Mimicking Portfolios, Economic Risk Premia, and Tests of Multi-Beta Models
Author(s): Pierluigi Balduzzi and Cesare Robotti
Source: *Journal of Business & Economic Statistics*, Vol. 26, No. 3 (Jul., 2008), pp. 354-368
Published by: Taylor & Francis, Ltd. on behalf of American Statistical Association
Stable URL: https://www.jstor.org/stable/27638994

Mimicking Portfolios, Economic Risk Premia, and Tests of Multi-Beta Models

Pierluigi BALDUZZI
Boston College, Finance Department, The Wallace E. Carroll School of Management, Chestnut Hill, MA 02467
(*balduzzp@bc.edu*)

Cesare ROBOTTI
Federal Reserve Bank of Atlanta, Research Department, 1000 Peachtree Street NE, Atlanta, GA 30309
(*cesare.robotti@atl.frb.org*)

We consider two formulations of the linear factor model (LFM) with nontraded factors. In the first formulation, LFM, risk premia and alphas are estimated by a cross-sectional regression of average returns on betas. In the second formulation, LFM*, the factors are replaced by their projections on the span of excess returns, and risk premia and alphas are estimated by time series regressions. We compare the two formulations and study the small-sample properties of estimates and test statistics. We conclude that the LFM* formulation should be considered in addition to, or even instead of, the more traditional LFM formulation.

KEY WORDS: Mimicking portfolios; Economic risk premia; Linear factor models.

1. INTRODUCTION

Most asset-pricing models explain the cross-section of expected returns in terms of exposures, or betas, to one or more risk factors. We call these models *multi-beta* models. In the consumption-based capital asset pricing model (C–CAPM) and the intertemporal capital asset pricing model (I–CAPM), one or more of the risk factors are "economic" and they are not themselves asset returns. The standard formulation of such asset-pricing models starts from a linear factor model (LFM),

$$r_t = \alpha + \beta^\top[\lambda + y_t - E(y_t)] + e_t, \qquad (1)$$

where r_t is an $N \times 1$ vector of returns in excess of the risk-free rate, α is the vector of deviations from the model, λ is the vector of economic risk premia, y_t is a $K \times 1$ vector of factor realizations, and e_t is an $N \times 1$ vector of residuals orthogonal to the factors. If the multi-beta model is correct, then $\alpha = 0$, and a unit-beta portfolio (i.e., any portfolio with a beta of 1 with respect to factor y_k and 0 with respect to all the other factors) earns a risk premium equal to λ_k.

The risk premia on the unit-beta portfolios can be immediately estimated as the coefficients (or averages of coefficients) of a cross-sectional regression (CSR) of average returns (or returns) on the betas. This CSR can be performed using various weighting schemes. Here we focus on two of these schemes. In the first scheme, the assets are unweighted. This corresponds to a CSR ordinary least squares (OLS) style, following the seminal articles of Black, Jensen, and Scholes (1972) and Fama and MacBeth (1973). This approach has been implemented in numerous empirical studies. In the second scheme, the assets are weighted by the inverse of the covariance matrix of the idiosyncratic components. This corresponds to a CSR generalized least squares (GLS) style.

An alternative formulation of the LFM is obtained when we replace the factors with the variable component of their projections onto the span of the excess returns, augmented with a constant (see Huberman, Kandel, and Stambaugh 1987). We have

$$y_t = \gamma_0^\star + (\gamma^\star)^\top r_t + \epsilon_t \qquad (2)$$

and $y_t^\star \equiv (\gamma^\star)^\top r_t$, where $\gamma^\star = \Sigma_{rr}^{-1}\Sigma_{ry}$. For the mimicking portfolios, $y_t^\star$, to exist we assume that $(\gamma^\star)^\top 1_N = \Sigma_{yr}\Sigma_{rr}^{-1}1_N \neq 0_K$, where 1_N is an $N \times 1$ vector of 1's and 0_K is a $K \times 1$ vector of 0's. This condition is equivalent to assuming that the global minimum-variance portfolio has positive systematic risk. We then have

$$r_t = \alpha^\star + (\beta^\star)^\top y_t^\star + e_t^\star. \qquad (3)$$

We denote this alternative formulation as LFM*.

The projection coefficients are the weights of portfolios whose returns have maximum (squared) correlations with the factors. Thus, these maximum-correlation portfolio weights are proportional to the hedging-portfolio weights of Merton (1973). Other examples of this second approach have been provided by Breeden (1979), Breeden, Gibbons, and Litzenberger (1989), Fama (1996), Lamont (2001), and Ferson, Siegel, and Xu (2005).

Table 1 presents results for the LFM and the LFM* formulations of the C–CAPM. (Details on the estimation procedures and the data used are given later in the article.) Table 1 shows how, depending on the linear factor model, parameter estimates and statistical inferences can differ substantially: The consumption risk premium is .24 and significant in the LFM (CSR–OLS) but only .01 and insignificant in the LFM*. Moreover, the test of overidentifying restrictions leads to a strong rejection of the C–CAPM in the LFM* formulation but only to a weak rejection in the CSR–OLS version of the LFM formulation. The literature does not explain the theoretical relation between these two sets of results, nor does it establish which representation is most likely to deliver the correct estimates and inference.

One contribution of this article is to clarify the relationships between the LFM and LFM* formulations and between unit-beta and maximum-correlation portfolios. We show that $\alpha = 0$ if and only if $\alpha^\star = 0$, that $\lambda = \lambda^\star$ if $\alpha = 0$ and if the factors are traded ($y_t = y_t^\star$), and that the estimates of α and $\alpha^\star$ coincide

Journal of Business & Economic Statistics
July 2008, Vol. 26, No. 3
DOI 10.1198/073500108000000042

Table 1. LFM versus LFM* representations

	LFM C–CAPM				LFM* C–CAPM	
	λ_{OLS}	λ_{GLS}	χ^2_{OLS}	χ^2_{GLS}	λ^*	χ^2
Estimate	.24	.09	17.94	23.92	.01	28.66
t-ratio	2.23	1.43			1.03	
p-value	.026	.152	.036	.004	.300	.001

NOTE: Here we estimate a C–CAPM economy. Parameter estimates of the LFM are obtained by cross-sectional regression, with $W = I$ (OLS) and $W = \hat{\Sigma}_{ee}^{-1}$ (GLS). Estimates of the parameters of the LFM* are obtained by time series regressions. Asymptotic standard errors and chi-squared statistics are obtained by GMM, assuming no serial dependence. The sample is from 1959:03–2002:11. We consider the 10 size portfolios as the test assets.

when the CSR is GLS style. We also show that when the unit-beta mimicking portfolios have minimum idiosyncratic risk, we can obtain the same Sharpe ratio from maximum-correlation portfolios and from unit-beta mimicking portfolios. Finally, we show how the construction of "mimicking portfolios" allows us to translate characteristic-based explanations of the cross-section of stock returns into risk-based explanations.

Following Hansen (1982) and Cochrane (2001), we show how both formulations of the linear factor model, and both estimation approaches (CSR and time series), can be cast within the unifying framework of the generalized method of moments (GMM). This approach allows us to derive the asymptotic properties of estimates and statistics without the need for restrictive distributional assumptions. We then construct a simulation exercise to assess the properties of the estimators for the two different formulations of the model. The simulations consider a one-factor (C–CAPM) and a two-factor (I–CAPM) economy, both iid and serially correlated data, different lengths of the sample (525 vs. 240 months), and different choices of test assets (10 size-sorted portfolios vs. the 25 size- and value-sorted portfolios of Fama and French 1993).

Summarizing our simulation exercises, the LFM* formulation performs better than the LFM–CSR–GLS formulation and similar to the LFM–CSR–OLS formulation in terms of bias of the risk premium estimates. In terms of root mean squared errors (RMSEs) of the risk premium estimates and power of the tests, the LFM* and LFM–CSR–GLS perform similarly and better than the LFM–CSR–OLS formulation. In the case of a noisy factor, the LFM* formulation works much better than either LFM formulation. Our small-sample analysis should help researchers draw the correct conclusions when facing different results from different methods. For example, an econometrician faced with the results of Table 1 should be more (less) inclined to trust the risk premium estimates (test statistics) resulting from the LFM* representation than from the LFM representation.

We also explore several extensions of the basic analysis summarized earlier. First, we consider the case of time-varying conditional moments and conditional versions of the LFM and LFM* formulations. In the presence of time-varying betas and risk premia, the LFM* formulation performs worse than the LFM formulation in terms of bias but better in terms of RMSE of the risk premium estimates. Second, we consider the approach of Lehmann and Modest (1988) to the construction of mimicking portfolios. Their approach generally does worse than the LFM* formulation, in terms of both bias and RMSE of the risk premium estimates. Third, we consider the correction for small-sample bias in the CSR risk premium estimates, suggested by Chen and Kan (2006). In several cases, the bias adjustment actually leads to worse bias, and the bias-adjusted CSR estimators always have higher RMSEs than the nonadjusted λ^* estimates. Fourth, we study the various methods in the case of long-horizon overlapping returns. Using long-horizon returns generally leads to higher biases and RMSEs, although in the case of the C–CAPM, these effects are more pronounced for the CSR risk premium estimates than for the λ^* estimates. Finally, we consider the implications of the restrictions from the LFM and LFM* formulations on covariance matrices, for the purpose of constructing minimum-variance portfolios. Interestingly, the restrictions from the LFM* formulation generally lead to portfolios with better properties than the portfolios restricted using the LFM formulation.

For the aforementioned reasons, we conclude that the LFM* formulation should be considered in addition to, or even instead of, the more traditional LFM formulation. This conclusion applies not only to the estimation and testing of multi-beta models, but also to the construction of mean–variance efficient portfolios.

The article is organized as follows. Section 2 provides a review of the literature. Section 3 presents the theoretical results. Section 4 derives the moment conditions used in estimation, and Section 5 derives the asymptotic properties of the GMM estimators. Section 6 discusses the case in which both traded and nontraded factors are present and the nontraded factor is observed with noise. Section 7 illustrates the data used to calibrate the simulation exercise. Section 8 describes the bootstrap experiment, Section 9 discusses the results of the simulation, and Section 10 explores some extensions of the analysis. The final section summarizes our findings, and the Appendix contains proofs of the analytical results.

2. RELATED LITERATURE

Several authors have studied the properties of CSR risk premium estimators, including Litzenberger and Ramaswamy (1979), Amsler and Schmidt (1985), MacKinlay (1987), Shanken (1992), Shanken and Zhou (2007), Affleck-Graves and Bradfield (1993), Kim (1995), and Chen and Kan (2006). Others have focused on the comparison between stochastic discount factor (SDF) and LFM representations in tests of multi-beta models (see, e.g., Cochrane 2001; Kan and Zhou 1999, 2001a; Jagannathan and Wang 2002).

The theoretical properties of different mimicking portfolios and formulations of the LFM have been discussed by Grinblatt and Titman (1987), Huberman et al. (1987), and Kandel and Stambaugh (1995). Grinblatt and Titman showed that if the arbitrage pricing theory holds, then the portfolios of maximum likelihood factor analysis can be combined to obtain a Sharpe ratio as high as the maximum Sharpe ratio attainable from all of the available assets. Huberman et al. showed how unit-beta portfolios and maximum-correlation portfolios can replace the nontraded factors in linear factor models. Kandel and Stambaugh showed how the CSR–GLS results, but not the CSR–OLS results, are invariant to the repackaging of the individual securities.

Our analysis focuses on risk exposures (i.e., betas) as the drivers of the cross-section of returns. Starting with Fama and French (1992), several authors have focused instead on "characteristics," such as a company's market capitalization and book-to-market ratio. Fama and French (1992) interpreted the characteristics as new indicators of risk exposure, and then later (1993) constructed long–short characteristic-sorted portfolios, showing that the exposures to the returns on those portfolios [Fama and French (FF) "factors"] explain the cross-section of average returns. The articles by Daniel and Titman (1997) and Brennan, Chordia, and Subrahmanyam (1998) challenged Fama and French's risk-based interpretation, and Ferson (1996) and Ferson, Sarkissian, and Simin (1999) questioned the very possibility of distinguishing between characteristic- and risk-based explanations.

The present article focuses on the comparison between two alternative formulations of a multi-beta model with nontraded factors. To the best of our knowledge, the only other author who has focused on this comparison is Kimmel (2003), who derived the asymptotic properties of estimates of risk premia on CSR–GLS unit-beta portfolios and on maximum-correlation portfolios. Kimmel's work differs from ours because he focused on asymptotics only, for the case of Gaussian iid returns, and did not address issues of size and power of the test statistics.

Also related to the present work are the work of Asgharian (2004) and Asgharian and Hansson (2005). Asgharian showed how testing a multi-beta model in its LFM or LFM* formulation can lead to different conclusions. In addition, he discussed the "portfolio" method of estimating factor realizations, in which assets with a high loading on a factor receive a positive weight, and assets with a low loading on a factor receive a negative weight (see, e.g., Chan, Karceski, and Lakonishok 1998; Fama and French 1993). Asgharian's work differs from ours because he did not study the theoretical relation between the LFM and LFM* representations and did not compare the small-sample properties of risk premium estimates and test statistics for the two formulations. Asgharian and Hansson considered the effects of replacing the original factors in the LFM with the factor realizations estimated using the portfolio method and an equally weighted scheme. Their work differs from ours because they focused on the implications of the portfolio method, rather than on the properties of the LFM and LFM* representations.

3. THEORETICAL RESULTS

3.1 LFM and LFM*

Result 1 establishes the relation between the two representations.

Result 1. Mispricing in the LFM* representation relates to mispricing in the LFM representation as follows:

$$\alpha^\star = [I - (\beta^\star)^\top (\gamma^\star)^\top]\alpha, \tag{4}$$

where $[I - (\beta^\star)^\top (\gamma^\star)^\top]$ is of rank $N - K$ with null space given by $\beta^\top \delta$, δ being a generic real-valued $N \times 1$ vector.

Thus, if $\alpha = 0$, then $\alpha^\star = 0$. On the other hand, if $\alpha^\star = 0$, then α is not necessarily 0, because the $N \times N$ matrix $[I - (\beta^\star)^\top (\gamma^\star)^\top]$ is of rank $N - K$. Yet the only case in which $\alpha^\star = 0$ and $\alpha \neq 0$ is when $\alpha = \beta^\top \delta$, but this means that the multi-beta model holds, although with parameters $\lambda + \delta$. Thus we conclude that $\alpha^\star = 0$ is a necessary and sufficient condition for $\alpha = 0$.

Result 2 obtains the relation between risk premia in the two formulations.

Result 2. We have

$$\lambda^\star = (\gamma^\star)^\top \alpha + \Sigma_{y^\star y^\star} \Sigma_{yy}^{-1} \lambda, \tag{5}$$

where $\lambda^\star = (\gamma^\star)^\top E(r_t)$.

Thus a sufficient condition for the risk premia in the two formulations to be the same is that $\alpha = 0$ and that the factors are traded (in which case $\Sigma_{y^\star y^\star} = \Sigma_{yy}$).

3.2 Cross-Sectional Regressions, Mimicking Portfolios, and Sharpe Ratios

The standard approach of running CSRs of average excess returns on betas leads to the risk-premium coefficients

$$\tilde{\lambda} = (\beta W \beta^\top)^{-1} \beta W E(r_t). \tag{6}$$

The CSR coefficients minimize the quadratic form

$$[E(r_t) - \beta^\top \tilde{\lambda}]^\top W [E(r_t) - \beta^\top \tilde{\lambda}]. \tag{7}$$

Two common choices for the weighting matrix W are $W = I$ and $W = \Sigma_{ee}^{-1}$. The first choice corresponds to an OLS regression; the second, to a GLS regression. Alternatively, we can choose $W = \Sigma_{rr}^{-1}$.

Interestingly, we obtain the same CSR coefficients for $W = \Sigma_{rr}^{-1}$ as for $W = \Sigma_{ee}^{-1}$. The intuition for this result is straightforward: Minimizing the total portfolio variance with a constraint on systematic risk is equivalent to minimizing the idiosyncratic portfolio variance with a constraint on systematic risk. This result proves to be useful when we consider an alternative interpretation of the CSR coefficients. First, we recognize that the $K \times N$ matrix

$$\tilde{\gamma}^\top = (\beta W \beta^\top)^{-1} \beta W \tag{8}$$

is a matrix of K sets of N portfolio weights. Each portfolio has unit beta with respect to the chosen factor and zero betas with respect to all other factors included in the LFM,

$$\tilde{\gamma}^\top \beta^\top = (\beta W \beta^\top)^{-1} \beta W \beta^\top = I. \tag{9}$$

On the other hand, it is straightforward to show how the CSR portfolio weights correspond to the solution of the minimization problem (see, e.g., Litzenberger and Ramaswamy 1979; Huberman et al. 1987)

$$\begin{aligned} &\min_{\gamma_k} \gamma_k{}^\top W^{-1} \gamma_k, \\ &\text{s.t.} \quad \beta \gamma_k = s_k, \end{aligned} \tag{10}$$

where γ_k is the $N \times 1$ vector of portfolio weights and s_k is a vector with the kth element equal to 1 and 0's elsewhere.

Thus, when $W = I$, the CSR portfolio weights minimize the length of the vector of portfolio weights, subject to the unit-beta constraint. When $W = \Sigma_{rr}^{-1}$, the CSR portfolio weights minimize the variance of the mimicking portfolio returns, subject to the unit-beta constraint. This second set of portfolio weights is equal to the set of portfolio weights implicit in the CSR–GLS coefficients $(\beta\Sigma_{ee}^{-1}\beta^\top)^{-1}\beta\Sigma_{ee}^{-1}E(r_t)$.

Note that the maximum-correlation portfolios are proportional to the solution of the minimization of $\gamma_k^\top\Sigma_{rr}\gamma_k$ with respect to γ_k, subject to the *single* constraint $\Sigma_{y_k r}\gamma_k = 1$. Thus both unit-beta CSR–GLS portfolios and maximum-correlation portfolios minimize the variance of the portfolio returns subject to constraints on covariances. The difference is that the unit-beta portfolios are subject to constraints on covariances with *all* of the other factors, whereas the maximum-correlation portfolios are subject only to a single constraint on the covariance with the factor being tracked.

One additional property of the unit-beta portfolio weights is that for $W = \Sigma_{rr}^{-1}$ (or $W = \Sigma_{ee}^{-1}$), and, in the case of a single factor, the weights are proportional to the maximum-correlation mimicking-portfolio weights; in fact,

$$\begin{aligned}\tilde{\gamma}^\top &= (\beta\Sigma_{rr}^{-1}\beta^\top)^{-1}\beta\Sigma_{rr}^{-1}\\ &= (\Sigma_{yr}\Sigma_{rr}^{-1}\Sigma_{ry}/\sigma_y^2)^{-1}\Sigma_{yr}\Sigma_{rr}^{-1}.\end{aligned} \tag{11}$$

Because $(\Sigma_{yr}\Sigma_{rr}^{-1}\Sigma_{ry}/\sigma_y^2)$ in the one-factor case is a scalar, $\tilde{\gamma}^\top$ here is proportional to $\Sigma_{yr}\Sigma_{rr}^{-1} = (\gamma^\star)^\top$. In the general case of K factors, on the other hand, the unit-beta mimicking-portfolio weights $\tilde{\gamma}$ are *linear combinations* of the maximum-correlation mimicking-portfolio weights $\gamma^\star$.

Note, however, that the foregoing discussion does not mean that in the presence of a single nontraded factor, the properties of the estimates of the risk premium on the CSR–GLS portfolio and on the maximum correlation portfolio are the same. In fact, although the maximum-correlation portfolios are constructed subject to a covariance constraint, the CSR–GLS portfolios are constructed subject to a *beta* constraint. In general, covariance and beta estimates have different properties that lead to different properties of the risk premium estimates. For example, in the case of a noisy factor, covariance estimates are consistent, whereas beta estimates are not (more on this in Sec. 6.2). This leads to different behaviors of the two risk premium estimates, even when only one factor is at work.

Result 3 shows how the alphas associated with the CSR premia equal the alphas of the LFM* representation when $W = \Sigma_{ee}^{-1}$ (or $W = \Sigma_{rr}^{-1}$).

Result 3. When $W = \Sigma_{ee}^{-1}$, $\alpha^\star = \tilde{\alpha}$.

Result 4 discusses the relation between Sharpe ratios attainable from different mimicking portfolios.

Result 4. When $W = \Sigma_{ee}^{-1}$, then the maximum squared Sharpe ratio from investment in the unit-beta mimicking portfolios equals the maximum squared Sharpe ratio from investment in the maximum-correlation portfolios.

This means that in the case of a single factor, the Sharpe ratios on the two types of mimicking portfolios are the same.

3.3 Mimicking Portfolios and Characteristic-Based Models

Our analysis focuses on multi-beta models, that is, risk-based models of expected returns. Yet the mimicking-portfolio approach allows us to revisit the formulation of characteristic-based models in terms of exposure to characteristic-sorted portfolios (e.g., Fama and French 1993). Indeed, Result 5 establishes the "observational equivalence" between characteristic-based and risk-based models; we can always formulate a characteristic-based model as a multi-beta model with traded factors.

Result 5. Assume that expected returns are linearly related to security characteristics

$$E(r_t) = H^\top\delta, \tag{12}$$

where H is a full-rank $K \times N$ matrix of characteristics and δ is a $K \times 1$ vector. We can always construct a set of K minimum-variance, unit-characteristic portfolios with returns $\bar{y}_t$, such that

$$E(\bar{y}_t) = \delta \tag{13}$$

and

$$\Sigma_{\bar{y}\bar{y}}^{-1}\Sigma_{\bar{y}r} \equiv \bar{\beta} = H. \tag{14}$$

One implication of Result 5 is that securities with the same characteristics also have the same exposures to unit-characteristic portfolio returns. This may explain why, after controlling for characteristics, average portfolio returns are not related to their exposures to the FF factors (Daniel and Titman 1997). Moreover, for Result 5 to hold, the unit-characteristic portfolios must have weights that are the coefficients of a GLS-style CSR of expected returns on characteristics. The FF factors are not constructed in such a way, which may explain why, even after controlling for the exposure to the FF factors, the cross-section of average stock returns still exhibits significant characteristic-driven patterns (Brennan et al. 1998).

Note that the observational equivalence between characteristic- and risk-based explanations also is demonstrated in the theoretical analyses of Ferson (1996) and Ferson et al. (1999). Ferson's result differs from ours, because he assumed the orthogonality between security characteristics and risk exposures, together with either a factor structure in returns or a time variation in the δ coefficients; for a comparison, we need only assume that Σ_{rr} is invertible. Ferson et al.'s analysis also differs from ours, because they considered unit-characteristic portfolios constructed by means of an *OLS-style* CSR and assumed a covariance matrix of returns proportional to the identity matrix.

4. MOMENT CONDITIONS

Here we illustrate the moment conditions imposed in estimating the two formulations of the linear factor model.

4.1 LFM

Let z_t denote the data and θ denote the parameter vector. The moment conditions, $E[f(z_t,\theta)] = 0$, for the LFM representation are given by

$$E(r_t - \beta_0 - \beta^\top y_t) = 0, \tag{15}$$

$$E[(r_t - \beta_0 - \beta^\top y_t) \otimes y_t] = 0, \tag{16}$$

and

$$E(r_t - \beta^\top\lambda) = 0. \tag{17}$$

Following Hansen (1982), we set a linear combination of the moment conditions equal to 0, that is, $a\hat{E}[f(z_t,\hat{\theta})]=0$, where

$$a=\begin{bmatrix} I_{N(K+1)} & 0_{N(K+1),N} \\ 0_{K,N(K+1)} & \hat{\beta}W \end{bmatrix}. \tag{18}$$

We consider two weighting matrices, $W=I$ and $W=\hat{\Sigma}_{ee}^{-1}$.

The resulting λ estimates are given by

$$\hat{\lambda}=(\hat{\beta}W\hat{\beta}^\top)^{-1}\hat{\beta}W\hat{E}(r_t). \tag{19}$$

For $W=\hat{\Sigma}_{ee}^{-1}$ (or $W=\hat{\Sigma}_{rr}^{-1}$), the resulting λ estimates coincide with those of the *optimal* GMM estimator (in the iid case),

$$a=\left.\frac{\partial\hat{E}[f(z_t,\theta)]^\top}{\partial\theta}\right|_{\theta=\hat{\theta}}\hat{\Sigma}_{ff}^{-1}, \tag{20}$$

conditional on $\hat{\beta}$.

Relative to the approach of using the optimal weighting matrix, as used by Jagannathan and Wang (2002) and Kan and Zhou (1999, 2001a), we see two main advantages to our approach. First, all estimates are obtained in closed-form, which is especially useful in the simulation exercise; second, the estimates of the risk premium parameters λ are the same as those obtained with the traditional CSR approach.

4.2 LFM*

In the LFM* formulation, the moment conditions are given by

$$E[y_t-\gamma_0^\star-(\gamma^\star)^\top r_t]=0, \tag{21}$$

$$E\{[y_t-\gamma_0^\star-(\gamma^\star)^\top r_t]\otimes r_t\}=0, \tag{22}$$

$$E[(\gamma^\star)^\top r_t-\lambda^\star]=0, \tag{23}$$

$$E[r_t-\alpha^\star-(\beta^\star)^\top y_t^\star]=0, \tag{24}$$

and

$$E\{[r_t-\alpha^\star-(\beta^\star)^\top y_t^\star]\otimes y_t^\star\}=0. \tag{25}$$

In this case, the estimates are obtained by exactly identified GMM ($a=I$).

One potential disadvantage of this second approach is that we now have a total of $K(N+1)+K+N(K+1)$ parameters to estimate, compared with $K+N(K+1)$ parameters for the LFM representation. Thus, one concern in the simulation analysis will be to see whether the larger number of estimated parameters affects the small-sample properties of the estimates.

5. ASYMPTOTICS

For both representations, the asymptotic covariance matrix of the estimates is given by

$$\mathrm{cov}(\hat{\theta})=\frac{1}{T}(ad)^{-1}aSa^\top[(ad)^{-1}]^\top, \tag{26}$$

where

$$d=\left.\frac{\partial\hat{E}[f(z_t,\theta)]}{\partial\theta'}\right|_{\theta=\hat{\theta}} \tag{27}$$

and

$$S=\sum_{j=-\infty}^{\infty}\hat{E}[f(z_t,\hat{\theta})f(z_{t-j},\hat{\theta})^\top]. \tag{28}$$

In the LFM* representation, the S-matrix does not have full rank, $K+KN+K+N+NK$, but instead has rank equal to $(K+KN+K+N+NK)-K-K^2$. The reason for this is that we can construct K linear combinations of the excess returns that exactly replicate the excess returns on the maximum-correlation mimicking portfolios. Thus for K linear combinations of the excess returns, the residuals of the regression on the excess returns on the maximum-correlation mimicking portfolios are identically 0. This also means that the K^2 products of these residuals with the excess returns on the maximum-correlation mimicking portfolios are 0. One way to get around this problem is to differentiate the set of the test assets from the set of assets used to construct the mimicking portfolios; indeed, this is the approach used in the maximum likelihood estimation of Breeden et al. (1998). Note, however, that in our GMM setting the singularity of S does not present a problem in estimation, because it does not need to be inverted. Moreover, the inversion of *portions* of the covariance matrix of the estimates is indeed possible; for example, we can invert the covariance matrix of the $\alpha^\star$ estimates to compute Wald-style statistics. Note that a noninvertible S matrix is not unique to our setting; Kan and Zhou (2001b) also obtained a noninvertible S matrix in the case of exactly identified GMM spanning tests using the stochastic discount factor approach. Peñaranda and Sentana (2004), in contrast, considered the case of *overidentified* GMM spanning tests when S is not invertible.

The covariance matrix of the estimates is used to test the joint significance of the $\alpha^\star$ estimates in the LFM* representation. In the LFM representation, on the other hand, the α estimates are the averages of the last set of K moment conditions. Thus the joint significance of the α estimates is tested based on the covariance matrix

$$\mathrm{cov}\{\hat{E}[f(z_t,\hat{\theta})]\}=\frac{1}{T}[I-d(ad)^{-1}a]S[I-d(ad)^{-1}a]^\top. \tag{29}$$

6. SPECIAL CASES

6.1 Traded Factors

When a subset y_{1t} of the factors y_t are traded, the analytical results of the previous sections, as well as the estimation approach, are changed slightly. This is because we introduce the restriction that the risk premium estimates for the traded factors are simply the time series averages of the traded factors (see, e.g., Shanken 1992). Let

$$r_{2t}\equiv r_t-\beta_1^\top y_{1t}, \tag{30}$$

where the β_1's are the coefficients of a projection of r_t onto the span of y_{1t}. In addition, let β_2 denote the coefficients of the projection of r_{2t} onto the augmented span of the nontraded factors and let $\gamma_2^\star$ denote the coefficients of the projection of the nontraded factors onto the augmented span of r_{2t}.

The LFM representation leads to the mimicking-portfolio coefficients

$$\tilde{\gamma}_2^\top=(\beta_2W\beta_2^\top)^{-1}\beta_2W \tag{31}$$

and the risk premia

$$\tilde{\lambda}_2 = (\beta_2 W \beta_2^\top)^{-1} \beta_2 W E(r_{2t}). \tag{32}$$

The LFM* representation leads to the risk premia

$$\lambda_2^\star = (\gamma_2^\star)^\top E(r_{2t}). \tag{33}$$

Thus all of the analytical results of the previous sections still hold, but with r_t replaced by r_{2t}, β replaced by β_2, y_t replaced by y_{2t}, and $\gamma^\star$ replaced by $\gamma_2^\star$.

For estimation, the same modifications are used. For the LFM specification, we have the orthogonality condition

$$E(r_t - \beta_1^\top y_{1t}) = 0, \tag{34}$$

combined with suitably modified versions of (15)–(17). In constructing the a matrix, $\hat{\beta}_2$ replaces $\hat{\beta}$, whereas W can equal either I or $\hat{\Sigma}_{ee}^{-1}$. For the LFM* formulation, we have the orthogonality condition (34) combined with suitably modified versions of (21)–(25).

6.2 Noisy Factors

It is straightforward to show that in the case of a single noisy factor, the CSR–OLS estimator leads to inconsistent estimates of the risk premium λ (see Kan and Zhang 1999 for a similar discussion). Assume that $\hat{y}_t = y_t + \epsilon_t$, where $E(y_t\epsilon_t) = 0$ and $E(r_t\epsilon_t) = 0$.

We have

$$\mathrm{P}\lim \hat{\beta} = \frac{\mathrm{var}(y_t)}{\mathrm{var}(\hat{y}_t)}\beta. \tag{35}$$

Consider the CSR λ estimator for $W = I$. We have

$$\mathrm{P}\lim \hat{\lambda} = \frac{\mathrm{var}(\hat{y}_t)}{\mathrm{Var}(y_t)}\lambda. \tag{36}$$

In other words, the attenuation bias in the β estimates translates into an upward bias in the λ estimates. On the other hand, it is immediate to show that the $\gamma^\star$ and $\lambda^\star$ estimates are still consistent.

7. DATA

We calibrate the simulation experiment using monthly data from 1959:03–2002:11 (525 observations). We consider two sets of test assets: 10 size-sorted portfolios [value-weighted (VW) NYSE–AMEX–NASDAQ decile portfolios, from CRSP] and the 25 FF portfolios (from French's web page). We consider the following three factors: log real per-capita consumption growth (nondurable goods and services, from Global Insight), the VW NYSE–AMEX–NASDAQ index return in excess of the risk-free rate, (from CRSP), and the change in the dividend yield on the composite S&P500 (from Global Insight). As a proxy for the risk-free rate, we use the nominal 1-month Treasury bill rate from CRSP.

8. BOOTSTRAP EXPERIMENT

Here we provide a detailed description of the bootstrap experiment.

Return-Generating Process. We simulate a one-factor C–CAPM economy and a two-factor I–CAPM economy. In the one-factor economy, the factor is the change in the log of real per-capita consumption, y_{ct},

$$r_t = \alpha + \beta_c^\top [y_{ct} - E(y_{ct}) + \lambda_c] + e_t. \tag{37}$$

In the two-factor economy, the traded factor is the excess return on the market, y_{mt}, and the nontraded factor is the change in the dividend yield, y_{dt},

$$r_t = \alpha + \beta_m^\top y_{mt} + \beta_d^\top [y_{dt} - E(y_{dt}) + \lambda_d] + e_t. \tag{38}$$

Calibration. We estimate betas, alphas, and risk premia for the LFM using the approach described in Sections 4 and 6, where we set $W = I$. For both economies, we consider 2 choices of reference parameters, corresponding to 2 choices of assets: the 10 size portfolios and the 25 FF portfolios.

Bootstrap. We jointly bootstrap the realizations of the factors and the realizations of the residuals of the LFM. This means that we are sampling from a distribution in which the factors and the residuals are orthogonal but may display higher-order dependence. Because we are drawing from the empirical distribution of factors and residuals, we are not imposing distributional assumptions; in particular, we are not imposing normality. We set $\alpha = 0$ in all simulations, with the exception of the analysis of power, where we set α to its estimated value. We consider two lengths of the data set: a full sample of 525 observations and a shorter sample of 240 observations (the first 240 observations in our sample). We implement both an iid and a block bootstrap (blocks of three monthly observations; see Cochrane 2001).

Empirical Analysis of Simulated Data. For each string of simulated data, we estimate the parameters of the LFM based on the two weighting matrices $W = I$ and $W = \Sigma_{ee}^{-1}$. Estimates of the parameters of the LFM* are obtained by an exactly identified GMM and do not require choosing a weighting matrix. Asymptotic statistics for each simulation are obtained assuming no serial correlation (for the iid bootstrap) or serial correlation of order three (for the block bootstrap), with Newey and West (1987) adjustment.

9. RESULTS

9.1 Estimates of Economic Risk Premia

9.1.1 C–CAPM. Table 2 reports results for the one-factor (C–CAPM) case for the λ and $\lambda^\star$ estimates. The table presents population values for the risk premia, average risk premium estimates, biases, and RMSEs. For ease of comparison, we report risk premia as percentages of the population standard deviation of the corresponding mimicking-portfolio excess return. Given the results of Section 3, this means that the reference values for the CSR–GLS estimates and the $\lambda^\star$ estimates are the same, whereas the reference values for CSR–OLS estimates and CSR–GLS/$\lambda^\star$ estimates differ. This is because the same population risk premium is divided by different population values of the standard deviations of the mimicking-portfolio excess returns. Biases and RMSEs are reported as percentages of the population standard deviation of the corresponding mimicking-portfolio excess return, and also as percentages of the population absolute values of the risk premia (in parentheses).

Table 2. λ versus λ*

	Finite-sample properties of λ								Finite-sample properties of λ*			
	Full sample				240 months				Full sample		240 months	
	Size		FF		Size		FF		Size	FF	Size	FF
	$W=I$	$W=sse$	$W=I$	$W=sse$	$W=I$	$W=sse$	$W=I$	$W=sse$	$W=I$	$W=sse$	$W=I$	$W=sse$
λ	11.57	15.86	11.73	17.96	11.57	15.86	11.73	17.96	15.86	17.96	15.86	17.96
$\hat{\lambda}$	12.31	12.98	12.12	11.17	13.08	10.44	12.43	7.35	15.75	16.99	15.59	16.01
Bias	.74	−2.88	.39	−6.79	1.51	−5.42	.70	−10.62	−.11	−.97	−.27	−1.95
	(6.40)	(−18.16)	(3.32)	(−37.81)	(13.05)	(−34.17)	(5.97)	(−59.08)	(−.69)	(−5.40)	(−1.70)	(−10.86)
RMSE	6.10	5.28	5.48	7.90	11.03	8.11	8.63	11.72	5.82	6.38	9.32	11.31
	(52.72)	(33.29)	(46.72)	(43.99)	(95.33)	(51.13)	(73.57)	(65.26)	(36.70)	(35.52)	(58.76)	(62.97)
t_{mean}	−.12	−.78	−.12	−1.83	−.18	−1.08	−.18	−2.39	−.12	−.13	−.12	−.20
t_{std}	.97	1.07	.97	1.17	.93	1.10	.93	1.24	.97	.92	.92	.89
$\frac{AsySE}{EmpStd}$	.97	.98	.98	.99	.96	.95	1.01	.94	1.05	1.09	1.09	1.11

NOTE: Here we simulate a one-factor C–CAPM economy 10,000 times under the null (with the α parameters set to 0). Parameter estimates of the LFM are obtained using GMM with $W=I$. In the bootstrap, we simulate jointly the factor and the residuals from the estimated LFM. We consider two lengths of the data set: the full sample of 525 observations (full sample) and a shorter sample of 240 observations (240 months). We consider 2 choices of assets: the 10 size portfolios (size) and the 25 FF portfolios (FF). When investigating the properties of estimates of the LFM, we consider estimates based on the two weighting matrices, $W=I$ and $W=see=\hat{\Sigma}_{ee}^{-1}$. Estimates of the parameters of the LFM* are obtained by exactly identified GMM. Asymptotic statistics are obtained assuming no serial correlation. The first row reports the population value of the risk premium as a percentage of the population standard deviation of the excess return on the corresponding mimicking portfolio. The second row reports the average value across simulations of the risk premium as a percentage of the population standard deviation of the excess return on the corresponding mimicking portfolio. The third and fourth rows report absolute (percentage) bias and absolute (percentage) RMSEs. The fifth and sixth rows report the means and standard deviation of the t-ratios across simulations. The seventh row reports the ratio between average asymptotic standard errors across simulations and empirical standard deviation.

Table 2 also reports averages and standard deviations of asymptotic t-ratios on the risk premium estimates. We compute t-ratios using the population risk premium as the reference value. Finally, the table reports the ratios between the average asymptotic standard error on the risk premium estimate and the standard deviation of the estimate across simulations.

LFM. First, consider the results for the estimates of λ. OLS estimates are biased away from 0, with biases ranging from .39 to 1.51 and percentage biases ranging from 3.32% to 13.05%. In contrast, GLS estimates are biased toward 0, with biases ranging from −10.62 to −2.88 and percentage biases ranging from −59.08% to −18.16%. In all instances, the absolute percentage bias is larger for the GLS estimates than for the OLS estimates. RMSEs are substantial for both estimators, ranging from 5.48 to 11.03 (46.72–95.33% of the true value) for the OLS estimator and from 5.28 to 11.72 (33.29–65.26% of the true value) for the GLS estimator. Percentage RMSEs are always larger for the OLS estimator than for the GLS estimator. As would be expected, reducing the length of the sample increases biases and RMSEs. Using the FF rather than the size portfolios leads to smaller (larger) biases and RMSEs for the OLS (GLS) estimates. In this exercise, as well as in the other exercises that use nonoverlapping data, going from the iid bootstrap to the block bootstrap (not reported in the table, but available in a separate appendix) makes little difference.

Asymptotic t-ratios have means substantially different from 0, ranging from −2.39 to −.12; the biases are always more pronounced for the GLS estimator than for the OLS estimator. The standard deviations of the asymptotic t-ratios also can deviate substantially from the theoretical value of 1, ranging from .93 to 1.24. The deviations are always more pronounced for the GLS estimator than for the OLS estimator. Finally, the ratios between asymptotic and empirical standard deviations are mainly <1 for both choices of weighting matrix, ranging from .95 to 1.01, with no clear superiority of one estimator over the other.

*LFM**. Results for the estimates of $\lambda^\star$ can be directly compared with the results for the CSR–GLS estimates, because the population Sharpe ratios for the two mimicking portfolios are the same. There is an advantage in using the LFM* formulation in terms of bias: Absolute biases of the $\lambda^\star$ estimates are always substantially smaller than the biases of the corresponding CSR–GLS estimates. For example, for the 240-month sample FF portfolios, the bias for the CSR–GLS estimator is −10.62, whereas that for the $\lambda^\star$ estimator is only −1.95. Note that under normality, the estimates of $\lambda^\star$ should be unbiased (Dickey 1967); Thus, deviations from normality (e.g., skewness and excess kurtosis) in our bootstrap data are responsible for the biases that we document. RMSEs are roughly of the same magnitude as those for the GLS estimates.

Percentage biases for the $\lambda^\star$ estimator range from −10.86% to −.69% and are higher or lower than the percentage biases for the CSR–OLS estimates depending on the choice of test assets. Percentage RMSEs for the $\lambda^\star$ estimator range from 35.52% to 62.97%, again higher or lower than the corresponding quantities for the CSR–OLS estimates depending on the choice of test assets. Thus the performance of the $\lambda^\star$ and CSR–OLS estimates in terms of bias and RMSE is roughly comparable.

T-ratios in the LFM* are negatively biased. Biases are similar to those for the OLS estimates and much less pronounced than those for the GLS estimates. The standard deviations of the t-ratios are consistently <1, also similar to the OLS estimates. Finally, the ratios between asymptotic standard errors and empirical standard errors are consistently >1, differing from the OLS and GLS estimates.

9.1.2 I–CAPM. We also performed our simulation analysis for the multifactor (I–CAPM) case. Results for this case are not reported in Table 2 but are available in a separate appendix. In this case we focus on the risk premia on the nontraded factor: changes in the dividend yield.

Consider first the simulation evidence for the LFM formulation. As in the one-factor case, the GLS estimates are biased toward 0 (the bias is positive, whereas the true value of the parameter is negative); in contrast, the bias in the OLS estimates changes sign depending on the set of assets. Again as in the one-factor case, percentage biases are more pronounced for the GLS estimates. Percentage RMSEs are substantial for both sets of estimates and of roughly similar magnitude. Consistent with the one-factor case, RMSEs are always larger for the OLS estimates than for the GLS estimates. Biases in t-ratios are now consistently positive for both estimators; as in the one-factor case, they are more pronounced for the GLS estimator. As in the one-factor case, standard deviations of t-ratios are <1 for the OLS estimator and >1 for the GLS estimator. Finally, the ratios between asymptotic and empirical standard errors of the estimates are mainly <1, also consistent with the one-factor case.

Consider now the simulation evidence for the LFM* formulation. As in the one-factor case, $\lambda^\star$ estimates are biased toward 0; thus biases are now positive. Again, biases are consistently smaller than for the corresponding GLS estimates. RMSEs are again comparable to those for the GLS estimates. Comparing $\lambda^\star$ estimates to CSR–OLS estimates, we see that the percentage biases are always more pronounced for the $\lambda^\star$ estimates. Percentage RMSEs, in contrast, are similar for the two estimators, with their relative size depending on the choice of test assets, just as in the one-factor case. Biases in t-ratios are positive, but less pronounced than for the CSR–OLS estimator and comparable to that for the CSR–GLS estimator, as in the one-factor case. The volatilities of the t-ratios are now consistently >1, and when comparing departures from the true value of 1, no one estimator clearly outperforms the other two estimators. Finally, the ratios between asymptotic and empirical standard errors of the estimates are always >1 and tend to depart from the reference value of 1, more so than the other two estimators, consistent with the one-factor case.

9.1.3 Summary and Discussion. In summary, the various scenarios herein considered paint a fairly consistent picture of the advantages and disadvantages of the various estimators. In terms of bias, the $\lambda^\star$ estimates always do better than the CSR–GLS estimates, whereas they often do worse than the CSR–OLS estimates. In terms of RMSE, the $\lambda^\star$ estimates always do better than the CSR–OLS estimates and perform similarly to the CSR–GLS estimates. In terms of average t-ratios, the $\lambda^\star$ estimates always do better than the CSR–GLS estimates and perform similarly to the CSR–OLS estimates.

We can compare the foregoing results with the results obtained by Chen and Kan (2006) in a similar setting. Chen and Kan analytically derived the small-sample bias and standard deviation of CSR estimates for the OLS and GLS cases under the assumption of iid Gaussian returns and factor, and they performed a simulation exercise for nonnormal (Student t) returns and factor. They considered the case in which the factor is calibrated to consumption growth and considered both size-sorted and size- and value-sorted test portfolios. Several of their results are consistent with ours. They found that absolute biases were larger for the GLS estimates than for the OLS estimates, that the GLS estimates were consistently biased toward 0, and that the sign of the bias for the OLS estimates could be either positive or negative, depending on the calibration. They found that the GLS estimates were less volatile than the OLS estimates, which is consistent with our finding that the RMSEs are smaller for the GLS estimates.

9.2 Changing Betas

In this section we further investigate the nature of the difference in results between the LFM and LFM* formulations. We replicate the one-factor case for the scenario with the full sample and size portfolios, but with the betas with respect to consumption growth scaled by a factor m. The variances of the idiosyncratic components are adjusted so that the variances of returns remain the same. Thus, as betas increase in absolute value, the percentages of return variance explained by the single-index model also increase.

The results, reported in Table 3, clearly show how the biases in the λ estimates tend to 0 as the consumption betas increase from $1/3$ to 3 times the original value. This pattern is especially pronounced for the CSR–GLS estimate, the bias of which goes from -11.22 to $-.43$ (from -68.62% to -3.36%). On the other hand, the bias in the $\lambda^\star$ estimates remains remarkably stable (between $-.32$ and $-.11$; -1.96% and $-.69\%$ of

Table 3. Changing betas: λ_1 ($W=I$), λ_2 ($W=\hat{\Sigma}_{ee}^{-1}$) versus $\lambda^\star$

	$m=1/3$	$m=1/2$	Base	$m=2$	$m=3$
λ_1	11.75	11.72	11.57	11.00	10.21
$\hat{\lambda}_1$	10.11	12.98	12.31	11.19	10.30
λ_2	16.35	16.27	15.86	14.48	12.81
$\hat{\lambda}_2$	5.13	8.46	12.98	13.62	12.38
$\lambda^\star$	16.35	16.27	15.86	14.48	12.81
$\hat{\lambda}^\star$	16.03	16.07	15.75	14.36	12.67
$Bias_{\lambda_1}$	-1.64	1.26	.74	.19	.09
	(-13.96)	(10.75)	(6.40)	(1.73)	(.88)
$Bias_{\lambda_2}$	-11.22	-7.81	-2.88	$-.85$	$-.43$
	(-68.62)	(-48.00)	(-18.16)	(-5.87)	(-3.36)
$Bias_{\lambda^\star}$	$-.32$	$-.20$	$-.11$	$-.11$	$-.14$
	(-1.96)	(-1.23)	$(-.69)$	$(-.76)$	(-1.10)
$RMSE_{\lambda_1}$	13.84	11.18	6.10	4.74	4.55
	(117.78)	(95.39)	(52.72)	(43.09)	(44.56)
$RMSE_{\lambda_2}$	11.81	8.88	5.28	4.41	4.36
	(72.23)	(54.58)	(33.29)	(3.46)	(34.04)
$RMSE_{\lambda^\star}$	11.75	8.57	5.82	4.74	4.46
	(71.87)	(52.67)	(36.70)	(32.73)	(34.82)
$(\frac{AsySE}{EmpStd})_{\lambda_1}$	1.19	1.06	.97	.99	.99
$(\frac{AsySE}{EmpStd})_{\lambda_2}$	.85	.93	.98	.98	.99
$(\frac{AsySE}{EmpStd})_{\lambda^\star}$	1.12	1.10	1.05	1.00	1.00

NOTE: Here we simulate a C–CAPM economy 10,000 times under the null (with the α parameters set to 0). Parameter estimates of the LFM are obtained using GMM with $W=I$ for different values of β. We denote these different values of β with $\beta_{adj}=m*\beta$, where m is a scalar. In the bootstrap, we jointly simulate the factor and the residuals from the estimated LFM. The adjusted returns that we construct have the same volatility as the original returns. The columns report simulation results for the cases of $m=1/3, 1/2, 1, 2, 3$. We investigate the properties of estimates of the LFM using the full sample of 525 observations, the 10 size portfolios (size), and the 2 weighting matrices, $W=I$ and $W=\hat{\Sigma}_{ee}^{-1}$. Estimates of the parameters of the LFM* are obtained by exactly identified GMM. Asymptotic statistics are obtained assuming no serial correlation. We report the population value of the risk premium as a percentage of the population standard deviation of the excess return on the corresponding mimicking portfolio, the average value across simulations of the risk premium as a percentage of the population standard deviation of the excess return on the corresponding mimicking portfolio, absolute (percentage) bias and absolute (percentage) RMSEs, mean and standard deviation of the t-ratios across simulations, and the ratio between average asymptotic standard errors across simulations and empirical standard deviation.

the population values) across values of m. As in the scenarios described in the previous section, the percentage RMSEs of the CSR–GLS and $\lambda^\star$ estimates are comparable, whereas the percentage RMSEs of the CSR–OLS estimator are consistently higher. RMSEs also decrease monotonically as the consumption betas increase (in absolute value), although they are still substantial for $m = 3$, ranging from 34.04% to 44.56%.

In summary, we conclude that the discrepancies in results across the different risk premium estimators tend to disappear as the factor explains more of the variance in returns.

9.3 Noisy Factor

Table 4 gives results for the one-factor case when the factor is observed with noise. We focus on one scenario only: full sample and size portfolios. Following Kan and Zhou (1999), we consider noise with standard deviations equal to .5, 1, 1.5, and 2 times the standard deviation of the factor.

A setting in which the issue of measurement error in the factor arises naturally is that of tests of heterogenous-agent models (see, e.g., Brav, Constantinides, and Geczy 2002; Jacobs and Wang 2004). In these studies, the realizations of the factors are the moments of the cross-sectional distribution of consumption growth. Because the cross-sections used in these studies are relatively small, the cross-sectional moments can be estimated with a great deal of noise.

The table presents reference values of the risk premia, biases, RMSEs, and ratios between average asymptotic standard errors and empirical standard deviations of the estimates. In this setting, the results dramatically favor the LFM* formulation over both the CSR–OLS and CSR–GLS approaches. Bias and RMSE monotonically increase with noise for the CSR λ estimates. For example, when the noise is twice as large as the signal, the biases for the CSR–OLS and CSR–GLS estimates are 48.95 and 18.54 (423.08% and 116.90%) and RMSEs are as high as 79.97 and 26.28 (691.18% and 165.70%). In contrast the bias for the $\lambda^\star$ estimates never exceeds .13 (.82%) in absolute value, with a maximum RMSE of 9.10 (57.38%).

Table 4. Noisy factor: λ_1 ($W = I$), λ_2 ($W = \hat{\Sigma}_{ee}^{-1}$) versus $\lambda^\star$

	No noise	$c = .50$	$c = 1.0$	$c = 1.50$	$c = 2.0$
λ_1	11.57	11.57	11.57	11.57	11.57
$\hat{\lambda}_1$	12.31	15.50	25.90	42.56	60.52
λ_2	15.86	15.86	15.86	15.86	15.86
$\hat{\lambda}_2$	12.98	15.43	21.44	28.35	34.40
$\lambda^\star$	15.86	15.86	15.86	15.86	15.86
$\hat{\lambda}^\star$	15.75	15.74	15.74	15.73	15.73
$Bias_{\lambda_1}$	.74	3.93	14.33	30.99	48.95
	(6.40)	(33.97)	(123.85)	(267.85)	(423.08)
$Bias_{\lambda_2}$	−2.88	−.43	5.58	12.49	18.54
	(−18.16)	(−2.71)	(35.18)	(78.75)	(116.90)
$Bias_{\lambda^\star}$	−.11	−.12	−.12	−.13	−.13
	(−.69)	(−.76)	(−.76)	(−.82)	(−.82)
$RMSE_{\lambda_1}$	6.10	8.87	21.69	46.62	79.97
	(52.72)	(76.66)	(187.47)	(402.94)	(691.18)
$RMSE_{\lambda_2}$	5.28	5.41	10.09	17.96	26.28
	(33.29)	(34.11)	(63.62)	(113.24)	(165.70)
$RMSE_{\lambda^\star}$	5.82	6.04	6.76	7.82	9.10
	(36.70)	(38.08)	(42.62)	(49.30)	(57.38)
$(\frac{AsySE}{EmpStd})_{\lambda_1}$	.97	.99	.99	1.04	1.10
$(\frac{AsySE}{EmpStd})_{\lambda_2}$	.98	1.00	1.01	1.00	.96
$(\frac{AsySE}{EmpStd})_{\lambda^\star}$	1.05	1.06	1.09	1.11	1.13

NOTE: Here we simulate a C–CAPM economy 10,000 times under the null (with the α parameters set to 0). Parameter estimates of the LFM are obtained using GMM with $W = I$. In the bootstrap, we jointly simulate the factor and the residuals from the estimated LFM. For each bootstrap replication, we add an iid Gaussian shock to the factor. The standard deviation of the shock is proportional (constant of proportionality equal to c) to the population standard deviation of the factor. The columns report simulation results for the case of no noise and the case of noise ($c = .5, 1, 1.5, 2.0$). We investigate the properties of estimates of the LFM using the full sample of 525 observations, the 10 size portfolios (size), and the 2 weighting matrices, $W = I$ and $W = \hat{\Sigma}_{ee}^{-1}$. Estimates of the parameters of the LFM* are obtained by exactly identified GMM. Asymptotic statistics are obtained assuming no serial correlation. We report the population value of the risk premium as a percentage of the population standard deviation of the excess return on the corresponding mimicking portfolio, the average value across simulations of the risk premium as a percentage of the population standard deviation of the excess return on the corresponding mimicking portfolio, absolute (percentage) bias and absolute (percentage) RMSEs, mean and standard deviation of the t-ratios across simulations, and the ratio between average asymptotic standard errors across simulations and empirical standard deviation.

9.4 Estimates of Betas

Here we analyze the finite-sample properties of β and $\beta^\star$. Note that the $\beta^\star$ estimates are subject to the errors-in-variables (EIV) problem arising from estimation of the weights of the maximum-correlation mimicking portfolio. We conduct simulations to assess the magnitude of this bias.

Table 5 reports results for the C–CAPM case. We report the same statistics as in Table 2, averaged across assets. For ease of comparison, both β and $\beta^\star$ estimates are scaled by the ratio between the population standard deviations of the factor and of the dependent variable; thus they have the dimension of correlation coefficients. Percentage biases and RMSEs are computed as average biases and RMSEs across assets divided by the population values of the betas, averaged across assets.

In the results for the estimates of β, biases are essentially nonexistent, although RMSEs are substantial (as high as 38.48%). Similarly unbiased on average, are the t-ratios. The volatilities of the t-ratios are close to 1, and the ratios between average the asymptotic standard errors and empirical standard errors are also close to 1.

In the results for the estimates of $\beta^\star$, biases are negative (the well-known "attenuation bias" associated with the EIV problem) and substantial, ranging from -29.67 to -12.18 (-53.30% to -17.53%). RMSEs also are substantial, as high as 31.63 (56.82%), although of the same order of magnitude as for the LFM specification. Finally, biases in t-ratios are negative and pronounced, and there are substantial discrepancies between small-sample and asymptotic volatilities of t-ratios.

We also perform our analysis for the case of the I–CAPM (results not reported in the table). In this case we focus on the betas for the nontraded factor. As in the single-factor case, biases and RMSEs are more pronounced for the $\beta^\star$ estimates than for the β estimates.

9.5 Size and Power of the Tests

We investigate, by simulation, the size and power properties of the Wald-style tests for the one-factor and two-factor models. Our main contribution is an analysis of the size and

Table 5. β versus $\beta^\star$

	Finite-sample properties of β				Finite-sample properties of $\beta^\star$			
	Full sample		240 months		Full sample		240 months	
	Size	FF	Size	FF	Size	FF	Size	FF
β	17.92	16.27	17.92	16.27	69.47	55.67	69.47	55.67
$\hat{\beta}$	17.80	16.26	17.70	16.29	57.29	37.00	46.50	26.00
Bias	−.12	−.01	−.22	.02	−12.18	−18.67	−22.97	−29.67
	(−.67)	(−.06)	(−1.23)	(−.12)	(−17.53)	(−33.54)	(−33.06)	(−53.30)
RMSE	4.22	4.24	6.25	6.26	17.73	21.69	27.88	31.63
	(23.55)	(26.06)	(34.88)	(38.48)	(25.52)	(38.96)	(40.13)	(56.82)
t_{mean}	−.01	−.02	−.01	−.02	−1.14	−1.83	−1.71	−2.99
t_{std}	1.02	1.02	1.03	1.03	1.15	1.31	1.31	1.72
$\frac{AsySE}{EmpStd}$	.98	.98	.98	.98	1.07	1.11	1.11	1.10

NOTE: Here we simulate a one-factor C–CAPM economy 10,000 times. Parameter estimates of the LFM are obtained using GMM with $W = I$. In the bootstrap, we jointly simulate the factor and the residuals from the estimated LFM. We consider 2 lengths of the data set: the full sample of 525 observations (full sample) and a shorter sample of 240 observations (240 months). We consider 2 choices of assets: the 10 size portfolios (size) and the 25 FF portfolios (FF). Estimates of the parameters of the LFM* are obtained by exactly identified GMM. Asymptotic statistics are obtained assuming no serial correlation. Both β and $\beta^\star$ estimates (average β's and $\beta^\star$'s across simulations and across assets) are scaled by the ratio between the population standard deviations of the factor and the excess return. We report population and average values of the betas, absolute (percentage) bias and absolute (percentage) RMSEs, mean and standard deviation of the t-ratios across simulations, and the ratio between average asymptotic standard errors across simulations and empirical standard deviation.

power properties of tests in the context of the LFM* representation.

9.5.1 Size. Table 6 reports the theoretical and actual sizes of the Wald-style test for the one-factor model the for LFM and LFM* specifications. First, consider the results for the LFM specification for the full sample. The actual size is generally close to the theoretical value, although there are systematic discrepancies. The CSR–OLS approach always leads to lower rejection rates than the CSR–GLS approach, and using the FF portfolios leads to higher rejection rates than using the size-sorted portfolios. Results are similar for the shorter sample of 240 months. Indeed, rejection rates are quite similar for the two sample lengths. Results are also similar for the block bootstrap (not reported in the table).

When we consider the LFM* specification, overall we see overrejections. Again, rejections are stronger for the FF portfolios. Rejection rates for the LFM* specification are always higher than rejection rates for the LFM specification for both the OLS and GLS approaches.

Results for the I–CAPM (not reported in the table) are similar to those for the one-factor specification.

9.5.2 Power. Table 7 (organized similarly to Table 6) reports rejection rates for the Wald test under the alternative, where the size is adjusted using the bootstrap results of Table 6. We compute the 10%, 5%, and 1% quantiles of the empirical distribution of the Wald statistic under the null and compute the percentage of times that the Wald statistic exceeds the corresponding quantile, when the economy is simulated under the alternative.

In the case of the LFM specification, the power of the tests is higher for $W = \hat{\Sigma}_{ee}^{-1}$, for the FF portfolio returns, and for the longer sample. For example, consider the case of the full sample and size-sorted portfolios. When the size is set at 1%, the actual rejection rate is 57% for the OLS approach; the rejection rate increases to 80% for the GLS approach, further increases to

Table 6. Size of the Wald test

	LFM								LFM*			
	Full sample				240 months				Full sample		240 months	
	Size		FF		Size		FF		Size	FF	Size	FF
	$W=I$	$W=sse$	$W=I$	$W=sse$	$W=I$	$W=sse$	$W=I$	$W=sse$	$W=I$	$W=sse$	$W=I$	$W=sse$
Average chi-squared	8.47	9.57	23.61	27.54	8.05	9.79	23.01	27.48	12.01	33.54	11.94	33.59
.50	.43	.54	.47	.67	.40	.57	.46	.68	.63	.82	.63	.81
.25	.21	.30	.24	.42	.19	.31	.23	.43	.40	.62	.39	.62
.10	.09	.14	.11	.23	.08	.14	.10	.22	.21	.42	.21	.42
.05	.05	.07	.06	.15	.04	.08	.05	.13	.14	.31	.13	.31
.025	.027	.040	.034	.092	.022	.044	.027	.075	.082	.227	.080	.227
.01	.01	.02	.01	.05	.01	.02	.01	.04	.04	.15	.04	.15

NOTE: Here we simulate a one-factor C–CAPM economy 10,000 times under the null (with the α parameters set to 0). Parameter estimates of the LFM are obtained using GMM with $W = I$. In the bootstrap, we jointly simulate the factor and the residuals from the estimated LFM. We consider two lengths of the data set: the full sample of 525 observations (full sample) and a shorter sample of 240 observations (240 months). We consider 2 choices of assets: the 10 size portfolios (size) and the 25 FF portfolios (FF). When we investigate the properties of estimates of the LFM, we consider estimates based on the two weighting matrices, $W = I$ and $W = see = \hat{\Sigma}_{ee}^{-1}$. Estimates of the parameters of the LFM* are obtained by exactly identified GMM. Asymptotic statistics are obtained assuming no serial correlation. We report the theoretical and actual sizes of the Wald test for the one-factor model. For the LFM and LFM* specifications. The LFM test statistic should be χ^2_{N-K} distributed, whereas the LFM* test statistic should be χ^2_N distributed.

Table 7. Power of the Wald test

	LFM								LFM*			
	Full sample				240 months				Full sample		240 months	
	Size		FF		Size		FF		Size	FF	Size	FF
	$W=I$	$W=sse$	$W=I$	$W=sse$	$W=I$	$W=sse$	$W=I$	$W=sse$	$W=I$	$W=sse$	$W=I$	$W=sse$
Average chi-squared	24.11	32.35	83.78	104.85	14.48	19.26	48.85	58.80	38.27	138.67	23.21	83.14
.10	.89	.97	1.00	1.00	.50	.64	.88	.98	.96	1.00	.62	.98
.05	.81	.93	.99	1.00	.35	.50	.82	.95	.92	1.00	.48	.96
.01	.57	.80	.98	1.00	.15	.25	.66	.83	.78	1.00	.22	.84

NOTE: Here we simulate a one-factor C–CAPM economy 10,000 times under the alternative ($\alpha \neq 0$). Parameter estimates of the LFM are obtained using GMM with $W = I$. In the bootstrap, we jointly simulate the factor and the residuals from the estimated LFM. We consider two lengths of the data set: the full sample of 525 observations (full sample) and a shorter sample of 240 observations (240 months). We consider 2 choices of assets: the 10 size portfolios (size) and the 25 FF portfolios (FF). When we investigate the properties of estimates of the LFM, we consider estimates based on the two weighting matrices, $W = I$ and $W = see = \hat{\Sigma}_{ee}^{-1}$. Estimates of the parameters of the LFM* are obtained by exactly identified GMM. Asymptotic statistics are obtained assuming no serial correlation. We report rejection rates for the Wald test under the alternative, where the size is adjusted using the bootstrap results of Table 6. We compute the 10%, 5%, and 1% quantiles of the empirical distribution of the Wald statistic under the null, then compute the percentage of times the Wald statistic exceeds the corresponding quantile when the economy is simulated under the alternative.

98% for the choice of FF test portfolios, and decreases to 15% when the sample is shortened to 240 months.

In the case of the LFM* specification, the power is also higher for the FF portfolio returns and for the longer sample. For example, in the case of the entire sample, for size portfolios, when the size is set at 1%, the rejection rate is 78%; the rejection rate increases to 100% for the FF portfolios, and falls to 22% for the shorter sample. The power of the LFM* specification is always higher than the power of the LFM specification and the OLS approach and comparable to the power of the GLS approach.

Results for the I–CAPM (not reported in the table) are similar. In the case of the LFM specification, the power of the tests is almost always higher for $W = \hat{\Sigma}_{ee}^{-1}$, although the differences are modest. Power also is higher for tests using the FF portfolios and for the longer sample. The power of the LFM* specification is comparable to the power of the LFM specification with the GLS approach and always higher than the power of the LFM specification with the OLS approach.

Thus, when it comes to power, the CSR–GLS and LFM* approaches have an advantage over the CSR–OLS approach, especially in the one-factor case. Moreover, there is a clear advantage to using the larger cross-section of the FF portfolios in both the one-factor and the two-factor cases.

10. EXTENSIONS

In this section we consider several extensions of our main analysis. First, we consider the case of time-varying conditional moments and conditional versions of the LFM and LFM* formulations. Second, we consider an alternative approach to the construction of mimicking portfolios, first suggested by Lehmann and Modest (1988). Third, we consider the performance of the correction for small-sample bias in risk premium estimates, suggested by Chen and Kan (2006). Fourth, we study the performance of the various methods when we consider long-horizon overlapping returns. Finally, we consider the implications of the restrictions from the LFM and LFM* formulations on covariance matrices for the purpose of portfolio construction. Details of the various exercises are in a separate appendix available on request. In what follows we provide summaries of the various setups and results.

10.1 Conditional Models

Here we consider conditional versions of the LFM and LFM* formulations. We allow for both risk premia and betas to be time-varying, but also specialize the analysis for cases in which betas or risk premia are constant. Following Ferson and Harvey (1999), we assume that the conditional betas and the conditional expectations of the factors are linear in the instruments. Also, following Ferson and Harvey (1991), we assume that the conditional risk premia also are linear in the instruments. For tractability, we assume that although the betas are time-varying, the factors are homoscedastic. The same assumption is made in, for example, the simulation analysis of Ferson et al. (2005). We focus on conditional versions of the I–CAPM, where the nontraded factor is the innovation in the dividend yield and lagged realizations of the dividend yield drive the time-varying betas and dividend yield risk premium. Conditional betas with respect to the traded factor, the market, also can vary as a function of the dividend yield.

Relative to the unconditional case, the properties of all estimators worsen considerably when the betas are allowed to vary. For example, the percentage bias of the CSR–OLS estimate increases from -3.69% to 79.88% and the percentage RMSE increases from 67.44% to 193.91%, where the population value of the Sharpe ratio is -7.94%. In contrast, when the betas are kept constant, the results are comparable to the base case illustrated in Table 2. In terms of the relative performance of the three estimators, the CSR–OLS estimator still displays the lowest bias, but the highest RMSE, across scenarios. The $\lambda^\star$ estimate displays the lowest percentage RMSE when the betas are time-varying. The CSR–GLS estimator has the lowest percentage RMSE when the betas are constant.

Interestingly, while the properties of the estimates worsen, the correlations between mimicking-portfolio returns and the factor can increase substantially. For example, in the case where betas and risk premia are both time-varying, the correlations for the conditional mimicking portfolios are 18.76% for CSR–OLS, 26.83% for CSR–GLS, and 26.50% for maximum correlation. For comparison, the respective correlations for the unconditional mimicking portfolios are 16.69%, 20.76%, and 20.76%.

10.2 Lehmann and Modest Portfolios

Lehmann and Modest (1988; LM hereinafter) argued that constructing unit-beta mimicking portfolios tends to place large weights on security returns associated with large estimated betas. Although this procedure is appropriate in the absence of measurement error, it is less appropriate when estimated betas reflect both the true betas and measurement error. Thus LM suggested constructing portfolios with minimum idiosyncratic risk, with betas of 0 with respect to all factors other than the factor being tracked and with the unit-beta constraint replaced by the constraint that the sum of portfolio weights equals 1. LM approximated the $\hat{\Sigma}_{ee}$ matrix with a diagonal matrix consisting of estimates of the idiosyncratic variances, $D_{\hat{\sigma}_e}$. We also consider a further approximation, in which all of the variances are the same, $I\hat{\sigma}_e^2$. Note that because the LM portfolios do not have a beta of 1, we divide the portfolio excess cash flows by the portfolio beta to obtain estimates of the factor risk premium. We perform this analysis for both the C–CAPM and the I–CAPM.

In general, the LM approach does worse than the LFM* formulation, both in terms of bias and RMSE. In the case of the C–CAPM, for example, the LM procedure leads to percentage biases of 8.12% ($W = I$) and 7.19% [$W = (D_{\hat{\sigma}_e})^{-1}$], with respective population values of the Sharpe ratios of 11.59% and 11.57%. For a comparison, the $\lambda^\star$ estimate has a bias of $-.69$%. Percentage RMSEs are 55.73% and 53.51% for the LM portfolios and 36.70% for the maximum-correlation portfolio.

10.3 Bias Corrections

Chen and Kan (2006) derived the small-sample distribution of the estimates of λ for both the OLS and the GLS cases under the assumption of Gaussian returns. In addition, they showed that the adjustments suggested by Litzenberger and Ramaswamy (1979) and Kim (1995), based on asymptotic results, lead to bias-adjusted estimators without finite moments.

Here we study the small-sample properties of the bias-adjusted estimators proposed by Chen and Kan (2006). We replicate the simulation analysis presented in Table 2 of the article for the case of the iid bootstrap. These estimators reduce the bias in 8 of our 16 cases. This result differs from that of Chen and Kan, who found that for the most part, the bias-adjusted estimates have less bias than the standard estimates in simulations. We attribute the difference in results to the fact that Chen and Kan assume that returns and factors are either multivariate normal or multivariate t-distributed, whereas we sample from the empirical distribution. Moreover, the RMSEs of the bias-adjusted estimates increase substantially, especially for the shorter sample of 240 months. For example, in the case of the C–CAPM, size portfolios, $W = I$, the RMSE is 95.33% for the standard estimate and 211.91% for the bias-adjusted estimate. This result is in line with the substantial increase in standard deviation of the bias-adjusted estimates documented by Chen and Kan, especially for the CSR–GLS case.

Comparing the λ and the $\lambda^\star$ estimates, the introduction of bias adjustment means that now there are cases in which the CSR–GLS estimator has less bias than the $\lambda^\star$ estimator. (Without a bias adjustment, the CSR–GLS estimator always has more bias than the $\lambda^\star$ estimator.) On the other hand, there are cases in which the $\lambda^\star$ estimator has less bias than the bias-adjusted CSR–OLS estimator. Finally, the $\lambda^\star$ estimates always have a lower RMSE than *both* the bias-adjusted CSR–OLS and CSR–GLS estimates.

In summary, in terms of bias, the bias-adjustment of Chen and Kan (2006) does not significantly alter the comparison between the LFM and LFM* formulations. In terms of RMSE, the bias adjustment makes the λ estimates always noisier than the $\lambda^\star$ estimates.

10.4 Long-Horizon Returns

Some authors have suggested that studying long-horizon returns may help better characterize the explanatory power of asset-pricing models, especially in the case of the C–CAPM (see, e.g., Lynch 1996; Daniel and Marshall 1997; Gabaix and Laibson 2001; Jagannathan and Wang 2007). Thus, we perform a simulation exercise where we simulate an iid economy at the monthly frequency, but study the inference on the factor risk premia using quarterly and annual returns. We perform the simulation for both the C–CAPM and the I–CAPM.

As would be expected, our results generally deteriorate relative to the case of monthly returns. Yet, at least in the case of the C–CAPM, the deterioration is much less marked for the $\lambda^\star$ estimates than for the CSR estimates. For example, consider the case of annual returns. The percentage biases in the CSR–OLS and CSR–GLS estimates are -10.43% and -66.55%, with respective population values of the Sharpe ratios of 11.57% and 15.86%. For comparison, the corresponding percentage biases in the case of monthly returns are 6.40% and -18.16%. (Given the assumption of iid returns, the population values are the same as in the case of annual returns.) Similarly, percentage RMSEs are 114.62% and 70.40% in the case of annual returns and 52.72% and 33.29% in the case of monthly returns. There is much less deterioration in the properties of the $\lambda^\star$ estimates. Percentage bias and RMSE are $-.57$% and 43.28% in the case of annual returns and $-.69$% and 36.70% in the case of monthly returns, with a population Sharpe ratio of 15.85%. Thus, using long-horizon returns may get around issues of misalignment between consumption decisions and returns in tests of the C–CAPM, but also may worsen considerably the properties of the risk premium estimators, particularly those based on the construction of unit-beta portfolios.

10.5 Portfolio Implications

The previous sections have focused on the implications of the LFM and LFM* representations for *mean* estimates of risk premia. It is now interesting to explore whether the two representations also have different implications for the *covariance* structure in asset returns for the purpose of constructing mean–variance–efficient portfolios. Specifically, we can take an "APT view" of the factors y_t and $y_t^\star$, and approximate the covariance matrices $\hat{\Sigma}_{ee}$ and $\hat{\Sigma}_{e^\star e^\star}$ with $D_{\hat{\sigma}_e}$ and $D_{\hat{\sigma}_{e^\star}}$. We can then use the resulting approximate covariance matrices as inputs for the construction of mean–variance–efficient portfolios.

We construct global minimum-variance (GMV) portfolios and evaluate their out-of-sample performance. Following the methodology of Jagannathan and Ma (2003), we estimate the covariance matrix of returns at the end of April of each year, using monthly returns from the previous 5 years. The GMV portfolios are held for 1 year. This procedure is repeated from 1964 to 2001. We consider two portfolios based on the restrictions of the LFM and the LFM* representations: a portfolio based on the sample covariance matrix and an equally weighted portfolio. In addition to the methodology of Jagannathan and Ma, we also consider the case in which the GMV portfolios are held for only 1 month and the exercise is repeated every month. Covariance matrices are estimated using either a lengthening sample (initial sample length 60 months) with a fixed starting date or moving windows of 60 and 360 months. In all experiments we consider two sets of base assets: the 10 size portfolios and the 25 FF portfolios.

Interestingly, in 9 of the 10 scenarios, the restrictions of the LFM* formulation lead to portfolios with lower volatility than the LFM formulation. Moreover, in six scenarios, the average return of the LFM* portfolios is higher than the average return of the LFM portfolios, whereas in the remaining four scenarios, average returns are essentially the same. Following Jagannathan and Ma (2003), we compute t-ratios on the differences between average returns and average squared returns. In eight of the nine cases in which the LFM* portfolios are less volatile than the LFM portfolios, the difference in average squared returns is significant (absolute t-ratio >2).

We also compare the performance of the LFM* portfolios and portfolios constructed with the unrestricted sample covariance matrix and also equally weighted portfolios. In the comparisons with the "unrestricted" portfolios, the LFM* portfolios always lead to higher standard deviation, and in 8 out of 10 cases the average return is also lower. Not surprisingly, the better performance of the unrestricted portfolios comes at the "cost" of more extreme positions; in eight cases, the average short interest in the unrestricted portfolios is higher than the average short interest in the LFM* portfolios. Comparing the equally weighted portfolios, the LFM* portfolios have a lower standard deviation in six cases and a higher average return in four cases.

In summary, the LFM* portfolios outperform the LFM portfolios, perform worse than the unrestricted portfolios, and perform similarly to the equally weighted portfolios.

11. CONCLUSIONS

In this article we have considered two alternative formulations of the linear factor model with nontraded factors. The first formulation is the traditional one, in which we estimate risk premia and alphas by means of a cross-sectional regression of average returns on betas. The second formulation replaces the factors with their projections onto the augmented span of excess returns. This second formulation requires only time series regressions for the estimation of risk premia and alphas.

First, we compared the theoretical properties of the two approaches. We provided some new results regarding the properties of factor models and mimicking portfolios and regarding the equivalence between characteristic- and risk-based explanations of the cross-section of stock returns. Then we performed a simulation exercise to assess the small-sample properties of the estimates. We found that the LFM* formulation of the model led to estimates of risk premia with less bias than, but similar RMSEs to, the CSR–GLS estimates. Because this pattern is robust to the different simulation scenarios considered, we believe that it is not an artifact of our experimental design. Moreover, although we do not cast our analysis within an explicit decision-theoretic framework, it is natural to think that the choice of the estimator depends on the loss function of the agent using the estimates. If the agent is concerned only with the first moment of the estimates, then the choice falls on the $\lambda^\star$ estimator. On the other hand, if the agent cares about first and second moments, then the choice depends on the trade-off between bias and variability. The RMSE ($=\sqrt{\text{variance}+\text{bias}^2}$) places equal weight on variance and squared bias. Thus, an agent equally concerned about variability and bias would be roughly indifferent between the two methods.

We also showed that the difference in bias between the two estimators is gradually reduced as the nontraded factor explains more of the variability of returns. On the other hand, in the case of a single factor observed with noise, the LFM* risk-premium estimate works much better than the CSR estimates in terms of both bias and RMSE.

Under the null, rejection rates are roughly correct for the LFM formulation but too high for the LFM* formulation. Thus testing the multi-beta model in its LFM* formulation requires an adjustment of critical values, based on the empirical distribution of the statistics under the null. With the bootstrap correction, the power of tests of multi-beta models is similar in the LFM* and CSR–GLS formulations and higher than in the CSR–OLS approach.

We also explored various extensions of our basic setup. For example, we investigated the effectiveness of the bias-adjustments of the CSR estimates proposed by Chen and Kan (2006) and used the LFM and LFM* formulations to impose restrictions on the covariance matrix of returns to construct minimum-variance portfolios. We found that in all scenarios, the $\lambda^\star$ estimates had lower RMSEs than the bias-adjusted CSR estimates of Chen and Kan. Moreover, we found that the restrictions from the LFM* formulation generally led to portfolios with better properties than the portfolios restricted using the LFM formulation. In light of these results, we conclude that when estimating risk premia, testing multi-beta models, and constructing mean–variance efficient portfolios, the LFM* formulation should be considered in addition to, or even instead of, the more traditional LFM formulation.

ACKNOWLEDGMENTS

The authors thank Torben Andersen, the editor, an anonymous associate editor, and an anonymous referee for valuable comments. The authors also thank seminar participants at the 2004 All Georgia Finance Conference, the 2006 AFA Meetings, and the 2006 Citigroup "Quant" Conference for useful comments. They also thank Wayne Ferson, Jay Shanken, Tim Simin (the discussant at the AFA Meetings), Paula Tkac, and Dan Waggoner for useful conversations and comments. Finally, a special thanks to Raymond Kan for providing several useful comments and for sharing his Matlab code for the bias-corrected estimators of Chen and Kan (2006).

APPENDIX: PROOFS

Proof of Result 1

We have that

$$\begin{aligned}E(r_t) &= \alpha^\star + (\beta^\star)^\top E(y^\star)\\ &= \alpha^\star + (\beta^\star)^\top (\gamma^\star)^\top E(r_t)\\ &= \alpha^\star + (\beta^\star)^\top (\gamma^\star)^\top \alpha + (\beta^\star)^\top (\gamma^\star)^\top \beta^\top \lambda. \quad (A.1)\end{aligned}$$

We also have

$$\begin{aligned}&(\beta^\star)^\top (\gamma^\star)^\top \beta^\top\\ &= \Sigma_{rr}\gamma^\star((\gamma^\star)^\top \Sigma_{rr}\gamma^\star)^{-1}(\gamma^\star)^\top \Sigma_{ry}\Sigma_{yy}^{-1}\\ &= \Sigma_{rr}\Sigma_{rr}^{-1}\Sigma_{ry}(\Sigma_{yr}\Sigma_{rr}^{-1}\Sigma_{rr}\Sigma_{rr}^{-1}\Sigma_{ry})^{-1}\Sigma_{yr}\Sigma_{rr}^{-1}\Sigma_{ry}\Sigma_{yy}^{-1}\\ &= \Sigma_{ry}(\Sigma_{yr}\Sigma_{rr}^{-1}\Sigma_{ry})^{-1}\Sigma_{yr}\Sigma_{rr}^{-1}\Sigma_{ry}\Sigma_{yy}^{-1}\\ &= \Sigma_{ry}\Sigma_{yy}^{-1}\\ &= \beta^\top. \quad (A.2)\end{aligned}$$

Thus we write

$$\begin{aligned}&E(r_t) - (\beta^\star)^\top (\gamma^\star)^\top \beta^\top \lambda\\ &= E(r_t) - \beta^\top \lambda\\ &= \alpha = \alpha^\star + (\beta^\star)^\top (\gamma^\star)^\top \alpha \quad (A.3)\end{aligned}$$

and

$$\alpha^\star = [I - (\beta^\star)^\top (\gamma^\star)^\top]\alpha. \quad (A.4)$$

It is straightforward to show that both $I - (\beta^\star)^\top(\gamma^\star)^\top$ and $(\beta^\star)^\top(\gamma^\star)^\top$ are idempotent matrices. We want to show that if $(\beta^\star)^\top(\gamma^\star)^\top$ is idempotent of rank K, then $I - (\beta^\star)^\top(\gamma^\star)^\top$ is idempotent of rank $N-K$. Note that $(\beta^\star)^\top(\gamma^\star)^\top$ can be written as $\Sigma_{ry}(\Sigma_{yr}\Sigma_{rr}^{-1}\Sigma_{ry})^{-1}\Sigma_{yr}\Sigma_{rr}^{-1}$. Given that Σ_{rr}^{-1} is of full rank,

$$\begin{aligned}&\operatorname{rank}[\Sigma_{ry}(\Sigma_{yr}\Sigma_{rr}^{-1}\Sigma_{ry})^{-1}\Sigma_{yr}\Sigma_{rr}^{-1}]\\ &\qquad = \operatorname{rank}[\Sigma_{ry}(\Sigma_{yr}\Sigma_{rr}^{-1}\Sigma_{ry})^{-1}\Sigma_{yr}]. \quad (A.5)\end{aligned}$$

Because $\Sigma_{yr}\Sigma_{rr}^{-1}\Sigma_{ry}$ is symmetric and invertible, $\operatorname{rank}[(\beta^\star)^\top \times \operatorname{rank}(\gamma^\star)^\top] = \operatorname{rank}[\Sigma_{ry}(\Sigma_{yr}\Sigma_{rr}^{-1}\Sigma_{ry})^{-1}\Sigma_{yr}] = \operatorname{rank}(\Sigma_{yr}) = K$. Now the null space of $I - (\beta^\star)^\top(\gamma^\star)^\top$ is the set of all $z \neq 0$ such that $[I - (\beta^\star)^\top(\gamma^\star)^\top]z = 0$. But $[I - (\beta^\star)^\top(\gamma^\star)^\top]z = 0$ if and only if $z = (\beta^\star)^\top(\gamma^\star)^\top z$, that is, z is an eigenvector of $(\beta^\star)^\top(\gamma^\star)^\top$ with eigenvalue equal to 1. Because $(\beta^\star)^\top(\gamma^\star)^\top$ is idempotent, the dimension of all such eigenvectors is equal to $\operatorname{rank}[(\beta^\star)^\top(\gamma^\star)^\top] = K$, which means that $\operatorname{rank}[I - (\beta^\star)^\top(\gamma^\star)^\top] = N-K$. Moreover, the null space of the idempotent matrix $I - (\beta^\star)^\top(\gamma^\star)^\top$ is spanned by linear combinations of the column vectors in $\beta^\top$. Because $\beta^\top$ is of rank K, the set of all $\alpha = \beta^\top\delta$ for $\delta \in \Re^K$ is of dimension K. This set is contained in the null space of $I - (\beta^\star)^\top(\gamma^\star)^\top$ and must be all of the null space.

Proof of Result 2

We have that

$$\begin{aligned}\lambda^\star &= (\gamma^\star)^\top E(r_t)\\ &= (\gamma^\star)^\top \alpha + (\gamma^\star)^\top \beta^\top \lambda\\ &= (\gamma^\star)^\top \alpha + \Sigma_{yr}\Sigma_{rr}^{-1}\Sigma_{ry}\Sigma_{yy}^{-1}\lambda. \quad (A.6)\end{aligned}$$

Note that

$$\begin{aligned}\operatorname{var}((\gamma^\star)^\top r_t) &= \Sigma_{y^\star y^\star}\\ &= \Sigma_{yr}\Sigma_{rr}^{-1}\Sigma_{rr}\Sigma_{rr}^{-1}\Sigma_{ry} = \Sigma_{yr}\Sigma_{rr}^{-1}\Sigma_{ry}. \quad (A.7)\end{aligned}$$

Thus

$$\begin{aligned}\lambda^\star &= (\gamma^\star)^\top \alpha + (\gamma^\star)^\top \beta^\top \lambda\\ &= (\gamma^\star)^\top \alpha + \Sigma_{y^\star y^\star}\Sigma_{yy}^{-1}\lambda. \quad (A.8)\end{aligned}$$

Proof of Result 3

Using the fact that $\tilde{\gamma}$ is the same for $W = \Sigma_{ee}^{-1}$ and $W = \Sigma_{rr}^{-1}$, we have

$$\begin{aligned}&\beta^\top(\beta\Sigma_{ee}^{-1}\beta^\top)^{-1}\beta\Sigma_{ee}^{-1}E(r_t)\\ &= \Sigma_{ry}\Sigma_{yy}^{-1}(\beta\Sigma_{rr}^{-1}\beta^\top)^{-1}\beta\Sigma_{rr}^{-1}E(r_t)\\ &= \Sigma_{ry}\Sigma_{yy}^{-1}(\Sigma_{yy}^{-1}\Sigma_{yr}\Sigma_{rr}^{-1}\Sigma_{ry}\Sigma_{yy}^{-1})^{-1}\Sigma_{yy}^{-1}\Sigma_{yr}\Sigma_{rr}^{-1}E(r_t)\\ &= \Sigma_{ry}(\Sigma_{yr}\Sigma_{rr}^{-1}\Sigma_{ry})^{-1}\Sigma_{yr}\Sigma_{rr}^{-1}E(r_t)\\ &= (\beta^\star)^\top\lambda^\star. \quad (A.9)\end{aligned}$$

This implies that when $W = \Sigma_{ee}^{-1}$, then $\beta^\top\tilde{\lambda} = (\beta^\star)^\top\lambda^\star$. Thus, we also have $\tilde{\alpha} = E(r_t) - \beta^\top\tilde{\lambda} = E(r_t) - (\beta^\star)^\top\lambda^\star = \alpha^\star$.

Proof of Result 4

Consider the unit-beta mimicking portfolios. The maximum squared Sharpe ratio from investment in the unit-beta mimicking portfolios is given by

$$\begin{aligned}&E(r_t)^\top\tilde{\gamma}^\top(\tilde{\gamma}^\top\Sigma_{rr}\tilde{\gamma})^{-1}\tilde{\gamma}E(r_t)\\ &\qquad = E(r_t)^\top W\beta^\top(\beta W\Sigma_{rr}W\beta^\top)^{-1}\beta W E(r_t). \quad (A.10)\end{aligned}$$

Using the decomposition $E(r_t) = \tilde{\alpha} + \beta^\top\tilde{\lambda}$, where $\beta W\tilde{\alpha} = 0$ by construction (with $\tilde{\alpha}$ the vector of residuals of a weighted least squares regression of the expected excess returns on the betas), the maximum squared Sharpe ratio simplifies to

$$\tilde{\lambda}^\top(\beta W\beta)(\beta W\Sigma_{rr}W\beta^\top)^{-1}(\beta W\beta^\top)\tilde{\lambda}. \quad (A.11)$$

Note that $(\beta\Sigma_{ee}^{-1}\beta^\top)^{-1}\beta\Sigma_{ee}^{-1} = (\beta\Sigma_{rr}^{-1}\beta^\top)^{-1}\beta\Sigma_{rr}^{-1}$, and the unit-beta portfolio weights are the same for the two choices of weighting matrix, $W = \Sigma_{ee}^{-1}$ and $W = \Sigma_{rr}^{-1}$. Substituting $W = \Sigma_{rr}^{-1}$ in the foregoing expression, we obtain

$$\tilde{\lambda}^\top\beta\Sigma_{rr}^{-1}\beta^\top\tilde{\lambda}. \quad (A.12)$$

Now consider the maximum-correlation mimicking portfolios. The maximum squared Sharpe ratio from investing in the maximum-correlation mimicking portfolios is

$$\begin{aligned}&E(y_t^\star)^\top\Sigma_{y^\star y^\star}^{-1}E(y_t^\star)\\ &\qquad = E(r_t)^\top\gamma^\star((\gamma^\star)^\top\Sigma_{rr}\gamma^\star)^{-1}(\gamma^\star)^\top E(r_t). \quad (A.13)\end{aligned}$$

Again using the decomposition $E(r_t) = \tilde{\alpha} + \beta^\top\tilde{\lambda}$ and the fact that $\Sigma_{yr}W\tilde{\alpha} = 0$ ($\tilde{\alpha} \equiv E(r_t) - \beta^\top\tilde{\lambda} = E(r_t) - \Sigma_{ry} \times (\Sigma_{yr}W\Sigma_{ry})^{-1}\Sigma_{yr}WE(r_t) = [I - \Sigma_{ry}(\Sigma_{yr}W\Sigma_{ry})^{-1}\Sigma_{yr}W]E(r_t)$), we have

$$E(y_t^\star)^\top \Sigma_{y^\star y^\star}^{-1} E(y_t^\star) = \tilde{\lambda}^\top \beta \Sigma_{rr}^{-1} \beta^\top \tilde{\lambda}. \tag{A.14}$$

Proof of Result 5

Consider unit-characteristic portfolios, which minimize $\gamma_k^\top W\gamma_k$ subject to $H\gamma_k = s_k$. The weights of these portfolios are the coefficients of a CSR of expected returns on the characteristics

$$\bar{\gamma}^\top = (HWH^\top)^{-1}HW. \tag{A.15}$$

It is immediate that

$$E(\bar{\gamma}^\top r) = E(\bar{y}) = \delta. \tag{A.16}$$

Moreover, if $W = \Sigma_{rr}^{-1}$, then the betas with respect to the returns on the unit-characteristic portfolios, $\bar{\beta}$, are equal to the characteristics themselves,

$$\bar{\beta} = (\bar{\gamma}^\top \Sigma_{rr}\bar{\gamma})^{-1}\bar{\gamma}^\top \Sigma_{rr} = H. \tag{A.17}$$

[Received September 2005. Revised April 2007.]

REFERENCES

Affleck-Graves, J. F., and Bradfield, D. J. (1993), "An Examination of the Power of Univariate Tests of the CAPM: A Simulation Approach," *Journal of Economics & Business*, 45, 17–33.

Amsler, C. E., and Schmidt, P. (1985), "A Monte Carlo Investigation of the Accuracy of Multivariate CAPM Tests," *Journal of Financial Economics*, 14, 359–375.

Asgharian, H. (2004), "A Comparative Analysis of Ability of Mimicking Portfolios in Representing the Background Factors," working paper, Lund University.

Asgharian, H., and Hansson, B. (2005), "A Critical Investigation of the Explanatory Role of Factor Mimicking Portfolios in Equilibrium Pricing," *Applied Financial Economics*, 15, 835–847.

Black, F., Jensen, M. C., and Scholes, M. (1972), "The Capital Asset Pricing Model: Some Empirical Tests," in *Studies in the Theory of Capital Markets*, ed. M. C. Jensen, New York: Praeger, pp. 79–121.

Brav, A., Constantinides, G. M., and Geczy, C. C. (2002), "Asset Pricing With Heterogeneous Consumers and Limited Participation: Empirical Evidence," *Journal of Political Economy*, 110, 793–824.

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

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

Brennan, M. J., Chordia, T., and Subrahmanyam, A. (1998), "Alternative Factor Specifications, Security Characteristics, and the Cross-Section of Expected Stock Returns," *Journal of Financial Economics*, 49, 345–373.

Chan, L. K. C., Karceski, J., and Lakonishok, J. (1998), "The Risk and Return From Factors," *Journal of Financial and Quantitative Analysis*, 33, 159–188.

Chen, R., and Kan, R. (2006), "Finite-Sample Analysis of Two-Pass Cross-Sectional Regressions," working paper, University of Toronto.

Cochrane, H. J. (2001), *Asset Pricing*, Princeton, NJ: Princeton University Press.

Daniel, K., and Marshall, D. (1997), "Equity Premium and Risk-Free-Rate Puzzles at Long Horizons," *Macroeconomic Dynamics*, 1, 452–484.

Daniel, K., and Titman, S. (1997), "Evidence on the Characteristics of Cross-Sectional Variation in Stock Returns," *Journal of Finance*, 52, 1–33.

Dickey, J. M. (1967), "Matricvariate Generalizations of the Multivariate t Distribution and the Inverted Multivariate t Distribution," *Annals of Mathematical Statistics*, 38, 511–518.

Fama, E. F. (1996), "Multifactor Portfolio Efficiency and Multifactor Asset Pricing," *Journal of Financial and Quantitative Analysis*, 31, 441–465.

Fama, E. F., and French, K. R. (1992), "The Cross-Section of Expected Stock Returns," *Journal of Finance*, 47, 427–465.

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

Fama, E. F., and MacBeth, J. D. (1973), "Risk, Return and Equilibrium," *Journal of Political Economy*, 81, 607–636.

Ferson, W. E. (1996), "Warning: Attribute-Sorted Portfolios Can Be Hazardous to Your Research," in *Modern Portfolio Theory and Its Applications*, eds. S. Saitou, K. Sawaki, and K. Kubota, Osaka, Japan: Center for Academic Societies, pp. 21–32.

Ferson, W. E., and Harvey, C. R. (1991), "The Variation of Economic Risk Premiums," *Journal of Political Economy*, 99, 385–415.

——— (1999), "Conditioning Variables and the Cross-Section of Stock Returns," *Journal of Finance*, 54, 1325–1360.

Ferson, W. E., Sarkissian, S., and Simin, T. (1999), "The Alpha Factor Asset Pricing Model: A Parable," *Journal of Financial Markets*, 2, 49–68.

Ferson, W. E., Siegel, A. F., and Xu, P. T. (2005), "Mimicking Portfolios With Conditioning Information," *Journal of Financial and Quantitative Analysis*, 41, 607–636.

Gabaix, X., and Laibson, D. (2001), "The 6D Bias and the Equity Premium Puzzle," in *NBER Macroeconomics Annual*, eds. B. Bernanke and K. Rogoff, Vol. 16, Cambridge, MA: MIT Press, pp. 257–311.

Grinblatt, M., and Titman, S. (1987), "The Relation Between Mean–Variance Efficiency and Arbitrage Pricing," *Journal of Business*, 60, 97–112.

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

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

Jacobs, K., and Wang, K. Q. (2004), "Idiosyncratic Consumption Risk and the Cross-Section of Asset Returns," *Journal of Finance*, 59, 2211–2252.

Jagannathan, R., and Ma, T. (2003), "Risk Reduction in Large Portfolios: Why Imposing the Wrong Constraints Helps," *Journal of Finance*, 58, 1651–1683.

Jagannathan, R., and Wang, Y. (2007), "Lazy Investors, Discretionary Consumption, and the Cross-Section of Stock Returns," *Journal of Finance*, 62, 1623–1661.

Jagannathan, R., and Wang, Z. (2002), "Empirical Evaluation of Asset Pricing Models: A Comparison of the SDF and Beta Methods," *Journal of Finance*, 57, 2337–2367.

Kan, R., and Zhang, C. (1999), "Two-Pass Tests of Asset Pricing Models With Useless Factors," *Journal of Finance*, 54, 204–235.

Kan, R., and Zhou, G. (1999), "A Critique of the Stochastic Discount Factor Methodology," *Journal of Finance*, 54, 1221–1248.

——— (2001a), "Empirical Asset Pricing: The Beta Method versus the Stochastic Discount Factor Method," working paper, University of Toronto.

——— (2001b), "Tests of Mean–Variance Spanning," working paper, University of Toronto.

Kandel, S., and Stambaugh, R. F. (1995), "Portfolio Inefficiency and the Cross-Section of Expected Returns," *Journal of Finance*, 50, 157–184.

Kim, D. (1995), "The Errors in the Variables Problem in the Cross-Section of Expected Stock Returns," *Journal of Finance*, 50, 1605–1634.

Kimmel, R. L. (2003), "Risk Premia in Linear Factor Models: Theoretical and Econometric Issues," working paper, Princeton University.

Lamont, O. (2001), "Economic Tracking Portfolios," *Journal of Econometrics*, 105, 161–184.

Lehmann, B., and Modest, D. M. (1988), "The Empirical Foundations of the Arbitrage Pricing Theory," *Journal of Financial Economics*, 21, 213–254.

Litzenberger, R. H., and Ramaswamy, K. (1979), "The Effect of Personal Taxes and Dividends on Capital Asset Prices," *Journal of Financial Economics*, 7, 163–195.

Lynch, A. (1996), "Decision Frequency and Synchronization Across Agents: Implications for Aggregate Consumption and Equity Return," *Journal of Finance*, 51, 1479–1498.

MacKinlay, A. C. (1987), "On Multivariate Tests of the CAPM," *Journal of Financial Economics*, 18, 341–372.

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

Newey, W., and West, K. (1987), "A Simple Positive Definite Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 703–708.

Peñaranda, F., and Sentana, E. (2004), "Spanning Tests in Return and Stochastic Discount Mean–Variance Frontiers: A Unifying Approach," working paper, CEMFI.

Shanken, J. (1992), "On the Estimation of Beta-Pricing Models," *Review of Financial Studies*, 5, 1–33.

Shanken, J., and Zhou, G. (2007), "Estimating and Testing Beta Pricing Models: Alternative Methods and Their Performance in Simulations," *Journal of Financial Economics*, 84, 40–86.