Higher-order Accuracy of Asymptotic F and t-tests for Over-identified GMM in Time Series*

Jungbin Hwang
Department of Economics,
University of Connecticut

Gonzalo Valdés
Departamento de Ingeniería Industrial y de Sistemas,
Universidad de Tarapacá

Dec, 2023

Abstract

This paper's first contribution is to develop a new higher-order expansion of test statistics for heteroskedasticity autocorrelation robust (HAR) inference in an over-identified two-step GMM framework. Our novel higher-order expansion rigorously shows the asymptotic F and t-tests (Hwang and Sun, 2017; Hwang and Valdés, 2022a), driven under the alternative fixed-smoothing asymptotics, are second-order correct under the conventional increasing-smoothing asymptotics. Building upon the new higher-order expansion of the GMM tests, this paper also provides a testing-oriented optimal choice of smoothing parameters that directly accounts for the degree of overidentification in HAR inferential problem. Our Monte Carlo simulations provide numerical evidences that are consistent with the theoretical results developed in this paper.

JEL Classification: C12, C13, C32

Keywords: Asymptotics, t and F tests, Fixed-smoothing asymptotics, Higher-order expansion, Genearlized Method of Moments, Heteroskedasticity Autocorrelated Robust

*Email: jungbin.hwang@uconn.edu and gvaldes@uta.cl.

For helpful comments, we thank the participants of the Advances in Econometrics (AiE) Conference and the Festschrift in honor of Joon Y. Park at Indiana University, September 2023. Gonzalo Valdés acknowledges financial support from ANID Fondecyt Iniciación 11220125.

1 Introduction

Generalized method of moments (GMM) is one of the most widely used statistical methods for the analysis of economic data and financial data (Hall, 2005). In time-series data, with an unknown form of dependence, an important initial contribution is made by Newey and West (1987) to implement the feasible two-step efficient estimation of GMM proposed by Hansen (1982). The key step for the efficient GMM with over-identification procedure is an estimation of the optimal weighting matrix, which is referred to as the long-run variance (LRV) of the moment conditions. Newey and West (1987) and Andrews (1994) propose LRV using the nonparametric kernel method.

However, it is well known that the two-step GMM procedure shows poor performance in a finite sample context. See, for example, the special issue in the Journal of Business & Economic Statistics (Christiano, 1996). Making an accurate inference of the GMM estimator in the time series setting becomes more challenging due to the considerable estimation uncertainty embodied in the nonparametric LRV estimator. The standard first-order normal approximation completely ignores the estimation uncertainty in LRV and does not reflect the finite sample situations properly. Phillips (2005) proposes a new type of LRV estimators using orthonormal series basis functions. Müller (2007) and Sun (2011, 2013) develop its more accurate first-order approximation, so called fixed-smoothing asymptotics, and show its superior finite-sample performances.

In the two-step GMM with over-identification, Sun (2014) and Hwang and Sun (2017) develop fixed-smoothing asymptotics and improve performance relative to the normal approximation. To be specific, Hwang and Sun (2017) propose a modification to the conventional GMM Wald test statistics, a function of the original test statistic and the usual J statistic for testing over-identificaitons. The modification leads to more accurate standard F and t-tests than the conventional GMM Wald testing with chi-square and normal critical values. A recent work by Hwang and Valdés (2022a) provide further finite-sample improvement by correcting the bias of the two-step GMM variance estimator.

The literature, however, show the finite-sample improvements using the fixed-smoothing asymptotics by Monte-Carlo simulations. To the best of our knowledge, there is no work in time-series GMM that rigorously prove the higher-order accuracy of the fixed-smoothing asymptotics that achieve more accurate finite sample null rejection in a precise technical sense. In this paper, we fill the technical gap in the literature and develop a higher-oder expansion of t-statistic of the two-step GMM. In addition to the higher-order expansion, we provide an optimal choice of smoothing parameters which directly focuses on heteroskedasticity autocorrelation robust (HAR) inferential problem in an over-identified GMM setting.

The higher-order properties of GMM estimator has been an active research area in econometrics. Hansen, Heaton, and Yaron (1996) first suggest continuously updating (CU) GMM to improve finite-sample performance of the two-step GMM. Its generalization to generalized empirical likelihood (GEL) method has been done in Smith (1997), Imbens, Spady, and Johnson (1998), Kitamura and Stutzer (1997), and Newey and Smith (2004). Importantly, Newey and Smith (2004) and Anatolyev (2005) prove improved higher-order bias and variance properties of GEL estimators in i.i.d setting and time-series setting, respectively. Still, the poor finite-sample performances of GEL remain issues, as pointed by extensive numerical work in Guggenberger and Hahn (2005) and Guggenberger (2008). Moreover, the results in Newey and Smith (2004) and Anatolyev (2005) do not lead to key inferential properties of GMM estimators, as they do not consider a higher-order expansion of finite sample null rejection probability of t-statistic.

In the context of bootstrapping GMM, Hall and Horowitz (1996) first prove the higher-order property of GMM inference via an Edgeworth expansion of the t-statistic, but they assume a less general finite dependence structure in time series. Inoue and Shintani (2005) relaxes a finite dependence structure in Hall and Horowitz (1996), but the proposed moving-block bootstrap method requires an additional choice of block length in practice. Sun and Phillips (2009) also provide an Edgeworth expansion in time-series GMM regression, but their approach is not applicable to the HAR inference with the OS-LRV considered in our paper.

Recently, a new type of inferential methods in time-series, so called HAR inference, develop the alternative first-order fixed-smoothing asymptotics (Phillips,2005; Müller, 2007; Sun et al.,2008). See also Lazarus et al. (2018) for its practical implementation. Sun et al. (2008), Sun (2011, 2013, 2014), and Zhang and Shao (2013) rigorously prove the higher-order accuracy of the alternative fixed-smoothing asymptotics in HAR inference. All of these papers, however, focus on the one-step GMM estimator, or exactly identified GMM model.

In this paper, we develop a higher-order expansion of the t-statistic considering the two-step GMM estimation in HAR inference. Technically, it is more challenging to account for the estimation uncertainty of the optimal weighting matrix in the two-step GMM framework. To derive the higher-order expansion, we follow a technique developed in a recent work by Hwang and Valdes (2022b). The analysis first decomposes the original t-statistic into an infeasible t-statistic and other error terms that do not involve the non-parametric estimation of the optimal weighting matrix. We then transform the higher-order approximation of the infeasible t-statistic into an inference in Gaussian over-identified location model. In the over-identified location model, we introduce a novel GLS-two-step t-statistic and show its asymptotic equivalence to the infeasible t-statistic. Using these findings, we are able to expand the finite-sample CDF of the original t-statistic around CDF of the chi-square distribution with one degree of freedom and characterize the leading higher-order terms.

Based on the higher-order expansion, we provide an optimal choice of smoothing parameters for HAR inference which directly focuses on the inferential problem in the over-identified GMM problem. Our formula generalizes what is previously suggested in Sun (2013) for the exact identified GMM, as it converges to the formula in Sun (2013) when there is no over-identification.

This paper generalizes a higher-order expansion in HAR inference pioneered in Sun, Phillips, and Jin (2008) and Sun (2011, 2013) to the over-identified two-step GMM framework. Our theoretical results show that the fixed-smoothing asymptotics in two-step GMM provides a higher-order refinement of the conventional GMM t-test using standard normal critical value. Our result also provides a new formula of optimal smoothing parameter in HAR inference in GMM, which directly focuses on inferential problems.

The rest of the paper is organized as follows. Section 2 describes a general two-step GMM in a time-series setting and derives its asymptotic property. Section **??** establishes a higher-order expansion for the test statistics in Gaussian overidentified location model. Section 4 consider selecting K to minimize the coverage probability error on the basis of this expansion in previous sections. Section 5 presents Monte Carlo simulation results, and Section 6 concludes. Proofs are presented in the Appendix.

2 Asymptotic F and t tests in Over-identified System

This paper considers the following over-identified GMM model in time-series and HAR inference, e.g., Sun (2014), Hwang and Sun (2017), and Hwang and Valdés (2020). We want to estimate a $d \times 1$ vector of parameter $\theta \in \Theta$ using a vector of observation $w_t \in \mathbb{R}^{d_w}$ at time t. The true parameter θ_0 is assumed to be an interior point of Θ. The moment condition is given as

$$Ef(w_t, \theta) = 0 \text{ iff } \theta = \theta_0,$$

where $f(w_t, \theta)$ is an $m \times 1$ vector of moment process which is linear in parameter vector θ. Leading example is a linear IV regression model in time series with $f(w_t, \theta) = z_t(y_t - x_t'\theta)$, where $w_t = (y_t, x_t', z_t')' \in \mathbb{R}^{d_y+d_x+d_z}$ consists of the observed data, e.g., Inoue and Shintani (2006) and Sun and Phillips (2009). With the true parameter $\theta = \theta_0$, the process $f(w_t, \theta_0)$ is stationary with zero mean. We allow $f(w_t, \theta_0)$ to have general autocorrelation of unknown forms and satisfy $\sum_{j=-\infty}^{\infty} \|Ef(w_t, \theta_0)f(w_{t-j}, \theta_0)'\| < \infty$ and some mixing conditions for the time-series CLT as follows

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T} f(w_t, \theta_0) \xrightarrow{d} N(0, \Sigma), \tag{1}$$

where $N(0, \Sigma)$ is a m-dimensional normal random vector. The $m \times m$ matrix $\Sigma = \Lambda\Lambda'$ is a positive definite long-run variance (LRV) of the moment process $f(w_t, \theta_0)$, i.e. $\Sigma = \sum_{j=-\infty}^{\infty} Ef(w_t, \theta_0)f(w_{t-j}, \theta_0)$. Also, we assume that $q = m - d > 0$, and the full rank of $G = G(\theta_0) = E[\partial f(w_t, \theta_0)/\partial\theta']$. Thus, the model is overidentified with a degree of over-identificaion q. Let $f_T(\theta) = T^{-1} \sum_{s=1}^{T} f(w_s, \theta)$ and define a one-step GMM estimator as

$$\hat{\theta}_1 = \arg\min_{\theta \in \Theta} f_T(\theta)' W_T^{-1} f_T(\theta),$$

where W_T is an initial weight matrix. Using this one-step estimator, the feasible efficient two-step GMM estimator in Hansen(1982) is defined as

$$\hat{\theta}_2 = \arg\min_{\theta \in \Theta} f_T(\theta)' \left[\Sigma_T(\hat{\theta}_1)\right]^{-1} f_T(\theta),$$

where

$$\Sigma_T(\hat{\theta}_1) = \frac{1}{K} \sum_{j=1}^{K} \left(\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \phi_j(\frac{t}{T}) f(w_t, \theta) \right) \left(\frac{1}{\sqrt{T}} \sum_{s=1}^{T} \phi_j(\frac{s}{T}) f(w_s, \theta) \right)',$$

and $\left\{\phi_j(r)\right\}_{j=1}^{K}$ are orthonormal basis functions on $L^2[0,1]$ satisfying $\int_0^1 \phi_j(r)\, dr = 0$. We parametrize the number of series terms, K, in such a way that it indicates the level of smoothing for both types of LRV estimators. By construction, $\Sigma_T(\hat{\theta}_1)$ is a quadratic heteroskedasticity autocorrelation robust (HAR) estimator for the LRV Ω. It is important to note that $\Sigma_T(\hat{\theta}_1)$ is a "centered" version of the LRV estimator, as it is based on the estimation of the demeaned moment process $f(w_t, \hat{\theta}_1) - f_T(\hat{\theta}_1)$. Under the conventional asymptotic theory, Hall (2000) shows that the demeaning procedure can potentially improve the power performance of the J-test using the HAR estimator. Also, this demeaned procedure plays an important role in fixed smoothing asymptotics as the random matrix limit of $\Sigma_T(\theta_0)$ is independent of the limiting distribution of $\sqrt{T} f_T(\theta_0)$ which is a normal distribution.

With $d = 1$, we define the following GMM t statistic for testing $H_0 : \theta = \theta_0$:

$$\mathbb{T}_T := \mathbb{T}_T(\hat{\theta}_2; \hat{\theta}_1) = \frac{\sqrt{T}(\hat{\theta}_2 - \theta_0)}{\sqrt{[G_T'(\hat{\theta}_1)\Sigma_T^{-1}(\hat{\theta}_1)G_T(\hat{\theta}_1)]^{-1}}},$$

and consider the following modified t and Wald statistics (Hwang and Sun (2017), Hwang and Valdés (2022a)):

$$\tilde{\mathbb{T}}_T := \tilde{\mathbb{T}}_T(\hat{\theta}_2; \hat{\theta}_1) = \sqrt{\frac{K-q}{K}} \cdot \frac{\mathbb{T}_T}{\sqrt{1 + \frac{1}{K}\mathbb{J}_T(\hat{\theta}_2; \hat{\theta}_1)}},$$

where $\mathbb{J}_T(\hat{\theta}_2; \hat{\theta}_1) = T f_T(\hat{\theta}_2)' \Sigma_T^{-1}(\hat{\theta}_1) f_T(\hat{\theta}_2)$ is the standard J statistic for testing the over-identifying restrictions. Similarly, the Wald test-statistics, $\mathbb{F}_T$ and $\tilde{\mathbb{F}}_T$, for the case of multiple testing parameters, i.e., $d \geq 2$, are defined as

$$\tilde{\mathbb{F}}_T := \tilde{\mathbb{F}}_T(\hat{\theta}_2; \hat{\theta}_1) = \left(\frac{K-q}{K}\right) \cdot \frac{\mathbb{F}_T}{1 + \frac{1}{K}\mathbb{J}_T(\hat{\theta}_2; \hat{\theta}_1)},$$

where $\mathbb{F}_T(\hat{\theta}_2; \hat{\theta}_1) = T(\hat{\theta}_2 - \theta_0)'[G_T'(\hat{\theta}_1)\Sigma_T^{-1}(\hat{\theta}_1) G_T(\hat{\theta}_1)]^{-1}(\hat{\theta}_2 - \theta_0)$.

For a fixed K as $T \to \infty$, Hwang and Sun (2017) show that

$$\tilde{\mathbb{T}}_T \xrightarrow{d} \mathcal{T}_{K-q} \text{ and } \tilde{\mathbb{F}}_T \xrightarrow{d} \mathcal{F}_{K-d-q-1}$$

The first contribution of this paper is to develop a higher-order expansion of the finite-sample probability of the GMM t-statistics

$$P(\tilde{\mathbb{T}}_T(\hat{\theta}_2; \hat{\theta}_1) \leq z),$$

as $T \to \infty$ such that $K/T \to 0$. The main object is the two-step GMM t-statistic, $\tilde{\mathbb{T}}_T(\hat{\theta}_2; \hat{\theta}_1)$, that depends on both first-step and two-step estimators with non-linear interactions, especially through $\Sigma_T^{-1}(\hat{\theta}_1)$. It is technically challenging to obtain a valid Edgeworth expansion for $\tilde{\mathbb{T}}_T(\hat{\theta}_2; \hat{\theta}_1)$. Sun and Phillips (2009) prove the same object with a linear IV-regression setting with kernel-LRV estimator for $\Sigma_T(\hat{\theta}_1)$. Sun (2013, 2014) develops the higher-order expansion of general GMM model, but they are only applicable to the first-step GMM estimator $\hat{\theta}_1$.To confront these technical challenges and develop a high order expansion in the over-identified GMM setting, we begin by the following stochastic approximation in Lemma **1**. We first introduce some high level assumptions.

Assumption 1 *$\hat{\theta}_1 = \theta_0 + O_p(T^{-1/2})$ and $G_T(\hat{\theta}_1) = G + O_p(T^{-1/2})$, where the terms of order $O_p(T^{-1/2})$ does not depend on K.*

Assumption 2 *As $K \to \infty$ and $T \to \infty$ such that $K/T \to 0$, we have that*

$$\Sigma_T(\theta_0) = \Sigma + O_p\left(\frac{K^2}{T^2} + \frac{1}{\sqrt{K}}\right); \quad (2)$$

$$\Upsilon_T(\theta_0) + \Upsilon_T'(\theta_0) = \Upsilon + \Upsilon' + O_p\left(\frac{K^2}{T^2} + \frac{1}{\sqrt{K}}\right), \quad (3)$$

where Σ is positive definite, and

$$\Upsilon_T(\theta) = \frac{1}{K}\sum_{j=1}^{K}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\phi_j\left(\frac{t}{T}\right)\frac{\partial f(w_t, \theta)}{\partial \theta}\right)\left(\frac{1}{\sqrt{T}}\sum_{s=1}^{T}\phi_j\left(\frac{s}{T}\right)f(w_s, \theta)\right)'.$$

Assumption 1 is made for convenience, and primitive sufficient conditions are available from any standard GMM asymptotic theory in time series, e.g. Hall (2005) and Sun (2013, 2014). One of the key requirements for the first condition is that the initial GMM weight matrix W_T do not depend on the unknown parameter value θ_0 and it converges to W with parametric rate of convergence of order $O_p(T^{-1/2})$. Assumption 2 is made for convenience, and sufficient conditions for (2) and (3) can be referred to HAR literature, as in Phillips (2005) and Sun (2013). They study the asymptotic results for the OS-LRV estimator assuming large $K \to \infty$ and $T \to \infty$ such that $K/T \to 0$.

Lemma 1 *Under Assumptions (1)–(2), we have that*

$$\mathbb{T}_T(\hat{\theta}_2; \hat{\theta}_1) = \mathcal{T}_T(\theta_0) + \psi_T + \psi_T^*,$$

where

$$\begin{aligned}\mathcal{T}_T(\theta_0) &= -\frac{\left[G'\Sigma_T^{-1}(\theta_0)G\right]^{-1} G'\Sigma_T^{-1}(\theta_0)\sqrt{T} f_T(\theta_0)}{\sqrt{[G'\Sigma_T^{-1}(\theta_0)G]^{-1}}}; \\ \psi_T &= O_p\left(\frac{1}{\sqrt{T}}\right); \\ \psi_T^* &= O_p\left(\frac{1}{\sqrt{T}}\left(\frac{1}{\sqrt{K}} + \frac{K^2}{T^2} + \frac{1}{\sqrt{T}}\right)\right),\end{aligned}$$

and ψ_T does not depend on K.

Assumption 3 *We have that $P(|\psi_T| > (\log T)/\sqrt{T}) = O(1/\sqrt{T})$ and $P\left(|\psi_T^*| > \delta_T/\log T\right) = o(\delta_T)$ for $\delta_T = 1/K + (K/T)^2$.*

Lemma 2 *Under Assumptions 1–3, and if $K \to \infty$ such that $K/T \to 0$, then we have that*

$$P\left(\mathbb{T}_T(\hat{\theta}_2; \hat{\theta}_1) \le z\right) = P\left(\mathcal{T}_T(\theta_0) \le z\right) + O\left(\frac{\log T}{\sqrt{T}}\right) + o\left(\frac{K^2}{T^2}\right) + o\left(\frac{1}{K}\right). \quad (4)$$

The result in Lemma 2 indicates that we can decompose the original t-statistic, $\mathbb{T}_T(\hat{\theta}_2; \hat{\theta}_1)$, into an infeasible one, $\mathcal{T}_T(\theta_0)$, and other error terms that do not involve the non-parametric estimation of the optimal weighting matrix. A recent work by Hwang and Valdés (2022a) develops a finite-sample corrected version of $\mathbb{T}_T(\hat{\theta}_2; \hat{\theta}_1)$ to correct the $O(\log T/\sqrt{T})$ error in (4). Still, it is challenging to fully account for the estimation uncertainty of the optimal weighting matrix in $\mathcal{T}_T(\theta_0)$. To confront this challenge, we follow a technique developed in Hwang and Sun (2018) and Hwang and Valdés (2022b) that transforms $\mathcal{T}_T(\theta_0)$ into an inference in Gaussian over-identified location model. To be more specific, let $G = \mathbb{U} \cdot \Psi \cdot \mathbb{V}'$ be a singular value decomposition (SVD) of G, where $\Psi' = (A, O) \in \mathbb{R}$, and A is a $d \times d$ diagonal matrix and O is a matrix of zeros. Also, we define the rotated moment conditions and the rotated long run variance

$$\begin{aligned}f^*(w_t, \theta_0) \quad &: \quad = \mathbb{U}'f(w_t, \theta_0) = (f_1^{*\prime}(w_t, \theta_0), f_2^{*\prime}(w_t, \theta_0))'; \quad (5) \\ \Omega \quad &= \quad \mathbb{U}'\Sigma\mathbb{U} = \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix},\end{aligned}$$

respectively, where $f_1^*(w_t,\theta_0) \in \mathbb{R}^d$ and $f_2^*(w_t,\theta_0) \in \mathbb{R}^q$. The corresponding long run variance estimator with the transformed moment condition is then naturally defined as

$$\begin{aligned}
\hat{\Omega} &= \mathbb{U}'\Sigma_T(\theta_0)\mathbb{U} \\
&= \frac{1}{K}\sum_{j=1}^{K}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\phi_j(\frac{t}{T})f^*(w_t,\theta_0)\right)\left(\frac{1}{\sqrt{T}}\sum_{s=1}^{T}\phi_j(\frac{s}{T})f^*(w_s,\theta)\right)' \\
&= \begin{pmatrix} \hat{\Omega}_{11} & \hat{\Omega}_{12} \\ \hat{\Omega}_{21} & \hat{\Omega}_{22} \end{pmatrix}.
\end{aligned}$$

The following theorem provides a useful expansion that relates the infeasible statistic, $\mathcal{T}_T(\theta_0)$, to the rotated moment conditions in (5).

Lemma 3 *The following results hold:*
(a) $-T^{-1/2}\sum_{t=1}^{T}\left[G'\Sigma_T^{-1}(\theta_0)G\right]^{-1}G'\Sigma_T^{-1}(\theta_0)f(w_t,\theta_0) = \frac{1}{\sqrt{T}}\mathbb{V}A^{-1}\sum_{t=1}^{T}\left[f_1^*(w_t,\theta_0) - \hat{\Omega}_{12}[\hat{\Omega}_{22}]^{-1}f_2^*(w_t,\theta_0)\right]$
(b) $G'\Sigma_T^{-1}(\theta_0)G = A^{-1}\mathbb{V}\hat{\Omega}_{11\cdot 2}A^{-1}\mathbb{V}$.

Let us define the following rotated moment process,

$$s_{1t} = f_1^*(w_t,\theta_0) \text{ and } s_{2t} = f_2^*(w_t,\theta_0),$$

and $\bar{s}_{1T} = T^{-1}\sum_{t=1}^{T}s_{1t}$ and $\bar{s}_{2T} = T^{-1}\sum_{t=1}^{T}s_{2t}$. Then, the results in Lemma 3 indicates that

$$-\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[G'\Sigma_T^{-1}(\theta_0)G\right]^{-1}G'\Sigma_T^{-1}(\theta_0)f(w_t,\theta_0) = \frac{1}{\sqrt{T}}\mathbb{V}A^{-1}\sum_{t=1}^{T}\left(s_{1t} - \hat{\Omega}_{12}[\hat{\Omega}_{22}]^{-1}s_{2t}\right),$$

and, with $d = 1$, we have that

$$\mathcal{T}_T(\theta_0) = \frac{\sqrt{T}\left(\bar{s}_{1T} - \hat{\Omega}_{12}[\hat{\Omega}_{22}]^{-1}\bar{s}_{2T}\right)}{\sqrt{\hat{\Omega}_{11\cdot 2}}}.$$

3 Higher-order Expansion of Two-step GMM-statistics

In this section, we follow Hwang and Valdés (2022b) and approach the finite-sample distribution of $\mathcal{T}_T(\theta_0)$ by considering the following over-identified location problem for testing H_0^*: $\theta_0^* = 0$ as below:

$$s_{1t} = \theta_0^* + v_{1t}; \tag{6}$$
$$s_{2t} = v_{2t} \tag{7}$$

for $t = 1,\ldots,T$, where $\theta_0^* = A\mathbb{V}'\theta_0$ and $\Omega = \mathbb{U}'\Sigma\mathbb{U}$. In the context of the over-identified location setting in (6) and (7), the one-step and the two-step GMM estimators are

$$\begin{aligned}
\hat{\theta}_1^* &= \bar{s}_{1T}; \\
\hat{\theta}_2^* &= T^{-1}\sum\left(s_{1t} - \hat{\Omega}_{12}\hat{\Omega}_{22}^{-1}s_{2t}\right) \\
&= \bar{s}_{1T} - \hat{\Omega}_{12}\hat{\Omega}_{22}^{-1}\bar{s}_{2T},
\end{aligned}$$

For simplicity, we assume that $d = 1$. Under $H_0 : \theta_0^* = 0$, the goal is equivalent to provide a finite-sample probability of the following two-step GMM statistics:

$$\frac{\sqrt{T}\left(\hat{\theta}_2^* - \theta_0^*\right)}{\sqrt{\hat{\Omega}_{11\cdot 2}}} = \frac{\sqrt{T}\left(\bar{s}_{1T} - \hat{\Omega}_{12}[\hat{\Omega}_{22}]^{-1}\bar{s}_{2T}\right)}{\sqrt{\hat{\Omega}_{11\cdot 2}}},$$

which is equivalent to the infeasible t-statistic $\mathcal{T}_T(\theta_0)$ at the original problem.

Assumption 4 (Linear Process) *Let $v_t = (v_{1t}', v_{2t}')' \in \mathbb{R}^{d+q}$ be a zero mean Gaussian linear process, i.e.,*

$$v_t = C(L)\epsilon_t = \sum_{j=0}^{\infty} C_j \epsilon_{t-j},$$

where $\epsilon_t = (\epsilon_{1t}, \epsilon_{2t}')' \overset{i.i.d}{\sim} N(0, I_{q+1})$, and C_j's are $(q+1) \times (q+1)$ matrix such that $E||\epsilon_t||^{\nu} < \infty$ for some $\nu \geq 4$, $\sum_{j=0}^{\infty} j^a ||C_j|| < \infty$ for $a > 3$, and $\Omega = C(1)C(1)' > 0$, where $C(1) = \sum_{j=0}^{\infty} C_j$.

Lemma 4 *Under Assumptions 1–4, we have that*

$$P(\tilde{\mathbb{T}}_T(\hat{\theta}_2; \hat{\theta}_1)^2 \leq \kappa \cdot z) = E[G_1\left(\tilde{\Xi}^{-1}\right)] + O\left(\frac{\log T}{\sqrt{T}}\right) + o\left(\frac{K^2}{T^2}\right) + o\left(\frac{1}{K}\right),$$

where $\kappa = ((K-q)/K)^{-1/2}$, $G_p(\cdot)$ indicates the cdf of χ_p^2 random variable, and $\tilde{\Xi}(\hat{\Omega})$ is defined to be

$$\tilde{\Xi}(\hat{\Omega}) = \Xi(\hat{\Omega}) \cdot \left(1 + \left(\frac{\chi_q^2}{K}\right) \cdot \left(\frac{Z_q}{\|Z_q\|}\right)' \Omega_{22}^{1/2\prime} \hat{\Omega}_{22}^{-1} \Omega_{22}^{1/2} \left(\frac{Z_q}{\|Z_q\|}\right)\right)^{-1}$$

for $Z_q \sim N(0, I_q)$ which is independent of χ_q^2, and

$$\Xi(\hat{\Omega}) = \left(1 + \frac{\left(\Omega_{12}\Omega_{22}^{-1} - \hat{\Omega}_{12}\hat{\Omega}_{22}^{-1}\right)\Omega_{22}\left(\Omega_{12}\Omega_{22}^{-1} - \hat{\Omega}_{12}\hat{\Omega}_{22}^{-1}\right)'}{\Omega_{11\cdot 2}}\right) \times \left(\frac{\Omega_{11\cdot 2}}{\hat{\Omega}_{11\cdot 2}}\right).$$

Define $\Gamma_u(h) = Ev_t v_{t-h}'$, and $\Gamma_{11,u}(h)$, $\Gamma_{12,u}(h)$, and $\Gamma_{22,u}(h)$ are similarly defined. Let

$$B = \begin{bmatrix} B_{11} & B_{12} \\ B_{12}' & B_{22} \end{bmatrix} := -\frac{\pi^2}{6} \begin{bmatrix} \sum_{h=-\infty}^{\infty} h^2 \Gamma_{11,u}(h) & \sum_{h=-\infty}^{\infty} h^2 \Gamma_{12,u}(h) \\ \sum_{h=-\infty}^{\infty} h^2 \Gamma_{12,u}'(h) & \sum_{h=-\infty}^{\infty} h^2 \Gamma_{22,u}(h) \end{bmatrix},$$

and

$$\tilde{B} = \Omega_{11\cdot 2}^{-1}(B_{11} - 2\Omega_{12}\Omega_{22}^{-1}B_{21} + \left(\Omega_{12}\Omega_{22}^{-1} \otimes \Omega_{12}\Omega_{22}^{-1}\right) vec(B_{22})).$$

From the theoretical results in Lemmas 4, the following Lemma leads to the higher order expansion of the over-identified location, which is the main result of this paper as follows.

Lemma 5 *Let Assumption 4 hold. Then, under the null hypothesis of $H_0 : \theta = \theta_0$, and, as $K, T \to \infty$ such that $K/T \to 0$, we have that*

$$\begin{aligned} P_{H_0}\left(|\tilde{\mathbb{T}}_T(\hat{\theta}_2; \hat{\theta}_1)|^2 \leq z\right) &= G_1(z) + \left(\frac{K^2}{T^2}\right)\tilde{B}G_1'(z)z - \left(\frac{q}{K}\right)G_1'(z)z + \frac{1}{K}G_1''(z)z^2 \\ &\quad + O\left(\frac{\log T}{\sqrt{T}}\right) + o\left(\frac{K^2}{T^2}\right) + o\left(\frac{1}{K}\right). \end{aligned}$$

Let $\mathcal{F}_{K-q}^{1-\alpha}$ be the $(1-\alpha)$ quantile of $\mathcal{F}_{K-q}$.We want to provide a higher-order expansion of

$$P_{H_0}\left(\left|\tilde{\mathbb{T}}_T(\hat{\theta}_T)\right|^2 > \mathcal{F}_{K-q}^{1-\alpha}\right).$$

Using a technical result proven in Hwang and Valdés (2022b), we show that

$$\mathcal{F}_{K-q}^{1-\alpha} = \chi^2_{1,1-\alpha} - \left(\frac{1}{K}\right)\frac{G_1''(\chi^2_{1,1-\alpha})}{G_1'(\chi^2_{1,1-\alpha})}\left(\chi^2_{1,1-\alpha}\right)^2 + o\left(\frac{1}{K}\right). \tag{8}$$

Using this result, the following theorem provides the high-order expansion of $\tilde{\mathbb{T}}_T(\hat{\theta}_2;\hat{\theta}_1)$, together with the F-critical value, which is the key finding of this paper.

Theorem 6 *Let Assumption 4 hold. and* $\tilde{\mathbb{F}}_{1-\alpha} = \tilde{\mathbb{T}}^2_{1-\alpha/2}$. *Then, under the null hypothesis of* $H_0 : \theta = \theta_0$, *and, as* $K, T \to \infty$ *such that* $K/T \to 0$, *we have that*

$$P_{H_0}\left(\left|\tilde{\mathbb{T}}_T(\hat{\theta}_2;\hat{\theta}_1)\right|^2 > \mathcal{F}_{K-q}^{1-\alpha}\right) = \alpha - \left(\frac{K^2}{T^2}\right)\left(G_1'(\chi^2_{1,1-\alpha})\chi^2_{1,1-\alpha}\right)\tilde{B} + O\left(\frac{\log T}{\sqrt{T}}\right) + o\left(\frac{K^2}{T^2}\right) + o\left(\frac{1}{K}\right).$$

Under the alternative fixed-K critical value, the result in Theorem 6 shows that the use of F critical value removes the error of order $O\left((1+q)/K\right)$ embodied in the standard GMM t-test via the normal (chi-square) critical value. In contrast, the null rejection probability of the standard GMM t-test using conventional normal critical value, $z_{1-\alpha/2}$, yields

$$\begin{aligned} P_{H_0}\left(\left|\mathbb{T}_T(\hat{\theta}_2;\hat{\theta}_1)\right| > z_{1-\alpha/2}\right) &= \alpha - \left(\frac{K^2}{T^2}\right)\left(G_1'(\chi^2_{1,1-\alpha})\chi^2_{1,1-\alpha}\right)\tilde{B} + O\left(\frac{(1+q)}{K}\right) \\ &\quad + O\left(\frac{\log T}{\sqrt{T}}\right) + o\left(\frac{K^2}{T^2}\right) + o\left(\frac{1}{K}\right), \end{aligned}$$

which is proved in Hwang and Valdés (2022b). In this sense, our result formally shows the second-order correctness of the alternative fixed-K asymptotics in two-step GMM framework.

For a general Wald statistic, recall that

$$\tilde{\mathbb{F}}_T := \tilde{\mathbb{F}}_T(\hat{\theta}_2;\hat{\theta}_1) = \left(\frac{K-q}{K}\right)\cdot\frac{\mathbb{F}_T}{1+\frac{1}{K}\mathbb{J}_T(\hat{\theta}_2;\hat{\theta}_1)},$$

where $\mathbb{F}_T(\hat{\theta}_2;\hat{\theta}_1) = T(\hat{\theta}_2-\theta_0)'[G_T'(\hat{\theta}_1)\Sigma_T^{-1}(\hat{\theta}_1)G_T(\hat{\theta}_1)]^{-1}(\hat{\theta}_2-\theta_0)$.In a similar way with the proof of Lemma 2, we can show that

$$\mathbb{F}_T(\hat{\theta}_2;\hat{\theta}_1) = \mathcal{F}_T(\theta_0) + \psi_T + \psi_T^*,$$

where $\mathcal{F}_T(\theta_0) = \left(G'\Sigma_T^{-1}(\theta_0)\sqrt{T}f_T(\theta_0)\right)'[G'\Sigma_T^{-1}(\theta_0)G]^{-1}\left(G'\Sigma_T^{-1}(\theta_0)\sqrt{T}f_T(\theta_0)\right)$.

Theorem 7 *Assume* $V \sim N(0,\tilde{\Sigma}_{2T})$. *Then, for* $K \geq d+q$ *and* $d=1$, *we have that* $P\left(|\tilde{\mathbb{F}}_T(\hat{\theta}_2;\hat{\theta}_1))| \leq z\right) = E\left[G_d\left(\tilde{\Xi}^{-1}\right)\right] + O\left(\frac{1}{T}\right)$, *where*

$$\tilde{\Xi}(\hat{\Lambda}) = \frac{\Xi}{1+\left(\frac{\chi_q^2}{K}\right)\cdot\left(\frac{Z_q}{\|Z_q\|}\right)'\Omega_{22}^{1/2\prime}\hat{\Omega}_{22}^{-1}\Omega_{22}^{1/2}\left(\frac{Z_q}{\|Z_q\|}\right)},$$

and

$$\begin{aligned}\Xi &= e'_\mu (I_d+\Phi)^{1/2\prime}\left(\Omega_{11\cdot 2}^{1/2}\hat{\Omega}_{11\cdot 2}^{-1}\Omega_{11\cdot 2}^{1/2}\right)(I_d+\Phi)^{1/2}e_\mu;\\ \Phi &= \Omega_{11\cdot 2}^{-1/2}\hat{\delta}\Omega_{22}\hat{\delta}'\Omega_{11\cdot 2}^{-1/2};\\ \hat{\delta} &= \Omega_{12}\Omega_{22}^{-1}-\hat{\Omega}_{12}\hat{\Omega}_{22}^{-1}.\end{aligned}$$

Also, $P_{H_0}\left(\left|\tilde{\mathbb{F}}_T(\hat{\theta}_2;\hat{\theta}_1)\right| > \tilde{\mathbb{F}}_{1-\alpha}\right) = \alpha - \left(\frac{K^2}{T^2}\right)\left(G'_d(\chi^2_{d,1-\alpha})\chi^2_{d,1-\alpha}\right)\tilde{B} + O\left(\frac{\log T}{\sqrt{T}}\right) + o\left(\frac{K^2}{T^2}\right) + o\left(\frac{1}{K}\right)$ *holds, where*

$$\tilde{B} = tr\left(B_{11}\Omega_{11\cdot 2}^{-1} - 2\Omega_{12}\Omega_{22}^{-1}B_{21}\Omega_{11\cdot 2}^{-1} + (\Omega_{12}\Omega_{22}^{-1}B_{22}\Omega_{22}^{-1}\Omega_{21})\Omega_{11\cdot 2}^{-1}\right)/d.$$

4 Testing-oriented Smoothing Parameter Selection for GMM

Suppose we are interested in testing the parameter θ_0. Using a conventional χ_1^α critical value, one can construct a two-sided confidence interval of the form $\{\theta : |\tilde{\mathbb{F}}_T(\hat{\theta}_2;\hat{\theta}_1)| \le \chi_p^\alpha\}$, and reject the null hypothesis $H_0 : R\theta_0 = r \in \mathbb{R}^p$ if and only if $\theta \notin \{\theta : |\tilde{\mathbb{F}}_T(\hat{\theta}_2;\hat{\theta}_1) \le \chi_p^{1-\alpha}\}$, where $1 - G_p\left(\chi_p^{1-\alpha}\right) = \alpha$. Then up to small order terms, $o(K^{-1})$, its Type-I error, or false rejection probability, is

$$\begin{aligned}&\left|P\left(\tilde{\mathbb{F}}_T(\hat{\theta}_2;\hat{\theta}_1) > \chi_p^\alpha\right) - \alpha\right|\\ =\ &\left|G'_p(\chi_p^{1-\alpha})\chi_p^{1-\alpha}\left(\frac{q}{K}\right) - G'_p(\chi_p^{1-\alpha})\chi_p^{1-\alpha}\left(\frac{K^2}{T^2}\right)\tilde{B} - G''_p(\chi_p^{1-\alpha})(\chi_p^{1-\alpha})^2\left(\frac{1}{K}\right)\right|\\ =\ &\left|\left(\frac{K^2}{T^2}\right)G'_p(\chi_p^{1-\alpha})\chi_p^{1-\alpha}\tilde{B} + \left(\frac{1}{K}\right)G''_p(\chi_p^{1-\alpha})(\chi_p^{1-\alpha})^2 - \left(\frac{q}{K}\right)G'_p(\chi_p^{1-\alpha})\chi_p^{1-\alpha}\right|\\ =\ &\left|\left(\frac{K^2}{T^2}\right)G'_p(\chi_p^{1-\alpha})\chi_p^{1-\alpha}\tilde{B} + \left(\frac{1}{K}\right)\left(G''_p(\chi_p^{1-\alpha})(\chi_p^{1-\alpha})^2 - qG'_p(\chi_p^{1-\alpha})\chi_p^{1-\alpha}\right)\right|\\ \le\ &\left(\frac{K^2}{T^2}\right)G'_p(\chi_p^{1-\alpha})\chi_p^{1-\alpha}\left|\tilde{B}\right| + \left(\frac{1}{K}\right)\left|G''_p(\chi_p^{1-\alpha})(\chi_p^{1-\alpha})^2 - qG'_p(\chi_p^{1-\alpha})\chi_p^{1-\alpha}\right|.\end{aligned}$$

Thus, to control the type-I error, we can choose optimal K^* to minimize the upper bound, and this leads us to F.O.C.

$$\left(\frac{2K^*_{\text{CPE}}}{T^2}\right)G'_p(\chi_p^{1-\alpha})\chi_p^{1-\alpha}\left|\tilde{B}\right| - \frac{1}{(K^*_{\text{CPE}})^2}\left|G''_p(\chi_p^{1-\alpha})(\chi_p^{1-\alpha})^2 - qG'_p(\chi_p^{1-\alpha})\chi_p^{1-\alpha}\right| = 0;$$

and we have that

$$\begin{aligned}K^*_{\text{CPE}} &= \left\lceil\left(\frac{\left|\left(G''_p(\chi_p^{1-\alpha})\chi_p^{1-\alpha} - qG'_p(\chi_p^{1-\alpha})\right)\right|}{2G'_p(\chi_p^{1-\alpha})\left|\tilde{B}\right|}\right)^{1/3}\cdot T^{2/3}\right\rceil\\ &= \left\lceil\left(\left|\frac{G''_p(\chi_p^{1-\alpha})\chi_p^{1-\alpha}}{2G'_p(\chi_p^{1-\alpha})\left|\tilde{B}\right|} - q\frac{G'_p(\chi_p^{1-\alpha})}{2G'_p(\chi_p^{1-\alpha})\left|\tilde{B}\right|}\right|\right)^{1/3}\cdot T^{2/3}\right\rceil\end{aligned}$$

where $\lceil\cdot\rceil$ is the ceiling function. Note that

$$\begin{aligned} \frac{G_p''(z)z}{2G_p'(z)} &= \frac{1}{2} \times \frac{-\frac{(z-p+2)}{2^{p/2+1}\Gamma(p/2)} z^{p/2-2} \exp\left(-\frac{z}{2}\right) z}{\frac{1}{2^{p/2}\Gamma(p/2)} z^{p/2-1} \exp\left(-\frac{z}{2}\right)} \\ &= -\frac{1}{4}(z-p+2) = \frac{1}{4}(p-z-2). \end{aligned}$$

Therefore, we have that

$$\begin{aligned} K_{\mathrm{CPE}}^* &= \left\lceil \left(\left| \frac{G_p''(\chi_p^{1-\alpha})\chi_p^{1-\alpha}}{2G_p'(\chi_p^{1-\alpha})\left|\tilde{B}\right|} - q\frac{G_p'(\chi_p^{1-\alpha})}{2G_p'(\chi_p^{1-\alpha})\left|\tilde{B}\right|} \right| \right)^{1/3} \cdot T^{2/3} \right\rceil \\ &= \left\lceil \left| \frac{(p-\chi_p^{1-\alpha}-2)-2q}{4\tilde{B}} \right|^{1/3} \cdot T^{2/3} \right\rceil . \end{aligned}$$

When $q=0$, i.e., the model is exactly identified with $\Omega_{12}=0$ and $B_{21}=0$, we have that

$$\begin{aligned} \tilde{B} &= \frac{1}{p} \cdot \operatorname{tr}\left(B_{11}\Omega_{11\cdot 2}^{-1} - 2\Omega_{12}\Omega_{22}^{-1}B_{21}\Omega_{11\cdot 2}^{-1} + (\Omega_{12}\Omega_{22}^{-1}B_{22}\Omega_{22}^{-1}\Omega_{21})\Omega_{11\cdot 2}^{-1} \right) \\ &= \frac{\operatorname{tr}(B_{11}\Omega_{11}^{-1})}{p}. \end{aligned}$$

Then, the optimal choice for K_{CPE}^* is equal to

$$\left\lceil \left(\frac{\left|\left(G_p''(\chi_p^{1-\alpha})\chi_p^{1-\alpha}\right)\right|}{2G_p'(\chi_p^{1-\alpha})\left|\tilde{B}\right|} \right)^{1/3} \cdot T^{2/3} \right\rceil ,$$

which is exactly the result for CPE-optimized K^* obtained in Sun (2013) in the context of the exactly identified GMM. Therefore, our result is a generalization of the optimal K-formula when we consider an over-identified GMM model. The optimal formula for K^* establishes a relation between the value of q and other model parameters such as Ω_{12}, Ω_{22}, and B_{21} in the over-identified GMM system.

5 Simulation Results

We follow the simulation design in Hwang and Sun (2017) and Hwang and Valdés (2022a) by considering the following linear structural model:

$$y_t = \alpha + x_{1,t}\beta_1 + x_{2,t}\beta_2 + x_{3,t}\beta_3 + \epsilon_{y,t},$$

where $x_{1,t}, x_{2,t}$, and $x_{3,t}$ are scalar regressors that are correlated with $\epsilon_{y,t}$. The unknown parameter vector is $\theta = (\alpha, \beta_1, \beta_2, \beta_3)' \in \mathbb{R}^d$ with $d=4$, and there are m instruments $z_{0,t}, z_{1,t}, ..., z_{m-1,t}$, with $z_{0,t} \equiv 1$. The reduced-form equations for $x_{1,t}, x_{2,t}$, and $x_{3,t}$ are given by

$$x_{j,t} = z_{j,t} + \sum_{i=d-1}^{m-1} z_{i,t} + \epsilon_{x_j,t} \text{ for } j \in \{1,2,3\}. \tag{9}$$

We assume that the $z_{i,t}$ for $i \geq 1$ follow an AR(1) process, that is, $z_{i,t} = \rho z_{i,t-1} + \sqrt{1-\rho^2} e_{z_i,t}$ where $(e^1_{zt}, ..., e^{m-1}_{zt})' \sim N(0, V_e)$. The diagonal elements of V_e are equal to 1 and the off-diagonal elements are equal to ψ. The data-generation process (DGP) for $\epsilon_t = (\epsilon_{yt}, \epsilon_{x_1t}, \epsilon_{x_2t}, \epsilon_{x_3t})'$ is the same as the DGP for $(z_{1,t}, ..., z_{m-1,t})'$ except for the dimensionality difference. Thus the parameter $\psi \in [0, 1)$ serves as a degree of endogeneity between the regressor $x_{j,t}$ and $\epsilon_{y,t}$. By construction, the vectors, ϵ_t and $(z_{1,t}, ..., z_{m-1,t})'$ are independent of each other. We consider the true parameters to be $\theta_0 = (0, 0, 0, 0)'$, $\rho \in \{0.5, 0.7, 0.9\}$, and $\psi = 0.5$. We consider the following null hypothesis of interest,

$$H_0 : \beta_1 = \beta_2 = \beta_3 = 0, \tag{10}$$

where the number of restricted parameters in R is $p = 3$ and the nominal significance level α is 5%. With $p = 3$, we consider $q = 5$ when $T = 200$ and $q = 7$ when $T = 200$.

Define $x_t = (1, x_{1,t}, x_{2,t}, x_{3,t})'$ and $z_t = (z_{0,t}, z_{1,t}, ..., z_{m-1,t})'$. Then we have that the m moment conditions are given by

$$E[f(v_t, \theta_0)] = E[z_t(y_t - x_t'\theta_0)] \in \mathbb{R}^m. \tag{11}$$

For the basis functions in the OS-HAR estimation, we use the orthonormal Fourier basis functions introduced in (**??**). For the choice of K^* in the OS LRV estimation, we employ the coverage probability optimal smoothing parameter, K^*_{CPE}. There are two versions of K^*_{CPE} in our simulations. The first one is

$$K^*_{\text{SUN}} = \left\lceil \left| \frac{(p - \chi_p^{1-\alpha} - 2)}{4B_{11}} \right|^{1/3} \cdot T^{2/3} \right\rceil, \tag{12}$$

which is proposed by Sun (2013) under the exact identified GMM. The second choice is a generalized version of the CPE-optimal K^* which is proposed in this paper in the presence of over-identification:

$$K^*_{\text{CPE}} = \left\lceil \left| \frac{(p - \chi_p^{1-\alpha} - 2) - 2q}{4\tilde{B}} \right|^{1/3} \cdot T^{2/3} \right\rceil. \tag{13}$$

The last choice of K^* is the AMSE-optimal formula in Phillips (2005) and Sun (2013):

$$K^*_{\text{MSE}} = \left\lceil \left(\frac{tr\left[(I_{m^2} + \mathbb{K}_{mm})(\Omega^* \otimes \Omega^*)\right]}{4vec(B^*)'vec(B^*)} \right)^{1/5} T^{4/5} \right\rceil. \tag{14}$$

$\lceil \cdot \rceil$ is the ceiling function, $\mathbb{K}_{mm}$ is the $m^2 \times m^2$ commutation matrix, and $\Omega^* = \sum_{j=-\infty}^{\infty} j^2 E u_t^* u_{t-j}^{*\prime}$ and $B^* = -\pi^2/6 \sum_{j=-\infty}^{\infty} j^2 E u_t^* u_{t-j}^{*\prime}$ with

$$u_t^* = R\left(G'\Sigma^{-1}G\right)^{-1} G'\Sigma^{-1} f(v_t, \theta_0).$$

The unknown parameters, $\tilde{B}$, B^*, Ω^*, and Σ in (12)–(14), can be either calibrated or data driven using the VAR(1) plug-in approach. Here, we use the data-driven VAR(1) plug-in approach, following Andrews (1991). The qualitative messages remain the same regardless of how the unknown parameters are obtained. The number of replications in all of our Monte Carlo simulations is $10,000$.

Under (10), we study the empirical rejection probability (ERP) of the Wald statistic, $F(\hat{\theta}_2)$, and the finite-sample corrected Wald statistic, $F_c^{\text{adj}}(\hat{\theta}_2)$, which use chi-square critical values derived from the conventional increasing-smoothing asymptotics. We also investigate performances of the Wald statistic with the J-statistic modification, $\tilde{F}(\hat{\theta}_2)$, as in Hwang and Sun (2017) and examine its finite-sample corrected version $\tilde{F}_c^{\text{adj}}(\hat{\theta}_2)$. Note that both $\tilde{F}(\hat{\theta}_2)$ and $\tilde{F}_c^{\text{adj}}(\hat{\theta}_2)$ use asymptotic F critical values which are derived under the fixed smoothing asymptotics, and they are proven to have higher-order accuracy compared to the conventional Wald statistics, $F(\hat{\theta}_2)$ and $F_c^{\text{adj}}(\hat{\theta}_2)$ that implement the chi-square critical values.

Tables 1 and 2 report the numerical results. In sum, our numerical findings for two-step GMM are consistent with the theoretical results developed in this paper, which first indicate that the conventional chi-square tests, $F(\hat{\theta}_2)$ and $F_c^{\text{adj}}(\hat{\theta}_2)$, have the large size distortion, which increases with both the error dependence and the number of over-identificaitons (instrumental variables). Second, the asymptotic F tests with the J-statistic modifications, $\tilde{F}(\hat{\theta}_2)$ and $\tilde{F}_c^{\text{adj}}(\hat{\theta}_2)$, eliminates the size distortion of the conventional Wald test to a great extent. Note that this paper rigorously proves the finite-sample accuracy of the asymptotic F-test, where the critical value from the F distribution is second-order correct under the conventional increasing-smoothing asymptotics. Lastly, Table 2 shows that the testing-oriented (CPE) selection of smoothing parameters with K^*_{SUN} and K^*_{HV}, yields less amount of Type-I error than using the popular AMSE selection rule, K^*_{MSE}. Also, comparing the performances of the two optimal smoothing parameter rules, K^*_{SUN} and K^*_{HV}, we find that the use of K^*_{HV} is more advantageous, because it explicitly takes account the over-identification.

6 Conclusion

In this paper, we develop a higher-order expansion of test statistic for heteroskedasticity auto-correlation robust (HAR) inference in an over-identified two-step GMM framework. Our novel higher-order expansion rigorously shows the asymptotic F and t-tests (Hwang and Sun, 2017; Hwang and Valdes, 2022a), driven under the alternative fixed-smoothing asymptotics, are second-order correct under the conventional fixed-smoothing asymptotics.

We also provide an optimal choice of smoothing parameters that directly accounts for the degree of over-identificaion in HAR inferential problem. Our formula for the optimal smoothing parameter generalizes the formula in Sun (2013) which works under the exact identified GMM, as both formulas coincide when there is no over-identification.

Table 1: Empirical rejection probabilities of two-step GMM tests using OS LRV at nominal level $\alpha = 0.05$ with AR(1) coefficient $\rho \in \{0.50, 0.70, 0.90\}$

Asymptotic GMM-χ_p^2 tests without J-statistic modification						
	AR(1) coefficient $\rho = 0.50$					
	$T = 100$ and $q = 5$			$T = 200$ and $q = 7$		
	K^*_{MSE}	K^*_{SUN}	K^*_{HV}	K^*_{MSE}	K^*_{SUN}	K^*_{HV}
$F(\hat{\theta}_2)$	0.3738	0.3492	0.3536	0.2673	0.2496	0.2576
$F_c^{adj}(\hat{\theta}_2)$	0.2779	0.2714	0.2712	0.2008	0.1946	0.1992
	AR(1) coefficient $\rho = 0.70$					
	$T = 100$ and $q = 5$			$T = 200$ and $q = 7$		
	K^*_{MSE}	K^*_{SUN}	K^*_{HV}	K^*_{MSE}	K^*_{SUN}	K^*_{HV}
$F(\hat{\theta}_2)$	0.5561	0.5343	0.5325	0.4196	0.4000	0.4024
$F_c^{adj}(\hat{\theta}_2)$	0.4094	0.4204	0.4117	0.3079	0.2989	0.3054
	AR(1) coefficient $\rho = 0.90$					
	$T = 100$ and $q = 5$			$T = 200$ and $q = 7$		
	K^*_{MSE}	K^*_{SUN}	K^*_{HV}	K^*_{MSE}	K^*_{SUN}	K^*_{HV}
$F(\hat{\theta}_2)$	0.8736	0.8400	0.8305	0.7924	0.7379	0.7331
$F_c^{adj}(\hat{\theta}_2)$	0.5584	0.6946	0.6660	0.5033	0.5711	0.5523

Note: $F(\hat{\theta}_2)$ and $F_c^{\text{adj}}(\hat{\theta}_2)$ report empirical null rejection probabilities for two-step GMM Wald tests for testing the null hypothesis in (10) with conventional chi-square critical values using uncorrected. The nominal size α is 5%, and the number of replications is 10,000.

Table 2: Empirical rejection probabilities of two-step GMM tests using OS LRV at nominal level $\alpha = 0.05$ with AR(1) coefficient $\rho \in \{0.50, 0.70, 0.90\}$

Asymptotic GMM-F tests with J-statistic modifications						
	AR(1) coefficient $\rho = 0.50$					
	$T = 100$ and $q = 5$			$T = 200$ and $q = 7$		
	K^*_{MSE}	K^*_{SUN}	K^*_{HV}	K^*_{MSE}	K^*_{SUN}	K^*_{HV}
$\tilde{F}(\hat{\theta}_2)$	0.1442	0.0916	0.0560	0.1072	0.0628	0.0673
$\tilde{F}_c^{adj}(\hat{\theta}_2)$	0.0930	0.0657	0.0385	0.0753	0.0432	0.0494
	AR(1) coefficient $\rho = 0.70$					
	$T = 100$ and $q = 5$			$T = 200$ and $q = 7$		
	K^*_{MSE}	K^*_{SUN}	K^*_{HV}	K^*_{MSE}	K^*_{SUN}	K^*_{HV}
$\tilde{F}(\hat{\theta}_2)$	0.2050	0.1134	0.0697	0.1661	0.0620	0.0794
$\tilde{F}_c^{adj}(\hat{\theta}_2)$	0.1283	0.0809	0.0462	0.1064	0.0421	0.0537
	AR(1) coefficient $\rho = 0.90$					
	$T = 100$ and $q = 5$			$T = 200$ and $q = 7$		
	K^*_{MSE}	K^*_{SUN}	K^*_{HV}	K^*_{MSE}	K^*_{SUN}	K^*_{HV}
$\tilde{F}(\hat{\theta}_2)$	0.3174	0.1989	0.0559	0.2633	0.0931	0.0455
$\tilde{F}_c^{adj}(\hat{\theta}_2)$	0.1760	0.1403	0.0306	0.1380	0.0601	0.0271

Note: $\tilde{F}(\hat{\theta}_2)$ and $\tilde{F}_c^{\text{adj}}(\hat{\theta}_2)$ indicate those for two-step GMM Wald tests with alternative F critical values and J-statistic modifications using uncorrected and corrected variance estimates, respectively. The nominal size α is 5%, and the number of replications is 10,000.

7 Appendix of Proofs

Proof of Lemma 1–3. The proof is the same as that of Lemma 1–3 of Hwang and Valdés (2022b), and they are omitted here. ■

Proof of Lemma 4. The J-statistic modified t-statistic in the over-identification model is

$$\tilde{\mathbb{T}}_T = \frac{\sqrt{T}(\hat{\theta}_2^* - \theta_0^*)}{\sqrt{\hat{\Omega}_{11\cdot 2}\left(1 + (\sqrt{T}\bar{y}_2)'\hat{\Omega}_{22}^{-1}(\sqrt{T}\bar{y}_2)/K\right)}}.$$

We define L_T and $\tilde{\Sigma}_{2T}$ as below:

$$\begin{aligned} L_T &= (1_T \otimes I_2) = \begin{pmatrix} I_2 \\ \vdots \\ I_2 \end{pmatrix} \in \mathbb{R}^{2T\times 2}; \\ \tilde{\Sigma}_{2T} &= Var((y_1', y_2', \ldots, y_T')) \\ &= Var(Y_T) \in \mathbb{R}^{2T\times 2T}. \end{aligned}$$

Without loss of generality, we assume that $q = 1$, we construct an infeasible GLS of the over-identified system, i.e.

$$\begin{aligned} \hat{\mu}_{GLS} &= \begin{pmatrix} \bar{y}_{1,GLS} \\ \bar{y}_{2,GLS} \end{pmatrix} = \left[(1_T \otimes I_2)'\left[\tilde{\Sigma}_{2T}\right]^{-1}(1_T \otimes I_2)\right]^{-1}\left[(1_T \otimes I_2)'\left[\tilde{\Sigma}_{2T}\right]^{-1} Y_T\right] \in \mathbb{R}^2 \\ &= \left[L_T'\left[\tilde{\Sigma}_{2T}\right]^{-1} L_T\right]^{-1}\left[L_T'\left[\tilde{\Sigma}_{2T}\right]^{-1} Y_T\right]. \end{aligned}$$

We define a GLS version of the two-step GMM estimator as below:

$$\begin{aligned} \hat{\theta}_{T,GLS} &= \bar{y}_{1,GLS} - \hat{\Omega}_{12}\hat{\Omega}_{22}^{-1}\bar{y}_{2,GLS} \\ &= \begin{pmatrix} 1 & -\hat{\Omega}_{12}\hat{\Omega}_{22}^{-1} \end{pmatrix}\begin{pmatrix} \bar{y}_{1,GLS} \\ \bar{y}_{2,GLS} \end{pmatrix} \\ &= \hat{\Pi}\hat{\mu}_{GLS}. \end{aligned}$$

and define the GLS-version of t-statistics:

$$\mathbb{T}_T(\hat{\theta}_{T,GLS}) = \frac{\sqrt{T}(\hat{\theta}_{T,GLS} - \theta_0)}{\sqrt{\hat{\Omega}_{11\cdot 2}}},$$

and the GLS version of the J-statistic

$$\begin{aligned} \mathbb{J}_{T,GLS}(\hat{\theta}_{T,GLS}) &= T g_{T,GLS}(\hat{\theta}_{T,GLS})'\hat{\Omega}^{-1} g_{T,GLS}(\hat{\theta}_{T,GLS}) \\ &= (\sqrt{T}\bar{y}_{2,GLS})'\hat{\Omega}_{22}^{-1}(\sqrt{T}\bar{y}_{2,GLS}), \end{aligned}$$

where

$$\hat{\Omega} = \begin{pmatrix} \hat{\Omega}_{11} & \hat{\Omega}_{12} \\ \hat{\Omega}_{21} & \hat{\Omega}_{22} \end{pmatrix} = \frac{1}{T}\sum_{t=1}^{T}\sum_{s=1}^{T} Q_K\left(\frac{t}{T}, \frac{s}{T}\right)(y_t - \bar{y})(y_s - \bar{y})'.$$

The GLS version of the J-statistic modified t-statistic is

$$\begin{aligned}\tilde{\mathbb{T}}_T(\hat{\theta}_{T,GLS}) &= \sqrt{\frac{K-q}{K}} \cdot \frac{\mathbb{T}_T(\hat{\theta}_{T,GLS})}{\sqrt{1+\mathbb{J}_{T,GLS}(\hat{\theta}_{T,GLS})/K}} \\ &= \frac{\sqrt{T}(\hat{\theta}_{T,GMM}-\theta_0)}{\sqrt{\hat{\Omega}_{11\cdot 2}\left(1+(\sqrt{T}\bar{y}_{2,GLS})'\hat{\Omega}_{22}^{-1}(\sqrt{T}\bar{y}_{2,GLS})/K\right)}}.\end{aligned}$$

We assume a linear process for $u_t = (u_{1t}', u_{2t}')' \in \mathbb{R}^{1+q}$, i.e.,

$$u_t = C(L)\epsilon_t = \sum_{j=0}^{\infty} C_j \epsilon_{t-j},$$

where $\epsilon_t = (\epsilon_{1t}, \epsilon_{2t}')' \overset{\text{i.i.d}}{\sim} N(0, I_{1+q})$, and C_j's are $(q+1)\times(q+1)$ matrix such that

$$\begin{aligned} E\|\epsilon_t\|^{\nu} < \infty \text{ for some } \nu \geq 4; \\ \sum_{j=0}^{\infty} j^a \|C_j\| < \infty \text{ for } a > 3; \\ \Omega = C(1)C(1)' > 0,\end{aligned}$$

where $C(1) = \sum_{j=0}^{\infty} C_j$. Assume $V \sim N(0, \tilde{\Sigma}_{2T})$. Then, for $K \geq d+q$ and $d=1$, Then, we can show that

$$\begin{aligned} P\left(|\mathbb{T}_T(\hat{\theta}_{T,GLS})|^2 \leq z^2\right) &= P\left(\left(\frac{\sqrt{T}(\hat{\theta}_{T,GLS}-\theta_0)}{\sqrt{\hat{\Omega}_{11\cdot 2}}}\right)^2 \leq z^2\right) \\ &= P\left(\Theta \times (1+\Phi_T) \times \left(\frac{\Omega_{11\cdot 2}}{\hat{\Omega}_{11\cdot 2}}\right) + O\left(\frac{1}{T}\right)\right) \\ &= E\left[G_1\left(z\Xi^{-1}\right)\right] + O\left(\frac{1}{T}\right),\end{aligned}$$

where

$$\begin{aligned}\Phi_T &= \frac{(\Omega_{12}\Omega_{22}^{-1} - \hat{\Omega}_{12}\hat{\Omega}_{22}^{-1})\Omega_{T,22}(\Omega_{12}\Omega_{22}^{-1} - \hat{\Omega}_{12}\hat{\Omega}_{22}^{-1})'}{\Omega_{T,11\cdot 2}}; \\ \Xi &= \left(1 + \frac{\left(\Omega_{12}\Omega_{22}^{-1} - \hat{\Omega}_{12}\hat{\Omega}_{22}^{-1}\right)\Omega_{22}\left(\Omega_{12}\Omega_{22}^{-1} - \hat{\Omega}_{12}\hat{\Omega}_{22}^{-1}\right)'}{\Omega_{11\cdot 2}}\right) \times \left(\frac{\Omega_{11\cdot 2}}{\hat{\Omega}_{11\cdot 2}}\right),\end{aligned}$$

and Θ is a $\chi^2(1)$ random variable which is independent of both $\hat{\Omega}_{11\cdot 2}$ and Φ_T. Now, we want to expand the probability of $\tilde{\mathbb{T}}_T(\hat{\theta}_{T,GLS})$, i.e.,

$$P\left(\left|\sqrt{\frac{K-q}{K}}\tilde{\mathbb{T}}_T(\hat{\theta}_{T,GLS})\right|^2 \leq z^2\right) = P\left(\left(\sqrt{\frac{K-q}{K}}\frac{\sqrt{T}(\hat{\theta}_{T,GMM}-\theta_0)}{\sqrt{\hat{\Omega}_{11\cdot 2}\left(1+(\sqrt{T}\bar{y}_{2,GLS})'\hat{\Omega}_{22}^{-1}(\sqrt{T}\bar{y}_{2,GLS})/K\right)}}\right)^2 \leq z^2\right).$$

Using $\sqrt{T}\bar{y}_{2,GLS} \sim N(0, \Omega_{T,22})$, and $\sqrt{T}\bar{y}_{2,GLS}$ and $\hat{\Omega}_{22}$ are independent, we have that

$$\begin{aligned}
&(\sqrt{T}\bar{y}_{2,GLS})'\hat{\Omega}_{22}^{-1}(\sqrt{T}\bar{y}_{2,GLS}) \\
= \;& \frac{\left\|\sqrt{T}\bar{y}_{2,GLS}\right\|^2}{\left(\left(\frac{\sqrt{T}\bar{y}_{2,GLS}}{\left\|\sqrt{T}\bar{y}_{2,GLS}\right\|}\right)'\hat{\Omega}_{22}^{-1}\left(\frac{\sqrt{T}\bar{y}_{2,GLS}}{\left\|\sqrt{T}\bar{y}_{2,GLS}\right\|}\right)\right)^{-1}}.
\end{aligned}$$

Since $\sqrt{T}\bar{y}_{2,GLS} \sim N(0, \Omega_{T,22})$, $\left\|\sqrt{T}\bar{y}_{2,GLS}\right\|^2$ is independent of $\sqrt{T}\bar{y}_{2,GLS}/\left\|\sqrt{T}\bar{y}_{2,GLS}\right\|$, the direction of $\sqrt{T}\bar{y}_{2,GLS}$. Hence, the numerator is independent of the denominator. Also, note that

$$\left\|\sqrt{T}\bar{y}_{2,GLS}\right\|^2 \overset{d}{=} \left\|\Omega_{T,22}^{1/2}Z_q\right\|^2 = Z_q'\Omega_{T,22}Z_q = ||Z_q||^2\left(\frac{Z_q'}{||Z_q||}\Omega_{T,22}\frac{Z_q}{||Z_q||}\right).$$

Let H be an orthonormal random matrix such that

$$H'\left(\frac{Z_q'}{||Z_q||}\right) = \begin{pmatrix} 1 & 0 & \cdots & 0 \end{pmatrix} = e_1.$$

Then, we have that

$$\begin{aligned}
\left(\frac{Z_q}{||Z_q||}\right)'\Omega_{T,22}\left(\frac{Z_q}{||Z_q||}\right) &= \left(H'\frac{Z_q}{||Z_q||}\right)'H'\Omega_{T,22}H\left(H'\frac{Z_q}{||Z_q||}\right) \\
&= e_1'(H'\Omega_{T,22}H)e_1,
\end{aligned}$$

which leads us to conclude that

$$\begin{aligned}
&\left\|\sqrt{T}\bar{y}_{2,GLS}\right\|^2 \overset{d}{=} ||Z_q||^2\left(\frac{Z_q'}{||Z_q||}\Omega_{T,22}\frac{Z_q}{||Z_q||}\right) \\
&\overset{d}{=} \chi_q^2 \cdot e_1'(H'\Omega_{T,22}H)e_1.
\end{aligned}$$

Similarly, we have that

$$\begin{aligned}
&\left(\frac{\sqrt{T}\bar{y}_{2,GLS}}{\left\|\sqrt{T}\bar{y}_{2,GLS}\right\|}\right)'\hat{\Omega}_{22}^{-1}\left(\frac{\sqrt{T}\bar{y}_{2,GLS}}{\left\|\sqrt{T}\bar{y}_{2,GLS}\right\|}\right) \\
&\overset{d}{=} \left(\frac{\Omega_{T,22}^{1/2}Z_q}{\sqrt{\|Z_q\| \cdot e_1'(H'\Omega_{T,22}H)e_1}}\right)'\hat{\Omega}_{22}^{-1}\left(\frac{\Omega_{T,22}^{1/2}Z_q}{\sqrt{\|Z_q\| \cdot e_1'(H'\Omega_{T,22}H)e_1}}\right) \\
= \;& \left(\frac{\Omega_{T,22}^{1/2}Z_q}{\|Z_q\|}\right)'\frac{\hat{\Omega}_{22}^{-1}}{e_1'(H'\Omega_{T,22}H)e_1}\left(\frac{\Omega_{T,22}^{1/2}Z_q}{\|Z_q\|}\right).
\end{aligned}$$

Combining this, we have that

$$
\begin{aligned}
(\sqrt{T}\bar{y}_{2,GLS})'\hat{\Omega}_{22}^{-1}(\sqrt{T}\bar{y}_{2,GLS}) &= \frac{\left\|\sqrt{T}\bar{y}_{2,GLS}\right\|^2}{\left(\left(\frac{\sqrt{T}\bar{y}_{2,GLS}}{\|\sqrt{T}\bar{y}_{2,GLS}\|}\right)'\hat{\Omega}_{22}^{-1}\left(\frac{\sqrt{T}\bar{y}_{2,GLS}}{\|\sqrt{T}\bar{y}_{2,GLS}\|}\right)\right)^{-1}} \\
&= \frac{||Z_q||^2(e_1'(H'\Omega_{T,22}H)e_1).}{\left(\left(\frac{\Omega_{T,22}^{1/2}Z_q}{\|Z_q\|}\right)'\frac{\hat{\Omega}_{22}^{-1}}{e_1'(H'\Omega_{T,22}H)e_1}\left(\frac{\Omega_{T,22}^{1/2}Z_q}{\|Z_q\|}\right)\right)^{-1}} \\
&= \chi_q^2\cdot\underbrace{\left(\frac{Z_q}{\|Z_q\|}\right)'\Omega_{T,22}^{1/2\prime}\hat{\Omega}_{22}^{-1}\Omega_{T,22}^{1/2}\left(\frac{Z_q}{\|Z_q\|}\right)}_{:=\Delta_T}.
\end{aligned}
$$

Thus, we have that

$$
\begin{aligned}
& P\left(\Theta\tilde{\Xi}\le z\right)+O\left(\frac{1}{T}\right) \\
: &= P\left(\Theta\times(1+\Phi_T)\times\left(\frac{\Omega_{11\cdot 2}}{\hat{\Omega}_{11\cdot 2}}\right)\left(\frac{1}{\left(1+(\sqrt{T}\bar{y}_{2,GLS})'\hat{\Omega}_{22}^{-1}(\sqrt{T}\bar{y}_{2,GLS})/K\right)}\right)+O\left(\frac{1}{T}\right)\right) \\
&= P\left(\Theta\times\left\{\frac{(1+\Phi)}{(1+\chi_q^2\cdot\Delta/K)}\times\left(\frac{\Omega_{11\cdot 2}}{\hat{\Omega}_{11\cdot 2}}\right)\right\}+O\left(\frac{1}{T}\right)\right),
\end{aligned}
$$

where

$$
\begin{aligned}
\tilde{\Xi}(\hat{\Lambda}) &= \left\{\left(1+\Phi(\hat{\Omega}_{11},\hat{\Omega}_{12},\hat{\Omega}_{22})\right)\times\left(\frac{\Omega_{11\cdot 2}}{\hat{\Omega}_{11\cdot 2}}\right)\times\frac{1}{\left(1+\chi_q^2\cdot\Delta(\hat{\Omega}_{22})/K\right)}\right\} \\
&= \left\{\frac{\left(1+\frac{(\Omega_{12}\Omega_{22}^{-1}-\hat{\Omega}_{12}\hat{\Omega}_{22}^{-1})\Omega_{T,22}(\Omega_{12}\Omega_{22}^{-1}-\hat{\Omega}_{12}\hat{\Omega}_{22}^{-1})'}{\Omega_{T,11\cdot 2}}\right)\times\left(\frac{\Omega_{11\cdot 2}}{\hat{\Omega}_{11\cdot 2}}\right)}{1+\left(\frac{\chi_q^2}{K}\right)\cdot\left(\frac{Z_q}{\|Z_q\|}\right)'\Omega_{22}^{1/2\prime}\hat{\Omega}_{22}^{-1}\Omega_{22}^{1/2}\left(\frac{Z_q}{\|Z_q\|}\right)}\right\} \\
&= \frac{\Xi}{1+\left(\frac{\chi_q^2}{K}\right)\cdot\left(\frac{Z_q}{\|Z_q\|}\right)'\Omega_{22}^{1/2\prime}\hat{\Omega}_{22}^{-1}\Omega_{22}^{1/2}\left(\frac{Z_q}{\|Z_q\|}\right)},
\end{aligned}
$$

Therefore, we conclude that

$$
\begin{aligned}
P\left(|\tilde{\mathbb{T}}_T(\hat{\theta}_{T,GLS})|\le z\right) &= E\left[G_1\left(z\tilde{\Xi}^{-1}(\hat{\Lambda})\right)\right]+O\left(\frac{1}{T}\right) \\
&= E\left[G_1\left(z\Xi^{-1}(\hat{\Omega}_{11},\hat{\Omega}_{12},\hat{\Omega}_{22})\cdot\left(1+\frac{\chi_q^2}{K}\Delta(\hat{\Omega}_{22})\right)\right)\right]+O\left(\frac{1}{T}\right),
\end{aligned}
$$

which leads to

$$
P(\tilde{\mathbb{T}}_T(\hat{\theta}_2;\hat{\theta}_1)^2\le\kappa\cdot z)=E[G_1\left(\tilde{\Xi}^{-1}\right)]+O\left(\frac{\log T}{\sqrt{T}}\right)+o\left(\frac{K^2}{T^2}\right)+o\left(\frac{1}{K}\right),
$$

as desired. ■

Proof of Lemma 5. We first want to expand the following main term

$$E\left[zG_1\left(\Xi^{-1}\cdot\left(1+\frac{\chi_q^2}{K}\cdot\Delta\right)\right)\right],$$

where

$$\begin{aligned}\Xi &= \Xi\left(\left[\hat{\Omega}_{11},\hat{\Omega}_{12},vec(\hat{\Omega}_{22})'\right]'\right);\\ \Delta(\hat{\Omega}_{22}) &= \left(\frac{Z_q}{\|Z_q\|}\right)'\Omega_{22}^{1/2\prime}\hat{\Omega}_{22}^{-1}\Omega_{22}^{1/2}\left(\frac{Z_q}{\|Z_q\|}\right).\end{aligned}$$

We want to expand the term $\Xi^{-1}(1+\chi_q^2/K\cdot\mathbb{D})$ at the true value of Ω:

$$\begin{aligned}&\Xi^{-1}(\hat{\Lambda})\cdot\left(1+\frac{\chi_q^2}{K}\cdot\Delta(\hat{\Lambda})\right)\\ \approx\ &\Xi^{-1}(\Lambda)\cdot\left(1+\frac{\chi_q^2}{K}\cdot\Delta(\Lambda)\right)+\nabla'\left\{\Xi^{-1}(\hat{\Lambda})\left(1+\frac{\chi_q^2}{K}\cdot\gneqq(\hat{\Lambda})\right)\right\}\Bigg|_{\hat{\Lambda}=\Lambda}(\hat{\Lambda}-\Lambda)\\ &+\frac{1}{2}(\hat{\Lambda}-\Lambda)'\left\{\nabla^2\Xi^{-1}(\hat{\Lambda})\left(1+\frac{\chi_q^2}{K}\cdot\gneqq(\hat{\Lambda})\right)\Bigg|_{\hat{\Lambda}=\Lambda}\right\}(\hat{\Lambda}-\Lambda)\end{aligned}$$

After some algebra which are available upon request, we can show that

$$\begin{aligned}P\left(|\tilde{\mathbb{T}}_T(\hat{\theta}_{T,GLS})|\le z\right) &= E\left[zG_1\left(\Xi^{-1}\cdot\left(1+\frac{\chi_q^2}{K}\cdot\gneqq\right)\right)\right]+O\left(\frac{1}{T}\right)\\ &= E\left[z\left(\left(1+\frac{\chi_q^2}{K}\right)+\tilde{L}+\tilde{Q}\right)\right]+O\left(\frac{1}{T}\right),\end{aligned}$$

where

$$\begin{aligned}\tilde{L} &= \nabla'\left\{\Xi^{-1}(\hat{\Lambda})\left(1+\frac{\chi_q^2}{K}\cdot\gneqq(\hat{\Lambda})\right)\right\}\Bigg|_{\hat{\Lambda}=\Lambda}(\hat{\Lambda}-\Lambda)\\ &= L\times\left(1+\frac{\chi_q^2}{K}\right)-\frac{\chi_q^2}{K}\cdot\frac{1}{\|Z_q\|^2}\begin{pmatrix}0 & 0 & \left[Z_q'\Omega_{22}^{-1/2\prime}\right]\otimes\left[Z_q'\Omega_{22}^{-1/2\prime}\right]\end{pmatrix}(\hat{\Lambda}-\Lambda)\\ &= L\times\left(1+\frac{\chi_q^2}{K}\right)-\frac{\chi_q^2}{K}\cdot\underbrace{\frac{1}{\|Z_q\|^2}\left[Z_q'\Omega_{22}^{-1/2\prime}\right]\otimes\left[Z_q'\Omega_{22}^{-1/2\prime}\right]\left(vec(\hat{\Omega}_{22})-vec(\Omega_{22})\right)}_{=l_1},\end{aligned}$$

and

$$
\begin{aligned}
\tilde{Q} &= \frac{1}{2}(\hat{\Lambda}-\Lambda)' \left\{ \nabla^2 \Xi^{-1}(\hat{\Lambda}) \left(1+\frac{\chi_q^2}{K}\cdot \gtreqqless(\hat{\Lambda})\right)\Bigg|_{\hat{\Lambda}=\Lambda} \right\} (\hat{\Lambda}-\Lambda) \\
&= \left(1+\frac{\chi_q^2}{K}\right) Q + \frac{1}{2}(\hat{\Lambda}-\Lambda)' S_2 (\hat{\Lambda}-\Lambda) + \frac{1}{2}(\hat{\Lambda}-\Lambda)' S_3 (\hat{\Lambda}-\Lambda) \\
&= \left(1+\frac{\chi_q^2}{K}\right) Q + 0 + \frac{\chi_q^2}{K}\cdot \underbrace{\frac{1}{2}\left(vec(\hat{\Omega}_{22})-vec(\Omega_{22})\right)' S_3 \left(vec(\hat{\Omega}_{22})-vec(\Omega_{22})\right)}_{=l_2}.
\end{aligned}
$$

We continue the equation

$$
\begin{aligned}
P\left(|\tilde{\mathbb{T}}_T(\hat{\theta}_{T,GLS})| \le z\right) &= E\left[zG_1\left(\Xi^{-1}\cdot\left(1+\frac{\chi_q^2}{K}\cdot \gtreqqless\right)\right)\right] + O\left(\frac{1}{T}\right)\left(\frac{K}{K-q}\right) \\
&= E\left[G_1\left(z\left(1+\frac{\chi_q^2}{K}+\tilde{L}+\tilde{Q}\right)\right)\right] + O\left(\frac{1}{T}\right) \\
&= EG_1(z) + E\left\{G_1'(z)z\left(\frac{\chi_q^2}{K}+\tilde{L}+\tilde{Q}\right)\right\} \\
&\quad +\frac{1}{2}E\left[G_1''(z)\left\{z\left(\frac{\chi_q^2}{K}+\tilde{L}+\tilde{Q}\right)\right\}^2\right] + \text{small orders} \\
&= G_1(z) + G_1'(z)z \\
&\quad \times \underbrace{E\left[\frac{\chi_q^2}{K} + \left\{L\times\left(1+\frac{\chi_q^2}{K}\right) - \frac{\chi_q^2}{K}\cdot l_1\right\} + \left\{\left(1+\frac{\chi_q^2}{K}\right)Q + \frac{\chi_q^2}{K}\cdot l_2\right\}\right]}_{:=(\mathrm{A})} \\
&\quad +\frac{1}{2}G_1''(z)z^2 \\
&\quad \times \underbrace{E\left[\left(\frac{\chi_q^2}{K} + \left\{L\times\left(1+\frac{\chi_q^2}{K}\right) - \frac{\chi_q^2}{K}\cdot l_1\right\} + \left\{\left(1+\frac{\chi_q^2}{K}\right)Q + \frac{\chi_q^2}{K}\cdot l_2\right\}\right)^2\right]}_{:=(\mathrm{B})} + \text{small orders}
\end{aligned}
$$

We provide an expression for the term (A) as follows:

$$
\begin{aligned}
& E\left[\frac{\chi_q^2}{K} + \left\{L\times\left(1+\frac{\chi_q^2}{K}\right) - \frac{\chi_q^2}{K}\cdot l_1\right\} + \left\{\left(1+\frac{\chi_q^2}{K}\right)Q + \frac{\chi_q^2}{K}\cdot l_2\right\}\right] \\
= {} & \frac{E\left[\chi_q^2\right]}{K} + E\left[\left(1+\frac{\chi_q^2}{K}\right)\right]E[L] - E\left[\frac{\chi_q^2}{K}\right]E[l_1] \\
& + E\left[\left(1+\frac{\chi_q^2}{K}\right)\right]E[Q] + E\left[\frac{\chi_q^2}{K}\right]E[l_2],
\end{aligned}
$$

where $E[\chi_q^2] = q$, and

$$\begin{aligned} E[l_1] &= \frac{K^2}{2T^2} vec(\Omega_{22}^{-1})' vec(B_{22})(1+o(1)) + o\left(\frac{K^2}{T^2}\right) + O\left(\frac{1}{T}\right); \\ E[l_2] &= \frac{3q}{2K}(1+o(1)) + o\left(\frac{K^2}{T^2}\right) + O\left(\frac{1}{T}\right). \end{aligned}$$

Using these results, we have that

$$\begin{aligned} & E\left[\frac{\chi_q^2}{K} + \left\{L \times \left(1 + \frac{\chi_q^2}{K}\right) - \frac{\chi_q^2}{K} \cdot l_1\right\} + \left\{\left(1 + \frac{\chi_q^2}{K}\right) Q + \frac{\chi_q^2}{K} \cdot l_2\right\}\right] \\ = & \frac{q}{K} + \left(1 + \frac{q}{K}\right) E[L] - \frac{q}{K}\left(\frac{K^2}{2T^2} vec(\Omega_{22}^{-1})' vec(B_{22})(1+o(1))\right) \\ & + \left(1 + \frac{q}{K}\right) E[Q] + \frac{q}{K}\frac{3q}{2K}(1+o(1)) + o\left(\frac{K^2}{T^2}\right) + O\left(\frac{1}{T}\right) \\ = & \frac{q}{K} + \left(1 + \frac{q}{K}\right)(E[L] + E[Q]) - \frac{q}{2K}\left(\frac{K^2}{T^2} vec(\Omega_{22}^{-1})' vec(B_{22}) - \frac{3q}{K}\right)(1+o(1)) + o\left(\frac{K^2}{T^2}\right) + O\left(\frac{1}{T}\right) \\ = & \frac{q}{K} + \left(1 + \frac{q}{K}\right)(E[L] + E[Q]) + o\left(\frac{1}{K}\right) + o\left(\frac{K^2}{T^2}\right) + O\left(\frac{1}{T}\right) \end{aligned}$$

Note that

$$\begin{aligned} E\left[\left(\frac{\chi_q^2}{K} + \tilde{L} + \tilde{Q}\right)^2\right] &= E\left[\left(\frac{\chi_q^2}{K}\right)^2\right] + E\left[\tilde{L}^2\right] + E[\tilde{Q}^2] + 2E\left[\left(\frac{\chi_q^2}{K}\right)\tilde{L}\right] + 2E\left[\left(\frac{\chi_q^2}{K}\right)\tilde{Q}\right] + 2E\left[\tilde{L}\tilde{Q}\right] \\ &= E\left[\left(\frac{\chi_q^2}{K}\right)^2\right] + E\left[\tilde{L}^2\right] + 2E\left[\left(\frac{\chi_q^2}{K}\right)\tilde{L}\right] + 2E\left[\left(\frac{\chi_q^2}{K}\right)\tilde{Q}\right] \\ & \quad + \text{small orders} \\ &= \frac{q(q+2)}{K^2} + \left\{E[L^2] \times \left(1 + \frac{q(q+2)}{K^2} + \frac{2q}{K}\right) + \left(\frac{q(q+2)}{K^2}\right) \times E[l_1^2] - 2E\left[L \cdot l_1\right] \times \left(\frac{q}{K} + \frac{q(q+2)}{K^2}\right)\right\} \\ & \quad + \frac{2q}{K} \cdot E\left[\tilde{L}\right] + \frac{2q}{K} E\left[\tilde{Q}\right] + \text{small orders} \\ &= \frac{q(q+2)}{K^2} + \left\{E[L^2] \times \left(1 + \frac{q(q+2)}{K^2} + \frac{2q}{K}\right) + \left(\frac{q(q+2)}{K^2}\right) \times E[l_1^2] - 2E\left[L \cdot l_1\right] \times \left(\frac{q}{K} + \frac{q(q+2)}{K^2}\right)\right\} \\ & \quad + \frac{2q}{K} \cdot \left(E\left[L \times \left(1 + \frac{\chi_q^2}{K}\right) - \frac{\chi_q^2}{K} \cdot l_1\right]\right) + \frac{2q}{K} \cdot E\left(\left(1 + \frac{\chi_q^2}{K}\right) Q + \frac{\chi_q^2}{K} \cdot l_2\right) + \text{small orders} \end{aligned}$$

We have that

$$\begin{aligned} E\left[\left(\frac{\chi_q^2}{K}\right)^2\right] &= \frac{1}{K^2} E\left[(\chi_q^2)^2\right] = \frac{1}{K^2}(Var(\chi_q^2) + \left(E\left[\chi_q^2\right]\right)^2) \\ &= \frac{1}{K^2}(2q + q^2) = \frac{q(q+2)}{K^2}, \end{aligned}$$

and

$$E\left[\left(1+\frac{\chi_q^2}{K}\right)^2\right] = 1+E\left[\left(\frac{\chi_q^2}{K}\right)^2\right]+2E\left[\frac{\chi_q^2}{K}\right]$$
$$= 1+\frac{q(q+2)}{K^2}+\frac{2q}{K},$$

and

$$E\left[\left(1+\frac{\chi_q^2}{K}\right)\cdot\left(\frac{\chi_q^2}{K}\right)\right]=\frac{q}{K}+\frac{q(q+2)}{K^2}.$$

We now want to express the terms $E[l_1^2]$ and $2E\left[L\cdot l_1\right]$. For $E[l_1^2]$, we have that

$$\begin{aligned}
E[l_1^2] &= E\left[\left(\frac{1}{\|Z_q\|^2}\left[Z_q'\Omega_{22}^{-1/2\prime}\right]\otimes\left[Z_q'\Omega_{22}^{-1/2\prime}\right]\left(vec(\hat{\Omega}_{22})-vec(\Omega_{22})\right)\right)^2\right] \\
&= E\left[\frac{1}{\|Z_q\|^2}\left[Z_q'\Omega_{22}^{-1/2\prime}\right]\otimes\left[Z_q'\Omega_{22}^{-1/2\prime}\right]\left(vec(\hat{\Omega}_{22})-vec(\Omega_{22})\right)\right. \\
&\quad \left.\times\left(vec(\hat{\Omega}_{22})-vec(\Omega_{22})\right)'\frac{1}{\|Z_q\|^2}\left[\Omega_{22}^{-1/2}Z_q\right]\otimes\left[\Omega_{22}^{-1/2}Z_q\right]\right] \\
&= \frac{2}{K}(1+o(1))+o\left(\frac{K^2}{T^2}\right)+O\left(\frac{1}{T}\right)
\end{aligned}$$

where the last equation follows by

$$\begin{aligned}
& E\left[\frac{1}{\|Z_q\|^2}\left[Z_q'\Omega_{22}^{-1/2\prime}\right]\otimes\left[Z_q'\Omega_{22}^{-1/2\prime}\right]\left(\frac{1}{K}(\Omega_{22}\otimes\Omega_{22})\right)\frac{1}{\|Z_q\|^2}\left[\Omega_{22}^{-1/2}Z_q\right]\otimes\left[\Omega_{22}^{-1/2}Z_q\right]\right] \\
& \times(1+o(1))+o\left(\frac{K^2}{T^2}\right)+O\left(\frac{1}{T}\right) \\
=\; & \frac{1}{K}E\left[\frac{1}{\|Z_q\|^2}\left[Z_q'\Omega_{22}^{1/2}\right]\otimes\left[Z_q'\Omega_{22}^{1/2}\right]\frac{1}{\|Z_q\|^2}\left[\Omega_{22}^{-1/2}Z_q\right]\otimes\left[\Omega_{22}^{-1/2}Z_q\right]\right]\times(1+o(1))+o\left(\frac{K^2}{T^2}\right)+O\left(\frac{1}{T}\right) \\
=\; & \frac{1}{K}E\left[\frac{1}{\|Z_q\|^4}\left[Z_q'Z_q\right]\otimes\left[Z_q'\Omega_{22}^{1/2}\Omega_{22}^{-1/2}Z_q\right]\right]\times(1+o(1))+o\left(\frac{K^2}{T^2}\right)+O\left(\frac{1}{T}\right) \\
=\; & \frac{1}{K}E\left[\frac{1}{\|Z_q\|^4}\left[\|Z_q\|^2\right]\otimes\left[\|Z_q\|^2\right]\right]\times(1+o(1))+o\left(\frac{K^2}{T^2}\right)+O\left(\frac{1}{T}\right) \\
=\; & \frac{1}{K}(1+o(1))+o\left(\frac{K^2}{T^2}\right)+O\left(\frac{1}{T}\right).
\end{aligned}$$

Next, we have that

$$\begin{aligned}
E[L\cdot l] &= E\left[\left(\frac{1}{\Omega_{11\cdot 2}}\times\left[1,-2\left(\Omega_{12}\Omega_{22}^{-1}\right),\left(\Omega_{12}\Omega_{22}^{-1}\otimes\Omega_{12}\Omega_{22}^{-1}\right)\right]\times(\Lambda-\hat{\Lambda})\right)\right. \\
&\quad \left.\times\begin{pmatrix} 0 & 0 & \left[Z_q'\Omega_{22}^{-1/2\prime}\right]\otimes\left[Z_q'\Omega_{22}^{-1/2\prime}\right]/\|Z_q\|^2 \end{pmatrix}\times(\Lambda-\hat{\Lambda})\right] \\
&= \frac{1}{2\Omega_{11\cdot 2}}E\left[\tilde{l}_{13}\times(vec(\hat{\Omega}_{22})-vec(\Omega_{22}))\right]
\end{aligned}$$

where

$$\tilde{l}_{13} = \frac{(\hat{\Omega}_{11}-\Omega_{11})}{\|Z_q\|^2}\left[Z_q'\Omega_{22}^{-1/2\prime}\right]\otimes\left[Z_q'\Omega_{22}^{-1/2\prime}\right] - \frac{2(\hat{\Omega}_{12}-\Omega_{12})\Omega_{22}^{-1}\Omega_{21}}{\|Z_q\|^2}\left[Z_q'\Omega_{22}^{-1/2\prime}\right]\otimes\left[Z_q'\Omega_{22}^{-1/2\prime}\right]$$
$$+\frac{(vec(\hat{\Omega}_{22})-vec(\Omega_{22}))'}{\|Z_q\|^2}\left(\Omega_{22}^{-1}\Omega_{21}\otimes\Omega_{22}^{-1}\Omega_{21}\right)\left[Z_q'\Omega_{22}^{-1/2\prime}\right]\otimes\left[Z_q'\Omega_{22}^{-1/2\prime}\right],$$

and we have that

$$E\left[\tilde{l}_{13}\times(vec(\hat{\Omega}_{22})-vec(\Omega_{22}))\right] = E[\tilde{l}_{13,1}]+E[\tilde{l}_{13,2}]+E[\tilde{l}_{13,3}],$$

where

$$E[\tilde{l}_{13,1}] = E\left[\frac{(\hat{\Omega}_{11}-\Omega_{11})}{\|Z_q\|^2}\left[Z_q'\Omega_{22}^{-1/2\prime}\right]\otimes\left[Z_q'\Omega_{22}^{-1/2\prime}\right](vec(\hat{\Omega}_{22})-vec(\Omega_{22}))\right]$$
$$E[\tilde{l}_{13,2}] = -E\left[\frac{2(\hat{\Omega}_{12}-\Omega_{12})\Omega_{22}^{-1}\Omega_{21}}{\|Z_q\|^2}\left[Z_q'\Omega_{22}^{-1/2\prime}\right]\otimes\left[Z_q'\Omega_{22}^{-1/2\prime}\right](vec(\hat{\Omega}_{22})-vec(\Omega_{22}))\right]$$
$$E[\tilde{l}_{13,3}] = E\left[\frac{(vec(\hat{\Omega}_{22})-vec(\Omega_{22}))'}{\|Z_q\|^2}\left(\Omega_{22}^{-1}\Omega_{21}\otimes\Omega_{22}^{-1}\Omega_{21}\right)\left[Z_q'\Omega_{22}^{-1/2\prime}\right]\otimes\left[Z_q'\Omega_{22}^{-1/2\prime}\right](vec(\hat{\Omega}_{22})-vec(\Omega_{22}))\right].$$

We now provide term by term by expressions for $\tilde{l}_{13,1}$, $\tilde{l}_{13,2}$, and $\tilde{l}_{13,3}$. For $E[\tilde{l}_{13,1}]$, we have that

$$\begin{aligned}
E[\tilde{l}_{13,1}] &= -E\left[\frac{(\hat{\Omega}_{11}-\Omega_{11})}{\|Z_q\|^2}\left[Z_q'\Omega_{22}^{-1/2\prime}\right]\otimes\left[Z_q'\Omega_{22}^{-1/2\prime}\right](vec(\hat{\Omega}_{22})-vec(\Omega_{22}))\right]\\
&= -\mathrm{tr}E\left[\frac{1}{\|Z_q\|^2}\left[Z_q'\Omega_{22}^{-1/2\prime}\right]\otimes\left[Z_q'\Omega_{22}^{-1/2\prime}\right](vec(\hat{\Omega}_{22})-vec(\Omega_{22}))(\hat{\Omega}_{11}-\Omega_{11})\right]\\
&= -E\left[\frac{1}{\|Z_q\|^2}\left[Z_q'\Omega_{22}^{-1/2\prime}\right]\otimes\left[Z_q'\Omega_{22}^{-1/2\prime}\right]\right]\times E\left[(vec(\hat{\Omega}_{22})-vec(\Omega_{22}))(\hat{\Omega}_{11}-\Omega_{11})\right]\\
&= -E\left[\frac{1}{\|Z_q\|^2}\left[Z_q'\Omega_{22}^{-1/2\prime}\right]\otimes\left[Z_q'\Omega_{22}^{-1/2\prime}\right]\right]\times\left(\frac{2}{K}(\Omega_{21}\otimes\Omega_{21})(1+o(1))\right)+o\left(\frac{K^2}{T^2}\right)+O\left(\frac{1}{T}\right)\\
&= -\frac{2}{K}E\left[\frac{1}{\|Z_q\|^2}\left(Z_q'\Omega_{22}^{-1/2\prime}\Omega_{21}\right)\otimes\left(Z_q'\Omega_{22}^{-1/2\prime}\Omega_{21}\right)\right](1+o(1))+o\left(\frac{K^2}{T^2}\right)+O\left(\frac{1}{T}\right)\\
&= -\frac{2}{K}E\left[\frac{1}{\|Z_q\|^2}\left(Z_q'\Omega_{22}^{-1/2\prime}\Omega_{21}\right)'\times\left(Z_q'\Omega_{22}^{-1/2\prime}\Omega_{21}\right)\right](1+o(1))\\
&\quad +o\left(\frac{K^2}{T^2}\right)+O\left(\frac{1}{T}\right)\\
&= -\frac{\Omega_{12}\Omega_{22}^{-1}\Omega_{12}}{K}(1+o(1))+o\left(\frac{K^2}{T^2}\right)+O\left(\frac{1}{T}\right).
\end{aligned}$$

Next, we have that

$$\begin{aligned}
E[\tilde{l}_{13,2}] &= -E\left[\frac{2(\hat{\Omega}_{12}-\Omega_{12})\Omega_{22}^{-1}\Omega_{21}}{\|Z_q\|^2}\left[Z_q'\Omega_{22}^{-1/2\prime}\right]\otimes\left[Z_q'\Omega_{22}^{-1/2\prime}\right](vec(\hat{\Omega}_{22})-vec(\Omega_{22}))\right] \\
&= -\mathrm{tr}E\left[\frac{2\Omega_{22}^{-1}\Omega_{21}}{\|Z_q\|^2}\left[Z_q'\Omega_{22}^{-1/2\prime}\right]\otimes\left[Z_q'\Omega_{22}^{-1/2\prime}\right](vec(\hat{\Omega}_{22})-vec(\Omega_{22}))(\hat{\Omega}_{12}-\Omega_{12})\right] \\
&= -\frac{2}{K}\mathrm{tr}E\left[\left((\Omega_{22}\otimes\Omega_{12})\left(I_{q^2}+\mathbb{K}_{qq}\right)\right)\frac{1}{\|Z_q\|^2}\left[\Omega_{22}^{-1/2}Z_q\right]\otimes\left[\Omega_{22}^{-1/2}Z_q\right]\Omega_{12}\Omega_{22}^{-1}\right] \\
&\quad\times(1+o(1))+o\left(\frac{K^2}{T^2}\right)+O\left(\frac{1}{T}\right),
\end{aligned}$$

and

$$\begin{aligned}
&-\frac{2}{K}\mathrm{tr}E\left[(\Omega_{22}\otimes\Omega_{12})\frac{1}{\|Z_q\|^2}\left[\Omega_{22}^{-1/2}Z_q\right]\otimes\left[\Omega_{22}^{-1/2}Z_q\right]\Omega_{12}\Omega_{22}^{-1}\right] \\
&= -\frac{2}{K}\mathrm{tr}E\left[\frac{1}{\|Z_q\|^2}\left(\Omega_{22}^{1/2}Z_q\otimes\Omega_{12}\Omega_{22}^{-1/2}Z_q\right)\Omega_{12}\Omega_{22}^{-1}\right] \\
&= -\frac{1}{K}\mathrm{tr}\Omega_{21}\Omega_{12}\Omega_{22}^{-1}=-\frac{1}{K}\Omega_{12}\Omega_{22}^{-1}\Omega_{21}.
\end{aligned}$$

Thus, we have that

$$E[\tilde{l}_{13,1}]=-\frac{2}{K}\Omega_{12}\Omega_{22}^{-1}\Omega_{21}(1+o(1))+o\left(\frac{K^2}{T^2}\right)+O\left(\frac{1}{T}\right)$$

For the term $\tilde{l}_{13,3}$, we have that

$$\begin{aligned}
\tilde{l}_{13,3} &= E\left[\frac{(vec(\hat{\Omega}_{22})-vec(\Omega_{22})'}{\|Z_q\|^2}\left(\Omega_{22}^{-1}\Omega_{21}\otimes\Omega_{22}^{-1}\Omega_{21}\right)\left(\left[Z_q'\Omega_{22}^{-1/2\prime}\right]\otimes\left[Z_q'\Omega_{22}^{-1/2\prime}\right]\right)(vec(\hat{\Omega}_{22})-vec(\Omega_{22}))\right] \\
&= \mathrm{tr}E\left[\frac{1}{\|Z_q\|^2}\left(\Omega_{22}^{-1}\Omega_{21}\otimes\Omega_{22}^{-1}\Omega_{21}\right)\left(\left[Z_q'\Omega_{22}^{-1/2\prime}\right]\otimes\left[Z_q'\Omega_{22}^{-1/2\prime}\right]\right)\right] \\
&\quad\times E\left[(vec(\hat{\Omega}_{22})-vec(\Omega_{22}))(vec(\hat{\Omega}_{22})-vec(\Omega_{22})'\right] \\
&= \mathrm{tr}\left(E\left[\frac{1}{\|Z_q\|^2}\left(\Omega_{22}^{-1}\Omega_{21}\otimes\Omega_{22}^{-1}\Omega_{21}\right)\left(\left[Z_q'\Omega_{22}^{-1/2\prime}\right]\otimes\left[Z_q'\Omega_{22}^{-1/2\prime}\right]\right)\right]\times(\Omega_{22}\otimes\Omega_{22})\left(I_{q^2}+\mathbb{K}_{qq}\right)(1+o(1))\right) \\
&\quad+o\left(\frac{K^2}{T^2}\right)+O\left(\frac{1}{T}\right) \\
&= \frac{1}{K}\mathrm{tr}\left(E\left[\frac{1}{\|Z_q\|^2}\left(\Omega_{22}^{-1}\Omega_{21}Z_q'\Omega_{22}^{1/2}\otimes\Omega_{22}^{-1}\Omega_{21}Z_q'\Omega_{22}^{1/2}\right)\right]\left(I_{q^2}+\mathbb{K}_{qq}\right)\right)(1+o(1))+o\left(\frac{K^2}{T^2}\right)+O\left(\frac{1}{T}\right).
\end{aligned}$$

We have that

$$\begin{aligned}
&E\mathrm{tr}\left[\frac{1}{\|Z_q\|^2}\left(\Omega_{22}^{-1}\Omega_{21}Z_q'\Omega_{22}^{1/2}\otimes\Omega_{22}^{-1}\Omega_{21}Z_q'\Omega_{22}^{1/2}\right)\right] \\
&= E\left[\frac{1}{\|Z_q\|^2}\mathrm{tr}(\Omega_{22}^{-1}\Omega_{21}Z_q'\Omega_{22}^{1/2})\times\mathrm{tr}(\Omega_{22}^{-1}\Omega_{21}Z_q'\Omega_{22}^{1/2})\right],
\end{aligned}$$

and

$$
\begin{aligned}
\mathrm{tr}(\Omega_{22}^{-1}\Omega_{21}Z_q'\Omega_{22}^{1/2}) &= \mathrm{tr}(\Omega_{21}Z_q'\Omega_{22}^{1/2}\Omega_{22}^{-1}) = \mathrm{tr}(\Omega_{21}Z_q'\Omega_{22}^{-1/2}) \\
&= \mathrm{tr}(Z_q'\Omega_{22}^{-1/2}\Omega_{21}) = (Z_q'\Omega_{22}^{-1/2}\Omega_{21})' \\
&= \Omega_{12}\Omega_{22}^{-1/2}Z_q,
\end{aligned}
$$

which leads us to have

$$
\begin{aligned}
& E\left[\frac{1}{\|Z_q\|^2}\mathrm{tr}(\Omega_{22}^{-1}\Omega_{21}Z_q'\Omega_{22}^{1/2}) \times \mathrm{tr}(\Omega_{22}^{-1}\Omega_{21}Z_q'\Omega_{22}^{1/2})\right] \\
= \; & E\left[\frac{1}{\|Z_q\|^2}(\Omega_{12}\Omega_{22}^{-1/2}Z_q) \times (Z_q'\Omega_{22}^{-1/2}\Omega_{21})\right] \\
= \; & \frac{1}{2}\Omega_{12}\Omega_{22}^{-1}\Omega_{21},
\end{aligned}
$$

and

$$
\begin{aligned}
\tilde{l}_{13,3} &= \frac{2}{K} \times \frac{1}{2}\Omega_{12}\Omega_{22}^{-1}\Omega_{21}(1+o(1)) + o\left(\frac{K^2}{T^2}\right) + O\left(\frac{1}{T}\right). \\
&= \frac{1}{K}\Omega_{12}\Omega_{22}^{-1}\Omega_{21}(1+o(1)) + o\left(\frac{K^2}{T^2}\right) + O\left(\frac{1}{T}\right).
\end{aligned}
$$

Combining previous results, we have that

$$
\begin{aligned}
E\left[L \cdot l_1\right] &= \frac{1}{2\Omega_{11\cdot 2}} E\left[\tilde{l}_{13} \times (vec(\hat{\Omega}_{22}) - vec(\Omega_{22}))\right] \\
&= \frac{1}{2\Omega_{11\cdot 2}}\left(E[\tilde{l}_{13,1}] + E[\tilde{l}_{13,2}] + E[\tilde{l}_{13,3}]\right) \\
&= \frac{1}{2\Omega_{11\cdot 2}}\left(\left(-\frac{\Omega_{12}\Omega_{22}^{-1}\Omega_{12}}{K}\right) + \left(-\frac{2}{K}\Omega_{12}\Omega_{22}^{-1}\Omega_{21}\right) + \left(\frac{1}{K}\Omega_{12}\Omega_{22}^{-1}\Omega_{21}\right)\right) \\
&\quad \times (1+o(1)) + o\left(\frac{K^2}{T^2}\right) + O\left(\frac{1}{T}\right) \\
&= \frac{1}{2\Omega_{11\cdot 2}}\frac{-2}{K}\Omega_{12}\Omega_{22}^{-1}\Omega_{12}(1+o(1)) + o\left(\frac{K^2}{T^2}\right) + O\left(\frac{1}{T}\right) \\
&= -\frac{\Omega_{12}\Omega_{22}^{-1}\Omega_{12}}{K\Omega_{11\cdot 2}}(1+o(1)) + o\left(\frac{K^2}{T^2}\right) + O\left(\frac{1}{T}\right).
\end{aligned}
$$

Using these results, we provide the expression for the term (B) as follows:

$$
\begin{aligned}
& E\left[\left(\frac{\chi_q^2}{K}+\left\{L\times\left(1+\frac{\chi_q^2}{K}\right)-\frac{\chi_q^2}{K}\cdot l_1\right\}+\left\{\left(1+\frac{\chi_q^2}{K}\right)Q+\frac{\chi_q^2}{K}\cdot l_2\right\}\right)^2\right] \\
= \; & \frac{q(q+2)}{K^2}+\left\{E[L^2]\times\left(1+\frac{q(q+2)}{K^2}+\frac{2q}{K}\right)+\left(\frac{q(q+2)}{K^2}\right)\times E[l_1^2]-2E\left[L\cdot l_1\right]\times\left(\frac{q}{K}+\frac{q(q+2)}{K^2}\right)\right\} \\
& +\frac{2q}{K}\cdot\left(E\left[L\times\left(1+\frac{\chi_q^2}{K}\right)-\frac{\chi_q^2}{K}\cdot l_1\right]\right)+\frac{2q}{K}\cdot E\left(\left(1+\frac{\chi_q^2}{K}\right)Q+\frac{\chi_q^2}{K}\cdot l_2\right)+\text{small orders} \\
= \; & \frac{q(q+2)}{K^2}+E[L^2]\times\left(1+\frac{q(q+2)}{K^2}+\frac{2q}{K}\right) \\
& +\left(\frac{q(q+2)}{K^2}\right)\times\frac{2}{K}(1+o(1))+o\left(\frac{K^2}{T^2}\right)+O\left(\frac{1}{T}\right) \\
& -2\times\left(-\frac{\Omega_{12}\Omega_{22}^{-1}\Omega_{12}}{K\Omega_{11\cdot 2}}(1+o(1))+o\left(\frac{K^2}{T^2}\right)+O\left(\frac{1}{T}\right)\right)\times\left(\frac{q}{K}+\frac{q(q+2)}{K^2}\right) \\
& +\frac{2q}{K}\cdot\left(1+\frac{q}{K}\right)E[L]-\frac{2q}{K}\cdot\left(\frac{q}{K}\right)E[l_1]+\frac{2q}{K}\left(1+\frac{q}{K}\right)E[Q]+\frac{2q}{K}\left(\frac{q}{K}\right)E[l_2]+\text{small orders.} \\
= \; & \left(1+\frac{2q}{K}\right)E[L^2]+\frac{2q}{K}\left(E[L]+E[Q]\right)+o\left(\frac{1}{K}\right)+o\left(\frac{K^2}{T^2}\right)+O\left(\frac{1}{T}\right)
\end{aligned}
$$

Finally, we have that

$$
\begin{aligned}
P\left(|\tilde{\mathbb{T}}_T(\hat{\theta}_{T,GLS})|\leq z\right) \; = \; & E\left[zG_1\left(\Xi^{-1}\cdot\left(1+\frac{\chi_q^2}{K}\cdot\gneqq\right)\right)\right]+O\left(\frac{1}{T}\right) \\
= \; & G_1(z)+G_1'(z)z\left(\frac{q}{K}\right)+\left(1+\frac{q}{K}\right)G_1'(z)z\left(\left(E[L]+E[Q]\right)\right) \\
& +\frac{1}{2}G_1''(z)z^2E[L^2]\times\left(1+\frac{2q}{K}\right)+\frac{1}{2}G_1''(z)z^2\left(\frac{2q}{K}\left(E[L]+E[Q]\right)\right) \\
& +o\left(\frac{1}{K}\right)+o\left(\frac{K^2}{T^2}\right)+O\left(\frac{1}{T}\right) \\
= \; & G_1(z)+G_1'(z)z\left(E[L]+E[Q]\right)+\frac{1}{2}G_1''(z)z^2E[L^2] \\
& +\left(\frac{q}{K}\right)G_1'(z)z+\left(\frac{q}{K}\right)G_1'(z)z\left(\left(E[L]+E[Q]\right)\right) \\
& +\left(\frac{2q}{K}\right)\frac{1}{2}G_1''(z)z^2E[L^2]+\left(\frac{2q}{K}\right)\frac{1}{2}G_1''(z)z^2\left(E[L]+E[Q]\right) \\
& +o\left(\frac{1}{K}\right)+o\left(\frac{K^2}{T^2}\right)+O\left(\frac{1}{T}\right)
\end{aligned}
$$

where

$$
\begin{aligned}
E[L] &= \frac{K^2}{T^2}\tilde{B}+o\left(\frac{K^2}{T^2}\right)+O\left(\frac{1}{T}\right); \\
E[Q] &= -\frac{2q}{K}(1+o(1))+o\left(\frac{K^2}{T^2}\right)+O\left(\frac{1}{T}\right),
\end{aligned}
$$

and

$$
\begin{aligned}
E[L^2] &= \frac{2}{K}(1+o(1)) + o\left(\frac{K^2}{T^2}\right) + O\left(\frac{1}{T}\right); \\
E[Q^2] &= o\left(\frac{1}{K} + \frac{K^2}{T^2}\right) + O\left(\frac{1}{T}\right),
\end{aligned}
$$

and thus

$$
\begin{aligned}
&\approx \Xi^{-1}(\Lambda)\cdot\left(1+\frac{\chi_q^2}{K}\cdot\Delta(\Lambda)\right) + \nabla'\left\{\Xi^{-1}(\hat{\Lambda})\left(1+\frac{\chi_q^2}{K}\cdot\gneqq(\hat{\Lambda})\right)\right\}\Bigg|_{\hat{\Lambda}=\Lambda}(\hat{\Lambda}-\Lambda) \\
&\quad +\frac{1}{2}(\hat{\Lambda}-\Lambda)'\left\{\nabla^2\Xi^{-1}(\hat{\Lambda})\left(1+\frac{\chi_q^2}{K}\cdot\gneqq(\hat{\Lambda})\right)\Bigg|_{\hat{\Lambda}=\Lambda}\right\}(\hat{\Lambda}-\Lambda)
\end{aligned}
$$

$$
\begin{aligned}
&\left(\Xi(\hat{\Lambda})\right)^{-1}\cdot\left(1+\frac{\chi_q^2}{K}\cdot\gneqq(\hat{\Lambda})\right) \\
&\approx \Xi^{-1}(\Lambda)\cdot\left(1+\frac{\chi_q^2}{K}\cdot\Delta(\Lambda)\right) + \nabla'\left\{\Xi^{-1}(\hat{\Lambda})\left(1+\frac{\chi_q^2}{K}\cdot\gneqq(\hat{\Lambda})\right)\right\}\Bigg|_{\hat{\Lambda}=\Lambda}(\hat{\Lambda}-\Lambda) \\
&\quad +\frac{1}{2}(\hat{\Lambda}-\Lambda)'\left\{\nabla^2\Xi^{-1}(\hat{\Lambda})\left(1+\frac{\chi_q^2}{K}\cdot\gneqq(\hat{\Lambda})\right)\Bigg|_{\hat{\Lambda}=\Lambda}\right\}(\hat{\Lambda}-\Lambda),
\end{aligned}
$$

where

$$
\begin{aligned}
&\nabla'\left\{\Xi^{-1}(\hat{\Lambda})\left(1+\frac{\chi_q^2}{K}\cdot\Delta(\hat{\Lambda})\right)\right\}\Bigg|_{\hat{\Lambda}=\Lambda}(\hat{\Lambda}-\Lambda) \\
&= \left\{\frac{1}{\Omega_{11\cdot 2}}\times\left[1,-2\left(\Omega_{12}\Omega_{22}^{-1}\right),\left(\Omega_{12}\Omega_{22}^{-1}\otimes\Omega_{12}\Omega_{22}^{-1}\right)\right]\times\left(1+\frac{\chi_q^2}{K}\right)\right. \\
&\quad \left. -\frac{\chi_q^2}{K}\cdot\begin{pmatrix} 0 & 0 & \left[Z_q'\Omega_{22}^{-1/2\prime}\right]\otimes\left[Z_q'\Omega_{22}^{-1/2\prime}\right]/\|Z_q\|^2 \end{pmatrix}\right\} \\
&\quad \times\begin{bmatrix} \hat{\Omega}_{11}-\Omega_{11}, \\ \hat{\Omega}_{21}-\Omega_{21} \\ vec(\hat{\Omega}_{22})-vec(\Omega_{22}) \end{bmatrix} \\
&= \left(1+\frac{\chi_q^2}{K}\right)\left(\frac{1}{\Omega_{11\cdot 2}}\times\left[1,-2\left(\Omega_{12}\Omega_{22}^{-1}\right),\left(\Omega_{12}\Omega_{22}^{-1}\otimes\Omega_{12}\Omega_{22}^{-1}\right)\right]\times\begin{bmatrix} \hat{\Omega}_{11}-\Omega_{11}, \\ \hat{\Omega}_{21}-\Omega_{21} \\ vec(\hat{\Omega}_{22})-vec(\Omega_{22}) \end{bmatrix}\right) \\
&\quad -\frac{\chi_q^2}{K}\cdot\begin{pmatrix} 0 & 0 & \left[Z_q'\Omega_{22}^{-1/2\prime}\right]\otimes\left[Z_q'\Omega_{22}^{-1/2\prime}\right]/\|Z_q\|^2 \end{pmatrix}\times\begin{bmatrix} \hat{\Omega}_{11}-\Omega_{11}, \\ \hat{\Omega}_{21}-\Omega_{21} \\ vec(\hat{\Omega}_{22})-vec(\Omega_{22}) \end{bmatrix}.
\end{aligned}
$$

Using this results, we have that Thus, we have that

$$\begin{aligned} P\left(|\tilde{\mathbb{T}}_T(\hat{\theta}_{T,GLS})| \le z\right) &= G_1(z) + G_1'(z)z\left(\frac{K^2}{T^2}\tilde{B} - \frac{2q}{K}(1+o(1)) + o\left(\frac{K^2}{T^2}\right) + O\left(\frac{1}{T}\right)\right) \\ &\quad + \frac{1}{2}G_1''(z)z^2\left(\frac{2}{K}(1+o(1)) + o\left(\frac{K^2}{T^2}\right) + O\left(\frac{1}{T}\right)\right) \\ &\quad + \left(\frac{q}{K}\right)G_1'(z)z(1+o(1)) + o\left(\frac{1}{K}\right) + o\left(\frac{K^2}{T^2}\right) + O\left(\frac{1}{T}\right) \\ &= G_1(z) + G_1'(z)z\frac{K^2}{T^2}\tilde{B} - G_1'(z)z\frac{q}{K}(1+o(1)) + \frac{1}{K}G_1''(z)z^2 \\ &\quad + o\left(\frac{1}{K}\right) + o\left(\frac{K^2}{T^2}\right) + O\left(\frac{1}{T}\right), \end{aligned}$$

which implies the following result

$$\begin{aligned} P_{H_0}\left(\left|\mathbb{T}_T(\hat{\theta}_2;\hat{\theta}_1)\right| \le z\right) &= G_1(z) + G_1'(z)z\tilde{B}\left(\frac{K^2}{T^2}\right) - 2G_1'(z)zq\left(\frac{1}{K}\right) + G_1''(z)z^2\left(\frac{1}{K}\right) \\ &\quad + O\left(\frac{\log T}{\sqrt{T}}\right) + o\left(\frac{K^2}{T^2}\right) + o\left(\frac{1}{K}\right), \end{aligned}$$

as desired. ■

Proof of Theorem 6.

Let $\tilde{\mathbb{F}}_{1-\alpha}$ be the $(1-\alpha)$ quantile of $\tilde{\mathbb{F}}_\infty = (K^{-1})(\chi_1^2/\chi_{K-q}^2)$ and $\chi_1^2 \perp \chi_{K-q}^2$. Note that

$$\tilde{\mathbb{F}}_\infty = K \cdot \frac{\chi_1^2}{\chi_{K-q}^2} = \mathcal{F}_{1,K-q} \cdot \left(\frac{K}{K-q}\right).$$

Using (8), we have that

$$\begin{aligned} \tilde{\mathbb{F}}_{1-\alpha} - \chi_{1,1-\alpha}^2 &= \left(\frac{K}{K-q} - 1\right)\chi_{1,1-\alpha}^2 - \left(\frac{1}{K-q}\right)\frac{G_1''(\chi_{1,1-\alpha}^2)}{G_1'(\chi_{1,1-\alpha}^2)}\left(\chi_{1,1-\alpha}^2\right)^2 + o\left(\frac{1}{K}\right) \\ &= \left(\frac{q}{K-q}\right)\chi_{1,1-\alpha}^2 - \left(\frac{1}{K-q}\right)\frac{G_1''(\chi_{1,1-\alpha}^2)}{G_1'(\chi_{1,1-\alpha}^2)}\left(\chi_{1,1-\alpha}^2\right)^2 + o\left(\frac{1}{K}\right) \\ &= O\left(\frac{1}{K}\right), \end{aligned}$$

which implies that

$$\begin{aligned} G_1\left(\tilde{\mathbb{F}}_{1-\alpha}\right) &= G_1\left(\chi_{1,1-\alpha}^2\right) + G_1'\left(\chi_{1,1-\alpha}^2\right)\left(\tilde{\mathbb{F}}_{1-\alpha} - \chi_{1,1-\alpha}^2\right) + o\left(\frac{1}{K}\right) \\ &= (1-\alpha) + G_1'\left(\chi_{1,1-\alpha}^2\right)\left(\tilde{\mathbb{F}}_{1-\alpha} - \chi_{1,1-\alpha}^2\right) + o\left(\frac{1}{K}\right); \\ G_1'(\tilde{\mathbb{F}}_{1-\alpha}) &= G_1'(\chi_{1,1-\alpha}^2) + G_1''(\chi_{1,1-\alpha}^2)\left(\tilde{\mathbb{F}}_{1-\alpha} - \chi_{1,1-\alpha}^2\right) + o\left(\frac{1}{K}\right) \\ &= G_1'(\chi_{1,1-\alpha}^2) + O\left(\frac{1}{K}\right) \\ G_1''(\tilde{\mathbb{F}}_{1-\alpha}) &= G_1''(\chi_{1,1-\alpha}^2) + O\left(\frac{1}{K}\right). \end{aligned}$$

Also, the result in Lemma 5 implies that

$$
\begin{aligned}
P\left(|\tilde{\mathbb{T}}_T(\hat{\theta}_2;\hat{\theta}_1)|^2 \le \tilde{\mathbb{F}}_{1-\alpha}\right) &= G_1\left(\tilde{\mathbb{F}}_{1-\alpha}\right) + G_1'(\tilde{\mathbb{F}}_{1-\alpha})\tilde{\mathbb{F}}_{1-\alpha}\frac{K^2}{T^2}\tilde{B} - G_1'(\tilde{\mathbb{F}}_{1-\alpha})\tilde{\mathbb{F}}_{1-\alpha}\frac{q}{K}(1+o(1)) \\
&\quad + \frac{1}{K}G_1''(\tilde{\mathbb{F}}_{1-\alpha})\tilde{\mathbb{F}}_{1-\alpha}^2 + o\left(\frac{1}{K}\right) + o\left(\frac{K^2}{T^2}\right) + O\left(\frac{1}{T}\right) \\
&= (1-\alpha) + G_1'\left(\chi^2_{1,1-\alpha}\right)\chi^2_{1,1-\alpha}\left(\left(\frac{q}{K-q}\right) - \frac{q}{K}\right) + \left(G_1''(\chi^2_{1,1-\alpha})\right)\left((\chi^2_{1,1-\alpha})^2\right) \\
&\quad \times\left(-\left(\frac{1}{K-q}\right) + \frac{1}{K}\right) + \left(G_1'(\chi^2_{1,1-\alpha})\right)\left(\chi^2_{1,1-\alpha}\right)\frac{K^2}{T^2}\tilde{B} \\
&\quad + o\left(\frac{1}{K}\right) + o\left(\frac{K^2}{T^2}\right) + O\left(\frac{1}{T}\right).
\end{aligned}
$$

Note that

$$
\begin{aligned}
\left(\frac{q}{K-q}\right) - \frac{q}{K} &= \frac{Kq-(K-q)q}{(K-q)K} \\
&= \frac{q^2}{(K-q)K} = O\left(\frac{1}{K^2}\right) = o\left(\frac{1}{K}\right),
\end{aligned}
$$

and similarly we have that

$$
-\left(\frac{1}{K-q}\right) + \frac{1}{K} = \frac{-K+K-q}{K(K-q)} = o\left(\frac{1}{K}\right).
$$

Thus, we have that

$$
\begin{aligned}
P\left(|\tilde{\mathbb{T}}_T(\hat{\theta}_{T,GLS})|^2 \le \tilde{\mathbb{F}}_{1-\alpha}\right) &= (1-\alpha) + G_1'\left(\chi^2_{1,1-\alpha}\right)\chi^2_{1,1-\alpha}\left(\left(\frac{q}{K-q}\right) - \frac{q}{K}\right) + \left(G_1''(\chi^2_{1,1-\alpha})\right)\left((\chi^2_{1,1-\alpha})^2\right) \\
&\quad \times\left(-\left(\frac{1}{K-q}\right) + \frac{1}{K}\right) + \left(G_1'(\chi^2_{1,1-\alpha})\right)\left(\chi^2_{1,1-\alpha}\right)\frac{K^2}{T^2}\tilde{B} \\
&\quad + o\left(\frac{1}{K}\right) + o\left(\frac{K^2}{T^2}\right) + O\left(\frac{1}{T}\right) \\
&= (1-\alpha) + \left(G_1'(\chi^2_{1,1-\alpha})\right)\left(\chi^2_{1,1-\alpha}\right)\frac{K^2}{T^2}\tilde{B} + o\left(\frac{1}{K}\right) + o\left(\frac{K^2}{T^2}\right) + O\left(\frac{1}{T}\right),
\end{aligned}
$$

which implies

$$
P_{H_0}\left(\left|\tilde{\mathbb{T}}_T(\hat{\theta}_2;\hat{\theta}_1)\right|^2 > \mathcal{F}^{1-\alpha}_{K-q}\right) = \alpha - \left(\frac{K^2}{T^2}\right)\left(G_1'(\chi^2_{1,1-\alpha})\chi^2_{1,1-\alpha}\right)\tilde{B} + O\left(\frac{\log T}{\sqrt{T}}\right) + o\left(\frac{K^2}{T^2}\right) + o\left(\frac{1}{K}\right),
$$

as desired. ■

References

[1] Andrews, D.W.K. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. Econometrica 59, 817–854.

[2] Anatolyev, S. (2005). GMM, GEL, serial correlation, and asymptotic bias. Econometrica, 73(3), 983-1002.

[3] Christiano, L. J. (1996). Special Section on Small-Sample Properties of Generalized Method of Moments (GMM) Associate Editor's Introduction. Journal of Business & Economic Statistics, 14(3), 261-261.

[4] Hall, A. R. (2005). Generalized method of moments. Oxford university press.

[5] Hall, P., & Horowitz, J. L. (1996). Bootstrap critical values for tests based on generalized-method-of-moments estimators. Econometrica: Journal of the Econometric Society, 891-916.

[6] Hansen, L. P., Heaton, J., & Yaron, A. (1996). Finite-sample properties of some alternative GMM estimators. Journal of Business & Economic Statistics, 14(3), 262-280.

[7] Hansen, L. P., & Singleton, K. J. (1982). Generalized instrumental variables estimation of nonlinear rational expectations models. Econometrica: Journal of the Econometric Society, 1269-1286.

[8] Hwang, J., & Sun, Y. (2017). Asymptotic F and t Tests in an Efficient GMM Setting. Journal of econometrics, 198(2), 277-295.

[9] Hwang, J., & Sun, Y. (2018). Should we go one step further? An accurate comparison of one-step and two-step procedures in a generalized method of moments framework. Journal of econometrics, 207(2), 381-405.

[10] Hwang, J., & Valdés, G. (2022a). Finite-sample corrected inference for two-step GMM in time series. Journal of Econometrics.

[11] Hwang, J., & Valdés, G. (2022b). Higher-order Expansion and Optimal HAR Inference in Over-identified Systems

[12] Kiefer, N. M., & Vogelsang, T. J. (2002). Heteroskedasticity-autocorrelation robust standard errors using the Bartlett kernel without truncation. Econometrica, 70(5), 2093-2095.

[13] Kiefer, N. M., & Vogelsang, T. J. (2005). A new asymptotic theory for heteroskedasticity-autocorrelation robust tests. Econometric Theory, 21(6), 1130-1164.

[14] Kitamura, Y., & Stutzer, M. (1997). An information-theoretic alternative to generalized method of moments estimation. Econometrica: Journal of the Econometric Society, 861-874.

[15] Lazarus, E., Lewis, D. J., & Stock, J. H. (2019). The size-power tradeoff in HAR inference. Available at SSRN 3436372.

[16] Lazarus, E., Lewis, D. J., Stock, J. H., & Watson, M. W. (2018). HAR Inference: Recommendations for Practice. Journal of Business & Economic Statistics, 36(4), 541-559.

[17] Müller, U. K. (2007). A theory of robust long-run variance estimation. Journal of Econometrics, 141(2), 1331-1352.

[18] Müller, U. K., & Watson, M. W. (2018). Long-Run Covariability. Econometrica, 86(3), 775-804.

[19] Newey, W. K., & Smith, R. J. (2004). Higher order properties of GMM and generalized empirical likelihood estimators. Econometrica, 72(1), 219-255.

[20] Phillips, P. C. (2005). Automated discovery in econometrics. Econometric Theory, 21(1), 3-20.

[21] Sun, Y. (2013). A heteroskedasticity and autocorrelation robust F test using an orthonormal series variance estimator. The Econometrics Journal, 16(1), 1-26.

[22] Sun, Y. (2014a). Let's fix it: Fixed-b asymptotics versus small-b asymptotics in heteroskedasticity and autocorrelation robust inference. Journal of Econometrics, 178, 659-677.

[23] Sun, Y. (2014b). Fixed-smoothing asymptotics in a two-step generalized method of moments framework. Econometrica, 82(6), 2327-2370.

[24] Sun, Y., Phillips, P. C., & Jin, S. (2008). Optimal bandwidth selection in heteroskedasticity–autocorrelation robust testing. Econometrica, 76(1), 175-194.

[25] Windmeijer, F. (2005). A finite sample correction for the variance of linear efficient two-step GMM estimators. Journal of econometrics, 126(1), 25-51.

[26] Andrews, D.W.K. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. Econometrica 59, 817–854.

[27] Anatolyev, S. (2005). GMM, GEL, serial correlation, and asymptotic bias. Econometrica, 73(3), 983-1002.

[28] Christiano, L. J. (1996). Special Section on Small-Sample Properties of Generalized Method of Moments (GMM) Associate Editor's Introduction. Journal of Business & Economic Statistics, 14(3), 261-261.

[29] Hall, A. R. (2005). Generalized method of moments. Oxford university press.

[30] Hall, P., & Horowitz, J. L. (1996). Bootstrap critical values for tests based on generalized-method-of-moments estimators. Econometrica: Journal of the Econometric Society, 891-916.

[31] Hansen, L. P., Heaton, J., & Yaron, A. (1996). Finite-sample properties of some alternative GMM estimators. Journal of Business & Economic Statistics, 14(3), 262-280.

[32] Hansen, L. P., & Singleton, K. J. (1982). Generalized instrumental variables estimation of nonlinear rational expectations models. Econometrica: Journal of the Econometric Society, 1269-1286.

[33] Hwang, J., & Sun, Y. (2017). Asymptotic F and t Tests in an Efficient GMM Setting. Journal of econometrics, 198(2), 277-295.

[34] Hwang, J., & Sun, Y. (2018). Should we go one step further? An accurate comparison of one-step and two-step procedures in a generalized method of moments framework. Journal of econometrics, 207(2), 381-405.

[35] Kiefer, N. M., & Vogelsang, T. J. (2002). Heteroskedasticity-autocorrelation robust standard errors using the Bartlett kernel without truncation. Econometrica, 70(5), 2093-2095.

[36] Kiefer, N. M., & Vogelsang, T. J. (2005). A new asymptotic theory for heteroskedasticity-autocorrelation robust tests. Econometric Theory, 21(6), 1130-1164.

[37] Kitamura, Y., & Stutzer, M. (1997). An information-theoretic alternative to generalized method of moments estimation. Econometrica: Journal of the Econometric Society, 861-874.

[38] Lazarus, E., Lewis, D. J., & Stock, J. H. (2019). The size-power tradeoff in HAR inference. Available at SSRN 3436372.

[39] Lazarus, E., Lewis, D. J., Stock, J. H., & Watson, M. W. (2018). HAR Inference: Recommendations for Practice. Journal of Business & Economic Statistics, 36(4), 541-559.

[40] Müller, U. K. (2007). A theory of robust long-run variance estimation. Journal of Econometrics, 141(2), 1331-1352.

[41] Müller, U. K., & Watson, M. W. (2018). Long-Run Covariability. Econometrica, 86(3), 775-804.

[42] Newey, W. K., & Smith, R. J. (2004). Higher order properties of GMM and generalized empirical likelihood estimators. Econometrica, 72(1), 219-255.

[43] Phillips, P. C. (2005). Automated discovery in econometrics. Econometric Theory, 21(1), 3-20.

[44] Sun, Y. (2013). A heteroskedasticity and autocorrelation robust F test using an orthonormal series variance estimator. The Econometrics Journal, 16(1), 1-26.

[45] Sun, Y. (2014a). Let's fix it: Fixed-b asymptotics versus small-b asymptotics in heteroskedasticity and autocorrelation robust inference. Journal of Econometrics, 178, 659-677.

[46] Sun, Y. (2014b). Fixed-smoothing asymptotics in a two-step generalized method of moments framework. Econometrica, 82(6), 2327-2370.

[47] Sun, Y., Phillips, P. C., & Jin, S. (2008). Optimal bandwidth selection in heteroskedasticity–autocorrelation robust testing. Econometrica, 76(1), 175-194.

[48] Windmeijer, F. (2005). A finite sample correction for the variance of linear efficient two-step GMM estimators. Journal of econometrics, 126(1), 25-51.