dot{\alpha} = (\dot{Z}'\dot{Z})^{-1}\dot{Z}'u = \dot{\xi}_{-\frac{1}{2}} + \dot{\xi}_{-1} + \dot{\xi}_{-\frac{3}{2}} + O_p(T^{-2}),$$
(A.23)

where  $\dot{\xi}_{-\frac{1}{2}}$ ,  $\dot{\xi}_{-1}$ , and  $\dot{\xi}_{-\frac{3}{2}}$  are defined below, and the suffixes show the order of magnitude in probability,

$$\begin{split} \dot{\xi}_{-\frac{1}{2}} &= \dot{Q}\bar{\dot{Z}}'(0,I_T)\nu = O_p(T^{-\frac{1}{2}}), \\ \dot{\xi}_{-1} &= \dot{Q}\sum_{i=1}^2 \nu'\dot{H}_i\nu Ne_{1,i} - \dot{Q}\dot{A}\bar{\dot{Z}}'(0,I_T)\nu = O_p(T^{-1}), \\ \dot{\xi}_{-\frac{3}{2}} &= -\dot{Q}\dot{A}\sum_{i=1}^2 \nu'\dot{H}_i\nu Ne_{1,i} - \dot{Q}\dot{B}\bar{\dot{Z}}'(0,I_T)\nu + \dot{Q}\dot{A}^2\bar{\dot{Z}}'(0,I_T)\nu = O_p(T^{-\frac{3}{2}}), \end{split}$$

where  $\dot{H} = \dot{G}'L_i(0, I_T)$ . Also, let  $\dot{W}_D \equiv \dot{Z}'\dot{Z} = \ddot{W}_D + \tilde{W}_D + \tilde{Z}'\tilde{Z}$ , where  $\ddot{W}_D = \mathbb{E}(\dot{Z})'\mathbb{E}(\dot{Z}) = \dot{Z}'\dot{Z} = O_p(T)$ ,  $\ddot{W}_D = \dot{Z}'\tilde{Z} + \dot{Z}'\dot{Z} = \dot{A}\dot{Q}^{-1} = O_p(T^{\frac{1}{2}})$ , and  $\ddot{Z}'\tilde{Z} = O_p(1)$ , with  $\dot{A}$  defined in Lemma C.1 in the Supplementary Online Appendix.

Using equation (A.23) in equation (3.6), we have

$$\frac{1}{D(\hat{\alpha},\tilde{\alpha})} = \left[ (\hat{\alpha} - \tilde{\alpha})' \dot{W}_D(\hat{\alpha} - \tilde{\alpha}) \right]^{-1} = \left[ \hat{\alpha}' \dot{P}' \dot{W}_D \dot{P} \hat{\alpha} \right]^{-1} \\
= \frac{1}{\dot{\phi}} \left[ 1 - \frac{1}{\dot{\phi}} \tilde{\phi}_{\frac{1}{2}} - \frac{2}{\dot{\phi}} \dot{\alpha}' \bar{P}' \bar{W}_D \bar{P} \dot{\xi}_{-\frac{1}{2}} - \frac{2}{\dot{\phi}} \dot{\alpha}' \bar{P}' \bar{W}_D \tilde{P}_{-\frac{1}{2}} \dot{\alpha} \right] + O_p(T^{-2}) \\
\equiv \underbrace{\frac{1}{\dot{\phi}}}_{O_p(T^{-1})} - \underbrace{\frac{1}{\dot{\phi}^2} \dot{D}_{\frac{1}{2}}}_{O_p(T^{-\frac{3}{2}})} + O_p(T^{-2}), \quad (A.24)$$

where  $\dot{D}_{\frac{1}{2}} = \tilde{\phi}_{\frac{1}{2}} + 2\left[\dot{\alpha}'\bar{P}'\bar{W}_D\bar{P}\dot{\xi}_{-\frac{1}{2}} + \dot{\alpha}'\bar{P}'\bar{W}_D\bar{P}_{-\frac{1}{2}}\dot{\alpha}\right] = O_p(T^{\frac{1}{2}}), \ \dot{\phi} = \dot{\alpha}'\bar{P}'\bar{W}_D\bar{P}\dot{\alpha} = O(T),$  $\tilde{\phi}_{\frac{1}{2}} = \dot{\alpha}'\bar{P}'\bar{W}_D\bar{P}\dot{\alpha} = O(T^{1/2}),$  and the last equality above holds by using the standard geometric expansion. The terms with order  $O_p(T^{-2})$  and smaller are dropped, because they will not enter in the calculation of the bias and MSE of the Stein-like shrinkage estimator up to the orders of interest.

Employing equations (A.23)-(A.24) in equation (4.16), we obtain

$$\begin{split} \dot{\dot{\alpha}} - \dot{\alpha} &= (\hat{\dot{\alpha}} - \dot{\alpha}) - \tau \Big[ \frac{1}{\dot{\phi}} - \frac{1}{\dot{\phi}^2} \dot{D}_{\frac{1}{2}} + O_p(T^{-2}) \Big] \Big[ \dot{\vec{P}} + \dot{\vec{P}}_{-\frac{1}{2}} + O_p(T^{-1}) \Big] \hat{\dot{\alpha}} \\ &= \dot{\zeta}_{-\frac{1}{2}} + \dot{\zeta}_{-1} + \dot{\zeta}_{-\frac{3}{2}} + O_p(T^{-2}), \end{split}$$
(A.25)

where  $\dot{\zeta}_{-\frac{1}{2}}, \dot{\zeta}_{-1}$  and  $\dot{\zeta}_{-\frac{3}{2}}$  are defined below

$$\dot{\zeta}_{-\frac{1}{2}} = \dot{\xi}_{-\frac{1}{2}} = O_p(T^{-\frac{1}{2}}),$$

$$\dot{\zeta}_{-1} = \dot{\xi}_{-1} - \frac{\tau}{\dot{\phi}} \dot{\bar{P}} \dot{\alpha} = O_p(T^{-1}),$$
$$\dot{\zeta}_{-\frac{3}{2}} = \dot{\xi}_{-\frac{3}{2}} - \frac{\tau}{\dot{\phi}} \dot{\bar{P}} \dot{\xi}_{-\frac{1}{2}} - \frac{\tau}{\dot{\phi}} \dot{\bar{P}}_{-\frac{1}{2}} \dot{\alpha} + \frac{\tau}{\dot{\phi}^2} \dot{D}_{\frac{1}{2}} \dot{\bar{P}} \dot{\alpha} = O_p(T^{-\frac{3}{2}}).$$

The bias of the Stein-like shrinkage estimator using the approximations in equation (A.25) up to order  $O(T^{-1})$  is

$$\mathbb{E}(\check{\dot{\alpha}}-\dot{\alpha}) = \mathbb{E}(\dot{\zeta}_{-\frac{1}{2}}+\dot{\zeta}_{-1}) = \mathbb{E}(\dot{\xi}_{-\frac{1}{2}}) + \mathbb{E}(\dot{\xi}_{-1}) - \frac{\tau}{\dot{\phi}}\dot{P}\dot{\alpha} = \dot{\Theta} - \frac{\tau}{\dot{\phi}}\dot{P}\dot{\alpha}, \tag{A.26}$$

where the last equality holds because of (C.8) and (C.9) in Lemma C.2 of the Supplementary Online Appendix.

The MSE of the Stein-like shrinkage estimator up to order  $O(T^{-2})$  is

$$\mathbb{E}\left[(\check{\alpha} - \dot{\alpha})(\check{\alpha} - \dot{\alpha})'\right] = \mathbb{E}(\dot{\Gamma}_{-1} + \dot{\Gamma}_{-\frac{3}{2}} + \dot{\Gamma}_{-2}),\tag{A.27}$$

where  $\dot{\Gamma}_{-1}, \, \dot{\Gamma}_{-\frac{3}{2}}$  and  $\dot{\Gamma}_{-2}$  are

$$\begin{split} \dot{\Gamma}_{-1} &= \dot{\zeta}_{-\frac{1}{2}} \dot{\zeta}'_{-\frac{1}{2}}, \\ \dot{\Gamma}_{-\frac{3}{2}} &= \dot{\zeta}_{-\frac{1}{2}} \dot{\zeta}'_{-1} + \dot{\zeta}_{-1} \dot{\zeta}'_{-\frac{1}{2}}, \\ \dot{\Gamma}_{-2} &= \dot{\zeta}_{-\frac{1}{2}} \dot{\zeta}'_{-\frac{3}{2}} + \dot{\zeta}_{-\frac{3}{2}} \dot{\zeta}'_{-\frac{1}{2}} + \dot{\zeta}_{-1} \dot{\zeta}'_{-1}, \end{split}$$

and we derive their expectations in the rest of the proof using Lemma B.1 and Lemma D.2.

$$\mathbb{E}(\dot{\Gamma}_{-1}) = \mathbb{E}(\dot{\zeta}_{-\frac{1}{2}}\dot{\zeta}_{-\frac{1}{2}}') = \mathbb{E}(\dot{\xi}_{-\frac{1}{2}}\dot{\xi}_{-\frac{1}{2}}') = \dot{Q}\Sigma, \tag{A.28}$$

where the last equality holds because of (D.12) in Lemma D.2.

$$\mathbb{E}(\dot{\Gamma}_{-\frac{3}{2}}) = \mathbb{E}(\dot{\zeta}_{-\frac{1}{2}}\dot{\zeta}_{-1}') + \mathbb{E}(\dot{\zeta}_{-1}\dot{\zeta}_{-\frac{1}{2}}') = \mathbb{E}(\dot{\xi}_{-1}\dot{\xi}_{-\frac{1}{2}}') + \mathbb{E}(\dot{\xi}_{-\frac{1}{2}}\dot{\xi}_{-1}'),$$
(A.29)

because

$$\mathbb{E}(\dot{\zeta}_{-1}\dot{\zeta}'_{-\frac{1}{2}}) = \mathbb{E}\left[\left(\dot{\xi}_{-1} - \frac{\tau}{\dot{\phi}}\dot{P}\dot{\alpha}\right)\dot{\xi}'_{-\frac{1}{2}}\right] = \mathbb{E}(\dot{\xi}_{-1}\dot{\xi}'_{-\frac{1}{2}}) - \frac{\tau}{\dot{\phi}}\dot{P}\dot{\alpha}\,\mathbb{E}(\dot{\xi}'_{-\frac{1}{2}}) = \mathbb{E}(\dot{\xi}_{-1}\dot{\xi}'_{-\frac{1}{2}}),\tag{A.30}$$

where the last equality holds by (D.8), and  $\mathbb{E}(\dot{\xi}_{-1}\dot{\xi}'_{-\frac{1}{2}})$  is given in (D.12). Also, we have

$$\mathbb{E}(\dot{\Gamma}_{-2}) = \mathbb{E}(\dot{\zeta}_{-\frac{1}{2}}\dot{\zeta}'_{-\frac{3}{2}}) + \mathbb{E}(\dot{\zeta}_{-\frac{3}{2}}\dot{\zeta}'_{-\frac{1}{2}}) + \mathbb{E}(\dot{\zeta}_{-1}\dot{\zeta}'_{-1})$$

$$= \mathbb{E}(\dot{\xi}_{-\frac{3}{2}}\dot{\xi}'_{-\frac{1}{2}}) + \mathbb{E}(\dot{\xi}_{-\frac{1}{2}}\dot{\xi}'_{-\frac{3}{2}}) + \mathbb{E}(\dot{\xi}_{-1}\dot{\xi}'_{-1}) + \frac{\tau^{2}}{\dot{\phi}^{2}}\bar{P}\dot{\alpha}\dot{\alpha}'\bar{P}' - \frac{\tau}{\dot{\phi}}\Big[\dot{\Theta}\dot{\alpha}'\bar{P}' + \bar{P}\dot{\alpha}\dot{\Theta}'\Big] \\ - \frac{\tau}{\dot{\phi}}\Big[\bar{P}\dot{Q}\Sigma + \Sigma\dot{Q}\bar{P}' + \mathbb{E}(\dot{P}_{-\frac{1}{2}}\dot{\alpha}\dot{\xi}'_{-\frac{1}{2}}) + \mathbb{E}(\dot{P}_{-\frac{1}{2}}\dot{\alpha}\dot{\xi}'_{-\frac{1}{2}})'\Big] + \frac{2\tau}{\dot{\phi}^{2}}\Big[\bar{P}\dot{\alpha}\dot{\alpha}'\bar{P}'\bar{W}_{D}\bar{P}\dot{Q}\Sigma \\ + \Sigma\dot{Q}\bar{P}'\bar{W}_{D}\bar{P}\dot{\alpha}\dot{\alpha}'\bar{P}' + \bar{P}\dot{\alpha}\dot{\alpha}'\bar{P}'\bar{W}_{D}\mathbb{E}(\dot{P}_{-\frac{1}{2}}\dot{\alpha}\dot{\xi}'_{-\frac{1}{2}}) + \mathbb{E}(\dot{P}_{-\frac{1}{2}}\dot{\alpha}\dot{\xi}'_{-\frac{1}{2}})'\bar{W}_{D}\bar{P}\dot{\alpha}\dot{\alpha}'\bar{P}'\Big] \\ + \frac{\tau}{\dot{\phi}^{2}}\Big[\bar{P}\dot{\alpha}\dot{\alpha}'\bar{P}'\mathbb{E}\left(\tilde{W}_{D}\bar{P}\dot{\alpha}\dot{\xi}'_{-\frac{1}{2}}\right) + \mathbb{E}\left(\dot{\xi}_{-\frac{1}{2}}\dot{\alpha}'\bar{P}'\tilde{W}_{D}\right)\bar{P}\dot{\alpha}\dot{\alpha}'\bar{P}'\Big], \tag{A.31}$$

where

$$\mathbb{E}(\dot{\zeta}_{-1}\dot{\zeta}'_{-1}) = \mathbb{E}(\dot{\xi}_{-1}\dot{\xi}'_{-1}) - \frac{\tau}{\dot{\phi}} \mathbb{E}(\dot{\xi}_{-1}\dot{\alpha}'\bar{P}') - \frac{\tau}{\dot{\phi}} \mathbb{E}(\bar{P}\dot{\alpha}\dot{\xi}'_{-1}) + \frac{\tau^2}{\dot{\phi}^2}\bar{P}\dot{\alpha}\dot{\alpha}'\bar{P}'$$

$$= \mathbb{E}(\dot{\xi}_{-1}\dot{\xi}'_{-1}) - \frac{\tau}{\dot{\phi}} \Big[ \mathbb{E}(\dot{\xi}_{-1})\dot{\alpha}'\bar{P}' + \bar{P}\dot{\alpha}\mathbb{E}(\dot{\xi}'_{-1}) \Big] + \frac{\tau^2}{\dot{\phi}^2}\bar{P}\dot{\alpha}\dot{\alpha}'\bar{P}', \tag{A.32}$$

 $\mathbb{E}(\dot{\xi}_{-1}) = \dot{\Theta}$ , by (D.9), and  $\mathbb{E}(\dot{\xi}_{-1}\dot{\xi}'_{-1})$  is given in (D.13). Also,

$$\mathbb{E}(\dot{\zeta}_{-\frac{3}{2}}\dot{\zeta}'_{-\frac{1}{2}}) = \mathbb{E}(\dot{\xi}_{-\frac{3}{2}}\dot{\xi}'_{-\frac{1}{2}}) - \frac{\tau}{\dot{\phi}}\dot{\bar{P}}\mathbb{E}(\dot{\xi}_{-\frac{1}{2}}\dot{\xi}'_{-\frac{1}{2}}) - \frac{\tau}{\dot{\phi}}\mathbb{E}(\dot{\bar{P}}_{-\frac{1}{2}}\dot{\alpha}\dot{\xi}'_{-\frac{1}{2}}) + \frac{\tau}{\dot{\phi}^{2}}\mathbb{E}(\dot{D}_{\frac{1}{2}}\dot{\bar{P}}\dot{\alpha}\dot{\xi}'_{-\frac{1}{2}}) \\
= \mathbb{E}(\dot{\xi}_{-\frac{3}{2}}\dot{\xi}'_{-\frac{1}{2}}) - \frac{\tau}{\dot{\phi}}\Big[\dot{\bar{P}}\dot{Q}\Sigma + \mathbb{E}(\dot{\bar{P}}_{-\frac{1}{2}}\dot{\alpha}\dot{\xi}'_{-\frac{1}{2}})\Big] + \frac{2\tau}{\dot{\phi}^{2}}\dot{\bar{P}}\dot{\alpha}\dot{\alpha}'\dot{\bar{P}}'\dot{W}_{D}\Big[\dot{\bar{P}}\dot{Q}\Sigma + \mathbb{E}(\dot{\bar{P}}_{-\frac{1}{2}}\dot{\alpha}\dot{\xi}'_{-\frac{1}{2}})\Big] \\
+ \frac{\tau}{\dot{\phi}^{2}}\dot{\bar{P}}\dot{\alpha}\dot{\alpha}'\dot{\bar{P}}'\mathbb{E}\left(\dot{\bar{W}}_{D}\dot{\bar{P}}\dot{\alpha}\dot{\xi}'_{-\frac{1}{2}}\right),$$
(A.33)

where  $\mathbb{E}(\dot{\xi}_{-\frac{3}{2}}\dot{\xi}'_{-\frac{1}{2}})$  is given in (D.14), and the last equality above holds by using

$$\mathbb{E}(\dot{D}_{\frac{1}{2}}\bar{P}\dot{\alpha}\dot{\xi}'_{-\frac{1}{2}}) = 2\bar{P}\dot{\alpha}\dot{\alpha}'\bar{P}'\bar{W}_{D}\mathbb{E}\left[\left(\bar{P}\dot{\xi}_{-\frac{1}{2}} + \tilde{P}_{-\frac{1}{2}}\dot{\alpha}\right)\dot{\xi}'_{-\frac{1}{2}}\right] + \bar{P}\dot{\alpha}\dot{\alpha}'\bar{P}'\mathbb{E}\left(\tilde{W}_{D}\bar{P}\dot{\alpha}\dot{\xi}'_{-\frac{1}{2}}\right) \\
= 2\bar{P}\dot{\alpha}\dot{\alpha}'\bar{P}'\bar{W}_{D}\bar{P}\dot{Q}\Sigma + 2\bar{P}\dot{\alpha}\dot{\alpha}'\bar{P}'\bar{W}_{D}\mathbb{E}(\tilde{P}_{-\frac{1}{2}}\dot{\alpha}\dot{\xi}'_{-\frac{1}{2}}) + \bar{P}\dot{\alpha}\dot{\alpha}'\bar{P}'\mathbb{E}\left(\tilde{W}_{D}\bar{P}\dot{\alpha}\dot{\xi}'_{-\frac{1}{2}}\right),$$
(A.34)

and using equation (D.11). Moreover, we find that

$$\mathbb{E}(\tilde{\dot{P}}_{-\frac{1}{2}}\dot{\alpha}\dot{\xi}'_{-\frac{1}{2}}) = \mathbb{E}\left\{\left(\bar{\dot{P}}-I\right)\dot{Q}\dot{A}\dot{Q}^{-1}\bar{\dot{P}}\dot{\alpha}\nu'(0,I_T)'\bar{\dot{Z}}\dot{Q}\right\} \\
= \left(\bar{\dot{P}}-I\right)\dot{Q}\left\{\sum_{i=1}^{2}e'_{1,i}N\bar{\dot{P}}\dot{\alpha}\bar{\dot{Z}}'L_i\dot{C}\Omega_u\bar{\dot{Z}}\dot{Q} + \sum_{i=1}^{2}Ne_{1,i}\dot{\alpha}'\bar{\dot{P}}'\bar{\dot{Z}}'L_i\dot{C}\Omega_u\bar{\dot{Z}}\dot{Q}\right\} \equiv \dot{\Psi},$$
(A.35)

and

$$\mathbb{E}(\tilde{W}_{D}\bar{P}\dot{\alpha}\dot{\xi}'_{-\frac{1}{2}}) = \mathbb{E}\left[\left[\sum_{l=1}^{2} \left(\bar{Z}'_{l}L_{l}\dot{G}\nu e'_{1,l}N + Ne_{1,l}\nu'\dot{G}'L_{l}\bar{Z}_{l}\right)\right]\bar{P}\dot{\alpha}\left[\nu'(0,I_{T})'\bar{Z} + \sum_{i=1}^{2}\nu'\dot{H}_{i}\nu e'_{1,i}N\right]\dot{Q}\right] \\ = \sum_{l=1}^{2}\left[(e'_{1,l}N\bar{P}\dot{\alpha})\bar{Z}'L_{l}\dot{C}\Omega_{u}\bar{Z}\dot{Q} + Ne_{1,l}\dot{\alpha}'\bar{P}'\bar{Z}'L_{l}\dot{C}\Omega_{u}\bar{Z}\dot{Q}\right] \equiv \dot{\Phi}.$$

(A.36)

By employing the results of equations (A.28),(A.29) and (A.31), in equation (A.27), we obtain the MSE of the Stein-like shrinkage estimator up to order  $O(T^{-2})$ , as

$$MSE(\check{\alpha}) = MSE(\hat{\alpha}) + \frac{\tau^2}{\dot{\phi}^2} \bar{P}\dot{\alpha}\dot{\alpha}'\bar{P}' - \frac{\tau}{\dot{\phi}} \Big[\bar{P}\dot{Q}\Sigma + \Sigma\dot{Q}\bar{P}' + \dot{\Theta}\dot{\alpha}'\bar{P}' + \bar{P}\dot{\alpha}\dot{\Theta}' + \dot{\Psi} + \dot{\Psi}'\Big] + \frac{2\tau}{\dot{\phi}^2} \Big[\bar{P}\dot{\alpha}\dot{\alpha}'\bar{P}'\bar{W}_D \Big[\bar{P}\dot{Q}\Sigma + \dot{\Psi}\Big] + \Big[\Sigma\dot{Q}\bar{P}' + \dot{\Psi}'\Big]\bar{W}_D\bar{P}\dot{\alpha}\dot{\alpha}'\bar{P}'\Big] + \frac{\tau}{\dot{\phi}^2} \Big[\bar{P}\dot{\alpha}\dot{\alpha}'\bar{P}'\dot{\Phi} + \dot{\Phi}'\bar{P}\dot{\alpha}\dot{\alpha}'\bar{P}'\Big].$$
(A.37)

Further, the risk of the Stein-like shrinkage estimator up to order  $O(T^{-1})$ , can be written as

Risk 
$$(\check{\alpha}) = \mathbb{E}\left[(\check{\alpha} - \dot{\alpha})'W(\check{\alpha} - \dot{\alpha})\right] = \operatorname{tr}\left[W\mathbb{E}\left[(\check{\alpha} - \dot{\alpha})(\check{\alpha} - \dot{\alpha})'\right]\right] = \operatorname{tr}\left[W\operatorname{MSE}\left(\check{\alpha}\right)\right]$$
  

$$= \operatorname{Risk}\left(\widehat{\alpha}\right) + \frac{\tau^{2}}{\dot{\phi}^{2}}\dot{\alpha}'\bar{P}'W\bar{P}\dot{\alpha} - \frac{2\tau}{\dot{\phi}}\left[\operatorname{tr}(W\bar{P}\dot{Q}\Sigma) + \dot{\alpha}'\bar{P}'W\dot{\Theta} + \operatorname{tr}(W\dot{\Psi})\right]$$

$$+ \frac{4\tau}{\dot{\phi}^{2}}\left[\dot{\alpha}'\bar{P}'\bar{W}_{D}\bar{P}\dot{Q}\Sigma W\bar{P}\dot{\alpha} + \dot{\alpha}'\bar{P}'\bar{W}_{D}\dot{\Psi}W\bar{P}\dot{\alpha}\right] + \frac{2\tau}{\dot{\phi}^{2}}\dot{\alpha}'\bar{P}'\dot{\Phi}W\bar{P}\dot{\alpha}.$$
(A.38)

This completes the proof of Theorem 4.

### Proof of Theorem 5:

Let  $\dot{z}_{T+1} = z_{T+1}N$ , similar to the arguments in the proof of Lemma D.1, we write  $\dot{z}_{T+1} = \bar{z}_{T+1} + \tilde{z}_{T+1}$  where  $\bar{z}_{T+1} = \mathbb{E}(\dot{z}_{T+1}) = \mathbb{E}(z_{T+1})N$ , and  $\tilde{z}_{T+1} = (\dot{z}_{T+1} - \bar{Z}_{T+1}) = \dot{g}\nu e'_{1,2}N$ , with  $\dot{g}$  defined as g after replacing  $\lambda_1 = 1$  and  $\lambda_2 = 1$ . Let  $\dot{W}_f = T\dot{z}'_{T+1}\dot{z}_{T+1} = \bar{W}_f + \tilde{W}_f$ , where  $\bar{W}_f = T\mathbb{E}(\dot{z}_{T+1})'\mathbb{E}(\dot{z}_{T+1}) = T\dot{z}'_{T+1}\dot{z}_{T+1} = O_p(T)$ , and  $\tilde{W}_f = T\dot{z}'_{T+1}\dot{g}\nu e'_{1,2}N + TNe_{1,2}\nu'\dot{g}'\dot{z}_{T+1} + T(\nu'\dot{g}'\dot{g}\nu - \operatorname{tr}(\dot{g}'\dot{g}\Omega_{\nu}))Ne_{1,2}e'_{1,2}N = O_p(T^{1/2}) + O_p(1).$ 

The one-step ahead MSFE of the Stein-like shrinkage estimator up to order  $O(T^{-1})$  is

$$\begin{split} \mathrm{MSFE}(\check{\alpha}_{2}) &= \mathbb{E}\left[(\check{\alpha} - \dot{\alpha})'\dot{W}_{f}(\check{\alpha} - \dot{\alpha})\right] = \mathrm{tr}\left[\bar{W}_{f}\,\mathrm{MSE}\,(\hat{\alpha})\right] + \mathrm{tr}\left[\mathbb{E}(\dot{\Gamma}_{-1}\tilde{W}_{f}) + \mathbb{E}(\dot{\Gamma}_{-\frac{3}{2}}\tilde{W}_{f}) + \mathbb{E}(\dot{\Gamma}_{-2}\tilde{W}_{f})\right] \\ &= \mathrm{MSFE}(\hat{\alpha}_{2}) + \frac{\tau^{2}}{\dot{\phi}^{2}}\dot{\alpha}'\bar{P}'\bar{W}_{f}\bar{P}\dot{\alpha} - \frac{2\tau}{\dot{\phi}}\left[\mathrm{tr}(\bar{W}_{f}\bar{P}\dot{Q}\Sigma) + \dot{\alpha}'\bar{P}'\bar{W}_{f}\dot{\Theta} + \mathrm{tr}(\bar{W}_{f}\dot{\Psi})\right] \\ &+ \frac{4\tau}{\dot{\phi}^{2}}\left[\dot{\alpha}'\bar{P}'\bar{W}_{D}\bar{P}\dot{Q}\Sigma\bar{W}_{f}\bar{P}\dot{\alpha} + \dot{\alpha}'\bar{P}'\bar{W}_{D}\dot{\Psi}\bar{W}_{f}\bar{P}\dot{\alpha}\right] + \frac{2\tau}{\dot{\phi}^{2}}\dot{\alpha}'\bar{P}'\dot{\Phi}\bar{W}_{f}\bar{P}\dot{\alpha} \end{split}$$

$$-\frac{2\tau}{\dot{\phi}}\operatorname{tr}\left[\mathbb{E}(\tilde{W}_{f}\dot{\xi}_{-\frac{1}{2}})\dot{\alpha}'\bar{P}'\right]$$

$$=\operatorname{MSFE}(\hat{\alpha}_{2})+\frac{\tau^{2}}{\dot{\phi}^{2}}\dot{\alpha}'\bar{P}'\bar{W}_{f}\bar{P}\dot{\alpha}-\frac{2\tau}{\dot{\phi}}\left[\operatorname{tr}(\bar{W}_{f}\bar{P}\dot{Q}\Sigma)+\dot{\alpha}'\bar{P}'\bar{W}_{f}\dot{\Theta}+\operatorname{tr}(\bar{W}_{f}\dot{\Psi})\right]$$

$$+\frac{4\tau}{\dot{\phi}^{2}}\left[\dot{\alpha}'\bar{P}'\bar{W}_{D}\bar{P}\dot{Q}\Sigma\bar{W}_{f}\bar{P}\dot{\alpha}+\dot{\alpha}'\bar{P}'\bar{W}_{D}\dot{\Psi}\bar{W}_{f}\bar{P}\dot{\alpha}\right]+\frac{2\tau}{\dot{\phi}^{2}}\dot{\alpha}'\bar{P}'\dot{\Phi}\bar{W}_{f}\bar{P}\dot{\alpha}-\frac{2\tau}{\dot{\phi}}\operatorname{tr}\left[\dot{\Upsilon}\dot{\alpha}'\bar{P}'\right],$$
(A.39)

where

$$\dot{\Upsilon} \equiv \mathbb{E}(\tilde{W}_{f}\dot{\xi}_{-\frac{1}{2}}) = T \mathbb{E}\left[ \left[ \dot{\bar{Z}}'_{T+1}\dot{g}\nu e'_{1,2}N + Ne_{1,2}\nu'\dot{g}'\bar{\bar{Z}}_{T+1} \right] \dot{Q}\bar{\bar{Z}}'(0,I_{T})\nu \right] + o_{p}(1) \\ = T \left[ \dot{\bar{Z}}'_{T+1}\dot{c}_{T}\Omega_{u}\bar{\bar{Z}}\dot{Q}Ne_{1,2} + Ne_{1,2}\operatorname{tr}(\dot{Q}\bar{\bar{Z}}'\Omega_{u}\dot{c}'_{T}\bar{\bar{Z}}_{T+1}) \right] + o_{p}(1).$$
(A.40)

This completes the proof of Theorem 5.

	$b_1:$	0.5	2	0.4	4	0.0	6	0.8	8
Exp.	$\delta$	$\mathrm{RMSFE}(\check{\alpha}_2)$	$RMSFE(\tilde{\alpha}_2)$	$\overline{\text{RMSFE}(\check{\alpha}_2)}$	$\operatorname{RMSFE}(\tilde{\alpha}_2)$	$\overline{\text{RMSFE}(\check{\alpha}_2)}$	$RMSFE(\tilde{\alpha}_2)$	$\overline{\text{RMSFE}(\check{\alpha}_2)}$	$\operatorname{RMSFE}(\tilde{\alpha}_2)$
	0.000	0.794	0.439	0.796	0.445	0.797	0.447	0.788	0.443
	0.250	0.816	0.529	0.841	0.647	0.828	0.650	0.839	0.665
<i>-</i> // 1	0.500	0.887	1.045	0.902	1.776	0.876	1.818	0.855	1.222
<i></i> #1	0.750	0.966	2.225	0.966	4.248	0.943	4.704	0.907	2.814
	1.000	0.988	3.602	0.984	7.454	0.971	8.720	0.948	5.231
	0.000	0.802	0.460	0.799	0.435	0.790	0.392	0.788	0.357
	0.250	0.824	0.577	0.836	0.675	0.818	0.651	0.827	0.609
#9	0.500	0.890	1.125	0.907	1.921	0.879	1.938	0.865	1.278
#4	0.750	0.964	2.310	0.967	4.401	0.945	4.803	0.916	2.900
	1.000	0.989	3.744	0.984	7.615	0.972	8.864	0.951	5.356
	0.000	0.819	0.505	0.814	0.499	0.786	0.438	0.765	0.396
	0.250	0.847	0.628	0.851	0.753	0.827	0.740	0.817	0.663
<i>4</i> 19	0.500	0.904	1.172	0.924	2.140	0.892	2.237	0.856	1.440
#3	0.750	0.970	2.337	0.971	4.600	0.952	5.350	0.914	3.268
	1.000	0.988	3.707	0.985	7.906	0.975	9.589	0.952	5.932
	0.000	0.862	0.644	0.858	0.646	0.837	0.549	0.824	0.472
	0.250	0.879	0.802	0.878	0.974	0.852	0.893	0.837	0.684
-#-1	0.500	0.931	1.424	0.938	2.390	0.911	2.543	0.882	1.637
#4	0.750	0.975	2.531	0.971	4.818	0.954	5.596	0.928	3.556
	1.000	0.989	3.948	0.984	8.214	0.974	10.038	0.955	6.404
	0.000	0.858	0.682	0.853	0.746	0.811	0.629	0.779	0.520
	0.250	0.887	0.904	0.889	1.264	0.854	1.241	0.818	0.903
45	0.500	0.943	1.633	0.949	2.966	0.923	3.344	0.880	2.164
#0	0.750	0.979	2.783	0.977	5.679	0.962	6.835	0.932	4.493
	1.000	0.989	4.237	0.987	9.376	0.978	11.665	0.960	7.700
	0.000	0.944	1.114	0.959	1.313	0.955	1.194	0.933	0.753
	0.250	0.953	1.326	0.959	1.689	0.956	1.617	0.941	1.002
#6	0.500	0.973	1.952	0.969	3.038	0.963	3.251	0.953	1.991
	0.750	0.986	2.944	0.979	5.365	0.972	6.172	0.964	3.776
	1.000	0.992	4.312	0.986	8.670	0.980	10.370	0.972	6.342

Table 2: Simulation results with k = 5

Note: This table reports the results of the RMSFE where the benchmark model is the unrestricted estimator. The first column shows the experiment numbers which represent the break specifications based on Table 1, and the second column is the break size in the slope coefficients. In the heading of the table, RMSFE( $\check{\alpha}_2$ ) = MSFE( $\check{\alpha}_2$ )/MSFE( $\hat{\alpha}_2$ ) shows the RMSFE of the shrinkage estimator, and RMSFE( $\check{\alpha}_2$ ) = MSFE( $\check{\alpha}_2$ )/MSFE( $\hat{\alpha}_2$ ) is the RMSFE of the restricted estimator.

	$b_1:$	0.2	2	0.4	4	0.	6	0.8	3
Exp.	$\delta$	$\mathrm{RMSFE}(\breve{\alpha}_2)$	$RMSFE(\tilde{\alpha}_2)$	$\overline{\text{RMSFE}(\check{\alpha}_2)}$	$RMSFE(\tilde{\alpha}_2)$	$\overline{\text{RMSFE}(\check{\alpha}_2)}$	$RMSFE(\tilde{\alpha}_2)$	$\overline{\text{RMSFE}(\breve{\alpha}_2)}$	$RMSFE(\tilde{\alpha}_2)$
	0.000	0.964	2.279	0.964	3.096	0.928	2.529	0.849	1.267
	0.250	0.981	2.670	0.980	4.032	0.956	3.738	0.898	2.042
	0.500	0.989	3.412	0.989	5.946	0.976	6.274	0.939	3.688
₩I	0.750	0.994	4.541	0.993	8.895	0.986	10.178	0.963	6.238
	1.000	0.996	6.074	0.996	12.909	0.992	15.417	0.977	9.744
	0.000	0.749	0.339	0.745	0.287	0.732	0.220	0.717	0.197
	0.250	0.783	0.404	0.802	0.436	0.809	0.460	0.820	0.534
#8	0.500	0.877	0.888	0.913	1.626	0.898	1.758	0.870	1.232
#0	0.750	0.973	2.108	0.978	4.110	0.960	4.601	0.925	2.933
	1.000	0.992	3.476	0.990	7.304	0.980	8.664	0.958	5.396
	0.000	0.931	0.865	0.963	0.959	0.952	0.914	0.905	0.761
	0.250	0.965	1.250	0.969	1.560	0.954	1.446	0.933	1.021
-#0	0.500	0.995	3.354	0.997	6.920	0.988	7.162	0.954	2.892
#9 -	0.750	1.000	7.611	1.000	16.127	0.997	18.475	0.986	8.864
	1.000	1.001	12.939	1.001	28.858	0.998	34.304	0.993	17.941
	0.000	0.875	0.577	0.873	0.569	0.879	0.578	0.874	0.574
	0.250	0.886	0.664	0.885	0.754	0.882	0.802	0.907	0.833
#10	0.500	0.922	1.249	0.933	2.277	0.918	2.529	0.912	1.689
#10	0.750	0.974	2.660	0.982	5.566	0.970	6.706	0.960	4.385
	1.000	0.991	4.375	0.995	9.757	0.987	12.200	0.988	8.386
	0.000	0.857	0.495	0.853	0.404	0.830	0.324	0.820	0.304
	0.250	0.859	0.520	0.861	0.556	0.868	0.624	0.890	0.726
#11	0.500	0.918	1.136	0.943	2.232	0.929	2.587	0.915	1.792
#11	0.750	0.984	2.571	0.985	5.441	0.974	6.666	0.954	4.653
	1.000	0.995	4.262	0.993	9.659	0.987	12.150	0.976	8.595
	0.000	0.991	0.908	0.997	0.976	0.993	0.950	0.986	0.837
	0.250	0.995	1.287	0.999	1.702	0.998	1.669	0.992	1.168
#12	0.500	1.000	3.850	1.001	8.888	1.003	10.062	1.000	4.404
₩14	0.750	1.000	9.250	1.000	21.014	1.002	25.997	1.003	14.672
	1.000	1.000	15.912	1.000	37.792	1.000	47.839	1.002	29.821

Table 2: Simulation results with k = 5 (Cont.)

	Estimated	break dat	tes						Unknow	vn break č	lates	
$\frac{1}{2}$	Shrinkage	Unrest.	Rest.	PPP	Troff	CV	WA	Pooled	CV	WA	Pooled	AveW
	Danal A.10	80-01-2020	.03									
	1 01161 V. 13	1707-T0.00	<b>c</b> 0.0									
Η	$7.248^{*}$	7.454	7.384	7.257	7.254	7.308	7.281	7.281	7.352	7.246	7.318	7.318
9	$8.123^{**}$	8.321	9.286	8.126	8.131	8.412	8.993	8.980	8.768	8.544	8.468	8.467
12	$9.126^{**}$	9.398	9.734	9.199	9.481	10.061	9.763	9.763	9.833	9.481	9.621	9.622
	Panel B: 19	90:01-2020	:03									
<del>, -</del>	7 12/**	7 520	7 410	7 340	7 330	7 408	7 301	606 2	7 715	01970	7 365	7 365
- U	+900 L	006 4	0 7 7 7 0	0101 L	000.1		100.1	6 JO 0	LOL L	710.1	0000-1 1	7 530
с,		000.1	0.004	001.1	011.1	1.491	17.0	0.200	101.1	enu.1	800.1	800.1 077 0
17	8.190**	8.588	8.913	8.389	8.723	9.405	8.979	8.989	8.911	8.549	8.551	8.552
	Panel C: 20	00:01-2020	):03									
1	$8.953^{**}$	9.162	8.960	8.968	8.962	9.060	9.036	9.036	9.068	8.956	9.048	9.048
9	$8.287^{*}$	8.484	10.530	8.293	8.304	8.867	9.985	9.960	9.924	9.462	9.400	9.400
12	$9.380^{**}$	9.735	9.576	9.536	10.035	10.771	9.807	9.861	10.773	9.393	10.135	10.136
	Panel D: 20	05:01-2020	:03									
1	$10.090^{**}$	10.302	10.098	10.112	10.102	10.231	10.137	10.138	10.302	10.091	10.232	10.232
9	$8.627^{*}$	8.830	9.066	$8.627^{*}$	8.657	8.789	8.973	8.977	10.454	8.792	9.302	9.302
12	$10.726^{**}$	11.138	11.027	10.939	11.651	12.705	11.407	11.489	12.760	10.903	12.030	12.032
Not Not head for a line with with a a.	e: This table ling of the the the unrestriation 2013), Troff oosed by Pee to different of post-break f ad *, respect	e reports table, Shr cted estim c, CV, W/ saran and ut-of-sam ut-of-sam tively.	the MSFI inkage sh nator, Res A, and Po I Pick (20 ple foreca according	<sup>2</sup> for differ ows the re st. shows <sup>-</sup> oled are th 11) with 7 st periods. to the Die	ent forecas sults for c the results ne method $\Gamma(1 - w_{\min}$ An asteri sbold–Mar	sting meth our propos s for the r s used in ] $_{1}$ + 1 wind isk denote ish denote iano test s	iods. <i>h</i> in ed Stein-li estricted e Pesaran an dows and forecast t statistic. 7	the first co like shrinka estimator, id Timme: $w_{\min} = 0$ . hat is sign The 5% an	alumn shor age estima PPP is th rmann (20 1. Panels ufficantly b d 10% sign	ws the for- tor, Unres te one pro- 07), and / A–D repo better thar nificance l	ecast horiz st. shows t posed by F AveW is th rt the MSI 1 that obta evels are d	on. In the he results besaran et te method FE results ined from enoted by

# Supplementary Online Appendix to "Shrinkage Estimation and Forecasting in Dynamic Regression Models under Structural Instability"

Ali Mehrabani

Department of Economics University of Kansas Shahnaz Parsaeian\* Department of Economics University of Kansas

#### Aman Ullah

Department of Economics University of California, Riverside

June 10, 2024

This supplement contains three parts. Appendix B provides the Lemmas used in the proof of Theorem 1, and the detailed derivations of the bias and MSE of the unrestricted estimator up to order  $T^{-1}$  and  $T^{-2}$ , respectively. Appendix C provides the Lemmas used in the proof of Theorem 4, and the detailed derivations of the bias and MSE of the unrestricted estimator up to order  $T^{-1}$  and  $T^{-2}$ , respectively, when the coefficient of the lagged dependent variable is unity. Appendix D reports the extra simulation results of the paper for k = 3 and k = 8 in Tables D1–D2. Also, the simulation results when the error term is generated from a normalized chi-squared distribution with two degrees of freedom are reported in Tables D3–D5.

<sup>\*</sup>Address correspondence to: Shahnaz Parsaeian, <u>sh.parsaeian@ku.edu</u>.

## B Appendix B

**Lemma B.1** Let  $A_1$ ,  $A_2$ ,  $A_3$ , and  $A_4$  be symmetric  $T \times T$  matrices and  $B_1$ ,  $B_2$ , and  $B_3$ be arbitrary  $T \times T$  matrices. If the random vector u be such that  $u \sim N(0, \Omega)$ , where  $\Omega$  is  $T \times T$ , then the following results hold:

$$\mathbb{E}\prod_{i=1}^{2} (u'B_{i}u) = tr(B_{1}\Omega) tr(B_{2}\Omega) + tr(B_{1}\Omega B_{2}\Omega) + tr(B_{1}\Omega B_{2}'\Omega);$$
(B.1)

$$\mathbb{E}\left[u'A_1u - tr(A_1\Omega)\right](u'B_1u) = 2 tr(A_1\Omega B_1\Omega);$$
(B.2)

$$\mathbb{E}(uu'B_1uu') = tr(B_1\Omega)\Omega + \Omega B_1\Omega + \Omega B'_1\Omega; \tag{B.3}$$

$$\mathbb{E}\prod_{i=1}^{3} (u'B_{i}u) = tr(B_{1}\Omega) tr(B_{2}\Omega) tr(B_{3}\Omega) + tr(B_{1}\Omega) \Big[ tr(B_{2}\Omega B_{3}\Omega) + tr(B_{2}\Omega B_{3}'\Omega) \Big] + tr(B_{2}\Omega) \Big[ tr(B_{1}\Omega B_{3}\Omega) + tr(B_{1}\Omega B_{3}'\Omega) \Big] + tr(B_{3}\Omega) [tr(B_{1}\Omega B_{2}\Omega) + tr(B_{1}\Omega B_{2}'\Omega)] + tr(B_{1}\Omega B_{2}\Omega B_{3}\Omega) + tr(B_{1}\Omega B_{2}\Omega B_{3}'\Omega) + tr(B_{1}\Omega B_{2}'\Omega B_{3}\Omega) + tr(B_{1}\Omega B_{2}'\Omega B_{3}'\Omega) + tr(B_{1}\Omega B_{3}'\Omega B_{2}'\Omega) + tr(B_{1}\Omega B_{3}\Omega B_{2}'\Omega) + tr(B_{1}\Omega B_{3}'\Omega B_{2}\Omega) + tr(B_{1}\Omega B_{3}\Omega B_{2}\Omega);$$
(B.4)

$$\mathbb{E}\left[u'A_{1}u - tr(A_{1}\Omega)\right]\prod_{i=1}^{2}(u'B_{i}u) = 2 tr(B_{1}\Omega) tr(A_{1}\Omega B_{2}\Omega) + 2 tr(B_{2}\Omega) tr(A_{1}\Omega B_{1}\Omega) + 2 tr(A_{1}\Omega B_{2}\Omega B_{1}\Omega) + 2 tr(A_{1}\Omega B_{2}'\Omega B_{1}\Omega) + 2 tr(A_{1}\Omega B_{1}\Omega B_{2}\Omega) + 2 tr(A_{1}\Omega B_{1}\Omega B_{2}'\Omega);$$
(B.5)

$$\mathbb{E}\prod_{i=1}^{2} (u'B_{i}u)uu' = \Omega tr(B_{1}\Omega) tr(B_{2}\Omega) + \Omega tr(B_{1}\Omega B_{2}\Omega) + \Omega tr(B_{1}\Omega B_{2}'\Omega) + tr(B_{1}\Omega)\Omega[B_{2} + B_{2}']\Omega + tr(B_{2}\Omega)\Omega[B_{1} + B_{1}']\Omega + \Omega B_{1}\Omega B_{2}\Omega + \Omega B_{1}\Omega B_{2}'\Omega + \Omega B_{1}'\Omega B_{2}\Omega + \Omega B_{1}'\Omega B_{2}'\Omega + \Omega B_{2}\Omega B_{1}\Omega + \Omega B_{2}'\Omega B_{1}\Omega + \Omega B_{2}\Omega B_{1}'\Omega + \Omega B_{2}'\Omega B_{1}'\Omega;$$
(B.6)

$$\mathbb{E}\left[u'A_{1}u - tr(A_{1}\Omega)\right](uu'B_{1}uu) = 2 tr(A_{1}\Omega B_{1}\Omega)\Omega + 2 tr(B_{1}\Omega)\Omega A_{1}\Omega + 2\Omega A_{1}\Omega B_{1}\Omega + 2\Omega B_{1}\Omega A_{1}\Omega + 2\Omega A_{1}\Omega B_{1}'\Omega + \Omega B_{1}'\Omega A_{1}\Omega;$$
(B.7)

$$\mathbb{E}\prod_{i=1}^{2} \left[ u'A_{i}u - tr(A_{i}\Omega) \right] (u'B_{1}u) = 2 tr(B_{1}\Omega) tr(A_{1}\Omega A_{2}\Omega) + 4 tr(A_{1}\Omega A_{2}\Omega B_{1}\Omega) + 4 tr(A_{1}\Omega A_{2}\Omega B_{1}'\Omega);$$
(B.8)

$$\mathbb{E}\prod_{i=1}^{2} \left[ u'A_{i}u - tr(A_{i}\Omega) \right] uu' = 2\Omega tr(A_{1}\Omega A_{2}\Omega) + 4\Omega A_{1}\Omega A_{2}\Omega + 4\Omega A_{2}\Omega A_{1}\Omega;$$
(B.9)

$$\mathbb{E}\prod_{i=1}^{2} (u'A_{i}u)(u'B_{i}u) = \prod_{i=1}^{2} tr(A_{i}\Omega) tr(B_{i}\Omega) + 2 tr(A_{1}\Omega) \Big[ tr(A_{2}\Omega B_{1}\Omega B_{2}\Omega) + tr(A_{2}\Omega B_{1}\Omega B_{2}'\Omega) \\ + tr(A_{2}\Omega B_{2}\Omega B_{1}\Omega) + tr(A_{2}\Omega B_{2}'\Omega B_{1}\Omega) \Big] + 2 tr(A_{2}\Omega) \Big[ tr(A_{1}\Omega B_{1}\Omega B_{2}\Omega) \\ + tr(A_{1}\Omega B_{1}\Omega B_{2}'\Omega) + tr(A_{1}\Omega B_{2}\Omega B_{1}\Omega) + tr(A_{1}\Omega B_{2}'\Omega B_{1}\Omega) \Big] \\ + 4 tr(B_{1}\Omega) \Big[ tr(A_{1}\Omega A_{2}\Omega B_{2}\Omega) + tr(A_{1}\Omega A_{2}\Omega B_{2}'\Omega) \Big] \\ + 4 tr(B_{1}\Omega) \Big[ tr(A_{1}\Omega A_{2}\Omega B_{2}\Omega) + tr(A_{1}\Omega A_{2}\Omega B_{2}'\Omega) \Big] \\ + 2 tr(A_{1}\Omega A_{2}\Omega) \Big[ tr(A_{1}\Omega A_{2}\Omega B_{1}\Omega) + tr(A_{1}\Omega A_{2}\Omega B_{1}'\Omega) \Big] + 4 tr(A_{1}\Omega B_{2}\Omega) tr(A_{2}\Omega B_{1}\Omega) + tr(A_{1}\Omega B_{2}\Omega) \Big] + 4 tr(A_{1}\Omega B_{2}\Omega) tr(A_{2}\Omega B_{2}\Omega) + tr(B_{1}\Omega B_{2}'\Omega) \Big]$$

$$+ 2 tr(A_1\Omega) tr(B_1\Omega) tr(A_2\Omega B_2\Omega) + 2 tr(A_1\Omega) tr(B_2\Omega) tr(A_2\Omega B_1\Omega)$$

$$+ 2 tr(A_2\Omega) tr(B_1\Omega) tr(A_1\Omega B_2\Omega) + 2 tr(A_2\Omega) tr(B_2\Omega) tr(A_1\Omega B_1\Omega)$$

$$+ 2 tr(B_1\Omega) tr(B_2\Omega) tr(A_1\Omega A_2\Omega) + 4 tr(A_1\Omega A_2\Omega B_1\Omega B_2\Omega)$$

$$+ 4 tr(A_1\Omega A_2\Omega B_1\Omega B_2'\Omega) + 4 tr(A_1\Omega B_2'\Omega B_1\Omega A_2\Omega) + 4 tr(A_1\Omega B_2\Omega B_1\Omega A_2\Omega)$$

$$+ 4 tr(A_1\Omega A_2\Omega B_2\Omega B_1\Omega) + 4 tr(A_1\Omega A_2\Omega B_2\Omega B_1'\Omega) + 4 tr(A_1\Omega B_1'\Omega B_2\Omega A_2\Omega)$$

$$+ 4 tr(A_1\Omega B_1\Omega B_2\Omega A_2\Omega) + 4 tr(A_1\Omega B_1\Omega A_2\Omega B_2\Omega) + 4 tr(A_1\Omega B_1\Omega A_2\Omega B_2'\Omega)$$

$$+ 4 tr(A_1\Omega B_1\Omega B_2\Omega A_2\Omega) + 4 tr(A_1\Omega B_1\Omega A_2\Omega B_2\Omega) + 4 tr(A_1\Omega B_1\Omega A_2\Omega B_2'\Omega)$$

$$+ 4 tr(A_1\Omega B_1\Omega B_2\Omega A_2\Omega) + 4 tr(A_1\Omega B_1\Omega A_2\Omega B_2\Omega) + 4 tr(A_1\Omega B_1\Omega A_2\Omega B_2'\Omega)$$

$$+ 4 tr(A_1\Omega B_1\Omega B_2\Omega A_2\Omega) + 4 tr(A_1\Omega B_1\Omega A_2\Omega B_2\Omega) + 4 tr(A_1\Omega B_1\Omega A_2\Omega B_2'\Omega)$$

$$+ 4 tr(A_1\Omega B_1\Omega B_2\Omega A_2\Omega) + 4 tr(A_1\Omega B_1\Omega A_2\Omega B_2\Omega) + 4 tr(A_1\Omega B_1\Omega A_2\Omega B_2'\Omega)$$

$$+ 4 tr(A_1\Omega B_1\Omega B_2\Omega A_2\Omega) + 4 tr(A_1\Omega B_1\Omega A_2\Omega B_2\Omega) + 4 tr(A_1\Omega B_1\Omega A_2\Omega B_2'\Omega)$$

$$+ 4 tr(A_1\Omega B_1\Omega B_2\Omega A_2\Omega) + 4 tr(A_1\Omega B_1\Omega A_2\Omega B_2\Omega) + 4 tr(A_1\Omega B_1\Omega A_2\Omega B_2'\Omega)$$

$$+ 4 tr(A_1\Omega B_1\Omega B_2\Omega A_2\Omega B_1\Omega) + 4 tr(A_1\Omega B_2\Omega A_2\Omega B_1\Omega);$$

$$(B.10)$$

$$\mathbb{E}\prod_{i=1}^{2} \left[ u'A_{i}u - tr(A_{i}\Omega) \right] (u'B_{i}u) = 4 tr(B_{1}\Omega) \left[ tr(A_{1}\Omega A_{2}\Omega B_{2}\Omega) + tr(A_{1}\Omega A_{2}\Omega B_{2}'\Omega) \right] + 4 tr(B_{2}\Omega) \left[ tr(A_{1}\Omega A_{2}\Omega B_{1}\Omega) + tr(A_{1}\Omega A_{2}\Omega B_{1}'\Omega) \right] + 2 tr(A_{1}\Omega A_{2}\Omega) \left[ tr(B_{1}\Omega B_{2}\Omega) + tr(B_{1}\Omega B_{2}'\Omega) \right] + 4 tr(A_{1}\Omega B_{1}\Omega) tr(A_{2}\Omega B_{2}\Omega) + 4 tr(A_{1}\Omega B_{2}\Omega) tr(A_{2}\Omega B_{1}\Omega) + 2 tr(B_{1}\Omega) tr(B_{2}\Omega) tr(A_{1}\Omega A_{2}\Omega) + 4 tr(A_{1}\Omega A_{2}\Omega B_{1}\Omega B_{2}\Omega) + 4 tr(A_{1}\Omega A_{2}\Omega B_{1}\Omega B_{2}'\Omega) + 4 tr(A_{1}\Omega B_{2}'\Omega B_{1}\Omega A_{2}\Omega) + 4 tr(A_{1}\Omega B_{2}\Omega B_{1}\Omega A_{2}\Omega) + 4 tr(A_{1}\Omega B_{1}'\Omega B_{2}\Omega A_{2}\Omega) + 4 tr(A_{1}\Omega B_{1}\Omega B_{2}\Omega A_{2}\Omega) + 4 tr(A_{1}\Omega B_{1}\Omega A_{2}\Omega B_{2}\Omega) + 4 tr(A_{1}\Omega B_{1}\Omega A_{2}\Omega B_{2}'\Omega) + 4 tr(A_{1}\Omega B_{1}\Omega A_{2}\Omega B_{2}\Omega) + 4 tr(A_{1}\Omega B_{1}\Omega A_{2}\Omega B_{2}'\Omega) + 4 tr(A_{1}\Omega B_{1}\Omega A_{2}\Omega B_{2}\Omega) + 4 tr(A_{1}\Omega B_{1}\Omega A_{2}\Omega B_{2}'\Omega) + 4 tr(A_{1}\Omega B_{1}\Omega A_{2}\Omega B_{1}\Omega) + 4 tr(A_{1}\Omega B_{1}\Omega A_{2}\Omega B_{2}'\Omega) + 4 tr(A_{1}\Omega B_{1}\Omega A_{2}\Omega B_{1}\Omega) + 4 tr(A_{1}\Omega B_{1}\Omega A_{2}\Omega B_{2}'\Omega) + 4 tr(A_{1}\Omega B_{1}\Omega A_{2}\Omega B_{1}\Omega)$$

$$\mathbb{E}\prod_{i=1}^{3} \left[ u'A_{i}u - tr(A_{i}\Omega) \right] (u'B_{1}u) = 8 tr(B_{1}\Omega) tr(A_{1}\Omega A_{2}\Omega A_{3}\Omega) + 4 tr(A_{1}\Omega A_{2}\Omega) tr(A_{3}\Omega B_{1}) + 4 tr(A_{1}\Omega A_{3}\Omega) tr(A_{2}\Omega B_{1}) + 4 tr(A_{1}\Omega B_{1}\Omega) tr(A_{2}\Omega A_{3}\Omega) + 8 tr(A_{1}\Omega A_{2}\Omega A_{3}\Omega B_{1}\Omega) + 8 tr(A_{1}\Omega A_{2}\Omega A_{3}\Omega B_{1}'\Omega) + 8 tr(A_{1}\Omega A_{2}\Omega B_{1}\Omega A_{3}\Omega) + 8 tr(A_{1}\Omega A_{3}\Omega B_{1}\Omega A_{2}\Omega) + 8 tr(A_{1}\Omega A_{3}\Omega A_{2}\Omega B_{1}\Omega) + 8 tr(A_{1}\Omega A_{3}\Omega A_{2}\Omega B_{1}'\Omega) + 8 tr(A_{1}\Omega A_{3}\Omega A_{2}\Omega B_{1}\Omega) + 8 tr(A_{1}\Omega A_{3}\Omega A_{2}\Omega B_{1}'\Omega) ; (B.12)$$

$$\mathbb{E}\prod_{i=1}^{3} \left[ u'A_i u - tr(A_i\Omega) \right] = 8 tr(A_1\Omega A_2\Omega A_3\Omega);$$
(B.13)

$$\mathbb{E}\prod_{i=1}^{4} \left[ u'A_{i}u - tr(A_{i}\Omega) \right] = 4 tr(A_{1}\Omega A_{2}\Omega) tr(A_{3}\Omega A_{4}\Omega) + 4 tr(A_{1}\Omega A_{3}\Omega) tr(A_{2}\Omega A_{4}\Omega) + 4 tr(A_{1}\Omega A_{4}\Omega) tr(A_{2}\Omega A_{3}\Omega) + 16 tr(A_{1}\Omega A_{2}\Omega A_{3}\Omega A_{4}\Omega) + 16 tr(A_{1}\Omega A_{2}\Omega A_{4}\Omega A_{3}\Omega) + 16 tr(A_{1}\Omega A_{3}\Omega A_{2}\Omega A_{4}\Omega);$$
(B.14)

Proof: The proof of (B.1), (B.4), and (B.10) for symmetric matrices are proved in Magnus (1978) and Ullah (2004).<sup>1</sup> Our results for arbitrary matrices  $B_1$ ,  $B_2$ , and  $B_3$  follow by noting that for example  $u'B_1u = \frac{1}{2}u'(B_1 + B'_1)u = u'\bar{B}_1u$ , where  $\bar{B}_1 = \frac{1}{2}(B_1 + B'_1)$  is symmetric. Hence, the results of (B.1), (B.4), and (B.10) can be easily verified. (B.2) follows from (B.1) and using  $\mathbb{E}(u'B_1u) = tr(B_1\Omega)$ .

For (B.3), consider its  $(i, j)^{th}$  element which is  $\mathbb{E}(u_i u' B_1 u u_j)$ . Let  $e_i$  be the  $i^{th}$  column of

<sup>&</sup>lt;sup>1</sup>Some terms for the moment of fourth order products of quadratic forms were missing in Ullah (2004) which are corrected in Bao and Ullah (2010) erratum.

 $I_T$ . Then,  $u_i = e'_i u$ , and

$$\mathbb{E}(u_i u' B_1 u u_j) = \mathbb{E}(u' e_j e'_i u u' B_1 u) = e'_i \Omega e_j \operatorname{tr}(B_1 \Omega) + e'_i \Omega B_1 \Omega e_j + e'_i \Omega B'_1 \Omega e_j,$$

where the last equality above holds by (B.1). Now, (B.3) can be easily verified. In a similar way, (B.6) can be verified from (B.4).

(B.5) follows directly from (B.1) and (B.4). (B.7) follows from (B.6) and (B.3). (B.8) can be easily verified from (B.1), (B.4), and using  $\mathbb{E}(u'B_1u) = tr(B_1\Omega)$ . (B.9) follows directly from (B.6), (B.3), and noting that  $\mathbb{E}(uu') = \Omega$ . (B.11) and (B.12) are derived from (B.1), (B.4), and (B.10). (B.13) is derived from (B.9), and (B.11). Lastly, (B.14) is derived from (B.12) and (B.13).

**Lemma B.2** Under the model defined in Section 2 of the paper, it can be verified that

$$E(Z'Z) = \bar{Z}'\bar{Z} + \sum_{i=1}^{2} tr(G'L_i G\Omega_{\nu}) e_{1i} e'_{1i}, \qquad (B.15)$$

$$Z'u = \bar{Z}'(0_{T\times 1}, I_T)\nu + \sum_{i=1}^{2} \nu' H_i \nu e_{1,i} = O_p(T^{\frac{1}{2}}),$$
(B.16)

$$(Z'Z)^{-1} = Q - QA - QB + Q(A+B)^2 + o_p(T^{-2}),$$
(B.17)

$$P = \bar{P} + \tilde{P}_{-\frac{1}{2}} + O_p(T^{-1}), \tag{B.18}$$

where  $Q = [E(Z'Z)]^{-1} = O(T^{-1}), \ \bar{P} = QR'(RQR')^{-1}R, \ P = (Z'Z)^{-1}R'(R(Z'Z)^{-1}R')^{-1}R, \ H_i = G'L_i(0, I_T), \ A = (\bar{Z}'\tilde{Z} + \tilde{Z}'\bar{Z})Q = \left[\bar{Z}'\sum_{i=1}^2 L_iG\nu e'_{1,i} + \sum_{i=1}^2 e_{1,i}\nu'G'L_i\bar{Z}\right]Q = O_p(T^{-\frac{1}{2}}), \ B = \left(\tilde{Z}'\tilde{Z} - \mathbb{E}(\tilde{Z}'\tilde{Z})\right)Q = \sum_{i=1}^2 \left[\nu'G'L_iG\nu - tr(G'L_iG\Omega_{\nu})\right]e_{1,i}e'_{1,i}Q = O_p(T^{-\frac{1}{2}}), \ and \ \tilde{P}_{-\frac{1}{2}} = [\bar{P} - I_{2(k+1)}]Q(A + B)R'(RQR')^{-1}R = O_p(T^{-\frac{1}{2}}). \ Further, \ we \ note \ that \ (0, I_T)\Omega_{\nu}G' = \Omega_uC', \ tr(H_i\Omega_{\nu}) = tr(\Omega_uC'L_i) = 0, \ tr(H_i\Omega_{\nu}H_j\Omega_{\nu}) = tr(\Omega_uC'L_i\Omega_uC'L_j) = 0, \ because \ the \ diagonal \ elements \ of \ C \ are \ zeros, \ and \ tr(H_i\Omega_{\nu}H'_j\Omega_{\nu}) = \sigma_i^2 \ tr(G'_iG_i\Omega_{\nu}) \ if \ i = j, \ and \ is \ equal \ to \ 0, \ if \ i = j, \ and \ is \ equal \ to \ 0, \ if \ i = j, \ and \ is \ equal \ to \ 0, \ if \ i = j, \ and \ is \ equal \ to \ 0, \ if \ i = j, \ and \ is \ equal \ to \ 0, \ if \ i = j, \ and \ is \ equal \ to \ 0, \ if \ i = j, \ and \ is \ equal \ to \ 0, \ if \ i = j, \ and \ is \ equal \ to \ 0, \ if \ i = j, \ and \ is \ equal \ to \ 0, \ if \ i = j, \ and \ is \ equal \ to \ 0, \ if \ i = j, \ and \ is \ equal \ to \ 0, \ if \ i = j, \ and \ is \ equal \ to \ 0, \ if \ i = j, \ and \ is \ equal \ to \ 0, \ if \ i = j, \ and \ is \ equal \ to \ 0, \ if \ i = j, \ and \ is \ equal \ to \ 0, \ if \ i = j, \ and \ is \ equal \ to \ 0, \ if \ i = j, \ and \ is \ equal \ in \ det \ interval \ interval$ 

 $i \neq j$ .

Proof: It can be easily verified, and we omit it.

**Lemma B.3** Under Assumption 1, the bias of the unrestricted estimator up to order  $T^{-1}$  is

$$Bias(\widehat{\alpha}) = \mathbb{E}(\widehat{\alpha} - \alpha) = -Q \sum_{i=1}^{2} \left[ \bar{Z}' L_i C \Omega_u \bar{Z} Q e_{1,i} + e_{1,i} tr(Q \bar{Z}' L_i C \Omega_u \bar{Z}) + 2\sigma_i^2 e_{1,i} e_{1,i}' Q e_{1,i} tr(L_i G \Omega_\nu G' L_i C) \right] \equiv \Theta,$$
(B.19)

where  $\overline{P} = QR'(RQR')^{-1}R = O(1)$ ,  $Q = [E(Z'Z)]^{-1} = O(T^{-1})$ ,  $C = (C'_1, C'_2)'$ , and the MSE of the unrestricted estimator up to order  $T^{-2}$  is

$$\begin{split} MSE\left(\widehat{\alpha}\right) &= Q\Sigma \end{split} \tag{B.20} \\ &+ \sum_{i=1}^{2} \sum_{j=1}^{2} Q\bar{Z}' \Big[ \sigma_{i}^{2} L_{i} G \Omega_{\nu} G' L_{j} - \sigma_{j}^{2} L_{i} C L_{j} C \Omega_{u} - 2 \Omega_{u} C' L_{i} L_{j} C \Omega_{u} - \Omega_{u} C' L_{i} \Omega_{u} C' L_{j} \Big] \bar{Z} Q e_{1,i} e_{1,j}' Q \\ &+ \sum_{i=1}^{2} \sum_{j=1}^{2} Q e_{1,i} e_{1,j}' Q \bar{Z}' \Big[ \sigma_{j}^{2} L_{i} G \Omega_{\nu} G' L_{j} - \sigma_{j}^{2} L_{i} C L_{j} C \Omega_{u} - 2 \Omega_{u} C' L_{i} L_{j} C \Omega_{u} - \sigma_{i}^{2} \Omega_{u} C' L_{i} C' L_{j} \Big] \bar{Z} Q \\ &+ \sum_{i=1}^{2} e_{1,i}' Q e_{1,i} Q \bar{Z}' \Big[ \sigma_{i}^{2} L_{i} G \Omega_{\nu} G' L_{i} - \sigma_{i}^{2} L_{i} C L_{i} C \Omega_{u} - \sigma_{i}^{2} \Omega_{u} C' L_{i} C' L_{i} \Big] \bar{Z} Q \\ &+ \sum_{i=1}^{2} \sum_{j=1}^{2} Q e_{1,i} e_{1,j}' Q \Big[ \sigma_{i}^{2} tr(Q \bar{Z}' L_{i} G \Omega_{\nu} G' L_{j} \bar{Z}) + tr(Q \bar{Z}' L_{i} C \Omega_{u} \bar{Z}) tr(Q \bar{Z}' L_{j} C \Omega_{u} \bar{Z}) \Big] \\ &+ \sum_{i=1}^{2} \sum_{j=1}^{2} Q e_{1,i} e_{1,j}' Q \Big[ tr(Q \bar{Z}' L_{i} C \Omega_{u} \bar{Z} Q \bar{Z}' L_{j} C \Omega_{u} \bar{Z}) - 2 tr(Q \bar{Z}' L_{i} C \Omega_{u} L_{j} C \Omega_{u} \bar{Z}) \Big] \\ &+ \sum_{i=1}^{2} e_{1,i}' Q e_{1,i} Q \bar{Z}' \Big[ L_{i} C \Omega_{u} + \Omega_{u} C' L_{i} \Big] \bar{Z} Q tr(Q \bar{Z}' L_{i} C \Omega_{u} \bar{Z}) \\ &+ \sum_{i=1}^{2} e_{1,i}' Q e_{1,i} Q \bar{Z}' \Big[ L_{i} C \Omega_{u} + \Omega_{u} C' L_{i} \Big] \bar{Z} Q tr(Q \bar{Z}' L_{i} C \Omega_{u} \bar{Z}) \\ &+ \sum_{i=1}^{2} e_{1,i}' Q \bar{Z}' L_{i} C \Omega_{u} \bar{Z} Q e_{1,i} Q \bar{Z}' \Big[ L_{i} C \Omega_{u} + \Omega_{u} C' L_{i} \Big] \bar{Z} Q \\ &+ \sum_{i=1}^{2} \sum_{j=1}^{2} Q \bar{Z}' L_{i} C \Omega_{u} \bar{Z} Q e_{1,i} Q \bar{Z}' \Big[ L_{i} C \Omega_{u} + \Omega_{u} C' L_{i} \Big] \bar{Z} Q \end{aligned}$$

$$\begin{split} &+ \sum_{i=1}^{2} \sum_{j=1}^{2} \left[ QZ'L_{i}C\Omega_{u}ZQe_{1,i}e_{1,j}'QZ'\Omega_{u}C'L_{j}ZQ + QZ'\Omega_{u}C'L_{j}ZQe_{1,i}e_{1,j}'QZ'\Omega_{u}C'L_{i}ZQ \right] \\ &+ \sum_{i=1}^{2} e_{1,i}'Qe_{1,i}Q\bar{Z}L_{i}C\Omega_{u}\bar{Z}Q\bar{Z}'L_{i}C\Omega_{u}\bar{Z}Q + \sum_{i=1}^{2} e_{1,i}'Qe_{1,i}Q\bar{Z}'\Omega_{u}C'L_{i}\bar{Z}Q\bar{Z}'\Omega_{u}C'L_{i}\bar{Z}Q \\ &+ \sum_{i=1}^{2} \sum_{j=1}^{2} Qe_{1,i}e_{1,j}'Q\bar{Z}' \left[ \Omega_{u}C'L_{j} + L_{i}C\Omega_{u} \right] \bar{Z}Qtt(Q\bar{Z}'L_{i}C\Omega_{u}\bar{Z}) \\ &+ \sum_{i=1}^{2} \sum_{j=1}^{2} Q\bar{Z}' \left[ L_{i}C\Omega_{u} + \Omega_{u}C'L_{j} \right] \bar{Z}Qe_{1,i}e_{1,j}'Qtt(Q\bar{Z}'L_{j}C\Omega_{u}\bar{Z}) \\ &+ \sum_{i=1}^{2} \sum_{j=1}^{2} Qe_{1,i}e_{1,j}'Q\bar{Z}' \left[ L_{i}C\Omega_{u} + \Omega_{u}C'L_{i} \right] \bar{Z}Q\bar{Z}'L_{j}C\Omega_{u}\bar{Z}Q \\ &+ \sum_{i=1}^{2} \sum_{j=1}^{2} Qe_{1,i}e_{1,j}'Q \bar{Z}' \left[ L_{i}C\Omega_{u} + \Omega_{u}C'L_{i} \right] \bar{Z}Q\bar{Z}'L_{j}C\Omega_{u}\bar{Z}Q \\ &+ \sum_{i=1}^{2} \sum_{j=1}^{2} Qe_{1,i}e_{1,j}'Q \left[ \bar{Z}'\Omega_{u}C'L_{i}\bar{Z}Q\bar{Z}'\Omega_{u}C'L_{j}\bar{Z}Q\bar{Z}'L_{j}C\Omega_{u}\bar{Z}Q \\ &+ \sum_{i=1}^{2} \sum_{j=1}^{2} Qe_{1,i}e_{1,j}'Q \left[ \bar{Z}'\Omega_{u}C'L_{i}\bar{Z}Q\bar{Z}'L_{i}C\Omega_{u}\bar{Z}Q \\ &+ \sum_{i=1}^{2} \sum_{j=1}^{2} Q\bar{Z}' \left[ L_{i}C\Omega_{u} + \Omega_{u}C'L_{i} \right] \bar{Z}Q\bar{Z}'L_{j}C\Omega_{u}\bar{Z}Q \\ &+ \sum_{i=1}^{2} \sum_{j=1}^{2} Q\bar{Z}'Q_{u}c'L_{i}\bar{Z}Q\bar{Z}' \left[ L_{j}C\Omega_{u} + \Omega_{u}C'L_{j} \right] \bar{Z}Qe_{1,i}e_{1,j}'Q \\ &+ \sum_{i=1}^{2} \sum_{j=1}^{2} Q\bar{Z}'\Omega_{u}C'L_{i}\bar{Z}Q\bar{Z}' \left[ L_{j}C\Omega_{u} + \Omega_{u}C'L_{j} \right] \bar{Z}Qe_{1,j}e_{1,j}'Q \\ &+ 2\sum_{i=1}^{2} \sum_{j=1}^{2} e_{1,i}'Qe_{1,i}Q \\ &+ 2\sum_{i=1}^{2} \sum_{j=1}^{2} e_{1,i}'Qe_{1,i}Q\bar{Z}'\Omega_{u}C'L_{i}\bar{Z}Qe_{1,i}e_{1,i}'Q \\ &+ Q\bar{Z}'L_{i}C\Omega_{u}\bar{Z}Qe_{1,i}e_{1,j}Q \\ &+ Q\bar{Z}'L_{i}Q\Omega_{u}\bar{Z}'\Omega_{u}C'L_{i}\bar{Z}Qe_{1,i}e_{1,i}'Q \\ &+ Qe_{1,j}e_{1,j}Q\bar{Z}'\Omega_{u}C'L_{i}\bar{Z}Q \\ &+ Qe_{1,j}e_{1,j}Q\bar{Z}'L_{i}Qe_{1,i}Q\bar{Z}'L_{i}Q\Omega_{u}\bar{Z}'L_{i}Q\Omega_{u}\bar{Z}'L_{i}Q\Omega_{u}Z'L_{i}Q\Omega_{u} \\ &+ 2\sum_{i=1}^{2} \sum_{j=1}^{2} \left[ 2e_{1,i}'Q\bar{Z}'L_{i}Q\Omega_{u}\bar{Z}'L_{i}Q\Omega_{u}\bar{Z}'L_{i}Q\Omega_{u}\bar{Z}'L_{i}Q\Omega_{u}\bar{Z}'L_{i}Q\Omega_{u}\bar{Z}'L_{i}Q\Omega_{u}Z'L_{i}Q\Omega_{u}Z'L_{i}ZQU_{u}Z'L_{i}Q\Omega_{u}Z'L_{i}Q\Omega_{u}Z'L_{i}Q\Omega_{u}Z'$$

$$\begin{split} & \left[ Qe_{1,i}e_{1,j}'Q + Qe_{1,j}e_{1,i}'Q \right] + \sum_{i=1}^{2} \sum_{j=1}^{2} tr(Q\bar{Z}'L_{i}C\Omega_{u}\bar{Z}) \left\{ \left[ 2e_{1,j}'Qe_{1,j} tr(L_{j}G\Omega_{\nu}G'L_{j}C\Omega_{u}) \right. \right. \\ & \left. + 2e_{1,i}'Qe_{1,i} tr(L_{j}G\Omega_{\nu}G'L_{i}C\Omega_{u}) \right] \left[ Qe_{1,i}e_{1,j}'Q + Qe_{1,j}e_{1,i}'Q \right] \right\} \\ & - 2\sum_{i=1}^{2} \sum_{j=1}^{2} e_{1,i}'Qe_{1,i} tr(G'L_{i}G\Omega_{\nu}G'L_{j}G\Omega_{\nu}) \left[ Qe_{1,i}e_{1,j}'Q + Qe_{1,j}e_{1,i}'Q \right] \\ & + 6\sum_{i=1}^{2} e_{1,i}'Qe_{1,i} tr(G'L_{i}G\Omega_{\nu}G'L_{i}G\Omega_{\nu})Qe_{1,i}e_{1,i}'Q \\ & - 2\sum_{i=1}^{2} \sum_{j=1}^{2} \sigma_{i}^{2}\sigma_{j}^{2}e_{1,i}'Qe_{1,i} \left[ Qe_{1,i}e_{1,j}'Q + Qe_{1,j}e_{1,i}'Q \right] \left[ tr(G\Omega_{\nu}G'L_{i}CL_{i}CL_{j}) + tr(G\Omega_{\nu}G'L_{i}CL_{j}CL_{i}) \right] \\ & - 2\sum_{i=1}^{2} \sum_{j=1}^{2} \sigma_{i}^{2}\sigma_{j}^{2}e_{1,i}'Qe_{1,i} \left[ Qe_{1,i}e_{1,j}'Q + Qe_{1,j}e_{1,i}'Q \right] \left[ tr(G\Omega_{\nu}G'L_{i}CL_{i}CL_{j}) + tr(G\Omega_{\nu}G'L_{i}CL_{j}CL_{i}) \right] \\ & - 2\sum_{i=1}^{2} \sum_{j=1}^{2} \sigma_{i}^{2}\sigma_{j}^{2}e_{1,i}'Qe_{1,i} \left[ Qe_{1,i}e_{1,j}'Q + Qe_{1,j}e_{1,i}'Q \right] \left[ tr(G\Omega_{\nu}G'L_{i}CL_{i}CL_{j}) + tr(G\Omega_{\nu}G'L_{i}CL_{j}CL_{i}) \right] \\ & - 2\sum_{i=1}^{2} \sum_{j=1}^{2} \sigma_{i}^{2}\sigma_{j}^{2}e_{1,i}'Qe_{1,i} tr(G\Omega_{\nu}G'L_{i}C'L_{i}CL_{j}) \left[ Qe_{1,i}e_{1,j}'Q + Qe_{1,j}e_{1,i}'Q \right] \\ & + \sum_{i=1}^{2} \sum_{j=1}^{2} \left\{ 4e_{1,i}'Qe_{1,i}e_{1,j}'Qe_{1,j} \left[ tr(L_{i}G\Omega_{\nu}G'L_{j}C\Omega_{u}) tr(L_{j}G\Omega_{\nu}G'L_{i}C\Omega_{u}) + tr(L_{i}G\Omega_{\nu}G'L_{i}C\Omega_{u}) tr(L_{i}G\Omega_{\nu}G'L_{i}C\Omega_{u}) \right] \\ & tr(L_{j}G\Omega_{\nu}G'L_{j}C\Omega_{u}) \left] Qe_{1,i}e_{1,j}'Q + 8(e_{1,i}'Qe_{1,i})^{2} tr(L_{j}G\Omega_{\nu}G'L_{i}C\Omega_{u}) tr(L_{i}G\Omega_{\nu}G'L_{i}C\Omega_{u}) \right] \\ & \left[ Qe_{1,i}e_{1,j}'Q + Qe_{1,j}'e_{1,i}'Q \right] \right\}, \end{aligned}$$

where 
$$\Sigma = diag(\sigma_1^2 I_{k+1}, \sigma_2^2 I_{k+1}).$$

Proof: Using the results of Lemma B.2 in equation (2.2) of the paper, we have

$$\widehat{\alpha} - \alpha = (Z'Z)^{-1}Z'u = \xi_{-\frac{1}{2}} + \xi_{-1} + \xi_{-\frac{3}{2}} + O_p(T^{-2}),$$
(B.22)

where  $\xi_{-\frac{1}{2}}$ ,  $\xi_{-\frac{1}{2}}$ , and  $\xi_{-\frac{3}{2}}$  are defined below, and the suffixes show the order of magnitude in probability,

$$\begin{aligned} \xi_{-\frac{1}{2}} &= Q\Big[\bar{Z}'(0,I_T)\nu + \sum_{i=1}^{2}\nu' H_i\nu e_{1,i}\Big] = O_p(T^{-\frac{1}{2}}),\\ \xi_{-1} &= -Q(A+B)\Big[\bar{Z}'(0,I_T)\nu + \sum_{i=1}^{2}\nu' H_i\nu e_{1,i}\Big] = O_p(T^{-1}),\\ \xi_{-\frac{3}{2}} &= Q(A+B)^2\Big[\bar{Z}'(0,I_T)\nu + \sum_{i=1}^{2}\nu' H_i\nu e_{1,i}\Big] = O_p(T^{-\frac{3}{2}}).\end{aligned}$$

The bias of the unrestricted estimator using the approximations in equation (B.22) up to order  $T^{-1}$  is

$$\mathbb{E}(\widehat{\alpha} - \alpha) = \mathbb{E}(\xi_{-\frac{1}{2}}) + \mathbb{E}(\xi_{-1}) = -Q \sum_{i=1}^{2} \left[ \bar{Z}' L_i C \Omega_u \bar{Z} Q e_{1,i} + e_{1,i} tr(Q \bar{Z}' L_i C \Omega_u \bar{Z}) + 2\sigma_i^2 e_{1,i} e_{1,i}' Q e_{1,i} tr(L_i G \Omega_\nu G' L_i C) \right],$$
(B.23)

where the last equality holds because

$$\mathbb{E}(\xi_{-\frac{1}{2}}) = \sum_{i=1}^{2} Q e_{1,i} \,\mathbb{E}(\nu' H_i \nu) = \sum_{i=1}^{2} Q e_{1,i} \,tr(L_i C \Omega_u) = 0, \tag{B.24}$$

as the diagonal elements of C are zero, and by using (B.2) in Lemma B.1,  $\mathbb{E}(\xi_{-1})$  is

$$\mathbb{E}(\xi_{-1}) = -Q \mathbb{E}\left\{ (A+B) \left[ \bar{Z}'(0,I_T) \nu + \sum_{i=1}^2 \nu' H_i \nu e_{1,i} \right] \right\}$$
  
=  $-Q \sum_{i=1}^2 \left[ \bar{Z}' L_i C \Omega_u \bar{Z} Q e_{1,i} + e_{1,i} \operatorname{tr}(Q \bar{Z}' L_i C \Omega_u \bar{Z}) + 2\sigma_i^2 e_{1,i} e_{1,i}' Q e_{1,i} \operatorname{tr}(L_i G \Omega_\nu G' L_i C) \right]$   
(B.25)

The MSE up to order  $T^{-2}$  is

$$\mathbb{E}\left[\left(\widehat{\alpha}-\alpha\right)(\widehat{\alpha}-\alpha)'\right] = \mathbb{E}\left(\xi_{-\frac{1}{2}}\xi_{-\frac{1}{2}}' + \xi_{-\frac{1}{2}}\xi_{-1}' + \xi_{-1}\xi_{-\frac{1}{2}}' + \xi_{-\frac{1}{2}}\xi_{-\frac{3}{2}}' + \xi_{-\frac{3}{2}}\xi_{-\frac{1}{2}}' + \xi_{-1}\xi_{-1}'\right).$$
(B.26)

We derive the expectations of the terms in (B.26) in the rest of the proof. Using (B.1) in Lemma B.1, we obtain

$$\mathbb{E}(\xi_{-\frac{1}{2}}\xi'_{-\frac{1}{2}}) = \mathbb{E}\left\{Q\Big[\bar{Z}'(0,I_T)\nu + \sum_{i=1}^{2}\nu'H_i\nu e_{1,i}\Big]\Big[\nu'(0,I_T)'\bar{Z} + \sum_{i=1}^{2}\nu'H_i\nu e'_{1,i}\Big]Q\right\}$$
$$= Q\bar{Z}'\Omega_u\bar{Z}Q + Q\sum_{i=1}^{2}e_{1,i}e'_{1,i}\operatorname{tr}(H_i\Omega_\nu H'_i\Omega_\nu)Q$$

$$= Q\bar{Z}'\Omega_u\bar{Z}Q + Q\sum_{i=1}^2 \sigma_i^2 e_{1,i} e'_{1,i} \operatorname{tr}(G_i G'_i \Omega_\nu)Q$$
  
$$= Q\Sigma, \qquad (B.27)$$

where the last equality holds because  $Q\bar{Z}'\Omega_u\bar{Z} = \Sigma - Q\sum_{i=1}^2 \sigma_i^2 e_{1,i} e'_{1,i} tr(G_iG'_i\Omega_\nu).$ 

Using (B.1), (B.3), and (B.5) in Lemma B.1, we obtain

$$\begin{split} \mathbb{E}(\xi_{-1}\xi'_{-\frac{1}{2}}) &= -Q \mathbb{E}\left\{ (A+B) \left[ \bar{Z}'(0,I_T) \nu + \sum_{i=1}^{2} \nu' H_i \nu e_{1,i} \right] \left[ \nu'(0,I_T)' \bar{Z} + \sum_{i=1}^{2} \nu' H_i \nu e'_{1,i} \right] \right\} Q \\ &= -Q \sum_{i=1}^{2} \sum_{j=1}^{2} \sigma_j^2 \bar{Z}' \left[ L_i G \Omega_{\nu} G' L_j + L_i C L_j C \Omega_u \right] \bar{Z} Q e_{1,i} e'_{1,j} Q \\ &- Q \sum_{i=1}^{2} \sigma_i^2 e'_{1,i} Q e_{1,i} Q \bar{Z}' \left[ L_i G \Omega_{\nu} G' L_i + L_i C L_i C \Omega_u \right] \bar{Z} Q \\ &- Q \sum_{i=1}^{2} \sum_{j=1}^{2} \sigma_j^2 e_{1,i} e'_{1,j} Q \left[ tr(Q \bar{Z}' L_i G \Omega_{\nu} G' L_j \bar{Z}) + tr(Q \bar{Z}' L_i C L_j C \Omega_u \bar{Z}) \right] \\ &- Q \sum_{i=1}^{2} \sum_{j=1}^{2} \sigma_j^2 e_{1,i} e'_{1,j} Q \left[ tr(Q \bar{Z}' L_i G \Omega_{\nu} G' L_j + L_i C L_j C \Omega_u \right] \bar{Z} Q \\ &- Q \sum_{i=1}^{2} \sum_{j=1}^{2} \sigma_j^2 e_{1,i} e'_{1,j} Q \bar{Z}' \left[ L_i G \Omega_{\nu} G' L_j + L_i C L_j C \Omega_u \right] \bar{Z} Q \\ &- 2Q \sum_{i=1}^{2} e_{1,i} e'_{1,i} Q \bar{Z}' \Omega_u C' L_i C \Omega_u \bar{Z} Q - 2Q \sum_{i=1}^{2} \sum_{j=1}^{2} e_{1,i} e'_{1,i} Q e_{1,i} e'_{1,j} Q \left[ \sigma_i^2 \sigma_j^2 tr(G \Omega_{\nu} G' L_i C L_i C L_j) \right] \\ &+ \sigma_i^2 \sigma_j^2 tr(G \Omega_{\nu} G' L_i C L_j C L_i) + \sigma_i^2 tr(G' L_i G \Omega_{\nu} G' L_j G \Omega_{\nu}) + \sigma_i^2 \sigma_j^2 tr(G \Omega_{\nu} G' L_i C L_i C L_j) \right]. \end{aligned} \tag{B.28}$$

Using (B.1), (B.3), and (B.5)-(B.9) in Lemma B.1, we obtain

$$\mathbb{E}(\xi_{-1}\xi'_{-1}) = Q \mathbb{E}\left\{ (A+B) \Big[ \bar{Z}'(0,I_T)\nu + \sum_{i=1}^2 \nu' H_i \nu e_{1,i} \Big] \Big[ \nu'(0,I_T)'\bar{Z} + \sum_{i=1}^2 \nu' H_i \nu e'_{1,i} \Big] (A+B)' \right\} Q$$

$$= \sum_{i=1}^2 \sigma_i^2 e'_{1,i} Q e_{1,i} Q \bar{Z}' L_i G \Omega_\nu G' L_i \bar{Z} Q + \sum_{i=1}^2 \sum_{j=1}^2 Q \bar{Z}' L_i C \Omega_u \bar{Z} Q e_{1,j} e'_{1,i} Q \bar{Z}' \Omega_u C' L_j \bar{Z} Q$$

$$+ \sum_{i=1}^2 \sum_{j=1}^2 Q \bar{Z}' L_i C \Omega_u \bar{Z} Q e_{1,i} e'_{1,j} Q \bar{Z}' \Omega_u C' L_j \bar{Z} Q + \sum_{i=1}^2 \sum_{j=1}^2 Q \bar{Z}' L_i C \Omega_u \bar{Z} Q e_{1,i} e'_{1,j} Q tr(Q \bar{Z}' L_j C \Omega_u \bar{Z})$$

$$\begin{split} &+\sum_{i=1}^{2}\sum_{j=1}^{2}Q\bar{Z}'L_{i}C\Omega_{u}\bar{Z}Q\bar{Z}'L_{j}C\Omega_{u}\bar{Z}Qe_{1,i}e_{1,j}'Q + \sum_{i=1}^{2}\sum_{j=1}^{2}\sigma_{i}^{2}Q\bar{Z}'L_{i}G\Omega_{\nu}G'L_{j}\bar{Z}Qe_{1,i}e_{1,j}'Q \\ &+\sum_{i=1}^{2}\sum_{j=1}^{2}Qe_{1,i}e_{1,j}'Q\bar{Z}'\Omega_{u}C'L_{j}\bar{Z}Q \ tr(Q\bar{Z}'L_{i}C\Omega_{u}\bar{Z}) + \sum_{i=1}^{2}\sum_{j=1}^{2}Qe_{1,i}e_{1,j}'Q\bar{Z}'\Omega_{u}C'L_{i}\bar{Z}Q\bar{Z}'\Omega_{u}C'L_{j}\bar{Z}Q \\ &+\sum_{i=1}^{2}\sum_{j=1}^{2}\sigma_{j}^{2}Qe_{1,i}e_{1,j}'Q\bar{Z}'L_{i}G\Omega_{\nu}G'L_{j}\bar{Z}Q + \sum_{i=1}^{2}\sum_{j=1}^{2}Qe_{1,i}e_{1,j}'Q \\ \left[tr(Q\bar{Z}'L_{i}C\Omega_{u}\bar{Z}) \ tr(Q\bar{Z}'L_{j}C\Omega_{u}\bar{Z}) + \sigma_{i}^{2} \ tr(Q\bar{Z}'L_{i}G\Omega_{\nu}G'L_{j}\bar{Z}) + tr(Q\bar{Z}'L_{i}C\Omega_{u}\bar{Z}Q\bar{Z}'L_{j}C\Omega_{u}\bar{Z})\right] \\ &+2\sum_{i=1}^{2}\sum_{j=1}^{2}e_{1,j}'Qe_{1,j} \ tr(L_{j}G\Omega_{\nu}G'L_{j}C\Omega_{u}) \\ \left[Q\bar{Z}'L_{i}C\Omega_{u}\bar{Z}Qe_{1,i}e_{1,j}'Q + Qe_{1,j}e_{1,i}'Q\bar{Z}'\Omega_{u}C'L_{i}\bar{Z}Q\right] \\ &+2\sum_{i=1}^{2}\sum_{j=1}^{2}e_{1,i}'Qe_{1,i} \ tr(L_{i}G\Omega_{\nu}G'L_{j}C\Omega_{u}) \\ \left[Q\bar{Z}'L_{i}C\Omega_{u}\bar{Z}Qe_{1,j}e_{1,j}'Q + Qe_{1,j}e_{1,j}'Q\bar{Z}'\Omega_{u}C'L_{i}\bar{Z}Q\right] \\ &+2\sum_{i=1}^{2}\sum_{j=1}^{2}e_{1,j}'Qe_{1,j} \ tr(Q\bar{Z}'L_{i}C\Omega_{u}\bar{Z}) \ tr(L_{j}G\Omega_{\nu}G'L_{j}C\Omega_{u}) \\ \left[Q\bar{Z}'L_{i}C\Omega_{u}\bar{Z}Qe_{1,j}e_{1,j}'Q + Qe_{1,j}e_{1,j}'Q\bar{Z}'\Omega_{u}C'L_{i}\bar{Z}Q\right] \\ &+2\sum_{i=1}^{2}\sum_{j=1}^{2}e_{1,j}'Qe_{1,j} \ tr(Q\bar{Z}'L_{i}C\Omega_{u}\bar{Z}) \ tr(L_{j}G\Omega_{\nu}G'L_{j}C\Omega_{u}) \\ \left[Qe_{1,i}e_{1,j}'Q + Qe_{1,j}e_{1,j}'Q\bar{Z}'\Omega_{u}C'L_{i}\bar{Z}Q\right] \\ &+2\sum_{i=1}^{2}\sum_{j=1}^{2}e_{1,j}'Qe_{1,j} \ tr(Q\bar{Z}'L_{i}C\Omega_{u}\bar{Z}) \ tr(L_{i}G\Omega_{\nu}G'L_{j}C\Omega_{u}) \\ \left[Qe_{1,i}e_{1,j}'Q + Qe_{1,j}e_{1,i}'Q\right] \qquad (B.29) \\ &+2\sum_{i=1}^{2}\sigma_{i}^{2}e_{1,i}'Q\bar{Z}'L_{i}Qe_{1,i}e_{1,i}'Q \ tr(G'L_{i}G\Omega_{\nu}G'L_{i}G\Omega_{\nu}) + 4\sum_{i=1}^{2}\sum_{j=1}^{2}e_{1,i}'Qe_{1,i}e_{1,j}'Qe_{1,j}Qe_{1,i}e_{1,j}'Q \\ \\ \left[tr(L_{i}G\Omega_{\nu}G'L_{j}C\Omega_{u}) \ tr(L_{j}G\Omega_{\nu}G'L_{i}C\Omega_{u}) + tr(L_{i}G\Omega_{\nu}G'L_{i}C\Omega_{u}) \ tr(L_{j}G\Omega_{\nu}G'L_{j}C\Omega_{u}) \\ \right] + o(T^{-2}). \end{aligned}$$

Also, Using (B.3), and (B.5)–(B.9) in Lemma B.1, we obtain

$$\begin{split} \mathbb{E}(\xi_{-\frac{3}{2}}\xi'_{-\frac{1}{2}}) &= Q \,\mathbb{E}\left\{ (A+B)^2 \Big[ \bar{Z}'(0,I_T)\nu + \sum_{i=1}^2 \nu' H_i \nu e_{1,i} \Big] \Big[ \nu'(0,I_T)'\bar{Z} + \sum_{i=1}^2 \nu' H_i \nu e'_{1,i} \Big] \right\} Q \\ &= \sum_{i=1}^2 \sum_{j=1}^2 e'_{1,i} Q \bar{Z}' L_j C \Omega_u \bar{Z} Q e_{1,j} Q \bar{Z}' L_i C \Omega_u \bar{Z} Q + \sum_{i=1}^2 \sum_{j=1}^2 \sigma_j^2 Q \bar{Z}' L_i G \Omega_\nu G' L_j \bar{Z} Q e_{1,i} e'_{1,j} Q \\ &+ \sum_{i=1}^2 \sum_{j=1}^2 Q \bar{Z}' L_i C \Omega_u \bar{Z} Q e_{1,j} e'_{1,i} Q \bar{Z}' L_j C \Omega_u \bar{Z} Q + \sum_{i=1}^2 e'_{1,i} Q e_{1,i} Q \bar{Z}' L_i C \Omega_u \bar{Z} Q \operatorname{tr}(Q \bar{Z}' L_i C \Omega_u \bar{Z}) \\ &+ \sum_{i=1}^2 \sigma_i^2 e'_{1,i} Q e_{1,i} Q \bar{Z}' L_i G \Omega_\nu G' L_i \bar{Z} Q + \sum_{i=1}^2 e'_{1,i} Q e_{1,i} Q \bar{Z}' L_i C \Omega_u \bar{Z} Q \\ &+ \sum_{i=1}^2 \sigma_i^2 e'_{1,i} Q e_{1,i} Q \bar{Z}' L_i G \Omega_\nu G' L_i \bar{Z} Q + \sum_{i=1}^2 e'_{1,i} Q e_{1,i} Q \bar{Z}' L_i C \Omega_u \bar{Z} Q \\ &+ \sum_{i=1}^2 \sum_{j=1}^2 \sigma_j^2 Q e_{1,i} e'_{1,j} Q \operatorname{tr}(Q \bar{Z}' L_j G \Omega_\nu G' L_i \bar{Z}) + \sum_{i=1}^2 \sum_{j=1}^2 Q e_{1,i} e'_{1,j} Q \bar{Z}' \Omega_u C' L_i \bar{Z} Q \bar{Z}' L_j C \Omega_u \bar{Z} Q \end{split}$$

$$+ \sum_{i=1}^{2} \sum_{j=1}^{2} Qe_{1,i}e'_{1,j}Q\bar{Z}'\Omega_{u}C'L_{j}\bar{Z}Q\bar{Z}'L_{i}C\Omega_{u}\bar{Z}Q + \sum_{i=1}^{2} \sum_{j=1}^{2} Qe_{1,i}e'_{1,j}Q\bar{Z}'L_{i}C\Omega_{\nu}\bar{Z}Q tr(Q\bar{Z}'L_{j}C\Omega_{u}\bar{Z})$$

$$+ \sum_{i=1}^{2} \sum_{j=1}^{2} \sigma_{j}^{2}Qe_{1,i}e'_{1,j}Q\bar{Z}'L_{i}G\Omega_{\nu}G'L_{j}\bar{Z}Q + \sum_{i=1}^{2} \sum_{j=1}^{2} Qe_{1,i}e'_{1,j}Q\bar{Z}'L_{i}C\Omega_{u}\bar{Z}Q\bar{Z}'L_{j}C\Omega_{u}\bar{Z}Q$$

$$+ 2\sum_{i=1}^{2} \sigma_{i}^{2}e'_{1,i}Qe_{1,i}Qe_{1,i}e'_{1,i}Q tr(G'L_{i}G\Omega_{\nu}G'L_{i}G\Omega_{\nu})$$

$$+ 8\sum_{i=1}^{2} \sum_{j=1}^{2} (e'_{1,i}Qe_{1,i})^{2}Qe_{1,i}e'_{1,j}Q tr(L_{i}G\Omega_{\nu}G'L_{i}C\Omega_{u}) tr(L_{j}G\Omega_{\nu}G'L_{i}C\Omega_{u})$$

$$+ 2\sum_{i=1}^{2} \sum_{j=1}^{2} e'_{1,i}Qe_{1,i}Q\bar{Z}'L_{i}C\Omega_{u}\bar{Z}Qe_{1,i}e'_{1,j}Q tr(L_{j}G\Omega_{\nu}G'L_{i}C\Omega_{u})$$

$$+ 4\sum_{i=1}^{2} \sum_{j=1}^{2} Qe_{1,i}e'_{1,i}Q\bar{Z}'L_{i}C\Omega_{u}\bar{Z}Qe_{1,i}e'_{1,j}Q tr(L_{j}G\Omega_{\nu}G'L_{i}C\Omega_{u})$$

$$+ 2\sum_{i=1}^{2} (e'_{1,i}Qe_{1,i})^{2}Q\bar{Z}'L_{i}C\Omega_{u}\bar{Z}Q tr(L_{i}G\Omega_{\nu}G'L_{i}C\Omega_{u})$$

$$+ 2\sum_{i=1}^{2} (e'_{1,i}Qe_{1,i}Qe_{1,i}e'_{1,i}Q\bar{Z}'L_{i}C\Omega_{u}\bar{Z}Q tr(L_{i}G\Omega_{\nu}G'L_{i}C\Omega_{u})$$

$$+ 2\sum_{i=1}^{2} (e'_{1,i}Qe_{1,i}Qe_{1,i}Qe_{1,i}e'_{1,i}Qtr(Q\bar{Z}'L_{i}C\Omega_{u}\bar{Z}) tr(L_{i}G\Omega_{\nu}G'L_{i}C\Omega_{u}) + o(T^{-2}).$$

By employing the results of equations (B.27)-(B.30) in equation (B.26), we obtain the MSE of the estimator up to order  $T^{-2}$ . This completes the proof of Lemma B.3.

### C Appendix C

**Lemma C.1** Under Assumption 1 (iii)–(vii), when the coefficients of the lagged dependent variable in the model (2.1) is equal to unity (be integrated of order 1, I(1)) or be integrated of higher order in both of the regimes, and the exogenous regressors contain non-stationary

components, it can be verified that

$$(\dot{Z}'\dot{Z})^{-1} = \dot{Q}[I_{2(k+1)} + \dot{A} + \dot{B}]^{-1} = \dot{Q} - \dot{Q}\dot{A} - \dot{Q}\dot{B} + \dot{Q}\dot{A}^2 + o_p(T^{-2}),$$
(C.1)

$$\dot{Z}'u = \ddot{\bar{Z}}'u + \ddot{\bar{Z}}'u = O_p(T^{1/2}) + O_p(1),$$
(C.2)

$$\dot{P} = \bar{P} + \tilde{P}_{-\frac{1}{2}} + o_p(T^{-1/2}), \tag{C.3}$$

where  $\dot{Q} \equiv (\dot{Z}'\dot{Z})^{-1} = O_p(T^{-1}), \ \dot{A} \equiv (\dot{Z}'\dot{Z} + \dot{Z}'\dot{Z})\dot{Q} = O_p(T^{-1/2}), \ and \ \dot{B} \equiv (\ddot{Z}'\dot{Z})\dot{Q} = O_p(T^{-1}), \ with \ \dot{Z} = \mathbb{E}(\dot{Z}), \ and \ \dot{Z} = \sum_{i=1}^{2} L_i \dot{G}\nu e'_{1,i}N, \ and \ \dot{G} \ is \ defined \ as \ G \ after \ setting \ \lambda_1 = 1, \ and \ \lambda_2 = 1. \ Also, \ \dot{P} = (\dot{Z}'\dot{Z})^{-1}R' \left(R(\dot{Z}'\dot{Z})^{-1}R'\right)^{-1}R, \ where \ \dot{P} = \dot{Q}R'(R\dot{Q}R')^{-1}R, \ and \ \dot{P}_{-\frac{1}{2}} = (\dot{P} - I_{2(k+1)})\dot{Q}\dot{A}\dot{Q}^{-1}\dot{P}.$ 

Proof: Similar to Section 2 of the paper, we decompose  $\dot{Z} = \dot{Z} + \tilde{Z}$ , where  $\dot{Z} = \mathbb{E}(\dot{Z}) = \mathbb{E}(Z)N = \bar{Z}N$ , and  $\tilde{\dot{Z}} = \dot{Z} - \dot{Z} = (Z - \bar{Z})N = \tilde{Z}N = \sum_{i=1}^{2} L_i \dot{G}\nu e'_{1,i}N$ . Hence, we have

$$\dot{Z}'\dot{Z} = \bar{\dot{Z}}'\bar{\dot{Z}} + \bar{\dot{Z}}'\tilde{\dot{Z}} + \tilde{\dot{Z}}'\bar{\dot{Z}} + \tilde{\dot{Z}}'\bar{\dot{Z}},$$

where  $\overline{\dot{Z}}'\overline{\dot{Z}} = O_p(T)$ ,  $\overline{\dot{Z}}'\widetilde{\dot{Z}} + \widetilde{\dot{Z}}'\overline{\dot{Z}} = O_p(T^{1/2})$ , and  $\widetilde{\dot{Z}}'\widetilde{\dot{Z}} = O_p(1)$ . Using this expansion, the results of Lemma C.1 can be verified easily.

**Lemma C.2** Under Assumption 1 (iii)–(vii), when the coefficients of the lagged dependent variable in the model (2.1) is equal to unity (be integrated of order 1, I(1)) or be integrated of higher order in both of the regimes, and the exogenous regressors contain non-stationary components, the bias of the rescaled unrestricted estimator up to order  $T^{-1}$  is

$$Bias(\hat{\alpha}) = \mathbb{E}(\hat{\alpha} - \dot{\alpha}) = -\dot{Q}\sum_{i=1}^{2} \left[ \bar{\dot{Z}}'L_i \dot{C}\Omega_u \bar{\dot{Z}} \dot{Q} N e_{1,i} + N e_{1,i} \operatorname{tr}(Q \bar{\dot{Z}}'L_i \dot{C}\Omega_u \bar{\dot{Z}}) \right] \equiv \dot{\Theta}, \qquad (C.4)$$

where  $\dot{C}$  is defined as C after replacing  $\lambda_1$  and  $\lambda_2$  with one. The MSE of the rescaled

unrestricted estimator up to order  $T^{-2}$  is

$$\begin{split} MSE\left(\hat{\hat{\mathbf{u}}}\right) &= \dot{Q}\Sigma + \sum_{i=1}^{2} \dot{Q}Ne_{1,i}e_{1,i}^{i}N\dot{Q}\sigma_{i}^{2} tr(\dot{G}_{i}\dot{G}_{i}^{i}\Omega_{\nu}) \\ &+ \sum_{i=1}^{2} \sum_{j=1}^{2} \dot{Q}\dot{Z}^{i} \left[\sigma_{i}^{2}L_{i}\dot{G}\Omega_{\nu}\dot{G}^{i}L_{j} - \sigma_{j}^{2}L_{i}\dot{C}L_{j}\dot{C}\Omega_{u} - 2\Omega_{u}\dot{C}^{i}L_{i}L_{j}\dot{C}\Omega_{u} - \Omega_{u}\dot{C}^{i}L_{i}\Omega_{u}\dot{C}^{i}L_{j}\right] \dot{Z}\dot{Q}Ne_{1,i}e_{1,j}^{i}N\dot{Q} \\ &+ \sum_{i=1}^{2} \sum_{j=1}^{2} \dot{Q}Ne_{1,i}e_{1,j}^{i}N\dot{Q}\dot{Z}^{i} \left[\sigma_{j}^{2}L_{i}\dot{G}\Omega_{\nu}\dot{G}^{i}L_{i} - \sigma_{i}^{2}L_{i}\dot{C}L_{j}\dot{C}\Omega_{u} - 2\Omega_{u}\dot{C}^{i}L_{i}L_{j}\dot{C}\Omega_{u} - \sigma_{i}^{2}\Omega_{u}\dot{C}^{i}L_{i}\dot{C}^{i}L_{j}\dot{Q} \\ &+ \sum_{i=1}^{2} c_{1,i}^{i}N\dot{Q}Ne_{1,i}\dot{Q}\dot{Z}^{i} \left[\sigma_{i}^{2}L_{i}\dot{G}\Omega_{\nu}\dot{G}^{i}L_{i} - \sigma_{i}^{2}L_{i}\dot{C}L_{i}\dot{C}\Omega_{u} - \sigma_{i}^{2}\Omega_{u}\dot{C}^{i}L_{i}\dot{C}\dot{Q}_{u} \\ &+ \sum_{i=1}^{2} c_{1,i}^{i}N\dot{Q}Ne_{1,i}\dot{Q}\dot{Z}^{i} \left[\sigma_{i}^{2}U\dot{Q}\dot{Z}^{i}L_{i}\dot{G}\Omega_{\nu}\dot{G}^{i}L_{j}\dot{Z}\dot{Q} \\ &+ \sum_{i=1}^{2} c_{1,i}^{2}\dot{Q}Ne_{1,i}e_{1,j}^{i}N\dot{Q} \left[\sigma_{i}^{2}tr(\dot{Q}\dot{Z}^{i}L_{i}\dot{G}\Omega_{\nu}\dot{G}^{i}L_{j}\dot{Z}\dot{Q} \\ &+ \sum_{i=1}^{2} \sum_{j=1}^{2} \dot{Q}Ne_{1,i}e_{1,j}^{i}N\dot{Q} \left[tr(\dot{Q}\dot{Z}^{i}L_{i}\dot{G}\Omega_{u}\dot{G}^{i}L_{j}\dot{Z}\dot{Q} \\ &+ \sum_{i=1}^{2} \sum_{j=1}^{2} \dot{Q}Ne_{1,i}e_{1,j}^{i}N\dot{Q} \left[tr(\dot{Q}\dot{Z}^{i}L_{i}\dot{G}\Omega_{u}\dot{Z}\dot{Q}\dot{Z}^{i}L_{j}\dot{G}\Omega_{u}\dot{Z}\dot{Q} \\ &+ \sum_{i=1}^{2} \sum_{j=1}^{2} \dot{Q}Ne_{1,i}e_{1,j}^{i}N\dot{Q} \dot{Z}^{i} \left[L_{i}\dot{C}\Omega_{u} + \Omega_{u}\dot{C}^{i}L_{i}\right] \ddot{Z}\dot{Q} \\ &+ \sum_{i=1}^{2} \sum_{j=1}^{2} \dot{Q}Ne_{1,i}\dot{Q} \dot{Z}^{i}L_{i}\dot{C}\Omega_{u}\dot{Z}\dot{Q}Ne_{1,i}\dot{Q} \dot{Z}^{i}L_{i}\dot{G}\Omega_{u}\dot{Z}\dot{Q} \\ &+ \sum_{i=1}^{2} \sum_{j=1}^{2} \dot{Q}\dot{Z}^{i}L_{i}\dot{C}\Omega_{u}\ddot{Z}\dot{Q}Ne_{1,i}\dot{Q} \dot{Z}^{i} \left[L_{i}\dot{C}\Omega_{u} + \Omega_{u}\dot{C}^{i}L_{i}\right] \ddot{Z}\dot{Q} \\ &+ \sum_{i=1}^{2} \sum_{j=1}^{2} \dot{Q}Ne_{1,i}\dot{Q} \dot{Z}^{i}L_{i}\dot{C}\Omega_{u}\ddot{Z}\dot{Q}\dot{Z}^{i}L_{i}\dot{C}\Omega_{u}\ddot{Z}\dot{Q} \\ &+ \sum_{i=1}^{2} \sum_{j=1}^{2} \dot{Q}Ne_{1,i}\dot{Q} \dot{Z}^{i}\dot{Q} \\ &+ \sum_{i=1}^{2} \sum_{j=1}^{2} \dot{Q}Ne_{1,i}\dot{Q} \dot{Q} \\ &+ \sum_{i=1}^{2} \sum_{j=1}^{2} \dot{Q}Ne_{1,i}\dot{Q} \\ &+ \sum_{i=1}^{2} \dot{Q}Ne_{$$

$$+\sum_{i=1}^{2}\sum_{j=1}^{2}\dot{Q}\bar{Z}'\Big[L_{i}\dot{C}\Omega_{u}+\Omega_{u}\dot{C}'L_{i}\Big]\bar{Z}\dot{Q}\bar{Z}'L_{j}\dot{C}\Omega_{u}\bar{Z}\dot{Q}Ne_{1,i}e_{1,j}'N\dot{Q}$$
  
$$+\sum_{i=1}^{2}\sum_{j=1}^{2}\dot{Q}\bar{Z}'\Omega_{u}\dot{C}'L_{i}\bar{Z}\dot{Q}\bar{Z}'\Big[L_{j}\dot{C}\Omega_{u}+\Omega_{u}\dot{C}'L_{j}\Big]\bar{Z}\dot{Q}Ne_{1,j}e_{1,i}'N\dot{Q}$$
  
$$-\dot{Q}\sum_{i=1}^{2}Ne_{1,i}e_{1,i}'N\dot{Q}\bar{Z}'\Omega_{u}\bar{Z}\dot{Q}-\dot{Q}\sum_{i=1}^{2}\bar{Z}'\Omega_{u}\bar{Z}\dot{Q}Ne_{1,i}e_{1,i}'N\dot{Q},$$
 (C.5)

where  $\Sigma = diag(\sigma_1^2 I_{k+1}, \sigma_2^2 I_{k+1}).$ 

Proof: Using Lemma C.1, we have

$$\tilde{\dot{\alpha}} - \dot{\alpha} = (\dot{Z}'\dot{Z})^{-1}\dot{Z}'u = \dot{\xi}_{-\frac{1}{2}} + \dot{\xi}_{-1} + \dot{\xi}_{-\frac{3}{2}} + O_p(T^{-2}),$$
(C.6)

where  $\dot{\xi}_{-\frac{1}{2}}$ ,  $\dot{\xi}_{-\frac{1}{2}}$ , and  $\dot{\xi}_{-\frac{3}{2}}$  are defined below, and the suffixes show the order of magnitude in probability,

$$\begin{split} \dot{\xi}_{-\frac{1}{2}} &= \dot{Q}\bar{\dot{Z}}'(0,I_T)\nu = O_p(T^{-\frac{1}{2}}), \\ \dot{\xi}_{-1} &= \dot{Q}\sum_{i=1}^2 \nu'\dot{H}_i\nu Ne_{1,i} - \dot{Q}\dot{A}\bar{\dot{Z}}'(0,I_T)\nu = O_p(T^{-1}), \\ \dot{\xi}_{-\frac{3}{2}} &= -\dot{Q}\dot{A}\sum_{i=1}^2 \nu'\dot{H}_i\nu Ne_{1,i} - \dot{Q}\dot{B}\bar{\dot{Z}}'(0,I_T)\nu + \dot{Q}\dot{A}^2\bar{\dot{Z}}'(0,I_T)\nu = O_p(T^{-\frac{3}{2}}), \end{split}$$

where  $\dot{H} = \dot{G}' L_i(0, I_T)$ .

The bias of the unrestricted estimator using the approximations in equation (C.6) up to order  $T^{-1}$  is

$$\mathbb{E}(\hat{\dot{\alpha}} - \dot{\alpha}) = \mathbb{E}(\dot{\xi}_{-\frac{1}{2}}) + \mathbb{E}(\dot{\xi}_{-1}) = -\dot{Q}\sum_{i=1}^{2} \left[ \bar{Z}'L_i \dot{C}\Omega_u \bar{Z}\dot{Q}e_{1,i} + e_{1,i} \operatorname{tr}(\dot{Q}\bar{Z}'L_i \dot{C}\Omega_u \bar{Z}) \right], \quad (C.7)$$

where the last equality holds because

$$\mathbb{E}(\dot{\xi}_{-\frac{1}{2}}) = 0, \tag{C.8}$$

and by using Lemma B.1,  $\mathbb{E}(\dot{\xi}_{-1})$  is

$$\mathbb{E}(\dot{\xi}_{-1}) = \sum_{i=1}^{2} \dot{Q} N e_{1,i} \mathbb{E}(\nu' \dot{H}_{i}\nu) - \dot{Q} \mathbb{E}\left\{\dot{A}\bar{Z}'(0,I_{T})\nu\right\}$$
  
$$= \sum_{i=1}^{2} \dot{Q} N e_{1,i} tr(L_{i}\dot{C}\Omega_{u}) - \dot{Q}\sum_{i=1}^{2} \left[\bar{Z}'L_{i}\dot{C}\Omega_{u}\bar{Z}\dot{Q}Ne_{1,i} + Ne_{1,i} tr(Q\bar{Z}'L_{i}C\Omega_{u}\bar{Z})\right] \quad (C.9)$$
  
$$= -\dot{Q}\sum_{i=1}^{2} \left[\bar{Z}'L_{i}\dot{C}\Omega_{u}\bar{Z}\dot{Q}Ne_{1,i} + Ne_{1,i} tr(Q\bar{Z}'L_{i}\dot{C}\Omega_{u}\bar{Z})\right] \equiv \dot{\Theta},$$

where  $\dot{C}$  has zeroes on and above its main diagonal and components unity below. Since the diagonal elements of  $\dot{C}$  are zero,  $tr(L_i\dot{C}\Omega_u) = 0$ .

The MSE up to order  $T^{-2}$  is

$$\mathbb{E}\left[(\hat{\dot{\alpha}} - \dot{\alpha})(\hat{\dot{\alpha}} - \dot{\alpha})'\right] = \mathbb{E}(\dot{\xi}_{-\frac{1}{2}}\dot{\xi}'_{-\frac{1}{2}} + \dot{\xi}_{-\frac{1}{2}}\dot{\xi}'_{-1} + \dot{\xi}_{-1}\dot{\xi}'_{-\frac{1}{2}} + \dot{\xi}_{-\frac{1}{2}}\dot{\xi}'_{-\frac{3}{2}} + \dot{\xi}_{-\frac{3}{2}}\dot{\xi}'_{-\frac{1}{2}} + \dot{\xi}_{-1}\dot{\xi}'_{-1}).$$
(C.10)

We derive the expectations of the terms in (C.10) in the rest of the proof using Lemma B.1. We obtain

$$\mathbb{E}(\dot{\xi}_{-\frac{1}{2}}\dot{\xi}'_{-\frac{1}{2}}) = \mathbb{E}\left\{\dot{Q}\ddot{Z}'(0,I_T)\nu\nu'(0,I_T)'\dot{Z}\dot{Q}\right\} = Q\dot{Z}'\Omega_u\dot{Z}\dot{Q} = \dot{Q}\Sigma,$$
(C.11)

where the last equality holds because  $\dot{Q}\dot{Z}'\Omega_u\dot{Z} = \dot{Q}\dot{Z}'\dot{Z}\Sigma = \Sigma$ .

$$\mathbb{E}(\dot{\xi}_{-1}\dot{\xi}'_{-\frac{1}{2}}) = \mathbb{E}\left\{\dot{Q}\sum_{i=1}^{2}\nu'\dot{H}_{i}\nu Ne_{1,i}\nu'(0,I_{T})'\bar{Z}\dot{Q}\right\} - \dot{Q}\mathbb{E}\left\{\dot{A}\bar{Z}'(0,I_{T})\nu\nu'(0,I_{T})'\bar{Z}\right\}\dot{Q} = 0.$$
(C.12)

$$\mathbb{E}(\dot{\xi}_{-1}\dot{\xi}'_{-1}) = \dot{Q} \mathbb{E}\left\{ \left[ \sum_{i=1}^{2} \nu' \dot{H}_{i} \nu N e_{1,i} - \dot{A} \bar{Z}'(0, I_{T}) \nu \right] \left[ e_{1,i}' N \sum_{i=1}^{2} \nu' \dot{H}_{i} \nu - \nu'(0, I_{T})' \bar{Z} \dot{A} \right] \right\} \dot{Q}$$

$$\begin{split} &= \sum_{i=1}^{2} \dot{Q} N e_{1,i} e_{1,i}' N \dot{Q} \sigma_{i}^{2} tr(\dot{G}_{i}\dot{G}_{i}'\Omega_{\nu}) \\ &= \dot{Q} \sum_{i=1}^{2} \sum_{j=1}^{2} \sigma_{j}^{2} \vec{Z}' \Big[ L_{i}\dot{G}\Omega_{\nu}\dot{G}'L_{j} + L_{i}\dot{C}L_{j}\dot{C}\Omega_{u} \Big] \ddot{Z}\dot{Q} N e_{1,i} e_{1,j}' N \dot{Q} \\ &- \dot{Q} \sum_{i=1}^{2} \sum_{j=1}^{2} \sigma_{j}^{2} N e_{1,i} e_{1,j}' N \dot{Q} \Big[ tr(\dot{Q}\ddot{Z}'L_{i}\dot{G}\Omega_{\nu}\dot{G}'L_{j}\ddot{Z}) + tr(\dot{Q}\ddot{Z}'L_{i}\dot{C}L_{j}\dot{C}\Omega_{u}\ddot{Z}) \Big] \\ &- \Big( \dot{Q} \sum_{i=1}^{2} \sum_{j=1}^{2} \sigma_{j}^{2} \ddot{Z}' \Big[ L_{i}\dot{G}\Omega_{\nu}\dot{G}'L_{j} + L_{i}\dot{C}L_{j}\dot{C}\Omega_{u} \Big] \ddot{Z}\dot{Q} N e_{1,i} e_{1,j}' N \dot{Q} \Big)' \\ &- \Big( \dot{Q} \sum_{i=1}^{2} \sum_{j=1}^{2} \sigma_{j}^{2} N e_{1,i} e_{1,j}' N \dot{Q} \Big[ tr(\dot{Q}\ddot{Z}'L_{i}\dot{G}\Omega_{\nu}\dot{G}'L_{j}\ddot{Z}) + tr(\dot{Q}\ddot{Z}'L_{i}\dot{C}L_{j}\dot{C}\Omega_{u}\ddot{Z}) \Big] \Big)' \\ &+ \sum_{i=1}^{2} \sigma_{i}^{2} c_{1,i}' N \dot{Q} N e_{1,i} e_{1,j}' N \dot{Q} \Big[ tr(\dot{Q}\ddot{Z}'L_{i}\dot{G}\Omega_{\nu}\dot{G}'L_{j}\ddot{Z}) + tr(\dot{Q}\ddot{Z}'L_{i}\dot{C}\Omega_{u}\ddot{Z}) N e_{1,j} e_{1,i}' N \dot{Q} \ddot{Z}'\Omega_{u}\dot{C}'L_{j}\ddot{Z}\dot{Q} \\ &+ \sum_{i=1}^{2} \sum_{j=1}^{2} \dot{Q}\ddot{Z}'L_{i}\dot{C}\Omega_{u}\ddot{Z}\dot{Q} N e_{1,i} e_{1,j}' N \dot{Q} \ddot{Z}'\Omega_{u}\dot{C}'L_{j}\ddot{Z}\dot{Q} + \sum_{i=1}^{2} \sum_{j=1}^{2} \dot{Q}\ddot{Z}'L_{i}\dot{C}\Omega_{u}\ddot{Z}\dot{Q} N e_{1,i} e_{1,j}' N \dot{Q} tr(\dot{Q}\ddot{Z}'L_{j}\dot{C}\Omega_{u}\ddot{Z}) \\ &+ \sum_{i=1}^{2} \sum_{j=1}^{2} \dot{Q}\ddot{Z}'L_{i}\dot{C}\Omega_{u}\ddot{Z}\dot{Q} X'L_{j}\dot{C}\Omega_{u}\ddot{Z}\dot{Q} N e_{1,i} e_{1,j}' N \dot{Q} \ddot{Z}'\Omega_{u}\dot{C}'L_{j}\ddot{Z}\dot{Q} \\ &+ \sum_{i=1}^{2} \sum_{j=1}^{2} \dot{Q}\ddot{Z}'L_{i}\dot{C}\Omega_{u}\ddot{Z}\dot{Q} \ddot{Z}'L_{j}\dot{C}\Omega_{u}\ddot{Z}\dot{Q} N e_{1,i} e_{1,j}' N \dot{Q} \dot{Z}'\Omega_{u}\dot{C}'L_{j}\ddot{Z}\dot{Q} \\ &+ \sum_{i=1}^{2} \sum_{j=1}^{2} \dot{Q}N e_{1,i} e_{1,j}' N \dot{Q} \ddot{Z}'\Omega_{u}\dot{C}'L_{j}\ddot{Z}\dot{Q} tr(\dot{Q}\ddot{Z}'L_{i}\dot{C}\Omega_{u}\ddot{Z}) \\ &+ \sum_{i=1}^{2} \sum_{j=1}^{2} \dot{Q}N e_{1,i} e_{1,j}' N \dot{Q} \ddot{Z}'\Omega_{u}\dot{C}'L_{i}\ddot{Z}\dot{Q} \ddot{Z}'\Omega_{u}\dot{C}'L_{j}\ddot{Z}\dot{Q} \\ &+ \sum_{i=1}^{2} \sum_{j=1}^{2} \dot{Q}N e_{1,i} e_{1,j}' N \dot{Q} \ddot{Z}'L_{i}\dot{G}\Omega_{u}\dot{G}'L_{j}\ddot{Z}\dot{Q} \\ &+ \sum_{i=1}^{2} \sum_{j=1}^{2} \dot{Q}N e_{1,i} e_{1,j}' N \dot{Q} \ddot{Z}'\Omega_{u}\dot{C}'L_{i}\ddot{Z}\dot{Q} \dot{Z}'\Omega_{u}\dot{C}'L_{j}\ddot{Z}\dot{Q} \\ &+ \sum_{i=1}^{2} \sum_{j=1}^{2} \dot{Q}N e_{1,i} e_{1,j}' N \dot{Q} \ddot{Z}'\Omega_{u}\dot{C}'L_{i}\ddot{Z}\dot{Q}\dot{Z}'\Omega_{u}\dot{Z}'\Omega_{u}\dot{Z}'\Omega_{u}\dot{Z}'\Omega_{u}\dot{Z}'\Omega_{u}\dot{Z}'\Omega_{u}\dot{Z}'\Omega_{u}\dot{Z}'\Omega_{u}\dot$$

Also, by Lemma B.1, we obtain

$$\mathbb{E}(\dot{\xi}_{-\frac{3}{2}}\dot{\xi}'_{-\frac{1}{2}}) = \dot{Q} \mathbb{E}\left\{ \left[ -\dot{A}\sum_{i=1}^{2}\nu'\dot{H}_{i}\nu Ne_{1,i} - \dot{B}\bar{Z}'(0,I_{T})\nu + \dot{A}^{2}\bar{Z}'(0,I_{T})\nu \right]\nu'(0,I_{T})'\bar{Z} \right\} \dot{Q}$$
$$= -\dot{Q}\sum_{i=1}^{2}\sigma_{i}^{2}e_{1,i}'N\dot{Q}Ne_{1,i}\dot{Q}\bar{Z}' \left[ L_{i}\dot{G}\Omega_{\nu}\dot{G}'L_{i} + L_{i}\dot{C}L_{i}\dot{C}\Omega_{u} \right]\bar{Z}\dot{Q}$$

$$\begin{split} &-\dot{Q}\sum_{i=1}^{2}\sum_{j=1}^{2}\sigma_{j}^{2}Ne_{1,i}e_{1,j}^{\prime}N\dot{Q}\bar{Z}^{\prime}\Big[L_{i}\dot{G}\Omega_{\nu}\dot{G}^{\prime}L_{j}+L_{i}\dot{C}L_{j}\dot{C}\Omega_{u}\Big]\bar{Z}\dot{Q} \\ &-2\dot{Q}\sum_{i=1}^{2}Ne_{1,i}e_{1,i}^{\prime}N\dot{Q}\bar{Z}^{\prime}\Omega_{u}\dot{C}^{\prime}L_{i}\dot{C}\Omega_{u}\bar{Z}\dot{Q}-\dot{Q}\sum_{i=1}^{2}Ne_{1,i}e_{1,i}^{\prime}N\dot{Q}\bar{Z}^{\prime}\Omega_{u}\bar{Z}\dot{Q} \\ &+\sum_{i=1}^{2}\sum_{j=1}^{2}e_{1,i}^{\prime}N\dot{Q}\bar{Z}^{\prime}L_{j}\dot{C}\Omega_{u}\bar{Z}\dot{Q}Ne_{1,j}\dot{Q}\bar{Z}^{\prime}L_{i}\dot{C}\Omega_{u}\bar{Z}\dot{Q}+\sum_{i=1}^{2}\sum_{j=1}^{2}\sigma_{j}^{2}\dot{Q}\bar{Z}^{\prime}L_{i}\dot{G}\Omega_{\nu}\dot{G}^{\prime}L_{j}\bar{Z}\dot{Q}Ne_{1,i}e_{1,j}^{\prime}N\dot{Q} \\ &+\sum_{i=1}^{2}\sum_{j=1}^{2}\dot{Q}\bar{Z}^{\prime}L_{i}\dot{C}\Omega_{u}\bar{Z}\dot{Q}Ne_{1,i}e_{1,i}^{\prime}N\dot{Q}\bar{Z}^{\prime}L_{j}\dot{C}\Omega_{u}\dot{Z}\dot{Q}+\sum_{i=1}^{2}e_{1,i}^{\prime}N\dot{Q}Ne_{1,i}\dot{Q}\bar{Z}^{\prime}L_{i}\dot{C}\Omega_{u}\bar{Z}\dot{Q} tr(\dot{Q}\bar{Z}^{\prime}L_{i}\dot{C}\Omega_{u}\bar{Z}\dot{Q}) \\ &+\sum_{i=1}^{2}\sigma_{i}^{2}\dot{Q}_{1,i}N\dot{Q}Ne_{1,i}\dot{Q}\bar{Z}^{\prime}L_{i}\dot{G}\Omega_{\nu}\dot{G}^{\prime}L_{i}\bar{Z}\dot{Q}+\sum_{i=1}^{2}e_{1,i}^{\prime}N\dot{Q}Ne_{1,i}\dot{Q}\bar{Z}^{\prime}L_{i}\dot{C}\Omega_{u}\bar{Z}\dot{Q} \\ &+\sum_{i=1}^{2}\sum_{j=1}^{2}\sigma_{j}^{2}\dot{Q}Ne_{1,i}e_{1,j}^{\prime}N\dot{Q}tr(\dot{Q}\bar{Z}^{\prime}L_{j}\dot{G}\Omega_{\nu}\dot{G}^{\prime}L_{i}\bar{Z})+\sum_{i=1}^{2}\sum_{j=1}^{2}\dot{Q}Ne_{1,i}e_{1,j}^{\prime}N\dot{Q}\bar{Z}^{\prime}L_{i}\dot{C}\Omega_{\nu}\bar{Z}\dot{Q} \\ &+\sum_{i=1}^{2}\sum_{j=1}^{2}\dot{Q}Ne_{1,i}e_{1,j}^{\prime}N\dot{Q}\bar{Z}^{\prime}\Omega_{u}\dot{C}^{\prime}L_{j}\bar{Z}\dot{Q}\bar{Z}^{\prime}L_{i}\dot{C}\Omega_{u}\bar{Z}\dot{Q}+\sum_{i=1}^{2}\sum_{j=1}^{2}\dot{Q}Ne_{1,i}e_{1,j}^{\prime}N\dot{Q}\bar{Z}^{\prime}L_{i}\dot{C}\Omega_{\nu}\bar{Z}\dot{Q}tr(\dot{Q}\bar{Z}^{\prime}L_{j}\dot{C}\Omega_{u}\bar{Z}) \\ &+\sum_{i=1}^{2}\sum_{j=1}^{2}\sigma_{j}^{2}\dot{Q}Ne_{1,i}e_{1,j}^{\prime}N\dot{Q}\bar{Z}^{\prime}L_{i}\dot{G}\Omega_{\nu}\dot{G}^{\prime}L_{j}\bar{Z}\dot{Q}+\sum_{i=1}^{2}\sum_{j=1}^{2}\dot{Q}Ne_{1,i}e_{1,j}^{\prime}N\dot{Q}\bar{Z}^{\prime}L_{i}\dot{C}\Omega_{\nu}\bar{Z}\dot{Q}tr(\dot{Q}\bar{Z}^{\prime}L_{j}\dot{C}\Omega_{u}\bar{Z}) \\ &+\sum_{i=1}^{2}\sum_{j=1}^{2}\sigma_{j}^{2}\dot{Q}Ne_{1,i}e_{1,j}^{\prime}N\dot{Q}\bar{Z}^{\prime}L_{i}\dot{G}\Omega_{\nu}\dot{G}^{\prime}L_{j}\bar{Z}\dot{Q}+\sum_{i=1}^{2}\sum_{j=1}^{2}\dot{Q}Ne_{1,i}e_{1,j}^{\prime}N\dot{Q}\bar{Z}^{\prime}L_{i}\dot{C}\Omega_{u}\bar{Z}\dot{Q}. \end{split} \end{split}$$

By employing the results of equations (C.11)-(C.14) in equation (C.10), we obtain the MSE of the estimator up to order  $T^{-2}$ . This completes the proof of Lemma C.2.

**Remark C.3** We note that the bias and MSE of the unrestricted estimator can be obtained using the bias and MSE of the rescaled unrestricted estimator as follows:

$$Bias(\widehat{\alpha}) = \mathbb{E}(\widehat{\alpha} - \alpha) = NBias(\widehat{\dot{\alpha}}),$$

$$MSE(\widehat{\alpha}) = \mathbb{E}\left[(\widehat{\alpha} - \alpha)(\widehat{\alpha} - \alpha)'\right] = N MSE(\widehat{\alpha})N,$$

where the  $j^{th}$  element of bias is approximated up to order  $T^{-1-g_j}$ , and the  $(j,l)^{th}$  element of MSE is approximated up to order  $T^{-2-g_j-g_l}$ .

# D Appendix D

This section contains further simulation results of the paper when k = 3, and k = 8, and the simulation results for  $k \in \{3, 5, 8\}$  when  $\epsilon_t \stackrel{\text{i.i.d.}}{\sim} (\chi^2(2) - 2)/2$ , where  $\chi^2(2)$  denotes the chi-squared distribution with two degrees of freedom.

	$b_1:$	0	.2	0.	4	0	.6	0.	.8
Exp.	$\delta$	$\overline{\text{RMSFE}}(\breve{\alpha}_{2})$	$_{2}$ )RMSFE( $\tilde{\alpha}_{2}$	$_{2})$ RMSFE( $\breve{\alpha}_{2}$	)RMSFE( $\tilde{\alpha}_2$	$(\breve{\alpha}_2)$ RMSFE $(\breve{\alpha}_2)$	$_{2}$ )RMSFE( $\tilde{\alpha}_{2}$	$(\breve{\alpha}_2)$ RMSFE $(\breve{\alpha}_2)$	$(\widetilde{\alpha}_2)$ RMSFE $(\widetilde{\alpha}_2)$
	0.000	0.806	0.391	0.810	0.392	0.813	0.406	0.808	0.410
	0.250	0.826	0.454	0.854	0.558	0.862	0.611	0.877	0.670
_//_1	0.500	0.864	0.742	0.897	1.265	0.884	1.330	0.892	1.049
<i></i> #1	0.750	0.932	1.614	0.957	3.380	0.937	3.664	0.918	2.283
	1.000	0.969	2.875	0.983	6.264	0.970	7.220	0.952	4.554
	0.000	0.805	0.397	0.814	0.393	0.819	0.376	0.814	0.375
	0.250	0.832	0.482	0.842	0.544	0.851	0.584	0.867	0.598
49	0.500	0.875	0.796	0.894	1.354	0.883	1.412	0.897	1.074
#4	0.750	0.930	1.650	0.955	3.486	0.937	3.745	0.925	2.371
	1.000	0.968	2.956	0.982	6.410	0.970	7.322	0.955	4.678
	0.000	0.860	0.516	0.852	0.495	0.829	0.448	0.830	0.458
	0.250	0.882	0.599	0.884	0.694	0.868	0.697	0.878	0.710
<i>-</i> # <b>9</b>	0.500	0.915	0.956	0.920	1.630	0.900	1.743	0.899	1.273
#3	0.750	0.956	1.868	0.968	3.870	0.949	4.455	0.928	2.892
π0	1.000	0.985	3.158	0.988	6.908	0.977	8.359	0.962	5.654
	0.000	0.900	0.645	0.903	0.671	0.904	0.652	0.914	0.649
	0.250	0.902	0.716	0.905	0.868	0.894	0.850	0.905	0.753
-#-1	0.500	0.926	1.134	0.938	1.999	0.923	2.156	0.920	1.520
#4	0.750	0.962	2.050	0.971	4.188	0.956	4.881	0.944	3.406
	1.000	0.985	3.332	0.985	7.201	0.974	8.813	0.966	6.361
	0.000	0.896	0.686	0.890	0.778	0.871	0.735	0.871	0.706
	0.250	0.910	0.845	0.911	1.208	0.889	1.243	0.892	1.046
#5	0.500	0.943	1.442	0.950	2.769	0.930	3.183	0.916	2.189
$\pi 0$	0.750	0.973	2.504	0.979	5.410	0.965	6.722	0.951	4.672
	1.000	0.991	3.940	0.990	8.983	0.981	11.423	0.975	8.453
	0.000	0.947	1.011	0.972	1.338	0.974	1.351	0.963	1.006
	0.250	0.954	1.161	0.969	1.583	0.970	1.607	0.965	1.184
#6	0.500	0.970	1.649	0.974	2.696	0.972	2.932	0.970	2.072
$\pi$ 0	0.750	0.983	2.483	0.982	4.677	0.977	5.404	0.976	3.740
	1.000	0.990	3.635	0.988	7.558	0.982	9.028	0.980	6.207

Table D1: Simulation results with k = 3

Note: This table reports the results of the RMSFE where the benchmark model is the unrestricted estimator. The first column shows the experiment numbers which represent the break specifications based on Table 1, and the second column is the break size in the slope coefficients. In the heading of the table,  $\text{RMSFE}(\check{\alpha}_2) = \text{MSFE}(\check{\alpha}_2)/\text{MSFE}(\widehat{\alpha}_2)$  shows the RMSFE of the shrinkage estimator, and  $\text{RMSFE}(\check{\alpha}_2) = \text{MSFE}(\check{\alpha}_2)/\text{MSFE}(\widehat{\alpha}_2)$  is the RMSFE of the restricted estimator.

	$b_1:$	0	.2	0	.4	0	.6	0	.8
Exp.	$\delta$	$RMSFE(\check{\alpha}_{2})$	$_{2}$ )RMSFE( $\tilde{\alpha}_{2}$	) RMSFE( $\breve{\alpha}_{2}$	$_{2})$ RMSFE $(\tilde{\alpha}_{2})$	) RMSFE( $\check{\alpha}_{2}$	$_{2})$ RMSFE( $\widetilde{\alpha}_{2}$	$RMSFE(\check{\alpha})$	$_{2}$ )RMSFE( $\widetilde{\alpha}_{2}$ )
	0.000	0.968	2.397	0.973	3.669	0.943	3.233	0.893	1.847
	0.250	0.982	2.847	0.987	4.720	0.970	4.653	0.925	2.843
117	0.500	0.993	3.649	0.994	6.696	0.986	7.365	0.957	4.826
<i>#1</i>	0.750	0.997	4.760	0.998	9.636	0.994	11.331	0.978	7.819
	1.000	0.998	6.216	0.999	13.518	0.997	16.605	0.989	11.894
	0.000	0.789	0.318	0.789	0.263	0.779	0.214	0.766	0.208
	0.250	0.807	0.349	0.827	0.365	0.846	0.422	0.868	0.560
#8	0.500	0.862	0.613	0.901	1.140	0.893	1.234	0.901	1.032
#0	0.750	0.939	1.527	0.970	3.256	0.952	3.551	0.931	2.409
	1.000	0.984	2.836	0.989	6.120	0.980	7.151	0.960	4.790
	0.000	0.942	0.844	0.967	0.932	0.958	0.899	0.891	0.655
	0.250	0.967	1.094	0.975	1.315	0.961	1.256	0.945	0.959
-#0	0.500	0.980	2.317	0.994	4.960	0.983	5.008	0.961	2.234
<i>#</i> 9	0.750	0.999	5.859	1.001	13.601	1.003	15.277	0.993	7.279
	1.000	1.002	10.845	1.001	24.773	1.001	29.547	1.001	15.702
	0.000	0.925	0.619	0.926	0.618	0.931	0.633	0.926	0.644
	0.250	0.925	0.661	0.934	0.764	0.937	0.834	0.957	0.894
#10	0.500	0.943	1.055	0.950	1.853	0.947	2.067	0.959	1.511
#10	0.750	0.968	2.244	0.987	5.305	0.987	6.425	0.987	3.972
	1.000	0.994	4.184	0.999	9.796	1.005	12.346	1.020	8.866
	0.000	0.917	0.528	0.908	0.434	0.904	0.382	0.905	0.429
	0.250	0.920	0.534	0.921	0.573	0.929	0.663	0.954	0.828
#11	0.500	0.936	0.955	0.952	1.927	0.948	2.205	0.959	1.567
#11	0.750	0.978	2.295	0.987	5.290	0.979	6.464	0.977	4.386
	1.000	0.996	4.116	0.996	9.655	0.992	12.410	0.995	9.597
	0.000	0.995	0.921	0.999	0.971	0.997	0.937	0.994	0.822
	0.250	0.997	1.156	1.000	1.471	0.999	1.461	0.997	1.150
#19	0.500	1.000	2.764	1.000	7.087	1.004	8.126	1.000	3.599
#14	0.750	1.000	7.918	1.000	20.866	1.001	25.534	1.008	13.715
	1.000	1.000	15.183	1.000	37.949	1.000	48.573	1.012	32.209

Table D1: Simulation results with k = 3 (Cont.)

	$b_1$ :	0	.2	0.	.4	0	.6	0	.8
Exp.	$\delta$	$RMSFE(\breve{\alpha}_2)$	$_{2}$ )RMSFE( $\widetilde{\alpha}_{2}$	$(\tilde{\alpha}_2)$ RMSFE $(\tilde{\alpha}_2)$	$(\widetilde{\alpha}_2)$ RMSFE $(\widetilde{\alpha}_2)$	$(\breve{\alpha}_2)$ RMSFE $(\breve{\alpha}_2)$	$_{2})$ RMSFE( $\widetilde{\alpha}_{2}$	) RMSFE( $\breve{\alpha}_2$	$_{2}$ )RMSFE( $\tilde{\alpha}_{2}$ )
	0.000	0.713	0.353	0.708	0.356	0.707	0.352	0.689	0.341
	0.250	0.757	0.496	0.772	0.640	0.763	0.629	0.751	0.560
// 1	0.500	0.892	1.263	0.904	2.126	0.864	2.114	0.799	1.178
<i></i> #1	0.750	0.975	2.556	0.964	4.669	0.942	5.107	0.882	2.567
	1.000	0.989	4.070	0.981	8.101	0.968	9.196	0.929	4.576
	0.000	0.724	0.390	0.739	0.407	0.718	0.330	0.705	0.264
	0.250	0.775	0.570	0.787	0.720	0.769	0.663	0.746	0.508
<i>4</i> 9	0.500	0.896	1.340	0.912	2.266	0.876	2.243	0.817	1.210
#2	0.750	0.976	2.673	0.966	4.814	0.944	5.241	0.890	2.637
	1.000	0.990	4.223	0.982	8.279	0.968	9.317	0.932	4.692
	0.000	0.731	0.435	0.736	0.440	0.701	0.362	0.654	0.281
	0.250	0.792	0.625	0.810	0.805	0.766	0.724	0.717	0.541
Щ2	0.500	0.913	1.419	0.926	2.405	0.886	2.436	0.805	1.307
#0	0.750	0.977	2.637	0.969	4.961	0.949	5.507	0.891	2.831
	1.000	0.989	4.155	0.983	8.449	0.972	9.755	0.935	4.998
	0.000	0.801	0.607	0.819	0.639	0.767	0.461	0.709	0.287
	0.250	0.849	0.841	0.862	1.062	0.818	0.915	0.762	0.538
-#-1	0.500	0.942	1.645	0.936	2.593	0.906	2.697	0.843	1.420
#4	0.750	0.978	2.837	0.969	5.166	0.951	5.803	0.905	2.994
	1.000	0.989	4.403	0.983	8.770	0.971	10.163	0.939	5.268
	0.000	0.803	0.655	0.809	0.704	0.743	0.510	0.674	0.329
	0.250	0.862	0.951	0.879	1.306	0.825	1.152	0.742	0.678
#5	0.500	0.949	1.786	0.947	2.997	0.916	3.149	0.842	1.722
#0	0.750	0.980	2.983	0.974	5.724	0.956	6.458	0.908	3.479
	1.000	0.989	4.580	0.985	9.501	0.974	11.094	0.942	5.958
	0.000	0.952	1.252	0.946	1.341	0.934	1.088	0.898	0.491
	0.250	0.961	1.508	0.952	1.818	0.941	1.621	0.913	0.753
46	0.500	0.975	2.195	0.965	3.305	0.955	3.387	0.933	1.630
#0	0.750	0.985	3.325	0.977	5.809	0.967	6.366	0.949	3.152
	1.000	0.992	4.889	0.985	9.348	0.977	10.616	0.961	5.330

Table D2: Simulation results with k = 8

	$b_1$ :	0	.2	0	.4	0	.6	0	.8
Exp.	$\delta$	$\overline{\text{RMSFE}}(\check{\alpha})$	$_{2})$ RMSFE( $\tilde{\alpha}_{2}$	) RMSFE( $\check{\alpha}_{2}$	$_{2})$ RMSFE( $\tilde{\alpha}_{2}$	) RMSFE( $\check{\alpha}_{i}$	$_{2})$ RMSFE( $\widetilde{\alpha}_{2}$	$(\breve{\alpha}_{2})$ RMSFE $(\breve{\alpha}_{2})$	$_{2}$ )RMSFE( $\widetilde{\alpha}_{2}$ )
	0.000	0.965	2.216	0.955	2.601	0.914	1.918	0.814	0.734
	0.250	0.979	2.603	0.971	3.476	0.946	2.947	0.874	1.262
117	0.500	0.988	3.398	0.982	5.399	0.968	5.265	0.921	2.490
<i>#1</i>	0.750	0.992	4.620	0.988	8.367	0.980	8.920	0.949	4.431
	1.000	0.995	6.286	0.992	12.392	0.987	13.866	0.966	7.097
	0.000	0.675	0.278	0.636	0.200	0.626	0.143	0.609	0.104
	0.250	0.712	0.348	0.720	0.402	0.730	0.418	0.721	0.406
48	0.500	0.886	1.060	0.927	1.975	0.896	1.988	0.832	1.145
#0	0.750	0.980	2.430	0.978	4.526	0.960	5.003	0.912	2.610
	1.000	0.992	3.958	0.989	7.977	0.979	9.100	0.948	4.665
	0.000	0.917	0.891	0.942	0.951	0.931	0.896	0.872	0.740
	0.250	0.961	1.495	0.962	1.853	0.940	1.640	0.908	1.005
-#0	0.500	1.000	4.263	0.995	7.937	0.982	8.093	0.933	2.997
<i>#</i> 9	0.750	1.002	8.762	0.998	17.688	0.991	19.787	0.967	8.114
	1.000	1.003	14.774	0.999	31.318	0.995	35.920	0.982	15.730
	0.000	0.769	0.454	0.773	0.451	0.772	0.441	0.783	0.470
	0.250	0.811	0.611	0.830	0.744	0.812	0.732	0.809	0.703
#10	0.500	0.916	1.434	0.925	2.491	0.897	2.639	0.848	1.598
#10	0.750	0.978	2.842	0.975	5.542	0.960	6.416	0.919	3.688
	1.000	0.991	4.582	0.987	9.618	0.979	11.507	0.957	6.569
	0.000	0.736	0.363	0.710	0.279	0.692	0.205	0.680	0.174
	0.250	0.769	0.454	0.783	0.516	0.784	0.546	0.781	0.567
#11	0.500	0.908	1.277	0.943	2.374	0.918	2.591	0.866	1.644
#11	0.750	0.983	2.730	0.983	5.412	0.969	6.305	0.931	3.762
	1.000	0.993	4.472	0.991	9.491	0.984	11.440	0.961	6.689
	0.000	0.980	0.931	0.989	0.967	0.986	0.931	0.972	0.799
	0.250	0.994	1.530	0.998	1.992	0.996	1.819	0.986	1.133
#19	0.500	1.000	4.697	1.000	9.412	1.002	10.063	0.997	4.189
#14	0.750	1.000	9.871	1.000	21.062	1.000	24.988	0.999	11.993
	1.000	1.000	16.721	1.000	37.273	1.000	45.047	1.000	23.052

Table D2: Simulation results with k = 8 (Cont.)

	$b_1:$	0	.2	0.	.4	0	.6	0	.8
Exp.	$\delta$	$\overline{\text{RMSFE}(\check{\alpha}_2)}$	$_{2}$ )RMSFE( $\tilde{\alpha}_{2}$	$(\breve{\alpha}_2)$ RMSFE $(\breve{\alpha}_2)$	$RMSFE(\tilde{\alpha}_{2})$	$_{2})$ RMSFE( $\breve{\alpha}_{2}$	$_{2}$ )RMSFE( $\tilde{\alpha}_{2}$	$_{2})$ RMSFE( $\breve{\alpha}_{2}$	$_{2}$ )RMSFE( $\tilde{\alpha}_{2}$ )
	0.000	0.814	0.275	0.791	0.298	0.800	0.274	0.795	0.311
	0.250	0.777	0.262	0.758	0.310	0.767	0.360	0.842	0.564
// 1	0.500	0.765	0.320	0.775	0.631	0.745	0.770	0.826	0.970
#1	0.750	0.801	0.988	0.818	2.120	0.797	2.125	0.807	1.992
	1.000	0.884	2.541	0.869	4.307	0.852	4.719	0.847	4.220
	0.000	0.813	0.312	0.791	0.299	0.811	0.270	0.808	0.302
	0.250	0.761	0.233	0.745	0.256	0.757	0.305	0.822	0.474
#2	0.500	0.749	0.378	0.756	0.623	0.746	0.814	0.839	1.008
#4	0.750	0.807	1.034	0.811	2.252	0.791	2.068	0.845	2.334
	1.000	0.874	2.550	0.872	5.151	0.860	4.918	0.861	4.383
	0.000	0.864	0.431	0.875	0.475	0.865	0.437	0.859	0.432
	0.250	0.871	0.520	0.862	0.583	0.864	0.663	0.881	0.705
#3	0.500	0.866	0.780	0.868	1.377	0.862	1.483	0.886	1.266
#3	0.750	0.905	1.709	0.924	3.724	0.881	3.783	0.887	2.667
	1.000	0.949	3.580	0.957	7.681	0.917	7.644	0.890	5.338
	0.000	0.920	0.611	0.935	0.645	0.934	0.608	0.941	0.613
	0.250	0.913	0.641	0.906	0.709	0.911	0.775	0.928	0.774
_11_ 1	0.500	0.918	0.986	0.929	1.883	0.916	2.060	0.928	1.631
#4	0.750	0.958	2.071	0.963	4.529	0.937	5.191	0.940	3.855
	1.000	0.976	3.665	0.973	8.141	0.955	9.407	0.944	6.856
	0.000	0.936	0.688	0.924	0.689	0.909	0.673	0.908	0.655
	0.250	0.938	0.797	0.930	0.999	0.918	1.030	0.919	0.923
#5	0.500	0.948	1.274	0.958	2.510	0.942	2.815	0.936	1.925
#5	0.750	0.975	2.518	0.980	5.673	0.965	6.759	0.944	4.243
	1.000	0.978	4.082	0.990	9.760	0.982	12.410	0.963	8.128
	0.000	0.909	0.761	0.956	0.988	0.970	0.934	0.969	0.742
	0.250	0.904	0.781	0.948	1.105	0.955	1.058	0.964	0.914
#6	0.500	0.885	0.961	0.951	2.005	0.945	2.077	0.960	1.725
77-0	0.750	0.953	2.105	0.953	3.839	0.946	4.101	0.956	3.247
	1.000	0.959	3.494	0.959	7.026	0.951	7.458	0.953	5.645

Table D3: Simulation results with k = 3 and non-normal errors

	$b_1:$	0	.2	0	.4	0	.6	0	.8
Exp.	$\delta$	$\overline{\text{RMSFE}(\breve{\alpha}_{2})}$	$_{2}$ )RMSFE( $\tilde{\alpha}_{2}$	) RMSFE( $\breve{\alpha}_{2}$	$_{2})$ RMSFE( $\widetilde{\alpha}_{2}$	$RMSFE(\check{\alpha}_{2})$	$_{2}$ )RMSFE( $\tilde{\alpha}_{2}$	$(\breve{\alpha}_{2})$ RMSFE $(\breve{\alpha}_{2})$	$_{2}$ )RMSFE( $\tilde{\alpha}_{2}$ )
	0.000	0.974	1.875	0.978	2.722	0.952	2.232	0.916	1.325
	0.250	0.985	2.201	0.989	3.232	0.974	3.038	0.932	1.841
// <del>-</del>	0.500	0.993	3.073	0.996	5.251	0.990	5.436	0.957	3.444
#(	0.750	0.997	4.426	0.999	8.639	0.996	9.653	0.970	6.251
	1.000	0.999	6.283	1.000	13.130	0.999	15.840	0.985	10.456
	0.000	0.815	0.274	0.778	0.201	0.799	0.184	0.784	0.171
	0.250	0.780	0.206	0.769	0.221	0.794	0.279	0.861	0.537
-448	0.500	0.784	0.308	0.768	0.569	0.776	0.703	0.872	1.107
#0	0.750	0.802	0.672	0.837	1.879	0.832	2.101	0.874	2.561
	1.000	0.924	2.454	0.931	4.833	0.907	5.015	0.895	4.843
	0.000	0.934	0.748	0.930	0.769	0.899	0.680	0.849	0.435
	0.250	0.896	0.815	0.889	1.027	0.894	1.124	0.870	0.774
-#0	0.500	0.949	2.433	0.922	4.408	0.887	3.947	0.870	1.893
#9	0.750	0.959	6.434	0.937	11.681	0.916	11.900	0.887	5.509
	1.000	0.964	13.573	0.957	23.173	0.944	22.821	0.917	12.623
	0.000	0.881	0.395	0.795	0.312	0.777	0.363	0.824	0.266
	0.250	0.855	0.341	0.864	0.367	0.889	0.366	0.824	0.325
#10	0.500	0.920	0.615	0.906	0.454	0.884	0.336	0.844	0.238
#10	0.750	0.955	0.819	0.957	0.898	0.940	0.688	0.885	0.373
	1.000	0.989	1.286	0.985	1.477	0.964	1.125	0.898	0.568
	0.000	0.878	0.316	0.847	0.203	0.876	0.263	0.833	0.252
	0.250	0.881	0.355	0.859	0.172	0.873	0.184	0.828	0.151
#11	0.500	0.913	0.372	0.931	0.280	0.905	0.221	0.889	0.160
77-11	0.750	0.942	0.607	0.978	0.707	0.962	0.495	0.927	0.300
	1.000	0.991	1.080	0.993	1.237	0.980	0.927	0.955	0.452
	0.000	0.954	0.816	0.930	0.725	0.821	0.574	0.874	0.400
	0.250	0.955	0.870	0.926	0.822	0.919	0.702	0.822	0.382
<i>#</i> 19	0.500	0.987	1.363	0.973	1.441	0.941	1.044	0.886	0.466
$\pi^{\perp 2}$	0.750	0.989	2.141	0.992	2.751	0.978	2.285	0.927	0.872
	1.000	0.996	3.267	0.994	4.372	0.989	4.165	0.950	1.566

Table D3: Simulation results with k = 3 and non-normal errors (Cont.)

	$b_1:$	0	.2	0	.4	0	.6	0	.8
Exp.	$\delta$	$\overline{\text{RMSFE}(\breve{\alpha}_2)}$	$_{2}$ )RMSFE( $\tilde{\alpha}_{2}$	$(\breve{\alpha}_2)$ RMSFE $(\breve{\alpha}_2)$	$_{2})$ RMSFE $(\widetilde{\alpha}_{2})$	$_{2})$ RMSFE( $\check{\alpha}_{2}$	$_{2}$ )RMSFE( $\tilde{\alpha}_{2}$	$_{2})$ RMSFE( $\breve{\alpha}_{2}$	$(\widetilde{\alpha}_2)$ RMSFE $(\widetilde{\alpha}_2)$
	0.000	0.807	0.415	0.782	0.363	0.775	0.354	0.753	0.331
	0.250	0.762	0.336	0.782	0.444	0.780	0.518	0.835	0.630
<i>_//_</i> 1	0.500	0.820	0.804	0.806	1.319	0.807	1.572	0.836	1.278
#1	0.750	0.863	1.727	0.918	4.104	0.875	4.177	0.857	2.825
	1.000	0.963	3.990	0.954	7.768	0.926	8.230	0.902	5.613
	0.000	0.793	0.400	0.771	0.337	0.783	0.335	0.797	0.328
	0.250	0.772	0.413	0.760	0.423	0.758	0.477	0.818	0.559
#2	0.500	0.787	0.756	0.811	1.400	0.818	1.527	0.842	1.301
$\pi^{2}$	0.750	0.880	1.916	0.916	4.044	0.893	4.493	0.877	3.139
	1.000	0.957	3.975	0.949	7.556	0.933	8.579	0.912	5.781
	0.000	0.842	0.514	0.813	0.447	0.809	0.456	0.805	0.421
	0.250	0.839	0.579	0.835	0.663	0.825	0.712	0.844	0.721
<i>4</i> 4.9	0.500	0.893	1.163	0.893	2.100	0.873	2.231	0.860	1.527
#3	0.750	0.959	2.578	0.953	5.217	0.927	5.552	0.886	3.322
#0	1.000	0.982	4.366	0.977	9.063	0.958	10.559	0.922	6.365
	0.000	0.883	0.627	0.884	0.611	0.861	0.482	0.860	0.432
	0.250	0.874	0.693	0.885	0.908	0.868	0.856	0.866	0.720
-44-4	0.500	0.922	1.339	0.930	2.353	0.915	2.699	0.899	1.829
#4	0.750	0.976	2.680	0.970	5.319	0.953	6.198	0.932	4.092
	1.000	0.983	4.222	0.980	9.053	0.969	11.061	0.946	7.313
	0.000	0.884	0.653	0.864	0.634	0.840	0.556	0.817	0.477
	0.250	0.895	0.819	0.894	1.048	0.865	1.025	0.848	0.827
#5	0.500	0.943	1.560	0.952	2.878	0.920	3.069	0.887	2.076
#0	0.750	0.979	2.847	0.980	6.055	0.960	6.943	0.924	4.576
	1.000	0.992	4.669	0.988	10.059	0.978	12.724	0.948	8.210
	0.000	0.936	0.979	0.957	1.119	0.958	0.913	0.942	0.535
	0.250	0.929	1.100	0.955	1.488	0.955	1.342	0.947	0.861
#6	0.500	0.955	1.743	0.966	2.848	0.960	2.985	0.955	1.985
#0	0.750	0.975	2.791	0.972	5.219	0.964	6.065	0.960	3.843
	1.000	0.983	4.419	0.977	8.702	0.971	10.328	0.966	6.649

Table D4: Simulation results with k = 5 and non-normal errors
	$b_1$ :	0	.2	0	.4	0	.6	0	.8
Exp.	$\delta$	$RMSFE(\breve{\alpha}_2)$	$_{2}$ )RMSFE( $\tilde{\alpha}_{2}$	) RMSFE( $\check{\alpha}_{2}$	$_{2}$ )RMSFE( $\tilde{\alpha}_{2}$	$RMSFE(\check{\alpha}_{i})$	$_{2})$ RMSFE( $\widetilde{\alpha}_{2}$	$(\breve{\alpha}_2)$ RMSFE $(\breve{\alpha}_2)$	$_{2}$ )RMSFE( $\widetilde{\alpha}_{2}$ )
	0.000	0.960	1.867	0.967	2.313	0.932	1.798	0.855	0.908
#7	0.250	0.975	2.116	0.980	2.880	0.957	2.586	0.894	1.395
	0.500	0.980	2.949	0.990	4.883	0.976	4.989	0.929	2.971
	0.750	0.995	4.276	0.995	8.032	0.987	9.064	0.956	5.518
	1.000	0.997	6.101	0.997	12.376	0.993	14.812	0.972	9.074
	0.000	0.781	0.309	0.756	0.223	0.757	0.202	0.768	0.202
<u>Ш9</u>	0.250	0.764	0.328	0.746	0.274	0.788	0.393	0.847	0.566
	0.500	0.785	0.591	0.822	1.217	0.828	1.361	0.864	1.297
#0	0.750	0.921	1.938	0.955	4.175	0.932	4.383	0.904	3.179
	1.000	0.982	3.916	0.969	7.587	0.961	8.574	0.934	5.840
	0.000	0.918	0.774	0.944	0.873	0.931	0.830	0.843	0.567
<b>#</b> 9	0.250	0.907	1.089	0.939	1.504	0.921	1.438	0.886	0.923
	0.500	0.972	3.413	0.966	6.614	0.954	7.180	0.901	2.740
	0.750	0.984	8.235	0.975	17.015	0.963	18.387	0.934	8.454
	1.000	0.987	15.435	0.983	30.283	0.976	34.073	0.953	17.278
	0.000	0.772	0.323	0.782	0.352	0.842	0.349	0.831	0.413
	0.250	0.811	0.351	0.819	0.414	0.806	0.317	0.815	0.320
#10	0.500	0.818	0.473	0.880	0.638	0.861	0.475	0.818	0.312
#10	0.750	0.948	0.972	0.955	1.395	0.933	1.188	0.872	0.583
	1.000	0.985	1.721	0.983	2.313	0.962	2.108	0.924	1.040
	0.000	0.785	0.320	0.787	0.210	0.780	0.204	0.754	0.150
	0.250	0.833	0.337	0.813	0.243	0.802	0.133	0.834	0.188
#11	0.500	0.885	0.535	0.913	0.521	0.894	0.350	0.878	0.237
#11	0.750	0.955	0.915	0.979	1.241	0.966	1.002	0.920	0.524
	1.000	0.989	1.547	0.990	2.133	0.981	1.939	0.952	0.960
	0.000	0.908	0.789	0.868	0.681	0.906	0.734	0.851	0.383
	0.250	0.934	0.919	0.912	0.848	0.934	0.906	0.758	0.356
#19	0.500	0.981	1.741	0.966	2.154	0.938	1.767	0.858	0.713
#14	0.750	0.991	3.044	0.988	4.594	0.971	3.786	0.920	1.515
	1.000	0.997	5.272	0.992	8.346	0.987	7.543	0.969	3.035

Table D4: Simulation results with k = 5 and non-normal errors (Cont.)

	$b_1:$	0.	.2	0	.4	0	.6	0	.8
Exp.	$\delta$	$\overline{\text{RMSFE}(\check{lpha}_2)}$	$_{2}$ )RMSFE( $\tilde{\alpha}_{2}$	$(\breve{\alpha}_2)$ RMSFE $(\breve{\alpha}_2)$	$_{2}$ )RMSFE( $\tilde{\alpha}_{2}$	) RMSFE( $\breve{\alpha}_{2}$	$_{2})$ RMSFE( $\tilde{\alpha}_{2}$	) RMSFE( $\breve{\alpha}_2$	$(\widetilde{\alpha}_2)$ RMSFE $(\widetilde{\alpha}_2)$
	0.000	0.745	0.374	0.700	0.328	0.721	0.342	0.715	0.317
#1	0.250	0.743	0.426	0.751	0.528	0.758	0.573	0.768	0.585
	0.500	0.812	0.986	0.867	2.003	0.842	2.124	0.812	1.297
	0.750	0.957	2.645	0.951	4.946	0.917	5.134	0.873	2.851
	1.000	0.982	4.490	0.971	8.614	0.952	9.759	0.919	5.182
	0.000	0.734	0.387	0.718	0.327	0.725	0.302	0.722	0.262
#9	0.250	0.725	0.443	0.766	0.615	0.759	0.582	0.770	0.523
	0.500	0.854	1.160	0.879	2.100	0.841	2.044	0.834	1.322
#4	0.750	0.941	2.609	0.954	5.055	0.927	5.372	0.885	2.935
	1.000	0.977	4.550	0.970	8.831	0.955	9.675	0.921	5.277
	0.000	0.754	0.457	0.731	0.374	0.734	0.393	0.706	0.313
#3	0.250	0.794	0.600	0.797	0.737	0.776	0.723	0.762	0.620
	0.500	0.876	1.315	0.915	2.527	0.876	2.554	0.821	1.469
	0.750	0.964	2.835	0.967	5.561	0.938	6.051	0.881	3.225
	1.000	0.987	4.676	0.981	9.561	0.965	10.757	0.924	5.645
	0.000	0.808	0.568	0.796	0.516	0.788	0.406	0.739	0.271
	0.250	0.832	0.736	0.856	0.993	0.831	0.912	0.796	0.585
<u></u> 1	0.500	0.905	1.506	0.940	2.751	0.901	2.798	0.860	1.603
#4	0.750	0.957	2.829	0.971	5.645	0.953	6.319	0.916	3.562
	1.000	0.990	4.721	0.983	9.552	0.971	11.067	0.937	5.997
	0.000	0.800	0.576	0.799	0.572	0.762	0.439	0.704	0.301
	0.250	0.844	0.827	0.866	1.117	0.826	1.015	0.766	0.637
45	0.500	0.935	1.694	0.948	3.016	0.917	3.157	0.843	1.726
#0	0.750	0.968	2.998	0.967	5.819	0.954	6.742	0.905	3.754
	1.000	0.990	4.785	0.985	10.028	0.973	11.887	0.938	6.457
	0.000	0.937	1.163	0.952	1.263	0.939	0.897	0.904	0.377
	0.250	0.942	1.404	0.956	1.774	0.945	1.509	0.921	0.709
46	0.500	0.969	2.143	0.967	3.322	0.958	3.366	0.941	1.722
#0	0.750	0.983	3.371	0.978	5.999	0.967	6.725	0.953	3.452
	1.000	0.989	5.052	0.981	9.685	0.975	11.118	0.962	6.078

Table D5: Simulation results with k = 8 and non-normal errors

	$b_1:$	0	.2	0	.4	0	.6	0	.8
Exp.	$\delta$	$RMSFE(\breve{\alpha}_{2})$	$_{2}$ )RMSFE( $\tilde{\alpha}_{2}$	) RMSFE( $\breve{\alpha}_2$	$_{2}$ )RMSFE( $\tilde{\alpha}_{2}$	$RMSFE(\check{\alpha}_{i})$	$_{2})$ RMSFE $(\widetilde{\alpha}_{2})$	$_{2})$ RMSFE( $\breve{\alpha}_{2}$	$_{2}$ )RMSFE( $\tilde{\alpha}_{2}$ )
	0.000	0.961	1.965	0.954	2.030	0.909	1.401	0.811	0.514
#7	0.250	0.982	2.330	0.973	2.774	0.944	2.170	0.864	0.929
	0.500	0.989	3.186	0.983	4.654	0.966	4.456	0.916	2.156
	0.750	0.993	4.489	0.989	7.808	0.979	8.263	0.944	4.167
	1.000	0.995	6.349	0.992	12.214	0.986	13.676	0.962	7.320
	0.000	0.695	0.278	0.678	0.221	0.666	0.169	0.678	0.139
Q	0.250	0.689	0.295	0.712	0.355	0.751	0.429	0.773	0.474
	0.500	0.791	0.842	0.894	1.892	0.867	1.940	0.852	1.337
#0	0.750	0.948	2.404	0.973	4.912	0.946	5.080	0.915	3.027
	1.000	0.989	4.359	0.985	8.683	0.972	9.699	0.943	5.443
	0.000	0.850	0.696	0.889	0.796	0.919	0.837	0.846	0.658
#9	0.250	0.942	1.416	0.947	1.861	0.924	1.648	0.882	0.970
	0.500	0.992	4.333	0.980	8.247	0.963	8.287	0.901	2.956
	0.750	0.995	9.503	0.987	19.112	0.976	20.419	0.945	8.408
	1.000	0.996	16.597	0.990	33.809	0.984	38.830	0.968	16.478
	0.000	0.738	0.331	0.684	0.257	0.676	0.206	0.685	0.315
	0.250	0.744	0.309	0.775	0.352	0.662	0.248	0.744	0.293
#10	0.500	0.857	0.769	0.879	0.932	0.858	0.766	0.763	0.375
#10	0.750	0.957	1.588	0.959	2.191	0.930	1.848	0.874	0.848
	1.000	0.989	2.675	0.981	3.952	0.965	3.597	0.909	1.502
	0.000	0.630	0.178	0.664	0.157	0.679	0.120	0.681	0.109
	0.250	0.628	0.139	0.721	0.224	0.722	0.155	0.738	0.138
#11	0.500	0.819	0.585	0.902	0.721	0.870	0.536	0.839	0.322
#11	0.750	0.968	1.470	0.979	1.896	0.957	1.529	0.915	0.742
	1.000	0.993	2.548	0.990	3.678	0.979	3.456	0.949	1.493
	0.000	0.894	0.820	0.838	0.732	0.920	0.835	0.847	0.654
	0.250	0.936	1.161	0.916	1.139	0.873	0.932	0.793	0.519
#19	0.500	0.978	2.505	0.961	3.311	0.938	2.803	0.873	1.088
#14	0.750	0.998	5.009	0.987	7.631	0.971	6.677	0.934	2.424
	1.000	1.002	8.752	0.994	13.868	0.984	13.495	0.942	4.594

Table D5: Simulation results with k = 8 and non-normal errors (Cont.)

	$b_1:$	0.2	0.4	0.6	0.8		0.2	0.4	0.6	0.8
Exp.	$\delta$	$\sqrt[\infty]{\pi} < 0$	$\%\tau < 0$	$\%\tau < 0$	$\%\tau < 0$	Exp.	$\sqrt[\%]{\tau} < 0$	$\%\tau < 0$	$\%\tau < 0$	$\%\tau < 0$
	0.000	6.200	3.160	4.000	8.340		42.000	26.980	6.780	0.340
	0.250	5.880	2.480	3.100	7.120		51.520	42.000	14.880	1.100
// 1	0.500	5.280	1.700	2.080	5.400		59.400	58.880	26.880	2.900
#1	0.750	4.840	1.360	1.260	3.920	#(	64.740	70.360	41.700	6.660
	1.000	4.520	1.120	0.780	2.880		68.380	77.000	54.680	12.380
	0.000	2.140	2.520	7.560	14.040		2.580	1.460	2.860	6.580
	0.250	1.760	1.320	5.080	11.580		2.520	1.100	1.960	5.200
#2	0.500	1.300	0.540	2.400	7.660	-48	2.320	0.680	1.140	3.880
#4	0.750	1.200	0.160	0.800	5.000	#0	2.200	0.500	0.660	2.920
	1.000	1.160	0.020	0.380	3.500		2.160	0.300	0.560	1.960
	0.000	10.120	4.980	1.780	1.960		6.020	2.260	2.120	6.880
#3	0.250	10.740	5.500	2.080	1.680	#9	5.540	1.780	1.320	4.380
	0.500	11.100	6.020	2.120	1.240		5.080	1.500	0.880	2.860
	0.750	11.320	5.840	2.340	1.160		4.660	1.220	0.720	1.900
	1.000	11.540	5.440	2.500	1.080		4.420	0.840	0.440	1.360
	0.000	0.040	0.060	0.900	6.460		3.700	4.640	5.300	6.480
	0.250	0.040	0.000	0.160	3.400	#10	3.220	4.160	4.820	5.940
-#-1	0.500	0.040	0.000	0.060	1.780		3.160	4.120	4.780	5.800
#4	0.750	0.040	0.000	0.020	1.080		3.300	4.200	4.820	5.360
	1.000	0.040	0.000	0.000	0.440		3.340	4.600	5.420	5.540
	0.000	5.420	0.760	0.020	0.000		1.960	2.280	3.320	4.200
	0.250	6.880	1.160	0.040	0.000		1.600	2.180	2.800	3.880
#5	0.500	8.020	1.560	0.120	0.000	#11	1.520	1.960	2.720	3.500
#0	0.750	8.940	2.060	0.320	0.000	#11	1.860	2.000	2.880	3.240
	1.000	9.980	2.680	0.600	0.020		2.120	2.280	3.200	3.240
	0.000	0.340	8.300	32.940	45.900		2.640	3.560	4.160	5.340
	0.250	0.040	2.240	16.900	34.940		2.180	2.960	3.440	4.540
#6	0.500	0.000	0.380	5.880	21.880	#12	2.260	2.800	3.320	3.820
π.ο	0.750	0.000	0.080	1.800	12.300	TT 14	2.300	3.000	3.240	3.740
	1.000	0.000	0.000	0.360	6.900		2.380	3.280	3.740	3.620

Table D6: Percentage of  $\tau < 0$  in Simulation results with k = 3

	$b_1:$	0.2	0.4	0.6	0.8		0.2	0.4	0.6	0.8
Exp.	$\delta$	$\sqrt[\infty]{\pi} < 0$	$\%\tau < 0$	$\%\tau < 0$	$\%\tau < 0$	Exp.	$\%\tau < 0$	$\%\tau < 0$	$\%\tau < 0$	$\%\tau < 0$
	0.000	3.320	0.800	1.260	5.260		12.060	3.880	0.420	0.040
	0.250	3.260	0.560	0.700	3.860		15.560	7.340	1.120	0.060
// 1	0.500	3.200	0.460	0.320	2.540	117	19.060	12.820	2.700	0.180
<i></i> #1	0.750	3.300	0.440	0.200	1.680	<i>#</i> (	22.720	18.720	5.800	0.360
	1.000	3.220	0.380	0.160	1.180		25.600	26.020	11.140	0.940
	0.000	0.660	0.420	2.560	8.080		0.560	0.220	0.880	4.000
	0.250	0.500	0.100	1.180	5.440		0.620	0.060	0.200	1.720
#2	0.750	0.520	0.000	0.300	2.120	-#8	0.600	0.060	0.080	1.080
#4	1.000	0.540	0.000	0.160	1.280	#0	0.620	0.040	0.060	0.720
	0.000	5.260	1.760	0.420	1.000		3.240	0.660	0.320	2.400
#3	0.250	5.380	1.800	0.420	0.700	#9	3.340	0.560	0.220	1.220
	0.500	5.640	1.940	0.420	0.400		3.360	0.440	0.120	0.660
	0.750	5.820	1.980	0.480	0.260		3.320	0.400	0.060	0.300
	1.000	5.980	2.080	0.500	0.160		3.240	0.380	0.080	0.180
	0.000	0.000	0.000	0.040	0.820		1.600	2.240	3.080	3.800
// 4	0.250	0.000	0.000	0.020	0.280	#10	1.300	1.660	2.680	3.300
	0.500	0.000	0.000	0.020	0.060		1.180	1.680	2.420	2.960
#4	0.750	0.000	0.000	0.000	0.020		1.060	1.540	2.180	2.620
	1.000	0.000	0.000	0.000	0.020		1.020	1.680	2.140	2.640
	0.000	0.580	0.040	0.000	0.000		0.700	0.980	1.620	2.320
	0.250	0.680	0.020	0.000	0.000		0.520	0.720	1.340	2.080
#5	0.500	0.740	0.020	0.000	0.000	-#11	0.440	0.620	1.080	1.660
#0	0.750	0.780	0.020	0.000	0.000	#11	0.380	0.680	0.940	1.320
	1.000	0.860	0.020	0.000	0.000		0.320	0.780	0.940	1.300
	0.000	0.020	0.540	5.340	16.120		0.980	1.220	1.940	2.700
	0.250	0.000	0.120	1.600	9.080		0.880	0.840	1.580	2.140
#6	0.500	0.000	0.000	0.480	4.820	#19	0.760	0.880	1.220	1.700
#0	0.750	0.000	0.000	0.120	2.220	#14	0.720	1.020	1.260	1.740
	1.000	0.000	0.000	0.040	1.140		0.780	1.060	1.300	1.520

Table D7: Percentage of  $\tau < 0$  in Simulation results with k=5

	$b_1$ :	0.2	0.4	0.6	0.8		0.2	0.4	0.6	0.8
Exp.	$\delta$	$\sqrt[\%]{\tau} < 0$	$\%\tau < 0$	$\%\tau < 0$	$\%\tau < 0$	Exp.	$\sqrt[\%]{\tau} < 0$	$\%\tau < 0$	$\%\tau < 0$	$\%\tau < 0$
	0.000	2.140	0.340	0.340	4.260		4.060	0.320	0.000	0.000
	0.250	2.100	0.300	0.140	2.600		4.580	0.480	0.000	0.000
// 1	0.500	2.120	0.260	0.040	1.820		5.180	0.880	0.020	0.000
#1	0.750	2.160	0.260	0.020	1.020	#(	5.520	1.320	0.080	0.000
	1.000	2.200	0.280	0.000	0.680		6.000	1.880	0.180	0.000
	0.000	0.520	0.100	0.740	5.660		0.160	0.020	0.260	3.160
#2	0.250	0.520	0.020	0.320	3.680		0.180	0.020	0.080	1.880
	0.500	0.400	0.020	0.100	2.000	#8	0.180	0.020	0.020	1.160
#4	0.750	0.380	0.000	0.080	1.360	#0	0.180	0.020	0.020	0.720
	1.000	0.440	0.000	0.020	0.800		0.160	0.020	0.000	0.400
	0.000	3.200	0.560	0.020	0.740		2.120	0.360	0.040	1.040
#3	0.250	3.400	0.640	0.040	0.360	#9	2.100	0.280	0.020	0.560
	0.500	3.460	0.680	0.040	0.160		2.140	0.280	0.000	0.240
	0.750	3.600	0.660	0.040	0.080		2.120	0.260	0.000	0.080
	1.000	3.700	0.700	0.040	0.040		2.160	0.280	0.000	0.000
	0.000	0.000	0.000	0.000	0.220		0.740	1.100	1.500	2.680
	0.250	0.000	0.000	0.000	0.060	#10	0.600	0.820	1.180	2.440
11 4	0.500	0.000	0.000	0.000	0.020		0.540	0.640	0.840	2.040
#4	0.750	0.000	0.000	0.000	0.000		0.560	0.680	0.840	1.760
	1.000	0.000	0.000	0.000	0.000		0.600	0.660	0.800	1.580
	0.000	0.180	0.000	0.000	0.000		0.220	0.280	0.540	1.680
	0.250	0.180	0.000	0.000	0.000		0.120	0.180	0.260	1.200
#5	0.500	0.180	0.000	0.000	0.000	-#11	0.140	0.240	0.220	0.860
#0	0.750	0.180	0.000	0.000	0.000	#11	0.160	0.200	0.160	0.640
	1.000	0.180	0.000	0.000	0.000		0.140	0.160	0.180	0.600
	0.000	0.000	0.000	0.480	5.660		0.420	0.540	0.780	1.780
	0.250	0.000	0.000	0.140	2.840		0.420	0.440	0.460	1.320
#6	0.500	0.000	0.000	0.040	1.360	#19	0.440	0.420	0.360	0.880
#0	0.750	0.000	0.000	0.000	0.700	#14	0.500	0.440	0.400	0.700
	1.000	0.000	0.000	0.000	0.280		0.540	0.480	0.500	0.700

Table D8: Percentage of  $\tau < 0$  in Simulation results with k=8

## References

- Bao, Y. and Ullah, A. (2010). "Expectation of quadratic forms in normal and nonnormal variables with applications." *Journal of Statistical Planning and Inference* 140, 1193-1205.
- Magnus, J.R. (1978). "The moments of products of quadratic forms in normal variables." Statistica Neerlandica, 132, 201–210.
- Ullah, A., 2004. Finite Sample Econometrics. Oxford: Oxford University Press.