Further Results on the Limiting Distribution of GMM Sample Moment Conditions

Nikolay Gospodinov*, Raymond Kan†, and Cesare Robotti ‡

First draft: May 2010
This draft: May 2012

*Department of Economics, Concordia University, 1455 de Maisonneuve Blvd. West, Montreal, Quebec, Canada H3G 1M8; E-mail:nikolay.gospodinov@concordia.ca

†Joseph L. Rotman School of Management, University of Toronto, 105 St. George Street, Toronto, Ontario, Canada M5S 3E6; E-mail:kan@chass.utoronto.ca

‡Research Department, Federal Reserve Bank of Atlanta, 1000 Peachtree St. N.E., Atlanta, Georgia, 30309, USA; E-mail:cesare.robotti@atl.frb.org

Further Results on the Limiting Distribution of GMM Sample Moment Conditions

Abstract

In this paper, we examine the limiting behavior of GMM sample moment conditions and point out an important discontinuity that arises in their asymptotic distribution. We show that the part of the scaled sample moment conditions that gives rise to degeneracy in the asymptotic normal distribution is T-consistent and has a non-standard limiting distribution. We derive the appropriate asymptotic (weighted chi-squared) distribution when this degeneracy occurs and show how to conduct asymptotically valid statistical inference. We also propose a new rank test that provides guidance on which (standard or non-standard) asymptotic framework should be used for inference. The finite-sample properties of the proposed asymptotic approximation are demonstrated using simulated data from some popular asset pricing models.

JEL classification: C13, C32, G12.

Keywords: generalized method of moments; asymptotic approximation; T-consistent estimator; weighted chi-square distribution; rank test.

1. INTRODUCTION

Over the past thirty years, the generalized method of moments (GMM) has established itself as arguably the most popular method for estimating economic models defined by a set of moment conditions. In his seminal paper, Hansen (1982) develops the asymptotic distributions of the GMM estimator, sample moment conditions, and test of over-identifying restrictions for possibly nonlinear models with sufficiently general dependence structure. This large sample theory proved to cover a large class of models and estimators that are of interest to researchers in economics and finance.

There are cases, however, in which the root-T convergence and asymptotic normality of the GMM sample moment conditions and estimators based on these moment conditions do not accurately characterize their limiting behavior. In particular, it is possible that some linear combinations of the GMM sample moment conditions have a degenerate distribution and the standard limiting tools for inference are inappropriate. For example, Gospodinov, Kan and Robotti (2010) demonstrate that some GMM estimators, which are functions of the sample moment conditions, are proportional to the GMM objective function and, hence, cannot be root-T consistent and asymptotically normally distributed for correctly specified models. This situation is directly related to the results in Lemma 4.1 and its subsequent discussion in Hansen (1982) which draw attention to the singularity of the covariance matrix of the sample moment conditions. However, to the best of our knowledge, the limiting behavior of the GMM sample moment conditions in the degenerate case, when the covariance matrix reduces to a matrix of zeros, has not been formally investigated in the literature.

In this paper, we study some linear combinations of the sample moment conditions that give rise to degeneracy and analyze their asymptotic behavior. Interestingly, we show that in this case, the scaled sample moment conditions evaluated at the GMM estimator are characterized by a non-standard limiting theory. More specifically, we demonstrate that the estimated GMM moment conditions converge to zero (the value implied by the population moment conditions) at rate T

and have an asymptotic weighted chi-squared distribution. Besides being of academic interest, our results prove to be useful for inference in asset pricing models. Furthermore, a similar degeneracy occurs in the asymptotic analysis of Lagrange multipliers in the context of generalized empirical likelihood (GEL) estimation which has gained increased popularity in econometrics in recent years.

Non-standard asymptotics (with an accelerated rate of convergence and non-normal asymptotic distribution) often arises when the parameter of interest is near or on the boundary of the parameter space. Examples of this phenomenon include autoregressive roots near or on the unit circle (Dickey and Fuller, 1979; Phillips, 1987), (near) unit root in moving average models (Davis and Dunsmuir, 1996), local-to-zero signal-to-noise ratio in time-varying parameter models (Nyblom, 1989) and zero variance in non-nested model comparison tests (Vuong, 1989). In these cases, the rate of convergence of the estimated parameter or test statistic of interest increases from root-T to T and the asymptotic distribution changes from normal to a weighted chi-squared type of distribution. Interestingly, our paper shows that similar non-standard asymptotics can be encountered in seemingly regular GMM problems. It should also be noted that the degeneracy of the asymptotic distribution has been analyzed in other contexts although the existing results differ considerably from ours. See, for example, Sargan's (1959, 1983) discussion of nonlinear instrumental variable models with possible singularity, Park and Phillips (1989) and Sims, Stock and Watson (1990) for the analysis of time series models with a unit root, and Phillips (2007) for models with slowly varying regression functions.

The rest of this paper is organized as follows. Section 2 introduces the general setup and discusses an asset pricing example that illustrates the discontinuity in the asymptotic approximation of the sample moment conditions. This section also provides the main theoretical results on the limiting behavior of linear combinations of sample moment conditions and presents an easy-to-implement rank test that determines which asymptotic approximation should be used. Section 3 reports simulation results based on a problem in empirical asset pricing and Section 4 concludes.

2. ASYMPTOTICS FOR GMM SAMPLE MOMENT CONDITIONS

2.1. NOTATION AND ANALYTICAL FRAMEWORK

Let $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ denote a $p \times 1$ parameter vector of interest with true value $\boldsymbol{\theta}_0$ that lies in the interior of the parameter space $\boldsymbol{\Theta}$ and $\mathbf{g}_t(\boldsymbol{\theta})$ be a known function $\{\mathbf{g} : \mathbb{R}^p \to \mathbb{R}^m, m > p\}$ of the data and $\boldsymbol{\theta}$ that satisfies the set of population orthogonality conditions

$$E[\mathbf{g}_t(\boldsymbol{\theta}_0)] = \mathbf{0}_m. \tag{1}$$

The GMM estimator of $\boldsymbol{\theta}_0$ is defined as

$$\hat{\boldsymbol{\theta}} = \text{argmin}_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \bar{\mathbf{g}}_T(\boldsymbol{\theta})'\mathbf{W}_T\bar{\mathbf{g}}_T(\boldsymbol{\theta}), \tag{2}$$

where $\mathbf{W}_T$ is an $m \times m$ positive definite weight matrix and $\bar{\mathbf{g}}_T(\boldsymbol{\theta}) = \frac{1}{T}\sum_{t=1}^{T} \mathbf{g}_t(\boldsymbol{\theta})$. The matrix $\mathbf{W}_T$ is allowed to be a fixed matrix that does not depend on the data and $\boldsymbol{\theta}$ (the identity matrix, for example), a matrix that depends on the data but not on $\boldsymbol{\theta}$, or a matrix that depends on the data and a preliminary consistent estimator of $\boldsymbol{\theta}_0$ as in the two-step and iterated GMM estimation. Given the first-order asymptotic equivalence of the two-step, iterated, and continuously-updated GMM estimators, our results below can be easily modified to accommodate the continuously-updated (one-step) GMM estimator.

Let $\mathbf{D}_T(\boldsymbol{\theta}) = \frac{1}{T}\sum_{t=1}^{T} \frac{\partial \mathbf{g}_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'}$, $\mathbf{D}(\boldsymbol{\theta}) = E\left[\frac{\partial \mathbf{g}_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'}\right]$ and make the following assumptions.

Assumption A: *Assume that*

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T} \mathbf{g}_t(\boldsymbol{\theta}_0) \xrightarrow{d} N(\mathbf{0}_m, \mathbf{V}), \tag{3}$$

where $\mathbf{V} = \sum_{j=-\infty}^{\infty} E\left[\mathbf{g}_t(\boldsymbol{\theta}_0)\,\mathbf{g}_{t+j}(\boldsymbol{\theta}_0)'\right]$ *is a finite positive definite matrix.*

Assumption B: *Assume that*

(i) $\mathbf{g}_t(\boldsymbol{\theta})$ *is continuous in* $\boldsymbol{\theta}$ *almost surely,* $E\left[\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} |\mathbf{g}_t(\boldsymbol{\theta})|\right] < \infty$, *and the parameter space* $\boldsymbol{\Theta}$ *is a compact subset of* $\mathbb{R}^p$,

(ii) *there exists a unique* $\boldsymbol{\theta}_0 \in \boldsymbol{\Theta}$ *such that* $E\left[\mathbf{g}_t\left(\boldsymbol{\theta}_0\right)\right] = \mathbf{0}_m$ *and* $E\left[\mathbf{g}_t\left(\boldsymbol{\theta}\right)\right] \neq \mathbf{0}_m$ *for all* $\boldsymbol{\theta} \neq \boldsymbol{\theta}_0$,

(iii) $\mathbf{W}_T \xrightarrow{p} \mathbf{W}$, *where* $\mathbf{W}$ *is a non-stochastic symmetric positive definite matrix,*

(iv) $\mathbf{D}_T\left(\boldsymbol{\theta}\right) \xrightarrow{p} \mathbf{D}\left(\boldsymbol{\theta}\right)$ *uniformly in* $\boldsymbol{\theta}$ *on some neighborhood of* $\boldsymbol{\theta}_0$ *and* $\mathbf{D}_0 \equiv \mathbf{D}\left(\boldsymbol{\theta}_0\right)$ *is of rank* p.

Assumption A is a high-level assumption that implicitly imposes restrictions on the data and the vector $\mathbf{g}_t(\boldsymbol{\theta})$. The validity of this assumption can either be verified in the particular context or it can be replaced by a set of explicit primitive conditions. Assumption A can be further strengthened to allow for more general dependence structure (see, for instance, Stock and Wright, 2000). Assumption B imposes sufficient conditions that ensure $\hat{\boldsymbol{\theta}} \xrightarrow{p} \boldsymbol{\theta}_0$ in the interior of the compact parameter space $\boldsymbol{\Theta}$. The uniform convergence and the full rank condition in Assumption B (iv) are required for establishing the asymptotic distributions of $\hat{\boldsymbol{\theta}}$ and $\bar{\mathbf{g}}_T(\hat{\boldsymbol{\theta}})$.

Under Assumptions A and B, $\sqrt{T}\bar{\mathbf{g}}_T(\hat{\boldsymbol{\theta}})$ is asymptotically normally distributed (Hansen, 1982, Lemma 4.1) with mean zero and a singular asymptotic covariance matrix

$$\boldsymbol{\Omega}_0 = [\mathbf{I}_m - \mathbf{D}_0(\mathbf{D}_0'\mathbf{W}\mathbf{D}_0)^{-1}\mathbf{D}_0'\mathbf{W}]\mathbf{V}[\mathbf{I}_m - \mathbf{D}_0(\mathbf{D}_0'\mathbf{W}\mathbf{D}_0)^{-1}\mathbf{D}_0'\mathbf{W}]'. \tag{4}$$

Furthermore, it can be easily seen that the asymptotic covariance matrix of $\sqrt{T}\mathbf{D}_0'\mathbf{W}\bar{\mathbf{g}}_T(\hat{\boldsymbol{\theta}})$ reduces to a $p \times p$ matrix of zeros which renders the asymptotic distribution of $\sqrt{T}\mathbf{D}_0'\mathbf{W}\bar{\mathbf{g}}_T(\hat{\boldsymbol{\theta}})$ degenerate. Provided that $\mathbf{W}_T$ is a consistent estimator of $\mathbf{W}$, a similar degeneracy occurs for $\sqrt{T}\mathbf{D}_0'\mathbf{h}_T(\hat{\boldsymbol{\theta}})$, where $\mathbf{h}_T(\hat{\boldsymbol{\theta}}) \equiv \mathbf{W}_T\bar{\mathbf{g}}_T(\hat{\boldsymbol{\theta}})$.

It is interesting to note that this type of asymptotic degeneracy extends to other setups and arises, for example, in the analysis of Lagrange multipliers in the GEL estimation of moment condition models. In the GEL framework, the estimator of the Lagrange multipliers associated with the moment conditions takes a similar form as $\mathbf{h}_T(\hat{\boldsymbol{\theta}})$ and has an asymptotic covariance matrix given by $\mathbf{V}^{-1} - \mathbf{V}^{-1}\mathbf{D}_0\left(\mathbf{D}_0'\mathbf{V}^{-1}\mathbf{D}_0\right)^{-1}\mathbf{D}_0'\mathbf{V}^{-1}$ (see, for instance, Smith, 1997). It is easy to see that premultiplying by $\mathbf{D}_0'$ reduces this asymptotic variance to a zero matrix. Therefore, the results that we present below can be adapted to deal with the possible asymptotic degeneracy of sample Lagrange multipliers in the GEL framework.

For our analysis, it is more convenient to rewrite the asymptotic normality result in terms of the nonzero parts of the covariance matrices of $\sqrt{T}\bar{\mathbf{g}}_T(\hat{\boldsymbol{\theta}})$ and $\sqrt{T}\mathbf{h}_T(\hat{\boldsymbol{\theta}})$. Let $\mathbf{Q}$ denote an $m \times (m-p)$ orthonormal matrix whose columns are orthogonal to $\mathbf{W}^{\frac{1}{2}}\mathbf{D}_0$. Then,

$$\mathbf{Q}\mathbf{Q}' = \mathbf{I}_m - \mathbf{W}^{\frac{1}{2}}\mathbf{D}_0(\mathbf{D}_0'\mathbf{W}\mathbf{D}_0)^{-1}\mathbf{D}_0'\mathbf{W}^{\frac{1}{2}}. \tag{5}$$

Lemma 1: *Under Assumptions A and B,*

$$\sqrt{T}\mathbf{Q}'\mathbf{W}^{\frac{1}{2}}\bar{\mathbf{g}}_T(\hat{\boldsymbol{\theta}}) \stackrel{d}{\to} N(\mathbf{0}_{m-p}, \mathbf{Q}'\mathbf{W}^{\frac{1}{2}}\mathbf{V}\mathbf{W}^{\frac{1}{2}}\mathbf{Q}) \tag{6}$$

and

$$\sqrt{T}\mathbf{Q}'\mathbf{W}^{-\frac{1}{2}}\mathbf{h}_T(\hat{\boldsymbol{\theta}}) \stackrel{d}{\to} N(\mathbf{0}_{m-p}, \mathbf{Q}'\mathbf{W}^{\frac{1}{2}}\mathbf{V}\mathbf{W}^{\frac{1}{2}}\mathbf{Q}). \tag{7}$$

The role of matrix $\mathbf{Q}$ in Lemma 1 is similar in spirit to the decomposition of Sowell (1996) in which the m vector of normalized population moment conditions $\mathbf{W}^{\frac{1}{2}}E[\mathbf{g}_t(\boldsymbol{\theta}_0)]$ is decomposed into p identifying restrictions used for the estimation of $\boldsymbol{\theta}$ that characterize the space of identifying restrictions and $m-p$ over-identifying restrictions that characterize the space of over-identifying restrictions. Lemma 1 shows that $\sqrt{T}\mathbf{Q}'\mathbf{W}^{\frac{1}{2}}\bar{\mathbf{g}}_T(\hat{\boldsymbol{\theta}})$ and $\sqrt{T}\mathbf{Q}'\mathbf{W}^{-\frac{1}{2}}\mathbf{h}_T(\hat{\boldsymbol{\theta}})$ have a non-degenerate asymptotic normal distribution. However, little is known about the limiting behavior of those linear combinations of $\bar{\mathbf{g}}_T(\hat{\boldsymbol{\theta}})$ or $\mathbf{h}_T(\hat{\boldsymbol{\theta}})$ that do not have an asymptotic normal distribution. The purpose of this paper is to establish the rate of convergence and asymptotic distributions of $\mathbf{D}_0'\mathbf{W}\bar{\mathbf{g}}_T(\hat{\boldsymbol{\theta}})$ and $\mathbf{D}_0'\mathbf{h}_T(\hat{\boldsymbol{\theta}})$. Before we present our main result, we provide an example to illustrate the discontinuous nature of the asymptotic analysis for linear combinations of $\bar{\mathbf{g}}_T(\hat{\boldsymbol{\theta}})$ or $\mathbf{h}_T(\hat{\boldsymbol{\theta}})$.

2.2. MOTIVATION: AN ASSET PRICING EXAMPLE

Let $y_t(\boldsymbol{\theta})$ be a candidate stochastic discount factor (SDF) at time t, where $\boldsymbol{\theta}$ is a p vector of the parameters of the SDF. Suppose we use m test assets to estimate the true SDF parameter vector

$\boldsymbol{\theta}_0$ as well as to test if the proposed SDF is correctly specified. Denote by $\mathbf{R}_t$ the payoffs of the m test assets at time t and by $\mathbf{q}$ the vector of the costs of the m test assets. Let

$$\mathbf{g}_t(\boldsymbol{\theta}) = \mathbf{R}_t y_t(\boldsymbol{\theta}) - \mathbf{q}. \tag{8}$$

If the model is correctly specified, we have $E[\mathbf{g}_t(\boldsymbol{\theta}_0)] = \mathbf{0}_m$. A popular method of estimating $\boldsymbol{\theta}_0$ is to choose $\boldsymbol{\theta}$ to minimize the sample squared Hansen-Jagannathan (HJ, 1997) distance, defined as

$$\hat{\delta}_T^2 = \min_{\boldsymbol{\theta}} \bar{\mathbf{g}}_T(\boldsymbol{\theta})' \mathbf{W}_T \bar{\mathbf{g}}_T(\boldsymbol{\theta}), \tag{9}$$

where $\mathbf{W}_T = \left(\frac{1}{T}\sum_{t=1}^T \mathbf{R}_t\mathbf{R}_t'\right)^{-1}$.

To determine whether the proposed SDF is correctly specified, we can examine the sample pricing errors of the m test assets, i.e., $\bar{\mathbf{g}}_T(\hat{\boldsymbol{\theta}})$, where $\hat{\boldsymbol{\theta}}$ is the vector of estimated parameters chosen to minimize the sample HJ-distance. Alternatively, let $\boldsymbol{\lambda}$ denote an $m \times 1$ vector of Lagrange multipliers associated with the population moment conditions (pricing constraints) and consider the estimated Lagrange multipliers

$$\hat{\boldsymbol{\lambda}} = \mathbf{W}_T \bar{\mathbf{g}}_T(\hat{\boldsymbol{\theta}}), \tag{10}$$

which are a transformation of the sample pricing errors. Hansen and Jagannathan (1997) show that if the proposed SDF does not price the test assets correctly, then it is possible to correct the mispricing of the SDF by subtracting $\boldsymbol{\lambda}'\mathbf{R}_t$ from $y_t(\boldsymbol{\theta})$. As a result, researchers are often interested in testing $H_0 : \lambda_i = 0$, i.e., in determining whether asset i is responsible for the proposed SDF to deviate from the true SDF.

Gospodinov, Kan and Robotti (2010) show that for a linear SDF, $\mathbf{q}'\hat{\boldsymbol{\lambda}} = -\hat{\delta}_T^2$, where $\hat{\delta}_T^2 = \bar{\mathbf{g}}_T(\hat{\boldsymbol{\theta}})'\mathbf{W}_T\bar{\mathbf{g}}_T(\hat{\boldsymbol{\theta}})$ is the squared sample HJ-distance. For the special case of $\mathbf{q} = [1,\ \mathbf{0}_{m-1}']'$ (i.e., the payoff of the first test asset is a gross return and the rest are excess returns), the estimate of the Lagrange multiplier associated with the first test asset, $\hat{\lambda}_1$, is T-consistent and shares the weighted chi-squared distribution of $\hat{\delta}_T^2$ under the assumption of a correctly specified model. This result is of practical importance since applied researchers often rely on the statistical significance of individual

Lagrange multipliers and pricing errors to determine whether an asset pricing model is correctly specified (Cochrane, 1996; Hodrick and Zhang, 2001). Similar problems arise in other asset pricing contexts: for example, in conducting inference on the pricing errors associated with traded factors (see Peñaranda and Sentana, 2010). More generally, as we show below,

$$T\mathbf{D}_0'\hat{\boldsymbol{\lambda}} \xrightarrow{d} -(\mathbf{I}_p \otimes \mathbf{v}_2')\mathbf{v}_1, \tag{11}$$

where $\mathbf{v}_1$ and $\mathbf{v}_2$ are jointly normally distributed vectors of random variables. As a result, any linear combination of $\hat{\boldsymbol{\lambda}}$ with a vector of weights that is in the span of the column space of $\mathbf{D}_0$ is also T-consistent with a non-standard (product of normals) asymptotic distribution.

It is interesting to note that a similar type of discontinuity in the asymptotic approximation and accelerated rate of convergence have been established by Park and Phillips (1989) and Sims, Stock and Watson (1990) in an AR(p) model, $p > 1$, with a unit root in the AR polynomial. In particular, these papers show that a linear combination of $\mathbf{W}_T\bar{\mathbf{g}}_T(\boldsymbol{\theta}_0)$ with a vector of weights $[\alpha_1, ..., \alpha_p]' \neq [\bar{\alpha}, ..., \bar{\alpha}]'$, for some nonzero constant $\bar{\alpha}$, is root-T and asymptotically normally distributed while a linear combination of $\mathbf{W}_T\bar{\mathbf{g}}_T(\boldsymbol{\theta}_0)$ with a vector of weights $[\alpha_1, ..., \alpha_p]' = [\bar{\alpha}, ..., \bar{\alpha}]'$ yields a T-consistent and asymptotically non-normally distributed estimator.

2.3. MAIN RESULTS

We now turn to deriving the asymptotic distributions of $\mathbf{D}_0'\mathbf{W}\bar{\mathbf{g}}_T(\hat{\boldsymbol{\theta}})$ and $\mathbf{D}_0'\mathbf{h}_T(\hat{\boldsymbol{\theta}})$. Due to the similarities in their structure, we first present the results for $\mathbf{D}_0'\mathbf{h}_T(\hat{\boldsymbol{\theta}})$ and discuss the case of $\mathbf{D}_0'\mathbf{W}\bar{\mathbf{g}}_T(\hat{\boldsymbol{\theta}})$ in the next subsection. The following additional assumption on the joint limiting behavior of $\hat{\mathbf{D}}_T = \mathbf{D}_T(\hat{\boldsymbol{\theta}})$ and $\mathbf{h}_T(\hat{\boldsymbol{\theta}})$ is needed to establish the asymptotic distribution of $\mathbf{D}_0'\mathbf{h}_T(\hat{\boldsymbol{\theta}})$.

Assumption C: *Assume that*

$$\sqrt{T}\begin{bmatrix} \mathrm{vec}(\mathbf{Q}'\mathbf{W}^{\frac{1}{2}}\hat{\mathbf{D}}_T) \\ \mathbf{Q}'\mathbf{W}^{-\frac{1}{2}}\mathbf{h}_T(\hat{\boldsymbol{\theta}}) \end{bmatrix} \xrightarrow{d} N(\mathbf{0}_{(m-p)(p+1)}, \boldsymbol{\Sigma}) \tag{12}$$

for some finite positive semidefinite matrix $\boldsymbol{\Sigma}$.

The asymptotic normality of the $m-p$ vector $\mathbf{Q}'\mathbf{W}^{-\frac{1}{2}}\mathbf{h}_T(\hat{\boldsymbol{\theta}})$ follows directly from Lemma 1. The main requirement is on the limiting behavior of the matrix $\hat{\mathbf{D}}_T$ which is, however, rather weak and rules out only some trivial cases. It is important to note that we do not need to impose any restriction on the rate of convergence of $\mathbf{W}_T$ apart from being a consistent estimator of $\mathbf{W}$ (Assumption B (iii)). In contrast, as we argue later, deriving the asymptotic distribution of $\mathbf{D}_0'\mathbf{W}\bar{\mathbf{g}}_T(\hat{\boldsymbol{\theta}})$ requires explicit assumptions on the rate of convergence of $\mathbf{W}_T$ that can differ for parametric and nonparametric heteroskedasticity and autocorrelation consistent (HAC) estimators.

We now state our main result in the following theorem.

Theorem 1: *Under Assumptions A, B, and C,*

$$T\mathbf{D}_0'\mathbf{h}_T(\hat{\boldsymbol{\theta}}) \xrightarrow{d} -(\mathbf{I}_p \otimes \mathbf{v}_2')\mathbf{v}_1, \tag{13}$$

where $\mathbf{v}_1$ *and* $\mathbf{v}_2$ *are* $(m-p)p$ *and* $(m-p)$ *vectors, respectively, and* $[\mathbf{v}_1',\ \mathbf{v}_2']' \sim N(\mathbf{0}_{(m-p)(p+1)}, \boldsymbol{\Sigma})$.

Proof. See Appendix.

In order to make the asymptotic approximation derived in Theorem 1 operational for conducting inference, we need an estimate of the covariance matrix $\boldsymbol{\Sigma}$. In the following, we provide explicit expressions that can be used for consistent estimation of the covariance matrix $\boldsymbol{\Sigma}$ in Theorem 1. Let $\mathbf{G}_T(\boldsymbol{\theta}) = \frac{1}{T}\sum_{t=1}^{T} \partial \text{vec}(\partial \mathbf{g}_t(\boldsymbol{\theta})/\partial\boldsymbol{\theta}')/\partial\boldsymbol{\theta}'$, $\mathbf{G}(\boldsymbol{\theta}) = \partial\text{vec}(\mathbf{D}(\boldsymbol{\theta}))/\partial\boldsymbol{\theta}'$, and $\mathbf{G}_0 = \mathbf{G}(\boldsymbol{\theta}_0)$.

Assumption D: *Assume that* $\mathbf{G}_T(\boldsymbol{\theta}) \xrightarrow{p} \mathbf{G}(\boldsymbol{\theta})$ *uniformly in* $\boldsymbol{\theta}$ *on some neighborhood of* $\boldsymbol{\theta}_0$, *where* $\mathbf{G}(\boldsymbol{\theta})$ *exists, is finite, and is continuous in* $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ *almost surely.*

In the following lemma, we provide the explicit form of the matrix $\boldsymbol{\Sigma}$.

Lemma 2. *Let* $\tilde{\mathbf{G}} = (\mathbf{I}_p \otimes \mathbf{Q}'\mathbf{W}^{\frac{1}{2}})\mathbf{G}_0$. *Under Assumptions A, B, and D, we have*

$$\boldsymbol{\Sigma} = \sum_{j=-\infty}^{\infty} E[\mathbf{d}_t\mathbf{d}_{t+j}'], \tag{14}$$

where $\mathbf{d}_t = [\mathbf{d}'_{1,t},\ \mathbf{d}'_{2,t}]'$ *and*

$$\begin{aligned} \mathbf{d}_{1,t} &= -\tilde{\mathbf{G}}(\mathbf{D}_0'\mathbf{W}\mathbf{D}_0)^{-1}\mathbf{D}_0'\mathbf{W}\mathbf{g}_t(\boldsymbol{\theta}_0) + \mathrm{vec}\left(\mathbf{Q}'\mathbf{W}^{\frac{1}{2}}\frac{\partial \mathbf{g}_t(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}'}\right), \qquad (15) \\ \mathbf{d}_{2,t} &= \mathbf{Q}'\mathbf{W}^{\frac{1}{2}}\mathbf{g}_t(\boldsymbol{\theta}_0). \qquad (16) \end{aligned}$$

Proof. See Appendix.

The consistent estimation of the long-run covariance matrix $\boldsymbol{\Sigma}$ can proceed by using a HAC estimator (see Andrews, 1991, for example) based on the sample counterparts of $\mathbf{d}_{1,t}$ and $\mathbf{d}_{2,t}$.

2.4. DISCUSSION

The result in Theorem 1 has important implications for the asymptotic distribution of a linear combination of $\mathbf{h}_T(\hat{\boldsymbol{\theta}})$ with a weighting vector $\boldsymbol{\alpha}$ which is in the span of the column space of $\mathbf{D}_0$. In particular, if $\boldsymbol{\alpha} = \mathbf{D}_0\tilde{\mathbf{c}}$ for a constant nonzero p vector $\tilde{\mathbf{c}}$, then we have

$$T\boldsymbol{\alpha}'\mathbf{h}_T(\hat{\boldsymbol{\theta}}) \xrightarrow{d} -\tilde{\mathbf{v}}_1'\mathbf{v}_2, \qquad (17)$$

where $\tilde{\mathbf{v}}_1$ is the limit of $\sqrt{T}\mathbf{Q}'\mathbf{W}^{\frac{1}{2}}\hat{\mathbf{D}}_T\tilde{\mathbf{c}}$, $[\tilde{\mathbf{v}}_1',\ \mathbf{v}_2']' \sim N(\mathbf{0}_{2(m-p)}, \tilde{\boldsymbol{\Sigma}})$ with $\tilde{\boldsymbol{\Sigma}} = \sum_{j=-\infty}^{\infty} E[\tilde{\mathbf{d}}_t\tilde{\mathbf{d}}_{t+j}']$, $\tilde{\mathbf{d}}_t = [\tilde{\mathbf{d}}_{1,t}',\ \mathbf{d}_{2,t}']'$ and $\tilde{\mathbf{d}}_{1,t} = (\tilde{\mathbf{c}}' \otimes \mathbf{Q}'\mathbf{W}^{\frac{1}{2}})\mathbf{G}_0(\mathbf{D}_0'\mathbf{W}\mathbf{D}_0)^{-1}\mathbf{D}_0'\mathbf{W}\mathbf{g}_t(\boldsymbol{\theta}_0) + \mathbf{Q}'\mathbf{W}^{\frac{1}{2}}\frac{\partial \mathbf{g}_t(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}'}\tilde{\mathbf{c}}$. Instead of expressing the asymptotic distribution as the inner product of two normal random vectors, the following lemma shows that we can alternatively express it as a linear combination of independent χ_1^2 random variables.

Lemma 3. *Suppose that* $\mathbf{z} = [\mathbf{z}_1',\ \mathbf{z}_2']'$, *where* $\mathbf{z}_1$ *and* $\mathbf{z}_2$ *are both* $n \times 1$ *vectors, is multivariate normally distributed*

$$\mathbf{z} \sim N(\mathbf{0}_{2n}, \boldsymbol{\Psi}), \qquad (18)$$

where $\boldsymbol{\Psi}$ *is a positive semidefinite matrix with rank* $l \leq 2n$. *Let* $\boldsymbol{\Psi} = \mathbf{S}\boldsymbol{\Upsilon}\mathbf{S}'$, *where* $\boldsymbol{\Upsilon}$ *is an* $l \times l$ *diagonal matrix of the nonzero eigenvalues of* $\boldsymbol{\Psi}$ *and* $\mathbf{S}$ *is a* $2n \times l$ *matrix of the corresponding* *eigenvectors. In addition, let*

$$\mathbf{\Gamma} = \mathbf{\Upsilon}^{\frac{1}{2}}\mathbf{S}' \begin{bmatrix} \mathbf{0}_{n\times n} & \frac{1}{2}\mathbf{I}_n \\ \frac{1}{2}\mathbf{I}_n & \mathbf{0}_{n\times n} \end{bmatrix} \mathbf{S}\mathbf{\Upsilon}^{\frac{1}{2}}. \tag{19}$$

Then,

$$\mathbf{z}_1'\mathbf{z}_2 \sim \sum_{i=1}^{k} \gamma_i x_i, \tag{20}$$

where the γ_i's are the $k \leq l$ nonzero eigenvalues of $\mathbf{\Gamma}$ and the x_i's are independent χ_1^2 random variables.

Proof. See Appendix.

This lemma shows that the inner product of two vectors of normal random variables (with mean zero) can always be written as a linear combination of independent chi-squared random variables. This result proves very useful since it allows us to adopt numerical procedures for obtaining the p-value of a weighted chi-squared test that are already available in the literature (Imhof, 1961; Davies, 1980; Lu and King, 2002). Furthermore, this result helps us to reconcile the form of the asymptotic approximation proposed in Theorem 1 with the weighted chi-squared distribution that arises in some special cases as in the asset pricing example.

Extending the result in Theorem 1 to cover the limiting behavior of $\mathbf{A}_0'\bar{\mathbf{g}}_T(\hat{\boldsymbol{\theta}})$, where $\mathbf{A}_0 = \mathbf{W}\mathbf{D}_0$, requires stronger conditions. Defining $\hat{\mathbf{A}}_T = \mathbf{W}_T\hat{\mathbf{D}}_T$, we need to replace Assumption C by assuming

$$\sqrt{T} \begin{bmatrix} \mathrm{vec}(\mathbf{Q}'\mathbf{W}^{-\frac{1}{2}}\hat{\mathbf{A}}_T) \\ \mathbf{Q}'\mathbf{W}^{\frac{1}{2}}\bar{\mathbf{g}}_T(\hat{\boldsymbol{\theta}}) \end{bmatrix} \xrightarrow{d} N(\mathbf{0}_{(m-p)(p+1)}, \mathbf{\Xi}) \tag{21}$$

for some finite positive definite matrix $\mathbf{\Xi}$. The conditions that (21) imposes on the mp vector $\mathrm{vec}(\hat{\mathbf{A}}_T - \mathbf{A}_0)$ can be best seen using the decomposition

$$\begin{aligned} \sqrt{T}(\hat{\mathbf{A}}_T - \mathbf{A}_0) &= \sqrt{T}(\mathbf{W}_T\hat{\mathbf{D}}_T - \mathbf{W}\mathbf{D}_0) \\ &= \sqrt{T}\mathbf{W}(\hat{\mathbf{D}}_T - \mathbf{D}_0) + \sqrt{T}(\mathbf{W}_T - \mathbf{W})\mathbf{D}_0 + \sqrt{T}(\mathbf{W}_T - \mathbf{W})(\hat{\mathbf{D}}_T - \mathbf{D}_0) \\ &= \sqrt{T}\mathbf{W}(\hat{\mathbf{D}}_T - \mathbf{D}_0) + \sqrt{T}(\mathbf{W}_T - \mathbf{W})\mathbf{D}_0 + \mathbf{o}_p(1). \end{aligned} \tag{22}$$

While the conditions for the matrix $\hat{\mathbf{D}}_T$ are easily satisfied (Assumption C), the requirement of root-T convergence for $\mathbf{W}_T$ rules out nonparametric HAC estimators (see Andrews, 1991, for example) but allows for some parametric HAC estimators (West, 1997). In general, this assumption requires that $\mathbf{W}_T$ is computed using a martingale difference sequence process or a dependent process for which the form of serial correlation is known. Then, under the assumption in (21), it can be shown, using similar arguments as in the proof of Theorem 1, that

$$T\mathbf{A}_0'\bar{\mathbf{g}}_T(\hat{\boldsymbol{\theta}}) \xrightarrow{d} -(\mathbf{I}_p \otimes \mathbf{u}_2')\mathbf{u}_1, \tag{23}$$

where $[\mathbf{u}_1', \ \mathbf{u}_2']' \sim N(\mathbf{0}_{(m-p)(p+1)}, \boldsymbol{\Xi})$. We should note that, under some regularity conditions on the kernel function and bandwidth parameter as in Andrews (1991) and Hall and Inoue (2003), a similar limiting representation as in (23) can be derived for nonparameteric HAC estimators but with a slower rate of convergence that depends on the smoothing parameter.

2.5. RANK RESTRICTION TEST

The speed of convergence of $\boldsymbol{\alpha}'\mathbf{h}_T(\hat{\boldsymbol{\theta}})$ and $\boldsymbol{\alpha}'\mathbf{W}\bar{\mathbf{g}}_T(\hat{\boldsymbol{\theta}})$ depends crucially on whether $\boldsymbol{\alpha}$ is in the column span of $\mathbf{D}_0$. In some cases, we know that $\boldsymbol{\alpha}$ is in the column span of $\mathbf{D}_0$ (for instance, in our asset pricing example) and we should rely on the non-standard asymptotics developed above to conduct statistical inference. In general, however, we do not know if $\boldsymbol{\alpha}$ is in the column span of $\mathbf{D}_0$ and we need to resort to pre-testing in order to determine which asymptotic framework should be used for the particular problem at hand. Below we propose a computationally attractive pre-test that determines if $\boldsymbol{\alpha}$ is in the span of the column space of $\mathbf{D}_0$.

Let $\mathbf{P}_{\boldsymbol{\alpha}}$ be an $m \times (m-1)$ orthonormal matrix whose columns are orthogonal to $\boldsymbol{\alpha}$ such that

$$\mathbf{P}_{\boldsymbol{\alpha}}\mathbf{P}_{\boldsymbol{\alpha}}' = \mathbf{I}_m - \boldsymbol{\alpha}(\boldsymbol{\alpha}'\boldsymbol{\alpha})^{-1}\boldsymbol{\alpha}'. \tag{24}$$

Also, let $\boldsymbol{\Pi} = \mathbf{P}_{\boldsymbol{\alpha}}'\mathbf{D}_0$. It turns out that determining if $\boldsymbol{\alpha}$ is in the span of the column space of $\mathbf{D}_0$ is equivalent to determining if $\boldsymbol{\Pi}$ is of reduced rank.

Under the null that $\mathbf{\Pi}$ is of (reduced) rank $p-1$, $H_0 : \text{rank}(\mathbf{\Pi}) = p-1$, there exists a nonzero p vector $\tilde{\mathbf{c}}$ such that $\mathbf{D}_0\tilde{\mathbf{c}} = \boldsymbol{\alpha}$, or equivalently (by premultiplying by $\mathbf{P}'_{\boldsymbol{\alpha}}$ and using the properties of $\mathbf{P}_{\boldsymbol{\alpha}}$) $\mathbf{\Pi}\tilde{\mathbf{c}} = \mathbf{0}_{m-1}$ with the normalization $\tilde{\mathbf{c}}'\tilde{\mathbf{c}} = 1$. As discussed in Cragg and Donald (1997), if $\mathbf{\Pi}$ has a reduced column rank of $p-1$, we can use an alternative normalization and express one column of this matrix, say $\boldsymbol{\pi}_j$, as a linear combination of the other columns, assuming that $\tilde{c}_j \neq 0$. Without any loss of generality, we can order this column first and define the rearranged partitioned matrix $\mathbf{\Pi} = [\boldsymbol{\pi}_1,\ \mathbf{\Pi}_2]$ such that

$$[\boldsymbol{\pi}_1,\ \mathbf{\Pi}_2]\begin{bmatrix} -1 \\ c_2 \\ \vdots \\ c_p \end{bmatrix} = \mathbf{0}_{m-1} \tag{25}$$

or

$$\mathbf{\Pi}_2\mathbf{c}_0 = \boldsymbol{\pi}_1, \tag{26}$$

for some vector $\mathbf{c}_0 = [c_2, ..., c_p]'$. This is equivalent to imposing a normalization on $\tilde{\mathbf{c}}$ such that its first element is -1. With such a normalization, $\mathbf{c}_0$ is uniquely defined provided that $\text{rank}(\mathbf{\Pi}) = p-1$.

Let $\hat{\mathbf{\Pi}}_T = \mathbf{P}'_{\boldsymbol{\alpha}}\hat{\mathbf{D}}_T$. Using Assumption C and the proof of Lemma 2, it can be shown that

$$\sqrt{T}\text{vec}(\hat{\mathbf{\Pi}}_T - \mathbf{\Pi}) \xrightarrow{d} N(\mathbf{0}_{(m-1)p}, \mathbf{M}), \tag{27}$$

where $\mathbf{M} = \sum_{j=-\infty}^{\infty} E[\tilde{\mathbf{m}}_t\tilde{\mathbf{m}}'_{t+j}]$ and

$$\tilde{\mathbf{m}}_t = -(\mathbf{I}_p \otimes \mathbf{P}'_{\boldsymbol{\alpha}})\mathbf{G}_0(\mathbf{D}'_0\mathbf{W}\mathbf{D}_0)^{-1}\mathbf{D}'_0\mathbf{W}\mathbf{g}_t(\boldsymbol{\theta}_0) + \text{vec}\left(\mathbf{P}'_{\boldsymbol{\alpha}}\frac{\partial \mathbf{g}_t(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}'}\right). \tag{28}$$

In practice, $\mathbf{M}$ is replaced by a HAC estimator based on the sample counterparts of $\mathbf{G}_0$, $\mathbf{D}_0$, $\mathbf{W}$ and $\mathbf{g}_t(\boldsymbol{\theta}_0)$. It should be stressed that the first term in the expression for $\tilde{\mathbf{m}}_t$ explicitly accounts for the estimation uncertainty in $\hat{\boldsymbol{\theta}}$ when $\mathbf{g}_t(\boldsymbol{\theta}_0)$ is a nonlinear function of $\boldsymbol{\theta}_0$. When $\mathbf{g}_t(\boldsymbol{\theta}_0)$ is linear in $\boldsymbol{\theta}_0$, the first term in the expression for $\tilde{\mathbf{m}}_t$ drops out ($\mathbf{G}_0 = \mathbf{0}_{p\times p}$) and the second term does not depend on $\boldsymbol{\theta}_0$. As a result, the asymptotic variance of $\hat{\mathbf{\Pi}}_T$ depends only on the data and not on $\boldsymbol{\theta}_0$.

Let $\mathbf{l}_T(\mathbf{c}) = \hat{\mathbf{\Pi}}_{2,T}\mathbf{c} - \hat{\boldsymbol{\pi}}_{1,T}$. Define the test statistic

$$LM = \min_{\mathbf{c}} T[\mathbf{l}_T(\mathbf{c})'\hat{\mathbf{\Lambda}}_T(\mathbf{c})^{-1}\mathbf{l}_T(\mathbf{c})], \tag{29}$$

where $\mathbf{\Lambda}(\mathbf{c}) = ([-1,\ \mathbf{c}'] \otimes \mathbf{I}_{m-1})\mathbf{M}([-1,\ \mathbf{c}']' \otimes \mathbf{I}_{m-1})$ and $\hat{\mathbf{\Lambda}}_T(\mathbf{c})$ denotes its consistent estimator obtained by substituting a HAC estimator of $\mathbf{M}$. The LM statistic tests the null hypothesis that $\mathbf{\Pi}$ is of rank $p-1$ against the alternative that $\mathbf{\Pi}$ is of full rank p. The following lemma shows that the rank test statistic LM is chi-squared distributed with $m-p$ degrees of freedom under the null.

Lemma 4. *Under Assumptions A to D, and* $H_0 : \text{rank}(\mathbf{\Pi}) = p-1$,

$$LM \xrightarrow{d} \chi^2_{m-p}. \tag{30}$$

Proof. See Appendix.

It is important to note that the rank test statistic in equation (29) is invariant to scaling of $\mathbf{c}$. Furthermore, we would like to emphasize that the minimization in (29) is with respect to only a $p-1$ vector $\mathbf{c}$, and the complexity of the minimization problem does not increase with m. Although it can be shown (proof is available upon request) that the LM statistic in (29) is numerically equivalent to the test statistic proposed by Cragg and Donald (1997), it offers substantial computational advantages over the highly dimensional optimization problem in Cragg and Donald's (1997) test. Finally, our simulation experiments show that the test in (30) enjoys excellent size and power properties.

3. ILLUSTRATION: LINEAR ASSET PRICING MODEL

In this section, we specialize our theoretical results to the linear specification of the asset pricing model in Section 2.2 and assess the accuracy of the proposed asymptotic approximation in this setup using a Monte Carlo simulation experiment. In particular, we evaluate the size of the weighted chi-squared test on the Lagrange multiplier associated with the first asset when $\mathbf{q} = [1,\ \mathbf{0}'_{m-1}]'$ (i.e., the payoff of the first asset is a gross return and the payoffs of the other assets are excess returns). We consider two model specifications that are calibrated to monthly data for the period January 1932 – December 2006. The first one is calibrated to the capital asset pricing model (CAPM) with the value-weighted market excess return as risk factor. For the CAPM, the returns on the test assets are the gross return on the risk-free asset and the excess returns on 10 size ranked portfolios. The second specification is calibrated to the three-factor model (FF3) of Fama and French (1993) with risk factors given by the value-weighted market excess return, the return difference between portfolios of small and large stocks, and the return difference between portfolios of high and low book-to-market ratios. For FF3, the returns on the test assets are the gross return on the risk-free asset and the excess returns on 25 size and book-to-market ranked portfolios. All data are obtained from Kenneth French's website. The SDFs of the CAPM and FF3 include an intercept term.

For each model, the factors and the returns on the test assets are drawn from a multivariate normal distribution. The covariance matrix of the factors and returns is chosen based on the covariance matrix estimated from the data. The mean return vector is chosen such that the asset pricing model holds exactly for the test assets. For each simulated set of returns and factors, the unknown parameters $\boldsymbol{\theta}_0$ of the linear SDF $y_t(\boldsymbol{\theta}_0) = \tilde{\mathbf{f}}_t'\boldsymbol{\theta}_0$, where $\tilde{\mathbf{f}}_t = [1,\ \mathbf{f}_t']'$ with $\mathbf{f}_t$ denoting the vector of risk factors, are estimated by minimizing the sample HJ-distance, which yields

$$\hat{\boldsymbol{\theta}} = (\hat{\mathbf{D}}_T'\mathbf{W}_T\hat{\mathbf{D}}_T)^{-1}(\hat{\mathbf{D}}_T'\mathbf{W}_T\mathbf{q}), \tag{31}$$

where $\hat{\mathbf{D}}_T = \frac{1}{T}\sum_{t=1}^{T}\mathbf{R}_t\tilde{\mathbf{f}}_t'$, $\mathbf{W}_T = \left(\frac{1}{T}\sum_{t=1}^{T}\mathbf{R}_t\mathbf{R}_t'\right)^{-1}$, and $\mathbf{q} = [1,\ \mathbf{0}_{m-1}']'$. The estimated Lagrange multipliers are given by

$$\hat{\boldsymbol{\lambda}} = \mathbf{W}_T\bar{\mathbf{g}}_T(\hat{\boldsymbol{\theta}}) = \mathbf{W}_T\left[\frac{1}{T}\sum_{t=1}^{T}\mathbf{R}_t y_t(\hat{\boldsymbol{\theta}}) - \mathbf{q}\right], \tag{32}$$

and we consider the first element $\hat{\lambda}_1$. The next lemma specializes the results in Theorem 1 and Lemma 2 to this setup. It should be noted that while Lemma 5 presents the limiting distribution of $T\hat{\lambda}_1$ for the HJ-distance case, a similar result holds for any weighting matrix W_T that converges in probability to a non-stochastic positive definite matrix W.

Lemma 5. *Let* $\mathbf{D}_0 = E[\mathbf{R}_t\tilde{\mathbf{f}}_t']$, $\mathbf{V} = \sum_{j=-\infty}^{\infty}E[(\mathbf{R}_t\tilde{\mathbf{f}}_t'\boldsymbol{\theta}_0 - \mathbf{q})(\mathbf{R}_{t+j}\tilde{\mathbf{f}}_{t+j}'\boldsymbol{\theta}_0 - \mathbf{q})']$, $\mathbf{W} = (E[\mathbf{R}_t\mathbf{R}_t'])^{-1}$, $\mathbf{Q}$ *denote an orthonormal matrix whose columns are orthogonal to* $\mathbf{W}^{\frac{1}{2}}\mathbf{D}_0$, *and suppose that As-*

sumptions A to D hold. Then,

$$T\hat{\lambda}_1 = T\mathbf{q}'\hat{\boldsymbol{\lambda}} \xrightarrow{d} -\sum_{i=1}^{m-p} \gamma_i x_i, \tag{33}$$

where the x_i's are independent χ^2_1 random variables and the γ_i's are the eigenvalues of the matrix $\mathbf{Q}'\mathbf{W}^{\frac{1}{2}}\mathbf{V}\mathbf{W}^{-\frac{1}{2}}\mathbf{Q}$.

PROOF. See Appendix.

In the analysis of the empirical size of our asymptotic approximation, the computed p-values from the weighted chi-squared distribution in (33) are compared to the 10%, 5%, and 1% theoretical sizes of the test. For a comparison, we also provide the empirical size of a standard normal test of $H_0 : \lambda_1 = 0$ used, for example, in Hodrick and Zhang (2001). The empirical rejection probabilities are computed based on 100,000 Monte Carlo replications.

Table I about here

For different sample sizes T, we report the simulation results for the two model specifications in Panels A and B of Table I. In Panel A, the weighted chi-squared distribution provides a very accurate approximation to the finite-sample behavior of $\hat{\lambda}_1$. In contrast, the standard normal test leads to severe size distortions and rejects the true null hypothesis about 92% of the time at the 5% significance level. In the case of 25 risky assets (Panel B), our approximation tends to over-reject for small sample sizes. This over-rejection is a well documented fact in empirical finance and occurs when the number of test assets m is large relative to the number of time series observations T (see, for instance, Ahn and Gadarowki, 2004). As T increases, the empirical size of the weighted chi-squared approximation approaches its nominal level. In contrast, the standard normal test always rejects the true null hypothesis 100% of the time and does not improve as T increases.

While the incorrect size of the normal test is expected from our theoretical analysis, the severity of these size distortions is somewhat surprising and deserves further attention. Note that the conventional t-statistic for testing $H_0 : \boldsymbol{\alpha}'\boldsymbol{\lambda} = 0$, where $\boldsymbol{\alpha} = \mathbf{D}_0\tilde{\mathbf{c}}$ and $\tilde{\mathbf{c}}$ is a nonzero p vector, is defined as

$$t_{\boldsymbol{\alpha}'\hat{\boldsymbol{\lambda}}} = \frac{\sqrt{T}\boldsymbol{\alpha}'\hat{\boldsymbol{\lambda}}}{\left(\boldsymbol{\alpha}'\mathbf{W}_T\hat{\boldsymbol{\Omega}}\mathbf{W}_T\boldsymbol{\alpha}\right)^{\frac{1}{2}}} = \frac{T\boldsymbol{\alpha}'\hat{\boldsymbol{\lambda}}}{\left(T\boldsymbol{\alpha}'\mathbf{W}_T\hat{\boldsymbol{\Omega}}\mathbf{W}_T\boldsymbol{\alpha}\right)^{\frac{1}{2}}}, \tag{34}$$

where $\hat{\boldsymbol{\Omega}}$ denotes a consistent estimator of $\boldsymbol{\Omega}_0$ in (4). Since $\sqrt{T}\boldsymbol{\alpha}'\hat{\boldsymbol{\lambda}}$ is asymptotically degenerate, it might be expected that the t-test will be undersized, which stands in contrast to the overrejections reported in Table I. The following lemma derives the asymptotic distribution of the t-statistic for testing $H_0 : \boldsymbol{\alpha}'\boldsymbol{\lambda} = 0$ and shows that the numerator $\sqrt{T}\boldsymbol{\alpha}'\hat{\boldsymbol{\lambda}}$ and the denominator $[\boldsymbol{\alpha}'\mathbf{W}_T\hat{\boldsymbol{\Omega}}\mathbf{W}_T\boldsymbol{\alpha}]^{\frac{1}{2}}$ shrink to zero at the same rate, rendering the limit of the ratio $t_{\boldsymbol{\alpha}'\hat{\boldsymbol{\lambda}}}$ a bounded random variable.

Lemma 6. *Suppose $\mathbf{R}_t$ and $\mathbf{f}_t$ are i.i.d. multivariate elliptically distributed with finite fourth moments and kurtosis parameter $\kappa = \mu_4/(3\sigma^4) - 1$, where σ^2 and μ_4 are the second and fourth central moments of the elliptical distribution. Let $\mathbf{H} = E[\tilde{\mathbf{f}}_t\tilde{\mathbf{f}}_t'] + \kappa\mathrm{Var}[\tilde{\mathbf{f}}_t]$. Then,*

$$t_{\boldsymbol{\alpha}'\hat{\boldsymbol{\lambda}}} \xrightarrow{d} r\sqrt{u} + \sqrt{1 - r^2}w, \tag{35}$$

where $r = -\tilde{\mathbf{c}}'\mathbf{H}\boldsymbol{\theta}_0/\sqrt{(\tilde{\mathbf{c}}'\mathbf{H}\tilde{\mathbf{c}})(\boldsymbol{\theta}_0'\mathbf{H}\boldsymbol{\theta}_0)}$, $u \sim \chi^2_{m-p}$, $w \sim N(0,1)$, and u and w are independent of each other.

Proof. See Appendix.

The result in Lemma 6 shows that the asymptotic distribution of the statistic $t_{\boldsymbol{\alpha}'\hat{\boldsymbol{\lambda}}}$ is a mixture of two random variables. Also, it is straightforward to show that the first and second moments of the asymptotic distribution in Lemma 6 are given by

$$E[t_{\boldsymbol{\alpha}'\hat{\boldsymbol{\lambda}}}] = r\frac{\sqrt{2}\Gamma\left(\frac{m-p+1}{2}\right)}{\Gamma\left(\frac{m-p}{2}\right)}, \tag{36}$$

where $\Gamma(\cdot)$ is the gamma function, and

$$E[t^2_{\boldsymbol{\alpha}'\hat{\boldsymbol{\lambda}}}] = 1 + r^2(m - p - 1). \tag{37}$$

When $\boldsymbol{\alpha} = \mathbf{q}$, we have that $\tilde{\mathbf{c}} = \boldsymbol{\theta}_0$, $r = -1$, and the test statistic of $H_0 : \lambda_1 = 0$ is asymptotically distributed as $-\sqrt{\chi^2_{m-p}}$. One important implication of this result is that the correct asymptotic distribution of the t-statistic of $\hat{\lambda}_1$ is miscentered compared to the standard normal approximation and the shift to the left increases with the degree of over-identification. For example, the means of this limiting distribution for the CAPM (with $m - p = 9$) and FF3 (with $m - p = 22$) are -2.92 and -4.53, respectively. These results clearly illustrate that the standard asymptotic inference can be grossly misleading.

However, Lemma 6 also suggests that there are situations in which the standard normal is still the appropriate limiting distribution of $t_{\boldsymbol{\alpha}'\hat{\boldsymbol{\lambda}}}$. For example, this occurs when $r = 0$ or, equivalently, $\tilde{\mathbf{c}}'\mathbf{H}\boldsymbol{\theta}_0 = 0$. Since the mixing coefficient r is consistently estimable, the asymptotic approximation in Lemma 6 conveniently bridges these two extremes ($r^2 = 0$ and $r^2 = 1$).

Next, we investigate the empirical size properties of the sequential test of $H_0 : \boldsymbol{\alpha}'\boldsymbol{\lambda} = 0$ which uses the LM rank test of $H_0 : \text{rank}(\boldsymbol{\Pi}) = p - 1$ from Section 2.5 as a pre-test. Recall that the rank test determines if the normal or the weighted χ^2 distribution theory should be used for testing the hypothesis of interest $H_0 : \boldsymbol{\alpha}'\boldsymbol{\lambda} = 0$. Table II reports the empirical size and power of the reduced rank test when $\boldsymbol{\alpha}$ is in the span of the column space of $\mathbf{D}_0$ ($\boldsymbol{\alpha} = \mathbf{q} = [1,\ \mathbf{0}'_{m-1}]'$) or not ($\boldsymbol{\alpha} = \mathbf{1}_m$). Overall, the rank test exhibits excellent size and power properties.

Table II about here

Table III presents the results on the empirical size of the sequential test of $H_0 : \boldsymbol{\alpha}'\boldsymbol{\lambda} = 0$ for $\boldsymbol{\alpha} = \mathbf{q} = [1,\ \mathbf{0}'_{m-1}]'$ and $\boldsymbol{\alpha} = \mathbf{1}_m$ by setting the nominal levels of the rank pre-test and second-stage hypothesis test to be equal to each other. When $\boldsymbol{\alpha} = \mathbf{q} = [1,\ \mathbf{0}'_{m-1}]'$, $\boldsymbol{\Pi}$ has a reduced rank and the second-stage test uses the weighted χ^2 limiting distribution for inference. As a result, the size of the sequential test (Panel A in Table III) is very similar to the size of the weighted χ^2 approximation reported in Table I. On the other hand, when $\boldsymbol{\alpha}$ is not in the column span of $\mathbf{D}_0$, the appropriate distribution theory is based on the conventional asymptotic normal approximation. Panel B in Table III reveals that in this case, the rank test successfully identifies the appropriate asymptotic framework and the empirical size of the sequential test is very close to its nominal level.

Table III about here

4. CONCLUSION

This paper derives some new results on the asymptotic distribution of linear combinations of GMM sample moment conditions. These results complement Lemma 4.1 of Hansen (1982) with the cases that give rise to singularity of the asymptotic covariance matrix and degeneracy of the asymptotic distribution. Interestingly, we establish that in these cases, the GMM sample moment conditions converge at rate T to their population analogs and follow a non-standard (weighted chi-squared) limiting distribution. Finally, we propose an easy-to-implement rank test to determine which asymptotic framework should be adopted for the particular problem at hand.

ACKNOWLEDGMENTS

We would like to thank the Editor (Jonathan Wright), an Associate Editor, two anonymous referees, Marine Carrasco, Francisco Peñaranda, Peter Phillips, Enrique Sentana, Chu Zhang and the seminar participants at Columbia University and University of British Columbia for helpful comments and suggestions. Gospodinov gratefully acknowledges financial support from Fonds de recherche sur la société et la culture (FQRSC), Institut de finance mathématique de Montréal (IFM2), and the Social Sciences and Humanities Research Council of Canada. Kan gratefully acknowledges financial support from the National Bank Financial of Canada and the Social Sciences and Humanities Research Council of Canada. The views expressed here are the authors' and not necessarily those of the Federal Reserve Bank of Atlanta or the Federal Reserve System.

REFERENCES

Ahn, S. C., and C. Gadarowski (2004) Small sample properties of the model specification test based on the Hansen-Jagannathan distance. *Journal of Empirical Finance* 11, 109–132.

Andrews, D. W. K. (1991) Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica* 59, 817–858.

Cochrane, J. H. (1996) A cross-sectional test of an investment-based asset pricing model. *Journal of Political Economy* 104, 572–621.

Cragg, J. G., and S. G. Donald (1997) Inferring the rank of a matrix. *Journal of Econometrics* 76, 223–250.

Davies, R. B. (1980) Algorithm AS 155: The distribution of a linear combination of χ^2 random variables. *Applied Statistics* 29, 323–333.

Davis, R. A., and W. T. M. Dunsmuir (1996) Maximum likelihood estimation for MA(1) processes with a root on or near the unit circle. *Econometric Theory* 12, 1–29.

Dickey, D. A., and W. A. Fuller (1979) Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association* 74, 427–431.

Fama, E. F., and K. R. French (1993) Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33, 3–56.

Gallant, R. A. (1987) *Nonlinear Statistical Models*, (Wiley, New York).

Gospodinov, N., R. Kan, and C. Robotti (2010) On the Hansen-Jagannathan distance with a no-arbitrage constraint. Working Paper, Federal Reserve Bank of Atlanta.

Hall, A. R., and A. Inoue (2003) The large sample behaviour of the generalized method of moments estimator in misspecified models. *Journal of Econometrics* 114, 361–394.

Hansen, L. P. (1982) Large sample properties of generalized method of moments estimators. *Econometrica* 50, 1029–1054.

Hansen, L. P., and R. Jagannathan (1997) Assessing specification errors in stochastic discount factor models. *Journal of Finance* 52, 557–590.

Hodrick, R., and X. Zhang (2001) Evaluating the specification errors of asset pricing models. *Journal of Financial Economics* 62, 327–376.

Imhof, J. P. (1961) Computing the distribution of quadratic forms in normal variables. *Biometrika* 48, 419–426.

Lu, Z. H., and M. L. King (2002) Improving the numerical technique for computing the accumulated distribution of a quadratic form in normal variables. *Econometric Reviews* 21, 149–165.

Maruyama, Y., and T. Seo (2003) Estimation of moment parameters in elliptical distributions. *Journal of Japan Statistical Society* 33, 215–229.

Nyblom, J. (1989) Testing for the constancy of parameters over time. *Journal of the American Statistical Association* 84, 223–230.

Peñaranda, F., and E. Sentana (2010) A unifying approach to the empirical evaluation of asset pricing models. Working Paper, Universitat Pompeu Fabra.

Park, J. Y., and P. C. B. Phillips (1989) Statistical inference in regressions with integrated processes: Part 2. *Econometric Theory* 5, 95–131.

Phillips, P. C. B. (1987) Towards a unified asymptotic theory for autoregression. *Biometrika* 74, 535–547.

Phillips, P. C. B. (2007) Regression with slowly varying regressors and nonlinear trends. *Econometric Theory* 23, 557–614.

Sargan, J. D. (1959) The estimation of relationships with autocorrelated residuals by the use of instrumental variables. *Journal of the Royal Statistical Society B* 21, 91–105.

Sargan, J. D. (1983) Identification and lack of identification. *Econometrica* 51, 1605–1633.

Sims, C. A., J. H. Stock, and M. W. Watson (1990) Inference in linear time series models with some unit roots. *Econometrica* 58, 113–144.

Smith, R. J. (1997) Alternative semi-parametric likelihood approaches to generalized method of moments estimation. *Economic Journal* 107, 503–519.

Sowell, F. B. (1996) Optimal tests for parameter instability in the generalized method of moments framework. *Econometrica* 64, 1085–1107.

Stock, J. H., and J. H. Wright (2000) GMM with weak identification. *Econometrica* 68, 1055–1096.

Vuong, Q. H. (1989) Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica* 57, 307–333.

West, K. D. (1997) Another heteroskedasticity and autocorrelation consistent covariance matrix estimator. *Journal of Econometrics* 76, 171–191.

APPENDIX: PROOFS OF THEOREMS AND LEMMAS

Proof of Theorem 1: Using the first-order condition $\hat{\mathbf{D}}_T'\mathbf{h}_T(\hat{\boldsymbol{\theta}}) = \mathbf{0}_p$, we express $T\mathbf{D}_0'\mathbf{h}_T(\hat{\boldsymbol{\theta}})$ as

$$\begin{aligned} T\mathbf{D}_0'\mathbf{h}_T(\hat{\boldsymbol{\theta}}) &= -T(\hat{\mathbf{D}}_T - \mathbf{D}_0)'\mathbf{h}_T(\hat{\boldsymbol{\theta}}) \\ &= -\sqrt{T}(\hat{\mathbf{D}}_T - \mathbf{D}_0)'\mathbf{W}^{\frac{1}{2}}(\mathbf{Q}\mathbf{Q}' + \mathbf{W}^{\frac{1}{2}}\mathbf{D}_0(\mathbf{D}_0'\mathbf{W}\mathbf{D}_0)^{-1}\mathbf{D}_0'\mathbf{W}^{\frac{1}{2}})\sqrt{T}\mathbf{W}^{-\frac{1}{2}}\mathbf{h}_T(\hat{\boldsymbol{\theta}}) \\ &= -\left[\sqrt{T}\hat{\mathbf{D}}_T'\mathbf{W}^{\frac{1}{2}}\mathbf{Q}\right]\left[\sqrt{T}\mathbf{Q}'\mathbf{W}^{-\frac{1}{2}}\mathbf{h}_T(\hat{\boldsymbol{\theta}})\right] + \sqrt{T}\left[\mathbf{Q}'\mathbf{W}^{\frac{1}{2}}\mathbf{D}_0\right]'\mathbf{Q}'\mathbf{W}^{-\frac{1}{2}}\sqrt{T}\mathbf{h}_T(\hat{\boldsymbol{\theta}}) \\ &\quad -\left[\sqrt{T}(\hat{\mathbf{D}}_T - \mathbf{D}_0)\right]'\mathbf{W}\mathbf{D}_0(\mathbf{D}_0'\mathbf{W}\mathbf{D}_0)^{-1}\left[\sqrt{T}\mathbf{D}_0'\mathbf{h}_T(\hat{\boldsymbol{\theta}})\right]. \end{aligned} \tag{A1}$$

Since $\mathbf{Q}'\mathbf{W}^{\frac{1}{2}}\mathbf{D}_0 = \mathbf{0}_{(m-p)\times p}$, $\sqrt{T}\mathbf{D}_0'\mathbf{h}_T(\hat{\boldsymbol{\theta}}) = \mathbf{o}_p(1)$ and $\sqrt{T}(\hat{\mathbf{D}}_T - \mathbf{D}_0) = \mathbf{O}_p(1)$, it follows that

$$T\mathbf{D}_0'\mathbf{h}_T(\hat{\boldsymbol{\theta}}) = -\left[\sqrt{T}\hat{\mathbf{D}}_T'\mathbf{W}^{\frac{1}{2}}\mathbf{Q}\right]\left[\sqrt{T}\mathbf{Q}'\mathbf{W}^{-\frac{1}{2}}\mathbf{h}_T(\hat{\boldsymbol{\theta}})\right] + \mathbf{o}_p(1). \tag{A2}$$

Using Assumption C, let $\sqrt{T}\text{vec}(\mathbf{Q}'\mathbf{W}^{\frac{1}{2}}\hat{\mathbf{D}}_T)$ converge to a vector of normal random variables $\mathbf{v}_1$. Similarly, using (7) in Lemma 1, let $\sqrt{T}\mathbf{Q}'\mathbf{W}^{-\frac{1}{2}}\mathbf{h}_T(\hat{\boldsymbol{\theta}})$ converge to a vector of normal random variables $\mathbf{v}_2$ and write the joint distribution of $[\mathbf{v}_1',\ \mathbf{v}_2']'$ as

$$[\mathbf{v}_1',\ \mathbf{v}_2']' \sim N\left(\mathbf{0}_{(m-p)(p+1)}, \boldsymbol{\Sigma}\right). \tag{A3}$$

Thus,

$$T\mathbf{D}_0'\mathbf{h}_T(\hat{\boldsymbol{\theta}}) = \text{vec}(T\mathbf{h}_T(\hat{\boldsymbol{\theta}})'\mathbf{D}_0) \xrightarrow{d} -(\mathbf{I}_p \otimes \mathbf{v}_2')\mathbf{v}_1. \tag{A4}$$

This completes the proof of Theorem 1. ■

Proof of Lemma 2: To obtain the asymptotic distribution of $\text{vec}(\hat{\mathbf{D}}_T - \mathbf{D}_0)$, define $\tilde{\mathbf{D}}_T = \frac{1}{T}\sum_{t=1}^{T}\partial\mathbf{g}_t(\boldsymbol{\theta}_0)/\partial\boldsymbol{\theta}'$ and write

$$\sqrt{T}\text{vec}(\hat{\mathbf{D}}_T - \mathbf{D}_0) = \sqrt{T}\text{vec}(\hat{\mathbf{D}}_T - \tilde{\mathbf{D}}_T) + \sqrt{T}\text{vec}(\tilde{\mathbf{D}}_T - \mathbf{D}_0). \tag{A5}$$

For the first term, we use the mean-value theorem to obtain

$$\begin{aligned} \sqrt{T}\text{vec}(\hat{\mathbf{D}}_T - \tilde{\mathbf{D}}_T) &= \mathbf{G}_0\sqrt{T}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) + \mathbf{o}_p(1) \\ &= -\mathbf{G}_0(\mathbf{D}_0'\mathbf{W}\mathbf{D}_0)^{-1}\mathbf{D}_0'\mathbf{W}\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\mathbf{g}_t(\boldsymbol{\theta}_0) + \mathbf{o}_p(1), \end{aligned} \tag{A6}$$

where the first equality follows from Assumption D and the second equality is ensured by the conditions imposed in Assumption B. For the second term, we have

$$\sqrt{T}\text{vec}(\tilde{\mathbf{D}}_T - \mathbf{D}_0) = \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[\text{vec}\left(\frac{\partial \mathbf{g}_t(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}'}\right) - \text{vec}(\mathbf{D}_0)\right]. \tag{A7}$$

Using expressions (A5), (A6), and (A7), we have

$$\begin{aligned}
&\sqrt{T}\text{vec}(\mathbf{Q}'\mathbf{W}^{\frac{1}{2}}\hat{\mathbf{D}}_T) = \sqrt{T}\text{vec}(\mathbf{Q}'\mathbf{W}^{\frac{1}{2}}(\hat{\mathbf{D}}_T - \mathbf{D}_0)) \\
= \;& (\mathbf{I}_p \otimes \mathbf{Q}'\mathbf{W}^{\frac{1}{2}})\sqrt{T}\text{vec}(\hat{\mathbf{D}}_T - \mathbf{D}_0) \\
= \;& -\tilde{\mathbf{G}}(\mathbf{D}_0'\mathbf{W}\mathbf{D}_0)^{-1}\mathbf{D}_0'\mathbf{W}\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\mathbf{g}_t(\boldsymbol{\theta}_0) + \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\text{vec}\left(\mathbf{Q}'\mathbf{W}^{\frac{1}{2}}\frac{\partial \mathbf{g}_t(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}'}\right) + \mathbf{o}_p(1) \quad \text{(A8)}
\end{aligned}$$

using that $\tilde{\mathbf{G}} = (\mathbf{I}_p \otimes \mathbf{Q}'\mathbf{W}^{\frac{1}{2}})\mathbf{G}_0$. Stacking the expression for $\sqrt{T}\text{vec}(\hat{\mathbf{D}}_T'\mathbf{W}^{\frac{1}{2}}\mathbf{Q})$ with $\mathbf{Q}'\mathbf{W}^{\frac{1}{2}}\mathbf{g}_t(\boldsymbol{\theta}_0)$, we have $\boldsymbol{\Sigma} = \sum_{j=-\infty}^{\infty} E[\mathbf{d}_t\mathbf{d}_{t+j}']$, where $\mathbf{d}_t = [\mathbf{d}_{1,t}',\ \mathbf{d}_{2,t}']'$ and

$$\mathbf{d}_{1,t} = -\tilde{\mathbf{G}}(\mathbf{D}_0'\mathbf{W}\mathbf{D}_0)^{-1}\mathbf{D}_0'\mathbf{W}\mathbf{g}_t(\boldsymbol{\theta}_0) + \text{vec}\left(\mathbf{Q}'\mathbf{W}^{\frac{1}{2}}\frac{\partial \mathbf{g}_t(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}'}\right), \tag{A9}$$

$$\mathbf{d}_{2,t} = \mathbf{Q}'\mathbf{W}^{\frac{1}{2}}\mathbf{g}_t(\boldsymbol{\theta}_0). \tag{A10}$$

This completes the proof of Lemma 2. ■

Proof of Lemma 3: Defining $\tilde{\mathbf{z}} = \mathbf{S}'\mathbf{z} \sim N(\mathbf{0}_l, \boldsymbol{\Upsilon})$, we can write

$$\mathbf{z}_1'\mathbf{z}_2 = \mathbf{z}'\begin{bmatrix}\mathbf{0}_{n\times n} & \frac{1}{2}\mathbf{I}_n \\ \frac{1}{2}\mathbf{I}_n & \mathbf{0}_{n\times n}\end{bmatrix}\mathbf{z} = \tilde{\mathbf{z}}'\mathbf{S}'\begin{bmatrix}\mathbf{0}_{n\times n} & \frac{1}{2}\mathbf{I}_n \\ \frac{1}{2}\mathbf{I}_n & \mathbf{0}_{n\times n}\end{bmatrix}\mathbf{S}\tilde{\mathbf{z}}. \tag{A11}$$

Let $\mathbf{e} = \boldsymbol{\Upsilon}^{-\frac{1}{2}}\tilde{\mathbf{z}} \sim N(\mathbf{0}_l, \mathbf{I}_l)$. Then, we can write

$$\mathbf{z}_1'\mathbf{z}_2 = \mathbf{e}'\boldsymbol{\Upsilon}^{\frac{1}{2}}\mathbf{S}'\begin{bmatrix}\mathbf{0}_{n\times n} & \frac{1}{2}\mathbf{I}_n \\ \frac{1}{2}\mathbf{I}_n & \mathbf{0}_{n\times n}\end{bmatrix}\mathbf{S}\boldsymbol{\Upsilon}^{\frac{1}{2}}\mathbf{e} = \mathbf{e}'\boldsymbol{\Gamma}\mathbf{e}. \tag{A12}$$

Since $\mathbf{e}$ is standard normal, it follows that

$$\mathbf{z}_1'\mathbf{z}_2 \sim \sum_{i=1}^{k}\gamma_i x_i, \tag{A13}$$

where the γ_i's are the $k \leq l$ nonzero eigenvalues of $\boldsymbol{\Gamma}$ and the x_i's are independent χ_1^2 random variables. This completes the proof of Lemma 3. ■

Proof of Lemma 4: Combining $\mathbf{l}_T(\mathbf{c}) = \hat{\mathbf{\Pi}}_{2,T}\mathbf{c} - \hat{\boldsymbol{\pi}}_{1,T} = \text{vec}(\hat{\mathbf{\Pi}}_{2,T}\mathbf{c} - \hat{\boldsymbol{\pi}}_{1,T}) = ([-1,\ \mathbf{c}'] \otimes \mathbf{I}_{m-1})\text{vec}(\hat{\mathbf{\Pi}}_T)$ and equation (27), we have

$$\sqrt{T}\mathbf{l}_T(\mathbf{c}_0) \xrightarrow{d} N(\mathbf{0}_{m-1}, \mathbf{\Lambda}(\mathbf{c}_0)), \tag{A14}$$

where $\mathbf{\Lambda}(\mathbf{c}_0) = ([-1,\ \mathbf{c}_0'] \otimes \mathbf{I}_{m-1})\mathbf{M}([-1,\ \mathbf{c}_0']' \otimes \mathbf{I}_{m-1})$. Let

$$\hat{\mathbf{c}} = \arg\min_{\mathbf{c}} \mathbf{l}_T(\mathbf{c})'\mathbf{\Lambda}_T^{-1}(\mathbf{c})\mathbf{l}_T(\mathbf{c}) \tag{A15}$$

be the estimator of $\mathbf{c}_0$. First, note that while the estimator $\hat{\mathbf{c}}$ depends on a preliminary estimator $\hat{\boldsymbol{\theta}}$ when $\mathbf{g}_t(\boldsymbol{\theta})$ is a nonlinear function of $\boldsymbol{\theta}$, the uncertainty associated with the estimation of $\boldsymbol{\theta}$ is already incorporated in $\mathbf{M}$. Also, from the asymptotic equivalence of the estimator in (A15) with the two-step GMM estimator, we have (Hansen, 1982)

$$\sqrt{T}(\hat{\mathbf{c}} - \mathbf{c}_0) = -[\mathbf{\Pi}_2'\mathbf{\Lambda}(\mathbf{c}_0)^{-1}\mathbf{\Pi}_2]^{-1}\mathbf{\Pi}_2'\mathbf{\Lambda}(\mathbf{c}_0)^{-1}\sqrt{T}\mathbf{l}_T(\mathbf{c}_0) + \mathbf{o}_p(1). \tag{A16}$$

Finally, using similar arguments as in Hansen (1982, Lemma 4.2), it follows that LM is asymptotically distributed as a chi-squared random variable with $(m-1) - (p-1) = m - p$ degrees of freedom. This completes the proof of Lemma 4. ■

Proof of Lemma 5: In the case of asset pricing models with a pricing constraint $\bar{\mathbf{g}}_T(\boldsymbol{\theta}) = \frac{1}{T}\sum_{t=1}^T \mathbf{R}_t y_t(\boldsymbol{\theta}) - \mathbf{q}$, the expressions for $\mathbf{d}_{1,t}$ and $\mathbf{d}_{2,t}$ in the covariance matrix $\mathbf{\Sigma} = \sum_{j=-\infty}^{\infty} E[\mathbf{d}_t\mathbf{d}_{t+j}']$ in Lemma 2 specialize to

$$\mathbf{d}_{1,t} = -\tilde{\mathbf{G}}(\mathbf{D}_0'\mathbf{W}\mathbf{D}_0)^{-1}\mathbf{D}_0'\mathbf{W}(\mathbf{R}_t y_t(\boldsymbol{\theta}_0) - \mathbf{q}) + (\mathbf{Q}'\mathbf{W}^{\frac{1}{2}}\mathbf{R}_t \otimes \mathbf{I}_p)\frac{\partial y_t(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}'}, \tag{A17}$$

$$\mathbf{d}_{2,t} = \mathbf{Q}'\mathbf{W}^{\frac{1}{2}}(\mathbf{R}_t y_t(\boldsymbol{\theta}_0) - \mathbf{q}), \tag{A18}$$

where $\mathbf{D}_0 = E\left[\mathbf{R}_t\frac{\partial y_t(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}'}\right]$ and $\mathbf{G}_0 = E\left[(\mathbf{R}_t \otimes \mathbf{I}_p)\frac{\partial^2 y_t(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}\partial \boldsymbol{\theta}'}\right]$.

For the special case of a linear SDF that prices the test assets correctly, these expressions can be further simplified and have the form

$$\mathbf{d}_{1,t} = \mathbf{Q}'\mathbf{W}^{\frac{1}{2}}\mathbf{R}_t \otimes \tilde{\mathbf{f}}_t, \tag{A19}$$

$$\mathbf{d}_{2,t} = \mathbf{Q}'\mathbf{W}^{\frac{1}{2}}\mathbf{R}_t\tilde{\mathbf{f}}_t'\boldsymbol{\theta}_0 \tag{A20}$$

since $\mathbf{G}_0$ is a null matrix and $\mathbf{Q}'\mathbf{W}^{\frac{1}{2}}\mathbf{q} = \mathbf{Q}'\mathbf{W}^{\frac{1}{2}}\mathbf{D}_0\boldsymbol{\theta}_0 = \mathbf{0}_{m-p}$ from the definition of $\mathbf{Q}$.

For the linear combination $T\boldsymbol{\alpha}'\mathbf{h}_T(\hat{\boldsymbol{\theta}})$, where $\boldsymbol{\alpha} = \mathbf{D}_0\tilde{\mathbf{c}}$ and $\mathbf{h}_T(\hat{\boldsymbol{\theta}}) = \hat{\boldsymbol{\lambda}}$, we have from the proof of Theorem 1 that

$$T\tilde{\mathbf{c}}'\mathbf{D}_0'\hat{\boldsymbol{\lambda}} = -\left[\sqrt{T}\mathbf{Q}'\mathbf{W}^{\frac{1}{2}}\hat{\mathbf{D}}_T\tilde{\mathbf{c}}\right]'\left[\sqrt{T}\mathbf{Q}'\mathbf{W}^{-\frac{1}{2}}\mathbf{h}_T(\hat{\boldsymbol{\theta}})\right] + o_p(1). \tag{A21}$$

It is straightforward to show using the results above that

$$\sqrt{T}\mathbf{Q}'\mathbf{W}^{\frac{1}{2}}\hat{\mathbf{D}}_T\tilde{\mathbf{c}} \xrightarrow{d} N\left(\mathbf{0}_{m-p}, \sum_{j=-\infty}^{\infty} E[\tilde{\mathbf{d}}_{1,t}\tilde{\mathbf{d}}_{1,t+j}']\right), \tag{A22}$$

where $\tilde{\mathbf{d}}_{1,t} = \mathbf{Q}'\mathbf{W}^{\frac{1}{2}}\mathbf{R}_t\tilde{\mathbf{f}}_t'\tilde{\mathbf{c}}$. When $\tilde{\mathbf{c}} = \boldsymbol{\theta}_0$, i.e., $\tilde{\mathbf{c}}'\mathbf{D}_0'\hat{\boldsymbol{\lambda}} = \mathbf{q}'\hat{\boldsymbol{\lambda}}$, we have $\tilde{\mathbf{d}}_{1t} = \mathbf{d}_{2,t}$ and it follows that

$$T\mathbf{q}'\hat{\boldsymbol{\lambda}} \xrightarrow{d} -\mathbf{v}_2'\mathbf{v}_2, \tag{A23}$$

which is a linear combination of $m-p$ independent chi-squared random variables with one degree of freedom. Since $\mathbf{v}_2 \sim N(\mathbf{0}_{m-p}, \mathbf{Q}'\mathbf{W}^{\frac{1}{2}}\mathbf{V}\mathbf{W}^{\frac{1}{2}}\mathbf{Q})$, the weights for the weighted χ^2 distribution are given by the eigenvalues of the matrix $\mathbf{Q}'\mathbf{W}^{\frac{1}{2}}\mathbf{V}\mathbf{W}^{\frac{1}{2}}\mathbf{Q}$. This completes the proof of Lemma 5. ■

Proof of Lemma 6: Using that $\boldsymbol{\alpha} = \mathbf{D}_0\tilde{\mathbf{c}}$, the numerator of the t-statistic $t_{\boldsymbol{\alpha}'\hat{\boldsymbol{\lambda}}}$ can be expressed as

$$\begin{aligned} T\tilde{\mathbf{c}}'\mathbf{D}_0'\hat{\boldsymbol{\lambda}} &= T(\tilde{\mathbf{c}}'\mathbf{D}_0' - \tilde{\mathbf{c}}'\hat{\mathbf{D}}_T')\hat{\boldsymbol{\lambda}} \\ &= T(\tilde{\mathbf{c}}'\mathbf{D}_0' - \tilde{\mathbf{c}}'\hat{\mathbf{D}}_T')\mathbf{W}^{\frac{1}{2}}\mathbf{Q}\mathbf{Q}'\mathbf{W}^{-\frac{1}{2}}\hat{\boldsymbol{\lambda}} + o_p(1) \\ &= -[\sqrt{T}\tilde{\mathbf{c}}'\hat{\mathbf{D}}_T'\mathbf{W}^{\frac{1}{2}}\mathbf{Q}][\sqrt{T}\mathbf{Q}'\mathbf{W}^{-\frac{1}{2}}\hat{\boldsymbol{\lambda}}] + o_p(1) \xrightarrow{d} -\mathbf{z}_1'\mathbf{z}_2, \end{aligned} \tag{A24}$$

where $\mathbf{z}_1$ is the limiting distribution of $\sqrt{T}\mathbf{Q}'\mathbf{W}^{\frac{1}{2}}\hat{\mathbf{D}}_T\tilde{\mathbf{c}}$ and $\mathbf{z}_2$ is the limiting distribution of $\sqrt{T}\mathbf{Q}'\mathbf{W}^{-\frac{1}{2}}\hat{\boldsymbol{\lambda}}$.

Next, the consistent estimator of $\boldsymbol{\Omega}_0$ in (4) is given by

$$\begin{aligned} \hat{\boldsymbol{\Omega}} &= [\mathbf{I}_m - \hat{\mathbf{D}}_T(\hat{\mathbf{D}}_T'\mathbf{W}_T\hat{\mathbf{D}}_T)^{-1}\hat{\mathbf{D}}_T'\mathbf{W}_T]\hat{\mathbf{V}}[\mathbf{I}_m - \hat{\mathbf{D}}_T(\hat{\mathbf{D}}_T'\mathbf{W}_T\hat{\mathbf{D}}_T)^{-1}\hat{\mathbf{D}}_T'\mathbf{W}_T]' \\ &= \mathbf{W}_T^{-\frac{1}{2}}[\mathbf{I}_m - \mathbf{W}_T^{\frac{1}{2}}\hat{\mathbf{D}}_T(\hat{\mathbf{D}}_T'\mathbf{W}_T\hat{\mathbf{D}}_T)^{-1}\hat{\mathbf{D}}_T'\mathbf{W}_T^{\frac{1}{2}}]\mathbf{W}_T^{\frac{1}{2}}\hat{\mathbf{V}}\mathbf{W}_T^{\frac{1}{2}}[\mathbf{I}_m - \mathbf{W}_T^{\frac{1}{2}}\hat{\mathbf{D}}_T(\hat{\mathbf{D}}_T'\mathbf{W}_T\hat{\mathbf{D}}_T)^{-1}\hat{\mathbf{D}}_T'\mathbf{W}_T^{\frac{1}{2}}]\mathbf{W}_T^{-\frac{1}{2}} \\ &= \mathbf{W}_T^{-\frac{1}{2}}\hat{\mathbf{Q}}\hat{\mathbf{Q}}'\mathbf{W}_T^{\frac{1}{2}}\hat{\mathbf{V}}\mathbf{W}_T^{\frac{1}{2}}\hat{\mathbf{Q}}\hat{\mathbf{Q}}'\mathbf{W}_T^{-\frac{1}{2}}, \end{aligned} \tag{A25}$$

where $\hat{\mathbf{V}}$ is a consistent estimator of $\mathbf{V}$ and $\hat{\mathbf{Q}}$ is an $m \times (m-p)$ orthonormal matrix with its columns orthogonal to $\mathbf{W}_T^{\frac{1}{2}}\hat{\mathbf{D}}_T$. As a result, we can rewrite the term inside the squared root of the denominator of the t-statistic as

$$T\boldsymbol{\alpha}'\mathbf{W}_T\hat{\boldsymbol{\Omega}}\mathbf{W}_T\boldsymbol{\alpha} = T\tilde{\mathbf{c}}'\mathbf{D}_0'\mathbf{W}_T^{\frac{1}{2}}\hat{\mathbf{Q}}\hat{\mathbf{Q}}'\mathbf{W}_T^{\frac{1}{2}}\hat{\mathbf{V}}\mathbf{W}_T^{\frac{1}{2}}\hat{\mathbf{Q}}\hat{\mathbf{Q}}'\mathbf{W}_T^{\frac{1}{2}}\mathbf{D}_0\tilde{\mathbf{c}}. \tag{A26}$$

Since

$$\begin{aligned}\sqrt{T}\hat{\mathbf{Q}}'\mathbf{W}_T^{\frac{1}{2}}\mathbf{D}_0\tilde{\mathbf{c}} &= \sqrt{T}\hat{\mathbf{Q}}'\mathbf{W}_T^{\frac{1}{2}}(\mathbf{D}_0 - \hat{\mathbf{D}}_T)\tilde{\mathbf{c}} = \sqrt{T}\mathbf{Q}'\mathbf{W}^{\frac{1}{2}}(\mathbf{D}_0 - \hat{\mathbf{D}}_T)\tilde{\mathbf{c}} + \mathbf{o}_p(1) \\ &= -\sqrt{T}\mathbf{Q}'\mathbf{W}^{\frac{1}{2}}\hat{\mathbf{D}}_T\tilde{\mathbf{c}} + \mathbf{o}_p(1) \xrightarrow{d} -\mathbf{z}_1,\end{aligned} \tag{A27}$$

it follows that

$$T\boldsymbol{\alpha}'\mathbf{W}_T\hat{\boldsymbol{\Omega}}\mathbf{W}_T\boldsymbol{\alpha} \xrightarrow{d} \mathbf{z}_1'\mathbf{Q}'\mathbf{W}^{\frac{1}{2}}\mathbf{V}\mathbf{W}^{\frac{1}{2}}\mathbf{Q}\mathbf{z}_1. \tag{A28}$$

Therefore, we have

$$t_{\boldsymbol{\alpha}'\hat{\boldsymbol{\lambda}}} \xrightarrow{d} -\frac{\mathbf{z}_1'\mathbf{z}_2}{[\mathbf{z}_1'\mathbf{Q}'\mathbf{W}^{\frac{1}{2}}\mathbf{V}\mathbf{W}^{\frac{1}{2}}\mathbf{Q}\mathbf{z}_1]^{\frac{1}{2}}}. \tag{A29}$$

The joint distribution of $\mathbf{z}_1$ and $\mathbf{z}_2$ is $N(\mathbf{0}_{2(m-p)}, \boldsymbol{\Sigma})$, where

$$\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix} = \sum_{j=-\infty}^{\infty} E[\tilde{\mathbf{d}}_t\tilde{\mathbf{d}}_{t+j}'], \tag{A30}$$

and

$$\tilde{\mathbf{d}}_t = \begin{bmatrix} \tilde{\mathbf{d}}_{1,t} \\ \tilde{\mathbf{d}}_{2,t} \end{bmatrix} = \begin{bmatrix} \mathbf{Q}'\mathbf{W}^{\frac{1}{2}}\mathbf{R}_t\tilde{\mathbf{f}}_t'\tilde{\mathbf{c}}, \\ \mathbf{Q}'\mathbf{W}^{\frac{1}{2}}\mathbf{R}_t\tilde{\mathbf{f}}_t'\boldsymbol{\theta}_0 \end{bmatrix}. \tag{A31}$$

Conditional on $\mathbf{z}_1$, we have

$$\mathbf{z}_2 \sim N(\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\mathbf{z}_1, \boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}). \tag{A32}$$

Noting that $\mathbf{Q}'\mathbf{W}^{\frac{1}{2}}\mathbf{V}\mathbf{W}^{\frac{1}{2}}\mathbf{Q} = \boldsymbol{\Sigma}_{22}$, we have conditional on $\mathbf{z}_1$,

$$t_{\boldsymbol{\alpha}'\hat{\boldsymbol{\lambda}}} \xrightarrow{d} N\left(-\frac{\mathbf{z}_1'\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\mathbf{z}_1}{(\mathbf{z}_1'\boldsymbol{\Sigma}_{22}\mathbf{z}_1)^{\frac{1}{2}}}, \frac{\mathbf{z}_1'(\boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12})\mathbf{z}_1}{\mathbf{z}_1'\boldsymbol{\Sigma}_{22}\mathbf{z}_1}\right). \tag{A33}$$

Let $\tilde{\mathbf{z}}_1 = \boldsymbol{\Sigma}_{11}^{-\frac{1}{2}}\mathbf{z}_1 \sim N(\mathbf{0}_{m-p}, \mathbf{I}_{m-p})$ and $w \sim N(0,1)$ be independent of each other. Then, we can write

$$t_{\boldsymbol{\alpha}'\hat{\boldsymbol{\lambda}}} \xrightarrow{d} -\frac{\tilde{\mathbf{z}}_1'\boldsymbol{\Sigma}_{11}^{\frac{1}{2}}\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-\frac{1}{2}}\tilde{\mathbf{z}}_1}{(\tilde{\mathbf{z}}_1'\boldsymbol{\Sigma}_{11}^{\frac{1}{2}}\boldsymbol{\Sigma}_{22}\boldsymbol{\Sigma}_{11}^{\frac{1}{2}}\tilde{\mathbf{z}}_1)^{\frac{1}{2}}} + \left[\frac{\tilde{\mathbf{z}}_1'(\boldsymbol{\Sigma}_{11}^{\frac{1}{2}}\boldsymbol{\Sigma}_{22}\boldsymbol{\Sigma}_{11}^{\frac{1}{2}} - \boldsymbol{\Sigma}_{11}^{\frac{1}{2}}\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{11}^{\frac{1}{2}})\tilde{\mathbf{z}}_1}{\tilde{\mathbf{z}}_1'\boldsymbol{\Sigma}_{11}^{\frac{1}{2}}\boldsymbol{\Sigma}_{22}\boldsymbol{\Sigma}_{11}^{\frac{1}{2}}\tilde{\mathbf{z}}_1}\right]^{\frac{1}{2}} w. \tag{A34}$$

If we assume that $\mathbf{R}_t$ and $\mathbf{f}_t$ are i.i.d. multivariate elliptically distributed, (A34) can be greatly simplified. Suppose that (X_i, X_j, X_k, X_l) follow a multivariate elliptical distribution with multivariate kurtosis parameter κ. Using Lemma 2 of Maruyama and Seo (2003), we have

$$\begin{aligned} E[X_iX_jX_kX_l] &= (1+\kappa)(\sigma_{ij}\sigma_{kl}+\sigma_{ik}\sigma_{jl}+\sigma_{il}\sigma_{jk})+\mu_i\mu_j\mu_k\mu_l \\ &\quad +\sigma_{ij}\mu_k\mu_l+\sigma_{ik}\mu_j\mu_l+\sigma_{il}\mu_j\mu_k+\sigma_{jk}\mu_i\mu_l+\sigma_{jl}\mu_i\mu_k+\sigma_{kl}\mu_i\mu_j, \quad \text{(A35)} \end{aligned}$$

where $\mu_i = E[X_i]$ and $\sigma_{ij} = \text{Cov}[X_i, X_j]$. If we let $\tilde{\mathbf{R}}_t = \mathbf{Q}'\mathbf{W}^{\frac{1}{2}}\mathbf{R}_t$, we can easily establish that

$$E[\tilde{\mathbf{R}}_t] = \mathbf{Q}'\mathbf{W}^{\frac{1}{2}}\mathbf{D}_0[1,\ \mathbf{0}'_{p-1}]' = \mathbf{0}_{m-p}, \quad \text{(A36)}$$

$$\text{Var}[\tilde{\mathbf{R}}_t] = E[\mathbf{Q}'\mathbf{W}^{\frac{1}{2}}\mathbf{R}_t\mathbf{R}_t'\mathbf{W}^{\frac{1}{2}}\mathbf{Q}] = \mathbf{Q}'\mathbf{W}^{\frac{1}{2}}\mathbf{W}^{-1}\mathbf{W}^{\frac{1}{2}}\mathbf{Q} = \mathbf{I}_{m-p}, \quad \text{(A37)}$$

$$\text{Cov}[\tilde{\mathbf{R}}_t, \tilde{\mathbf{f}}_t'] = E[\mathbf{Q}'\mathbf{W}^{\frac{1}{2}}\mathbf{R}_t\tilde{\mathbf{f}}_t'] = \mathbf{Q}'\mathbf{W}^{\frac{1}{2}}\mathbf{D}_0 = \mathbf{0}_{(m-p)\times p}. \quad \text{(A38)}$$

Then, using (A35), we obtain

$$\begin{aligned} \boldsymbol{\Sigma}_{12} &= E[\tilde{\mathbf{d}}_{1,t}\tilde{\mathbf{d}}_{2,t}'] \\ &= (1+\kappa)\text{Cov}[\tilde{\mathbf{c}}'\tilde{\mathbf{f}}_t, \boldsymbol{\theta}_0'\tilde{\mathbf{f}}_t]\mathbf{I}_{m-p} + E[\tilde{\mathbf{c}}'\tilde{\mathbf{f}}_t]E[\boldsymbol{\theta}_0'\tilde{\mathbf{f}}_t]\mathbf{I}_{m-p} \\ &= (E[(\tilde{\mathbf{c}}'\tilde{\mathbf{f}}_t)(\boldsymbol{\theta}_0'\tilde{\mathbf{f}}_t)] + \kappa\text{Cov}[\tilde{\mathbf{c}}'\tilde{\mathbf{f}}_t, \boldsymbol{\theta}_0'\tilde{\mathbf{f}}_t])\mathbf{I}_{m-p} \\ &= (\tilde{\mathbf{c}}'\mathbf{H}\boldsymbol{\theta}_0)\mathbf{I}_{m-p}, \quad \text{(A39)} \end{aligned}$$

where $\mathbf{H} = E[\tilde{\mathbf{f}}_t\tilde{\mathbf{f}}_t'] + \kappa\text{Var}[\tilde{\mathbf{f}}_t]$. Similarly, we have $\boldsymbol{\Sigma}_{11} = (\tilde{\mathbf{c}}'\mathbf{H}\tilde{\mathbf{c}})\mathbf{I}_{m-p}$ and $\boldsymbol{\Sigma}_{22} = (\boldsymbol{\theta}_0'\mathbf{H}\boldsymbol{\theta}_0)\mathbf{I}_{m-p}$. Therefore, when $\mathbf{R}_t$ and $\mathbf{f}_t$ are i.i.d. multivariate elliptically distributed, we can simplify the asymptotic distribution of $t_{\boldsymbol{\alpha}'\hat{\boldsymbol{\lambda}}}$ to

$$t_{\boldsymbol{\alpha}'\hat{\boldsymbol{\lambda}}} \stackrel{d}{\to} -\frac{\tilde{\mathbf{c}}'\mathbf{H}\boldsymbol{\theta}_0}{\sqrt{(\tilde{\mathbf{c}}'\mathbf{H}\tilde{\mathbf{c}})(\boldsymbol{\theta}_0'\mathbf{H}\boldsymbol{\theta}_0)}}\sqrt{\tilde{\mathbf{z}}_1'\tilde{\mathbf{z}}_1} + \left[1 - \frac{(\tilde{\mathbf{c}}'\mathbf{H}\boldsymbol{\theta}_0)^2}{(\tilde{\mathbf{c}}'\mathbf{H}\tilde{\mathbf{c}})(\boldsymbol{\theta}_0'\mathbf{H}\boldsymbol{\theta}_0)}\right]^{\frac{1}{2}} w = r\sqrt{u} + \sqrt{1-r^2}w, \quad \text{(A40)}$$

where $r = -(\tilde{\mathbf{c}}'\mathbf{H}\boldsymbol{\theta}_0)/\sqrt{(\tilde{\mathbf{c}}'\mathbf{H}\tilde{\mathbf{c}})(\boldsymbol{\theta}_0'\mathbf{H}\boldsymbol{\theta}_0)}$ and $u = \tilde{\mathbf{z}}_1'\tilde{\mathbf{z}}_1$ is a χ^2_{m-p} random variable which is independent of w. This completes the proof of Lemma 6. ■

Table I
Empirical Sizes of the tests of $H_0 : \lambda_1 = 0$

Panel A: CAPM

	Standard Normal			Weighted χ^2		
	Level of Significance			Level of Significance		
T	10%	5%	1%	10%	5%	1%
150	0.978	0.929	0.689	0.144	0.082	0.022
300	0.977	0.925	0.682	0.121	0.065	0.015
450	0.976	0.923	0.679	0.115	0.060	0.014
600	0.976	0.924	0.679	0.111	0.057	0.013
750	0.975	0.923	0.679	0.109	0.057	0.012
900	0.976	0.923	0.680	0.107	0.055	0.011

Panel B: FF3

	Standard Normal			Weighted χ^2		
	Level of Significance			Level of Significance		
T	10%	5%	1%	10%	5%	1%
150	1.000	1.000	1.000	0.284	0.189	0.072
300	1.000	1.000	1.000	0.178	0.105	0.031
450	1.000	1.000	0.999	0.151	0.084	0.022
600	1.000	1.000	0.999	0.138	0.074	0.018
750	1.000	1.000	0.999	0.130	0.070	0.016
900	1.000	1.000	0.999	0.125	0.067	0.015

The table presents the actual probabilities of rejection for the asymptotic tests of $H_0 : \lambda_1 = 0$ with different levels of significance, assuming that the factors and returns are generated from a multivariate normal distribution. We consider two model specifications that are calibrated to monthly data for the period January 1932 – December 2006. The model specification in Panel A is calibrated to the capital asset pricing model (CAPM). The model specification in Panel B is calibrated to the three-factor model of Fama and French (FF3, 1993). The results for different number of time series observations (T) are based on 100,000 simulations.

Table II
Empirical Size and Power of the Rank Test

Panel A: $\boldsymbol{\alpha} = \mathbf{q} = [1,\ \mathbf{0}'_{m-1}]'$

	CAPM			FF3		
	Level of Significance			Level of Significance		
T	10%	5%	1%	10%	5%	1%
150	0.095	0.044	0.007	0.069	0.024	0.001
300	0.098	0.048	0.009	0.093	0.044	0.007
450	0.099	0.050	0.009	0.098	0.047	0.008
600	0.099	0.049	0.010	0.099	0.047	0.009
750	0.100	0.050	0.010	0.100	0.049	0.009
900	0.099	0.050	0.010	0.100	0.050	0.009

Panel B: $\boldsymbol{\alpha} = \mathbf{1}_m$

	CAPM			FF3		
	Level of Significance			Level of Significance		
T	10%	5%	1%	10%	5%	1%
150	0.999	0.997	0.965	0.977	0.913	0.531
300	1.000	1.000	1.000	1.000	1.000	1.000
450	1.000	1.000	1.000	1.000	1.000	1.000
600	1.000	1.000	1.000	1.000	1.000	1.000
750	1.000	1.000	1.000	1.000	1.000	1.000
900	1.000	1.000	1.000	1.000	1.000	1.000

The table presents the actual probabilities of rejection for the LM rank test of $H_0 : \text{rank}(\boldsymbol{\Pi}) = p-1$ with different levels of significance, assuming that the factors and returns are generated from a multivariate normal distribution. We consider two model specifications (CAPM and FF3) that are calibrated to monthly data for the period January 1932 – December 2006. Panel A presents the empirical size of the rank test for $\boldsymbol{\alpha} = \mathbf{q} = [1,\ \mathbf{0}'_{m-1}]'$. Panel B reports the empirical power of the rank test for $\boldsymbol{\alpha} = \mathbf{1}_m$. The results for different number of time series observations (T) are based on 100,000 simulations.

Table III
Empirical Size of the Sequential Test

Panel A: $\boldsymbol{\alpha} = \mathbf{q} = [1,\ \mathbf{0}'_{m-1}]'$

	CAPM			FF3		
	Level of Significance			Level of Significance		
T	10%	5%	1%	10%	5%	1%
150	0.145	0.082	0.022	0.284	0.189	0.072
300	0.121	0.065	0.015	0.178	0.105	0.031
450	0.115	0.060	0.014	0.151	0.085	0.022
600	0.111	0.058	0.013	0.138	0.074	0.018
750	0.109	0.057	0.012	0.130	0.070	0.016
900	0.107	0.055	0.011	0.125	0.067	0.015

Panel B: $\boldsymbol{\alpha} = \mathbf{1}_m$

	CAPM			FF3		
	Level of Significance			Level of Significance		
T	10%	5%	1%	10%	5%	1%
150	0.123	0.067	0.022	0.177	0.124	0.130
300	0.110	0.057	0.012	0.132	0.072	0.018
450	0.106	0.054	0.011	0.121	0.065	0.014
600	0.104	0.053	0.011	0.116	0.061	0.013
750	0.104	0.052	0.011	0.112	0.059	0.013
900	0.103	0.051	0.011	0.110	0.057	0.012

The table presents the actual probabilities of rejection for the sequential test (that includes a reduced rank pre-test) of $H_0 : \lambda_1 = 0$ with different levels of significance, assuming that the factors and returns are generated from a multivariate normal distribution. The nominal levels of the rank pre-test and the second-stage test of $H_0 : \lambda_1 = 0$ are set equal to each other. We consider two model specifications (CAPM and FF3) that are calibrated to monthly data for the period January 1932 – December 2006. Panel A presents the empirical size of the sequential test for $\boldsymbol{\alpha} = \mathbf{q} = [1,\ \mathbf{0}'_{m-1}]'$, i.e., $\boldsymbol{\alpha}$ is in the span of the column space of $\mathbf{D}_0$. Panel B reports the empirical size of the sequential test for $\boldsymbol{\alpha} = \mathbf{1}_m$, i.e., $\boldsymbol{\alpha}$ is not in the span of the column space of $\mathbf{D}_0$. The results for different number of time series observations (T) are based on 100,000 simulations.

FURTHER RESULTS ON THE LIMITING DISTRIBUTION OF GMM SAMPLE MOMENT CONDITIONS

Nikolay Gospodinov, Raymond Kan, and Cesare Robotti

Supplementary Material

SIMULATION SETUP

This appendix contains some additional simulation results regarding the properties of the standard normal test, the weighted χ^2 test, the LM rank test, and the sequential test considered in the paper. In the simulation experiment, the factors ($\mathbf{f}$) and the returns ($\mathbf{R}$) on the test assets for the CAPM (1 factor and 11 test asset returns) and FF3 (3 factors and 26 test asset returns) are drawn from a multivariate normal distribution with a covariance matrix estimated from the data. The mean return vector is chosen such that the asset pricing model holds exactly for the test assets. For each simulated set of returns and factors, the unknown parameters $\boldsymbol{\theta}_0$ of the linear SDF $y(\boldsymbol{\theta}_0) = \tilde{\mathbf{f}}'\boldsymbol{\theta}_0$, where $\tilde{\mathbf{f}} = (1,\ \mathbf{f}')'$, are estimated by minimizing the sample HJ-distance, which yields

$$\hat{\boldsymbol{\theta}} = (\hat{\mathbf{D}}_T'\mathbf{W}_T\hat{\mathbf{D}}_T)^{-1}(\hat{\mathbf{D}}_T'\mathbf{W}_T\mathbf{q}), \tag{1}$$

where $\hat{\mathbf{D}}_T = \frac{1}{T}\sum_{t=1}^T \mathbf{R}_t\tilde{\mathbf{f}}_t'$, $\mathbf{W}_T = \left(\frac{1}{T}\sum_{t=1}^T \mathbf{R}_t\mathbf{R}_t'\right)^{-1}$, and $\mathbf{q} = [1,\ \mathbf{0}_{m-1}']'$. The estimated Lagrange multipliers are given by

$$\hat{\boldsymbol{\lambda}} = \mathbf{W}_T\left[\frac{1}{T}\sum_{t=1}^{T}\mathbf{R}_t y_t(\hat{\boldsymbol{\theta}}) - \mathbf{q}\right], \tag{2}$$

where $y_t(\hat{\boldsymbol{\theta}}) = \tilde{\mathbf{f}}_t'\hat{\boldsymbol{\theta}}$.

We consider linear combinations of sample Lagrange multipliers with different choices of an $m \times 1$ nonzero weighting vector $\boldsymbol{\alpha}$, i.e., $\boldsymbol{\alpha}'\hat{\boldsymbol{\lambda}}$. Let matrix $\mathbf{Q}_c$ denote the null space of the p vector $E[\tilde{\mathbf{f}}_t\tilde{\mathbf{f}}_t']\boldsymbol{\theta}_0$ and $\mathbf{Q}_c^1$ be the first column of $\mathbf{Q}_c$. Also, let $\boldsymbol{\Pi} = \mathbf{P}_{\boldsymbol{\alpha}}'\mathbf{D}_0$, where $\mathbf{P}_{\boldsymbol{\alpha}}$ is an $m \times (m-1)$ orthonormal matrix whose columns are orthogonal to $\boldsymbol{\alpha}$. In Tables I through IV, we analyze the empirical sizes of four tests – (i) standard normal test of $H_0 : \boldsymbol{\alpha}'\boldsymbol{\lambda} = 0$, (ii) weighted χ^2 test of $H_0 : \boldsymbol{\alpha}'\boldsymbol{\lambda} = 0$, (iii) LM rank test of $H_0 : \text{rank}(\boldsymbol{\Pi}) = p - 1$, and (iv) sequential test of $H_0 : \boldsymbol{\alpha}'\boldsymbol{\lambda} = 0$ with a pre-test of $H_0 : \text{rank}(\boldsymbol{\Pi}) = p - 1$, using three choices of $\boldsymbol{\alpha}$:

1. $\boldsymbol{\alpha} = \mathbf{q} = [1\ ,\mathbf{0}_{m-1}']'$,
2. $\boldsymbol{\alpha} = \mathbf{D}_0\mathbf{1}_p$,
3. $\boldsymbol{\alpha} = \mathbf{D}_0\mathbf{Q}_c^1$.

We also analyze the statistical properties of the rank and sequential tests when $\boldsymbol{\alpha}$ in not in the span of the column space of $\mathbf{D}_0$. Specifically, in Table V, we analyze the empirical power of the rank test for $\boldsymbol{\alpha} = \mathbf{1}_m$ and $\boldsymbol{\alpha} = \sqrt{m}\mathbf{q} + \mathbf{1}_m$. In Table VI, we report results for the empirical size of the sequential test for $\boldsymbol{\alpha} = \mathbf{1}_m$ and $\boldsymbol{\alpha} = \sqrt{m}\mathbf{q} + \mathbf{1}_m$. The empirical rejection probabilities are computed based on 100,000 Monte Carlo replications.

STANDARD NORMAL TEST

Panels A and B of Table I show that the use of the standard normal test leads to severe over-rejections when $\boldsymbol{\alpha}$ is in the span of the column space of $\mathbf{D}_0$. By contrast, the normal test behaves well in Panel C. These simulation results can be explained using the theoretical results in Lemma 6 in the paper. In particular, in Panel A we have $\boldsymbol{\alpha} = \mathbf{q}$ and $r = -1$, and the t-test is asymptotically distributed as $-\sqrt{\chi^2_{m-p}}$. In Panel B, the squared t-test follows a mixture of two independent chi-squared random variables with $m-p$ and one degrees of freedom. Finally, in Panel C, $\boldsymbol{\alpha}$ is set such that $r_2 = 0$ (and $r = 0$) and the t-test follows a standard normal distribution which explains why the t-test works well in this setup.[1]

WEIGHTED χ^2 TEST

In Table II, we report the empirical size of the weighted χ^2 test. For the CAPM, our asymptotic approximation works very well even for relatively small sample sizes. For FF3, we need a larger T for the asymptotic approximation to work well. This is a well-known problem in empirical asset pricing that arises when the number of test assets m is large relative to T (see, e.g., Ahn and Gadarowski, 2004).

RANK TEST

Tables III and V report the empirical size and power of the rank test. Overall, the test has excellent size and power properties. Some modest under-rejections only occur for FF3 when $T =$ 150.

SEQUENTIAL TEST

In Tables IV and VI, we analyze the empirical size of the sequential test (that includes a reduced rank pre-test) of $H_0 : \lambda_1 = 0$ when $\boldsymbol{\alpha}$ is in the span of the column space of $\mathbf{D}_0$ and when $\boldsymbol{\alpha}$ is not. The sequential test we consider has the following structure. If we reject the null of reduced rank,

[1] Note that our conclusions are not affected by the particular choice of the column of $\mathbf{Q}_c$ (the matrix described in the simulation setup).

then we use the normal test in the second stage; otherwise, we use the weighted chi-squared test. Acceptance and rejection of $H_0 : \boldsymbol{\alpha}'\boldsymbol{\lambda} = 0$ is based on the outcome of the second test. Let η_1 be the asymptotic size of the rank restriction test and η_2 be the asymptotic size of either the normal test or the weighted chi-squared test used in the second stage.

When $\boldsymbol{\alpha}$ is in the span of the column space of $\mathbf{D}_0$ (Table IV), the rank restriction test will accept the null of reduced rank with probability $1-\eta_1$ (asymptotically). Therefore, the probability of using the normal test in the second stage is η_1. Unconditionally, the normal test will reject with probability $p_1 \geq \eta_2$ (in our simulation setup) and the weighted chi-squared test will reject with probability η_2. Therefore, if the two tests are independent, the size of the sequential test is given by

$$\eta_1 p_1 + (1-\eta_1)\eta_2 \geq \eta_2.$$

In general, the two tests are dependent because both the rank restriction test and the test of $H_0 : \boldsymbol{\alpha}'\boldsymbol{\lambda} = 0$ are specification tests. In this case, we can only establish an upper bound on the probability of rejection of the sequential test, which is given by

$$\eta_1 + \eta_2.$$

When $\boldsymbol{\alpha}$ is not in the span of the column space of $\mathbf{D}_0$ (Table VI), the rank restriction test will reject the null of reduced rank with probability one (asymptotically), so the normal test will be chosen in the second stage. As a result, the asymptotic size of the sequential test is simply η_2.

The results in Tables IV and VI (which are obtained by setting the asymptotic sizes of the first and second tests equal to each other, i.e., $\eta_1 = \eta_2$) show that the proposed sequential test tends to behave well in our simulation setup.

REFERENCES

[1] Ahn, S. C., and C. Gadarowski (2004): "Small Sample Properties of the Model Specification Test Based on the Hansen-Jagannathan Distance," *Journal of Empirical Finance*, 11, 109–132.

Table I
Empirical Size of the Standard Normal Test

Panel A: $\boldsymbol{\alpha} = \mathbf{q} = [1\ , \mathbf{0}'_{m-1}]'$

	CAPM			FF3		
	Level of Significance			Level of Significance		
T	10%	5%	1%	10%	5%	1%
150	0.978	0.929	0.689	1.000	1.000	1.000
300	0.977	0.925	0.682	1.000	1.000	1.000
450	0.976	0.923	0.679	1.000	1.000	0.999
600	0.976	0.924	0.679	1.000	1.000	0.999
750	0.975	0.923	0.679	1.000	1.000	0.999
900	0.976	0.923	0.680	1.000	1.000	0.999

Panel B: $\boldsymbol{\alpha} = \mathbf{D}_0\mathbf{1}_p$

	CAPM			FF3		
	Level of Significance			Level of Significance		
T	10%	5%	1%	10%	5%	1%
150	0.968	0.910	0.661	1.000	1.000	0.998
300	0.965	0.907	0.650	1.000	1.000	0.998
450	0.964	0.904	0.650	1.000	1.000	0.998
600	0.965	0.905	0.647	1.000	1.000	0.998
750	0.966	0.904	0.648	1.000	1.000	0.998
900	0.965	0.904	0.648	1.000	1.000	0.997

Panel C: $\boldsymbol{\alpha} = \mathbf{D}_0\mathbf{Q}_c^1$

	CAPM			FF3		
	Level of Significance			Level of Significance		
T	10%	5%	1%	10%	5%	1%
150	0.129	0.071	0.017	0.187	0.115	0.037
300	0.114	0.059	0.013	0.141	0.079	0.020
450	0.109	0.056	0.012	0.127	0.068	0.017
600	0.107	0.055	0.012	0.120	0.063	0.015
750	0.106	0.053	0.011	0.117	0.062	0.014
900	0.105	0.053	0.011	0.115	0.060	0.013

Table II
Empirical Size of the Weighted χ^2 Test

Panel A: $\boldsymbol{\alpha} = \mathbf{q} = [1\ , \mathbf{0}'_{m-1}]'$

	CAPM			FF3		
	Level of Significance			Level of Significance		
T	10%	5%	1%	10%	5%	1%
150	0.144	0.082	0.022	0.284	0.189	0.072
300	0.121	0.065	0.015	0.178	0.105	0.031
450	0.115	0.060	0.014	0.151	0.084	0.022
600	0.111	0.057	0.013	0.138	0.074	0.018
750	0.109	0.057	0.012	0.130	0.070	0.016
900	0.107	0.055	0.011	0.125	0.067	0.015

Panel B: $\boldsymbol{\alpha} = \mathbf{D}_0\mathbf{1}_p$

	CAPM			FF3		
	Level of Significance			Level of Significance		
T	10%	5%	1%	10%	5%	1%
150	0.124	0.068	0.018	0.209	0.137	0.052
300	0.111	0.058	0.013	0.136	0.077	0.021
450	0.109	0.057	0.012	0.123	0.066	0.015
600	0.106	0.054	0.012	0.115	0.061	0.014
750	0.105	0.054	0.011	0.112	0.058	0.013
900	0.104	0.054	0.012	0.112	0.058	0.012

Panel C: $\boldsymbol{\alpha} = \mathbf{D}_0\mathbf{Q}_c^1$

	CAPM			FF3		
	Level of Significance			Level of Significance		
T	10%	5%	1%	10%	5%	1%
150	0.132	0.072	0.018	0.185	0.111	0.034
300	0.116	0.061	0.013	0.138	0.076	0.019
450	0.109	0.056	0.012	0.124	0.067	0.016
600	0.108	0.055	0.012	0.119	0.062	0.014
750	0.108	0.054	0.011	0.115	0.060	0.013
900	0.105	0.053	0.010	0.111	0.059	0.013

Table III
Empirical Size of the Rank Test

Panel A: $\boldsymbol{\alpha} = \mathbf{q} = [1\ , \mathbf{0}'_{m-1}]'$

	CAPM			FF3		
	Level of Significance			Level of Significance		
T	10%	5%	1%	10%	5%	1%
150	0.095	0.044	0.007	0.069	0.024	0.001
300	0.098	0.048	0.009	0.093	0.044	0.007
450	0.099	0.050	0.009	0.098	0.047	0.008
600	0.099	0.049	0.010	0.099	0.047	0.009
750	0.100	0.050	0.010	0.100	0.049	0.009
900	0.099	0.050	0.010	0.100	0.050	0.009

Panel B: $\boldsymbol{\alpha} = \mathbf{D}_0 \mathbf{1}_p$

	CAPM			FF3		
	Level of Significance			Level of Significance		
T	10%	5%	1%	10%	5%	1%
150	0.096	0.045	0.007	0.072	0.026	0.001
300	0.099	0.047	0.009	0.093	0.043	0.007
450	0.100	0.050	0.010	0.098	0.046	0.008
600	0.100	0.050	0.010	0.098	0.048	0.008
750	0.101	0.050	0.010	0.100	0.048	0.009
900	0.101	0.050	0.010	0.100	0.050	0.009

Panel C: $\boldsymbol{\alpha} = \mathbf{D}_0 \mathbf{Q}_c^1$

	CAPM			FF3		
	Level of Significance			Level of Significance		
T	10%	5%	1%	10%	5%	1%
150	0.084	0.036	0.004	0.048	0.015	0.001
300	0.093	0.044	0.007	0.079	0.033	0.004
450	0.097	0.046	0.008	0.088	0.039	0.006
600	0.097	0.046	0.008	0.091	0.043	0.007
750	0.097	0.047	0.008	0.094	0.044	0.008
900	0.097	0.048	0.009	0.095	0.045	0.008

Table IV
Empirical Size of the Sequential Test
When $\boldsymbol{\alpha}$ is in the Span of the Column Space of $\mathbf{D}_0$

Panel A: $\boldsymbol{\alpha} = \mathbf{q} = [1\ , \mathbf{0}'_{m-1}]'$

	CAPM			FF3		
	Level of Significance			Level of Significance		
T	10%	5%	1%	10%	5%	1%
150	0.145	0.082	0.022	0.284	0.189	0.072
300	0.121	0.065	0.015	0.178	0.105	0.031
450	0.115	0.060	0.014	0.151	0.085	0.022
600	0.111	0.058	0.013	0.138	0.074	0.018
750	0.109	0.057	0.012	0.130	0.070	0.016
900	0.107	0.055	0.011	0.125	0.067	0.015

Panel B: $\boldsymbol{\alpha} = \mathbf{D}_0\mathbf{1}_p$

	CAPM			FF3		
	Level of Significance			Level of Significance		
T	10%	5%	1%	10%	5%	1%
150	0.141	0.072	0.018	0.210	0.137	0.052
300	0.146	0.072	0.014	0.145	0.080	0.021
450	0.149	0.075	0.015	0.143	0.073	0.016
600	0.149	0.074	0.015	0.142	0.072	0.015
750	0.149	0.074	0.015	0.145	0.072	0.015
900	0.149	0.075	0.015	0.147	0.074	0.015

Panel C: $\boldsymbol{\alpha} = \mathbf{D}_0\mathbf{Q}_c^1$

	CAPM			FF3		
	Level of Significance			Level of Significance		
T	10%	5%	1%	10%	5%	1%
150	0.119	0.067	0.017	0.180	0.110	0.034
300	0.103	0.055	0.012	0.130	0.073	0.019
450	0.095	0.050	0.012	0.116	0.063	0.015
600	0.094	0.049	0.011	0.110	0.058	0.014
750	0.093	0.048	0.010	0.106	0.056	0.013
900	0.091	0.047	0.010	0.102	0.055	0.012

Table V
Empirical Power of the Rank Test

Panel A: $\boldsymbol{\alpha} = \mathbf{1}_m$

	CAPM			FF3		
	Level of Significance			Level of Significance		
T	10%	5%	1%	10%	5%	1%
150	0.999	0.997	0.965	0.977	0.913	0.531
300	1.000	1.000	1.000	1.000	1.000	1.000
450	1.000	1.000	1.000	1.000	1.000	1.000
600	1.000	1.000	1.000	1.000	1.000	1.000
750	1.000	1.000	1.000	1.000	1.000	1.000
900	1.000	1.000	1.000	1.000	1.000	1.000

Panel B: $\boldsymbol{\alpha} = \sqrt{m}\mathbf{q} + \mathbf{1}_m$

	CAPM			FF3		
	Level of Significance			Level of Significance		
T	10%	5%	1%	10%	5%	1%
150	0.999	0.997	0.965	0.974	0.904	0.508
300	1.000	1.000	1.000	1.000	1.000	1.000
450	1.000	1.000	1.000	1.000	1.000	1.000
600	1.000	1.000	1.000	1.000	1.000	1.000
750	1.000	1.000	1.000	1.000	1.000	1.000
900	1.000	1.000	1.000	1.000	1.000	1.000

Table VI
Empirical Size of the Sequential Test
When $\boldsymbol{\alpha}$ is not in the Span of the Column Space of $\mathbf{D}_0$

Panel A: $\boldsymbol{\alpha} = \mathbf{1}_m$

	CAPM			FF3		
	Level of Significance			Level of Significance		
T	10%	5%	1%	10%	5%	1%
150	0.123	0.067	0.022	0.177	0.124	0.130
300	0.110	0.057	0.012	0.132	0.072	0.018
450	0.106	0.054	0.011	0.121	0.065	0.014
600	0.104	0.053	0.011	0.116	0.061	0.013
750	0.104	0.052	0.011	0.112	0.059	0.013
900	0.103	0.051	0.011	0.110	0.057	0.012

Panel B: $\boldsymbol{\alpha} = \sqrt{m}\mathbf{q} + \mathbf{1}_m$

	CAPM			FF3		
	Level of Significance			Level of Significance		
T	10%	5%	1%	10%	5%	1%
150	0.124	0.065	0.022	0.211	0.151	0.142
300	0.110	0.057	0.012	0.151	0.086	0.023
450	0.108	0.054	0.011	0.134	0.073	0.018
600	0.106	0.053	0.011	0.126	0.067	0.016
750	0.105	0.052	0.011	0.119	0.063	0.015
900	0.104	0.052	0.010	0.116	0.061	0.013