

Mimicking Portfolios of Macroeconomic Factors

Frank Kleibergen[†] Zhaoguo Zhan[‡]

December 20, 2014

Abstract

Mimicking portfolios of macroeconomic factors are commonly constructed by projecting these factors on a set of base assets. Current understanding is that compared to macroeconomic factors, their mimicking portfolios contain more relevant information and less noise for asset pricing. In this paper, we show that when factors are in fact useless, their constructed mimicking portfolios may still be spuriously favored by asset pricing tests in the Fama and MacBeth (1973) two-pass procedure. Our findings imply that empirical results based on mimicking portfolios of macroeconomic factors need to be taken with caution, since mimicking portfolios of useless or nearly useless factors may similarly yield such results.

JEL Classification: G12

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

[†]Department of Economics, Brown University, Box B, Providence, RI 02912, United States. Email: Frank_Kleibergen@brown.edu. Homepage: http://www.econ.brown.edu/fac/Frank_Kleibergen.

[‡]Department of Economics, School of Economics and Management, Tsinghua University, Beijing, China, 100084. Email: zhanzhg@sem.tsinghua.edu.cn.

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 beta's are small, the 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 a 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 common way of constructing their mimicking portfolios is to project these factors on the 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 Muir *et al.* (2013).

On the one hand, besides other potential advantages,¹ using mimicking portfolios instead of macroeconomic factors appears able to bypass the weak statistical correlation issue studied in Kleibergen (2009). By projecting macroeconomic factors on the base assets, it is natural that the resulting mimicking portfolios exhibit improved correlation with asset returns so their beta's are amplified. In existing studies (e.g., Vassalou 2003, Muir *et al.* 2013), 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 appear relevant for asset pricing, the risk premia estimator in the FM procedure has a non-standard distribution, which jeopardizes the *t*-test on risk premia. The underlying reason is that when the beta's of macroeconomic factors are small, the beta's of their mimicking portfolios can be spuriously large. Most strikingly, for factors that are completely useless, we find that their mimicking portfolios may still induce significant *t*-statistics on risk premia and large cross-sectional R^2 , both of which spuriously support the background useless factors.

Put differently, our findings 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 port-

¹Such as extending the available data sets (Ang *et al.*, 2006), providing extra restrictions for testing purposes (Huberman *et al.* 1987).

folios 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.

Furthermore, our concern on mimicking portfolios of macroeconomic factors presented in this paper also applies to other risk factors that are only minorly correlated with asset returns. In general, if mimicking portfolios result from the projection of factors with small beta's on base assets, then empirical findings based on such mimicking portfolios need to be taken with caution.

To effectively convey the message of this paper, we proceed as follows. In Section 2, we present a Monte Carlo study, which shows that mimicking portfolios of useless factors can be spuriously favored in the FM two-pass procedure. In Section 3, we provide the analytical results to explain the seemingly astonishing findings in Section 2. Section 4 discusses robust inference methods for risk premia under mimicking portfolios. These robust methods are based on the robust tests of risk premia for factors in Kleibergen (2009), and we modify these tests so that they can be used for mimicking portfolios. Some illustrative empirical examples are also contained in Section 4. Section 5 concludes the paper. Proofs and technical details are provided in the Appendix.

Throughout the paper, we use the following notation: $vec(A)$ stands for the vectorization of the matrix A , i.e., when $A = (a_1, \dots, a_n)$, $vec(A) = (a'_1, \dots, a'_n)'$ and $vecinv_n((a'_1, \dots, a'_n)') = A$, $P_A = A(A'A)^{-1}A'$, $M_A = I - P_A$, I is the identity matrix. \xrightarrow{p} and \xrightarrow{d} stand for convergence in probability and convergence in distribution, respectively.

2 Motivation

As a starting point, we present the outcome of a simple simulation experiment to illustrate the main point of this paper, i.e., mimicking portfolios may spuriously show support for asset pricing, even when their background factors are useless.

2.1 Simulation Design

In the simulation experiment, the data of asset returns is drawn from a normal distribution, i.e., $R_t \sim NID(\mu_R, V_{RR})$, $t = 1, 2, \dots, T$, where the mean μ_R and covariance V_{RR} are calibrated to data of the commonly used twenty-five size and book-to-market sorted portfolios. We set $T = 500$, which is close to the typical sample size used in practice.

Similarly, the data of each artificial risk factor used in the simulation experiment is drawn from a normal distribution with mean and variance calibrated to real data on consumption growth, and we independently generate up to three factors in this manner. Since these factors are simulated independent of returns, they correspond to useless factors in the asset pricing literature.

Our real data sets used for the calibration of mean and variance of returns and factors are adopted from Lettau and Ludvigson (2001).

2.2 Useless Factors and Their Mimicking Portfolios in FM

With the simulated returns and useless factors, we conduct the FM two-pass estimation and report the following statistics in Panel A of Table 1.

	Panel A: Useless Factors				Panel B: Mimicking Portfolios					
	p -value	Frequency of $ t > 1.96$			R^2	p -value	Frequency of $ t > 1.96$			R^2
		(1)	(2)	(3)			(1)	(2)	(3)	
$k = 1$	0.496	0.635			0.184	0	0.662			0.184
2	0.455	0.688	0.704		0.352	0	0.719	0.732		0.352
3	0.412	0.704	0.704	0.711	0.480	0	0.751	0.749	0.753	0.480

Note: p -value results from the rank test of Kleibergen and Paap (2006), which tests that the rank of the beta matrix is $k - 1$ when k factors or their mimicking portfolios are used. The frequency that t -statistic on risk premia in the Fama and MacBeth (1973) two-pass procedure exceeds 1.96 in absolute value is reported. The R^2 is the OLS R^2 in the second pass cross-sectional regression. Each row in this table corresponds to a linear factor model with $k = 1, 2$ or 3 factors, respectively. The number of Monte Carlo replications is 10000. The data generation process is described in the main text.

First, we report the p -value of a rank test on the beta matrix in the first pass time series

regression of the FM two-pass procedure. Specifically, when k factors ($k = 1, 2$ or 3) are used for a linear factor model, we test the null hypothesis that the rank of the corresponding beta matrix is $k - 1$ by the rank test in Kleibergen and Paap (2006). The resulting p -value of this test is presented. A large p -value implies that we could not reject the null, i.e., the beta matrix is unlikely to have full rank, which indicates the statistical quality of factors is under doubt. Second, we report the outcome of the conventional t -test on risk premia in the second pass cross-sectional regression of the FM two-pass procedure. In particular, we document whether risk premium of each of the k factors is significant at the 5% level. Third, we also report the cross-sectional ordinary least squares (OLS) R^2 , the popular measure of the goodness-of-fit of the model. The reported numbers in Table 1 result from the average of 10000 Monte Carlo replications.

In addition, we regress each simulated useless factor on the set of the simulated asset returns to construct its mimicking portfolio (see, e.g., Huberman *et al.* 1987, Breeden *et al.* 1989). We then use these constructed mimicking portfolios to replace useless factors, and conduct the FM two-pass estimation again. Furthermore, we similarly report the aforementioned statistics in Panel B of Table 1, which are now computed with mimicking portfolios of useless factors. In contrast, the statistics in Panel A are computed with useless factors, as we explained above. Our purpose is to compare the statistics in these two panels.

When useless factors are used in two-pass tests, Panel A of Table 1 shows that their risk premia can be significant with a large probability mass and the cross-sectional R^2 could also appear large as the number of factors increases. However, neither the significant risk premia nor the large R^2 can serve as convincing evidence to support the corresponding factors, as emphasized by the existing literature. For example, it is now well-known that that useless factors can induce significant t -statistics for risk premia, while Lewellen *et al.* (2010) and Kleibergen and Zhan (2014) warn that cross-sectional R^2 can be large under useless factors. In order to detect such factors, a rank test can serve as a diagnostic tool, since large p -values in Panel A imply that the full rank condition of the beta matrix for the FM methodology is

not satisfied under useless factors.

The focus of this paper is on Panel B of Table 1, which, similar to Panel A, shows that mimicking portfolios of useless factors may also spuriously yield significant risk premia and large R^2 . However, different from the large p -values of the rank test under useless factors in Panel A, the p -values under mimicking portfolios in Panel B are all close to zero. A rank test may thus incorrectly favor mimicking portfolios resulting from useless factors, in the sense that these mimicking portfolios appear closely related to asset returns so they seemingly contain the relevant information for asset pricing.

2.3 Remarks

This paper is mainly motivated by Table 1. As illustrated by this table, the spurious results induced by useless factors, which are discussed in Kan and Zhang (1999), Kleibergen (2009), Kleibergen and Zhan (2014), etc., prevail under mimicking portfolios of these factors. In addition, the sharp difference in p -values across the two panels of Table 1 indicates that the statistical quality of useless factors substantially differs from that of their mimicking portfolios. As a result, it is worth investigating why these two statistically different objects could similarly induce spurious risk premia and large R^2 in the commonly used FM two-pass procedure.

To further evaluate the findings in Table 1, we also conduct the simulation experiment when factors used in the FM procedure are not useless but are closely related to asset returns. For convenience, we call such factors useful factors.

Specifically, we follow Bai and Ng (2006) (see also Connor and Korajczyk 1988) to construct up to three factors that are closely related to asset returns. These factors result from the largest three principal components of our simulated asset returns. To account for the fact that risk factors used in practice are unlikely to be exactly equal to principal components, we add random noise (drawn from standard normal) to principal components, and consider the resulting sum as the proxy for observed useful factors.

Table 2: Useful Factors and Their Mimicking Portfolios in Fama and MacBeth (1973)

	Panel A: Useful Factors					Panel B: Mimicking Portfolios				
	p -value	Frequency of $ t > 1.96$			R^2	p -value	Frequency of $ t > 1.96$			R^2
		(1)	(2)	(3)			(1)	(2)	(3)	
$k = 1$	0	0.192			0.049	0	0.304			0.049
2	0	0.661	0.994		0.674	0	0.739	0.995		0.674
3	0	0.764	0.622	0.959	0.810	0	0.830	0.699	0.974	0.810

Note: p -value results from the rank test of Kleibergen and Paap (2006), which tests that the rank of the beta matrix is $k - 1$ when k factors or their mimicking portfolios are used. The frequency that t -statistic on risk premia in the Fama and MacBeth (1973) two-pass procedure exceeds 1.96 in absolute value is reported. The R^2 is the OLS R^2 in the second pass cross-sectional regression. Each row in this table corresponds to a linear factor model with $k = 1, 2$ or 3 factors, respectively. The number of Monte Carlo replications is 10000. The data generation process is described in the main text.

Table 2 presents the outcome when useful factors (Panel A) or their mimicking portfolios (Panel B) are used in the FM procedure. The reported numbers in Table 2 are computed in the same manner as in Table 1, but we now employ useful factors for Table 2 instead of useless factors for Table 1. All p -values of the rank test in Table 2 are found to be close to zero, since useful factors as well as their mimicking portfolios are closely related to asset returns. By comparing Table 1 with Table 2, it can be seen that useless factors and their mimicking portfolios induce significant risk premia and large R^2 that are often comparable to those induced by useful factors.

From the empirical perspective, the simulation outcome presented in this section effectively casts doubt on the existing asset pricing studies that rely on mimicking portfolios of macroeconomic factors. It is common that macroeconomic factors are only minorly correlated with asset returns,² thus these factors are almost indistinguishable from useless factors in terms of their correlation with asset returns in finite samples. Given that mimicking portfolios of useless factors may be spuriously supported in asset pricing tests as indicated by Table 1, empirical findings that are based on the mimicking portfolios of macroeconomic factors are thus up to further scrutiny.

²See, e.g., Bai and Ng (2006), Kleibergen and Zhan (2014).

3 Analytical Results

In this section, we proceed to explore why mimicking portfolios of useless or nearly useless factors can be spuriously favored, as illustrated by the simulation experiment above. To facilitate the analysis, we set up notation first.

3.1 Notation

Consider the $(2k + n + \tilde{n}) \times 1$ vector

$$\begin{pmatrix} F_t \\ G_t \\ R_t \\ \tilde{R}_t \end{pmatrix}, \text{ with mean } \begin{pmatrix} \mu_F \\ \mu_G \\ \mu_R \\ \mu_{\tilde{R}} \end{pmatrix} \text{ and covariance } \begin{pmatrix} V_{FF} & V_{FG} & V_{FR} & V_{F\tilde{R}} \\ V_{GF} & V_{GG} & V_{GR} & V_{G\tilde{R}} \\ V_{RF} & V_{RG} & V_{RR} & V_{R\tilde{R}} \\ V_{\tilde{R}F} & V_{\tilde{R}G} & V_{\tilde{R}R} & V_{\tilde{R}\tilde{R}} \end{pmatrix}, \quad (1)$$

where F_t is the $k \times 1$ vector of risk factors, G_t is the $k \times 1$ vector of mimicking portfolios for F_t , R_t is the $n \times 1$ vector of the returns on the base assets that span the asset space, and \tilde{R}_t is the $\tilde{n} \times 1$ vector of the returns on test assets used in the FM two-pass procedure. We assume stationarity and ergodicity of $(F'_t, G'_t, R'_t, \tilde{R}'_t)'$, $t = 1, \dots, T$, so its mean and covariance can be consistently estimated by sample counterparts.

In existing empirical studies, base assets used for constructing mimicking portfolios could differ from test assets used for the FM two-pass procedure.³ Consequently in our setup, R_t does not necessarily coincide with \tilde{R}_t , which is in line with the current practice.

Since mimicking portfolios of macroeconomic factors are commonly constructed by projecting factors on base assets, G_t can be written as (see, e.g., Huberman *et al.* 1987):

$$G_t = V_{FR}V_{RR}^{-1}R_t \quad (2)$$

where $V_{FR}V_{RR}^{-1}$ is unknown but can be consistently estimated. We denote the feasible version

³For convenience, our simulation study in Section 2 sets $R_t = \tilde{R}_t$ though.

of G_t by \hat{G}_t :

$$\hat{G}_t = \hat{V}_{FR} \hat{V}_{RR}^{-1} R_t \quad (3)$$

where \hat{V}_{FR} , \hat{V}_{RR} are the sample counterparts of V_{FR} and V_{RR} respectively.⁴

With the constructed mimicking portfolios \hat{G}_t from R_t , the widely used FM two-pass methodology using \tilde{R}_t as the return on test assets is as follows. Derive the estimator of beta's of mimicking portfolios (denoted by $\hat{\beta}_G$ below) by the first pass time series regression of \tilde{R}_t on \hat{G}_t with an intercept, i.e.,

$$\hat{\beta}_G = \hat{V}_{\tilde{R}R} \hat{V}_{RR}^{-1} \hat{V}_{RF} (\hat{V}_{FR} \hat{V}_{RR}^{-1} \hat{V}_{RF})^{-1} \quad (4)$$

and derive risk premia by the second pass cross-sectional regression of $\hat{\mu}_{\tilde{R}} = \frac{1}{T} \sum_{t=1}^T \tilde{R}_t$ on $\hat{\beta}_G$ with an intercept ($\iota_{\tilde{n}}$ is the $\tilde{n} \times 1$ vector of ones), i.e.,

$$\begin{pmatrix} \hat{\lambda}_1 \\ \hat{\lambda}_G \end{pmatrix} = \left[(\iota_{\tilde{n}} : \hat{\beta}_G)' (\iota_{\tilde{n}} : \hat{\beta}_G) \right]^{-1} (\iota_{\tilde{n}} : \hat{\beta}_G)' \hat{\mu}_{\tilde{R}} \quad (5)$$

where $\hat{\lambda}_1$ is the estimated zero-beta risk premium, $\hat{\lambda}_G$ is the estimated risk premia for k mimicking portfolios.

3.2 Beta

In the lemma below, we provide the limiting behavior of $\hat{\beta}_G$, the beta estimator under mimicking portfolios.

Lemma 1. *Let $\beta = V_{RF} V_{FF}^{-1}$, $\hat{\beta} = \hat{V}_{RF} \hat{V}_{FF}^{-1}$, and $\sqrt{T} \text{vec}(\hat{\beta} - \beta) \xrightarrow{d} \psi_\beta$ from Assumption 1 in the Appendix,⁵ where ψ_β is the $nk \times 1$ normally distributed random vector.*

⁴Throughout the paper, for mean μ and variance V , we use $\hat{\mu}$ and \hat{V} to denote its sample analogy. E.g.: $\hat{\mu}_X = \frac{1}{T} \sum_{t=1}^T X_t$, $\hat{V}_{XY} = \frac{1}{T} \sum_{t=1}^T (X_t - \hat{\mu}_X)(Y_t - \hat{\mu}_Y)'$.

⁵The asymptotic normality of $\hat{\beta}$ holds under mild conditions. See the Appendix for further details.

1. When $\beta = 0$:

$$T^{-1/2} \hat{\beta}_G \xrightarrow{d} V_{\tilde{R}R} V_{RR}^{-1} \Psi_\beta (\Psi_\beta' V_{RR}^{-1} \Psi_\beta)^{-1} V_{FF}^{-1} \quad (6)$$

where $\Psi_\beta = \text{vecinv}_k(\psi_\beta)$.

2. When $\beta = B/\sqrt{T}$, where B is a fixed full rank matrix, the limiting behavior of $T^{-1/2} \hat{\beta}_G$ is characterized by:

$$V_{\tilde{R}R} V_{RR}^{-1} (B + \Psi_\beta) [(B + \Psi_\beta)' V_{RR}^{-1} (B + \Psi_\beta)]^{-1} V_{FF}^{-1} \quad (7)$$

3. When β is a fixed full rank matrix:

$$\hat{\beta}_G \xrightarrow{p} \tilde{\beta}_G \quad (8)$$

where $\tilde{\beta}_G = V_{\tilde{R}G} V_{GG}^{-1} = V_{\tilde{R}R} V_{RR}^{-1} V_{RF} (V_{FR} V_{RR}^{-1} V_{RF})^{-1}$.

Proof. See the Appendix. □

Lemma 1 shows that, when factors are uncorrelated with base assets (case 1: $\beta = 0$) or only minorly correlated with base assets (case 2: $\beta = B/\sqrt{T}$), the beta estimator $\hat{\beta}_G$ using mimicking portfolios of these factors does not converge to a fixed matrix or remain close to zero. Instead, the limiting behavior of $\hat{\beta}_G$ is characterized by a random matrix whose magnitude is $O_p(T^{1/2})$, compared to $o_p(T^{-1/2})$ of β . In contrast, only when β is sizeable so factors are closely correlated with base assets (case 3: β is a fixed full rank matrix), $\hat{\beta}_G$ converges to its estimand $\tilde{\beta}_G$.

It is worth emphasizing that traditional large sample inference in the FM methodology requires that β is a full rank matrix, an assumption that is not realistic when factors are only minorly correlated with base assets. Instead, we adopt $\beta = B/\sqrt{T}$ from the weak-instrument assumption made in econometrics (see, e.g., Staiger and Stock 1997), in order to better model the small magnitude of β for macroeconomic factors. A similar treatment can

be found in Kleibergen (2009) and Kleibergen and Zhan (2014).

For macroeconomic factors whose β 's are small or equal to zero, Lemma 1 suggests that the estimated beta's using their mimicking portfolios, $\hat{\beta}_G$, could be spuriously large. These large values of $\hat{\beta}_G$, however, may not represent the true risk information of macroeconomic factors reflected by β . Consequently, further estimation of risk premia based on the spurious $\hat{\beta}_G$ is expected to be problematic, as we show later.⁶

3.3 Risk Premia

In the theorem below, we continue to provide the limiting distribution of the risk premia estimator in (5) under mimicking portfolios.

Theorem 1. *Under Assumption 1, the large sample behavior of the FM two-pass risk premia estimator in (5) is characterized as follows.*

1. When $\beta = 0$:

$$\begin{aligned} & \begin{pmatrix} \sqrt{T}(\hat{\lambda}_1 - \frac{1}{a}b'\mu_{\tilde{R}}) \\ T\hat{\lambda}_G - \sqrt{T}V_{FF}\Psi'_\beta V_{RR}^{-1}\Psi_\beta(\Psi'_\beta M_{\iota_{\tilde{n}}}\Psi_\beta)^{-1}\Psi'_\beta V_{RR}^{-1}V_{R\tilde{R}}M_{\iota_{\tilde{n}}}\mu_{\tilde{R}} \end{pmatrix} \\ & \xrightarrow{d} \begin{pmatrix} \frac{1}{a}b' \\ V_{FF}\Psi'_\beta V_{RR}^{-1}\Psi_\beta(\Psi'_\beta M_{\iota_{\tilde{n}}}\Psi_\beta)^{-1}\Psi'_\beta V_{RR}^{-1}V_{R\tilde{R}}M_{\iota_{\tilde{n}}} \end{pmatrix} \psi_{\tilde{R}} \end{aligned} \quad (9)$$

where $a = \iota'_{\tilde{n}} M_{V_{\tilde{R}\tilde{R}} V_{RR}^{-1} \Psi_\beta} \iota_{\tilde{n}}$, $b = M_{V_{\tilde{R}\tilde{R}} V_{RR}^{-1} \Psi_\beta} \iota_{\tilde{n}}$.

2. When $\beta = \frac{1}{\sqrt{T}}B$, where B is a fixed full rank matrix:

⁶The non-standard behavior of $\hat{\beta}_G$ in Lemma 1 also explains why it is associated with low p -values of the rank test in Table 1. When factors have zero or small β 's, mimicking portfolios of these factors drive $\hat{\beta}_G$ to a random matrix with larger magnitude. The non-convergence and larger magnitude of $\hat{\beta}_G$ thus jeopardize the rank test which relies on a consistent estimator of the tested matrix.

$$\begin{aligned} & \begin{pmatrix} \sqrt{T}(\hat{\lambda}_1 - \frac{1}{a}b'\mu_{\tilde{R}}) \\ T\hat{\lambda}_G - \sqrt{T}V_{FF}(B + \Psi_\beta)'V_{RR}^{-1}(B + \Psi_\beta)[(B + \Psi_\beta)'M_{\iota_{\tilde{n}}}(B + \Psi_\beta)]^{-1}(B + \Psi_\beta)'V_{RR}^{-1}V_{R\tilde{R}}M_{\iota_{\tilde{n}}}\mu_{\tilde{R}} \end{pmatrix} \\ & \xrightarrow{d} \begin{pmatrix} \frac{1}{a}b' \\ V_{FF}(B + \Psi_\beta)'V_{RR}^{-1}(B + \Psi_\beta)[(B + \Psi_\beta)'M_{\iota_{\tilde{n}}}(B + \Psi_\beta)]^{-1}(B + \Psi_\beta)'V_{RR}^{-1}V_{R\tilde{R}}M_{\iota_{\tilde{n}}} \end{pmatrix} \psi_{\tilde{R}} \quad (10) \end{aligned}$$

where $a = \iota_{\tilde{n}}'M_{V_{\tilde{R}\tilde{R}}V_{RR}^{-1}(B+\Psi_\beta)}\iota_{\tilde{n}}$, $b = M_{V_{\tilde{R}\tilde{R}}V_{RR}^{-1}(B+\Psi_\beta)}\iota_{\tilde{n}}$.

3. When β is a fixed full rank matrix and $(\iota_{\tilde{n}}:\tilde{\beta}_G)$ has full rank:

$$(\hat{\lambda}_1, \hat{\lambda}'_G) \xrightarrow{p} (\lambda_1, \lambda'_G) \quad (11)$$

$$\text{where } (\lambda_1, \lambda'_G)' = \left[(\iota_{\tilde{n}}:\tilde{\beta}_G)'(\iota_{\tilde{n}}:\tilde{\beta}_G) \right]^{-1} (\iota_{\tilde{n}}:\tilde{\beta}_G)'\mu_{\tilde{R}}.$$

Proof. See the Appendix. \square

When macroeconomic factors have zero or small β 's, Theorem 1 shows that the risk premia estimator $\hat{\lambda}_G$ using mimicking portfolios of such factors has a non-standard distribution that differs from the normal distribution. This non-standard distribution thus jeopardizes the conventional t -test on risk premia in the FM procedure, since the validity of t -test relies on normality of the risk premia estimator.

Existing standard error adjustment methods for the risk premia estimator, such as those in Shanken (1992), Jagannathan and Wang (1998) and Kan *et al.* (2013), do not fully resolve the malfunction of the t -test on the risk premia of mimicking portfolios. Although standard errors of risk premia provided by these methods account for estimation error in factor loadings, model misspecification, etc., these methods still rely on the full rank condition of β that is at risk under macroeconomic factors.

Our simulation study in Section 2 has shown that the non-standard distribution of the risk premia estimator in Theorem 1 tends to drive the conventional FM t -statistic to significant values, when mimicking portfolios are constructed from factors that are useless.

Consequently, inference on risk premia using mimicking portfolios and the t -test should be taken with caution.

In the corollary below, we provide the non-standard behavior of the FM t -statistic under the simulation setup in Section 2, which helps explain the malfunction of the t -test for risk premia. In particular, the t -statistic for risk premia of mimicking portfolios tends to be large, even when mimicking portfolios result from completely useless factors.

Corollary 1. *When $\beta = 0$, $R_t = \tilde{R}_t$ and $k = 1$, consider $t = \frac{\hat{\lambda}_G}{\sqrt{\widehat{AVar}(\hat{\lambda}_G)/T}}$, where $\widehat{AVar}(\hat{\lambda}_G)$ is the $(2, 2)$ element of $(H'H)^{-1}H'\hat{\Sigma}H(H'H)^{-1}$, with $H = (\iota_n; \hat{\beta}_G)$, $\hat{\Sigma} = \hat{V}_{RR} - \hat{V}_{RF}(\hat{V}_{FR}\hat{V}_{RR}^{-1}\hat{V}_{RF})^{-1}\hat{V}_{FR}$. As T gets large, the t -statistic goes to infinity in absolute value.*

Proof. See the Appendix. □

3.4 R -squared

Other than risk premia, the cross-sectional OLS R^2 in the FM two-pass procedure is also commonly reported in empirical studies.

With risk factors F_t and test assets \tilde{R}_t , the expression of the OLS R^2 reads

$$\begin{aligned} R_F^2 &= \frac{\hat{\mu}'_{\tilde{R}} P_{M_{\iota_{\tilde{n}}}\hat{\beta}} \hat{\mu}_{\tilde{R}}}{\hat{\mu}'_{\tilde{R}} M_{\iota_{\tilde{n}}} \hat{\mu}_{\tilde{R}}} \\ &= \frac{\hat{\mu}'_{\tilde{R}} M_{\iota_{\tilde{n}}} \hat{\beta} (\hat{\beta}' M_{\iota_{\tilde{n}}} \hat{\beta})^{-1} \hat{\beta}' M_{\iota_{\tilde{n}}} \hat{\mu}_{\tilde{R}}}{\hat{\mu}'_{\tilde{R}} M_{\iota_{\tilde{n}}} \hat{\mu}_{\tilde{R}}}, \end{aligned} \quad (12)$$

where $\hat{\beta} = \hat{V}_{\tilde{R}F} \hat{V}_{FF}^{-1}$. Note that $\hat{\beta}$ differs from $\hat{\beta} = \hat{V}_{RF} \hat{V}_{FF}^{-1}$ used above, unless $R_t = \tilde{R}_t$.

Similarly, with mimicking portfolios \hat{G}_t and \tilde{R}_t , the expression of the OLS R^2 reads

$$R_G^2 = \frac{\hat{\mu}'_{\tilde{R}} M_{\iota_{\tilde{n}}} \hat{\beta}_G (\hat{\beta}_G' M_{\iota_{\tilde{n}}} \hat{\beta}_G)^{-1} \hat{\beta}_G' M_{\iota_{\tilde{n}}} \hat{\mu}_{\tilde{R}}}{\hat{\mu}'_{\tilde{R}} M_{\iota_{\tilde{n}}} \hat{\mu}_{\tilde{R}}}, \quad (13)$$

where $\hat{\beta}_G = \hat{V}_{\tilde{R}R} \hat{V}_{RR}^{-1} \hat{V}_{RF} (\hat{V}_{FR} \hat{V}_{RR}^{-1} \hat{V}_{RF})^{-1}$.

3.4.1 $R_t = \tilde{R}_t$

When base assets are used as test assets, an observation from Table 1 and 2 is that the cross-sectional OLS R^2 under mimicking portfolios coincides with its counterpart under factors. The simple derivation below shows that this coincidence is not accidental.

When $\tilde{R}_t = R_t$, $\hat{\beta}_G = \hat{V}_{RF}(\hat{V}_{FR}\hat{V}_{RR}^{-1}\hat{V}_{RF})^{-1} = \hat{\beta}[\hat{V}_{FF}(\hat{V}_{FR}\hat{V}_{RR}^{-1}\hat{V}_{RF})^{-1}]$, so $\hat{\beta}_G$ equals $\hat{\beta}$ scaled by an invertible matrix. Using this fact in (13), it is straightforward to see that we have the numerical equivalence of the two R -squared's:

$$R_F^2 = R_G^2 \quad (14)$$

The equivalence of R_F^2 and R_G^2 implies that, same as the R -squared under factors, large values of the R -squared under mimicking portfolios also need to be interpreted with caution. Lewellen *et al.* (2010) warn that in various scenarios, R -squared under factors could be spuriously large. In addition, Kleibergen and Zhan (2014) provide the distribution of the R -squared under useless and nearly useless factors, which can take large values with a sizeable probability mass. These existing critics of R -squared under factors thus also apply to the R -squared under mimicking portfolios, given their equivalence in (14).

3.4.2 $R_t \neq \tilde{R}_t$

More generally, when base assets and test assets differ, we provide the distribution of R_G^2 in the theorem below. In particular, we focus on the scenario that the background macroeconomic factors for mimicking portfolios are only minorly correlated with base assets.

Theorem 2. *When $\beta = B/\sqrt{T}$, the behavior of R_G^2 in (13) is in large samples characterized by*

$$\frac{(\mu_{\tilde{R}} + \frac{1}{\sqrt{T}}\psi_{\tilde{R}})' P_{M_{\tilde{R}}} [V_{\tilde{R}\tilde{R}} V_{\tilde{R}\tilde{R}}^{-1} (B + \Psi_{\beta})] (\mu_{\tilde{R}} + \frac{1}{\sqrt{T}}\psi_{\tilde{R}})}{(\mu_{\tilde{R}} + \frac{1}{\sqrt{T}}\psi_{\tilde{R}})' M_{\tilde{R}} (\mu_{\tilde{R}} + \frac{1}{\sqrt{T}}\psi_{\tilde{R}})}, \quad (15)$$

where $\sqrt{T}(\hat{\mu}_{\tilde{R}} - \mu_{\tilde{R}}) \xrightarrow{d} \psi_{\tilde{R}}$, $\sqrt{T}vec(\hat{\beta} - \beta) \xrightarrow{d} \psi_{\beta}$, $\Psi_{\beta} = vecinv_k(\psi_{\beta})$, $\psi_{\tilde{R}}$ and ψ_{β} are $\tilde{n} \times 1$,

$nk \times 1$ dimensional independent random vectors whose elements are normally distributed.

Proof. See the Appendix. □

The large sample behavior of R_G^2 in Theorem 2 is similar to the large sample result for R_F^2 provided by Kleibergen and Zhan (2014).⁷ In fact, when $R_t = \tilde{R}_t$, Theorem 2 coincides with the result for R_F^2 in Kleibergen and Zhan (2014). Hence similar to R_F^2 , large values of R_G^2 could result from mimicking portfolios of macroeconomic factors which are in fact useless or nearly useless. As a result, these large R_G^2 's should not be used as evidence to support the corresponding factors or mimicking portfolios.

4 Further Discussion

When mimicking portfolios result from macroeconomic factors with small β 's, our analysis above shows that estimating risk premia of mimicking portfolios by the FM two-pass procedure is unreliable. To remedy this problem, we discuss robust inference methods on risk premia of mimicking portfolios in this section.

4.1 Specification: Beta or Covariance

With \tilde{R}_t for test assets and G_t for mimicking portfolios, the beta of mimicking portfolios is commonly defined as $\tilde{\beta}_G = V_{\tilde{R}G} V_{GG}^{-1}$. However, when $G_t = V_{FR} V_{RR}^{-1} R_t$ so mimicking portfolios are constructed by projecting factors on base assets, one subtle issue related to $\tilde{\beta}_G$ arises.

The subtle problem is that, under useless factors with $\beta = 0$, $\tilde{\beta}_G$ becomes undefined. This is because of $V_{GG} = \text{var}(G_t) = V_{FR} V_{RR}^{-1} V_{RF} = 0$ under useless factors. A similar problem arises when factors have small β 's. When β is small, $\tilde{\beta}_G$ can become so large that its estimator $\hat{\tilde{\beta}}_G$ is spurious, as shown in Lemma 1. The spurious $\hat{\tilde{\beta}}_G$ further induces the malfunction of the risk premia estimation in the second pass of the FM procedure, as shown in Theorem 1.

⁷See Theorem 3 in Kleibergen and Zhan (2014).

For the reason above, it is inappropriate to use the beta specification for the two-pass tests with mimicking portfolios. Instead, the alternative specification (see, e.g., Kan *et al.* 2013) that uses the covariance instead of beta turns out to be more feasible.

Consider the covariance specification for the second pass cross-sectional regression:

$$\mu_{\tilde{R}} = \iota_{\tilde{n}}\lambda_1 + V_{\tilde{R}G}\lambda_G \quad (16)$$

where $V_{\tilde{R}G} = V_{\tilde{R}R}V_{RR}^{-1}V_{RF}$, given $G_t = V_{FR}V_{RR}^{-1}R_t$. Unlike $\tilde{\beta}_G$, the covariance $V_{\tilde{R}G}$ is well defined, not matter whether the factors have zero β 's or not. If β of factors is zero, then $V_{\tilde{R}G}$ is also zero, while $\tilde{\beta}_G$ is not defined. Put differently, using the covariance specification in (16) instead of the beta specification allows us to specify meaningful risk premia λ_G for mimicking portfolios, particularly when their background factors are associated with small β 's.

Note that when base assets and tests assets coincide so $R_t = \tilde{R}_t$, $V_{\tilde{R}G}$ reduces to V_{RF} , and the covariance specification for mimicking portfolios in (16) is also the covariance specification for factors:

$$\mu_R = \iota_n\lambda_1 + V_{RF}\lambda_G \quad (17)$$

This further implies $\lambda_G = V_{FF}^{-1}\lambda_F$, where λ_F denotes risk premia of factors in the standard β specification $\mu_R = \iota_n\lambda_1 + \beta\lambda_F$.

4.2 Robust Inference

Our interest now is to conduct inference on λ_G , the risk premia of mimicking portfolios in the covariance specification of (16). Note that when factors are only minorly correlated with assets, the FM two-pass methodology similarly fails for (16).⁸ Although estimation of risk premia is unreliable when β 's are small, we could instead test $H_0 : \lambda_G = \lambda_{G,0}$.

⁸The reason is the same as the one in Kleibergen (2009), i.e., $V_{\tilde{R}G}$ is small when factors are weak, just like β is small when factors are weak.

For easier exposition, we focus on a simplified setup with $R_t = \tilde{R}_t$, and omit the $\iota_{\tilde{n}}\lambda_1$ term in (16).⁹ In this simplified setup, (16) as well as (17) reduces to:

$$\mu_R = V_{RF}\lambda_G \quad (18)$$

which implies that the linear factor model (without zero-beta return) reads:

$$R_t = V_{RF}[V_{FF}^{-1}(F_t - \mu_F) + \lambda_G] + u_t \quad (19)$$

where u_t is the error term.

Under $H_0 : \lambda_G = \lambda_{G,0}$, consider the least squares estimator of V_{RF} :

$$\hat{V}_{RF,r} = \sum_{t=1}^T R_t(\hat{V}_{FF}^{-1}\bar{F}_t + \lambda_{G,0})' \left[\sum_{t=1}^T (\hat{V}_{FF}^{-1}\bar{F}_t + \lambda_{G,0})(\hat{V}_{FF}^{-1}\bar{F}_t + \lambda_{G,0})' \right]^{-1} \quad (20)$$

where $\bar{F}_t = F_t - \hat{\mu}_F$.

Theorem 3. Under $H_0 : \lambda_G = \lambda_{G,0}$ and Assumption 2 in the Appendix,

1.

$$\sqrt{T} \begin{pmatrix} \hat{\mu}_R - \hat{V}_{RF,r}\lambda_{G,0} \\ \text{vec}(\hat{V}_{RF,r} - V_{RF}) \end{pmatrix} \xrightarrow{d} \begin{pmatrix} \tilde{\psi}_R \\ \tilde{\psi}_\beta \end{pmatrix} \quad (21)$$

where $\tilde{\psi}_R$ and $\tilde{\psi}_\beta$ are $n \times 1$ and $nk \times 1$ independently and normally distributed random vectors with mean 0 and covariance matrices $[1 - \lambda'_{G,0}(V_{FF}^{-1} + \lambda_{G,0}\lambda'_{G,0})^{-1}\lambda_{G,0}] \otimes \tilde{\Omega}$ and $(V_{FF}^{-1} + \lambda_{G,0}\lambda'_{G,0}) \otimes \tilde{\Omega}$, $\tilde{\Omega} = \text{var}(\tilde{\epsilon}_t) : n \times n$, $\tilde{\epsilon}_t = u_t + V_{RF}V_{FF}^{-1}\bar{v} + V_{RF}(V_{FF}^{-1} - \hat{V}_{FF}^{-1})\bar{F}_t$, $\bar{v} = \frac{1}{T} \sum_{t=1}^T (F_t - \mu_F)$.

2. The result in (21) implies the factor Anderson and Rubin (1949) statistic, denoted by

⁹This can be achieved by multiplying $M_{\iota_{\tilde{n}}}$ on both sides of (16). See also Kleibergen (2009). The general setup with $R_t \neq \tilde{R}_t$ can be analyzed similarly.

FAR below, is asymptotically chi-squared distributed, i.e.,

$$FAR(\lambda_{G,0}) = \frac{T}{1 - \lambda'_{G,0}(\hat{V}_{FF}^{-1} + \lambda_{G,0}\lambda'_{G,0})^{-1}\lambda_{G,0}} (\hat{\mu}_R - \hat{V}_{RF,r}\lambda_{G,0})' \hat{\hat{\Omega}}^{-1} (\hat{\mu}_R - \hat{V}_{RF,r}\lambda_{G,0})$$

$$\xrightarrow{d} \chi_n^2 \quad (22)$$

where $\hat{\hat{\Omega}} = \frac{1}{T-k} \sum_{t=1}^T [\bar{R}_t - \hat{V}_{RF,r}(\hat{V}_{FF}^{-1}\bar{F}_t + \lambda_{G,0})][\bar{R}_t - \hat{V}_{RF,r}(\hat{V}_{FF}^{-1}\bar{F}_t + \lambda_{G,0})]'$, with $\bar{R}_t = R_t - \hat{\mu}_R$.

Proof. See the Appendix. □

Theorem 3 shows that the asymptotic distribution of the FAR statistic does not depend on the magnitude of β . As a result, unlike the FM t -test, the FAR test for testing $H_0 : \lambda_G = \lambda_{G,0}$ has correct size, no matter β is small or sizeable. A confidence set of λ_G can thus be constructed by inverting the FAR test. Whether mimicking portfolios contain pricing information can be inferred by constructing such a confidence set.

It is worth emphasizing that the FAR statistic in (22) is a modified version of the FAR statistic in Kleibergen (2009). In particular, Kleibergen (2009) provides robust tests (including FAR, the Lagrange multiplier (LM) test, the conditional likelihood ratio (CLR) test, etc.) for testing λ_F , risk premia of factors, while this paper intends to test λ_G , risk premia of mimicking portfolios. Nevertheless, when $\lambda_G = V_{FF}^{-1}\lambda_F$, modifying the tests for λ_F so that they can be used for λ_G is straightforward. Theorem 3 provides an example for modifying the FAR test. Other robust test statistics of Kleibergen (2009), such as LM and CLR, can be updated in a similar manner.

4.3 Empirical Examples

In this section, we further illustrate the discussion above by employing two risk factors as specific examples: the consumption growth “ Δc ” for the conventional consumption Capital Asset Pricing Model, and the leverage factor “ lev ” recently suggested by Muir *et al.* (2013).

In our empirical application, we use the same data set for consumption growth as in Lettau and Ludvigson (2001), and the sample period is 1963Q4-1998Q3, so $T = 140$. Similarly, we use the same data for the leverage factor as in Muir *et al.* (2013), and the sample period is 1968Q1-2009Q4, so $T = 168$.

For base assets, we follow Muir *et al.* (2013) to consider seven assets: the excess returns of the six Fama-French portfolios on size and book-to-market (“*SL*”, “*SM*”, “*SH*”, “*BL*”, “*BM*”, “*BH*”), plus the momentum factor “*Mom*”. These seven assets are widely acknowledged for their ability to span the asset space. The data for these assets is available at French’s online data library.¹⁰

To construct mimicking portfolios, we project consumption growth and the leverage factor on the seven assets described above. The estimated coefficients $\widehat{V}_{FR}\widehat{V}_{RR}^{-1}$, together with the associated t -statistics, in this projection are presented in Table 3. As indicated by Table 3, these two factors appear only minorly correlated with base assets, since most of the coefficients are not statistically significant. For comparison purposes, we also report the average coefficients of mimicking portfolios and the frequency of $|t| > 1.96$ corresponding to the useless factor described in Section 2.

With the constructed mimicking portfolios, we proceed to compute their beta’s and risk premia. In addition, we also derive the beta’s and risk premia for the original factors, to facilitate the comparison.

For beta’s, Table 4 shows that the estimated β_G (under $R_t = \tilde{R}_t$, $\beta_G = \tilde{\beta}_G$) is much larger than β for the two factors Δc and lev , and the estimator of β_G is highly significant. This is consistent with the current understanding that mimicking portfolios of macroeconomic factors appear more relevant for asset pricing, compared to their background factors.

In regard of risk premia, we use three different test assets: besides the base assets of seven portfolios, we also consider the twenty-five Fama-French portfolios on size and book-

¹⁰Monthly data is available at “http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html”, which is then compounded to construct the quarterly data. Excess returns result from the difference of raw returns and risk-free returns.

Table 3: Weights of Mimicking Portfolios with Seven Assets

	Δc		lev		$useless$	
	Coef.	t -stat	Coef.	t -stat	Coef.	$ t > 1.96$
SL	-0.02	-1.75	-0.57	-2.32	0.00	0.053
SM	0.02	0.73	1.24	2.57	0.00	0.055
SH	0.01	0.60	-0.43	-1.14	0.00	0.056
BL	0.03	2.40	-0.22	-0.74	0.00	0.052
BM	-0.02	-1.28	-0.10	-0.26	0.00	0.048
BH	-0.01	-0.36	0.56	1.73	0.00	0.058
Mom	-0.01	-0.90	0.43	2.81	0.00	0.053

Note: Estimation of mimicking portfolios is implemented by the time series regression of factors (Δc , lev , $useless$) on the seven assets: six Fama-French portfolios on size and book-to-market (SL : small and low; SM : small and medium; SH : small and high; BL : big and low; BM : big and medium; BH : big and high), and the momentum factor (Mom). The $useless$ factor described in Section 2 is used for comparison. Its coefficients and frequency of $|t| > 1.96$ result from the average of 2000 Monte Carlo replications.

to-market, as well as these twenty-five portfolios augmented by thirty industry portfolios suggested by Lewellen *et al.* (2010). Table 4 shows that under mimicking portfolios, t -statistics on risk premia are more likely to be significant. This is consistent with the view that inference on risk premia could be improved by using mimicking portfolios rather than macroeconomic factors.

However, the last three rows in Table 4 also show that for the useless factor, its mimicking portfolio also likely induces significant risk premia. The underlying reason, as indicated by the comparison of β and β_G in Table 4, is due to the spurious amplification of β , i.e., although the true β for the useless factor is zero, the estimated beta of mimicking portfolios can be spuriously large. This large beta further spuriously induces seemingly informative risk premia, as we discussed in Section 3.

5 Conclusion

This paper reveals the downside of using mimicking portfolios of macroeconomic factors in the Fama and MacBeth (1973) two-pass procedure. We show that, when these factors have small

Table 4: Beta, Risk Premia of Factors and Mimicking Portfolios

	Δc				lev			
	β		β_G		β		β_G	
	Coef.	t -stat	Coef.	t -stat	Coef.	t -stat	Coef.	t -stat
	λ_F		λ_G		λ_F		λ_G	
<i>SL</i>	5.02	2.08	55.31	8.31	0.07	0.89	0.54	2.45
<i>SM</i>	4.83	2.48	53.26	11.01	0.13	2.03	0.98	6.05
<i>SH</i>	4.97	2.52	54.75	11.34	0.13	1.84	1.01	5.38
<i>BL</i>	4.02	2.56	44.26	11.63	0.06	1.07	0.43	2.99
<i>BM</i>	2.59	1.99	28.53	7.82	0.10	2.11	0.73	6.32
<i>BH</i>	3.34	2.36	36.83	10.08	0.12	2.33	0.93	7.15
<i>Mom</i>	-1.47	-1.33	-16.21	-4.72	0.05	1.13	0.40	3.17
6 FF + <i>Mom</i>	0.02	0.13	0.00	0.18	11.57	0.90	1.55	1.18
25 FF	0.25	0.97	0.03	1.12	13.84	2.36	2.36	3.25
25 FF + 30 Ind.	0.16	1.17	0.03	1.67	5.79	1.67	1.20	2.07
	<i>useless</i>							
	β		β_G					
	Coef.	$ t > 1.96$	Coef.	$ t > 1.96$				
	λ_F		λ_G					
<i>SL</i>	-0.01	0.045	0.18	0.684				
<i>SM</i>	-0.01	0.044	0.20	0.688				
<i>SH</i>	0.00	0.048	0.29	0.681				
<i>BL</i>	0.00	0.054	0.20	0.694				
<i>BM</i>	-0.01	0.056	0.07	0.688				
<i>BH</i>	0.00	0.049	0.21	0.683				
<i>Mom</i>	0.01	0.045	0.05	0.715				
6 FF + <i>Mom</i>	-0.01	0.004	0.00	0.147				
25 FF	-0.01	0.078	0.00	0.439				
25 FF + 30 Ind.	-0.01	0.042	0.00	0.226				

Note: The tests assets include: six Fama-French portfolios on size and book-to-market (*SL*: small and low; *SM*: small and medium; *SH*: small and high; *BL*: big and low; *BM*: big and medium; *BH*: big and high), and the momentum factor (*Mom*); twenty-five Fama-French portfolios on size and book-to-market; twenty-five Fama-French portfolios on size and book-to-market augmented by thirty industry portfolios.

beta's, their mimicking portfolios are associated with beta's that are spuriously large. These spurious beta's of mimicking portfolios thus do not necessarily reflect the risk information contained in the macroeconomic factors, and they further induce a non-standard behavior of the risk premia estimator. Conventional statistics used in the Fama and MacBeth (1973) procedure such as the t -statistic on risk premia and the cross-sectional R^2 may incorrectly favor macroeconomic factors even when they are useless, since both statistics can be spuriously large. To remedy the inference problem on risk premia of mimicking portfolios, we suggest extending the robust tests in Kleibergen (2009) for factors to mimicking portfolios.

Appendix

A: Preliminary

We first state one preliminary result that will be used later for the proof in the appendix.

Consider the stationary and ergodic vector below:

$$\begin{bmatrix} F_t \\ R_t \\ \tilde{R}_t \end{bmatrix}, \quad t = 1, \dots, T,$$

where F_t is the $k \times 1$ vector of risk factors, R_t is the $n \times 1$ vector of the returns on base assets that are assumed to span the return space, and \tilde{R}_t is the $\tilde{n} \times 1$ vector of the returns on test assets that is allowed to be different from R_t .

Let $\beta = V_{RF}V_{FF}^{-1}$ and $\mu_R, \mu_{\tilde{R}}$ be the mean of R_t and \tilde{R}_t . Denote their estimators by $\hat{\beta} = \hat{V}_{RF}\hat{V}_{FF}^{-1}$, $\hat{\mu}_R = \frac{1}{T} \sum_{t=1}^T R_t$, $\hat{\mu}_{\tilde{R}} = \frac{1}{T} \sum_{t=1}^T \tilde{R}_t$, respectively. Kleibergen (2009) (see also Shanken 1992) shows the asymptotic normality and independence of $\hat{\mu}_R$ and $\hat{\beta}$:

$$\sqrt{T} \begin{pmatrix} \hat{\mu}_R - \mu_R \\ \text{vec}(\hat{\beta} - \beta) \end{pmatrix} \xrightarrow{d} \begin{pmatrix} \psi_R \\ \psi_\beta \end{pmatrix}$$

where ψ_R, ψ_β are $n \times 1, nk \times 1$ normally distributed random vectors, and ψ_R is independent of ψ_β .

For our purpose, the asymptotic result above is extended to the following:

$$\sqrt{T} \begin{pmatrix} \hat{\mu}_{\tilde{R}} - \mu_{\tilde{R}} \\ \text{vec}(\hat{\beta} - \beta) \end{pmatrix} \xrightarrow{d} \begin{pmatrix} \psi_{\tilde{R}} \\ \psi_\beta \end{pmatrix}$$

where $\psi_{\tilde{R}}$ is the $\tilde{n} \times 1$ random vector, which is also normally distributed from the central limit theorem, and independent of ψ_β .

The independence of $\psi_{\tilde{R}}$ and ψ_{β} can be shown by the projection:

$$\tilde{R}_t - \mu_{\tilde{R}} = V_{\tilde{R}R} V_{RR}^{-1} (R_t - \mu_R) + e_t$$

where e_t is the projection error that is uncorrelated with R_t . This projection equation further implies

$$\sqrt{T}(\hat{\mu}_{\tilde{R}} - \mu_{\tilde{R}}) = V_{\tilde{R}R} V_{RR}^{-1} \sqrt{T}(\hat{\mu}_R - \mu_R) + \frac{1}{\sqrt{T}} \sum_{t=1}^T e_t$$

where $\sqrt{T}(\hat{\mu}_R - \mu_R)$ is asymptotically normal and independent of ψ_{β} by Kleibergen (2009); while $\frac{1}{\sqrt{T}} \sum_{t=1}^T e_t$ is also asymptotically normal by the central limit theorem and uncorrelated with ψ_{β} , since e_t is orthogonal to the return space. This implies the asymptotic independence of $\psi_{\tilde{R}}$ and ψ_{β} .

Instead of making upper-level assumptions that imply the asymptotic distribution for $\hat{\mu}_{\tilde{R}}$ and $\hat{\beta}$, we simply assume:

Assumption 1. For the stationary and ergodic vector $(F'_t, R'_t, \tilde{R}'_t)'$, with $\beta = V_{RF} V_{FF}^{-1}$, $\hat{\beta} = \hat{V}_{RF} \hat{V}_{FF}^{-1}$, $\mu_{\tilde{R}} = \mathbb{E}(\tilde{R}_t)$, $\hat{\mu}_{\tilde{R}} = \frac{1}{T} \sum_{t=1}^T \tilde{R}_t$:

$$\sqrt{T} \begin{pmatrix} \hat{\mu}_{\tilde{R}} - \mu_{\tilde{R}} \\ \text{vec}(\hat{\beta} - \beta) \end{pmatrix} \xrightarrow{d} \begin{pmatrix} \psi_{\tilde{R}} \\ \psi_{\beta} \end{pmatrix}$$

where $\psi_{\tilde{R}}$, ψ_{β} are $\tilde{n} \times 1$, $nk \times 1$ normally distributed random vectors with zero mean, and $\psi_{\tilde{R}}$ is independent of ψ_{β} . The variance of $\psi_{\tilde{R}}$ is $V_{\tilde{R}\tilde{R}}$, while the variance of ψ_{β} is provided in, e.g., Kleibergen (2009).

B: Proof for Lemma 1

Proof. With $\hat{\beta} = \hat{V}_{RF} \hat{V}_{FF}^{-1}$, $\hat{\beta}_G$ can be rewritten as

$$\hat{\beta}_G = \hat{V}_{\tilde{R}R} \hat{V}_{RR}^{-1} \hat{\beta} \hat{V}_{FF} (\hat{V}_{FF} \hat{\beta}' \hat{V}_{RR}^{-1} \hat{\beta} \hat{V}_{FF})^{-1} = \hat{V}_{\tilde{R}R} \hat{V}_{RR}^{-1} \hat{\beta} (\hat{\beta}' \hat{V}_{RR}^{-1} \hat{\beta})^{-1} \hat{V}_{FF}^{-1}$$

and

$$T^{-1/2}\hat{\hat{\beta}}_G = \hat{V}_{\tilde{R}R}\hat{V}_{RR}^{-1}\sqrt{T}\hat{\beta}(\sqrt{T}\hat{\beta}'\hat{V}_{RR}^{-1}\sqrt{T}\hat{\beta})^{-1}\hat{V}_{FF}^{-1}$$

where $\hat{V}_{\tilde{R}R} \xrightarrow{p} V_{\tilde{R}R}$, $\hat{V}_{RR} \xrightarrow{p} V_{RR}$, $\hat{V}_{FF} \xrightarrow{p} V_{FF}$.

1. Given $\sqrt{T}vec(\hat{\beta} - \beta) \xrightarrow{d} \psi_\beta$, if $\beta = 0$, $\sqrt{T}\hat{\beta} \xrightarrow{d} \Psi_\beta = vecinv_k(\psi_\beta)$. By continuous mapping:

$$T^{-1/2}\hat{\hat{\beta}}_G \xrightarrow{d} V_{\tilde{R}R}V_{RR}^{-1}\Psi_\beta(\Psi_\beta'V_{RR}^{-1}\Psi_\beta)^{-1}V_{FF}^{-1}$$

2. We rephrase $\sqrt{T}vec(\hat{\beta} - \beta) \xrightarrow{d} \psi_\beta$ as:

$$\hat{\beta} = \beta + \frac{\Psi_\beta}{\sqrt{T}} + o_p(T^{-1/2})$$

When $\beta = B/\sqrt{T}$, $\sqrt{T}\hat{\beta}$ is characterized by $B + \Psi_\beta$, so $T^{-1/2}\hat{\hat{\beta}}_G$ is characterized by:

$$V_{\tilde{R}R}V_{RR}^{-1}(B + \Psi_\beta) [(B + \Psi_\beta)'V_{RR}^{-1}(B + \Psi_\beta)]^{-1}V_{FF}^{-1}$$

3. When β is a fixed full rank matrix:

$$\hat{\hat{\beta}}_G = \hat{V}_{\tilde{R}R}\hat{V}_{RR}^{-1}\hat{V}_{RF}(\hat{V}_{FR}\hat{V}_{RR}^{-1}\hat{V}_{RF})^{-1} \xrightarrow{p} V_{\tilde{R}R}V_{RR}^{-1}V_{RF}(V_{FR}V_{RR}^{-1}V_{RF})^{-1} = V_{\tilde{R}G}V_{GG}^{-1}$$

since $V_{\tilde{R}G} = V_{\tilde{R}R}V_{RR}^{-1}V_{RF}$, $V_{GG} = V_{FR}V_{RR}^{-1}V_{RF}$ under $G_t = V_{FR}V_{RR}^{-1}R_t$.

□

C: Proof for Theorem 1

Proof.

$$\begin{pmatrix} \hat{\lambda}_1 \\ \hat{\lambda}_G \end{pmatrix} = \left[(\iota_{\tilde{n}}'\hat{\hat{\beta}}_G)(\iota_{\tilde{n}}'\hat{\hat{\beta}}_G) \right]^{-1} (\iota_{\tilde{n}}'\hat{\hat{\beta}}_G)' \hat{\mu}_{\tilde{R}} = \begin{pmatrix} (\iota_{\tilde{n}}'M_{\hat{\hat{\beta}}_G}\iota_{\tilde{n}})^{-1}\iota_{\tilde{n}}'M_{\hat{\hat{\beta}}_G}\hat{\mu}_{\tilde{R}} \\ (\hat{\hat{\beta}}_G'M_{\iota_{\tilde{n}}}\hat{\hat{\beta}}_G)^{-1}\hat{\hat{\beta}}_G'M_{\iota_{\tilde{n}}}\hat{\mu}_{\tilde{R}} \end{pmatrix}$$

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

Further more, since $\hat{\beta}_G = \hat{V}_{RR} \hat{V}_{RR}^{-1} \hat{\beta} (\hat{\beta}' \hat{V}_{RR}^{-1} \hat{\beta})^{-1} \hat{V}_{FF}^{-1}$, we have $M_{\hat{\beta}_G} = M_{\hat{V}_{RR} \hat{V}_{RR}^{-1} \hat{\beta}}$, where $\hat{\beta} = \beta + \frac{\Psi_\beta}{\sqrt{T}} + o_p(1/\sqrt{T})$.

1. If $\beta = 0$, $T^{-1/2} \hat{\beta}_G \xrightarrow{d} V_{RR} V_{RR}^{-1} \Psi_\beta (\Psi'_\beta V_{RR}^{-1} \Psi_\beta)^{-1} V_{FF}^{-1}$ from Lemma 1.

$$\begin{aligned} \begin{pmatrix} \hat{\lambda}_1 \\ \hat{\lambda}_G \end{pmatrix} &= \begin{pmatrix} (\iota'_{\tilde{n}} M_{\hat{\beta}_G} \iota_{\tilde{n}})^{-1} \iota'_{\tilde{n}} M_{\hat{\beta}_G} \hat{\mu}_{\tilde{R}} \\ (\hat{\beta}'_G M_{\iota_{\tilde{n}}} \hat{\beta}_G)^{-1} \hat{\beta}'_G M_{\iota_{\tilde{n}}} \hat{\mu}_{\tilde{R}} \end{pmatrix} \\ &\simeq \begin{pmatrix} (\iota'_{\tilde{n}} M_{V_{RR} V_{RR}^{-1} \Psi_\beta} \iota_{\tilde{n}})^{-1} \iota'_{\tilde{n}} M_{V_{RR} V_{RR}^{-1} \Psi_\beta} (\mu_{\tilde{R}} + \frac{\psi_{\tilde{R}}}{\sqrt{T}}) \\ \frac{1}{\sqrt{T}} V_{FF} \Psi'_\beta V_{RR}^{-1} \Psi_\beta (\Psi'_\beta M_{\iota_{\tilde{n}}} \Psi_\beta)^{-1} \Psi'_\beta V_{RR}^{-1} V_{R\tilde{R}} M_{\iota_{\tilde{n}}} (\mu_{\tilde{R}} + \frac{\psi_{\tilde{R}}}{\sqrt{T}}) \end{pmatrix} \end{aligned}$$

which implies

$$\begin{aligned} &\begin{pmatrix} \sqrt{T}(\hat{\lambda}_1 - \frac{1}{a} b' \mu_{\tilde{R}}) \\ T \hat{\lambda}_G - \sqrt{T} V_{FF} \Psi'_\beta V_{RR}^{-1} \Psi_\beta (\Psi'_\beta M_{\iota_{\tilde{n}}} \Psi_\beta)^{-1} \Psi'_\beta V_{RR}^{-1} V_{R\tilde{R}} M_{\iota_{\tilde{n}}} \mu_{\tilde{R}} \end{pmatrix} \\ &\xrightarrow{d} \begin{pmatrix} \frac{1}{a} b' \\ V_{FF} \Psi'_\beta V_{RR}^{-1} \Psi_\beta (\Psi'_\beta M_{\iota_{\tilde{n}}} \Psi_\beta)^{-1} \Psi'_\beta V_{RR}^{-1} V_{R\tilde{R}} M_{\iota_{\tilde{n}}} \end{pmatrix} \psi_{\tilde{R}} \end{aligned}$$

where $a = \iota'_{\tilde{n}} M_{V_{RR} V_{RR}^{-1} \Psi_\beta} \iota_{\tilde{n}}$, $b = M_{V_{RR} V_{RR}^{-1} \Psi_\beta} \iota_{\tilde{n}}$.

2. If $\beta = B/\sqrt{T}$, by the derivation similar to above, we get:

$$\begin{aligned} &\begin{pmatrix} \sqrt{T}(\hat{\lambda}_1 - \frac{1}{a} b' \mu_{\tilde{R}}) \\ T \hat{\lambda}_G - \sqrt{T} V_{FF} (B + \Psi_\beta)' V_{RR}^{-1} (B + \Psi_\beta) [(B + \Psi_\beta)' M_{\iota_{\tilde{n}}} (B + \Psi_\beta)]^{-1} (B + \Psi_\beta)' V_{RR}^{-1} V_{R\tilde{R}} M_{\iota_{\tilde{n}}} \mu_{\tilde{R}} \end{pmatrix} \\ &\xrightarrow{d} \begin{pmatrix} \frac{1}{a} b' \\ V_{FF} (B + \Psi_\beta)' V_{RR}^{-1} (B + \Psi_\beta) [(B + \Psi_\beta)' M_{\iota_{\tilde{n}}} (B + \Psi_\beta)]^{-1} (B + \Psi_\beta)' V_{RR}^{-1} V_{R\tilde{R}} M_{\iota_{\tilde{n}}} \end{pmatrix} \psi_{\tilde{R}} \end{aligned}$$

where $a = \iota'_{\tilde{n}} M_{V_{RR} V_{RR}^{-1} (B + \Psi_\beta)} \iota_{\tilde{n}}$, $b = M_{V_{RR} V_{RR}^{-1} (B + \Psi_\beta)} \iota_{\tilde{n}}$.

3. When β is a fixed full rank matrix, $\hat{\beta}_G \xrightarrow{p} \tilde{\beta}_G$ from Lemma 1, $\hat{\mu}_{\tilde{R}} \xrightarrow{p} \mu_{\tilde{R}}$, and:

$$\begin{aligned} \begin{pmatrix} \hat{\lambda}_1 \\ \hat{\lambda}_G \end{pmatrix} &= \left[(\iota_{\tilde{n}} \dot{\hat{\beta}}_G)' (\iota_{\tilde{n}} \dot{\hat{\beta}}_G) \right]^{-1} (\iota_{\tilde{n}} \dot{\hat{\beta}}_G)' \hat{\mu}_{\tilde{R}} \\ &\xrightarrow{p} \left[(\iota_{\tilde{n}} \dot{\tilde{\beta}}_G)' (\iota_{\tilde{n}} \dot{\tilde{\beta}}_G) \right]^{-1} (\iota_{\tilde{n}} \dot{\tilde{\beta}}_G)' \mu_{\tilde{R}} \\ &\equiv \begin{pmatrix} \lambda_1 \\ \lambda_G \end{pmatrix} \end{aligned}$$

given $(\iota_{\tilde{n}} \dot{\tilde{\beta}}_G)$ has full rank.

□

D: Proof for Corollary 1

Proof. When $\beta = 0$, $R_t = \tilde{R}_t$, Theorem 1(1) implies that $\sqrt{T}\hat{\lambda}_G$ is characterized by

$$V_{FF}\Psi'_\beta V_{RR}^{-1}\Psi_\beta(\Psi'_\beta M_{\iota_n}\Psi_\beta)^{-1}\Psi'_\beta M_{\iota_n}\mu_{\tilde{R}}$$

To show t -statistic goes to infinity in absolute value, it now suffices to show $\widehat{AVar}(\hat{\lambda}_G)$ goes to zero as T gets large.

Note that $\hat{\Sigma} = \hat{V}_{RR} - \hat{V}_{RF}(\hat{V}_{FR}\hat{V}_{RR}^{-1}\hat{V}_{RF})^{-1}\hat{V}_{FR}$ and under $\beta = 0$:

$$\hat{\Sigma} \xrightarrow{d} V_{RR} - \Psi_\beta(\Psi'_\beta V_{RR}^{-1}\Psi_\beta)^{-1}\Psi'_\beta = O_p(1)$$

So we can write $\hat{\Sigma} = \hat{\Sigma}^{1/2}\hat{\Sigma}^{1/2'}$, where $\hat{\Sigma}^{1/2} = O_p(1)$.

With $H = (\iota_n \dot{\hat{\beta}}_G)$, we rewrite $(H'H)^{-1}H'\hat{\Sigma}^{1/2}$ as:

$$(H'H)^{-1}H'\hat{\Sigma}^{1/2} = \begin{pmatrix} (\iota'_n M_{\hat{\beta}_G} \iota_n)^{-1} \iota'_n M_{\hat{\beta}_G} \hat{\Sigma}^{1/2} \\ (\hat{\beta}'_G M_{\iota_n} \hat{\beta}_G)^{-1} \hat{\beta}'_G M_{\iota_n} \hat{\Sigma}^{1/2} \end{pmatrix}$$

Note that $(\hat{\beta}'_G M_{\iota_n} \hat{\beta}_G)^{-1} \hat{\beta}'_G M_{\iota_n} \hat{\Sigma}^{1/2} = o_p(1)$, as $\hat{\beta}_G = O_p(T^{1/2})$ when $\beta = 0$. This implies the $(2, 2)$ element of $(H'H)^{-1} H' \hat{\Sigma} H (H'H)^{-1}$, which is the inner product of $(\hat{\beta}'_G M_{\iota_n} \hat{\beta}_G)^{-1} \hat{\beta}'_G M_{\iota_n} \hat{\Sigma}^{1/2}$ when $k = 1$, goes to zero as T increases. \square

E: Proof for Theorem 2

Proof. Since $\hat{\beta}_G = \hat{V}_{\tilde{R}R}^{-1} \hat{V}_{RR}^{-1} \hat{\beta} (\hat{\beta}' \hat{V}_{RR}^{-1} \hat{\beta})^{-1} \hat{V}_{FF}^{-1}$, we have $M_{\iota_{\tilde{n}} \hat{\beta}_G} = M_{\iota_{\tilde{n}} \hat{V}_{\tilde{R}R} \hat{V}_{RR}^{-1} \hat{\beta}}$, which is characterized by $M_{\iota_{\tilde{n}} V_{\tilde{R}R} V_{RR}^{-1} (B + \Psi_{\beta})}$. The proof is then completed by using the independence of $\psi_{\tilde{R}}$ and ψ_{β} in Assumption 1. \square

F: Proof for Theorem 3

To remove the unknown μ_F , we rewrite the linear factor model as:

$$R_t = V_{RF} (V_{FF}^{-1} \bar{F}_t + \lambda_G) + u_t + V_{RF} V_{FF}^{-1} \bar{v}$$

where $\bar{v} = \frac{1}{T} \sum_{t=1}^T (F_t - \mu_F)$. To further remove the unknown V_{FF} , we continue to rewrite this model as:

$$\begin{aligned} R_t &= V_{RF} (\hat{V}_{FF}^{-1} \bar{F}_t + \lambda_G) + u_t + V_{RF} V_{FF}^{-1} \bar{v} + V_{RF} (V_{FF}^{-1} - \hat{V}_{FF}^{-1}) \bar{F}_t \\ &= V_{RF} (\hat{V}_{FF}^{-1} \bar{F}_t + \lambda_G) + \tilde{\epsilon}_t \end{aligned}$$

where $\tilde{\epsilon}_t = u_t + V_{RF} V_{FF}^{-1} \bar{v} + V_{RF} (V_{FF}^{-1} - \hat{V}_{FF}^{-1}) \bar{F}_t$.

Similar to the assumption in Kleibergen (2009), we now assume a central limit theorem.

Assumption 2. When T becomes large,

$$\frac{1}{\sqrt{T}} \sum_{t=1}^T \left(\begin{pmatrix} 1 \\ F_t \end{pmatrix} \otimes (R_t - V_{RF} (\hat{V}_{FF}^{-1} \bar{F}_t + \lambda_G)) \right) \xrightarrow{d} \begin{pmatrix} \tilde{\varphi}_R \\ \tilde{\varphi}_{\beta} \end{pmatrix} \quad (23)$$

with $\tilde{\varphi}_R : n \times 1$, $\tilde{\varphi}_\beta : nk \times 1$, $(\tilde{\varphi}'_R, \tilde{\varphi}'_\beta)' \sim N(0, \tilde{V})$, and

$$\tilde{V} = Q \otimes \tilde{\Omega} \quad (24)$$

with $Q = \begin{pmatrix} 1 & \mu'_F \\ \mu_F & V_{FF} + \mu_F \mu'_F \end{pmatrix} = E \left[\begin{pmatrix} 1 \\ F_t \end{pmatrix} \begin{pmatrix} 1 \\ F_t \end{pmatrix}' \right] : (k+1) \times (k+1)$, $\tilde{\Omega} = \text{var}(\tilde{\epsilon}_t) : n \times n$, and $V_{FF} = \text{var}(F_t) : k \times k$.

Proof. Pre-multiply (23) by

$$\begin{pmatrix} 1 & 0 \\ \left[\frac{1}{T} \sum_{t=1}^T (\hat{V}_{FF}^{-1} \bar{F}_t + \lambda_G)(\hat{V}_{FF}^{-1} \bar{F}_t + \lambda_G)' \right]^{-1} (\lambda_G - \hat{V}_{FF}^{-1} \bar{F}) & \left[\frac{1}{T} \sum_{t=1}^T (\hat{V}_{FF}^{-1} \bar{F}_t + \lambda_G)(\hat{V}_{FF}^{-1} \bar{F}_t + \lambda_G)' \right]^{-1} \hat{V}_{FF}^{-1} \end{pmatrix} \otimes I_n$$

which has the limit

$$\begin{pmatrix} 1 & 0 \\ (V_{FF}^{-1} + \lambda_G \lambda'_G)^{-1} (\lambda_G - V_{FF}^{-1} \mu_F) & (V_{FF}^{-1} + \lambda_G \lambda'_G)^{-1} V_{FF}^{-1} \end{pmatrix} \otimes I_n$$

For the left of \xrightarrow{d} in (23), we have:

$$\begin{aligned} & \frac{1}{\sqrt{T}} \sum_{t=1}^T \left(\begin{pmatrix} 1 \\ \left[\frac{1}{T} \sum_{t=1}^T (\hat{V}_{FF}^{-1} \bar{F}_t + \lambda_G)(\hat{V}_{FF}^{-1} \bar{F}_t + \lambda_G)' \right]^{-1} (\hat{V}_{FF}^{-1} \bar{F}_t + \lambda_G) \end{pmatrix} \otimes (R_t - V_{RF}(\hat{V}_{FF}^{-1} \bar{F}_t + \lambda_G)) \right) \\ &= \sqrt{T} \begin{pmatrix} \bar{R} - V_{RF} \lambda_G \\ \text{vec}(\hat{V}_{RF,r} - V_{RF}) \end{pmatrix} \end{aligned}$$

For the right of \xrightarrow{d} in (23), we have:

$$\left(\begin{pmatrix} 1 & 0 \\ (V_{FF}^{-1} + \lambda_G \lambda'_G)^{-1} (\lambda_G - V_{FF}^{-1} \mu_F) & (V_{FF}^{-1} + \lambda_G \lambda'_G)^{-1} V_{FF}^{-1} \end{pmatrix} \otimes I_n \right) \begin{pmatrix} \tilde{\varphi}_R \\ \tilde{\varphi}_\beta \end{pmatrix} = \begin{pmatrix} \tilde{\xi}_R \\ \tilde{\xi}_\beta \end{pmatrix}$$

where $\tilde{\xi}_R = \tilde{\varphi}_R$, $\tilde{\xi}_\beta = [(V_{FF}^{-1} + \lambda_G \lambda'_G)^{-1} V_{FF}^{-1} \otimes I_n][\tilde{\varphi}_\beta - (\mu_F \otimes I_n)\tilde{\varphi}_R + (V_{FF} \lambda_G \otimes I_n)\tilde{\varphi}_R]$.

Combine the left and the right pieces above, we have:

$$\sqrt{T} \begin{pmatrix} \bar{R} - V_{RF} \lambda_G \\ \text{vec}(\widehat{V}_{RF,r} - V_{RF}) \end{pmatrix} \xrightarrow{d} \begin{pmatrix} \tilde{\xi}_R \\ \tilde{\xi}_\beta \end{pmatrix} \sim N(0, \tilde{Q}(\lambda_G) \otimes \tilde{\Omega})$$

with

$$\begin{aligned} \tilde{Q}(\lambda_G) &= \begin{pmatrix} 1 & 0 \\ (V_{FF}^{-1} + \lambda_G \lambda'_G)^{-1}(\lambda_G - V_{FF}^{-1} \mu_F) & (V_{FF}^{-1} + \lambda_G \lambda'_G)^{-1} V_{FF}^{-1} \end{pmatrix} Q \\ &\quad \begin{pmatrix} 1 & 0 \\ (V_{FF}^{-1} + \lambda_G \lambda'_G)^{-1}(\lambda_G - V_{FF}^{-1} \mu_F) & (V_{FF}^{-1} + \lambda_G \lambda'_G)^{-1} V_{FF}^{-1} \end{pmatrix}' \\ &= \begin{pmatrix} 1 & \lambda'_G (V_{FF}^{-1} + \lambda_G \lambda'_G)^{-1} \\ (V_{FF}^{-1} + \lambda_G \lambda'_G)^{-1} \lambda_G & (V_{FF}^{-1} + \lambda_G \lambda'_G)^{-1} \end{pmatrix} \end{aligned}$$

Note that

$$\bar{R} - \widehat{V}_{RF,r} \lambda_G = \bar{R} - V_{RF} \lambda_G - (\widehat{V}_{RF,r} - V_{RF}) \lambda_G$$

$$\sqrt{T}(\bar{R} - \widehat{V}_{RF,r} \lambda_G) \xrightarrow{d} \tilde{\xi}_R - (\lambda'_G \otimes I_n) \tilde{\xi}_\beta$$

which is independent of $\tilde{\xi}_\beta$, since

$$\begin{pmatrix} 1 & -\lambda'_G \\ 0 & I_k \end{pmatrix} \tilde{Q}(\lambda_G) \begin{pmatrix} 1 & 0 \\ -\lambda_G & I_k \end{pmatrix} = \begin{pmatrix} 1 - \lambda'_G (V_{FF}^{-1} + \lambda_G \lambda'_G)^{-1} \lambda_G & 0 \\ 0 & (V_{FF}^{-1} + \lambda_G \lambda'_G)^{-1} \end{pmatrix}$$

So

$$\sqrt{T} \begin{pmatrix} \bar{R} - \widehat{V}_{RF,r} \lambda_G \\ \text{vec}(\widehat{V}_{RF,r} - V_{RF}) \end{pmatrix} = \sqrt{T} \begin{pmatrix} I_n & -\lambda'_G \otimes I_n \\ 0 & I_{nk} \end{pmatrix} \begin{pmatrix} \bar{R} - V_{RF} \lambda_G \\ \text{vec}(\widehat{V}_{RF,r} - V_{RF}) \end{pmatrix} \xrightarrow{d} \begin{pmatrix} \tilde{\psi}_R \\ \tilde{\psi}_\beta \end{pmatrix}$$

where $\tilde{\psi}_R = \tilde{\xi}_R - (\lambda'_G \otimes I_n) \tilde{\xi}_\beta$ and $\tilde{\psi}_\beta = \tilde{\xi}_\beta$, which are independent and normally distributed. \square

References

- ANDERSON, T. and RUBIN, H. (1949). Estimation of the parameters of a single equation in a complete system of stochastic equations. *Annals of Mathematical Statistics*, pp. 46–63.
- ANG, A., HODRICK, R. J., XING, Y. and ZHANG, X. (2006). The cross-section of volatility and expected returns. *Journal of Finance*, **61** (1), 259–299.
- AVRAMOV, D. and CHORDIA, T. (2006). Asset pricing models and financial market anomalies. *Review of Financial Studies*, **19** (3), 1001–1040.
- BAI, J. and NG, S. (2006). Evaluating latent and observed factors in macroeconomics and finance. *Journal of Econometrics*, **131** (1), 507–537.
- BREEDEN, D. T. (1979). An intertemporal asset pricing model with stochastic consumption and investment opportunities. *Journal of Financial Economics*, **7** (3), 265–296.
- , GIBBONS, M. R. and LITZENBERGER, R. H. (1989). Empirical tests of the consumption-oriented CAPM. *Journal of Finance*, **44** (2), 231–262.
- CONNOR, G. and KORAJCZYK, R. (1988). Risk and return in an equilibrium APT: Application of a new test methodology. *Journal of Financial Economics*, **21** (2), 255–289.
- FAMA, E. and MACBETH, J. (1973). Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy*, **81** (3), 607–636.
- HUBERMAN, G., KANDEL, S. and STAMBAUGH, R. F. (1987). Mimicking portfolios and exact arbitrage pricing. *Journal of Finance*, **42** (1), 1–9.
- JAGANNATHAN, R. and WANG, Z. (1996). The conditional CAPM and the cross-section of expected returns. *Journal of Finance*, pp. 3–53.
- and — (1998). An asymptotic theory for estimating beta-pricing models using cross-sectional regression. *Journal of Finance*, **53** (4), 1285–1309.
- KAN, R., ROBOTTI, C. and SHANKEN, J. (2013). Pricing model performance and the two-pass cross-sectional regression methodology. *Journal of Finance*, **68** (6), 2617–2649.

- and ZHANG, C. (1999). Two-pass tests of asset pricing models with useless factors. *Journal of Finance*, pp. 203–235.
- KLEIBERGEN, F. (2009). Tests of risk premia in linear factor models. *Journal of Econometrics*, **149** (2), 149–173.
- and PAAP, R. (2006). Generalized reduced rank tests using the singular value decomposition. *Journal of Econometrics*, **133** (1), 97–126.
- and ZHAN, Z. (2014). Unexplained factors and their effects on second pass R-squared's. *Journal of Econometrics*, forthcoming.
- LAMONT, O. A. (2001). Economic tracking portfolios. *Journal of Econometrics*, **105** (1), 161–184.
- LETTAU, M. and LUDVIGSON, S. (2001). Resurrecting the (C) CAPM: A cross-sectional test when risk premia are time-varying. *Journal of Political Economy*, **109** (6), 1238–1287.
- LEWELLEN, J., NAGEL, S. and SHANKEN, J. (2010). A skeptical appraisal of asset pricing tests. *Journal of Financial Economics*, **96** (2), 175–194.
- LI, Q., VASSALOU, M. and XING, Y. (2006). Sector investment growth rates and the cross section of equity returns. *Journal of Business*, **79** (3), 1637–1665.
- MUIR, T., ADRIAN, T. and ETULA, E. (2013). Financial intermediaries and the cross-section of asset returns. *Journal of Finance*, forthcoming.
- SHANKEN, J. (1992). On the estimation of beta-pricing models. *Review of Financial Studies*, pp. 1–33.
- STAIGER, D. and STOCK, J. (1997). Instrumental variables regression with weak instruments. *Econometrica*, pp. 557–586.
- VASSALOU, M. (2003). News related to future GDP growth as a risk factor in equity returns. *Journal of Financial Economics*, **68** (1), 47–73.