

# Identification robust inference on risk premia of mimicking portfolios of macroeconomic factors

Frank Kleibergen\*      Zhaoguo Zhan†

September 28, 2016

## Abstract

Mimicking portfolios of macroeconomic factors are commonly constructed by projecting these factors on a set of base assets. We show that when macroeconomic factors are associated with small betas, the beta estimator using their mimicking portfolios has a non-standard limiting distribution. The non-standard behavior of the beta estimator jeopardizes inference on risk premia in the commonly used Fama and MacBeth (1973) two-pass procedure. To remedy this problem, we propose a test for risk premia on mimicking portfolios. The validity of this test does not depend on the magnitude of the betas. Simulation evidence suggests that the proposed test performs well in size and power. We use the test to analyze the risk premium on the leverage factor of Adrian et. al. (2014). Our results indicate that the leverage factor is a weak factor which leads to substantially different results for its risk premium.

JEL Classification: G12

Keywords: asset pricing; macroeconomic factor; mimicking portfolio; non-standard distribution; robust inference

---

\*Amsterdam School of Economics, University of Amsterdam, Roetersstraat 11, 1018 WB Amsterdam, The Netherlands. Email: f.r.kleibergen@uva.nl.

†Department of Economics, Finance and Quantitative Analysis, Coles College of Business, Kennesaw State University, GA 30144, USA. Email: zzhan@kennesaw.edu.

# 1 Introduction

An intersection of macroeconomics and finance is that a large group of macroeconomic factors have been found useful for financial asset pricing. These factors include, e.g., consumption growth in Breeden et al. (1989), labor income growth in Jagannathan and Wang (1996), consumption-wealth ratio in Lettau and Ludvigson (2001), GDP growth in Vassalou (2003), investment growth in Li et al. (2006), among many others. The prevalence of these macroeconomic factors has recently led to a sizeable and growing literature that scrutinizes their usefulness.

One major concern on macroeconomic factors is that their minor correlation with asset returns (see, e.g., Bai and Ng 2006) invalidates conventional inference methods used in asset pricing studies, such as the  $t$ -test on risk premia in the Fama and MacBeth (1973) (FM) two-pass methodology. Consequently, empirical support for macroeconomic factors based on conventional inference methods is up to careful scrutiny. An early contribution along this line is Kan and Zhang (1999), who show that the  $t$ -test in the FM methodology can spuriously support useless factors that are independent of asset returns. More recently, Kleibergen (2009) further warns that when factors are only weakly correlated with asset returns so their betas are small, inference on risk premia based on the FM two-pass procedure is also spurious.

Instead of using macroeconomic factors themselves, their mimicking portfolios are also widely used to replace these factors in asset pricing studies. The theoretical support for such practice is provided by Breeden (1979) and Huberman et al. (1987), who establish that factors can be replaced by their mimicking portfolios for asset pricing tests. In terms of macroeconomic factors, the norm of constructing their mimicking portfolios is to project these factors on a set of base assets that span the asset space. This projection is implemented by regressing factors on base assets in a time series regression. The resulting mimicking portfolios after projection are also called maximum correlation portfolios. See, e.g., Breeden et al. (1989), Lamont (2001), Vassalou (2003), Avramov and Chordia (2006) and Adrian et al. (2014).

On the one hand, besides other potential advantages,<sup>1</sup> using mimicking portfolios instead of macroeconomic factors appears able to bypass the weak statistical correlation issue studied in Kleibergen (2009), etc. By projecting macroeconomic factors on base assets, it is natural that the resulting mimicking portfolios exhibit improved correlation with asset returns so their betas can be amplified. In existing studies (e.g., Vassalou 2003, Adrian et al. 2014), this projection is commonly interpreted as a way to remove the noise in macroeconomic factors while keeping only their relevant information for asset pricing. Alongside this interpretation, inference on risk premia using mimicking portfolios is thus believed to be more informative, compared to that using the original macroeconomic factors.

On the other hand, using mimicking portfolios instead of macroeconomic factors has its own costs, which are rarely discussed in the existing literature. In this paper, we reveal the consequences of using mimicking portfolios in the FM two-pass procedure, when their background macroeconomic factors are only minorly correlated with asset returns. We show that although these mimicking portfolios may appear relevant for asset pricing, the risk premia estimator in the FM two-pass procedure has a non-standard distribution, which jeopardizes the  $t$ -test on risk premia. The underlying reason is that when betas of macroeconomic factors are small, the beta estimator of their mimicking portfolios is associated with non-standard limiting behavior, which further puts the  $t$ -test under doubt.

Our findings, therefore, suggest that using mimicking portfolios instead of macroeconomic factors for asset pricing does not necessarily imply improved inference on risk premia or more generally, improved performance of asset pricing tests. Ironically, since macroeconomic factors are almost indistinguishable from useless factors in terms of their minor correlation with asset returns in finite samples, empirical findings based on mimicking portfolios could be driven by the noise occurred in the construction of such portfolios, rather than the relevant pricing information contained in macroeconomic factors and preserved in the construction of their mimicking portfolios.

---

<sup>1</sup>Such as extending the available data sets (Ang et al. 2006), providing the extra restriction for testing purposes (Huberman et al. 1987).

Since the conventional  $t$ -test on risk premia of mimicking portfolios resulting from macroeconomic factors is under doubt, we propose a novel test in this paper. Unlike the  $t$ -test, the asymptotic size of our test does not depend on the magnitude of the betas, so it remains trustworthy when macroeconomic factors are associated with small betas. The suggested test is in line with the robust tests proposed by Kleibergen (2009) for risk premia of factors, and the extension to the mimicking portfolio setting is new. The performance of the test is examined by Monte Carlo simulation.

For illustrative purposes, we apply the proposed test to the leverage factor model from Adrian et al. (2014). Following Adrian et al. (2014), we construct mimicking portfolios by regressing the leverage factor on seven base assets: excess returns of the six Fama-French portfolios on size and book-to-market, plus the momentum factor. By inverting the proposed test, we further obtain confidence intervals for the risk premium of the constructed mimicking portfolios, with the usage of various test assets. We show that these confidence intervals differ substantially from those that result from inverting the conventional  $t$ -test, which is consistent with the fact that the leverage factor is associated with a small beta.

Related to our paper, there exists a sizeable literature that examines identification and inference issues in asset pricing. Besides those related articles we cite above, see, e.g., Burnside (2016), Balduzzi and Robotti (2008). Unlike this paper, Burnside (2016) does not focus on mimicking portfolios, but shows that when factors are weak, lack of identification exists in linear stochastic discount factor models. The rank test of Kleibergen and Paap (2006) is consequently suggested in Burnside (2016) for model diagnostics. Balduzzi and Robotti (2008) do not focus on the FM two-pass procedure, but argue that the use of mimicking portfolios improves inference on risk premia estimated by time series regressions. Furthermore, Balduzzi and Robotti (2008) do not consider the consequence of weak factors for inference on risk premia.

The paper proceeds as follows. In Section 2, we describe a linear factor model with observed factors and present the limiting behavior of the beta estimator under constructed mimicking portfolios. In Section 3, we propose the robust test for risk premia of mimicking portfolios and

provide simulation evidence for its size and power. Our application to data from Adrian et al. (2014) is also contained in Section 3. Section 4 concludes the paper. Proofs and technical details are provided in the Appendix.

Throughout the paper, we use the following notation:  $\mu$  is used for mean, while  $V$  is for covariance;  $\hat{\mu}$  and  $\hat{V}$  are the sample analogs of  $\mu$  and  $V$ , respectively;  $\iota_N$  is the  $N \times 1$  dimensional vector of ones,  $vec(A)$  stands for the column vectorization of a matrix  $A$ ,  $P_A = A(A'A)^{-1}A'$ ,  $M_A = I - P_A$ ,  $I$  is the identity matrix; “ $\xrightarrow{p}$ ” and “ $\xrightarrow{d}$ ” stand for convergence in probability and convergence in distribution, respectively.

## 2 Weak identification with mimicking portfolios

### 2.1 Linear factor model and observed factors

Asset returns are to a large extent explained by a small number of factors, see e.g. Merton (1973), Ross (1976), Roll and Ross (1980), Chamberlain and Rothschild (1983) and Connor and Korajczyk (1988, 1989). We therefore use a linear factor model for asset returns  $R_t$  ( $N \times 1$ ) with  $k$  (unobserved) factors  $F_t$  ( $k \times 1$ ):

$$R_t = \iota_N \lambda_0 + \beta(\bar{F}_t + \lambda_F) + v_t \quad (1)$$

where  $\bar{F}_t = F_t - \mu_F$ ,  $v_t$  is the mean zero error term,  $t = 1, \dots, T$ . Equation (1) implies the moment condition for the risk premia, i.e.,

$$E(R_t) = \iota_N \lambda_0 + \beta \lambda_F \quad (2)$$

with  $\lambda_0$ : the zero- $\beta$  return,  $\lambda_F$ : the  $k \times 1$  vector of factor risk premia.

Moreover, there exists a partition of  $R_t = (R'_{1t}, R'_{2t})'$ , with  $R_{1t} : N_1 \times 1$ ,  $R_{2t} : N_2 \times 1$ ,  $N_1 + N_2 = N$ . From now on, we consider  $R_{1t}$  as the returns on test assets, and  $R_{2t}$  as the

returns on base assets that are used for constructing mimicking portfolios.<sup>2</sup> Correspondingly,  $\beta = (\beta'_1, \beta'_2)'$  with  $\beta_1 : N_1 \times k$ ,  $\beta_2 : N_2 \times k$ , and  $v_t = (v'_{1t}, v'_{2t})'$  with  $v_{1t} : N_1 \times 1$ ,  $v_{2t} : N_2 \times 1$ .

Let  $G_t$  ( $m \times 1$ ) be the observed proxy for the unobserved factor  $F_t$ , i.e.,

$$\bar{F}_t = \delta \bar{G}_t + u_t \quad (3)$$

where  $\bar{G}_t = G_t - \mu_G$ ,  $u_t$  is the error term. Note that  $\delta$  ( $k \times m$ ) reflects the quality of the approximation of  $F_t$  by  $G_t$ . Ideally  $F_t$  is observed so  $G_t$  coincides with it and  $\delta = I_k$ ,  $u_t = 0$ . On the other hand, if  $\delta$  is approximately equal to zero, then  $F_t$  is poorly proxied by  $G_t$ .

Plugging (3) into (1), and if the vector of risk premia of  $G_t$ , denoted by  $\lambda_G$ , is defined by the moment condition  $E(R_t) = \iota_N \lambda_0 + \beta \delta \lambda_G$ , then  $\lambda_F$  and  $\lambda_G$  are also related by  $\delta$  since  $\lambda_F = \delta \lambda_G$ .<sup>3</sup> In addition, (1) is re-written as

$$R_t = \iota_N \lambda_0 + \beta \delta (\bar{G}_t + \lambda_G) + e_t \quad (4)$$

where  $e_t = (e'_{1t}, e'_{2t})' = \beta u_t + v_t$ , and  $\Sigma = \beta \Lambda \beta' + \Omega$  with  $\Sigma = \text{var}(e_t)$ ,  $\Lambda = \text{var}(u_t)$ ,  $\Omega = \text{var}(v_t)$ .

We make the following assumptions for the model described above.

**Assumption 1**  $\frac{1}{\sqrt{T}} \sum_{t=1}^T \left( \begin{pmatrix} 1 \\ F_t \end{pmatrix} \otimes (R_t - \iota_N \lambda_0 - \beta(\bar{F}_t + \lambda_F)) \right) \xrightarrow{d} \begin{pmatrix} \varphi_R \\ \varphi_F \end{pmatrix}$ , where  $\begin{pmatrix} \varphi_R \\ \varphi_F \end{pmatrix} \sim N(0, Q_F \otimes \Omega)$ ,  $Q_F = \begin{pmatrix} 1 & \mu'_F \\ \mu_F & V_{FF} + \mu_F \mu'_F \end{pmatrix}$ ,  $\Omega = \text{var}((v'_{1t}, v'_{2t})') = \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix}$ .

**Assumption 2** All the covariance of  $R_{1t}$  and  $R_{2t}$  is captured by the factors so  $\Omega_{12} = \Omega'_{21} = 0$  and  $\text{cov}(R_{1t}, R_{2t}) = \beta_1 V_{FF} \beta'_2$ .

---

<sup>2</sup>We acknowledge that in practice test assets and base assets need not be in the same frequency or in the same time period. For convenience, we focus on the same frequency and same time period setting, although the message of this paper can be similarly extended to more generalized settings.

<sup>3</sup>Note that  $\lambda_G$  is not always well-defined: if  $G_t$  is useless with  $\delta = 0$ , then  $\lambda_G$  is not defined.

**Assumption 3**  $\frac{1}{\sqrt{T}} \sum_{t=1}^T \left( \begin{pmatrix} 1 \\ G_t \end{pmatrix} \otimes (R_t - \iota_N \lambda_0 - \beta(\bar{F}_t + \lambda_F)) \right) \xrightarrow{d} \begin{pmatrix} \varphi_R \\ \varphi_G \end{pmatrix}$ , where  $\begin{pmatrix} \varphi_R \\ \varphi_G \end{pmatrix} \sim N(0, Q_G \otimes \Omega)$ ,  $Q_G = \begin{pmatrix} 1 & \mu'_G \\ \mu_G & V_{GG} + \mu_G \mu'_G \end{pmatrix}$ .

Assumption 1 is a central limit theorem that is also imposed in Kleibergen (2009). Assumption 2 is a necessary condition for the consistency of the mimicking portfolio estimator of  $\beta_1$ , which we show below.<sup>4</sup> Assumption 3 is a corollary of Assumption 1, with  $G_t$  replacing the unknown  $F_t$ .

Assumptions 1-3 just result from the model in (1) and (3) with i.i.d. mean zero errors and finite variance.

## 2.2 Beta of mimicking portfolios

In accordance with common practice, we construct mimicking portfolios by projecting  $G_t$  on base assets  $R_{2t}$ . The resulting feasible mimicking portfolios are then written as:

$$\widehat{V}_{GR_2} \widehat{V}_{R_2 R_2}^{-1} R_{2t} \quad (5)$$

where  $\widehat{V}_{GR_2} \widehat{V}_{R_2 R_2}^{-1}$  is the sample counterpart of the infeasible  $V_{GR_2} V_{R_2 R_2}^{-1}$ , and it can be obtained by regressing  $G_t$  on  $R_{2t}$  in a time series regression.

With  $R_{1t}$  as test assets, the beta estimator for the mimicking portfolios in (5) reads:

$$\widehat{\beta}_1 = \widehat{V}_{R_1 R_2} \widehat{V}_{R_2 R_2}^{-1} \widehat{V}_{R_2 G} (\widehat{V}_{GR_2} \widehat{V}_{R_2 R_2}^{-1} \widehat{V}_{R_2 G})^{-1} \quad (6)$$

**Theorem 1** *When Assumptions 1-3 hold, the limiting behavior of  $\widehat{\beta}_1$  can be described as follows.*

1. *When  $\delta$  is fixed and the number of elements of  $G$  equals the number of elements of  $F$ , so*

---

<sup>4</sup>When test assets  $R_{1t}$  are base assets  $R_{2t}$ , Assumption 2 is dropped.

$\delta$  is a square invertible matrix:

$$\hat{\beta}_1 \xrightarrow{p} \beta_1 V_{FF} \delta^{-1} V_{GG}^{-1}$$

and when  $G_t = F_t$ ,  $\delta = I_k : \hat{\beta}_1 \xrightarrow{p} \beta_1$ .

2. When  $\delta = d/\sqrt{T}$ , with  $d$  a fixed full rank matrix:

$$T^{-\frac{1}{2}} \hat{\beta}_1 \xrightarrow{d} \beta_1 V_{FF} \beta_2' (\beta_2 V_{FF} \beta_2' + \Omega_{22})^{-1} (\beta_2 (dV_{GG} + \psi_{uG}) + \psi_{v_2G}) \\ \left[ (\beta_2 (dV_{GG} + \psi_{uG}) + \psi_{v_2G})' (\beta_2 V_{FF} \beta_2' + \Omega_{22})^{-1} (\beta_2 (dV_{GG} + \psi_{uG}) + \psi_{v_2G}) \right]^{-1}$$

where  $\frac{1}{\sqrt{T}} \sum_{t=1}^T (\bar{G}_t \otimes u_t) \xrightarrow{d} \text{vec}(\psi_{uG})$ ,  $\frac{1}{\sqrt{T}} \sum_{t=1}^T (\bar{G}_t \otimes v_{2t}) \xrightarrow{d} \text{vec}(\psi_{v_2G})$ .

**Proof.** see Appendix A. ■

The first part of Theorem 1 shows that the beta estimator of the mimicking portfolios is a consistent estimator for the beta of the factors when mimicking portfolios are constructed from accurate proxies of the underlying unobserved factors. The proof of Theorem 1 in Appendix A also shows that the covariance of test assets and base assets must be fully captured by the factors to render the beta estimator consistent as implied by Assumption 2. Put differently, if Assumption 2 is not satisfied, then the estimated beta of the mimicking portfolios does not converge to the beta of the factors, even when the true underlying factors are used to construct the mimicking portfolios.

The motivation of this paper comes from the second part of Theorem 1, i.e., when mimicking portfolios result from observed factors that are only poor (weak) proxies for underlying factors. It is known that macroeconomic factors commonly exhibit minor correlation with asset returns, so they are likely to be poor proxies for underlying factors, as reflected by  $\delta = d/\sqrt{T}$ . In this scenario, Theorem 1 shows that the beta estimator of the mimicking portfolios is increasing with the sample size and has a non-standard distribution. In other words, when the betas of the macroeconomic factors are small, the mimicking portfolios of such factors can be associated



with betas that are spuriously large.

It is worth emphasizing that  $\delta = d/\sqrt{T}$  is adopted from the weak-instrument assumption made in econometrics (see, e.g., Staiger and Stock 1997), in order to appropriately reflect the case when  $G_t$  is a poor proxy for  $F_t$ . A similar treatment can be found in Kleibergen (2009) and Kleibergen and Zhan (2015).

In the commonly used FM two-pass procedure, the risk premia of the mimicking portfolios are obtained by regressing the sample average of the test assets  $\hat{\mu}_{R_1} = \frac{1}{T} \sum_{t=1}^T R_{1t}$  on  $\hat{\beta}_1$  and an intercept, i.e.,

$$\begin{pmatrix} \hat{\lambda}_0 \\ \hat{\lambda}_G \end{pmatrix} = \left[ (\iota_{N_1} \vdots \hat{\beta}_1)' (\iota_{N_1} \vdots \hat{\beta}_1) \right]^{-1} (\iota_{N_1} \vdots \hat{\beta}_1)' \hat{\mu}_{R_1} \quad (7)$$

where  $\hat{\lambda}_G$  ( $m \times 1$ ) denotes the estimated risk premia for the mimicking portfolios that are constructed from  $G_t$ .

If  $\hat{\beta}_1$  has a non-standard distribution, it is natural to expect that the performance of  $\hat{\lambda}_G$  is under doubt, as stated in the corollary below.

**Corollary 1** *When Assumptions 1-3 hold, the limiting behavior of  $\hat{\lambda}_G$  can be described as follows.*

1. *When  $\delta$  is fixed and the number of elements of  $G$  equals the number of elements of  $F$ , so  $\delta$  is a square invertible matrix:*

$$\hat{\lambda}_G \xrightarrow{p} V_{GG} \delta' V_{FF}^{-1} \lambda_F$$

*and when  $G_t = F_t$ ,  $\delta = I_k : \hat{\lambda}_G \xrightarrow{p} \lambda_F$ .*

2. *When  $\delta = d/\sqrt{T}$ :*

$$\sqrt{T} \hat{\lambda}_G \xrightarrow{d} (\Psi'_{\beta_1} M_{L_{N_1}} \Psi_{\beta_1})^{-1} \Psi'_{\beta_1} M_{L_{N_1}} \mu_{R_1} + \frac{1}{\sqrt{T}} (\Psi'_{\beta_1} M_{L_{N_1}} \Psi_{\beta_1})^{-1} \Psi'_{\beta_1} M_{L_{N_1}} \psi_{R_1}$$

where  $\psi_{R_1}$  is from Assumption 1 with  $\psi_R = (\psi'_{R_1}, \psi'_{R_2})'$ ,  $M_{\iota_{N_1}}$  is a projection matrix with  $M_{\iota_{N_1}} = I_{N_1} - \iota_{N_1}(\iota'_{N_1} \iota_{N_1})^{-1} \iota'_{N_1}$ ,  $\Psi_{\beta_1} \equiv \beta_1 V_{FF} \beta_2' (\beta_2 V_{FF} \beta_2' + \Omega_{22})^{-1} (\beta_2 (dV_{GG} + \psi_{uG}) + \psi_{v_2G}) \left[ (\beta_2 (dV_{GG} + \psi_{uG}) + \psi_{v_2G})' (\beta_2 V_{FF} \beta_2' + \Omega_{22})^{-1} (\beta_2 (dV_{GG} + \psi_{uG}) + \psi_{v_2G}) \right]^{-1}$ , see Theorem 1.

**Proof.** see Appendix B. ■

In line with Theorem 1, Corollary 1 also contains two cases. In the ideal first case, where accurate proxies of the underlying factors are used for constructing mimicking portfolios,  $\hat{\lambda}_G$  is a consistent risk premia estimator. On the other hand in the second case where weak factors are observed, as reflected by  $\delta = d/\sqrt{T}$ , because of the non-standard behavior of  $\hat{\beta}_1$ ,  $\hat{\lambda}_G$  also has a non-standard limiting distribution.

Due to the non-standard behavior of  $\hat{\lambda}_G$  in Corollary 1 under mimicking portfolios of weak factors, the conventional FM  $t$ -statistic for testing risk premia does not have an asymptotic standard normal distribution. Consequently, Corollary 1 implies that the conventional FM  $t$ -test on risk premia is under doubt, when it is used for mimicking portfolios of macroeconomic factors.

### 2.3 Covariance specification

Instead of the beta specification above, the alternative covariance specification suggests that with mimicking portfolios in (5), we could also use as an estimator for  $\beta_1$ :

$$\hat{\beta}_1 = \hat{V}_{R_1 R_2} \hat{V}_{R_2 R_2}^{-1} \hat{V}_{R_2 G}$$

Different from  $\hat{\beta}_1$ , the inverse part  $(\hat{V}_{GR_2} \hat{V}_{R_2 R_2}^{-1} \hat{V}_{R_2 G})^{-1}$  in  $\hat{\beta}_1$  is omitted in  $\hat{\beta}_1$ . This corresponds to the so-called covariance specification of risk premia, see e.g., Kan et al. (2013).

**Theorem 2** *When Assumptions 1-3 hold, the limiting behavior of  $\hat{\beta}_1$  can be described as follows.*

1. When  $\delta$  is fixed and the number of elements of  $G$  equals the number of elements of  $F$ , so  $\delta$  is a square invertible matrix:

$$\hat{\beta}_1 \xrightarrow{p} \beta_1 V_{FF} \beta_2' (\beta_2 V_{FF} \beta_2' + \Omega_{22})^{-1} \beta_2 \delta V_{GG}$$

and when the number of elements of  $\beta_2$  is large,  $\hat{\beta}_1 \xrightarrow{p} \beta_1 \delta V_{GG}$ .

2. When  $\delta = d/\sqrt{T}$ :

$$T^{\frac{1}{2}} \hat{\beta}_1 \xrightarrow{d} \beta_1 V_{FF} \beta_2' (\beta_2 V_{FF} \beta_2' + \Omega_{22})^{-1} (\beta_2 (dV_{GG} + \psi_{uG}) + \psi_{v_2G})$$

and when the number of elements of  $\beta_2$  is large,  $T^{\frac{1}{2}} \hat{\beta}_1 \xrightarrow{d} \beta_1 dV_{GG} + \beta_1 (\psi_{uG} + V_{FF} \beta_2' (\beta_2 V_{FF} \beta_2' + \Omega_{22})^{-1} \psi_{v_2G})$ . The specifications of  $\psi_{uG}$  and  $\psi_{v_2G}$  are stated in Theorem 1.

**Proof.** see Appendix C. ■

Unlike Theorem 1, the second part of Theorem 2 shows that the large sample behavior of the weak factor mimicking portfolio's beta estimator is now comparable to that of the pure weak factor's beta estimator, see Kleibergen (2009) and Kleibergen and Zhan (2015). Put differently, under the covariance specification, the beta estimator of the mimicking portfolios (denoted by  $\hat{\beta}_1$ ) does not suffer from the exaggeration problem of its counterpart (denoted by  $\hat{\beta}_1$ ) which exists under the beta specification, as shown by Theorem 1.

However, it is known that small betas jeopardize risk premia estimation in the FM methodology, see e.g. Kleibergen (2009). Consequently, a malfunction of the FM two-pass procedure also exists in the covariance specification, albeit for opposite reasons of the beta specification: if mimicking portfolios are constructed from weak factors, the beta estimator is large in magnitude in the beta specification, and small in the covariance specification, both could induce failure of risk premia estimation in the FM methodology. In order to resolve this issue, we develop a new inference method for risk premia of mimicking portfolios, which is presented later on.

## 2.4 Simulation study

To further illustrate the results in Theorem 1 and 2, we conduct a simple simulation experiment.

Asset returns are generated from the factor model (1), with  $T = 1000$  and  $k = 1$ . Specifically,  $F_t \sim NID(0, V_{FF})$ ,  $v_t \sim NID(0, \Omega)$ , where  $V_{FF}$  is calibrated from the market portfolio in Fama-French (1993),  $\Omega$  is calibrated from a regression of  $N_1$  industry portfolios and  $N_2$  size and book-to-market sorted portfolios on the market portfolio. We consider  $N_1 = N_2$  for convenience in the data generation process (d.g.p.) and set them equal to 1, 2, 3, 4 and 5, as reported in the first column of Table 1. The values of the parameters  $\lambda_0$ ,  $\lambda_F$  and  $\beta$  used in d.g.p. result from the Fama-MacBeth (1973) two-pass procedure using the described portfolios.<sup>5</sup>

In the simulation exercise, we consider two different types of observed factors which are encountered in empirical studies. The first case is the one of observed factors that are good/strong proxies for the underlying factors  $F_t$ . Here we use  $G_t = F_t$ . The second case is the one of observed factors that are poor/weak proxies of the underlying factors. We therefore use observed factors  $G_t$  that are independently generated as  $N(0, V_{FF})$  distributed random variables so the observed factor is completely useless for asset returns. In Panel A of Table 1, we present the outcome of the simulation study for the strong factor case, while the useless factor case is in Panel B.

Both the beta specification and the covariance specification for constructing the beta estimator are considered in our simulation study. In addition, for each specification, we consider three different manners of using the simulated factors. Under “Fac” in Table 1, the simulated factor is directly used for estimating  $\beta_1$  in the time series regression using the simulated asset returns  $R_{1t}$ . Under “MP” in Table 1, we construct mimicking portfolios of simulated factors as in (5); with the constructed mimicking portfolios, we proceed to compute their beta estimator  $\hat{\beta}_1$  (beta specification) or  $\hat{\beta}_1$  (covariance specification); furthermore, the asymptotic variances of such estimators are derived by first-order asymptotics (see Appendix D for details). Under

---

<sup>5</sup>The sample period used for calibration is Jan 1927 - Dec 2015. The data is downloaded from French’s online data library, “[http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data\\_library.html](http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html)”.

“MP as Fac” in Table 1, we treat constructed mimicking portfolios in the same manner as observed factors for beta estimation (hence the estimation error contained in the mimicking portfolios is ignored when computing the variance of their beta estimator), which is common in existing empirical studies.

We apply the rank test of Kleibergen and Paap (2006) for the estimator of  $\beta_1$  in the various scenarios described above. In particular, we test whether the estimand has reduced rank  $k - 1$  and document the rejection frequency of the null at the nominal 5% and 10% level using standard  $\chi^2$  critical values. Since  $\beta_1$  has full rank under strong factors, we expect that the rank test will strongly reject the null. This is in line with the rejection frequencies reported in Panel A of Table 1. In fact, we show that the rejection frequencies based on 2000 Monte Carlo replications are all equal to one, so the outcome of the rank test shows support for the strong factor, as expected.

On the contrary, under useless factors, we expect the rejection frequency of the rank test to be close to the nominal size, since useless factors are associated with zero beta’s so the null holds. This is in line with the rejection frequencies reported under “Fac” in Panel B of Table 1. Under “MP”, the rank test appears conservative, since the rejection frequencies quickly decrease from nominal sizes as  $N_1$  and  $N_2$  increase.<sup>6</sup>

What is astonishing in Panel B of Table 1 lies in the columns of “MP as Fac”, i.e., mimicking portfolios are naively treated as alternatives of factors and used for rank tests. Panel B shows that the rejection frequencies under “MP as Fac” are very large (albeit decrease as  $N_1$  and  $N_2$  increase). Consequently, mimicking portfolios of useless factors may signal strong factor pricing in rank tests, if they are treated in the same manner as factors. Theorem 1 and 2 show that such treatment is improper, since the beta estimator under mimicking portfolios does not have the same limiting distribution as under factors.

---

<sup>6</sup>There are two reasons that may help explain the conservative performance. First, the beta estimator of mimicking portfolios may have a non-standard distribution (as in Theorem 1), so first-order asymptotics do not approximate its variance well, which further jeopardizes the rank test. Second, as  $N_1$  and  $N_2$  increase, larger sample sizes are generally needed to estimate the covariance matrix of asset returns, a component that appears when beta’s are computed using mimicking portfolios.

Table 1: Rejection Frequencies of the Rank Test by Monte Carlo

Panel A: Strong Factor												
	Beta Specification						Covariance Specification					
	Fac		MP		MP as Fac		Fac		MP		MP as Fac	
	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
5	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Panel B: Useless Factor												
	Beta Specification						Covariance Specification					
	Fac		MP		MP as Fac		Fac		MP		MP as Fac	
	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
1	0.051	0.107	0.060	0.112	1.000	1.000	0.050	0.106	0.059	0.112	1.000	1.000
2	0.053	0.105	0.014	0.031	0.936	0.949	0.051	0.102	0.013	0.030	0.935	0.949
3	0.050	0.104	0.005	0.012	0.898	0.916	0.048	0.101	0.005	0.012	0.895	0.914
4	0.052	0.102	0.003	0.011	0.879	0.899	0.050	0.097	0.002	0.009	0.878	0.897
5	0.050	0.090	0.001	0.002	0.841	0.864	0.046	0.085	0.000	0.002	0.837	0.860

Note: This table reports the rejection frequencies of the Kleibergen and Paap (2006) rank test of the null  $H_0 : \text{rank}(\beta_1) = k - 1$  at the nominal 5% and 10% respectively, based on the average of 2000 replications. The d.g.p. is described in the main text, with  $T = 1000$  and  $k = 1$ . Column 1 lists five choices of  $N_1 = N_2$ : 1, 2, 3, 4, or 5. We consider two types of observed factors: Panel A for a strong factor and Panel B for a useless factor in the single factor model. For “Fac”: the observed factor is used for beta estimation. For “MP”: the mimicking portfolio of the observed factor is used for the beta estimation, and the variance of the beta estimator is derived by first-order asymptotics (see Appendix D). For “MP as Fac”: the mimicking portfolio of the observed factor is used for the beta estimation, as if the mimicking portfolio is an observed factor.

Overall, Table 1 suggests that the rank test can serve as a diagnostic tool for the quality of the factors, if implemented properly so correcting for the estimation error that results from using mimicking portfolios. When factors are strong (weak), the null is rejected (accepted with the probability close to the nominal size), if these factors are used for beta estimation. When mimicking portfolios are used for beta estimation, Table 1 suggests that the rank test outcome needs to be taken with caution. In particular, if mimicking portfolios are improperly treated as alternative to factors, then the rank test may spuriously favor useless factors.

### 2.4.1 Sensitivity to the strength of the factor structure

Instead of calibrating  $\Omega$  to the estimated  $\widehat{\Omega}$  as described above, we also consider two alternatives in our simulation experiments:  $\Omega = 0.04\widehat{\Omega}$  and  $\Omega = 25\widehat{\Omega}$ . The strength of the factor structure alters when the magnitude of  $\Omega$  changes with  $\Omega = 0.04\widehat{\Omega}$  having a very strong factor structure and  $\Omega = 25\widehat{\Omega}$  a weak factor structure. The other settings remain unchanged in the simulation. Our purpose is to re-examine the result reported in Table 1, as the strength of the factor structure changes.

Table 2 and Table 3 present the updated results, for  $\Omega = 0.04\widehat{\Omega}$  and  $\Omega = 25\widehat{\Omega}$ , respectively. It is found that the strength of the factor structure does not alter the two main findings conveyed in Table 1: (i) for rank testing, factors seem to perform better than mimicking portfolios; (ii) improperly treating mimicking portfolios as factors may yield to spurious rank test outcomes that favor poor factors.

## 3 Robust inference

The previous section shows that the beta estimator under mimicking portfolios is problematic when their underlying factors are weak. This jeopardizes risk premia estimation in the commonly used Fama-MacBeth (1973) methodology. To resolve this issue, we propose a robust inference procedure on risk premia of mimicking portfolios.

Table 2: Rejection Frequencies of the Rank Test by Monte Carlo ( $\Omega = 0.04\hat{\Omega}$ )

Panel A: Strong Factor												
	Beta Specification						Covariance Specification					
	Fac		MP		MP as Fac		Fac		MP		MP as Fac	
	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
5	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Panel B: Useless Factor												
	Beta Specification						Covariance Specification					
	Fac		MP		MP as Fac		Fac		MP		MP as Fac	
	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
1	0.063	0.110	0.066	0.116	1.000	1.000	0.062	0.110	0.064	0.116	1.000	1.000
2	0.053	0.103	0.016	0.035	0.956	0.960	0.051	0.102	0.016	0.034	0.955	0.960
3	0.055	0.111	0.005	0.014	0.915	0.931	0.051	0.103	0.005	0.015	0.914	0.931
4	0.052	0.100	0.006	0.009	0.911	0.921	0.047	0.095	0.005	0.009	0.908	0.920
5	0.045	0.091	0.001	0.004	0.884	0.902	0.042	0.086	0.001	0.004	0.881	0.900

Note: This table reports the rejection frequencies of the Kleibergen and Paap (2006) rank test of the null  $H_0 : \text{rank}(\beta_1) = k - 1$  at the nominal 5% and 10% respectively, based on the average of 2000 replications. The d.g.p. is described in the main text, with  $T = 1000$  and  $k = 1$ ,  $\Omega = 0.04\hat{\Omega}$ . Column 1 lists five choices of  $N_1 = N_2$ : 1, 2, 3, 4, or 5. We consider two types of observed factors: Panel A for a strong factor and Panel B for a useless factor in the single factor model. For “Fac”: the observed factor is used for beta estimation. For “MP”: the mimicking portfolio of the observed factor is used for the beta estimation, and the variance of the beta estimator is derived by first-order asymptotics (see Appendix D). For “MP as Fac”: the mimicking portfolio of the observed factor is used for the beta estimation, as if the mimicking portfolio is an observed factor.



Table 3: Rejection Frequencies of the Rank Test by Monte Carlo ( $\Omega = 25\widehat{\Omega}$ )

Panel A: Strong Factor												
	Beta Specification						Covariance Specification					
	Fac		MP		MP as Fac		Fac		MP		MP as Fac	
	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
1	1.000	1.000	0.465	0.583	0.429	0.556	1.000	1.000	0.319	0.486	0.426	0.554
2	1.000	1.000	0.663	0.760	0.636	0.742	1.000	1.000	0.489	0.648	0.629	0.742
3	1.000	1.000	0.688	0.782	0.666	0.761	1.000	1.000	0.495	0.650	0.656	0.758
4	1.000	1.000	0.878	0.930	0.856	0.914	1.000	1.000	0.642	0.793	0.848	0.911
5	1.000	1.000	0.922	0.953	0.902	0.936	1.000	1.000	0.635	0.805	0.896	0.935

Panel B: Useless Factor												
	Beta Specification						Covariance Specification					
	Fac		MP		MP as Fac		Fac		MP		MP as Fac	
	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
1	0.055	0.105	0.003	0.014	0.429	0.556	0.054	0.105	0.004	0.015	0.426	0.554
2	0.055	0.100	0.001	0.004	0.335	0.443	0.051	0.099	0.000	0.003	0.330	0.439
3	0.057	0.107	0.000	0.000	0.252	0.344	0.055	0.103	0.000	0.000	0.247	0.337
4	0.055	0.098	0.001	0.001	0.275	0.367	0.050	0.092	0.001	0.001	0.266	0.358
5	0.046	0.094	0.000	0.000	0.237	0.328	0.044	0.086	0.000	0.000	0.224	0.320

Note: This table reports the rejection frequencies of the Kleibergen and Paap (2006) rank test of the null  $H_0 : \text{rank}(\beta_1) = k - 1$  at the nominal 5% and 10% respectively, based on the average of 2000 replications. The d.g.p. is described in the main text, with  $T = 1000$  and  $k = 1$ ,  $\Omega = 25\widehat{\Omega}$ . Column 1 lists five choices of  $N_1 = N_2$ : 1, 2, 3, 4, or 5. We consider two types of observed factors: Panel A for a strong factor and Panel B for a useless factor in the single factor model. For “Fac”: the observed factor is used for beta estimation. For “MP”: the mimicking portfolio of the observed factor is used for the beta estimation, and the variance of the beta estimator is derived by first-order asymptotics (see Appendix D). For “MP as Fac”: the mimicking portfolio of the observed factor is used for the beta estimation, as if the mimicking portfolio is an observed factor.

### 3.1 Scaled risk premia under mimicking portfolios

Specifically, we suggest to test the risk premium  $\lambda_{G,cov}$  on mimicking portfolios using the moment condition:

$$E(R_{1t}) = \iota_{N_1} \lambda_0 + V_{R_1 R_2} V_{R_2 R_2}^{-1} V_{R_2 G} V_{GG}^{-1} \lambda_{G,cov} \quad (8)$$

where  $V_{R_1 R_2} V_{R_2 R_2}^{-1} V_{R_2 G}$  is the covariance of the test assets  $R_{1t}$  with the (infeasible) mimicking portfolios  $V_{GR_2} V_{R_2 R_2}^{-1} R_{2t}$ , so  $\lambda_{G,cov}$  is the scaled risk premium in the covariance specification (scaled by  $V_{GG}^{-1}$ ).<sup>7</sup> If the base assets in  $R_{2t}$  span the test assets in  $R_{1t}$ , then  $V_{R_1 R_2} V_{R_2 R_2}^{-1} V_{R_2 G}$  reduces to  $V_{R_1 G}$  and  $\lambda_{G,cov}$  equals  $\lambda_G$  in (4). For instance, in the special case that  $R_{1t} = R_{2t}$ , so test and base assets coincide,  $\lambda_{G,cov}$  reduces to  $\lambda_G$ . If so, inference on  $\lambda_{G,cov}$  reduces to inference on  $\lambda_G$ , which has been resolved by Kleibergen (2009). Since it is common that  $R_{2t}$  does not fully span  $R_{1t}$ ,  $\lambda_{G,cov}$  is consequently not necessarily equal to  $\lambda_G$ . We thus proceed to consider inference on  $\lambda_{G,cov}$  defined in (8).

As a starting point, we remove  $\lambda_0$  since our interest lies in  $\lambda_{G,cov}$ . This is done by removing the return on the  $N_1$ -th test asset and taking all other test asset returns in deviation from the return on the  $N_1$ -th asset. Equation (8) is then re-written as

$$\begin{aligned} E(\mathcal{R}_{1t}) &= V_{\mathcal{R}_1 R_2} V_{R_2 R_2}^{-1} V_{R_2 G} V_{GG}^{-1} \lambda_{G,cov} \\ &= \Gamma \beta_2 \delta \lambda_{G,cov} \end{aligned} \quad (9)$$

where  $\Gamma = V_{\mathcal{R}_1 R_2} V_{R_2 R_2}^{-1}$ ,  $\mathcal{R}_{1t} = R_{1t,1:(N_1-1)} - \iota_{N_1-1} R_{1t,N_1}$ , with  $R_{1t} = (R'_{1t,1:(N_1-1)}, R'_{1t,N_1})'$ , and  $V_{R_2 G} V_{GG}^{-1} = \beta_2 \delta$  as in (4).

---

<sup>7</sup>If  $V_{GG}$  is normalized to  $I_m$ , then  $\lambda_{G,cov}$  is the risk premium of the mimicking portfolios in the covariance specification. Furthermore,  $\lambda_{G,cov}$  could also be viewed as the scaled risk premium of mimicking portfolios in the beta specification: since  $E(R_{1t}) = \iota_{N_1} \lambda_0 + V_{R_1 R_2} V_{R_2 R_2}^{-1} V_{R_2 G} (V_{GR_2} V_{R_2 R_2}^{-1} V_{R_2 G})^{-1} (V_{GR_2} V_{R_2 R_2}^{-1} V_{R_2 G}) V_{GG}^{-1} \lambda_{G,cov}$ ,  $(V_{GR_2} V_{R_2 R_2}^{-1} V_{R_2 G}) V_{GG}^{-1} \lambda_{G,cov}$  is the risk premium in the beta specification. If the projection error during the construction of mimicking portfolios is negligible, i.e.,  $V_{GR_2} V_{R_2 R_2}^{-1} V_{R_2 G} \approx V_{GG}$ , then  $\lambda_{G,cov}$  approximately equals the risk premium of mimicking portfolios in the beta specification.

### 3.2 Mimicking portfolio Anderson-Rubin test

To conduct inference on  $\lambda_{G,cov}$ , we state the joint behavior of  $\bar{\mathcal{R}}_1$ ,  $\hat{B}_2$  and  $\hat{\Gamma}$  in Theorem 3, with  $\bar{\mathcal{R}}_1 = \frac{1}{T} \sum_{t=1}^T \mathcal{R}_{1t}$ ,  $\hat{B}_2 = \hat{V}_{R_2G} \hat{V}_{GG}^{-1}$  and  $\hat{\Gamma} = \hat{V}_{\mathcal{R}_1R_2} \hat{V}_{R_2R_2}^{-1}$ .

**Theorem 3** *When Assumptions 1-3 hold:*

$$\sqrt{T} \begin{pmatrix} \bar{\mathcal{R}}_1 - \Gamma \beta_2 \delta \lambda_{G,cov} \\ vec(\hat{B}_2 - \beta_2 \delta) \\ vec(\hat{\Gamma} - \Gamma) \end{pmatrix} \xrightarrow{d} \begin{pmatrix} \psi_1 \\ vec(\psi_2) \\ vec(\psi_3) \end{pmatrix} \sim N(0, \mathbb{W}) \quad (10)$$

and

$$\mathbb{W} = \begin{pmatrix} V_{\mathcal{R}_1\mathcal{R}_1} & 0 & 0 \\ 0 & V_{GG}^{-1} \otimes \Sigma_{22} & C' \\ 0 & C & V_{R_2R_2}^{-1} \otimes (V_{\mathcal{R}_1\mathcal{R}_1} - V_{\mathcal{R}_1R_2} V_{R_2R_2}^{-1} V_{R_2\mathcal{R}_1}) \end{pmatrix}$$

with

$$C = K_{(N_1-1)N_2} (V_{\mathcal{R}_1F} V_{FF}^{-1} \delta \otimes V_{R_2R_2}^{-1} \Sigma_{22}) + (V_{R_2R_2}^{-1} \beta_2 \delta \otimes \Sigma_{12}) - (V_{R_2R_2}^{-1} \otimes \Gamma) (I_{N_2^2} + K_{N_2N_2}) (\beta_2 \delta \otimes \Sigma_{22}),$$

$K$  is the commutation matrix and  $\Sigma_{12} = cov(\mathbf{e}_{1t}, \mathbf{e}_{2t})$  with  $\mathbf{e}_{1t} = \mathcal{R}_{1t} - V_{\mathcal{R}_1F} V_{FF}^{-1} \delta (\tilde{G}_t + \lambda_G)$ .

**Proof.** see Appendix E. ■

Theorem 3 implies that the limiting distribution of  $\sqrt{T}(\bar{\mathcal{R}}_1 - \hat{\Gamma} \hat{B}_2 \lambda_{G,cov})$  is normal.

**Corollary 2** *When Assumptions 1-3 hold and  $\Gamma$  has a full rank value:*

$$\sqrt{T}(\bar{\mathcal{R}}_1 - \hat{\Gamma} \hat{B}_2 \lambda_{G,cov}) \xrightarrow{d} \psi_1 - \Gamma \psi_2 \lambda_{G,cov} - \psi_3 \beta_2 \delta \lambda_{G,cov} \sim N(0, \mathbb{V}), \quad (11)$$

with

$$\begin{aligned} \mathbb{V} = & V_{\mathcal{R}_1\mathcal{R}_1} + (\lambda'_{G,cov} V_{GG}^{-1} \lambda_{G,cov} \otimes \Gamma \Sigma_{22} \Gamma') + \\ & ((\beta_2 \delta \lambda_{G,cov})' V_{R_2R_2}^{-1} (\beta_2 \delta \lambda_{G,cov}) \otimes (V_{\mathcal{R}_1\mathcal{R}_1} - V_{\mathcal{R}_1R_2} V_{R_2R_2}^{-1} V_{R_2\mathcal{R}_1})) - \\ & \{[(\beta_2 \delta \lambda_{G,cov})' \otimes I_{N_1-1}] C(\lambda_{G,cov} \otimes \Gamma')\} - \{[(\beta_2 \delta \lambda_{G,cov})' \otimes I_{N_1-1}] C(\lambda_{G,cov} \otimes \Gamma')\}' \end{aligned}$$

**Proof.** see Appendix F. ■

Since the unobserved factor  $F_t$  affects both  $R_1$  and  $R_2$ ,  $\Gamma$  has a full rank value. We do not make a full rank assumption on  $\delta$  so we allow for weak correlation between observed and unobserved factors. Using Corollary 2, it is straightforward to obtain an asymptotic result to test  $H_0 : \lambda_{G,cov} = \lambda_{G,cov,0}$ .

**Corollary 3 (MPAR)** Under  $H_0 : \lambda_{G,cov} = \lambda_{G,cov,0}$ ,

$$MPAR(\lambda_{G,cov,0}) = T(\bar{\mathcal{R}}_1 - \hat{\Gamma}\hat{B}_2\lambda_{G,cov,0})'\hat{\mathbb{V}}^{-1}(\bar{\mathcal{R}}_1 - \hat{\Gamma}\hat{B}_2\lambda_{G,cov,0}) \xrightarrow{d} \chi_{N_1-1}^2 \quad (12)$$

with

$$\begin{aligned} \hat{\mathbb{V}} &= \hat{V}_{\mathcal{R}_1\mathcal{R}_1} + (\lambda'_{G,cov,0}\hat{V}_{GG}^{-1}\lambda_{G,cov,0} \otimes \hat{\Gamma}\hat{\Sigma}_{22}\hat{\Gamma}') + \\ &\quad ((\hat{B}_2\lambda_{G,cov,0})'\hat{V}_{R_2R_2}^{-1}(\hat{B}_2\lambda_{G,cov,0}) \otimes (\hat{V}_{\mathcal{R}_1\mathcal{R}_1} - \hat{V}_{\mathcal{R}_1R_2}\hat{V}_{R_2R_2}^{-1}\hat{V}_{R_2\mathcal{R}_1})) - \\ &\quad [(\hat{B}_2\lambda_{G,cov,0})' \otimes I_{N_1-1}]\hat{C}(\lambda_{G,cov,0} \otimes \hat{\Gamma}') - (\lambda'_{G,cov,0} \otimes \hat{\Gamma})\hat{C}'[(\hat{B}_2\lambda_{G,cov,0}) \otimes I_{N_1-1}] \\ \hat{C} &= K_{(N_1-1)N_2}(\hat{B}_1 \otimes \hat{V}_{R_2R_2}^{-1}\hat{\Sigma}_{22}) + (\hat{V}_{R_2R_2}^{-1}\hat{B}_2 \otimes \hat{\Sigma}_{12}) - (\hat{V}_{R_2R_2}^{-1} \otimes \hat{\Gamma})(I_{N_2^2} + K_{N_2N_2})(\hat{B}_2 \otimes \hat{\Sigma}_{22}) \end{aligned}$$

and  $\hat{\Sigma}_{12} = \frac{1}{T} \sum_{t=1}^T \hat{e}_{1t}\hat{e}'_{2t}$ ,  $\hat{\Sigma}_{22} = \frac{1}{T} \sum_{t=1}^T \hat{e}_{2t}\hat{e}'_{2t}$ ,  $\hat{e}_{2t} = R_{2t} - \bar{R}_2 - \hat{B}_2(G_t - \hat{\mu}_G)$ ,  $\hat{e}_{1t} = \mathcal{R}_{1t} - \bar{\mathcal{R}}_1 - \hat{B}_1(G_t - \hat{\mu}_G)$ ,  $\hat{B}_1 = \hat{V}_{\mathcal{R}_1G}\hat{V}_{GG}^{-1}$ .

**Proof.** Results from Corollary 2, with  $\hat{\mathbb{V}}$  a consistent estimator for  $\mathbb{V}$ . ■

We refer to this statistic as the Mimicking Portfolio Anderson-Rubin (MPAR) statistic, since it is in line with the Anderson and Rubin (1949) statistic for robust inference and it is proposed using mimicking portfolios. Similar tests in the factor rather than mimicking portfolio setting include the Factor Anderson Rubin (FAR) test proposed in Kleibergen (2009) and the Hotelling-type statistic proposed in Beaulieu et al. (2013).<sup>8</sup>

Because of Corollary 3, a test of  $H_0 : \lambda_{G,cov} = \lambda_{G,cov,0}$  that rejects  $H_0$ , if  $MPAR(\lambda_{G,cov,0})$  exceeds the  $1-\alpha$  quantile of the  $\chi^2$  distribution with  $N_1-1$  degrees of freedom has an asymptotic

<sup>8</sup>Kleibergen (2009) also proposes several other tests (e.g., a score test). Our paper focuses on extending the FAR statistic to the MPAR statistic, because extending other tests appears, because of the structure of the covariance matrix in Theorem 3, difficult.

size that equals  $\alpha$  irrespective of the value of  $\delta$ . Values of  $\lambda_{G,cov,0}$  that are not rejected by this test thus constitute a  $100(1 - \alpha)\%$  confidence set of  $\lambda_{G,cov}$ .

### 3.3 Size of the MPAR statistic in simulations

To examine the size of the proposed MPAR test, we conduct a simple simulation experiment.

The d.g.p. is similar to the one used in Section 2. Specifically, asset returns are generated from Equation (1), with  $T = 1000$  and  $k = 1$ . In addition,  $F_t$  and  $v_t$  are generated as independent  $NID(0, V_{FF})$  and  $NID(0, \Omega)$  random variables, with  $V_{FF}$  calibrated from the market portfolio in Fama-French (1993) and  $\Omega$  calibrated from the regression of  $N_1$  industry and  $N_2$  size and book-to-market sorted portfolios on the market portfolio. We consider  $N_1 = N_2$  for convenience and set them equal to 5, 10, 15, 20 and 25, as reported in the first column of Table 4. The values of  $\lambda_0$ ,  $\lambda_F$  and  $\beta$  used in d.g.p. result from the Fama-MacBeth (1973) two-pass procedure using the portfolios described above.

The observed  $G_t$  is simulated as follows:  $G_t = \delta \cdot F_t + \sqrt{1 - \delta^2} \cdot NID(0, V_{FF})$ , so  $\delta$  coincides with the specification in (3) and reflects the quality of  $G_t$  for approximating  $F_t$ . When  $\delta$  is close to zero,  $G_t$  is a weak factor which becomes stronger for an increasing value of  $\delta$ . We consider a sequence of values for  $\delta$ :  $\delta \in \{0.01, 0.25, 0.50, 0.75, 0.99\}$ , which covers a wide range of settings of  $G_t$  as the observed proxy for  $F_t$ .

With the simulated asset returns and  $G_t$ , we conduct the MPAR test on the risk premia of the mimicking portfolios as described in Corollary 3. The resulting sizes of the test are reported in Table 4.

Table 4 shows that the MPAR test does not severely over reject in any of the settings. Its rejection frequency approximately equals the size of the test in many instances. Table 4 shows that the MPAR test is conservative when the factors are weak. It indicates that the covariance matrix estimator  $\widehat{\mathbf{V}}$  is on average too large in this simulation setting. Using the properties of the partitioned inference and that  $\widehat{\Gamma}$  does not depend on the strength of the factors, the too large value of  $\widehat{\mathbf{V}}$  can be attributed to the realized values of  $\widehat{B}_1$  and  $\widehat{B}_2$  being away from their

Table 4: Actual Sizes of the MPAR Test by Monte Carlo,  $T = 1000$

	$\delta = 0.01$		$\delta = 0.25$		$\delta = 0.50$		$\delta = 0.75$		$\delta = 0.99$	
	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
$N_1, N_2 = 5$	0.050	0.105	0.059	0.105	0.058	0.103	0.057	0.104	0.057	0.104
$N_1, N_2 = 10$	0.020	0.051	0.052	0.108	0.051	0.108	0.050	0.108	0.051	0.108
$N_1, N_2 = 15$	0.018	0.047	0.064	0.115	0.063	0.116	0.063	0.116	0.063	0.116
$N_1, N_2 = 20$	0.002	0.008	0.063	0.127	0.064	0.128	0.063	0.126	0.062	0.126
$N_1, N_2 = 25$	0.001	0.002	0.064	0.124	0.064	0.127	0.063	0.127	0.063	0.126

Note: The reported sizes are rejection frequencies of the MPAR test for  $H_0 : \lambda_{G,cov} = \lambda_{G,cov,0}$  at the nominal 5% and 10% respectively, based on the average of 2000 replications. The d.g.p. is described in the main text, with  $T = 1000$ .

expected values (zero). This explains also why the under rejection increases when  $N_1$  and  $N_2$  get larger. When we increase the sample size, the estimates of  $\hat{B}_1$  and  $\hat{B}_2$  become more precise and the under rejection of the MPAR test disappears. This is shown in Table 5 where we use a sample size of 1000,000. Table 5 shows that the MPAR test is now size correct in all instances.

Table 5: Actual Sizes of the MPAR Test by Monte Carlo,  $T = 1000,000$

	$\delta = 0.01$		$\delta = 0.25$		$\delta = 0.50$		$\delta = 0.75$		$\delta = 0.99$	
	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
$N_1, N_2 = 5$	0.053	0.102	0.059	0.098	0.058	0.097	0.058	0.097	0.058	0.097
$N_1, N_2 = 10$	0.060	0.112	0.057	0.111	0.058	0.110	0.058	0.111	0.059	0.110
$N_1, N_2 = 15$	0.055	0.115	0.056	0.110	0.057	0.111	0.058	0.110	0.058	0.109
$N_1, N_2 = 20$	0.057	0.104	0.054	0.098	0.055	0.099	0.054	0.098	0.054	0.098
$N_1, N_2 = 25$	0.056	0.104	0.053	0.107	0.053	0.107	0.054	0.107	0.054	0.107

Note: The reported sizes are rejection frequencies of the MPAR test for  $H_0 : \lambda_{G,cov} = \lambda_{G,cov,0}$  at the nominal 5% and 10% respectively, based on the average of 2000 replications. The d.g.p. is described in the main text, with  $T = 1000,000$ .

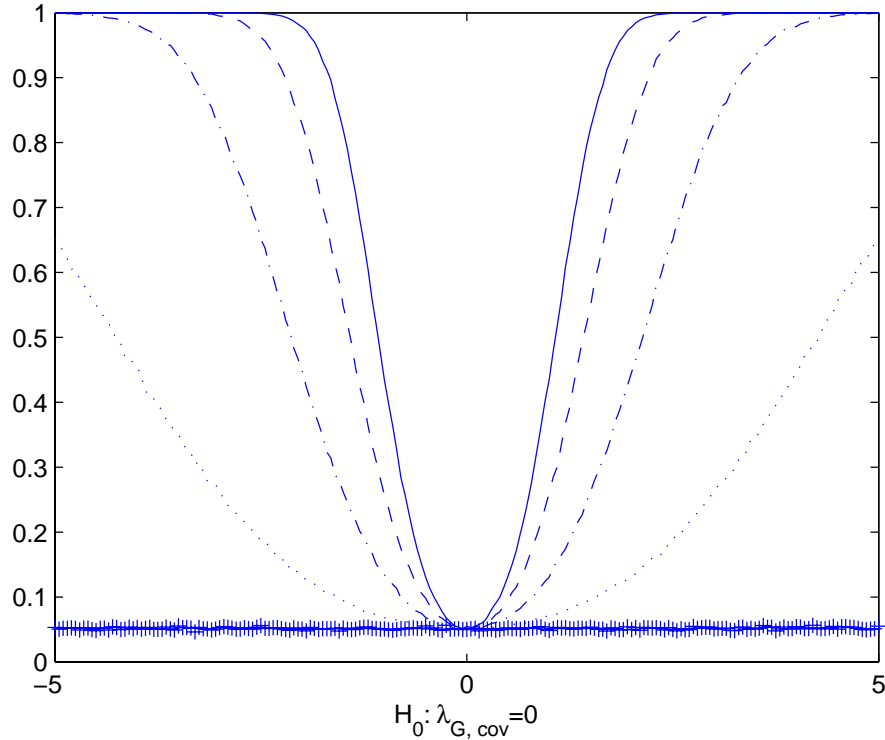
Tables 4 and 5 show that the behavior of the MPAR test accords with Corollary 3 since the under rejection reported in Table 4 results from the estimation of the covariance matrix  $\mathbb{V}$ . The size of a test is defined as the maximal rejection frequency under the null hypothesis so under rejection does not indicate size distortion.

### 3.4 Power of the MPAR test

To illustrate the power of the MPAR test, we consider a sequence of  $\lambda_{G,cov}$  (between  $-5$  and  $5$ ) in the d.g.p. described above with  $T = 1000$  and  $N_1 = N_2 = 5$ .<sup>9</sup> We then test  $H_0 : \lambda_{G,cov} = 0$  at the 5% level using the generated data. The resulting rejection frequencies of the MPAR test are in Figure 1.

We use a range of values of  $\delta$  to reflect the different qualities of the observed factors. Figure 1 has five power plots of the MPAR test, corresponding to  $\delta = 0.01$  (plus),  $\delta = 0.25$  (dotted),  $\delta = 0.50$  (dash-dot),  $\delta = 0.75$  (dashed) and  $\delta = 0.99$  (solid line), respectively.

Figure 1: Power Plots of the MPAR Test



Note: This figure presents the power plot of the MPAR test for  $H_0 : \lambda_{G,cov} = 0$  at the 5% level, with  $\delta = 0.01$  (plus),  $\delta = 0.25$  (dotted),  $\delta = 0.50$  (dash-dot),  $\delta = 0.75$  (dashed),  $\delta = 0.99$  (solid line). The d.g.p. is described in the main text, with  $T = 1000$  and  $N_1 = N_2 = 5$ .

<sup>9</sup>Note that  $\beta_1 \lambda_F = V_{R_1 R_2} V_{R_2 R_2}^{-1} V_{R_2 G} V_{G G}^{-1} \lambda_{G,cov}$ , so  $\lambda_F = (\beta_1' \beta_1)^{-1} \beta_1' V_{R_1 R_2} V_{R_2 R_2}^{-1} V_{R_2 G} V_{G G}^{-1} \lambda_{G,cov}$  is used for d.g.p., and we use 20,000 Monte Carlo replications for the power plots.

Figure 1 shows that the MPAR test has good power, as  $\delta$  gets large. This is as expected, since a larger value of  $\delta$  implies more informative factors and thus mimicking portfolios. Consequently, when confidence sets of risk premia are to be constructed by inverting the MPAR test, wider confidence sets indicate that the corresponding factors are less informative, while narrower sets signal that the factors are stronger for asset pricing. Note that when  $\delta = 0.01$  so the factor is close to being useless, Figure 1 shows that the corresponding power plot is close to the nominal 5%.

We further compare the MPAR test with the FAR test in Kleibergen (2009). As stated above, in the special case that test assets and base assets coincide so  $R_{1t} = R_{2t}$ , the mimicking portfolio risk premia  $\lambda_{G,cov}$  coincides with the factor risk premia  $\lambda_G$ . In this scenario, both MPAR and FAR tests are applicable, and we plot their power curves in Figure 2. As shown by Figure 2, the power of the MPAR test is almost identical to that of the FAR test.

### 3.5 Application

We illustrate practical usage of the MPAR test by employing it to the leverage factor model proposed by Adrian et al. (2014). Specifically, Adrian et al. (2014) consider a leverage factor “*LevFac*” in a single factor model, so  $k = 1$ . Their empirical study uses data for the leverage factor from 1968Q1-2009Q4, so  $T = 168$ . For illustrative purposes, we adopt the same data set.<sup>10</sup> For base assets, we follow Adrian et al. (2014) who consider seven assets, so  $N_2 = 7$ : the excess returns of the six Fama-French portfolios on size and book-to-market (“*BL*”, “*BM*”, “*BH*”, “*SL*”, “*SM*”, “*SH*”), plus the momentum factor “*Mom*”. These seven assets are widely acknowledged for their ability to span the asset space.<sup>11</sup> To construct mimicking portfolios, we project the leverage factor on these seven base assets.

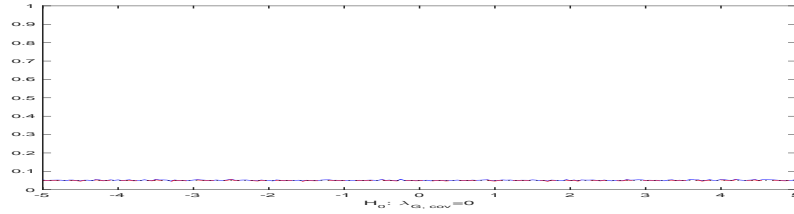
We obtained estimates for the coefficients (denoted by  $\widehat{V}_{GR_2} \widehat{V}_{R_2 R_2}^{-1}$  in Section 2) when projecting the leverage factor “*LevFac*” on the base assets (“*BL*”, “*BM*”, “*BH*”, “*SL*”, “*SM*”,

<sup>10</sup>We thank the authors for making the data publicly available.

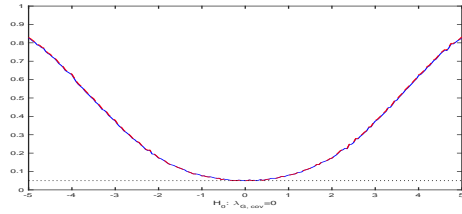
<sup>11</sup>Monthly data for these assets is available at French’s online data library, which is then compounded to construct the quarterly data. Excess returns result from the difference of raw returns and risk-free returns.



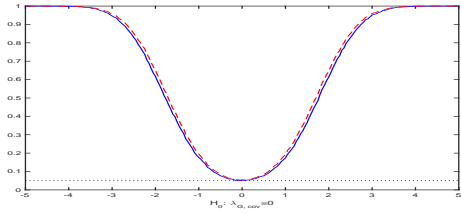
Figure 2: Power Plots of the MPAR Test and the FAR Test



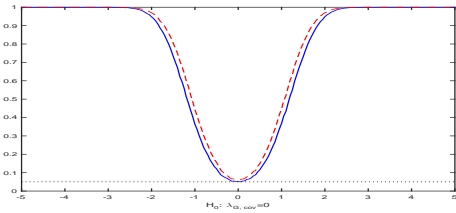
(a)  $\delta = 0.01$



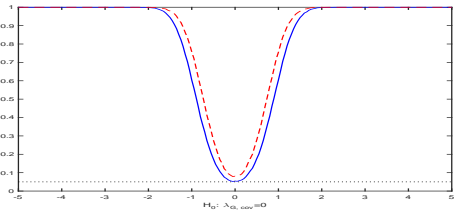
(b)  $\delta = 0.25$



(c)  $\delta = 0.50$



(d)  $\delta = 0.75$



(e)  $\delta = 0.99$

Note: This figure presents the power plots of the MPAR test (solid) and the FAR test (dashed) for  $H_0 : \lambda_{G,cov} = 0$  at the 5% level, with  $\delta = 0.01$  in (a),  $\delta = 0.25$  in (b),  $\delta = 0.50$  in (c),  $\delta = 0.75$  in (d) and  $\delta = 0.99$  in (e). The d.g.p. is described in the main text, with  $T = 1000$  and  $N_1 = N_2 = 5$ .

“*SH*”, “*Mom*”) equal to:  $(-0.22, -0.10, 0.56, -0.57, 1.24, -0.43, 0.43)$ . These coefficients (after normalizing their sum to one) are almost identical to those in Adrian et al. (2014) for the reported weights of the mimicking portfolio.

Adrian et al. (2014) show the mimicking portfolio of the leverage factor performs well in various asset pricing tests, e.g., it is associated with large R-squareds and low intercepts in cross-sectional regressions. Adrian et al. (2014) do not study the risk premium associated with the mimicking portfolio in cross-sectional regressions, while we focus on it. Specifically, we are interested in the risk premium of the mimicking portfolio “*LevMP*” that results from projecting “*LevFac*” on the seven base assets. We use as test assets the commonly used twenty-five Fama-French portfolios on size and book-to-market (25 FF). Since both test assets and base assets are on size and book-to-market, they are likely driven by the same underlying factors, as in the model setup in (1). In addition, since  $T = 168$  is relatively small, while our simulation study suggests that the size of the MPAR test is more reliable under small  $N_1$  and  $N_2$  for small  $T$ , we divide the twenty-five FF portfolios into five equal sized groups (denoted by I - V in the first column of Table 6), and use each group as test assets with  $N_1 = 5$ .<sup>12</sup>

To gauge the statistical quality of the leverage factor, we conduct the Kleibergen and Paap (2006) rank test, which is commonly used as a diagnostic tool in the asset pricing literature. The null of the rank test is that the leverage factor beta has reduced rank, and the resulting  $p$  values are found to be 0.09, 0.01, 0.26, 0.27, 0.08 for I - V, respectively.<sup>13</sup> Since most  $p$  values exceed 5%, the leverage factor appears to be weakly correlated with asset returns and its beta is small. Consequently, the conventional  $t$ -test on the risk premium of the leverage factor is under doubt.

Table 6 starts out with the FM two-pass methodology. When the leverage factor *LevFac* is tested using the FM procedure, Table 6 reports that the estimated risk premium is positive and

---

<sup>12</sup>Groups I - V are built in the order of 25 FF, i.e., the first 5 portfolios of 25 FF make Group I, ..., the last 5 portfolios of 25 FF make Group V.

<sup>13</sup>In accordance with the FAR and MPAR tests, we similarly remove the last asset and take all other asset returns in deviation from the return on the last asset when implementing the rank test. This is one way of conducting a rank test on  $(\iota_N \vdots \beta)$  which is the regressor matrix in the second pass of the FM procedure.

the conventional FM  $t$ -statistics are significant in I and II at the 95% significance level. Similarly, when the mimicking portfolio  $LevMP$  is tested in the FM procedure, Table 6 shows positive risk premium, associated with slightly larger FM  $t$ -statistics compared to the factor counterparts. These results thus appear to support the leverage factor for asset pricing. If the Shanken (1992) correction is adopted, however, none of the  $t$ -statistics remain significant. Furthermore, since the construction error in mimicking portfolios also contributes to the variance of the risk premia estimator,  $t$ -statistics under “ $LevMP$ ” are expected to further decrease, if this error is taken into account. See Jiang et al. (2015).

Table 6: Risk Premium of the Leverage Factor ( $LevFac$ ) and its Mimicking Portfolio ( $LevMP$ )

	$LevFac$				$LevMP$			
	$t$ -test			FAR test	$t$ -test			MPAR test
	Coef.	FM $t$	Shanken $t$	95% C.I.	Coef.	FM $t$	Shanken $t$	95% C.I.
I.	22.35	2.11	1.30	$(-\infty, -126.02] \cup [13.96, \infty)$	4.11	3.41	0.81	$(-\infty, -104.78] \cup [16.82, \infty)$
II.	11.78	2.10	0.80	$[-1.23, 98.81]$	2.73	2.56	0.54	$(-\infty, -25.57] \cup [-2.43, \infty)$
III.	16.59	1.64	0.98	$(-\infty, -43.11] \cup [1.67, \infty)$	2.32	2.32	0.46	$(-\infty, -191.49] \cup [3.08, \infty)$
IV.	8.32	1.12	0.54	$(-\infty, \infty)$	1.01	1.32	0.20	$(-\infty, \infty)$
V.	4.85	0.88	0.33	$(-\infty, \infty)$	0.80	0.93	0.16	$(-\infty, \infty)$

Note:  $LevFac$  stands for the leverage factor suggested in Adrian et al. (2014), while  $LevMP$  stands for the constructed mimicking portfolio of the leverage factor. Test assets are from twenty-five Fama-French portfolios on size and book-to-market (25 FF) in the sample period of 1968Q1-2009Q4, and we divide them into 5 groups of test assets, I-V, so each group contains 5 portfolios. Three tests on risk premia are employed, namely, the conventional FM- $t$  test for the Fama and MacBeth (1973) methodology, the Factor Anderson-Rubin (FAR) test of Kleibergen (2009), and the proposed Mimicking Portfolio Anderson-Rubin (MPAR) test in this paper. The Shanken (1992) correction is also considered for the  $t$ -test.

It is now known that the FM methodology and its associated  $t$ -test are doubtful when the factor is associated with a small beta. As an alternative to the  $t$ -test for the factor risk premia, the Factor Anderson Rubin (FAR) test is proposed by Kleibergen (2009). Unlike the FM  $t$ -test, the FAR test is size-correct since its limiting distribution does not depend on the quality of the observed factors. Table 6 presents the 95% confidence sets of the risk premium that results from inverting the FAR test. These sets are found to be substantially different from those obtained by inverting the FM  $t$ -test. Such differences put the quality of the leverage factor

under doubt, see also Kleibergen (2009) for a further comparison of the FM- $t$  test and the FAR test. Furthermore, in Appendix G, we present the  $p$ -value plots of the FAR test which show how we obtained the 95% FAR confidence sets reported in Table 6.

The last column of Table 6 shows the 95% confidence sets that result from the MPAR test. In line with the conventional FM  $t$ -test, it shows rejection of the null hypothesis of a zero risk premium for  $LevMP$  for I and III; however, unlike the FM  $t$ -test, it does not reject a zero risk premium in II. In IV and V, identical to the FAR test but unlike the FM  $t$ -test, we find that no information about the risk premium is contained in the leverage factor or its mimicking portfolio. In Appendix G, we also present the  $p$ -value plots of the MPAR test, which helps to explain the 95% MPAR confidence sets reported in Table 6.

Overall, Table 6 indicates that inference on the risk premium of the leverage factor may substantially change, when robust tests (FAR or MPAR) are employed.<sup>14</sup> This is consistent with the fact that the leverage factor is only weakly correlated with asset returns and thus likely to be a weak proxy for the underlying factor(s).

## 4 Conclusions

We document the threats involved in using mimicking portfolios of macroeconomic factors in the Fama and MacBeth (1973) two-pass procedure. When these factors have small betas, we show that their mimicking portfolios have betas that are spurious. These spurious betas induce non-standard behavior of the risk premia estimator so conventional  $t$ -tests on risk premia become unreliable.

A rank test on beta is used in the literature to serve as a diagnostic tool for the quality of the factors. We, however, find that the outcome of the rank test needs to be taken with caution when using mimicking portfolios. This results from the estimation error in the mimicking portfolio

---

<sup>14</sup>Note that the factor risk premia defined in (2)(4) and the mimicking portfolio risk premia defined in (8) are not identical in this application. Thus we do not expect the FAR confidence set for factor risk premia to coincide with the MPAR confidence set for the mimicking portfolio risk premia.

which we have to account for in the rank test. It implies a more challenging expression for the covariance matrix estimator employed in the rank test. When we do not account for this estimation error, the rank test performs poorly. When we account for it, the rank test still has some issues when the covariance matrix becomes of large dimension but generally works well.

Instead of gauging the quality of factors or mimicking portfolios, inference methods are available for analyzing risk premia that are reliable irrespective of the quality of the factors. These methods are robust in the sense that their limiting distributions do not depend on the quality of factors as reflected by the magnitude of the betas. To the best of our knowledge, the method we propose here is the first one which deals with mimicking portfolios. Robust methods do exist for tests on risk premia in the standard factor pricing setting, see Kleibergen (2009) and Beaulieu et al. (2013). This clearly indicates the need for our developed methods for which the empirical relevance is further emphasized by our application to the risk premium on the leverage factor from Adrian et al. (2014).

## References

- [1] Tobias Adrian, Erkki Etula, and Tyler Muir. Financial intermediaries and the cross-section of asset returns. *Journal of Finance*, 69(6):2557–2596, 2014.
- [2] TW Anderson and H. Rubin. Estimation of the parameters of a single equation in a complete system of stochastic equations. *Annals of Mathematical Statistics*, pages 46–63, 1949.
- [3] Andrew Ang, Robert J Hodrick, Yuhang Xing, and Xiaoyan Zhang. The cross-section of volatility and expected returns. *Journal of Finance*, 61(1):259–299, 2006.
- [4] Doron Avramov and Tarun Chordia. Asset pricing models and financial market anomalies. *Review of Financial Studies*, 19(3):1001–1040, 2006.
- [5] J. Bai and S. Ng. Evaluating latent and observed factors in macroeconomics and finance. *Journal of Econometrics*, 131(1):507–537, 2006.
- [6] Pierluigi Balduzzi and Cesare Robotti. Mimicking portfolios, economic risk premia, and tests of multi-beta models. *Journal of Business and Economic Statistics*, 26(3):354–368, 2008.
- [7] Marie-Claude Beaulieu, Jean-Marie Dufour, and Lynda Khalaf. Identification-robust estimation and testing of the zero-beta capm. *Review of Economic Studies*, 80(3):892–924, 2013.
- [8] Douglas T Breeden. An intertemporal asset pricing model with stochastic consumption and investment opportunities. *Journal of Financial Economics*, 7(3):265–296, 1979.
- [9] Douglas T Breeden, Michael R Gibbons, and Robert H Litzenberger. Empirical tests of the consumption-oriented CAPM. *Journal of Finance*, 44(2):231–262, 1989.
- [10] Craig Burnside. Identification and inference in linear stochastic discount factor models with excess returns. *Journal of Financial Econometrics*, 14(2):295–330, 2016.

- [11] G. Chamberlain and M. Rothschild. Arbitrage, Factor Structure, and Mean-Variance Analysis on Large Asset Markets. *Econometrica*, 51:1281–1302, 1983.
- [12] G. Connor and R.A. Korajczyk. Risk and return in an equilibrium APT: Application of a new test methodology. *Journal of Financial Economics*, 21(2):255–289, 1988.
- [13] G. Connor and R.A. Korajczyk. An intertemporal equilibrium beta pricing model. *Review of Financial Studies*, 2(3):373–392, 1989.
- [14] E.F. Fama and K. French. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1):3–56, 1993.
- [15] E.F. Fama and J.D. MacBeth. Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy*, 81(3):607–636, 1973.
- [16] Gur Huberman, Shmuel Kandel, and Robert F Stambaugh. Mimicking portfolios and exact arbitrage pricing. *Journal of Finance*, 42(1):1–9, 1987.
- [17] R. Jagannathan and Z. Wang. The conditional CAPM and the cross-section of expected returns. *Journal of Finance*, pages 3–53, 1996.
- [18] L. Jiang, Raymond Kan, and Z. Zhan. Asset pricing tests with mimicking portfolios. 2015.
- [19] R. Kan and C. Zhang. Two-pass tests of asset pricing models with useless factors. *Journal of Finance*, pages 203–235, 1999.
- [20] Raymond Kan, Cesare Robotti, and Jay Shanken. Pricing model performance and the two-pass cross-sectional regression methodology. *Journal of Finance*, 68(6):2617–2649, 2013.
- [21] F. Kleibergen. Tests of risk premia in linear factor models. *Journal of Econometrics*, 149(2):149–173, 2009.

- [22] F. Kleibergen and R. Paap. Generalized reduced rank tests using the singular value decomposition. *Journal of Econometrics*, 133(1):97–126, 2006.
- [23] F. Kleibergen and Z. Zhan. Unexplained factors and their effects on second pass R-squared's. *Journal of Econometrics*, 189:101–116, 2015.
- [24] Owen A Lamont. Economic tracking portfolios. *Journal of Econometrics*, 105(1):161–184, 2001.
- [25] M. Lettau and S. Ludvigson. Resurrecting the (C) CAPM: A cross-sectional test when risk premia are time-varying. *Journal of Political Economy*, 109(6):1238–1287, 2001.
- [26] Q. Li, M. Vassalou, and Y. Xing. Sector investment growth rates and the cross section of equity returns. *Journal of Business*, 79(3):1637–1665, 2006.
- [27] R.C. Merton. An intertemporal capital asset pricing model. *Econometrica*, pages 867–887, 1973.
- [28] R. Roll and S.A. Ross. An empirical investigation of the arbitrage pricing theory. *Journal of Finance*, pages 1073–1103, 1980.
- [29] S. Ross. The Arbitrage Theory of Capital Asset Pricing. *Journal of Economic Theory*, 13:341–360, 1976.
- [30] J. Shanken. On the estimation of beta-pricing models. *Review of Financial Studies*, pages 1–33, 1992.
- [31] D. Staiger and J.H. Stock. Instrumental variables regression with weak instruments. *Econometrica*, pages 557–586, 1997.
- [32] Maria Vassalou. News related to future GDP growth as a risk factor in equity returns. *Journal of Financial Economics*, 68(1):47–73, 2003.



# Appendix

## A. Proof of Theorem 1

**Proof.** Let's start with  $\widehat{V}_{R_1R_2}$ ,  $\widehat{V}_{R_2R_2}$  and  $\widehat{V}_{R_2G}$ . Assumptions 1 and 2 imply that:

$$\begin{aligned}\widehat{V}_{R_1R_2} &\xrightarrow{p} \beta_1 V_{FF} \beta_2' \\ \widehat{V}_{R_2R_2} &\xrightarrow{p} \beta_2 V_{FF} \beta_2' + \Omega_{22}\end{aligned}$$

where we used that  $\widehat{V}_{RR} \xrightarrow{p} \beta V_{FF} \beta' + \Omega$  and  $\Omega_{12} = 0$ .

The convergence of  $\widehat{V}_{R_2G}$  is a bit more tricky. Note that:

$$R_t = \iota_N \lambda_0 + \beta[(\delta \bar{G}_t + u_t) + \lambda_F] + v_t = \iota_N \lambda_0 + \beta \lambda_F + \beta \delta \bar{G}_t + \beta u_t + v_t$$

When  $\delta$  is fixed and the number of elements of  $G$  equals the number of elements of  $F$ , so  $\delta$  is a square invertible matrix:

$$\widehat{V}_{R_2G} \xrightarrow{p} \beta_2 \delta V_{GG}$$

so

$$\begin{aligned}\hat{\beta}_1 &\xrightarrow{p} \beta_1 V_{FF} \beta_2' (\beta_2 V_{FF} \beta_2' + \Omega_{22})^{-1} \beta_2 \delta V_{GG} [V_{GG} \delta' \beta_2' (\beta_2 V_{FF} \beta_2' + \Omega_{22})^{-1} \beta_2 \delta V_{GG}]^{-1} \\ &= \beta_1 V_{FF} \delta^{-1'} V_{GG}^{-1},\end{aligned}$$

and when  $G_t = F_t$ ,  $\delta = I_k$ ,  $\hat{\beta}_1 \xrightarrow{p} \beta_1$ .

When  $\delta = d/\sqrt{T}$ :

$$\sqrt{T} \widehat{V}_{R_2G} \xrightarrow{d} \beta_2 (dV_{GG} + \psi_{uG}) + \psi_{v_2G}$$

where  $\frac{1}{\sqrt{T}} \sum_{t=1}^T (\bar{G}_t \otimes u_t) \xrightarrow{d} \text{vec}(\psi_{uG})$ ,  $\frac{1}{\sqrt{T}} \sum_{t=1}^T (\bar{G}_t \otimes v_{2t}) \xrightarrow{d} \text{vec}(\psi_{v_2G})$ , so

$$\begin{aligned}T^{-\frac{1}{2}} \hat{\beta}_1 &\xrightarrow{d} \beta_1 V_{FF} \beta_2' (\beta_2 V_{FF} \beta_2' + \Omega_{22})^{-1} (\beta_2 (dV_{GG} + \psi_{uG}) + \psi_{v_2G}) \\ &\quad \left[ (\beta_2 (dV_{GG} + \psi_{uG}) + \psi_{v_2G})' (\beta_2 V_{FF} \beta_2' + \Omega_{22})^{-1} (\beta_2 (dV_{GG} + \psi_{uG}) + \psi_{v_2G}) \right]^{-1}.\end{aligned}$$

■

## B. Proof of Corollary 1

**Proof.**

$$\begin{pmatrix} \hat{\lambda}_0 \\ \hat{\lambda}_G \end{pmatrix} = \left[ (\iota_{N_1} : \hat{\beta}_1)' (\iota_{N_1} : \hat{\beta}_1) \right]^{-1} (\iota_{N_1} : \hat{\beta}_1)' \hat{\mu}_{R_1} = \begin{pmatrix} (\iota'_{N_1} M_{\hat{\beta}_1} \iota_{N_1})^{-1} \iota'_{N_1} M_{\hat{\beta}_1} \hat{\mu}_{R_1} \\ (\hat{\beta}'_1 M_{\iota_{N_1}} \hat{\beta}_1)^{-1} \hat{\beta}'_1 M_{\iota_{N_1}} \hat{\mu}_{R_1} \end{pmatrix}$$

where from Assumption 1:  $\hat{\mu}_{R_1} = \mu_{R_1} + \frac{\psi_{R_1}}{\sqrt{T}} + o_p(1/\sqrt{T})$ , while the limiting behavior of  $\hat{\beta}_1$  is provided by Theorem 1.

The strong factor case:  $\hat{\beta}_1 \xrightarrow{p} \beta_1 V_{FF} \delta^{-1'} V_{GG}^{-1}$ ,  $\hat{\mu}_{R_1} \xrightarrow{p} \iota_{N_1} \lambda_0 + \beta_1 \lambda_F$ , so

$$\begin{aligned} \hat{\lambda}_G &= (\hat{\beta}'_1 M_{\iota_{N_1}} \hat{\beta}_1)^{-1} \hat{\beta}'_1 M_{\iota_{N_1}} \hat{\mu}_{R_1} \\ &\xrightarrow{p} ((\beta_1 V_{FF} \delta^{-1'} V_{GG}^{-1})' M_{\iota_{N_1}} (\beta_1 V_{FF} \delta^{-1'} V_{GG}^{-1}))^{-1} (\beta_1 V_{FF} \delta^{-1'} V_{GG}^{-1})' M_{\iota_{N_1}} (\iota_{N_1} \lambda_0 + \beta_1 \lambda_F) \\ &= ((\beta_1 V_{FF} \delta^{-1'} V_{GG}^{-1})' M_{\iota_{N_1}} (\beta_1 V_{FF} \delta^{-1'} V_{GG}^{-1}))^{-1} (\beta_1 V_{FF} \delta^{-1'} V_{GG}^{-1})' M_{\iota_{N_1}} \beta_1 \lambda_F \\ &= V_{GG} \delta' V_{FF}^{-1} \lambda_F \end{aligned}$$

The weak factor case with  $\delta = d/\sqrt{T}$ : for convenience, we write  $T^{-\frac{1}{2}} \hat{\beta}_1 \xrightarrow{d} \Psi_{\beta_1}$ , where  $\Psi_{\beta_1}$  is the non-standard distribution in Theorem 1, so

$$\begin{aligned} \hat{\lambda}_G &= (\hat{\beta}'_1 M_{\iota_{N_1}} \hat{\beta}_1)^{-1} \hat{\beta}'_1 M_{\iota_{N_1}} \hat{\mu}_{R_1} \\ &\simeq \frac{1}{\sqrt{T}} (\Psi'_{\beta_1} M_{\iota_{N_1}} \Psi_{\beta_1})^{-1} \Psi'_{\beta_1} M_{\iota_{N_1}} (\mu_{R_1} + \frac{\psi_{R_1}}{\sqrt{T}}) \end{aligned}$$

which implies

$$T \hat{\lambda}_G - \sqrt{T} (\Psi'_{\beta_1} M_{\iota_{N_1}} \Psi_{\beta_1})^{-1} \Psi'_{\beta_1} M_{\iota_{N_1}} \mu_{R_1} \xrightarrow{d} (\Psi'_{\beta_1} M_{\iota_{N_1}} \Psi_{\beta_1})^{-1} \Psi'_{\beta_1} M_{\iota_{N_1}} \psi_{R_1}$$

■

## C. Proof of Theorem 2

**Proof.** To derive the limiting behavior of  $\hat{\beta}_1$ , we first study  $\widehat{V}_{R_1 R_2}$ ,  $\widehat{V}_{R_2 R_2}$  and  $\widehat{V}_{R_2 G}$ :

$$\begin{aligned} \widehat{V}_{R_1 R_2} &\xrightarrow{p} \beta_1 V_{FF} \beta'_2 \\ \widehat{V}_{R_2 R_2} &\xrightarrow{p} \beta_2 V_{FF} \beta'_2 + \Omega_{22} \end{aligned}$$

Consider

$$V_{R_2 R_2} = \beta_2 A \beta'_2 + \beta_{2,\perp} B \beta'_{2,\perp} + \beta_2 C \beta'_{2,\perp} + \beta_{2,\perp} D \beta'_2$$

with  $A = V_{FF} + (\beta'_2\beta_2)^{-1}\beta'_2\Omega_{22}\beta_2(\beta'_2\beta_2)^{-1}$ ,  $B = (\beta'_{2,\perp}\beta_{2,\perp})^{-1}\beta'_{2,\perp}\Omega_{22}\beta_{2,\perp}(\beta'_{2,\perp}\beta_{2,\perp})^{-1}$ ,  $C = (\beta'_2\beta_2)^{-1}\beta'_2\Omega_{22}\beta_{2,\perp}(\beta'_{2,\perp}\beta_{2,\perp})^{-1}$ ,  $D = (\beta'_{2,\perp}\beta_{2,\perp})^{-1}\beta'_{2,\perp}\Omega_{22}\beta_2(\beta'_2\beta_2)^{-1}$  so the inverse of  $V_{R_2R_2}$  can be specified as

$$V_{R_2R_2}^{-1} = \beta_2\tilde{A}\beta'_2 + \beta_{2,\perp}\tilde{B}\beta'_{2,\perp} + \beta_2\tilde{C}\beta'_{2,\perp} + \beta_{2,\perp}\tilde{D}\beta'_2$$

1. When  $\delta$  is fixed and the number of elements of  $G$  equals the number of elements of  $F$ , so  $\delta$  is a square invertible matrix:

$$\widehat{V}_{R_2G} \xrightarrow{p} \beta_2\delta V_{GG}$$

so

$$\hat{\beta}_1 \xrightarrow{p} \beta_1 V_{FF}\beta'_2(\beta_2 V_{FF}\beta'_2 + \Omega_{22})^{-1}\beta_2\delta V_{GG}.$$

Since we only need the inverse in the direction of  $\beta_2$ , we can now use that when the number of elements of  $\beta_2$  increases:

$$V_{FF}\beta'_2(\beta_2 V_{FF}\beta'_2 + \Omega_{22})^{-1}\beta_2 \approx I$$

since  $\tilde{A} \approx (\beta'_2\beta_2)^{-1}V_{FF}(\beta'_2\beta_2)^{-1}$  when the number of elements of  $\beta_2$  is large, so

$$\hat{\beta}_1 \xrightarrow{p} \beta_1\delta V_{GG}$$

2. When  $\delta = d/\sqrt{T}$ :

$$\sqrt{T}\widehat{V}_{R_2G} \xrightarrow{d} \beta_2(dV_{GG} + \psi_{uG}) + \psi_{v_2G}$$

so

$$T^{\frac{1}{2}}\hat{\beta}_1 \xrightarrow{d} \beta_1 V_{FF}\beta'_2(\beta_2 V_{FF}\beta'_2 + \Omega_{22})^{-1}(\beta_2(dV_{GG} + \psi_{uG}) + \psi_{v_2G})$$

When the number of elements of  $\beta_2$  increases,  $V_{FF}\beta'_2(\beta_2 V_{FF}\beta'_2 + \Omega_{22})^{-1}\beta_2 \approx I$ , so:

$$T^{\frac{1}{2}}\hat{\beta}_1 \approx \beta_1 dV_{GG} + \beta_1(\psi_{uG} + V_{FF}\beta'_2(\beta_2 V_{FF}\beta'_2 + \Omega_{22})^{-1}\psi_{v_2G})$$

■

## D. Asymptotic Variance of Beta Estimators for Rank Testing

### D0. Joint Behavior of $\widehat{V}_{R_2G}$ , $\widehat{V}_{R_1R_2}$ and $\widehat{V}_{R_2R_2}$

Here we present the joint behavior of  $\widehat{V}_{R_2G}$ ,  $\widehat{V}_{R_1R_2}$  and  $\widehat{V}_{R_2R_2}$ , since they make the beta estimator of mimicking portfolios.

We use the following notation:

$$\frac{1}{\sqrt{T}} \sum_{t=1}^T \begin{pmatrix} \bar{G}_t \otimes e_{1t} \\ \bar{G}_t \otimes e_{2t} \\ e_{1t} \otimes e_{2t} - \text{vec}(\Sigma_{12}) \\ e_{2t} \otimes e_{2t} - \text{vec}(\Sigma_{22}) \\ \bar{G}_t \otimes \bar{G}_t - \text{vec}(V_{GG}) \end{pmatrix} \xrightarrow{d} \begin{pmatrix} \text{vec}(\psi_{e_1G}) \\ \text{vec}(\psi_{e_2G}) \\ \text{vec}(\psi_{e_1e_2}) \\ \text{vec}(\psi_{e_2e_2}) \\ \text{vec}(\psi_{GG}) \end{pmatrix},$$

$\widehat{V}_{R_2G}$ ,  $\widehat{V}_{R_1R_2}$  and  $\widehat{V}_{R_2R_2}$  can be rewritten as follows.

$$\begin{aligned} \widehat{V}_{R_2G} &= \frac{1}{T} \sum_{t=1}^T \bar{R}_{2t} \bar{G}'_t \\ &= \frac{1}{T} \sum_{t=1}^T (\beta_2 \delta \bar{G}_t + e_{2t}) \bar{G}'_t \\ &= \beta_2 \delta V_{GG} + \beta_2 \delta \left( \frac{1}{T} \sum_{t=1}^T \bar{G}_t \bar{G}'_t - V_{GG} \right) + \frac{1}{T} \sum_{t=1}^T e_{2t} \bar{G}'_t \end{aligned}$$

$$\begin{aligned} \widehat{V}_{R_1R_2} &= \frac{1}{T} \sum_{t=1}^T \bar{R}_{1t} \bar{R}'_{2t} \\ &= \frac{1}{T} \sum_{t=1}^T (\beta_1 \delta \bar{G}_t + e_{1t}) (\beta_2 \delta \bar{G}_t + e_{2t})' \\ &= \beta_1 \delta V_{GG} \delta' \beta'_2 + \beta_1 \delta \left( \frac{1}{T} \sum_{t=1}^T \bar{G}_t \bar{G}'_t - V_{GG} \right) \delta' \beta'_2 + \beta_1 \delta \left( \frac{1}{T} \sum_{t=1}^T \bar{G}_t e'_{2t} \right) \\ &\quad + \left( \frac{1}{T} \sum_{t=1}^T e_{1t} \bar{G}'_t \right) \delta' \beta'_2 + \frac{1}{T} \sum_{t=1}^T e_{1t} e'_{2t} - \Sigma_{12} + \Sigma_{12} \end{aligned}$$

$$\begin{aligned} \widehat{V}_{R_2R_2} &= \frac{1}{T} \sum_{t=1}^T \bar{R}_{2t} \bar{R}'_{2t} \\ &= \frac{1}{T} \sum_{t=1}^T (\beta_2 \delta \bar{G}_t + e_{2t}) (\beta_2 \delta \bar{G}_t + e_{2t})' \\ &= \beta_2 \delta V_{GG} \delta' \beta'_2 + \beta_2 \delta \left( \frac{1}{T} \sum_{t=1}^T \bar{G}_t \bar{G}'_t - V_{GG} \right) \delta' \beta'_2 + \beta_2 \delta \left( \frac{1}{T} \sum_{t=1}^T \bar{G}_t e'_{2t} \right) \\ &\quad + \left( \frac{1}{T} \sum_{t=1}^T e_{2t} \bar{G}'_t \right) \delta' \beta'_2 + \frac{1}{T} \sum_{t=1}^T e_{2t} e'_{2t} - \Sigma_{22} + \Sigma_{22} \end{aligned}$$

The convergence of  $\widehat{V}_{R_2G}$ ,  $\widehat{V}_{R_2R_1}$  and  $\widehat{V}_{R_2R_2}$  is thus characterized by:

$$\begin{aligned} \sqrt{T} (\widehat{V}_{R_2G} - \beta_2 \delta V_{GG}) &\xrightarrow{d} \beta_2 \delta \psi_{GG} + \psi_{e_2G} && \text{denoted by } U_{R_2G} \\ \sqrt{T} \left( \widehat{V}_{R_1R_2} - \beta_1 \delta V_{GG} \delta' \beta'_2 - \Sigma_{12} \right) &\xrightarrow{d} \beta_1 \delta \psi_{GG} \delta' \beta'_2 + \beta_1 \delta \psi'_{e_2G} + \psi_{e_1G} \delta' \beta'_2 + \psi_{e_1e_2} && \text{denoted by } U_{R_1R_2} \\ \sqrt{T} \left( \widehat{V}_{R_2R_2} - \beta_2 \delta V_{GG} \delta' \beta'_2 - \Sigma_{22} \right) &\xrightarrow{d} \beta_2 \delta \psi_{GG} \delta' \beta'_2 + \beta_2 \delta \psi'_{e_2G} + \psi_{e_2G} \delta' \beta'_2 + \psi_{e_2e_2} && \text{denoted by } U_{R_2R_2} \end{aligned}$$

and in our simulation setup, their variances read:

$$\text{vec}(U_{R_2G}) \sim N(0, V_{GG} \otimes (V_{R_2R_2} + \beta_2 \delta V_{GG} \delta' \beta'_2)) \equiv W_{1,1})$$

$$vec(U_{R_1 R_2}) \sim N(0, V_{R_2 R_2} \otimes (V_{R_1 R_1} + V_{R_1 R_2} V_{R_2 R_2}^{-1} V_{R_2 R_1})) \equiv W_{2,2}$$

$$vec(U_{R_2 R_2}) \sim N(0, 2V_{R_2 R_2} \otimes V_{R_2 R_2} \equiv W_{3,3})$$

In addition, their covariances read:

$$\begin{aligned} cov(vec(U_{R_1 R_2}), vec(U_{R_2 G})) &= 2\beta_2 \delta V_{GG} \otimes \beta_1 \delta V_{GG} \delta' \beta_2' + K_{N_1 N_2} (\beta_1 \delta V_{GG} \otimes \Sigma_{22}) + \beta_2 \delta V_{GG} \otimes \Sigma_{12} \\ &\equiv W_{2,1} = W_{1,2}' \end{aligned}$$

$$\begin{aligned} cov(vec(U_{R_2 R_2}), vec(U_{R_2 G})) &= 2\beta_2 \delta V_{GG} \otimes \beta_2 \delta V_{GG} \delta' \beta_2' + (I_{N_2^2} + K_{N_2 N_2}) (\beta_2 \delta V_{GG} \otimes \Sigma_{22}) \\ &\equiv W_{3,1} = W_{1,3}' \end{aligned}$$

$$\begin{aligned} cov(vec(U_{R_1 R_2}), vec(U_{R_2 R_2})) &= 2V_{R_2 R_2} \otimes V_{R_1 R_2} \\ &\equiv W_{2,3} = W_{3,2}' \end{aligned}$$

To summarize:

$$\sqrt{T} vec \begin{pmatrix} \widehat{V}_{R_2 G} - V_{R_2 G} \\ \widehat{V}_{R_1 R_2} - V_{R_1 R_2} \\ \widehat{V}_{R_2 R_2} - V_{R_2 R_2} \end{pmatrix} = vec \begin{pmatrix} U_{R_2 G} \\ U_{R_1 R_2} \\ U_{R_2 R_2} \end{pmatrix} \xrightarrow{d} N(0, W_U), W_U = \begin{pmatrix} W_{1,1} & W_{1,2} & W_{1,3} \\ W_{2,1} & W_{2,2} & W_{2,3} \\ W_{3,1} & W_{3,2} & W_{3,3} \end{pmatrix}$$

This result will be useful, when we derive asymptotic variance of the beta estimator with mimicking portfolios. Note that  $W_U$  can be consistently estimated by data, so there exists  $\widehat{W}_U \xrightarrow{p} W_U$ . Specifically:

$$\begin{aligned} \widehat{W}_{1,1} &= \widehat{V}_{GG} \otimes (\widehat{V}_{R_2 R_2} + \widehat{B}_2 \widehat{V}_{GG} \widehat{B}_2') \\ \widehat{W}_{2,2} &= \widehat{V}_{R_2 R_2} \otimes (\widehat{V}_{R_1 R_1} + \widehat{V}_{R_1 R_2} \widehat{V}_{R_2 R_2}^{-1} \widehat{V}_{R_2 R_1}) \\ \widehat{W}_{3,3} &= 2\widehat{V}_{R_2 R_2} \otimes \widehat{V}_{R_2 R_2} \\ \widehat{W}_{2,1} = \widehat{W}_{1,2}' &= 2\widehat{B}_2 \widehat{V}_{GG} \otimes \widehat{B}_1 \widehat{V}_{GG} \widehat{B}_2' + K_{N_1 N_2} (\widehat{B}_1 \widehat{V}_{GG} \otimes \widehat{\Sigma}_{22}) + \widehat{B}_2 \widehat{V}_{GG} \otimes \widehat{\Sigma}_{12} \\ \widehat{W}_{3,1} = \widehat{W}_{1,3}' &= 2\widehat{B}_2 \widehat{V}_{GG} \otimes \widehat{B}_2 \widehat{V}_{GG} \widehat{B}_2' + (I_{N_2^2} + K_{N_2 N_2}) (\widehat{B}_2 \widehat{V}_{GG} \otimes \widehat{\Sigma}_{22}) \\ \widehat{W}_{2,3} = \widehat{W}_{3,2}' &= 2\widehat{V}_{R_2 R_2} \otimes \widehat{V}_{R_1 R_2} \end{aligned}$$

## D1. Beta Specification with Factors

Let  $G_t$  be the observed factors,  $R_{1t}$  be test assets. Consider the estimand

$$V_{R_1G}V_{GG}^{-1}$$

and the estimator

$$\widehat{V}_{R_1G}\widehat{V}_{GG}^{-1}$$

Then:

$$\sqrt{T}vec(\widehat{V}_{R_1G}\widehat{V}_{GG}^{-1} - V_{R_1G}V_{GG}^{-1}) \xrightarrow{d} N(0, V_{GG}^{-1} \otimes (V_{R_1R_1} - V_{R_1G}V_{GG}^{-1}V_{GR_1}))$$

This result allows us to conduct the rank test. That is, the estimator  $\widehat{V}_{R_1G}\widehat{V}_{GG}^{-1}$  and its estimated variance  $\widehat{V}_{GG}^{-1} \otimes (\widehat{V}_{R_1R_1} - \widehat{V}_{R_1G}\widehat{V}_{GG}^{-1}\widehat{V}_{GR_1})$  are used for the rank test.

## D2. Beta Specification with Mimicking Portfolios

Let  $G_t$  be the observed factors,  $R_{1t}$  be test assets. Define

$$\tilde{\beta}_1 = V_{R_1R_2}V_{R_2R_2}^{-1}V_{R_2G}(V_{GR_2}V_{R_2R_2}^{-1}V_{R_2G})^{-1}$$

and the estimator

$$\hat{\tilde{\beta}}_1 = \widehat{V}_{R_1R_2}\widehat{V}_{R_2R_2}^{-1}\widehat{V}_{R_2G}(\widehat{V}_{GR_2}\widehat{V}_{R_2R_2}^{-1}\widehat{V}_{R_2G})^{-1}$$

Using that  $\widehat{V}_{R_1R_2} = V_{R_1R_2} + \frac{1}{\sqrt{T}}U_{R_1R_2}$ ,  $\widehat{V}_{R_2R_2} = V_{R_2R_2} + \frac{1}{\sqrt{T}}U_{R_2R_2}$ ,  $\widehat{V}_{R_2G} = V_{R_2G} + \frac{1}{\sqrt{T}}U_{R_2G}$ :

$$\begin{aligned} \widehat{V}_{R_1R_2}\widehat{V}_{R_2R_2}^{-1}\widehat{V}_{R_2G} &= \left[ V_{R_1R_2} + \frac{1}{\sqrt{T}}U_{R_1R_2} \right] \left[ V_{R_2R_2} + \frac{1}{\sqrt{T}}U_{R_2R_2} \right]^{-1} \left[ V_{R_2G} + \frac{1}{\sqrt{T}}U_{R_2G} \right] \\ &\approx V_{R_1R_2}V_{R_2R_2}^{-1}V_{R_2G} + \\ &\quad \frac{1}{\sqrt{T}} \left[ U_{R_1R_2}V_{R_2R_2}^{-1}V_{R_2G} + V_{R_1R_2}V_{R_2R_2}^{-1}U_{R_2G} - V_{R_1R_2}V_{R_2R_2}^{-1}U_{R_2R_2}V_{R_2R_2}^{-1}V_{R_2G} \right] \end{aligned}$$

and similarly,

$$\begin{aligned} \widehat{V}_{GR_2}\widehat{V}_{R_2R_2}^{-1}\widehat{V}_{R_2G} &= \left[ V_{R_2G} + \frac{1}{\sqrt{T}}U_{R_2G} \right]' \left[ V_{R_2R_2} + \frac{1}{\sqrt{T}}U_{R_2R_2} \right]^{-1} \left[ V_{R_2G} + \frac{1}{\sqrt{T}}U_{R_2G} \right] \\ &\approx V_{GR_2}V_{R_2R_2}^{-1}V_{R_2G} + \\ &\quad \frac{1}{\sqrt{T}} \left[ U'_{R_2G}V_{R_2R_2}^{-1}V_{R_2G} + V'_{R_2G}V_{R_2R_2}^{-1}U_{R_2G} - V'_{R_2G}V_{R_2R_2}^{-1}U_{R_2R_2}V_{R_2R_2}^{-1}V_{R_2G} \right] \end{aligned}$$

Use the expansion of the inverse,  $\hat{\tilde{\beta}}_1$  is rewritten as:

$$\begin{aligned}\hat{\tilde{\beta}}_1 &\approx \tilde{\beta}_1 \\ &+ \frac{1}{\sqrt{T}} \left[ U_{R_1 R_2} V_{R_2 R_2}^{-1} V_{R_2 G} + V_{R_1 R_2} V_{R_2 R_2}^{-1} U_{R_2 G} - V_{R_1 R_2} V_{R_2 R_2}^{-1} U_{R_2 R_2} V_{R_2 R_2}^{-1} V_{R_2 G} \right] (V_{G R_2} V_{R_2 R_2}^{-1} V_{R_2 G})^{-1} \\ &- \frac{1}{\sqrt{T}} \tilde{\beta}_1 \left[ U'_{R_2 G} V_{R_2 R_2}^{-1} V_{R_2 G} + V'_{R_2 G} V_{R_2 R_2}^{-1} U_{R_2 G} - V'_{R_2 G} V_{R_2 R_2}^{-1} U_{R_2 R_2} V_{R_2 R_2}^{-1} V_{R_2 G} \right] (V_{G R_2} V_{R_2 R_2}^{-1} V_{R_2 G})^{-1}\end{aligned}$$

So

$$\begin{aligned}\sqrt{T} \text{vec}(\hat{\tilde{\beta}}_1 - \tilde{\beta}_1) &\approx \\ &[(V_{G R_2} V_{R_2 R_2}^{-1} V_{R_2 G})^{-1} \otimes (V_{R_1 R_2} V_{R_2 R_2}^{-1} - \tilde{\beta}_1 V_{G R_2} V_{R_2 R_2}^{-1}) - (V_{G R_2} V_{R_2 R_2}^{-1} V_{R_2 G})^{-1} V_{G R_2} V_{R_2 R_2}^{-1} \otimes \tilde{\beta}_1 \cdot K_{N_2 K}] \text{vec}(U_{R_2 G}) \\ &+ [(V_{G R_2} V_{R_2 R_2}^{-1} V_{R_2 G})^{-1} V_{G R_2} V_{R_2 R_2}^{-1} \otimes I_{N_1}] \text{vec}(U_{R_1 R_2}) \\ &+ [(V_{G R_2} V_{R_2 R_2}^{-1} V_{R_2 G})^{-1} V_{G R_2} V_{R_2 R_2}^{-1} \otimes (\tilde{\beta}_1 V_{G R_2} V_{R_2 R_2}^{-1} - V_{R_1 R_2} V_{R_2 R_2}^{-1})] \text{vec}(U_{R_2 R_2})\end{aligned}$$

where  $K_{N_2 K}$  is a commutation matrix such that  $\text{vec}(U'_{R_2 G}) = K_{N_2 K} \text{vec}(U_{R_2 G})$ .

Consequently, in order to further derive variance of  $\hat{\tilde{\beta}}_1$ , we use the joint behavior of  $U_{R_2 G}$ ,  $U_{R_1 R_2}$  and  $U_{R_2 R_2}$  (which has been derived earlier, see the detail of  $W_U$  and  $\widehat{W}_U$  in the D0 sub-section):

$$\text{vec} \begin{pmatrix} U_{R_2 G} \\ U_{R_1 R_2} \\ U_{R_2 R_2} \end{pmatrix} \xrightarrow{d} N(0, W_U), W_U = \begin{pmatrix} W_{1,1} & W_{1,2} & W_{1,3} \\ W_{2,1} & W_{2,2} & W_{2,3} \\ W_{3,1} & W_{3,2} & W_{3,3} \end{pmatrix}$$

Combining all these pieces, the asymptotic variance of  $\sqrt{T} \text{vec}(\hat{\tilde{\beta}}_1 - \tilde{\beta}_1)$  reads:

$$\mathbf{v}_{\hat{\tilde{\beta}}_1} W_U \mathbf{v}'_{\hat{\tilde{\beta}}_1}$$

where  $\mathbf{v}_{\hat{\tilde{\beta}}_1}$  is defined as

$$\begin{pmatrix} \left[ (V_{G R_2} V_{R_2 R_2}^{-1} V_{R_2 G})^{-1} \otimes (V_{R_1 R_2} V_{R_2 R_2}^{-1} - \tilde{\beta}_1 V_{G R_2} V_{R_2 R_2}^{-1}) - (V_{G R_2} V_{R_2 R_2}^{-1} V_{R_2 G})^{-1} V_{G R_2} V_{R_2 R_2}^{-1} \otimes \tilde{\beta}_1 \cdot K_{N_2 K} \right]' \\ \left[ (V_{G R_2} V_{R_2 R_2}^{-1} V_{R_2 G})^{-1} V_{G R_2} V_{R_2 R_2}^{-1} \otimes I_{N_1} \right]' \\ \left[ (V_{G R_2} V_{R_2 R_2}^{-1} V_{R_2 G})^{-1} V_{G R_2} V_{R_2 R_2}^{-1} \otimes (\tilde{\beta}_1 V_{G R_2} V_{R_2 R_2}^{-1} - V_{R_1 R_2} V_{R_2 R_2}^{-1}) \right]' \end{pmatrix}'$$

with  $\hat{\mathbf{v}}_{\hat{\tilde{\beta}}_1}$  equals

$$\begin{pmatrix} \left[ (\widehat{V}_{G R_2} \widehat{V}_{R_2 R_2}^{-1} \widehat{V}_{R_2 G})^{-1} \otimes (\widehat{V}_{R_1 R_2} \widehat{V}_{R_2 R_2}^{-1} - \hat{\tilde{\beta}}_1 \widehat{V}_{G R_2} \widehat{V}_{R_2 R_2}^{-1}) - (\widehat{V}_{G R_2} \widehat{V}_{R_2 R_2}^{-1} \widehat{V}_{R_2 G})^{-1} \widehat{V}_{G R_2} \widehat{V}_{R_2 R_2}^{-1} \otimes \hat{\tilde{\beta}}_1 \cdot K_{N_2 K} \right]' \\ \left[ (\widehat{V}_{G R_2} \widehat{V}_{R_2 R_2}^{-1} \widehat{V}_{R_2 G})^{-1} \widehat{V}_{G R_2} \widehat{V}_{R_2 R_2}^{-1} \otimes I_{N_1} \right]' \\ \left[ (\widehat{V}_{G R_2} \widehat{V}_{R_2 R_2}^{-1} \widehat{V}_{R_2 G})^{-1} \widehat{V}_{G R_2} \widehat{V}_{R_2 R_2}^{-1} \otimes (\hat{\tilde{\beta}}_1 \widehat{V}_{G R_2} \widehat{V}_{R_2 R_2}^{-1} - \widehat{V}_{R_1 R_2} \widehat{V}_{R_2 R_2}^{-1}) \right]' \end{pmatrix}'$$

This result allows us to conduct the rank test. That is, the estimator  $\hat{\beta}_1$  and its estimated variance  $\hat{\mathbf{v}}_{\hat{\beta}_1} \widehat{W}_U \hat{\mathbf{v}}'_{\hat{\beta}_1}$  are used for the rank test.

### D3. Covariance Specification with Factors

Let  $G_t$  be the observed factors,  $R_{1t}$  be test assets. Consider the estimand

$$V_{R_1G}$$

and the estimator

$$\widehat{V}_{R_1G}$$

Then:

$$\sqrt{T} \text{vec}(\widehat{V}_{R_1G} - V_{R_1G}) \xrightarrow{d} N(0, V_{GG} \otimes (V_{R_1R_1} + V_{R_1G} V_{GG}^{-1} V_{GR_1}))$$

This result allows us to conduct the rank test. That is, the estimator  $\widehat{V}_{R_1G}$  and its estimated variance  $\widehat{V}_{GG} \otimes (\widehat{V}_{R_1R_1} + \widehat{V}_{R_1G} \widehat{V}_{GG}^{-1} \widehat{V}_{GR_1})$  are used for the rank test.

### D4. Covariance Specification with Mimicking Portfolios

Consider the estimator

$$\hat{\beta}_1 = \widehat{V}_{R_1R_2} \widehat{V}_{R_2R_2}^{-1} \widehat{V}_{R_2G}$$

Using that  $\widehat{V}_{R_1R_2} = V_{R_1R_2} + \frac{1}{\sqrt{T}} U_{R_1R_2}$ ,  $\widehat{V}_{R_2R_2} = V_{R_2R_2} + \frac{1}{\sqrt{T}} U_{R_2R_2}$ ,  $\widehat{V}_{R_2G} = V_{R_2G} + \frac{1}{\sqrt{T}} U_{R_2G}$ , we can specify the above estimator as

$$\begin{aligned} \hat{\beta}_1 &= \widehat{V}_{R_1R_2} \widehat{V}_{R_2R_2}^{-1} \widehat{V}_{R_2G} \\ &= \left[ V_{R_1R_2} + \frac{1}{\sqrt{T}} U_{R_1R_2} \right] \left[ V_{R_2R_2} + \frac{1}{\sqrt{T}} U_{R_2R_2} \right]^{-1} \left[ V_{R_2G} + \frac{1}{\sqrt{T}} U_{R_2G} \right] \\ &= V_{R_1R_2} V_{R_2R_2}^{-1} V_{R_2G} + \\ &\quad \frac{1}{\sqrt{T}} \left[ U_{R_1R_2} V_{R_2R_2}^{-1} V_{R_2G} + V_{R_1R_2} V_{R_2R_2}^{-1} U_{R_2G} - V_{R_1R_2} V_{R_2R_2}^{-1} U_{R_2R_2} V_{R_2R_2}^{-1} V_{R_2G} \right] + o_p(T^{-1/2}). \end{aligned}$$

Here we used an expansion of the inverse. Hence

$$\sqrt{T}(\hat{\beta}_1 - V_{R_1R_2} V_{R_2R_2}^{-1} V_{R_2G}) \approx V_{R_1R_2} V_{R_2R_2}^{-1} U_{R_2G} + U_{R_1R_2} V_{R_2R_2}^{-1} V_{R_2G} - V_{R_1R_2} V_{R_2R_2}^{-1} U_{R_2R_2} V_{R_2R_2}^{-1} V_{R_2G}$$

Consequently, in order to further derive variance of  $\hat{\beta}_1$ , we use the joint behavior of  $U_{R_2G}$ ,  $U_{R_1R_2}$  and  $U_{R_2R_2}$  (which has been derived earlier, see the detail of  $W_U$  and  $\widehat{W}_U$  in the D0



sub-section):

$$vec \begin{pmatrix} U_{R_2G} \\ U_{R_1R_2} \\ U_{R_2R_2} \end{pmatrix} \xrightarrow{d} N(0, W_U), W_U = \begin{pmatrix} W_{1,1} & W_{1,2} & W_{1,3} \\ W_{2,1} & W_{2,2} & W_{2,3} \\ W_{3,1} & W_{3,2} & W_{3,3} \end{pmatrix}$$

Combining all these pieces, the asymptotic variance of  $\sqrt{T}vec(\hat{\beta}_1 - V_{R_1R_2}V_{R_2R_2}^{-1}V_{R_2G})$  reads:

$$\mathbf{v}_{\hat{\beta}_1} W_U \mathbf{v}'_{\hat{\beta}_1}$$

where  $\mathbf{v}_{\hat{\beta}_1}$  is defined as

$$\begin{pmatrix} I_m \otimes V_{R_1R_2}V_{R_2R_2}^{-1} & V_{GR_2}V_{R_2R_2}^{-1} \otimes I_{N_1} & -V_{GR_2}V_{R_2R_2}^{-1} \otimes V_{R_1R_2}V_{R_2R_2}^{-1} \end{pmatrix}$$

with  $\hat{\mathbf{v}}_{\hat{\beta}_1}$  equals

$$\begin{pmatrix} I_m \otimes \hat{V}_{R_1R_2}\hat{V}_{R_2R_2}^{-1} & \hat{V}_{GR_2}\hat{V}_{R_2R_2}^{-1} \otimes I_{N_1} & -\hat{V}_{GR_2}\hat{V}_{R_2R_2}^{-1} \otimes \hat{V}_{R_1R_2}\hat{V}_{R_2R_2}^{-1} \end{pmatrix}$$

This result allows us to conduct the rank test. That is, the estimator  $\hat{\beta}_1$  and its estimated variance  $\hat{\mathbf{v}}_{\hat{\beta}_1}\hat{W}_U\hat{\mathbf{v}}'_{\hat{\beta}_1}$  are used for the rank test.

## E. Proof of Theorem 3

**Proof.** For  $\bar{\mathcal{R}}_1$ , by the central limit theorem, we have:

$$\sqrt{T}(\bar{\mathcal{R}}_1 - \Gamma\beta_2\delta\lambda_{G,cov}) \xrightarrow{d} \psi_1 \sim N(0, V_{\mathcal{R}_1\mathcal{R}_1})$$

$\hat{B}_2$  results from linear regression of  $R_{2t}$  on  $G_t$ , so:

$$\sqrt{T}vec(\hat{B}_2 - \beta_2\delta) \xrightarrow{d} vec(\psi_2) \sim N(0, V_{GG}^{-1} \otimes \Sigma_{22})$$

where  $\Sigma_{22}$  is the covariance matrix of residuals.

Similarly for  $\hat{\Gamma}$  that results from linear regression of  $\mathcal{R}_{1t}$  on  $R_{2t}$ :

$$\sqrt{T}vec(\hat{\Gamma} - \Gamma) \xrightarrow{d} vec(\psi_3) \sim N(0, V_{R_2R_2}^{-1} \otimes (V_{\mathcal{R}_1\mathcal{R}_1} - V_{R_1R_2}V_{R_2R_2}^{-1}V_{R_2\mathcal{R}_1}))$$

Note that  $\psi_1$  and  $\psi_2$  are independent (see Kleibergen 2009). By the same logic,  $\psi_1$  and  $\psi_3$  are also independent. So what remains to derive is  $cov(vec(\psi_3), vec(\psi_2))$ .

For  $\psi_3$ , using that  $\hat{V}_{\mathcal{R}_1 R_2} = V_{\mathcal{R}_1 R_2} + \frac{1}{\sqrt{T}} U_{\mathcal{R}_1 R_2}$ ,  $\hat{V}_{R_2 R_2} = V_{R_2 R_2} + \frac{1}{\sqrt{T}} U_{R_2 R_2}$ ,

$$\begin{aligned}\hat{\Gamma} &= \hat{V}_{\mathcal{R}_1 R_2} \hat{V}_{R_2 R_2}^{-1} \\ &= \left( V_{\mathcal{R}_1 R_2} + \frac{1}{\sqrt{T}} U_{\mathcal{R}_1 R_2} \right) \left( V_{R_2 R_2} + \frac{1}{\sqrt{T}} U_{R_2 R_2} \right)^{-1} \\ &= V_{\mathcal{R}_1 R_2} V_{R_2 R_2}^{-1} + \frac{1}{\sqrt{T}} \left[ U_{\mathcal{R}_1 R_2} V_{R_2 R_2}^{-1} - V_{\mathcal{R}_1 R_2} V_{R_2 R_2}^{-1} U_{R_2 R_2} V_{R_2 R_2}^{-1} \right] + o_p\left(\frac{1}{\sqrt{T}}\right)\end{aligned}$$

so, with  $\Gamma = V_{\mathcal{R}_1 R_2} V_{R_2 R_2}^{-1}$ ,

$$\sqrt{T} \text{vec}(\hat{\Gamma} - \Gamma) \xrightarrow{d} (V_{R_2 R_2}^{-1} \otimes I_{N_1-1}) \text{vec}(U_{\mathcal{R}_1 R_2}) - (V_{R_2 R_2}^{-1} \otimes V_{\mathcal{R}_1 R_2} V_{R_2 R_2}^{-1}) \text{vec}(U_{R_2 R_2}) = \text{vec}(\psi_3).$$

where, as in Appendix D,<sup>15</sup>

$$\begin{aligned}U_{\mathcal{R}_1 R_2} &= \beta_1 \delta \psi_{GG} \delta' \beta_2' + \beta_1 \delta \psi'_{e_2 G} + \psi_{e_1 G} \delta' \beta_2' + \psi_{e_1 e_2} \\ U_{R_2 R_2} &= \beta_2 \delta \psi_{GG} \delta' \beta_2' + \beta_2 \delta \psi'_{e_2 G} + \psi_{e_2 G} \delta' \beta_2' + \psi_{e_2 e_2}\end{aligned}$$

For  $\psi_2$ :  $\psi_2 = \psi_{e_2 G} V_{GG}^{-1}$ . So

$$\begin{aligned}\text{cov}(\text{vec}(U_{\mathcal{R}_1 R_2}), \text{vec}(\psi_2)) &= K_{(N_1-1)N_2} (\beta_1 \delta \otimes \Sigma_{22}) + \beta_2 \delta \otimes \Sigma_{12} \\ \text{cov}(\text{vec}(U_{R_2 R_2}), \text{vec}(\psi_2)) &= (I_{N_2^2} + K_{N_2 N_2}) (\beta_2 \delta \otimes \Sigma_{22})\end{aligned}$$

which implies that  $\text{cov}(\text{vec}(\psi_3), \text{vec}(\psi_2))$  equals

$$K_{(N_1-1)N_2} (\beta_1 \delta \otimes V_{R_2 R_2}^{-1} \Sigma_{22}) + V_{R_2 R_2}^{-1} \beta_2 \delta \otimes \Sigma_{12} - (V_{R_2 R_2}^{-1} \otimes V_{\mathcal{R}_1 R_2} V_{R_2 R_2}^{-1}) (I_{N_2^2} + K_{N_2 N_2}) (\beta_2 \delta \otimes \Sigma_{22}) \equiv C$$

■

## F. Proof of Corollary 2

**Proof.** Rewrite  $\bar{\mathcal{R}}_1 - \hat{\Gamma} \hat{B}_2 \lambda_{G, cov}$  as follows.

$$\begin{aligned}& \bar{\mathcal{R}}_1 - \hat{\Gamma} \hat{B}_2 \lambda_{G, cov} \\ &= \bar{\mathcal{R}}_1 - \Gamma \beta_2 \delta \lambda_{G, cov} - (\hat{\Gamma} \hat{B}_2 - \Gamma \beta_2 \delta) \lambda_{G, cov} \\ &= \bar{\mathcal{R}}_1 - \Gamma \beta_2 \delta \lambda_{G, cov} - (\hat{\Gamma} - \Gamma) \beta_2 \delta \lambda_{G, cov} - \hat{\Gamma} (\hat{B}_2 - \beta_2 \delta) \lambda_{G, cov} \\ &= \bar{\mathcal{R}}_1 - \Gamma \beta_2 \delta \lambda_{G, cov} - (\hat{\Gamma} - \Gamma) \beta_2 \delta \lambda_{G, cov} - \Gamma (\hat{B}_2 - \beta_2 \delta) \lambda_{G, cov} - (\hat{\Gamma} - \Gamma) (\hat{B}_2 - \beta_2 \delta) \lambda_{G, cov}\end{aligned}$$

<sup>15</sup>Notation: here  $\beta_1 \delta$  and  $e_1$  are the beta and error that correspond to  $\mathcal{R}_1$  under  $G_t$ , while  $\beta_2 \delta$  and  $e_2$  are the beta and error that correspond to  $R_1$ . Similarly,  $\Sigma_{12} = \text{cov}(e_1, e_2)$ , while  $\Sigma_{22} = \text{cov}(e_2, e_2)$ .

Applying Theorem 3, we get  $\sqrt{T}(\bar{\mathcal{R}}_1 - \hat{\Gamma}\hat{B}_2\lambda_{G,cov}) \xrightarrow{d} \psi_1 - \Gamma\psi_2\lambda_{G,cov} - \psi_3\beta_2\delta\lambda_{G,cov}$ . Since  $\Gamma$  has a full rank value, the last element of the above expression converges to zero for every value of  $\delta$  since  $\sqrt{T}(\hat{\Gamma} - \Gamma) \xrightarrow{d} \psi_3$ .

For the analytical expression of the resulting variance, we use the elements in  $\mathbb{W}$  from Theorem 3. Specifically, for the three terms in  $\psi_1 - \Gamma\psi_2\lambda_{G,cov} - \psi_3\beta_2\delta\lambda_{G,cov}$ , their variances read:  $V_{\mathcal{R}_1\mathcal{R}_1}$ ,  $\lambda'_{G,cov}V_{GG}^{-1}\lambda_{G,cov} \otimes \Gamma\Sigma_{22}\Gamma'$  and  $(\beta_2\delta\lambda_{G,cov})'V_{R_2R_2}^{-1}(\beta_2\delta\lambda_{G,cov}) \otimes (V_{\mathcal{R}_1\mathcal{R}_1} - V_{\mathcal{R}_1R_2}V_{R_2R_2}^{-1}V_{R_2\mathcal{R}_1})$ . In addition, since  $\psi_2$  and  $\psi_3$  are not independent, we also consider the resulting covariance:

$$cov(vec(\psi_3\beta_2\delta\lambda_{G,cov}), vec(\Gamma\psi_2\lambda_{G,cov})) = [(\beta_2\delta\lambda_{G,cov})' \otimes I_{N_1-1}]C(\lambda_{G,cov} \otimes \Gamma')$$

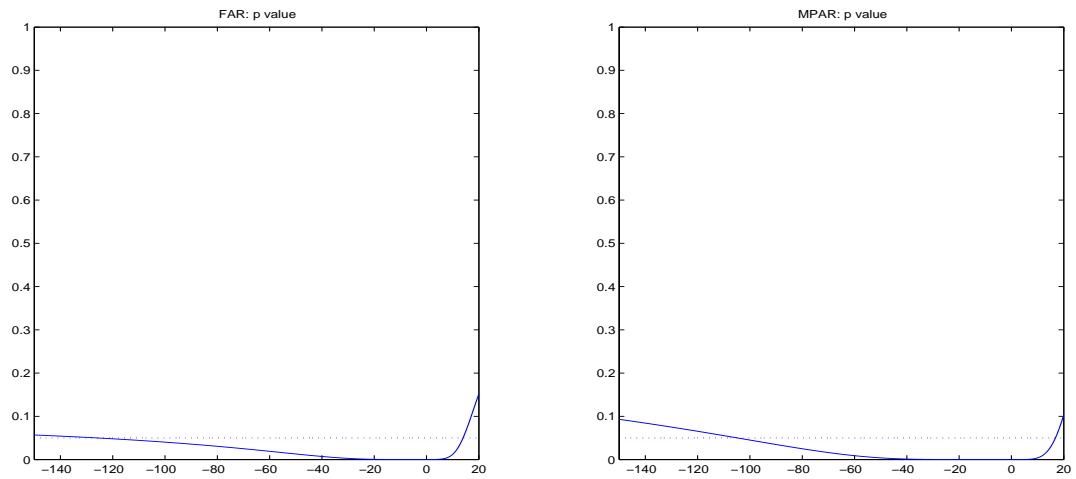
So the asymptotic variance of  $\sqrt{T}(\bar{\mathcal{R}}_1 - \hat{\Gamma}\hat{B}_2\lambda_{G,cov})$  results from combining the variance and covariance terms above:

$$\begin{aligned} & V_{\mathcal{R}_1\mathcal{R}_1} + \lambda'_{G,cov}V_{GG}^{-1}\lambda_{G,cov} \otimes \Gamma\Sigma_{22}\Gamma' + (\beta_2\delta\lambda_{G,cov})'V_{R_2R_2}^{-1}(\beta_2\delta\lambda_{G,cov}) \otimes (V_{\mathcal{R}_1\mathcal{R}_1} - V_{\mathcal{R}_1R_2}V_{R_2R_2}^{-1}V_{R_2\mathcal{R}_1}) \\ & - \{[(\beta_2\delta\lambda_{G,cov})' \otimes I_{N_1-1}]C(\lambda_{G,cov} \otimes \Gamma')\} - \{[(\beta_2\delta\lambda_{G,cov})' \otimes I_{N_1-1}]C(\lambda_{G,cov} \otimes \Gamma')\}' \end{aligned}$$

■

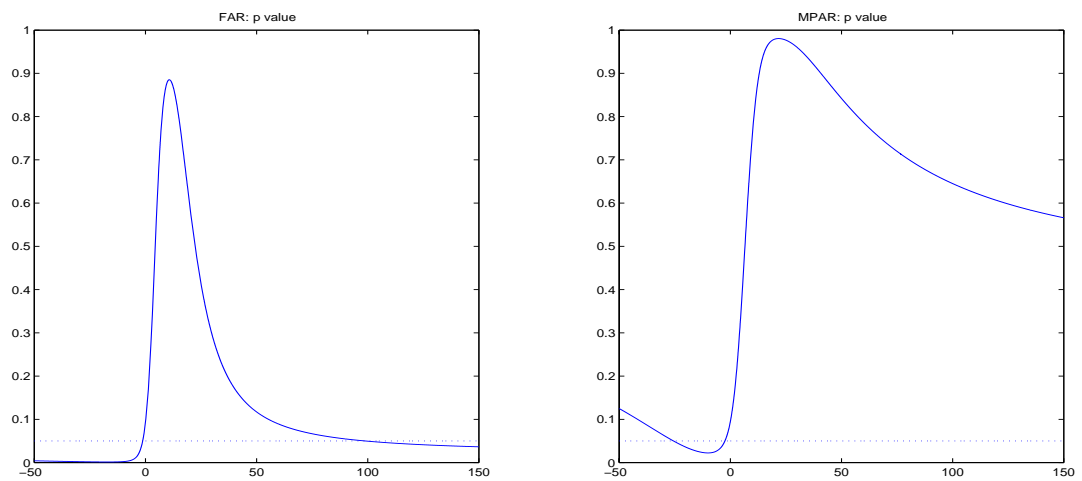
## G. $p$ -value of FAR and MPAR tests for Table 6, Figures 3-7

Figure 3:  $p$ -value of FAR and MPAR tests for Table 6 - I



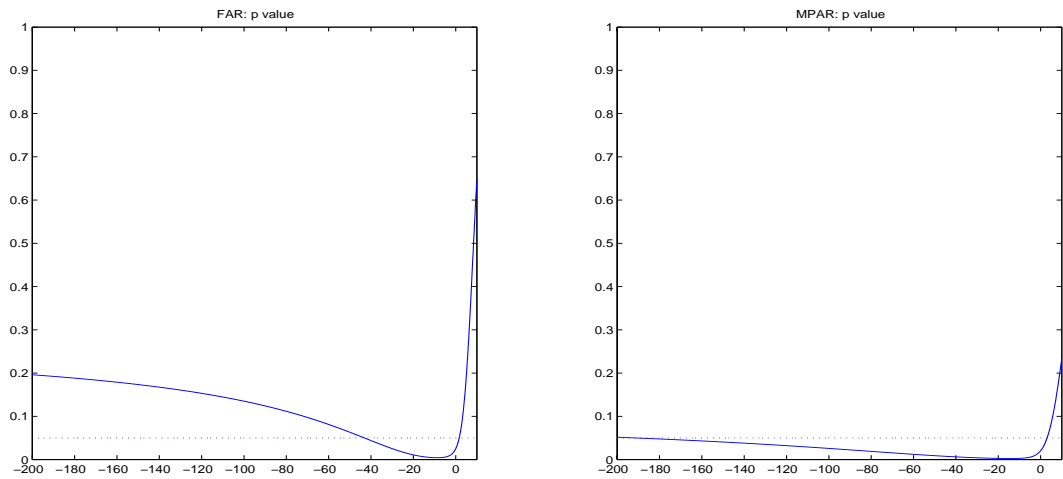
Note: This figure presents the  $p$ -values (solid) of the FAR (left) and MPAR (right) tests for testing risk premium equals the corresponding value on the x-axis, using test assets in I. The 5% line is also plotted for benchmark (dotted).

Figure 4:  $p$ -value of FAR and MPAR tests for Table 6 - II



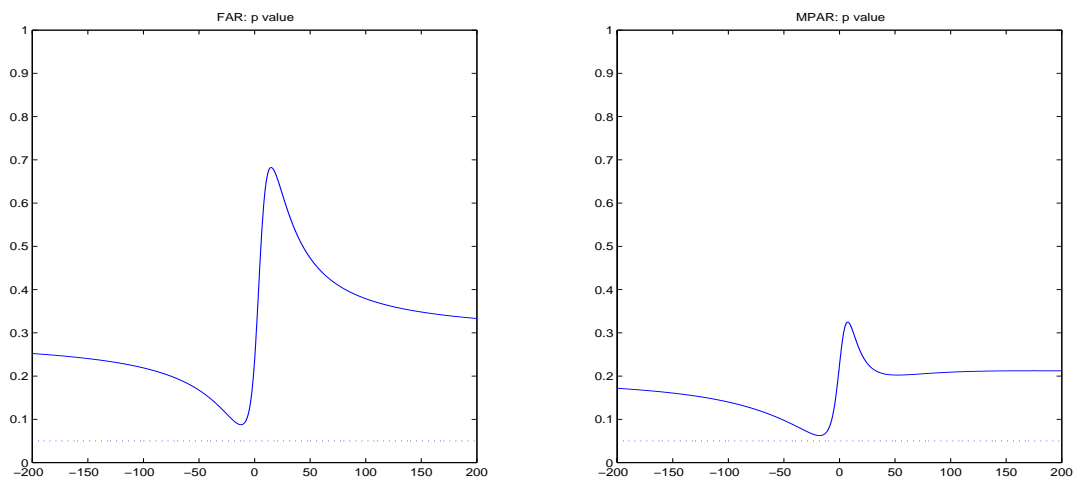
Note: This figure presents the  $p$ -values (solid) of the FAR (left) and MPAR (right) tests for testing risk premium equals the corresponding value on the x-axis, using test assets in II. The 5% line is also plotted for benchmark (dotted).

Figure 5:  $p$ -value of FAR and MPAR tests for Table 6 - III



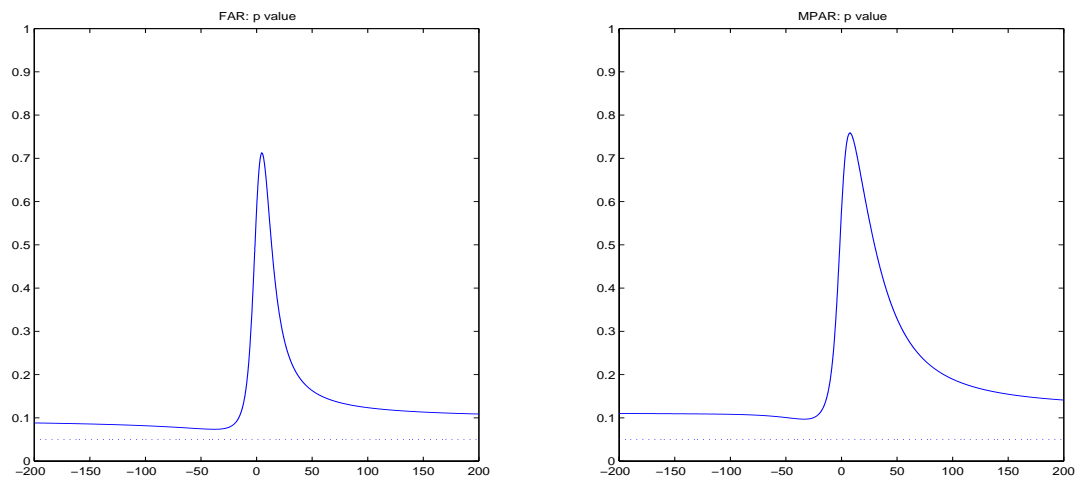
Note: This figure presents the  $p$ -values (solid) of the FAR (left) and MPAR (right) tests for testing risk premium equals the corresponding value on the x-axis, using test assets in III. The 5% line is also plotted for benchmark (dotted).

Figure 6:  $p$ -value of FAR and MPAR tests for Table 6 - IV



Note: This figure presents the  $p$ -values (solid) of the FAR (left) and MPAR (right) tests for testing risk premium equals the corresponding value on the x-axis, using test assets in IV. The 5% line is also plotted for benchmark (dotted).

Figure 7:  $p$ -value of FAR and MPAR tests for Table 6 - V



Note: This figure presents the  $p$ -values (solid) of the FAR (left) and MPAR (right) tests for testing risk premium equals the corresponding value on the x-axis, using test assets in V. The 5% line is also plotted for benchmark (dotted).