The Korean Economic Review
Volume 37, Number 2, Summer 2021, 399-417.
DOI : 10.22841/kerdoi.2021.37.2.007

Efficient Mimicking Portfolios in Asset Pricing Tests

Jinyong Kim* · Kun Ho Kim** · Jeong Hwan Lee***

The classic cross-sectional regression (CSR) and mimicking portfolio (MIM) procedures estimate factor risk premia on a test asset span and the resulting tests of asset pricing models are performed with reduced degrees of freedom. Although we can restrict the risk premia of traded factors to equal expected returns, imposing such restrictions on nontraded factors is difficult, which may prevent full performance evaluation. We suggest testing with efficient MIMs that project factors onto a return space spanned by test assets and benchmark traded factors. The generalized method of moments (GMM) tests show that this approach generates more powerful tests and fair comparison against a benchmark model.

JEL Classification: G12
Keywords: Asset Pricing Test, Nontraded Factor, Efficient Mimicking Portfolio, Power

I. Introduction

The two-pass cross-sectional regression (CSR) and mimicking portfolio (MIM) procedures are classic methods for the cross-sectional tests of asset pricing models. The CSR method, developed by Black et al. (1972) and Fama and MacBeth (1973), estimates factor risk premia and pricing errors from regressions of test asset expected returns on betas. The MIM formulation constructs portfolios of test assets to be maximally correlated with factors, which are then used in place of original factors in model tests, according to Huberman et al. (1987) and Breeden et al. (1989). These

Received: June 24, 2020. Revised: Dec. 21, 2020. Accepted: Feb. 25, 2021.
* Corresponding and First Author, Associate Professor, School of Economics, University of Seoul, 163 Seoulsiripdae-ro, Dongdaemun-gu, Seoul 02504, Korea, Phone: +82-2-6490-2074, email: jinyongkim0409@uos.ac.kr
** Co-Author, Assistant Professor, Katz School of Science and Health, Yeshiva University, 215 Lexington Avenue, New York, NY 10016, USA, Phone: +1-646-592-4717, email: kunho.kim@yu.edu
*** Co-Author, Assistant Professor, College of Economics and Finance, Hanyang University, 222 Wangsimni-ro, Seongdong-gu, Seoul 04763, Korea, Phone: +82-2-2220-1036, email: jeonglee@hanyang.ac.kr

standard test procedures estimate the factor risk premia on a return space spanned by test assets so that the model tests are performed with degrees of freedom reduced by the number of factors.

If a model is composed of traded factors only, as the capital asset pricing model (CAPM) of Sharpe (1964) and Lintner (1965), the factors are additionally required so that the factor risk premia are equal to their expected returns and the tests are performed with full degrees of freedom. However, for models with nontraded factors, such as the consumption capital asset pricing model (CCAPM) of Breeden (1979) or the intertemporal capital asset pricing model (ICAPM) of Merton (1973), imposing those restrictions is difficult, and the factor premia are estimated as portfolios of test assets. Accordingly, restrictions imposed on traded factors only leads to testing different candidate models on unequal return spaces that result in unfair comparison. If we test those models on a test asset span to maintain an equal space, as the standard procedures typically do, then the tests are likely to ignore the necessary restrictions on traded factors.

To resolve this problem, we conduct asset pricing tests using an efficient mimicking portfolio (EMIM) approach. Although we are based on the MIM procedure considering that the MIM test is proven to be more powerful than the CSR test by Balduzzi and Robotti (2008), the EMIM approach projects nontraded factors not onto a test asset span but onto a combined span of test assets and traded factors of a benchmark model, which is called the benchmark span. Specifically, we first analyze the small-sample properties by comparing the size and power of the asset pricing tests using the MIM and EMIM procedures under the framework of the generalized method of moments (GMM) of Hansen (1982), which allows us to estimate and test the models without requiring distributional assumptions and to account for all sources of sampling variabilities in estimating parameters. We test two multifactor extensions of the standard CCAPM to explain the firm size and book-to-market (B/M) sorted portfolios compared to the Fama-French (FF) three-factor model (Fama and French, 1993), which is composed of three traded factors, such as the market, size, and B/M factors. Our simulation studies indicate that the EMIM procedure provides a more powerful test than the MIM procedure does as the length of sample period increases.

We then apply the MIM and EMIM procedures to perform the in-sample tests of the candidate models in explaining the cross-sections of the size and B/M sorted portfolios. To further highlight the difference in the in-sample tests between the MIM and EMIM procedures with a relatively small number of cross-sections, we also apply the MIM and EMIM procedures to foreign currency portfolios following the work of Lustig et al. (LRV, 2011). Our examination of the LRV model also shows evidence suspecting the power of the standard asset pricing tests. Although the MIM test does not reject the null hypothesis of zero pricing errors, the EMIM test performed with full degrees of freedom rejects the model. This finding confirms that the EMIM approach provides a more powerful test than the MIM approach does.

In the application of the MIM to test the single-factor CCAPM, Breeden et al. (1989) added the market factor to their test assets to construct the consumption factor-mimicking portfolio, but no formal discussion is provided regarding the augmentation of the test asset span. The idea of including benchmark factors was also suggested by Barillas and Shanken (2017). The authors claimed that, in comparing performance of asset pricing models with traded factors, what matters is not the test assets but factors in other models. Our work is distinguished from Barillas and Shanken (2017) in that we mainly focus on models with nontraded factors, where projection of the factors should be onto the return space generated by benchmark traded factors and test assets for the models to be fairly compared to the benchmark model.[1]

The remainder of this paper is organized as follows. Section 2 provides a short review of the standard asset pricing tests and the implication of reduced degrees of freedom. The motivation of EMIM as a potential remedy to impose full restrictions on pricing errors is also provided in this section. Section 3 tests several existing asset pricing models by applying the MIM and EMIM procedures under the GMM framework. A bootstrap study to compare the size and power properties of the MIM and EMIM tests is performed following Balduzzi and Robotti (2008). The in-sample tests of the candidate models under the two procedures are also performed in this section. Section 4 applies the MIM and EMIM tests to the foreign currency portfolios. Section 5 concludes the paper.

II. Motivating EMIM

We consider an asset pricing model of the form

$$E(R_t) = \beta\lambda, \tag{1}$$

where $R_t = [R_{1,t} \cdots R_{N,t}]'$ is a vector of N excess returns with the expected returns $E(R_t)$ is, and λ is a vector of factor risk premia. An $N \times K$ matrix of betas, β, is the coefficients from the regressions of R_t on K risk factors $f_t = [f_{1,t} \cdots f_{K,t}]'$ as

$$R_t = a + \beta f_t + \varepsilon_t . \tag{2}$$

[1] Barillas and Shanken (2017) also pointed out that “test assets are relevant insofar as they are needed to identify the mimicking portfolio returns.”

The CSR procedure runs the first-pass time-series regressions to obtain the beta estimates, $\hat{\beta}$. In the second pass, the sample means of R_t, μ_R, is run on $\hat{\beta}$ to obtain the least-squares estimates $\hat{\lambda} = (\hat{\beta}'W\hat{\beta})^{-1}\hat{\beta}'W\mu_R$ with a weighting matrix W, where the ordinary least-squares (OLS) CSR corresponds to $W = I_N$, an $N \times N$ identity matrix. The sample estimates of pricing errors α are then derived as

$$\hat{\alpha} = \mu_R - \hat{\beta}\hat{\lambda} = (I_N - \hat{\beta}(\hat{\beta}'W\hat{\beta})^{-1}\hat{\beta}'W)\mu_R, \tag{3}$$

and the model is then tested using the squared pricing error statistic as

$$\hat{\alpha}'V_{\hat{\alpha}}^{+}\hat{\alpha} \sim \chi^2_{N-K}, \tag{4}$$

where $V_{\hat{\alpha}}^{+}$ is a generalized inverse of the covariance matrix of $\hat{\alpha}$, $V_{\hat{\alpha}}$, given that it is singular with idempotent $I_N - (\hat{\beta}'W\hat{\beta})^{-1}\hat{\beta}'W$ in (3). Throughout the paper, for any random variables x and y, we denote the variance of x by V_x and the covariance between x and y by V_{xy}.

Alternatively, f_t is projected onto a set of M base assets R_t^b to construct factor-mimicking portfolios $R_t^m = \omega R_t^b$ with $\omega = V'_{R^b f}V_{R^b}^{-1}$, and R_t^m is employed in place of f_t to test the model, where R_t^b is required to span R_t by satisfying $R_t = \Gamma' R_t^b$ with $M \times N$ matrix Γ of rank N. Given that R_t^m is portfolio returns, the risk premia should equal $E(R_t^m)$ and pricing errors α_m are directly estimated from the time-series regressions

$$R_t = \alpha_m + \beta_m R_t^m + \varepsilon_t^m, \tag{5}$$

so that the model is tested with the estimated pricing errors by

$$\hat{\alpha}'_m V_{\hat{\alpha}_m}^{+}\hat{\alpha}_m \sim \chi^2_{\min(N, M-K)}, \tag{6}$$

which is the time-series test of Gibbons et al. (GRS, 1989). The standard MIM test corresponds to selecting $R_t^b = R_t$ and $\Gamma = I_N$. Then $\hat{\alpha}$ from the generalized least-squares (GLS) CSR with $W = V_R^{-1}$ and $\hat{\alpha}_m$ from the MIM procedure turn out to be identical.

Equations (4) and (6) show that the CSR and MIM tests are performed with reduced $N-K$ degrees of freedom under the null hypothesis that pricing errors for N test assets are jointly zero. The implication of reduced degrees of freedom can be highlighted by testing a model with a single nontraded factor f_t. The resulting condition for zero pricing errors under the CSR and MIM procedures imposes a restriction that, for $i, j = 1, \cdots, N$,

$$\frac{E(R_{i,t})}{V_{R_1 f}} = \frac{E(R_{j,t})}{V_{R_2 f}}, \ \forall i \neq j, \tag{7}$$

indicating that satisfying the equal ratios of expected returns to covariance with f_t among the test assets to claim the successful model performance is sufficient, without further restricting the level of the ratios. As the factor premium is estimated on a span of R_t, the mean-variance ratio of the factor premium is automatically equalized once the ratios of the test assets are equal at any level.

To demonstrate an additional restriction when the factor is a traded return, further suppose that a benchmark model has a traded factor f_t^* that is linearly independent of R_t. The pricing errors are then directly estimated by the intercepts from the time-series regressions of R_t on f_t^*, and the condition of zero pricing errors for the benchmark model becomes

$$\frac{E(R_{i,t})}{V_{R_i f^*}} = \frac{E(f_t^*)}{V_{f^*}}, \ \forall i = 1, \cdots, N, \tag{8}$$

so that the ratios of expected returns of R_t to their covariance with f_t^* should be equal to the mean-variance ratio of f_t^*. This condition indicates that the test assets' risk-return ratios are further restricted by the factor risk-return ratio to claim a successful performance.

The difference between (7) and (8) arises because the models are tested on a different return space. Condition (8) is imposed on an $N+1$-dimensional return space spanned by R_t and f_t^* as bases. Contrastingly, Condition (7) implies that the model is tested on an N-dimensional return space spanned by R_t and f_t is projected on that space, so that the test is performed with reduced $N-1$ degrees of freedom. Accordingly, when the performances of the candidate models with the same number of factors are evaluated in accordance with the standard test procedures on the basis of all required restrictions, those with fewer traded factors may have an advantage with less strict requirements under the null hypothesis.

To resolve this problem, we propose the EMIM test by projecting nontraded risk factors onto a benchmark span generated by augmenting a test asset span with benchmark traded factors. Suppose that, with N test assets, we intend to evaluate performance of an asset pricing model with K nontraded factors, such as macroeconomic factors, in comparison with a benchmark model with K traded factors. With EMIM constructed by projecting the nontraded factors onto the $N+K$ dimensional benchmark span, the candidate and the benchmark models are tested with full N degrees of freedom on the common return space. The single-factor model in Section 2 is revisited to highlight the implication of EMIM estimated by regressing f_t on $R_t^b = [R_t \ f_t^*]$ to construct $R_t^{m*} = \omega^* R_t^b$. The model is then tested with R_t^{m*} in place of f_t by time-series regressions

$$R_t = \alpha_{m*} + \beta'_{m*} R_t^{m*} + \varepsilon_t^{m*}, \tag{9}$$

where the pricing error test with the statistic $\hat{\alpha}'_{m*} V_{\hat{\alpha}_{m*}}^{-1} \hat{\alpha}_{m*}$ is performed with χ_N^2. The condition of zero pricing errors imposed by EMIM then becomes

$$\frac{E(R_{i,t})}{V_{R_i R^{m*}}} = \frac{E(R_t^{m*})}{V_{R^{m*}}}, \ \forall i = 1, \cdots, N, \tag{10}$$

indicating that the expected return-covariance ratios of R_t should be equal to the mean-variance ratio of R_t^{m*}, which is also the requirement for the benchmark model.

III. GMM Tests

In this section, we test existing asset pricing models with nontraded factors that are known to explain the size-B/M portfolios, by applying the MIM and EMIM procedures under the GMM framework. We first describe the candidate models and data and then describe the moment conditions for the GMM estimation and tests. We then report the in-sample estimation results and those of the bootstrap experiments to examine the size and power properties of the MIM and EMIM tests.

3.1. Model and Data

We reexamine the performances of two consumption-based asset pricing models that comprise three nontraded factors: The first model is the Lettau and Ludvigson's (LL, 2001) scaled CCAPM with the consumption-wealth ratio, which is a cointegration residual between the logarithms of consumption, asset wealth, and labor income, as a conditioning factor. The second model is the Lustig and Van Nieuwerburgh's (LVN, 2005) scaled CAPM with the housing collateral ratio, which is a cointegration residual between the logarithms of real estate wealth, and aggregate income, is used as a conditioning variable. The FF three-factor model, which can best explain the size-B/M portfolios, is adopted as a benchmark model for a comparison in performance. We consider only the models with three nontraded factors and maintain the FF three-factor model as a benchmark to avoid such issues as comparing models with different numbers of factors and the arbitrary selection of return bases.

These candidate models include macroeconomic or nontraded factors that are motivated by the equilibrium asset pricing models. Kleibergen and Zhan's (2018) raised a concern regarding using MIM in place of macroeconomic factors when those factors are only weakly correlated with asset returns. Specifically, MIM may misleadingly support the factors, as the betas on MIM can be spuriously large, which indicates that macroeconomic factors tend to be closely connected with the useless factor problem as noted by Kan and Zhang (1999). Kleibergen and Zhan's (2020) suggested asset pricing tests that are robust to the problems of weak identification and limited time series sample size. However, the focus of this paper is neither on the usefulness of nontraded factors per se, nor on comparing the performance of model tests between the CSR and MIM methods. Rather, this paper concentrates on investigating the potential role of imposing a common benchmark span not only comprising test assets but also benchmark traded factors to provide a more powerful test of model performance.

The sample period is from the first quarter of 1963 to the fourth quarter of 2002, which is selected to follow the period of Lewellen et al. (LNS, 2010) as closely as possible.[2] The quarterly consumption growth and the consumption-wealth ratio data are obtained from Sydney Ludvigson's website.[3] The factor data for the LVN model and the housing collateral ratio, are obtained from Stijn Van Nieuwerburgh's website.[4] In addition to the FF 25 size-B/M portfolios as the primary test assets, we add the FF 10 industry-sorted portfolios to the size-B/M portfolios following LNS' prescription to mitigate the factor structure of test assets. We collect the return data for the FF 25 size-B/M portfolios and the 10 industry-sorted portfolios, the three FF factors, and the risk-free rate from Kenneth French's website.[5]

3.2. GMM Estimation

Let x_t denote the data and θ denote the vector of model parameters. The moment conditions for the MIM and EMIM formulations, $E[u(x_t,\theta)]=0$ with a vector-valued function $u(x_t,\theta)$, are as follows:

$$E[f_t-\omega_0-\omega' R_t^b]=0\,, \tag{11}$$

[2] Although the sample period of LNS is from the first quarter of 1963 to the fourth quarter of 2002, our sample period is up to the fourth quarter of 2002 mainly due to the data availability of the housing collateral ratio.

[3] See http://www.econ.nyu.edu/user/ludvigsons/.

[4] See http://pages.stern.nyu.edu/~svnieuwe/.

[5] See http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

$$E[(f_t - \omega_0 - \omega' R_t^b) \otimes R_t^b] = 0, \tag{12}$$

$$E[R_t - \alpha^* - \beta^{*\prime} f_t^*] = 0, \tag{13}$$

$$E[(R_t - \alpha^* - \beta^{*\prime} f_t^*) \otimes f_t^*] = 0, \tag{14}$$

where the estimates are obtained using the exactly identified GMM, as pointed out by Balduzzi and Robotti (2008). In the above conditions, $f_t^* = R_t^m$ when $R_t^b = R_t$, and $f_t^* = R_t^{m*}$ when R_t^b is a combined set of R_t and benchmark traded factors. The asymptotic covariance matrix of the estimates is then derived by

$$V_\theta = \frac{1}{T} d^{-1} S (d^{-1})', \tag{15}$$

where

$$d = \left. \frac{\partial E[u(x_t, \theta)]}{\partial \theta} \right|_{\theta = \hat{\theta}}, \tag{16}$$

and

$$S = \sum_{j=-\infty}^{\infty} E[u(x_t, \hat{\theta}) u(x_{t-j}, \hat{\theta})'], \tag{17}$$

to test the joint significance of $\hat{\alpha}^*$. As discussed in Balduzzi and Robotti (2008), despite that S does not have full rank in the MIM tests, we avoid this singularity problem by constructing the Wald test statistics for $\hat{\alpha}^*$ by inverting portions of S, by which $V_{\hat{\alpha}^*}$ becomes invertible and the tests are conducted with χ_N^2.

3.3. Size and Power

In this section, we investigate the size and power properties of the MIM and EMIM tests for the candidate models to compare the finite-sample performance of the two procedures. We begin by describing the bootstrap experiment to evaluate the statistical size and power of the tests following Balduzzi and Robotti's (2008) return-generating process and simulation procedure. Specifically, on the basis of the composition of traded factors $f_{T,t}$ and nontraded factors $f_{N,t}$ for a selected model, the returns are simulated by the following process:

$$R_t = \alpha + \beta_T' f_{T,t} + \beta_N' [f_{N,t} - E(f_{Nt}) + \lambda_N] + e_t, \tag{18}$$

where α, β_T, β_N, and λ_N are calibrated using the OLS CSR procedure. Given the estimates of the model parameters, we jointly bootstrap the realizations of model factors and the residuals from the OLS CSR to simulate the return data. To address the concern raised by Kleibergena and Zhan (2018) that minor correlations between asset returns and macroeconomic factors may lead to large estimation errors, we scale up the magnitude of variations in the nontraded factors up to four times as large as their original levels in the simulation. Consequently, we can also limit the size distortion and enhance the reliability of our simulation results for the power properties. Although the results are not reported here, different scaling of macroeconomic factors (such as to three or five times) does not alter the size and power properties significantly.

Table 1 provides the size of the MIM and EMIM tests for the selected models, where the top panel reports the results with the size-B/M portfolios and the bottom panel reports those with the size-B/M and industry portfolios. Considering that our in-sample tests are performed with 156 quarterly data, we adopt three different lengths of time horizon as $T = 156$, 312, and 624. With the simulated returns under the null hypothesis of $\alpha = 0$, the MIM and EMIM parameters are then estimated using the exactly identified GMM, and the Wald test statistics for the pricing errors are obtained assuming no serial correlations. The simulations are iterated 10,000 times, by which the sizes of the MIM and EMIM tests are evaluated at the 10%, 5%, and 1% significance levels. For each model, the rejection rates from the MIM tests are reported in the column on the left, where those from the EMIM tests are reported in the column on the right.

Table 1 indicates that the MIM and EMIM tests tend to show over-rejections, especially in the short term. For the size-B/M portfolios, the rejection rates of MIM and EMIM are in the range of 0.44–0.57 at the 10% level and 0.20 to 0.32 at the 1% level, with $T = 156$. These over-rejections are consistent with the results of the bootstrap experiments in Balduzzi and Robotti (2008), where MIM consistently shows higher rejection rates than OLS and GLS CSR. Our results also indicate that EMIM often generates higher rejection rates than MIM does in the short sample, and the short sample rejection rates become even higher when the 10 industry portfolios are added to the set of test assets, as presented in the bottom panel. However, as the time length increases, the rejection rates tend to approach the nominal significance levels. With the length of $T = 624$, the rejection rates are in the range of 0.03–0.07 for the size-B/M portfolios and 0.05 to 0.07 for the size-B/M and industry portfolios at the 1% level.[6]

[6] We confirm that the size distortion disappears and the EMIM-base test is performed with correct size as T becomes sufficiently large.

[Table 1] Size of the GMM test

		LL		LVN		FF	
		MIM	EMIM	MIM	EMIM	MIM	EMIM
		25 size-B/M					
10%	156	0.436	0.486	0.483	0.518	0.571	0.566
	312	0.267	0.283	0.265	0.279	0.340	0.319
	624	0.196	0.194	0.174	0.174	0.247	0.198
5%	156	0.343	0.389	0.387	0.421	0.478	0.478
	312	0.182	0.199	0.185	0.194	0.257	0.235
	624	0.126	0.125	0.111	0.106	0.167	0.133
1%	156	0.195	0.229	0.239	0.259	0.317	0.317
	312	0.082	0.088	0.085	0.088	0.133	0.115
	624	0.051	0.047	0.038	0.034	0.072	0.048
		25 size-B/M + 10 industry					
10%	156	0.613	0.650	0.667	0.683	0.722	0.735
	312	0.329	0.363	0.368	0.375	0.414	0.423
	624	0.200	0.223	0.221	0.226	0.244	0.248
5%	156	0.526	0.558	0.586	0.598	0.647	0.657
	312	0.241	0.264	0.273	0.276	0.312	0.331
	624	0.128	0.145	0.143	0.146	0.167	0.169
1%	156	0.361	0.388	0.428	0.432	0.491	0.505
	312	0.117	0.129	0.143	0.142	0.173	0.187
	624	0.046	0.055	0.056	0.054	0.068	0.067

Note: This table reports the simulation experiments to examine the size of the GMM test under the standard mimicking portfolio (MIM) and efficient mimicking portfolio (EMIM) procedures, for the LL, LVN, and FF models. The simulations are repeated 10,000 times, and the rejection rates are calculated. All models are simulated under the alternative hypothesis that α equals to zero; all other parameters are fixed at the estimated values obtained using the OLS CSR. The top panel reports the results with the 25 size-B/M portfolios, and the bottom panel reports the results with the 25 size-B/M and 10 industry-sorted portfolios, as the test assets. The rejection rates are calculated for the time periods of 156, 312, and 624 quarters by using 10%, 5%, and 1% quantiles of the theoretical distribution of the Wald test statistics under the null hypothesis generated from the GMM tests with MIM and EMIM. LL is the Lettau–Ludvigson model, LVN is the Lustig-Van Nieuwerburgh model, and FF is the Fama-French model.

To compare the statistical power of the MIM and EMIM tests, we calculate the rejection rates for the Wald test statistics under an alternative hypothesis on the basis of the empirical distributions obtained from the size experiments in Table 1 to adjust for the effect of size distortions. As an alternative hypothesis, we set all the alphas in (18) to the 30% scale of the estimated values from the in-sample OLS CSR, which we deliberately choose to avoid the trivial case of 100% rejection rates for all the models by selecting the original scale and to make an effective power comparison between the MIM and EMIM. All the models are again simulated 10,000 times under the alternative hypothesis.

Table 2 provides the size-corrected power of the MIM and EMIM tests for the selected models with $T = 156$, 312, and 624 at the 10%, 5%, and 1% significance levels, where the top panel reports the results for the size-B/M portfolios and the bottom panel reports those for the size-B/M and industry portfolios as the test assets. For each model, the rejection rate from the MIM test is reported in the column on the left, and that of the EMIM test is reported in the column on the right. Table 2 shows that, although EMIM does not show any power advantage for $T = 156$, the EMIM tests tend to gain power faster than the MIM tests as the time length

[Table 2] Power of the GMM test

		LL		LVN		FF	
		MIM	EMIM	MIM	EMIM	MIM	EMIM
		25 size-B/M					
10%	156	0.174	0.156	0.179	0.160	0.254	0.242
	312	0.434	0.414	0.432	0.436	0.518	0.534
	624	0.827	0.838	0.849	0.883	0.891	0.926
5%	156	0.094	0.082	0.096	0.080	0.152	0.143
	312	0.282	0.265	0.270	0.283	0.378	0.380
	624	0.731	0.742	0.736	0.786	0.818	0.862
1%	156	0.018	0.014	0.018	0.013	0.038	0.031
	312	0.101	0.090	0.090	0.083	0.167	0.162
	624	0.463	0.472	0.470	0.522	0.599	0.683
		25 size-B/M + 10 industry					
10%	156	0.281	0.277	0.320	0.288	0.355	0.332
	312	0.760	0.778	0.770	0.768	0.810	0.818
	624	0.997	0.998	0.998	0.997	0.998	0.999
5%	156	0.160	0.155	0.199	0.164	0.232	0.201
	312	0.620	0.631	0.636	0.627	0.691	0.704
	624	0.991	0.993	0.992	0.993	0.994	0.995
1%	156	0.032	0.035	0.056	0.047	0.075	0.058
	312	0.311	0.305	0.335	0.334	0.430	0.405
	624	0.945	0.954	0.949	0.958	0.960	0.968

Note: This table reports the simulation experiments to examine the size-adjusted power of the GMM test under the standard mimicking portfolio (MIM) and efficient mimicking portfolio (EMIM) procedures, for the LL, LVN, and FF models. The size is adjusted using the bootstrap results in Table 2. The simulations 10,000 are repeated times, and the rejection rates are calculated. All models are simulated under the alternative hypothesis that α equals to their estimated values; all other parameters are fixed at the estimated values obtained using thr OLS CSR. The top panel reports the results with the 25 size-B/M portfolios, and the bottom panel reports the results with the 25 size-B/M and 10 industry-sorted portfolios, as the test assets. The rejection rates are calculated for the time periods of 156, 312, and 624 quarters by using 10%, 5%, and 1% quantiles of the empirical distribution of the Wald test statistics generated from the GMM tests with MIM and EMIM in the size experiments. LL is the Lettau-Ludvigson model, LVN is the Lustig-Van Nieuwerburgh model, and FF is the Fama-French model.

increases, and the power of the EMIM tests becomes superior to that of the MIM tests for $T = 624$. These results indicate that, although deciding on one test that uniformly provides better power over the other test is difficult, the EMIM procedure can generally provide a more powerful test than the MIM procedure if we have reasonably sufficient length of time-series data. Although the results are not reported here, we check the robustness of our results to alternative alphas by selecting different scales of the in-sample estimated values. Our finding that the power of the EMIM test becomes superior to that of the MIM test, especially as the time length increases, remains consistent with the alternative alpha values.

3.4. In-sample Tests

Table 3 presents the results of the in-sample Wald tests with MIM in the panel on the left and those with EMIM in the panel on the right for the three selected models. The top panel reports the results of the tests with the size-B/M portfolios, where the bottom panel reports the results with the size-B/M and industry portfolios, as the test assets. All the candidate models are strongly rejected by the in-sample Wald tests, regardless of whether the models are tested with MIM or EMIM. This is somewhat inconsistent with the results of LNS, in which most of the models with nontraded factors are not rejected at the standard significance levels. This difference indicates that MIM provides more accurate risk premium estimates in a

[Table 3] In-sample GMM test with the size-B/M portfolios

	MIM		EMIM	
	Wald	p –value	Wald	p –value
		25 size-B/M		
LL	75.606	0.000	62.913	0.000
LVN	107.221	0.000	97.922	0.000
FF	61.815	0.000	80.578	0.000
		25 size-B/M + 10 industry		
LL	172.242	0.000	163.313	0.000
LVN	207.042	0.000	208.416	0.000
FF	141.613	0.000	159.900	0.000

Note: This table reports the Wald test statistics and p –values from the GMM estimation with the standard mimicking portfolio (MIM) and efficient mimicking portfolio (EMIM) procedures. The top panel reports the results with the 25 size-B/M portfolios, and the bottom panel reports the results with the 25 size-B/M and 10 industry-sorted portfolios, as the test assets. The benchmark traded factors are the FF three factors. The p-values are obtained from χ^2_{22} for the size-B/M portfolios and from χ^2_{32} for the size-B/M and industry portfolios in MIM, and χ^2_{25} for the size-B/M portfolios and from χ^2_{35} for the size-B/M and industry portfolios in EMIM, respectively. LL is the Lettau-Ludvigson model, LVN is the Lustig-Van Nieuwerburgh model, and FF is the Fama-French model. The sample period is from the first quarter of 1963 to the fourth quarter of 2002.

small sample, which is closely related to LNS' recommendation for the magnitude of the cross-sectional slopes also consistent with Balduzzi and Robotti's (2008) results.

While we present the enhanced power in the previous section, the in-sample results do not show any advantage of the EMIM-based tests simply because all the models are strongly rejected by the MIM and EMIM procedures. These results are likely to arise given that the correction of three degrees of freedom out of 25 for the size-B/M portfolios and 35 for the size-B/M and industry portfolios may fall short of making a significant difference, and suggest that comparison of the two procedures based on a smaller number of cross-sections should be able to make a distinction in their effects on the asset pricing tests more clearly. This issue is further discussed in the next section.

IV. Test with Foreign Currency Portfolios

This section considers another example of cross-sectional tests with foreign currency portfolios, which have attracted much attention in recent literature. Given that a correction for relatively small degrees of freedom in a parsimonious model will not noticeably impact performance evaluation with a large number of cross-sections, resolving the issue of reduced degrees of freedom may not significantly change asset pricing tests. However, the test with foreign currency portfolios is particularly compelling, as the number of portfolios to test in the cross-section is typically small. Employing a benchmark span in this case to achieve a test with full degrees of freedom can significantly impact evaluating the model performance.

We follow LRV's test of the common factor model to explain foreign currency returns with the currency portfolios sorted by forward discounts. LRV built six currency portfolios, from the lowest-interest rate to the highest-interest rate currencies, and demonstrated that the first two principal components of the portfolios, the average excess return, and slope factors, can explain most of the variations in the currency portfolio returns. The two factors are interpreted as such that the average return factor captures the home risk premium for dollar risk and the slope factor reflects the carry trade risk premium for global risk. Menkhoff et al. (2012) then proposed that global risk can also be captured by the global equity volatility factor. We adopt a two-factor model on the basis of these findings, which is composed of the average return and global equity volatility factors, and perform the cross-sectional test for the currency portfolios sorted by the interest rate. The monthly data from November 1983 to December 2009 ($T = 314$) for the six portfolio returns and the model factors are obtained from Hanno Lustig's website.[7]

[7] See http://sites.google.com/site/lustighanno/data.

We highlight the effect of full degrees of freedom on the test by making one modification in the selection of test assets. Of the two factors, the average return factor is obviously replicated by taking an average of the six currency portfolios. Therefore, we eliminate one mid-interest rate currency portfolio with the highest correlation among the pairs of the six portfolios from the test assets. This modification maintains the five-dimensional test asset span without mitigating the currency portfolios' cross-sectional variation, and the empirical results are robust to a selection of an eliminated portfolio as long as the portfolios of the highest and the lowest interest rate currencies are maintained in the test assets. Accordingly, we test the cross-sections of the five-currency portfolio returns with the two risk factors, namely, the average return factor of the six currency portfolios and the nontraded global equity volatility factor. The average return factor is then included to construct a benchmark span as another basis to the five test assets, and one more return basis must be added to achieve the test with full five degrees of freedom. Although no specific benchmark model exists and the test is only applied to the single model, we include the average return of the developed-country currency portfolio as an additional basis for a benchmark span, which is also available from Lustig's website. This portfolio does not arbitrarily deviate from the existing test assets and has already been used for the robustness test of the LRV's currency factor model.[8] Thus, the test on the benchmark span, constructed by the five test assets and two selected additional bases, can effectively reveal the implications of the full degrees of freedom.

Table 4 presents the results of the GRS tests in (6) for the two-factor model to explain the cross-section of five foreign currency portfolios. The panels on the left and right provide the GRS test statistics and p–values from the MIM and EMIM procedures, respectively. Under the MIM procedure, the GRS test performed with the χ_3^2 distribution does not reject the model at the standard significance level,

[Table 4] Test with the foreign currency portfolios

MIM		EMIM	
GRS	p–value	GRS	p–value
5.37	0.15	12.47	0.03

Note: This table reports the GRS test statistics and p–values for Lustig et al.'s (2011) common factor model with the five foreign currency portfolios sorted by interest rate, under the standard mimicking portfolio (MIM) and efficient mimicking portfolio (EMIM) procedures. The benchmark span for EMIM is constructed by adding the average return factor and the developed-country average return to the test assets. The p-values are derived from the χ_3^2 (test asset span) and χ_5^2 (benchmark span) distributions, respectively. The data are monthly, and the sample period is from January 1983 to December 2009.

[8] Our results are robust to using individual currency returns as alternative bases to construct a benchmark span.

suggesting that the factor model successfully prices the five currency portfolios. However, this performance is not maintained once the model is evaluated with EMIM, where the p–value is 0.03 for the GRS test with the χ_5^2 distribution so that the model is rejected at the 5% significance level, demonstrating that requiring the average return factor to be exactly priced and including the developed-country currency portfolio return as an additional basis generate a different evaluation in the model test.

These results indicate that, although a model performs well in explaining the cross-section of test assets when it is tested on the test asset span with reduced degrees of freedom, the model may not fully capture the cross-section once it is tested by imposing full pricing restrictions on the benchmark span. Additionally, the effects of testing with full degrees of freedom can be more clearly observed when the number of cross-sections is smaller, as with foreign currency portfolios. Finally, rather than arguing that the selected models perform poorly, our primary conclusion is that constructing a benchmark span on which candidate models are tested with full degrees of freedom can enhance the statistical power of the asset pricing tests, and provide a fair comparison of cross-sectional performance.

V. Conclusions

We perform the standard CSR and MIM procedures to test asset pricing models with nontraded factors estimate factor risk premia on a test asset span and the resulting tests with reduced degrees of freedom. For models with traded factors, restrictions for the factor risk premia to equal their expected returns can be further imposed. However, we cannot impose such restriction on nontraded factor models so that factor risk premia are estimated as portfolios of test assets, which result in either comparing candidate models with different factor compositions on unequal return spaces or not fully imposing necessary restrictions to evaluate model performance.

In this paper, we suggest the EMIM approach where all the model factors are projected onto a benchmark span that is generated by combining test assets and benchmark traded factors. On the benchmark span, competing models can be compared fairly with full restrictions, which suggest that the EMIM procedure may provide a more powerful test than the standard MIM procedure. We investigate the effect of adopting EMIM on the asset pricing tests by examining the size and power properties of the MIM and EMIM tests for existing asset pricing models that explain the size-B/M portfolios under the GMM framework. Our results provide evidence that the EMIM tests are more powerful than the MIM tests with higher rejection rates under the alternative hypotheses based on the data. We also examine the foreign currency risk model by employing the EMIM approach, which is relevant with a small number of cross-sections, and find that maintaining a favorable performance is difficult once it is evaluated with EMIM on the benchmark span. This again suggests that EMIM can increase the statistical power of asset pricing tests.

References

Balduzzi, P., and C. Robotti (2008), "Mimicking Portfolios, Economic Risk Premia, and Tests of Multi-beta Models," *Journal of Business and Economic Statistics*, 26, 354–368.

Barillas, F., and J. Shanken (2017), "Which Alpha?" *Review of Financial Studies*, 30, 1316–1338.

Black, F., M. Jensen, and M. Scholes (1972), "The Capital Asset Pricing Model: Some Empirical Tests," in Michael Jensen, eds., Studies in the Theory of Capital Markets, Praeger, New York.

Breeden, D. (1979), "An Intertemporal Asset Pricing Model with Stochastic Consumption and Investment Opportunities," *Journal of Financial Economics*, 7, 265–296.

Breeden, D., M. Gibbons, and R. Litzenberger (1989), Empirical Test of the Consumption-oriented CAPM," *Journal of Finance*, 44, 231–262.

Fama, E., and K. French (1993), "Common Risk Factors in the Returns on Stocks and Bonds," *Journal of Financial Economics*, 33, 3–56.

Fama, E., and J. MacBeth (1973), "Risk, Return, and Equilibrium: Empirical Tests," *Journal of Political Economy*, 71, 607–636.

Gibbons, M., S. Ross, and J. Shanken (1989), "A Test of the Efficiency of a Given Portfolio," *Econometrica*, 57, 1121–1152.

Hansen, L. (1982), "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50, 1029–1054.

Huberman, G., S. Kandel, and R. Stambaugh (1987), Mimicking Portfolios and Exact Arbitrage Pricing," *Journal of Finance*, 42, 1–9.

Kan, R., and C. Zhang (1999), "Two-pass Tests of Asset Pricing Models with Useless Factors," *Journal of Finance*, 54, 203–235.

Kleibergen, F., and Z. Zhan (2018), "Identification-robust Inference on Risk Premia of Mimicking Portfolios of Non-traded Factors," *Journal of Financial Econometrics*, 16, 155–190.

Kleibergen, F., and Z. Zhan (2020), "Robust Inference for Consumption-based Asset Pricing," *Journal of Finance*, 75, 501–550.

Lettau, M., and S. Ludvigson (2001), "Resurrecting the (C)CAPM: A Cross-sectional Test when Risk Premia are Time-varying," *Journal of Political Economy*, 109, 1238–1287.

Lewellen, J., S. Nagel, and J. Shanken (2010), "A Skeptical Appraisal of Asset Pricing Tests," *Journal of Financial Economics*, 96, 175–194.

Lintner, J. (1965), "The Valuation of Risky Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets," *Review of Economics and Statistics*, 47, 13–37.

Lustig, H., N. Roussanov, and A. Verdelhan (2011), "Common Risk Factors in Currency Markets," *Review of Financial Studies*, 24, 3731–3777.

Lustig, H.m and S. Van Nieuwerburgh (2005), "Housing Collateral, Consumption Insurance, and Risk Premia: An Empirical Perspective," *Journal of Finance*, 60, 1167–1219.

Menkhoff, L., L. Sarno, M. Schmeling, and A. Schrimpf (2012), "Carry Trades and Global

Foreign Exchange Volatility," *Journal of Finance*, 67, 681–718.

Merton, R. (1973), "An Intertemporal Capital Asset Pricing Model," *Econometrica*, 41, 867–887.

Sharpe, W. (1964), "Capital Asset Prices: A Theory of Market Equilibrium," *Journal of Finance*, 19, 425–442.

자산가격모형 평가에 있어서의 효율적 모방 포트폴리오 방법

김 진 용* · 김 건 호** · 이 정 환***

논문초록 횡단면 회귀분석 및 모방 포트폴리오 방법을 이용한 기존의 자산가격모형 성과 평가 과정에서는 모형 요인의 리스크 프리미엄이 테스트 자산만의 포트폴리오로 추정됨에 따라 자유도가 감소되어 통계적 검정이 이루어진다. 모형 요인이 수익률 변수인 경우 리스크 프리미엄이 기대수익률과 같다는 추가적인 제약을 부여할 수 있으나, 요인이 비수익률 변수인 경우에는 이러한 제약을 부여할 수 없어 모형의 성과에 대한 완전한 비교 평가가 어려워진다. 본 논문에서는 평가 대상 모형의 위험 요인을 테스트 자산 및 성과기준 모형의 수익률 요인에 의해 생성되는 확장된 수익률 공간에 투영하는 효율적 모방 포트폴리오 방법을 제안한다. 일반화적률법에 의한 모형의 성과 평가에 있어서 효율적 모방 포트폴리오 방법이 검정력을 높이고 기준 모형과의 비교를 보다 공정하게 수행할 수 있음을 보인다.

핵심 주제어: 자산가격모형, 비수익률 요인, 효율적 모방 포트폴리오, 검정력
경제학문헌목록 주제분류: G12

투고 일자: 2020. 6. 24. 심사 및 수정 일자: 2020. 12. 21. 게재 확정 일자: 2021. 2. 25.

* 교신저자 및 제1저자, 서울시립대학교 경제학부 부교수, e-mail: jinyongkim0409@uos.ac.kr

** 공동저자, 조교수, Katz School of Science and Health, Yeshiva University, e-mail: kunho.kim@yu.edu

*** 공동저자, 한양대학교 경제금융대학 조교수, e-mail: jeonglee@hanyang.ac.kr