eman ta zabal zazu

Universidad del País Vasco   Euskal Herriko Unibertsitatea

ESSAYS ON ESTIMATING AND TESTING ASSET PRICING MODELS

By

Martín Lozano

Departamento de Fundamentos del Análisis Económico II

Bilbao, Bizkaia. Septiembre 2009

To Míldred

❀

Table of Contents

List of Tables

List of Figures

Abstract

Asset pricing models are concerned with determining the expected returns of assets whose payoffs are risky. These financial models analyze the relationship between risk and expected return, and address the crucial question of how to value risk.

Empirical finance widely adopts either the classical *Beta* method or the *stochastic discount factor* (SDF) method for the estimation and evaluation of asset-pricing models. It is common for researchers to select one approach over the other and consequently, certain specific areas of the literature appear to favor one method over the other. However, only recently have there been attempts to empirically evaluate the two approaches, and even though the generalized consensus is that there are no significant differences between them, we find that this is not always the case.

One of the most relevant and original implication of our results is that if we are interested on making inference on a multi-factor model estimator(s), we should prefer the Beta method over the SDF method. Conversely, if we are primarily interested on making inference on the sampling pricing error or Jensen's alpha, the SDF method should be implemented. We argue that previous studies are conducted under fairly simple conditions which are not sufficient to raise substantial divergences.

Our results have an extensive list of practical implications in empirical asset pricing and financial econometrics since both researchers and practitioners are interested on retrieve efficient estimators when estimating asset pricing models. In this work, we formulate a set of recommendations based on empirical analysis and finite sample properties which will lead to more accurate hypothesis tests and calculations.

■

Acknowledgements

I sincerely thank first and foremost, my advisor, Gonzalo Rubio, for his many suggestions and constant support during this research. He quickly became for me the role model of a successful researcher in the field. Without his supervision, this work would not have been a reality.

For providing financial support I would like to thank to the Universidad del País Vasco (Econometrics and Statistic Department, and Department of Foundations of Economic Analysis II), the BBVA Foundation, the Delegation of the Basque Government in México, the Sixth European Community Framework Programme, the Marie Curie Program, and the University of Manchester.

I owe my deepest gratitude to Miguel Ángel Martínez (Universidad del País Vasco) whose thoughtful advise often served to give me a sense of direction during my PhD studies, and Fernando Tusell (Universidad del País Vasco) for his empathy and also for introduce me into LaTeX and computational programming. I would like to show my warm and sincere gratitude to Stuart Hyde (Manchester Business School), who has been a cornerstone in my professional development. For expressed his interest in my work and shared with me his knowledge I thank Ian Garrett and Massimo Guidolin (Manchester Business School).

For administrative support I thank to Inés García, Ana Franco, Isabel Urrutia, Karmele Largatxa, Itzel Olague, Graham Cockerill and specially to Mark Greenwood for useful advice with financial databases. I am indebted to Manchester Business School for providing me the access to the North West Grid computational facilities, which represented an invaluable tool in the last part of this research.

For many insightful and helpful comments, I thank Alfonso Novales (Universidad Complutense de Madrid). For taking time off to give me some computational advice I thank to John Cochrane (University of Chicago). To the helpful comments of an anonymous Journal of Empirical Finance referee. To Raymond Kan (University of Toronto) for kindly sharing complementary econometric notes on Kan and Zhou [66]. This thesis has benefitted greatly from distinguished seminars discussants: José Emilio Farinós (Universitat de València), Rosa Rodríguez and Genaro Sucarrat (Universidad Carlos III de Madrid), Chien-Ting Lin (University of Adelaide), and Sheng Guo (Florida International University). Also, I owe a huge debt of gratitude to Pedro Serrano (Universidad Carlos III de Madrid), who has been a source of valuable advice throughout my doctoral studies.

More people were crucial to the successful completion of this project like Cristina Mazón and Lola Robles (Universidad Complutense de Madrid); Eva Ferreira, Victoria Esteban and Susan Orbe (Universidad del País Vasco); Vicente Meneu (Universitat de València). Last but not least, I would like to thank to all my dear *QF* colleagues, but especially to Ana C., Ana G., Laura, Miguel, Jorge, Agueda and Tania.

I thank God for making possible to complete this project and giving me the opportunity to meet all this talented and friendly people far from home.

I am heartily grateful to my loving parents, my brother and his growing family, who have been a constant source of emotional, moral and of course financial support. Without them this work would never have come into existence (literally), and nothing I can say can do justice to how I feel about their support, for which I am eternally grateful.

I am also indebted to my mother-in-law who encouraged me throughout these years.

My final, and most heartfelt acknowledgment, go to my wise and beautiful wife, Míldred, for her selfless sacrifices, and for supporting me in every way possible. She has been my pillar, my joy and my guiding light. I thank her because without her support, this thesis would never have started much less finished. Her strategic vision and extraordinary developed sixth sense have been, and will be, crucial in our life. For all that, and for being everything I am not (and even more), she has my everlasting ♡.

Bilbao, Bizkaia — Martín Lozano
September 9, 2009

Introduction

The asset pricing models play an important role in modern price theory and financial economics because they help, among other things, to measure portfolio risk and the return an investor can expect for taking that risk.

These models are widely used in empirical analysis because of the abundance of available financial statistical data, and because they can be applied extensively in practical research and decision making processes. For instance, researchers and practitioners may employ asset pricing models for calculations of costs of capital associated with investment and takeover decisions, which it is a recurrent task in areas such as accountability and corporate finance. Also, asset pricing models are used in comparative analysis of the success of different investors or performance evaluation of investment funds. They are even applied in judicial inquires related to court decisions regarding compensation to expropriated firms whose shares are not listed on the stock market. Actually, the list of applications is endless.

Any econometric estimation technique has associated statistical errors, referred as the standard error of an estimator. Generally speaking, a large amount of standard error makes the estimator unreliable, and a relative small standard error is usually a sign of estimator precision. Then, the efficiency of the econometric method becomes relevant for researchers and practitioners because as far as the chosen technique delivers more precise estimators, the calculations and hypothesis tests results will be more trustworthy. For example, in order to evaluate an investment project, managers have to discount the project's future cash flows. To do so, they require a projection of the cash flows and an estimation of a discount rate which may be obtained from an asset pricing model. It follows that if this estimation is not accurate enough, the evaluation of the investment project will not be assertive either.

Roughly speaking, there are two main inputs necessary to estimate an asset pricing model, the factors and the expected returns of an asset or portfolio. We can associate the expected returns of a portfolio to what we are interested to explain, whereas the factors represent the explanatory variables. The actual interaction between the inputs depends on the theoretical foundations of the model.

On the other hand, there are two main outputs when estimating an asset pricing model, the risk premium and the pricing errors. The risk premium estimators measure the statistical importance of the model's incorporated factors at determining the expected return of the financial asset. For example, if we want to test whether the factor called *size of the firm* play an important role at determining the expected return of the financial asset, we should test whether the size-risk-premium is statistically significant. The pricing error measure overall how well the model explains the expected return of the financial asset; it helps to examine how good the model is as an approximation of the real world.

We can classify the relevant econometric techniques to estimate asset pricing models into two main categories, the *traditional* Beta and the *general and relatively new* Stochastic Discount Factor (SDF) methods. The Beta method usually involves the two-pass cross-sectional regression procedure, maximum likelihood, among others. The SDF method requires to choose a functional form for the discount factor, and traditionally involves the generalized method of moments procedure. Regardless of the pricing model and application, researchers and practitioner either chooses the Beta or the SDF method to conduct calculations and perform hypothesis tests. Naturally, their results and tests are subject to the statistical properties of the estimators, which at the same time are subject to the econometric method.

In the past few years, there have been concerns that, compared to the classical Beta method, the generality of the SDF method comes at the cost of loosing efficiency in the parameter estimation. This concern has motivated an interesting discussion, particularly in the Journal of Finance. Basically, the main and current consensus is that both methodologies provide virtually equal efficient estimators (i.e. risk premiums and pricing errors).

Since the comparison is not a trivial exercise, the current discussion have been criticized and debated. Actually, we have found that current studies are limited in their applicability since their empirical results are valid exclusively for the single-factor model case. The single factor model was introduced in the literature in the 50's and it is still an important benchmark, nevertheless recent models such as multi-factor pricing models are by far the most commonly used in current empirical applications.

Our main hypothesis is to test whether divergences between the methodologies emerge under more realistic conditions. There is no reason to expect that previous results will remain when estimating multi-factor models or when testing other portfolio formations which allow a greater return dispersion. Particularly, we argue that current studies are valid but incomplete because their results are not extensive in a multi-factor framework.

Our main objective is find out under which circumstances – if any – both methodologies may lead to differences in terms of the estimator properties; and consequently justify the use of a specific method over the other. This justification will depend on whether the main interest relies on the risk premium or on the pricing errors.

By justifying the use of an specific method over the other, we intend to cover a gap in the current empirical literature and contribute to the debate between the differences of both methodologies. As pointed out before by other authors, this comparison can be so important that it might change the course of our empirical research on asset pricing models.

In order to achieve such objective we present four chapters. First chapter explore in detail the econometrics of evaluating and estimating asset pricing models. This chapter is based on some of the most influential work in this field and makes emphasis on the risk premium and pricing errors estimation. One particular feature of this first chapter is that facilitates the consecutive programming. Second chapter apply the entire theoretical review to the empirical analysis of two dominant multi-factor models namely Fama-French and Carhart. One remarkable finding is that we introduce a third model empirically motivated called RUH which apparently outperform them. Furthermore, we find that Fama-French model is consistently rejected by the data while Carhart model is not. In addition, the size factor is usually statistically equal to zero regardless of the method and model employed.

Third and fourth chapters are advocated to cover our main objective. Third chapter analyze models and methods using historical dataset, we find evidence that suggest that SDF method achieve lower pricing errors than Beta method. On the other hand, their specifications tests show evidence in favor of RUH model because the likeliness of not rejecting the null is greater than in Fama-French and CAPM models. We also find that double-sorted portfolios are hardest to price compared with single-sorted portfolios, and this difference is correlated with the higher dispersion on the test portfolios' average returns. Our results suggest that the Beta method actually do better than the SDF method at estimating risk premiums.

Fourth chapter emphasize on the formal comparison of the methods with artificial data. In particular, we elaborate a finite sample analysis using simulations, in this way our results are directly comparable to the current works on the comparison of the Beta and SDF methods. One of our original contributions of this chapter is that the efficiency of the methodologies is sensible to the number and to the statistical properties of the factors employed. Furthermore, we find that each methodology tend to favor the efficiency of one particular asset pricing output, so that no method fully dominates the other in terms of deliver efficient estimators. A remarkable contribution of this chapter rely on the comparison of alternative discount factor representations commonly used in the literature. Even though third and fourth chapters have different approaches, their results and implications are similar since we find out that once we compare the methodologies under more complex setups differences clearly emerge.

There are other kinds of models such as the non-linear, often used for pricing derivative securities like futures and options; however, they can be frequently linearized and treated as multi-factor models as well. Therefore, our results regarding multi-factor models imply a considerably wider range of applicability compared to previous studies.

■

Chapter 1

A methodology review

§

Asset pricing models are concerned with determining the expected returns of assets whose payoffs are risky. Explicitly, these models analyze the relationship between risk and expected return, and address the crucial question of how to value risk. This review summarizes some of the methodology currently available for estimating and evaluating Beta and stochastic discount factor (SDF) models such as time-series regression, cross-sectional regression, Fama-MacBeth procedure, and the generalized method of moments (GMM).

1.1 Introduction

There is a large literature on econometric techniques to estimate and evaluate asset pricing models. This review is based on some of the most influential work in this

[§]An earlier version of this work was presented at Instituto Complutense de Análisis Económico (Madrid, May 2006); and at the IV Workshop in Quantitative Finance (València, June 2006), for helpful comments I thank José Emilio Farinós (Universitat de València). Martín Lozano gratefully acknowledges financial assistance from Consolidate Research Team 9/UPV-00038.321-15094, Econometrics and Statistics Department, Departamento de Fundamentos del Análisis Económico II at University of the Basque Country (Euskal Herriko Unibertsitatea), and Fundación BBVA.

field such as Cochrane [25], Ferson [34, 35], Ferson and Jagannathan [38], Campbell, Lo and MacKinlay [17], Shanken [93], Velu and Zhou [101], Ogaki [86], Jagannathan, Skoulakis and Wang [55, 54], and Marín and Rubio [81]. Each econometric technique focuses on the same questions: how to estimate parameters, how to calculate standard errors of the estimated parameters, how to calculate standard errors of the pricing errors, and how to test the model, usually with a quadratic test statistic of the form $\widehat{\alpha}' V^{-1} \widehat{\alpha}$, where $\widehat{\alpha}$ represent an estimator of pricing errors, and V is a weighting matrix. As Cochrane [25] resume, all the techniques come down to one of the two basic ideas: time-series regression or cross-sectional regression[1]. Time-series regression turns out to be a limiting case of cross-sectional regression. The stochastic discount factor representation, in which the value of an asset p is equal to the expected value of its payoff x and the discount factor m, such as $p = E(mx)$, turns out to be almost identical to cross-sectional regressions. The formulas for parameter estimates, standard errors, and test statistics are all strikingly similar.

1.2 Time-series estimation and evaluation

1.2.1 Regressions

Let R_t^{ei} be a vector of $i = 1, ..., n$ asset returns in excess (e) of the risk-free rate. To reduce notational complexity, we assume that there is one vector of economy-wide pervasive risk factors f_t. In this context, Black, Jensen and Scholes [9] suggested the following approach for time-series estimation and evaluation:

$$R_t^{ei} = \alpha_i + \beta_i f_t + \varepsilon_{it}. \tag{1.2.1}$$

[1]Usually known as the CRS method.

The model states that expected returns are linear in the betas, taking the expected value of 1.2.1: $E(R^{ei}) = \beta_i E(f)$.

When the economy-wide factor f_t is the return on a portfolio of traded assets, we call it a traded factor. An example of a traded factor would be the return on the value-weighted portfolio of stocks used in empirical studies like in Sharpe's [95] Capital Asset Pricing Model (CAPM). Examples of nontraded factors can be found in Chen, Roll and Ross [21], who use the growth rate of industrial production and the rate of inflation, and Breeden, Gibbons and Litzenberger [12], who use the growth rate in per capita consumption as a factor.

Here we will suppose that the factor f is also an excess return as R^{ei}, consequently the model applies to the factor as well, so $E(f) = 1 \times \lambda$. Comparing the model (1.2.1) and the expectation of the time-series regression $E(R^{ei}) = \beta_i E(f)$, we see that the model has one and only one implication for the data: *all the regression intercepts* α_i *should be zero.* Thus, the regression intercepts are equal to the pricing errors.

Given this fact, Black, Jensen and Scholes [9] suggested a natural strategy for estimation and evaluation: run time-series regressions (1.2.1) for each test asset i. Given that the factor is an excess return, the estimate of the factor risk premium $\widehat{\lambda}$ is just the sample mean of the factor:

$$\widehat{\lambda} = E_T(f). \tag{1.2.2}$$

This restriction allows us to use the sample mean of the factor as an estimator of the risk premium. If the factor is not traded, this restriction does not hold, and we have to estimate the risk premium using returns on traded assets. We focus on the case where the factor is traded, although we also consider the case where the factor is not traded.

1.2.2 Standard errors: OLS and GMM

We will show two versions of standard errors that came out from two assumptions about the residuals ε_{it} at (1.2.1): assuming serially uncorrelated and homoskedastic errors; and assuming serial correlation and conditional heteroskedasticity.[2] Correcting OLS standard errors for econometric problems is *not* the same thing as GLS. When errors do not obey the OLS assumptions, OLS is consistent, and often more robust than GLS, but its standard errors need to be corrected.

We start by deriving the OLS standard errors that correct for econometric problems. In general, OLS picks parameters β to minimize the variance of the residual σ_ε^2:

$$\min_{\{\beta\}} E_T\left[(y_t - \beta' x_t)^2\right], \text{ where: } y_t - \beta' x_t = \varepsilon_t.$$

We find β from the first-order condition, which states that the residual ε_t is orthogonal to the right-hand variable x_t:

$$g_T(\beta) = E_T\left[x_t\left(y_t - x_t'\beta\right)\right] = 0.$$

This condition is exactly identified – the number of moments g equals the number of parameters (β). Thus, we set the sample moments exactly to zero and there is no weighting matrix ($a = I$). We can also solve for the estimate analytically

$$\widehat{\beta} = \left[E_T\left(x_t x_t'\right)\right]^{-1} E_T\left(x_t y_t\right).$$

[2]Thanks to John Cochrane (University of Chicago) for kindly sharing complimentary notes for facilitating the programming of this section 1.2.2.

This is the familiar OLS formula.

At this moment we need some distribution theory. Hansen [46, theorem 3.1] states that the asymptotic distribution of the GMM estimate is:

$$\sqrt{T}\left(\widehat{b}-b\right)\longrightarrow N\left[0,(ad)^{-1}aSa'(ad)^{-1\prime}\right], \tag{1.2.3}$$

where: $d\equiv E\left[\frac{\partial f}{\partial b'}(x_t,b)\right]=\frac{\partial g_T(b)}{\partial b'}$.

Precisely d is defined as the population moment in the identity which we estimate in sample by the equality, where

$$a\equiv \text{plim } a_T,$$
$$S\equiv\sum_{j=-\infty}^{\infty}E\left[f(x_t,b)\,,\ f(x_t,b)'\right].$$

In practical terms, this means to use

$$var\left(\widehat{b}\right)=\frac{1}{T}(ad)^{-1}aSa'(ad)^{-1\prime}. \tag{1.2.4}$$

Coming back to our task, according to (1.2.3), the rest of the ingredients for the general standard errors formula are

$$d=-E\left(x_tx_t'\right),$$
$$f(x_t,\beta)=x_t\left(y_t-x_t'\beta\right)=x_t\varepsilon_t.$$

Taking $a=I$, equation (1.2.3) gives a formula for OLS standard errors,

$$var_{GMM}\left(\widehat{\beta}\right)=\frac{1}{T}E\left(x_tx_t'\right)^{-1}\left[\sum_{j=-\infty}^{\infty}E\left(\varepsilon_tx_tx_{t-j}'\varepsilon_{t-j}\right)\right]E\left(x_tx_t'\right)^{-1}. \tag{1.2.5}$$

So, by mapping OLS regressions in to the GMM framework, we derive a formula (1.2.5) for OLS standard errors that correct for autocorrelation and conditional heteroskedasticity of the errors. In order to find the standard errors we only have to take the square root of the diagonal of (1.2.5).

Taking (1.2.5) and assuming serially uncorrelated and homoskedasticity errors (usual OLS assumptions) such that

$$\begin{aligned} E\left(\varepsilon_t \mid x_t, x_{t-1}, ..., \varepsilon_{t-1}, \varepsilon_{t-2}...\right) &= 0, \\ E\left(\varepsilon_t^2 \mid x_t, x_{t-1}, ..., \varepsilon_{t-1}, \varepsilon_{t-2}...\right) &= constant = \sigma_\varepsilon^2. \end{aligned}$$

The first assumption means that only the $j = 0$ term enters the sum

$$\sum_{j=-\infty}^{\infty} E\left(\varepsilon_t x_t x_{t-j}' \varepsilon_{t-j}\right) = E\left(\varepsilon_t x_t x_t' \varepsilon_t\right) = E\left(\varepsilon_t^2 x_t x_t'\right).$$

The second assumption means that

$$E\left(\varepsilon_t^2 x_t x_t'\right) = E\left(\varepsilon_t^2\right) E\left(x_t x_t'\right) = \sigma_\varepsilon^2 E\left(x_t x_t'\right).$$

Hence, equation (1.2.5) reduces to the well known expression,

$$var\left(\widehat{\beta}\right) = \tfrac{1}{T} E\left(x_t x_t'\right)^{-1} \left[\sigma_\varepsilon^2 E\left(x_t x_t'\right)\right] E\left(x_t x_t'\right)^{-1},$$

$$var_{OLS}\left(\widehat{\beta}\right) = \frac{1}{T} \sigma_\varepsilon^2 E\left(x_t x_t'\right)^{-1}. \tag{1.2.6}$$

Thus, equation (1.2.6) is a particular case of (1.2.5).

1.2.3 Tests statistics

We can use t-tests to check whether the pricing errors α are in fact zero. These distributions are usually presented for the case that the regression errors in (1.2.1) are uncorrelated and homoskedastic, but (1.2.5) show easily how to calculate standard errors for arbitrary error covariance structures.

We also want to know whether all the pricing errors are *jointly* equal to zero. This requires us to go beyond standard formulas for the regression (1.2.1) taken alone, as we want to know the joint distribution of α estimates from separate regressions running side by side but with errors correlated across assets $E\left(\varepsilon_t^i \varepsilon_t^j\right) \neq 0$.

χ^2 statistic

The classic form of these tests assumes no autocorrelation or heteroskedasticity. Dividing the $\widehat{\alpha}$ regression coefficients by their variance-covariance matrix leads to a χ^2 test,

$$T\left[1+\left(\frac{E_T\left(f\right)}{\widehat{\sigma}\left(f\right)}\right)^2\right]^{-1}\widehat{\alpha}'\widehat{\Sigma}^{-1}\widehat{\alpha}\sim\chi_N^2, \tag{1.2.7}$$

where $E_T\left(f\right)$ denotes sample mean, $\widehat{\sigma}^2\left(f\right)$ denotes sample variance, $\widehat{\alpha}$ is a vector of the estimated intercepts, $\widehat{\alpha}=\left[\widehat{\alpha}_1 \;\; \widehat{\alpha}_2 \; ... \; \widehat{\alpha}_N\right]'$, $\widehat{\Sigma}$ is the residual covariance matrix, i.e., the sample estimate of $E\left(\varepsilon_t\varepsilon_t'\right)=\Sigma$, where $\varepsilon_t=\left[\varepsilon_t^1 \;\; \varepsilon_t^2 \; ... \; \varepsilon_t^N\right]'$, $N=$ number of assets.

As usual when testing hypotheses about regression coefficients, expression (1.2.7) is valid asymptotically. The asymptotic distribution theory assumes that $\sigma^2\left(f\right)$ and Σ have converged to their probability limits; therefore, it is asymptotically valid even though the factor is stochastic and Σ is estimated, but ignores those sources of variation in a finite sample. It does not require that the errors are normal, relying on the central limit theorem so that $\widehat{\alpha}$ is normal.

From (1.2.7), and following the classic form of these tests, we can derive the variance-covariance matrix of $\widehat{\alpha}$:

$$var\left(\widehat{\alpha}\right) = T\left[1+\left(\frac{E_T(f)}{\widehat{\sigma}(f)}\right)^2\right]^{-1}\widehat{\Sigma}.$$

We now show the derivation of the χ^2 statistic and distributions with general errors. The last approach at (1.2.7) allows us to generate straightforwardly the required corrections for autocorrelated and heteroskedastic disturbances. MacKinlay and Richardson [80] advocate generalized method of moments (GMM) approaches to regression tests in this way. It also serves to remind us that GMM and the stochastic discount factor representation $p = E\left(mx\right)$ are not necessary paired; one can do a GMM estimate of an expected return-beta model, too. The mechanisms are only slightly different than we did to generate distributions for OLS regression coefficients, since we keep track of N OLS regressions simultaneously.

First, we write the equations for all N assets together in vector form,

$$R_t^e = \alpha + \beta f_t + \varepsilon_t.$$

Then, we use the usual OLS moments to estimate the coefficients,

$$g_T\left(b\right) = \begin{bmatrix} E_T\left(R_t^e - \alpha - \beta f_t\right) \\ E_T\left[\left(R_t^e - \alpha - \beta f_t\right) f_t\right]\end{bmatrix} = E_T\left(\begin{bmatrix}\varepsilon_t \\ f_t\varepsilon_t\end{bmatrix}\right) = 0.$$

These moments exactly identify the parameters, so the a matrix in $a g_T\left(\widehat{b}\right) = 0$ is the identity matrix. Solving, the GMM estimates are of course the OLS estimates,

$$\begin{aligned}\widehat{\alpha} &= E_T\left(R_t^e\right) - \widehat{\beta}E_T\left(f_t\right), \\ \widehat{\beta} &= \frac{E_T\left[\left(R_t^e - E_T\left(R_t^e\right)\right)f_t\right]}{E_T\left[\left(f_t - E_T\left(f_t\right)\right)f_t\right]} = \frac{cov_T\left(R_t^e, f_t\right)}{var_T\left(f_t\right)}.\end{aligned}$$

The d matrix in the general GMM formula (1.2.3) is

$$d \equiv \tfrac{\partial g_T(b)}{\partial b'} = -\begin{bmatrix} I_N & I_N E\left(f_t\right) \\ I_N E\left(f_t\right) & I_N E\left(f_t^2\right)\end{bmatrix} = -\begin{bmatrix} 1 & E\left(f_t\right) \\ E\left(f_t\right) & E\left(f_t^2\right)\end{bmatrix} \otimes I_N.$$

where I_N is an $N \times N$ identity matrix. The S matrix is

$$S = \sum_{j=-\infty}^{\infty} \begin{bmatrix} E\left(\varepsilon_t \varepsilon_{t-j}'\right) & E\left(\varepsilon_t \varepsilon_{t-j}' f_{t-j}\right) \\ E\left(f_t \varepsilon_t \varepsilon_{t-j}'\right) & E\left(f_t \varepsilon_t \varepsilon_{t-j}' f_{t-j}\right)\end{bmatrix}.$$

Using the GMM variance formula (1.2.4) with $a = I$, we have

$$var\left(\begin{bmatrix}\widehat{\alpha} \\ \widehat{\beta}\end{bmatrix}\right) = \frac{1}{T} d^{-1} S d^{-1\prime}. \tag{1.2.8}$$

At this point, we are done. The upper left-hand corner of (1.2.8) gives us $var\left(\widehat{\alpha}\right)$ and the test we are looking for is

$$\widehat{\alpha}' var\left(\widehat{\alpha}\right)^{-1} \widehat{\alpha} \sim \chi_N^2. \tag{1.2.9}$$

The standard formulas make this expression prettier by assuming that the errors are uncorrelated over time and not heteroskedastic. These assumptions simplify the S matrix, as for the standard OLS formula. If we assume that f and ε are independent as well as orthogonal, $E\left(f\varepsilon\varepsilon'\right) = E\left(f\right)E\left(\varepsilon\varepsilon'\right)$ and $E\left(f^2\varepsilon\varepsilon'\right) = E\left(f^2\right)E\left(\varepsilon\varepsilon'\right)$. If we assume that the errors are independent over time as well, we lose all the lead and lag terms. Then S matrix simplifies to

$$S = \begin{bmatrix} E\left(\varepsilon_t \varepsilon_t'\right) & E\left(\varepsilon_t \varepsilon_t'\right) E\left(f_t\right) \\ E\left(f_t\right) E\left(\varepsilon_t \varepsilon_t'\right) & E\left(\varepsilon_t \varepsilon_t'\right) E\left(f_t^2\right) \end{bmatrix} = \begin{bmatrix} 1 & E\left(f_t\right) \\ E\left(f_t\right) & E\left(f_t^2\right) \end{bmatrix} \otimes \Sigma.$$

Now we can plug into (1.2.8). Using $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$ and $(A \otimes B)(C \otimes D) = AC \otimes BD$, we obtain

$$var\left(\begin{bmatrix} \widehat{\alpha} \\ \widehat{\beta} \end{bmatrix}\right) = \tfrac{1}{T}\left(\begin{bmatrix} 1 & E\left(f_t\right) \\ E\left(f_t\right) & E\left(f_t^2\right) \end{bmatrix}^{-1} \otimes \Sigma\right).$$

Evaluating the inverse,

$$var\left(\begin{bmatrix} \widehat{\alpha} \\ \widehat{\beta} \end{bmatrix}\right) = \tfrac{1}{T}\tfrac{1}{var(f)} \begin{bmatrix} E\left(f_t^2\right) & -E\left(f_t\right) \\ -E\left(f_t\right) & 1 \end{bmatrix} \otimes \Sigma.$$

We are interested in the top left corner, where the pricing errors are. Using the variance property $E\left(f^2\right) = E\left(f\right)^2 + var\left(f\right)$,

$$var\left(\widehat{\alpha}\right) = T\left[1 + \left(\tfrac{E(f)}{\sigma(f)}\right)^2\right]^{-1} \Sigma.$$

This is the same formula as (1.2.7). Note that there is no reason to assume that errors are i.i.d. or independent of the factors. By simply calculating (1.2.8) we can easily construct standard errors and test statistics that do not require these assumptions.

Gibbons-Ross-Shanken test

As usual in a regression context, we can derive a finite-sample F distribution for the hypothesis that a set of parameters are jointly zero,

$$\frac{T-N-K}{N}\left[1 + \left(\frac{E_T\left(f\right)}{\widehat{\sigma}\left(f\right)}\right)^2\right]^{-1} \widehat{\alpha}'\widehat{\Sigma}^{-1}\widehat{\alpha} \sim F_{N,T-N-K}, \tag{1.2.10}$$

where K is the number of factors.

Expression (1.2.10) is known as the Gibbons, Ross and Shanken [41] or GRS test statistic. The F distribution recognizes sampling variation in $\widehat{\Sigma}$, which is not included in (1.2.7). This distribution requires that errors ε are normal as well as uncorrelated and homoskedastic. With normal errors, the $\widehat{\alpha}$ are normal and $\widehat{\Sigma}$ is an independent Wishart (the multivariate version of a χ^2), so the ratio is F. This distribution is exact in a finite sample.

All these tests have a very intuitive form. The basic part of the test is a quadratic form in the pricing errors $\widehat{\alpha}'\widehat{\Sigma}^{-1}\widehat{\alpha}$. If there were no βf in the model, then the $\widehat{\alpha}$ would simply be the sample mean of the regression errors ε_t. Assuming i.i.d. ε_t, the variance of their sample mean is just $\frac{1}{T\Sigma}$. Thus, if we knew Σ, then $T\widehat{\alpha}'\Sigma^{-1}\widehat{\alpha}$ would be a sum of squared sample means divided by their variance-covariance matrix, which would have an asymptotic χ^2_N distribution, or a finite-sample χ^2_N distribution if the ε_t are normal. But we have to estimate Σ, which is why the finite-sample distribution is F rather than χ^2.

1.3 Cross-sectional estimation and evaluation

We can fit $E\left(R^{ei}\right) = \beta_i'\lambda + \alpha_i$ by running a cross-sectional regression of average returns on the betas. This technique can be used whether the factor is a return or not. Skoulakis [96] and Jagannathan, Skoulakis, and Wang [55] provide excellent syntheses of the two-pass CSR (cross-sectional regression) methodology.

In this section, we discuss OLS and GLS cross-sectional regressions, and show formulas for the standard errors of λ estimators, and a χ^2 test whether the pricing errors α are jointly zero. As in the previous section we derive the distributions as an instance of GMM, and show how to implement the same approach for autocorrelated and heteroskedastic errors.

1.3.1 Regressions

Start again with the K factor model, the asset pricing model under the beta representation is given by

$$E\left(R^{ei}\right) = \beta_i'\lambda, \qquad i = 1, 2, ...N.$$

The central economic question is *why average returns vary across portfolios*; expected returns of an asset should be high if that asset has high betas or a large exposure of factors that carry high risk premia.

1.3.2 Ordinary least squares

The model says that average returns should be proportional to betas. Given this facts, a natural idea is to run a cross-sectional regression, but first, we have to find the estimates of the betas from time-series regressions. Rewriting equation 1.2.1:

$$R_t^{ei} = a_i + \beta_i' f_t + \varepsilon_t^i, \qquad t = 1, 2, ..., T \text{ for each } i.$$

Then estimate the factor risk premia λ from a regression across portfolios of average returns on the betas,

$$E_T\left(R^{ei}\right) = \beta_i'\lambda + \alpha_i, \qquad i = 1, 2, ..., N. \tag{1.3.1}$$

β are the right-hand variables, λ are the regression coefficients, and the cross-sectional regression residuals α_i are the pricing errors. This procedure is also known as a *two-pass* regression estimate, because one estimates first time-series and then cross-sectional regressions.

We can run (1.3.1) with or without a constant. The theory says that the constant or zero-beta excess return should be zero. One can impose this restriction or estimate a constant and see if it turns out to be small (we evaluate both alternatives in Chapter 2). The usual trade-off between efficiency (impose the null as much as possible to get efficient estimates) and robustness applies.

The OLS cross-sectional risk premium and pricing errors estimates are

$$\widehat{\lambda}_{OLS}^{\text{cross-section}} = \left(\beta'\beta\right)^{-1} \beta' E_T\left(R^e\right), \tag{1.3.2}$$

$$\widehat{\alpha}_{OLS}^{\text{cross-section}} = E_T\left(R^e\right) - \widehat{\lambda}\beta. \tag{1.3.3}$$

Again, (1.3.2) can be estimated with or without a constant.

Next, we need a distribution theory for the estimated parameters (1.3.2) and (1.3.3) in order to find their standard errors. The most natural place to start is with the standard OLS distribution formulas. We start with the traditional assumption that the true errors are i.i.d. over time, and independent of the factors.

In an OLS regression $Y = X\beta + u$ and $E\left(u\ u'\right) = \Omega$, the variance of the β estimate is $\left(X'X\right)^{-1} X'\Omega X \left(X'X\right)^{-1}$. The residual covariance matrix is:
$\left(I - X\left(X'X\right)^{-1} X'\right) \Omega \left(I - X\left(X'X\right)^{-1} X'\right)'$.

To apply these formulas we need the error covariance in the cross-sectional regression, $cov\left(\alpha, \alpha'\right)$. With the traditional assumption that the factors and errors are i.i.d. over time, the answer is $cov\left(\alpha, \alpha'\right) = \frac{1}{T}\left(\beta\Sigma_f\beta' + \Sigma\right)$, where $\Sigma_f \equiv cov\left(f_t, f_t'\right)$ and $\Sigma = cov\left(\varepsilon_t, \varepsilon_t'\right)$. To see this, start with $\alpha = E_T\left(R^e\right) - \beta\lambda$. With $R_t^e =$

$a + \beta f_t + \varepsilon_t$, we have $E_T\left(R_t^e\right) = a + \beta E_T\left(f_t\right) + E_T\left(\varepsilon_t\right)$. Under the null that the model is correct, so $E_T\left(R^e\right) = a + \beta E\left(f\right) = \beta\lambda$, consequently, we have $cov\left(\alpha, \alpha'\right) = cov\left[E_T\left(R^e\right), E_T\left(R^e\right)'\right] = \frac{1}{T}\left(\beta\Sigma_f\beta' + \Sigma\right)$. It is important to note that this covariance is not the same as the covariance of the *estimated* α in the cross sectional regression.

Then, the conventional OLS formulas for the covariance matrices of OLS estimates and residuals, accounting for correlated errors, give us

$$\sigma^2_{OLS}\left(\widehat{\lambda}\right) = \frac{1}{T}\left[\left(\beta'\beta\right)^{-1}\beta'\Sigma\beta\left(\beta'\beta\right)^{-1} + \Sigma_f\right], \tag{1.3.4}$$

$$cov_{OLS}\left(\widehat{\alpha}\right) = \frac{1}{T}\left[I - \beta\left(\beta'\beta\right)^{-1}\beta'\right]\Sigma\left[I - \beta\left(\beta'\beta\right)^{-1}\beta'\right]'. \tag{1.3.5}$$

Now, we test whether all pricing errors are zero with the overidentifying restrictions statistic

$$\widehat{\alpha}' cov\left(\widehat{\alpha}\right)\widehat{\alpha} \sim \chi^2_{N-K}. \tag{1.3.6}$$

Note that the distribution is χ^2_{N-K}, and not χ^2_N as in (1.2.7), because the covariance matrix is singular. The singularity and the extra terms in (1.3.5) result from the fact that the λ coefficient was estimated along the way, and means that we have to use a generalized inverse.

1.3.3 Generalized least squares

Since the residuals in the cross-sectional regression (1.3.1) are correlated with each other, the standard advice is to run a GLS cross-sectional regression rather than OLS, using $E\left(\alpha\alpha'\right) = \frac{1}{T}\left(\Sigma + \beta\Sigma_f\beta'\right)$ as the error covariance matrix.

The GLS cross-sectional risk premium and pricing error estimates are

$$\widehat{\lambda}_{GLS}^{\text{cross-section}} = \left(\beta'\Sigma^{-1}\beta\right)^{-1}\beta'\Sigma^{-1}E_T\left(R^e\right), \tag{1.3.7}$$

$$\widehat{\alpha}_{GLS}^{\text{cross-section}} = E_T\left(R^e\right) - \widehat{\lambda}\beta, \tag{1.3.8}$$

which are analogous to expressions (1.3.2) and (1.3.3).

Again, we could estimate (1.3.7) with or without a constant.

The GLS formula is in fact

$$\widehat{\lambda} = \left[\beta'\left(\beta\Sigma_f^{-1}\beta' + \Sigma\right)^{-1}\beta\right]^{-1}\beta'\left(\beta\Sigma_f^{-1}\beta' + \Sigma\right)^{-1}E_T\left(R^e\right).$$

However, as shown in the appendix, we can drop the $\beta\Sigma_f^{-1}\beta'$ term.

The standard regression formulas give the variance of (1.3.7) and (1.3.8) estimates as

$$\sigma_{GLS}^2\left(\widehat{\lambda}\right) = \frac{1}{T}\left[\left(\beta'\Sigma^{-1}\beta\right)^{-1} + \Sigma_f\right], \tag{1.3.9}$$

$$cov_{GLS}\left(\widehat{\alpha}\right) = \frac{1}{T}\left[\Sigma - \beta\left(\beta'\Sigma^{-1}\beta\right)^{-1}\beta'\right]. \tag{1.3.10}$$

The GLS regression should improve efficiency, i.e., give more precise estimates. However, Σ may be hard to estimate and to invert, especially if the cross section N is large. One may well choose the robustness of OLS over the asymptotic statistical advantages of GLS. We will review this issue in Chapter 2.

A GLS regression can be understood as a transformation of the space of returns, to focus attention on the statistically most informative portfolios. Finding (say, by Choleski decomposition) a matrix C such that $CC' = \Sigma^{-1}$, the GLS regression is the same as an OLS regression of $CE_T(R^e)$ on $C\beta$, i.e., of testing the model on the portfolios CR^e. The statistically most informative portfolios are those with the lowest residual variance Σ. But this asymptotic statistical theory assumes that the covariance matrix has converged to its true value.

As usual, we could test the hypothesis that all the α are equal to zero with (1.3.6). Though the appearance of the statistic is the same, the covariance matrix is smaller, reflecting the greater power of the GLS test. The expression is

$$T\widehat{\alpha}'\Sigma^{-1}\widehat{\alpha} \sim \chi^2_{N-K}. \tag{1.3.11}$$

A formal derivation of this test can be found in the appendix.

Shanken correction

In applying standard OLS formulas to a cross-sectional regression, we assume that the right-hand variables β are fixed. The β in the cross-sectional regression are not fixed, of course, but are estimated in the times series regression (see equation 1.3.1). This turn out to matter, even asymptotically. In this subsection, we derive the correct asymptotic standard errors. With the simplifying assumption that the errors ε are i.i.d. over time and independent of the factors, the result is

$$\sigma^2_{Shanken}\left(\widehat{\lambda}_{OLS}\right) = \frac{1}{T}\left[\left(\beta'\beta\right)^{-1}\beta'\Sigma\beta\left(\beta'\beta\right)^{-1}\left(1+\lambda'\Sigma_f^{-1}\lambda\right)+\Sigma_f\right], \tag{1.3.12}$$

$$\sigma^2_{Shanken}\left(\widehat{\lambda}_{GLS}\right) = \frac{1}{T}\left[\left(\beta'\Sigma^{-1}\beta\right)^{-1}\left(1+\lambda'\Sigma_f^{-1}\lambda\right)+\Sigma_f\right]. \tag{1.3.13}$$

Where Σ_f is the variance-covariance matrix of the factors. This correction is due to Shanken [92]. Comparing these standard errors to (1.3.4) and (1.3.9), we see that there is a multiplicative correction $\left(1+\lambda'\Sigma_f^{-1}\lambda\right)$.

The asymptotic variance-covariance matrix of the pricing errors is

$$cov_{Shanken}\left(\widehat{\alpha}_{OLS}\right)=\frac{1}{T}\left[I_N-\beta\left(\beta'\beta\right)^{-1}\beta'\right]\Sigma\left[I_N-\beta\left(\beta'\beta\right)^{-1}\beta'\right]\left(1+\lambda'\Sigma_f^{-1}\lambda\right), \tag{1.3.14}$$

$$cov_{Shanken}\left(\widehat{\alpha}_{GLS}\right)=\frac{1}{T}\left[\Sigma-\beta\left(\beta'\Sigma^{-1}\beta\right)^{-1}\beta'\right]\left(1+\lambda'\Sigma_f^{-1}\lambda\right). \tag{1.3.15}$$

Comparing these results to (1.3.5) and (1.3.10), we see the same multiplicative correction.

We can form the asymptotic χ^2 test of the pricing errors by dividing pricing errors by their variance-covariance matrix, $\widehat{\alpha}cov\left(\widehat{\alpha}\right)\widehat{\alpha}'$. Following (1.3.11), we can simplify this result for the GLS pricing errors resulting in

$$T\left(1+\lambda'\Sigma_f^{-1}\lambda\right)\widehat{\alpha}'_{GLS}\Sigma^{-1}\widehat{\alpha}_{GLS}\sim\chi^2_{N-K}. \tag{1.3.16}$$

Cochrane [25] show that, when using annual data, this term is too large to ignore. However, the mean and variance both scale with horizon, so the Sharpe ratio scales with the square root of horizon. Therefore, for a monthly interval the multiplicative term is quite small, so ignoring it makes little difference. In our estimation, this hypothesis holds very well, mainly when estimating the three factor model.

Comparing (1.3.16) to (1.2.7) and (1.2.10), we see basically the same statistic. The only difference is that by estimating λ from imposing $\lambda = E\left(f\right)$, the cross-sectional regression loses degrees of freedom equal to the number of factors K.

GMM

In this subsection, we will derive the formulas that do not require i.i.d. errors. This subsection has benefited from the work of Jagannathan, Skoulakis and Wang [54].

The easy and elegant way to account for the effects of generated regressors such as the β in the cross-sectional regression is to map it into GMM. Then, we treat the moments that generate the regressors β at the same time as the moments that generate the cross-sectional regression coefficient λ, and the covariance matrix S between the two sets of moments captures the effects of generating the regressors on the standard error of the cross-sectional regression coefficients. Comparing this straightforward derivation with the difficulty of Shanken's [92] paper that originally derived the corrections for $\widehat{\lambda}$, and noting that Shanken did not go on to find the formulas (1.3.14) that allow a test of pricing errors is a nice argument for the simplicity and power of the GMM framework.

To keep the algebra manageable, let's treat the case of a single factor f for now. The moments are

$$g_T(b) = \begin{bmatrix} E\left(R_t^e - a - \beta f_t\right) \\ E\left[\left(R_t^e - a - \beta f_t\right) f_t\right] \\ E\left(R^e - \beta\lambda\right) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

The top two moment conditions exactly identify a and β as the time-series OLS estimates[3]. The bottom moment condition is the asset pricing model. It is general overidentified in a sample, since there is only one extra parameter λ and N extra moment conditions. If we use the weighting vector β' on this condition, we obtain the OLS cross-sectional estimate of λ. If we use a weighting vector $\beta'\Sigma^{-1}$, we obtain the

[3]Note a not α. The time-series intercept is not necessarily equal to the pricing error in a cross-sectional regression.

GLS cross-sectional estimate of λ. To accommodate both cases, we use a weighting vector γ', and then substitute $\gamma' = \beta'$ or $\gamma' = \beta'\Sigma^{-1}$ at the end. However, once we abandon i.i.d. errors[4], the GLS cross-sectional regression weighted by Σ^{-1} is no longer the optimal estimate. Once we recognize that the errors do not obey classical assumptions, and if we want efficient estimates, we might as well calculate the correct and fully efficient estimates. Having decided on a cross-sectional regression, the efficient estimates of the previous set of moments are $d'S^{-1}g_T(a, \beta, \lambda) = 0$.

The standard errors for $\widehat{\lambda}$ come straight from the general GMM standard error formula (1.2.4). The $\widehat{\alpha}$ are not parameters, but are the last N moments. Their covariance matrix is thus given by the Hansen's [46, Lemma 4.1]

$$\sqrt{T}g_T\left(\widehat{b}\right) \rightarrow N\left[0, \left(I - d\left(ad\right)^{-1}a\right)S\left(I - d\left(ad\right)^{-1}a\right)'\right].$$

All we have to do now is map the problem into the GMM notation. The parameter vector is

$$b' = \left[a' \quad \beta' \quad \lambda\right].$$

The d matrix is the sensitivity of the moment conditions to the parameters,

$$d = \frac{\partial g_T}{\partial b'} = \begin{bmatrix} -I_N & -I_N E\left(f\right) & 0 \\ -I_N E\left(f\right) & -I_N E\left(f^2\right) & 0 \\ 0 & -\lambda I_N & -\beta \end{bmatrix}.$$

The S matrix is the long-run covariance matrix of the moments,

[4] i.i.d. stands for independent and identical distribution.

$$
\begin{aligned}
S &= \sum_{j=-\infty}^{\infty} E\left(\begin{bmatrix} R_t^e - a - \beta f_t \\ \left(R_t^e - a - \beta f_t\right) f_t \\ R_t^e - \beta\lambda \end{bmatrix} \begin{bmatrix} R_{t-j}^e - a - \beta f_{t-j} \\ \left(R_{t-j}^e - a - \beta f_{t-j}\right) f_{t-j} \\ R_{t-j}^e - \beta\lambda \end{bmatrix}'\right), \\
S &= \sum_{j=-\infty}^{\infty} E\left(\begin{bmatrix} \varepsilon_t \\ \varepsilon_t f_t \\ \beta\left(f_t - Ef\right) + \varepsilon_t \end{bmatrix} \begin{bmatrix} \varepsilon_{t-j} \\ \varepsilon_{t-j} f_{t-j} \\ \beta\left(f_{t-j} - Ef\right) + \varepsilon_{t-j} \end{bmatrix}'\right).
\end{aligned}
$$

In the second expression, we have used the regression model and the restriction under the null that $E\left(R_t^e\right) = \beta\lambda$. In calculations, of course, one could simply estimate the first expression.

We have the elements to calculate GMM standard error formula (1.2.4) and formula for the covariance of moments showed before. Now with a vector f, the moments are

$$
ad = \begin{bmatrix} I_N \otimes I_{K+1} & \\ & \gamma' \end{bmatrix} \begin{bmatrix} E_T\left(R_t^e - a - \beta f_t\right) \\ E_T\left[\left(R_t^e - a - \beta f_t\right) \otimes f_t\right] \\ E\left(R^e - \beta\lambda\right) \end{bmatrix} = 0,
$$

Where $\beta_i = N \times 1$, $\gamma' = \beta'$ for OLS and $\gamma' = \beta'\left(\Sigma^{-1}\right)$ for GLS.

Note that the GLS estimate is *not* the efficient GMM estimate when returns are not i.i.d. The efficient GMM estimate is $d'S^{-1}g_T = 0$. The d matrix is

$$
d = \frac{\partial g_T}{\partial(\alpha' \; \beta' \; \lambda')} = -\begin{bmatrix} 1 & E\left(f'\right) \\ E\left(f\right) & E\left(ff'\right) \\ 0 & \lambda \end{bmatrix} \otimes I_N \begin{bmatrix} 0 \\ 0 \\ \beta \end{bmatrix}.
$$

We can recover the classic formulas (1.3.12), (1.3.13), (1.3.14), (1.3.15) by adding the assumption that the errors are i.i.d. and independent of the factors, and that the factors are uncorrelated over time as well. The assumption that the errors and factors are uncorrelated over time means we can ignore the lead and lag terms. Thus, the top left corner of S is $E\left(\varepsilon_t\varepsilon_t'\right) = \Sigma$. The assumption that the errors are independent from the factors f_t simplifies the terms in which ε_t and f_t are multiplied: $E\left(\varepsilon_t\left(\varepsilon_t' f_t\right)\right) = E\left(f\right)\Sigma$ for example. The result is

$$S = \begin{bmatrix} \Sigma & E\left(f\right)\Sigma & \Sigma \\ E\left(f\right)\Sigma & E\left(f^2\right)\Sigma & E\left(f\right)\Sigma \\ \Sigma & E\left(f\right)\Sigma & \beta\beta'\sigma^2\left(f\right) + \Sigma \end{bmatrix}.$$

Multiplying a, d, S together as specified by the GMM formula for the covariance matrix of parameters (1.2.4), we obtain the covariance matrix of all the parameters, and its $(3,3)$ element gives the variance of $\widehat{\lambda}$.

Once again, there is really no need to make the assumption that the errors are i.i.d. and especially that they are conditional heteroskedastic. It is quite easy to estimate an S matrix that does not impose those conditions and calculate standard errors. They will not have the pretty analytic form given above, but they will more closely report the true sampling uncertainty of the estimate. Furthermore, if one is really interested in efficiency, the GLS cross-sectional estimate should use the spectral density matrix as weighting matrix applied to all the moments rather than the Σ^{-1} applied only to the pricing errors.

1.4 Fama-MacBeth estimation and evaluation

We introduce the Fama-MacBeth procedure for running cross-sectional regression and calculating standard errors that correct for cross-sectional correlation. This section is partially based on Jagannathan and Wang [57], who illustrate the asymptotic distribution theory for the two-stage cross-sectional regression method and the FamaMacBeth procedure[5]; and Cochrane [25].

1.4.1 Regression

Fama and MacBeth [32] suggest an alternative procedure for running cross-sectional regressions, and for producing standard errors and test statistics.

First, we find beta estimates with a time-series regression. Fama and MacBeth use rolling 5-years regressions, but one can also use the technique with full-sample betas as we actually did in this work. Second, instead of estimating a single cross-sectional regression with the sample averages, we now run a cross-sectional regression *at each time period*, i.e.,

$$R_t^{ei} = \beta_i' \lambda_t + \alpha_{it}, \qquad i = 1, 2, ..., N \qquad \text{for each } t. \tag{1.4.1}$$

1.4.2 Estimators

Then, Fama and MacBeth suggest that we estimate λ and α_i as the average of the cross-sectional regression estimates,

$$\widehat{\lambda} = \tfrac{1}{T}\sum_{t=1}^{T} \widehat{\lambda}_t, \qquad \widehat{\alpha}_i = \tfrac{1}{T}\sum_{t=1}^{T} \widehat{\alpha}_{it}.$$

Most importantly, they suggest that we use the standard deviations of the cross-sectional regression estimates to generate the sampling errors for these estimates,

$$\sigma^2\left(\widehat{\lambda}\right) = \tfrac{1}{T^2}\sum_{t=1}^{T}\left(\widehat{\lambda}_t - \widehat{\lambda}\right)^2, \qquad \sigma^2\left(\widehat{\alpha}_i\right) = \tfrac{1}{T^2}\sum_{t=1}^{T}\left(\widehat{\alpha}_{it} - \widehat{\alpha}_i\right)^2.$$

[5] In particular, they show that without the assumption of conditional homoskedasticity, previously imposed by Shanken [92], a general asymptotic distribution theory for the two-stage cross-sectional regression method shows that the standard errors produced by the FamaMacBeth procedure do not necessarily overstate the precision of the risk premium estimates.

Note that it is $\frac{1}{T^2}$ because we are finding standard errors of sample means, $\frac{\sigma^2}{T}$.

This is an intuitively appealing procedure. Sampling error is, after all, about how statistic would vary from one sample to the next if we repeated the observations. We cannot do that with only one sample, but why not cut the sample in half, and deduce how a statistic would vary from one sample to the next from how it varies from the first half to the sample to the next half? Proceeding, why not cut the sample in fourths, eights, and so on? The Fama-MacBeth procedure carries this idea to its logical conclusion, using the variation in the statistic $\widehat{\lambda}_t$ over time to deduce its variation across samples.

We are used to deducing the sampling variance of the sample mean of a series x_t by looking at the variation of x_t through time in sample, using $\sigma^2(\overline{x}) = \frac{\sigma^2(x)}{T} = \frac{1}{T^2}\sum_t (x_t - \overline{x})^2$. The Fama-MacBeth technique just applies this idea to the slope and pricing error estimates. The formula assumes that the time series is not autocorrelated, but one could easily extend the idea to estimates $\widehat{\lambda}_t$ that are correlated over time by using a long-run variance matrix, i.e., estimate

$$\sigma^2\left(\widehat{\lambda}\right) = \frac{1}{T}\sum_{j=-\infty}^{\infty} cov_T\left(\widehat{\lambda}_t, \widehat{\lambda}_{t-j}\right).$$

1.4.3 Tests statistics

It is natural to use this sampling theory to test whether all the pricing errors are jointly zero as we have done before. Denote by α the vector of pricing errors across assets. We could estimate the covariance matrix of the sample pricing errors by

$$
\begin{aligned}
\widehat{\alpha} &= \frac{1}{T}\sum_{t=1}^{T}\widehat{\alpha}_t, \\
cov(\widehat{\alpha}) &= \frac{1}{T^2}\sum_{t=1}^{T}\left(\widehat{\alpha}_t - \widehat{\alpha}\right)\left(\widehat{\alpha}_t - \widehat{\alpha}\right)',
\end{aligned}
$$

or a general version that accounts for correlation over time, and finally use the test

$$
\widehat{\alpha}_t' pinv\left(cov(\widehat{\alpha})\right)\widehat{\alpha}_t \sim \chi^2_{N-K}, \tag{1.4.2}
$$

where *pinv* is the Moore-Penrose pseudoinverse of a matrix.

1.5 General method of moments / Stochastic discount factor - estimation and evaluation

The use of the stochastic discount factor (SDF) method for econometric evaluation of asset pricing models has become common in the recent empirical finance literature. A SDF has the following property: The value of a financial asset equals the expected value of the product of the payoff on the asset and the SDF. An asset pricing model identifies a particular SDF that is a function of observable variables and model parameters. For example, a linear factor pricing model identifies a specific linear function of the factors as a SDF. The SDF method involves estimating the asset pricing model using its SDF representation and the generalized method of moments (GMM). As Cochrane [25] points out, the SDF method is sufficiently general that it can be used for analysis of linear as well as nonlinear asset-pricing models, including pricing models for derivative securities.

The SDF representation can be traced back to Dybvig and Ingersoll [28], who derive the SDF representation for the CAPM. Ingersoll [53] derives the SDF representation for a number of theoretical asset pricing models[6]. Hansen and Jagannathan [48] and [49] develop diagnostic tests for asset pricing models based on the SDF representation. Ferson [35], Campbell, Lo, and MacKinlay [17], Velu and Zhou [101], Marín and Rubio [81], and Cochrane [25] provide introductions to the stochastic discount factor framework.

One can summarize asset pricing by two equations:

$$
\begin{aligned}
p_t &= E_t\left(m_{t+1} x_{t+1}\right), \\
m_{t+1} &= f\left(\text{data, parameters}\right),
\end{aligned}
$$

where p_t is the current price of the security, E_T is the conditional expectation given information up to time point t, m_{t+1} is the stochastic discount factor, and x_{t+1} is the payoff of the asset at time point $t+1$.

This approach allows us to conveniently separate the setup of specifying economic assumptions of the model (second equation) from the step of deciding which kind of empirical representation to pursue or understand.

The development of the generalized method of moments (GMM) by Hansen [46] has had a major impact on empirical research in finance, especially in the area of asset pricing because it allows for conditional heteroscedasticity, serial correlation, and non-normal distributions. GMM has made econometric evaluation of asset-pricing models

[6]However, he does not use the term "stochastic discount factor" — Hansen and Richard [50] coined the term.

possible under more realistic assumptions regarding the nature of the stochastic process governing the temporal evolution of exogenous variables.

The GMM approach is a natural fit for a discount factor formulation of asset pricing theories, since we just use sample moments in the place of population moments. As we will describe, there is no singular GMM estimate and test. As Cochrane [25] indicates, GMM is a large canvas and a big set of paints and brushes; a flexible tool for doing all kind of sensible things to the data.[7] See for example Hall [44] for an extensive exposition and applications of the GMM method in finance and econometrics.

In this section we will review the estimation and testing of linear discount factor models expressed as $p = E(mx)$, $m = b'f$ (for notational convention, time subscripts are usually deleted). This form naturally suggests a GMM approach using the pricing errors as moments. The resulting estimates look a lot like the regression estimates of the pasts sections.

1.5.1 Introduction to the GMM methodology

GMM can be applied in exactly the same way as described earlier to estimate the asset-pricing model parameters and test the overidentifying restrictions implied by the asset-pricing model using its SDF representation.

[7] Although it is well known that the GMM estimates are no more efficient than the maximum likelihood estimates, the advantages of the maximum likelihood estimates vanishes if one does not know the joint distribution of the returns and the factors. If we make the wrong distribution assumption, the maximum likelihood estimates can be biased, while the GMM does not suffer from the same problem. This point is well explained by Hansen and Singleton [51]. For further details on the ML approach see Gibbons [40].

First-stage estimators: $W = I$

If we could, we would pick b to make every element of $g_T(b) = 0$ – to have the model price perfectly in sample. However, there are usually more moment conditions (returns times instruments) than there are parameters. There should be, because theories with as many free parameters as facts (moments) are vacuous. Thus, we choose b to make the pricing errors $g_T(b)$ as small as possible, by minimizing the quadratic form,

$$\min_{\{b\}} \left[g_T(b)' W g_T(b) \right] . \tag{1.5.1}$$

W is a positive definite *weighting matrix* that tells us how much attention to pay to each moment, or how to trade off doing well in pricing one asset or linear combination of assets versus doing well in pricing another. When imposing $W = I$, GMM treats all portfolios symmetrically, and the objective is to minimize the sum of squared pricing errors.

The result of making such simplification (use the identity as the weighting matrix) is what we will call *first-stage estimators*. This estimator is consistent and asymptotically normal.

Using the identity matrix, as a *prespecified* weighting matrix, weights the initial choice of assets or portfolios equally in estimation and evaluation. This choice has a particular advantage with large systems in which S is nearly singular, as it avoids most of the problems associated with inverting a near-singular S matrix. Many empirical asset pricing studies use OLS cross-sectional regressions, which are the same thing as first-stage GMM estimate with an identity weighting matrix.

Thus, the first-stage estimators assumes no serial correlation and regression errors independent of right-hand variables.

Second-stage estimators: $W = S^{-1}$

This second-stage estimate picks a weighting matrix based on *statistical* considerations. Some assets returns may have much more variance than others, as we will show in next chapters. For those assets, the sample mean $g_T = E_T\left(m_t R_t - 1\right)$ will be a much less accurate measurement of the population mean $E\left(mR - 1\right)$, since the sample mean will vary more from sample to sample. Hence, it seems like a good idea to pay less attention to pricing errors from assets with high variance of $m_t R_t - 1$. One could implement this idea by using a W matrix composed of inverse variances of $E_T\left(m_t R_t - 1\right)$ on the diagonal. More generally, since asset returns are correlated, one might think of using the covariance matrix of $E_T\left(m_t R_t - 1\right)$. This weighting matrix pays most attention to *linear combinations* of moments about which data set at hand has the most information. This idea is exactly the same as heteroskedasticity and cross-correlation corrections that lead you from OLS to GLS in linear regressions.

The covariance matrix of $g_T = E_T\left(u_{t+1}\right)$ is the variance of a sample mean. Exploiting the assumption that $E\left(u_t\right) = 0$ and that u_t is stationary so $E\left(u_1 u_2\right) = E\left(u_t u_{t+1}\right)$ depends only on the time interval between the two u's, we have

$$\begin{aligned}
var\left(g_T\right) &= var\left(\frac{1}{T}\sum_{t=1}^{T} u_{t+1}\right), \\
var\left(g_T\right) &= \frac{1}{T^2}\left[T E\left(u_t u_t'\right) + (T-1)\left(E\left(u_t u_{t-1}'\right) + E\left(u_t u_{t+1}'\right)\right) + ...\right].
\end{aligned}$$

As $T \to \infty$, $\frac{(T-j)}{T} \to 1$, so $var\left(g_T\right) \to \frac{1}{T}\sum_{j=-\infty}^{\infty} E\left(u_t u_{t-j}'\right) = \frac{1}{T} S$.

The last equality denotes S, known for other reasons as the *spectral density matrix at frequency zero* of u_t[8].

This fact suggest that a good weighting matrix might be the inverse of S. In fact, Hansen [46] shows formally that the choice

$$W = S^{-1}, \qquad S \equiv \sum_{j=-\infty}^{\infty} E\left(u_t u'_{t-j}\right),$$

is the statistically optimal weighting matrix, meaning that it produces estimates with lowest asymptotic variance.

Different values of j can change dramatically the second-stage estimates as shown in (1.5.1), so we choose $j = 0$ and $j = 12$. Note that when $j = 0$ (0 lag estimate) we allow conditional heteroskedasticity, but no time-series correlation of residuals; and when $j = 12$ (12 lag, Newey-West estimate) is a correction of one year autocorrelation. Note that neither 0 nor 12 is the optimal number of lags, these fixed values for j are intended to compare two states for the second-stage estimates.

The first- and second-stage estimates should remind us of standard linear regression models. We start with OLS regression. If the errors are not i.i.d., the OLS estimates are consistent, but not efficient. If we want efficient estimates, we can use the OLS estimates to obtain a series of residuals, estimate a variance-covariance matrix of residuals, and then do GLS. GLS is also consistent and more efficient, meaning that the sampling variation in the estimated parameters is lower. Hall [44] illustrate a more detailed discussion about GMM first and second-stage properties.

[8]Precisely, S so defined is the variance-covariance matrix of g_T for fixed b. The actual variance-covariance matrix of g_T must take into account the fact that we choose b to set a linear combination of the g_T to zero in each sample.

1.5.2 Second moment matrix as the weighting matrix in second-stage estimators

Another example of prespecified economically interesting weighting matrix is the second moment matrix of returns and factors, advocated by Hansen and Jagannathan [49]. Hence, for this subsection 1.5.2 we will refer to S as the second moment matrix $E(xx') = cov(x) + E(x)E(x)'$.

Writing the model as $m = a - b'f$, we cannot separately identify a and b so we have to choose some normalization. The choice is entirely one of convenience; lack of identification means precisely that the pricing errors do not depend on the choice of normalization. The easiest choice is $a = 1$. Then $g_T(b) = -E_T(mR^e) = -E_T(R^e) + E(R^e f')b$. We have $d = \frac{\partial g_T(b)}{\partial b'} = E(R^e f')$, the second moment matrix of returns and factors. The first-order condition to minimize (1.5.1) is

$$d'W[E_T(R^e) - db] = 0,$$

Looking at (1.5.1), the first stage has $W = I$, the second stage has $W = S^{-1}$. Since this is a linear model, we can solve analytically for the GMM estimate, and it is

$$\begin{aligned} \text{First stage} &: \widehat{b}_1 = (d'd)^{-1} d'E_T(R^e), \qquad (1.5.2) \\ \text{Second stage} &: \widehat{b}_2 = \left(d'S^{-1}d\right)^{-1} d'S^{-1}E_T(R^e). \end{aligned}$$

The GMM estimate is a cross-sectional regression of mean excess returns on second moment matrix with factors. We find the distribution theory from the usual GMM standard error formulas in equations (1.2.3) and (1.2.4), this will help us on deriving the standard errors of $\widehat{b}_1$ and $\widehat{b}_2$. In the first stage, $a = d'$:

$$\begin{aligned} \text{First stage} &: \quad cov\left(\widehat{b}_1\right) = \frac{1}{T}\left(d'd\right)^{-1} d'Sd\left(d'd\right)^{-1}, \\ \text{Second stage} &: \quad cov\left(\widehat{b}_2\right) = \frac{1}{T}\left(d'S^{-1}d\right)^{-1}. \end{aligned} \tag{1.5.3}$$

For testing, we also need the covariance matrix of the pricing errors $g_T\left(\widehat{b}\right)$:

$$\begin{aligned} \text{First stage} &: \quad Tcov\left[g_T\left(\widehat{b}\right)\right] = \left(I - d\left(d'd\right)^{-1} d'\right) S \left(I - d\left(d'd\right)^{-1} d'\right), \\ \text{Second stage} &: \quad Tcov\left[g_T\left(\widehat{b}\right)\right] = S - d\left(d'S^{-1}d\right)^{-1} d'. \end{aligned} \tag{1.5.4}$$

These are obvious analogues to the standard regression formulas for the covariance matrix of regression residuals. The model test is

$$\begin{aligned} \text{First stage} &: \quad g_T\left(\widehat{b}\right)' cov\left(g_T\right)^{-1} g_T\left(\widehat{b}\right) \sim \chi^2_{(N-K)}, \\ \text{Second stage} &: \quad Tg_T\left(\widehat{b}\right)' S^{-1} g_T\left(\widehat{b}\right) \sim \chi^2_{(N-K)}. \end{aligned} \tag{1.5.5}$$

As usual, the test is a quadratic form in the vector of pricing errors. This test is known as the test of the model's overidentifying restrictions or Hansen's J-statistic [46], and it is often used as a specification test to examine whether the data are consistent with the model. When the linear factor pricing model holds, the J_T statistic converges to a central χ^2 distribution as T becomes large.

Notice, however, that there are two ways to get small value of the J_T statistic: generate small pricing errors with a high degree of precision, or generate a large pricing errors with even higher standard deviation of those errors. In the empirical work presented in the following sections, it will be important to distinguish between these two situations.

One interesting and, to the best of our knowledge, original result is to point out that both first and second-stage tests in 1.5.5 lead to the same value [9]. This equality does not hold when taking the covariance as the weighting matrix in the second-stage estimators, even though the numerical values are similar. Neither in OLS and GLS respective tests in a cross-sectional analysis, even though similarities arise when T is sufficiently long. Thus, it only holds when taking the second moment matrix as in equation 1.5.5.

The main implication that both first and second-stage tests in 1.5.5 lead to the same value is that regardless of testing the model with first or second-stage estimators, the result or p-value will be the same. Naturally, this does not imply a failure of the test. We know that second-stage estimators will lead to higher pricing errors in order to achieve more efficient estimators. Therefore, first stage estimators leads to lower pricing errors than second-stage estimators, and the weighting matrix in 1.5.5 compensate the test in such way that the value of both stages turns out to be the same.

1.5.3 Covariance as the weighting matrix in second-stage estimators

In this subsection we follow Cochrane [25] methodology, which have been also followed in recent works such as in Kan, Robotti, and Shanken [65]. The main idea es that we can obtain a cross-sectional regression of mean excess returns on *covariances*, which are just a heartbeat away from betas, by choosing the normalization $a = 1 + b'E(f)$ rather than $a = 1$. Then, the model is $m = 1 - b'[f - E(f)]$ with mean $E(m) = 1$.

[9] This issue has been briefly discussed with John Cochrane.

The pricing errors are $g_T(b) = E_T(mR^e) = E_T(R^e) - E_T\left(R^e\widetilde{f'}\right)b$, where we denote $\widetilde{f'} \equiv f - E(f)$. We have $d = \frac{\partial g_T(b)}{\partial b'} = E\left(R^e\widetilde{f'}\right)$, which now denotes the covariance matrix of returns and factors. The first-order condition to minimize (1.5.1) is now $-d'W\left[E_T(R^e) - db\right] = 0$. Then, the GMM estimates of b are

$$\begin{aligned} \text{First stage} &: \ \widehat{b}_1 = (d'd)^{-1} d' E_T(R^e), \\ \text{Second stage} &: \ \widehat{b}_2 = \left(d'S^{-1}d\right)^{-1} d'S^{-1}E_T(R^e). \end{aligned} \tag{1.5.6}$$

The GMM estimate is a cross-sectional regression of expected excess returns on the covariance between returns and factors. Naturally, the model says that expected excess returns should be proportional to the covariance between returns and factors, and we estimate that relation by a linear regression. The standard errors and variance of the pricing errors are the same as in (1.5.3) and (1.5.4), with d now representing the covariance matrix.

We must bear in mind that the mean of the factor $E(f)$ is estimated (as well as b), and the distribution theory should recognize sampling variation induced by this fact as we did for the fact that betas are generated regressors in the cross-sectional regressions. We can write the model as

$$\begin{aligned} E\left[R^e\left(1 - (f - Ef)'\,b\right)\right] &= 0, \\ E(R^e) &= cov(R^e f')\,b. \end{aligned}$$

The moments g are

$$g_T = \begin{bmatrix} E_T\left[R^e - R^e(f' - Ef')\,b\right] \\ E_T(f - Ef) \end{bmatrix}.$$

Ef represents the unknown parameter mean of the factors. The d matrix giving the derivative of moments with respect to parameters (b', Ef') is

$$d = \begin{bmatrix} -E\left(R^e \widetilde{f}'\right) & E\left(R^e\right) b' \\ 0 & -I_K \end{bmatrix},$$

where $\widetilde{f} = f - Ef$ and K is the number of factors.

We choose as a first stage estimate an OLS cross-sectional regression. The a matrix

$$a_T = \begin{bmatrix} E_T\left(\widetilde{f} R^{e\prime}\right) & 0 \\ 0 & -I_K \end{bmatrix},$$

generates the cross-sectional regression estimate of b in the first row and $Ef = E_T\left(f\right)$ in the second row,

$$\begin{aligned} \widehat{b} &= \left[C'C\right]^{-1} C' E_T\left(R^e\right), \\ E\left(f\right) &= E_T\left(f\right), \end{aligned}$$

where $C \equiv E\left(R^e \widetilde{f}'\right)$ denotes the covariance matrix of returns and factors.

To find the standard errors, we now just plug in to the general GMM formulas. The general formula is

$$cov \begin{pmatrix} \widehat{b} \\ \widehat{E}f \end{pmatrix} = \frac{1}{T}\left(ad\right)^{-1} aSa'\left(ad\right)^{-1\prime}. \tag{1.5.7}$$

Filling the pieces, the S matrix is

$$S = \sum_{j=-\infty}^{\infty} E \begin{bmatrix} u_t u'_{t-j} & u_t \widetilde{f}'_{t-j} \\ \widetilde{f}'_t u'_{t-j} & \widetilde{f}'_t \widetilde{f}'_{t-j} \end{bmatrix} = \begin{bmatrix} S_{uu} & S_{uf'} \\ S_{f'u} & S_{ff} \end{bmatrix},$$
$$u_t \equiv R^e_t \left(1 - \widetilde{f}'_t b\right).$$

The covariance of pricing errors $cov(g_T(\widehat{b}))$ follows from the usual general formula. The first stage standard errors look a lot like Shanken correction.

$$
\begin{aligned}
ad &= \begin{bmatrix} -C' & 0 \\ 0 & -I_K \end{bmatrix} \begin{bmatrix} -C & E\left(R^e\right) b' \\ 0 & -I_K \end{bmatrix} \\
ad &= \begin{bmatrix} C'C & -C'E\left(R^e\right) b' \\ 0 & I_K \end{bmatrix} \\
(ad)^{-1} &= \begin{bmatrix} \left(C'C\right)^{-1} & \left(C'C\right)^{-1} C'E\left(R^e\right) b' \\ 0 & I_K \end{bmatrix} \\
(ad)^{-1} a &= -\begin{bmatrix} \left(C'C\right)^{-1} C' & \left(C'C\right)^{-1} C'E\left(R^e\right) b' \\ 0 & I_K \end{bmatrix}.
\end{aligned}
$$

Under the null, asymptotically, we will have $E\left(R^e\right) = Cb$, so we can simplify the formula now with that substitution.

$$
(ad)^{-1} a = -\begin{bmatrix} \left(C'C\right)^{-1} C' & bb' \\ 0 & I_K \end{bmatrix}.
$$

We are interested in the top left and bottom right elements of

$$
\tfrac{1}{T} (ad)^{-1} a \begin{bmatrix} S_{uu'} & S_{uf'} \\ S_{f'u} & S_{ff'} \end{bmatrix} a' (ad)^{-1\prime}.
$$

The bottom right element is thus $var\left[E_T\left(f\right)\right] = \frac{1}{T} S_{ff}$.

This is the standard formula for the variance of the sample mean. The top left element is

$$
var\left(\widehat{b}\right) =
$$

$$
\tfrac{1}{T}\left(\left(C'C\right)^{-1} C'S_{uu'}C\left(C'C\right)^{-1} + bb'S_{ff'}bb' + \left(C'C\right)^{-1} C'S_{uf'}bb' + bb'S_{f'u}C\left(C'C\right)^{-1}\right).
$$

This equation reminds an important issue about the correction for cross-sectional regressions of average returns on betas. The first term is the same standard error we derived ignoring sampling variation in the sample mean Ef, and looks like the usual formula for OLS regression of expected returns on covariances C, corrected for residual covariation $S_{uu'}$. The second term is a lot like the term Σ_f in the cross-sectional regression formulas. The remaining terms add the effects of the fact that the sample mean must be estimated, as the extra terms in the Shanken formula correct for the fact that the betas had to be estimated. The presence of $b'S_{ff}b$ in the formulas is a lot like $\lambda'\Sigma_f^{-1}\lambda = \left(b'\Sigma_f^{-1}\right)\Sigma_f^{-1}\left(b\Sigma_f^{-1}\right)$ in the OLS regression formula.

The general pricing error test also has a simple form for the first-stage estimate. Use the general formula,

$$T cov\left[g_T\left(\widehat{b}\right)\right] = \left(I - d\,(ad)^{-1}\,a\right) S \left(I - d\,(ad)^{-1}\,a\right)'.$$

We have $(ad)^{-1}\,a$. Then

$$\begin{aligned}
I - (ad)^{-1}\,a &= I - \begin{bmatrix} C & E\left(R^e\right)b' \\ 0 & -I_K \end{bmatrix} \begin{bmatrix} \left(C'C\right)^{-1}C' & bb' \\ 0 & I_K \end{bmatrix} \\
&= \begin{bmatrix} I - C\left(C'C\right)^{-1}C' & \left(E\left(R^e\right) - Cb\right)b' \\ 0 & 0 \end{bmatrix} \\
&= \begin{bmatrix} I - C\left(C'C\right)^{-1}C' & 0 \\ 0 & 0 \end{bmatrix}.
\end{aligned}$$

Under the null, $E\left(R^e\right) - Cb = 0$, so the top right term vanishes. Thus, only the top $N \times N$ block survives. We should have expected this, as the $E\left(f\right)$ moments are set to zero in every sample. The top left part, in which we are interested, gives us

$$cov\left(\widehat{\alpha}_1\right) = \tfrac{1}{T}\left[I - C\left(C'C\right)^{-1}C'\right] S_{uu'} \left[I - C\left(C'C\right)^{-1}C'\right]'.$$

Thus, the pricing error test statistic is not affected by the fact that the factor mean $E(f)$ is estimated. This is natural, again, since the $E(f)$ moments are set to zero in each sample.

We might set up a second stage GMM by using the standard weighting matrix on the first set of moments only. In other words, if we wanted estimates b corresponding to: min $\left[E_T(mR^e)' S_{uu'}^{-1} E_T(mR^e)\right]$, we would set

$$a_T = \begin{bmatrix} E_T\left(\widetilde{f} R^{e\prime}\right) S_{uu'}^{-1} & 0 \\ 0 & -I_K \end{bmatrix}.$$

This mirrors what we did with the GLS cross-sectional regression. But this is *not* the efficient estimate. The efficient estimate is formed by $a = d'S^{-1}$, i.e.,

$$a = \begin{bmatrix} -E\left(\widetilde{f} R^{e\prime}\right) & 0 \\ bE(R^{e\prime}) & -I_K \end{bmatrix} \begin{bmatrix} S_{uu'} & S_{uf'} \\ S_{f'u} & S_{ff'} \end{bmatrix}^{-1}.$$

The two approaches are not the same. Intuitively, the fact that we must estimate Ef spills in to the optimal set of moments for estimating b. This estimate will allow us to estimate $Ef \neq E_T(f)$ if so doing helps a lot on the other moments, and that is what an efficient estimate should do.

The second-stage efficient estimate is

$$\begin{aligned} d'S^{-1} g_T \begin{bmatrix} b \\ f \end{bmatrix} &= 0, \\ d'S^{-1} \begin{bmatrix} E_T\left[R^e - R^e(f' - Ef')\, b\right] \\ E_T(f - Ef) \end{bmatrix} &= 0. \end{aligned}$$

The top term is quadratic in $Ef'b$, so we cannot solve it directly. It should be straightforward to solve a quadratic equation in b, Ef, but it is equivalent to minimization the following expression,

$$\min_{\{b,Ef\}} \left[g_T\left(b, Ef\right)' S^{-1} g_T\left(b, Ef\right) \right], \tag{1.5.8}$$

which can be solved using a numeric method. In this work, we use the MATLAB `fminsearch` function in order to minimize (1.5.8).

1.5.4 Efficient GMM, iterated to convergence.

Another possibility is *estimating* the spectral density matrix, in other words, use the optimal weighting matrix, instead of take prespecified weighting matrix on the second-stage estimators, as we advocate in pasts subsections.

The iterated GMM estimator using the optimal weighting matrix may present two important, and related, practical problems. First, if the covariance matrix for the iterated GMM estimator is poorly measured, hence the estimator will put too much weight on moments that spuriously appear to be measured precisely. Second, the iterated estimator may place too much weight on portfolios that are economically uninteresting, in the sense that they are composed of extreme short and long positions in some of the assets.[10] One straightforward solution to overcome this potential problems could be to include restrictions or exit flags which avoids unusual results.

The fact that S matrix changes with the model leads to a subtle trap. One model may improve a $J_T = g_T' S^{-1} g_T$ statistic because it blows up the estimates of S, rather than by making any progress in lowering the pricing errors g_T. As Cochrane [25] indicates, no one would formally use a comparison of J_T tests across models to compare them. This is one of the reasons because it is recommended to use a common weighting matrix W for comparing models like those presented before.

[10] See Chapman [20].

The optimal weighting matrix S depends on *population* moments, as well as on the parameters b. Work back trough the definitions,

$$S = \sum_{j=-\infty}^{\infty} E\left(u_t u'_{t-j}\right), \qquad u_t \equiv \left(m_t\left(b\right) x_t - p_{t-1}\right).$$

In order to estimate this matrix we estimate the population moments by their sample counterparts. Thus, use a first-stage b estimates and the data to construct sample versions of the definition of S. This procedure can produce a consistent estimate of the *true* spectral density matrix.

There are at least two alternatives to the second-stage procedure: *iteration* and *one step*. Hansen and Singleton [51] describe the above *two-step* procedure, and it has become popular for that reason. Two alternative procedures may perform better in practice, i.e., may result asymptotically equivalent estimates with better small-sample properties.

Estimating the spectral density matrix: Iterate

The second-stage estimate $\widehat{b}_2$ will not imply the same spectral density as the first stage. It might seem appropriate that the same estimate of b and of the spectral density be consistent, i.e., to find a fixed point of $\widehat{b} = \min_{\{b\}}\left[g_T\left(b\right)' S^{-1}\left(b_1\right) g_T\left(b\right)\right]$. One way to search for such a fixed point is to iterate: find b_2 from

$$\widehat{b}_2 = \min_{\{b\}}\left[g_T\left(b\right)' S^{-1}\left(b_1\right) g_T\left(b\right)\right],$$

where b_1 is a first-stage estimate, held fixed in the minimization over b_2.

Then use $\widehat{b}_2$ to find $S\left(\widehat{b}_2\right)$, find

$$\widehat{b}_3 = \min_{\{b\}}\left[g_T\left(b\right)' S^{-1}\left(\widehat{b}_2\right) g_T\left(b\right)\right],$$

and so on. There is no fixed-point theorem that such iterations will converge, but they often do. Ferson and Foerster [36] find that iteration gives better small-sample performance than two-stage GMM in Monte Carlo experiments. This procedure is also likely to produce estimates that do not depend on the initial weighting matrix.

Estimating the spectral density matrix: Pick b and S simultaneously

It is *not* true[11] that S must be held fixed as one searches for b. Instead, one can use a new $S(b)$ for each value of b. Explicitly, one can estimate b by

$$\min_{\{b\}} \left[g_T(b)' S^{-1}(b)\, g_T(b) \right]. \tag{1.5.9}$$

The estimates produced by this simultaneous search will be not numerically the same in a finite sample as the two-step or iterated estimates.

For our purposes, we will use this simultaneous procedure in the next two chapters in order to estimate the optimal weighting matrix.

Final note on prespecified vs. efficient GMM

The estimator from a prespecified weighting such as the identity matrix has three advantages over efficient GMM according to Ferson and Foerster [36].

First, given this prespecified weighting matrix, this estimates match the mean of the stochastic discount factor, and minimizes the sum of squared pricing errors on the Fama-French portfolios, given each portfolio equal weight. Thus, this choice of weighting matrix forces the model to explain the size effect and the value premium. Efficient GMM on the other hand minimizes the sum of squared pricing errors on

[11] See Cochrane [25].

weighted combinations of the portfolios, focusing on linear combinations of returns that have low variance, and often ignoring the value and size effect if they are hard to price, see Julliard and Parker [62]. In practice, efficient GMM prices rather unusual combination of portfolios, with extreme long and short positions.

Second, because GMM with a prespecified weighting matrix such as I tries to price the same portfolios as one vary S, measures of fit and specification tests are more comparable across different models (different S) than efficient GMM.

Third, GMM with a prespecified weighing matrix has a superior small sample properties.

At the end, we want good estimates of an approximate model, not efficient estimates of an exact model. Efficient GMM can do a poor job of that task because pays attention to well-measured linear combinations of moments, guided by S, not "large" or "economically interesting" moments. In other words, we want a GMM estimate of the approximate factor model that explains most of the variance of expected returns, not the one that minimizes the best measured, even if tiny, moments.

1.6 Conclusions

This econometric review is based on some of the most influential work in this field such as Cochrane [25], Ferson [34, 35], Ferson and Jagannathan [38], Campbell, Lo and MacKinlay [17], Shanken [93], Velu and Zhou [101], Ogaki [86], Jagannathan, Skoulakis and Wang [55, 54], and Marín and Rubio [81]. The organization of this introductory chapter will facilitate consecutive empirical exercises. Therefore it would not be uncommon that next chapters include continuous references to some specific sections and equations.

Nonetheless, next chapters will include a brief methodology section which help to follow the corresponding empirical results. Furthermore, adding a abbreviated methodology section contribute to have independent chapters.

It is quite difficult to include all possible ways to estimate and test asset pricing models. Here, we include a wide range of possibilities in which the same problem can be solved. However, next chapter include even more possibilities. For example, chapter 3 examine the weighted least squares, and chapter 4 study the GMM methodology in Beta models.

■

1.7 Appendix

1.7.1 Derivation of the χ^2 test.

We could test the hypothesis that all the α are equal to zero with a test that does not require a generalized inverse. Define, say by Choleski decomposition, a matrix C such that $CC' = \Sigma^{-1}$. Now, find the covariance matrix of $\sqrt{T}C'\widehat{\alpha}$:

$$cov\left(\sqrt{T}C'\widehat{\alpha}\right) = C'\left((CC')^{-1} - \beta\left(\beta' CC'\beta\right)^{-1}\beta'\right)C = I - \delta\left(\delta'\delta\right)^{-1}\delta',$$

where $\delta = C'\beta$.

In sum, $\widehat{\alpha}$ is asymptotically normal so $\sqrt{T}C'\widehat{\alpha}$ is asymptotically normal too, $cov\left(\sqrt{T}C'\widehat{\alpha}\right)$ is an independent matrix with rank $N-K$; therefore $T\widehat{\alpha}'CC'\widehat{\alpha} = T\widehat{\alpha}'\Sigma^{-1}\widehat{\alpha}$ is χ^2_{N-K}.

Chapter 2

Estimating and evaluating the Fama-French & Carhart models

§

We attempt to answer a classic empirical question in asset pricing: How do time-series regression, cross-sectional regression, Fama-MacBeth procedure, and GMM/SDF compare when applied to a test of linear factor models such as the Fama-French and Carhart models? We find that those econometric methods produce practically the same results for this classic exercise. In our sample of 871 monthly observations and 25 test portfolios, the pricing errors are jointly significant in Fama-French model, while in Carhart model are not.

Our findings support that efficient weighting matrix on GMM blow up standard errors rather than improve pricing errors, and with a prespecified weighting matrix

§An earlier version of this work was presented at Instituto Complutense de Análisis Económico (Madrid, May 2006); and at the IV Workshop in Quantitative Finance (València, June 2006), for helpful comments I thank José Emilio Farinós (Universitat de València). Martín Lozano gratefully acknowledges financial assistance from Consolidate Research Team 9/UPV-00038.321-15094, Econometrics and Statistics Department, Departamento de Fundamentos del Análisis Económico II at University of the Basque Country (Euskal Herriko Unibertsitatea), and Fundación BBVA.

we give up asymptotic efficiency but still obtaining consistent and more robust estimators.

Finally, according to our empirical results, we propose a slightly different specification that works somewhat better than Fama-French and Carhart models on explaining cross-sectional returns.

2.1 Introduction

The linear factor models, as the Fama-French [30, 31], and Carhart [18], are by far the most common in empirical asset pricing, and there is a large literature on econometric techniques to estimate and evaluate them. Each technique focuses on the same question: how to estimate parameters, how to calculate standard errors of the estimated parameters, how to calculate standard errors of the pricing errors, and how to test the model, usually with a test statistic of the form $\widehat{\alpha}' V^{-1} \widehat{\alpha}$. All the techniques come down to one of the two basic ideas: time-series regression or cross-sectional regression. Time-series regression turns out to be a limiting case of cross-sectional regression. The GMM, $p = E\left(mx\right)$ approach turns out to be almost identical to cross-sectional regressions. The formulas for parameter estimates, standard errors, and test statistics are all strikingly similar.

The GMM/discount factor, time-series, and cross-sectional regression procedures and distribution theory are similar but not identical. Cross-sectional regressions on betas are not the same thing as cross-sectional regressions on second moments. Cross-sectional regressions weighted by the residual covariance matrix are not the same thing as cross-sectional regressions weighted by the spectral density matrix. GLS cross-sectional regressions and second-stage GMM have a theoretical efficiency advantage over OLS cross-sectional regressions and first-stage GMM, but how important is this advantage, and is it outweighed by worse finite-sample performance?

The GMM/stochastic discount factor approach is still a *new* procedure (see Cochrane [25]). Thus, it is important to verify that it produces similar results and well-behaved test statistics in the setups of the classic regression tests[1]. To address these questions, we apply the various methods to a classic empirical question. *How do time-series regression, cross-sectional regression, Fama-MacBeth procedure, and GMM/SDF compare when applied to a test of the Fama-French and Carhart models?* We find that three methods produce practically the same results for this classic exercise. They produce almost exactly the same estimates, standard errors, t-statistics, and χ^2 statistics that the pricing errors are jointly zero.

The balance of this work is organized as follows. Section 2.2 describes the models, the test portfolios and the factors. Sections 2.3, 2.4 and 2.5 presents the empirical results. Each of these sections sets the basic econometric theory and their corresponding empirical results for Fama-French and Carhart models. Section 2.6 presents conclusions.

2.2 The models and data

2.2.1 Fama-French 3 factor model

Explaining cross-sectional differences in asset expected returns is one of the great challenges of modern finance. Asset pricing theory recognized at least since Merton

[1]Before the advent of GMM, the primary econometric tool in the asset-pricing area in finance was the maximum likelihood method originally proposed by Gibbons [40] and further explored by Shanken [91], which is often implemented using linear or nonlinear regression methods.

[82, 83], the theoretical possibility, indeed probability, that we should need factors, state variables, or sources of priced risk beyond movements in the market portfolio in order to explain why some average returns are higher than others. The Fama-French model is one of the most popular multifactor model

that now dominate empirical research. In 1993, Fama and French presents the model; and in 1996 they give an excellent summary, and also show how the three-factor model performs in evaluating expected return puzzles beyond the size[2] and value effects that motivated it.

Value stocks have market values that are small relative to the accountant's book value. That is because book values essentially track past investment expenditures. Book value is a better divisor for individual-firm price than are dividends or earnings, which can be negative. This category of stocks has given large average returns. *Growth* stocks are the opposite of value and have had low average returns. Since low prices relative to dividends, earnings, or book value forecast *times* when the market return will be high, it is natural to suppose that these same signals forecast categories of stocks that will do well; the *value effect* is the cross-sectional analogy to price-ratio predictability in the time series.

High average returns are consistent with the CAPM, if these categories of stocks have high sensitivities to the market, high betas. However, small and especially value stocks seem to have abnormally high returns even after accounting for the market beta. Conversely, *growth* stocks seem to do systematically worse than their CAPM betas suggest.

Figure 2.1 shows the value-size puzzle for the period 1932-2002, here, stocks are

[2]Banz [6] first documented the size effect by showing that small firms had higher risk-adjusted returns than large firms for the 1936-1977 period.

sorted into portfolios based on size and book-to-market ratio. As we can see, the highest portfolios have almost three times the average excess return of the lowest portfolios, and this variation has nothing at all to do with market betas.

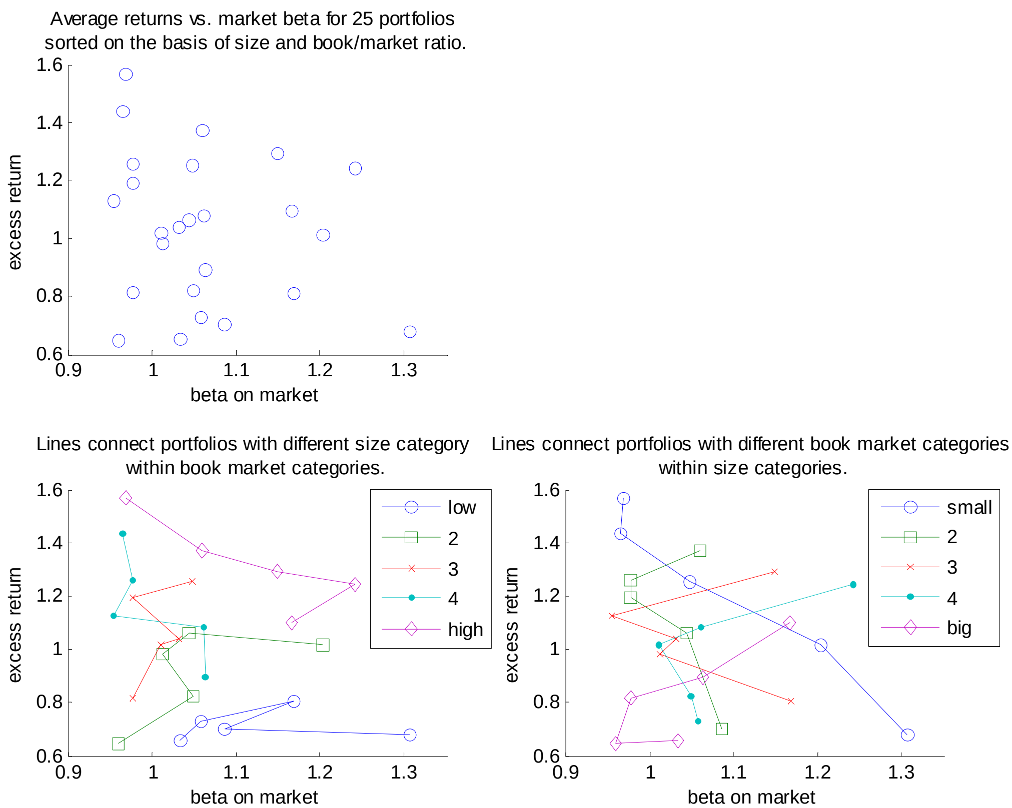

Figure 2.1: Representation of the size and value anomalies.

In the two bottom graphs (see Figure 2.1 again), we dig a little deeper to diagnose the problem, by connecting portfolios that have different size within the same book-to-market category, and different book-to-market within size category. We can see that portfolios with high book-to-market ratios have bigger returns; and small portfolios have also bigger returns[3]. Because of this value effect, the CAPM faces difficulties when confronted with these portfolios.[4]

To explain these patterns in average returns, Fama and French advocate a multifactor model with the market return (*rmrf*), the return of the small less big stocks (*smb*) and the return of high book-to-market minus low book-to-market stocks (*hml*) as three factors. They show that variation in average returns of the 25 size and book-to-market portfolios can be explained by varying loadings (betas) on the latter two factors. All their portfolios have betas close to one on the market portfolio, which is consistent with recent studies such as Grauer and Janmaat [42, Table 2]. Thus, market beta explains the average return difference between stocks and bonds, but not across categories of stocks.

2.2.2 Carhart 4 factor model

Although the benefits of the three-factor model are acknowledged, the Fama-French model has been subject to further improvement.

At every moment there is a most-studied anomaly, and momentum is that anomaly[5]. It is not explained by the Fama-French three factor model. The past losers have low prices and tend to move with value stocks. Hence the Fama-French model predicts

[3]Since the size effect disappeared in 1980, it is likely that almost the whole story can be told with book-to-market effects alone.

[4]The CAPM of Sharpe [95] and Lintner [75] faces some severe empirical difficulties. Specifically, Basu [7], Banz [6], Reinganum [88], and Jegadeesh and Titman [59] among others show that the CAPM fails to explain the returns of several equity investment strategies based on accounting data or past returns. Some authors explain the failures of the CAPM with nonrisk-based explanations such as biases in the empirical methodology, see for example Lo and MacKinlay [76], MacKinlay [79] and Kothari, Shanken and Sloan [69], or investor irrationality. While others take a rational view and explain differences in return with differences in risk (see Fama and French [31], Cochrane [24] and Lettau and Ludvigson [71]).

[5]See Jegadeesh and Titman [59] for a complete exposition of the momentum strategy.

they should have *high* average returns, not *low* average returns. Momentum stocks move together as do value and small stocks, so a momentum factor works to explain momentum portfolio returns. To put in other way, since a string of good returns gives a high price, it is not surprising that individual stocks that do well for a long time (and reach a high price) subsequently do poorly, and stocks that do poorly for a long time (and reach a low price, market value, or market to book ratio) subsequently do well.

A momentum factor is more palatable as a performance attribution factor. If we run fund returns on factors including momentum, we may be able to say that a fund did well by following a mechanical momentum strategy rather than by stock-picking ability, leaving aside why a momentum strategy should work. Carhart [18] uses it in this way; we will call Carhart model to a Fama-French three factor model plus momentum. Momentum (*umd* fourth factor) is really a new way of looking at an old phenomenon, the small apparent predictability of monthly individual stock returns.

We focus on the so-called anomalies and factors, such as the size effect (Banz [6]), the value premium (Fama and French), and the momentum anomaly (Jegadeesh and Titman [59]) because they provide the most empirically successful multifactor models and have attracted much industry as well as academic attention.

2.2.3 Data description

We thank Kenneth French for making the Fama-French portfolios and factors data available on his Web page:
http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/index.html

The complete database are 852 monthly observations (71 years) from January

1932 to December 2002 for four factors and 25 test portfolios. Most estimations and tests were done with MATLAB, version 7.

Test portfolios

The question of which test assets or test portfolios to use is not trivial. In fact, studies such as Roll [89] argues that forming portfolios can potentially impair asset pricing tests by reducing cross-sectional variation in some characteristics of the test assets[6]. In this chapter we choose a set of test portfolios with more return dispersion than the traditionally used; however, there are other portfolio formations (discussed in next chapters) which allow even more cross-sectional variation.[7]

We use the monthly returns on the 25 Fama and French [29] portfolios and construct excess returns as these return on a three-month Treasury bill. We study these returns because the Fama-French portfolios have a large dispersion in average returns that is relatively stable in subsamples, see Julliard and Parker [62], and because they have been used extensively to evaluate asset pricing models. These portfolios are designed to focus on two features of average returns: the *size effect*, firms with small market value have, on average, higher returns; and the *value premium*, firms with high book values relative to market equity have, on average, higher returns.

The 25 Fama-French portfolios are the intersections of five portfolios formed on size (market equity) and five portfolios formed on the ratio of book equity to market

[6]In Brennan, Chordia and Subrahmanyam [13] perform cross-sectional regressions on individual assets in order to increase the cross-sectional variation. The cost of this approach is discussed in Black, Jensen and Scholes [9], basically it increases the error in variables bias.

[7]This extension in the following chapters is important since according to Lewellen, Nagel, and Shanken [72] it is difficult to differentiate models that have been developed to explain the cross-sectional returns of the 25 size and book-to-market portfolios using traditional methods, because these models tend to have small pricing errors for the test assets by construction.

Table 2.1: Mean returns and standard deviation for the 25 size-value portfolios

Mean returns					
	low (growth stocks)	**2**	**3**	**4**	**high** (value stocks)
small	0.681	1.015	**1.253**	**1.439**	**1.567**
2	0.702	1.063	1.193	**1.258**	**1.373**
3	0.810	0.985	1.037	1.129	**1.292**
4	0.732	0.823	1.019	1.082	1.243
big	**0.655**	**0.646**	0.818	0.893	1.097
Standard deviation					
small	12.468	10.690	9.010	8.679	9.696
2	8.093	7.802	7.459	7.489	8.611
3	7.732	6.539	6.735	6.691	8.470
4	6.142	6.268	6.292	6.997	8.973
big	5.374	5.112	5.601	6.696	8.359

Data range: January 1932-December 2002.
Book-to-market goes from low to high; size goes from small to big.

equity. We denote a portfolio by the rank of its market equity and then the rank of its book-to-market ratio so that portfolio 15 is the smallest quintile of stocks by market equity and the largest quintile of stocks by book-to-market. We will also call portfolio *small-high*, small for size and high for book-to-market ratio.

Here are the 25 portfolios' mean returns and standard deviation. We organize this type of tables into 5×5 blocks, with small to big size on the vertical axis and low to high book-to-market on the horizontal axis. It will be easy to see the size and value effect in the following table, as we did before in Figure 2.1.

As shown in Figure 2.2, which collect the information of the upper panel of Table 2.1, small firms have the biggest expected returns, they are usually riskier and investors ask them for an extra premium. Firms with high book-to-market ratio have an accountant valuation that is not according to their market value. Those firms

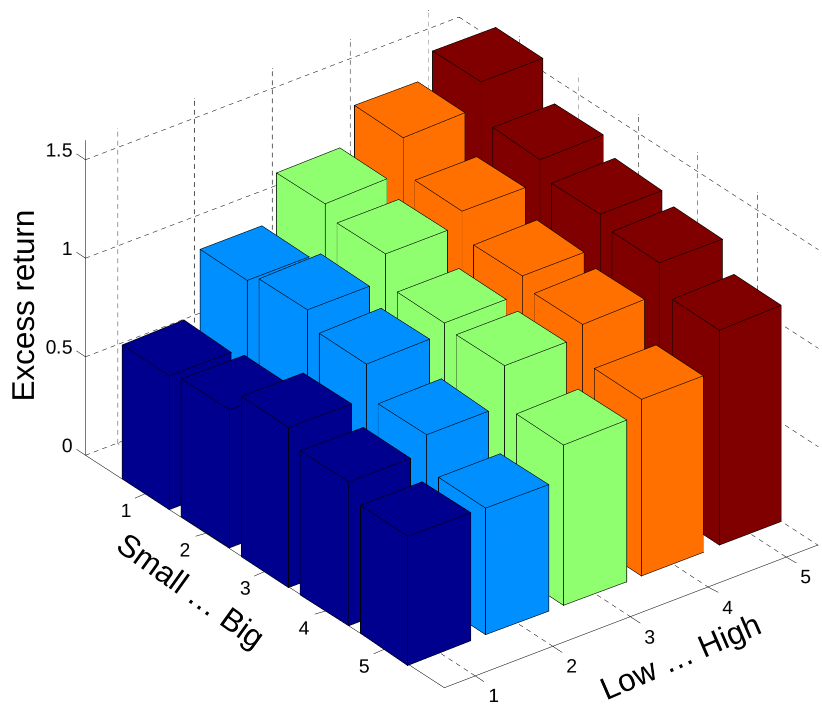

Figure 2.2: Average returns for 25 size-value portfolios.

are especially affected by economic crisis and investors will demand them an extra premium. These two effects together are located at the top-right cluster in Table 2.1 (in bold). Conversely, big-growth stocks have the lowest mean returns (bottom-left cluster).

We must keep in mind the numbers here, the returns spread is from 0.655 (portfolio 51) to 1.567 (portfolio 15), that is 1 per cent month spread in average returns, which is about 12 per cent per year. This annual return is relatively large, so there is a return spread in these 25 portfolios to explain. Other portfolio formations have even bigger monthly return spread like the size-momentum portfolios, which has 1.8 per cent for the same time period. In the appendix, Figure 2.13 we show a similar plot to Figure 2.2 for the 25 size-momentum portfolios.

Factors: *rmrf*, *smb*, *hml* & *umd*

The construction of the Fama-French factors is described in Fama and French [31]. In particular, the factors used in Fama-French model are market, size and value, denoted as *rmrf*, *smb*, *hml* respectively. On the other hand, in Carhart model are market, size, value and momentum: *rmrf*, *smb*, *hml*, *umd*.

The excess return on the market (*rmrf*), is the value-weight return on all NYSE, AMEX, and NASDAQ stocks (from the Center for Research in Security Prices, CRSP) minus the one-month Treasury bill rate (from Ibbotson Associates).

The *smb* monthly factor is computed as the average return for the smallest 30% of stocks minus the average return of the largest 30% of stocks in that month. A positive *smb* in a month indicates that small cap stocks out-performed large cap stocks in that month. A negative *smb* in a given month indicates the large caps outperformed. In this way, the *smb* factor will try to capture the size effect we described on Table 2.1.

Constructed in a fashion similar to that of *smb*, *hml* is computed as the average return for the 50% of stocks with the highest book-to-market ratio minus the average return of the 50% of stocks with the lowest book-to-market ratio each month. A positive *hml* in a month indicates that value stocks outperformed growth stocks in that month. A negative *hml* in a given month indicates the growth stocks outperformed. Hence, the *hml* factor will try to explain the value effect on Table 2.1.

Table 2.2: Market, size, value and momentum factors descriptive statistics

	rmrf	***smb***	***hml***	***umd***
Mean	0.7096	0.2780	0.4637	0.6934
Standard deviation	5.3160	3.4043	3.6310	4.6530
$\frac{\text{Mean}}{\text{Standard deviation}}$	0.133 48	8.1661×10^{-2}	0.127 71	0.149 02
rmrf	1	0.3496	0.1896	−0.3050
smb	0.3496	1	0.0924	−0.1644
hml	0.1896	0.0924	1	−0.3857
umd	−0.3050	−0.1644	−0.3857	1

Data range: January 1932-December 2002.

Momentum factor *umd* (up minus down) is the average return on the two high prior return portfolios minus the average return on the two low prior return portfolios.

The size, value and momentum effects are said to be anomalies in the sense that these cross-sectional patterns of portfolios' returns are not explained by beta risk as the covariance between market return and portfolio return (divided by market variance).

All means are positives. Thus, small caps stocks out-performed large cap-stocks in average, this is consistent with the pattern showed at Table 2.1. Moreover, value stocks out-performed growth stocks on average, this is also consistent with the returns of our 25 test portfolios. The positive value of momentum is like saying that winners continue to win, and the losers continue to loose.

Note that *smb* and *hml* factors are close to be orthogonal, and that is good since we are interested on explaining the size and value effect. Fama and French also find that *hml* and *smb* do not explain momentum.

2.3 Empirical results: Beta representation

2.3.1 Time-series

We estimate time-series regression of (1.2.1) by OLS. Standard errors are calculated in two ways, the classics OLS standard errors (1.2.6); and the GMM standard errors (1.2.5) using 0 and 12 lags. Also, both asymptotic χ^2 test (1.2.7) and GRS F test (1.2.10) are calculated in this section.

Jensen's alphas

The alphas presented in Table 2.3 are not *all* that small, at least for the Fama-French model. Even though, must of them (approximately 68%) are less than 0.1% per month, that is interesting because the phenomenon was a 1% per month spread[8] and the residuals or alphas are $\frac{1}{10}$ that size. Still, other alphas are more than two standard errors from zero. For instance, take a look of the *small-low* or 11 portfolio with alpha equal to -0.864 which has a wide standard error (as shown in appendix, Table 2.26) which goes from 0.177 when considering GMM standard errors to 0.253 when considering OLS standard errors.

Note that not only the *small-low* portfolio has a large alpha, it is statistically large too, according to the t-statistics shown in the appendix, Table 2.31. The largest of 25 t-statistics should not be much over 2, so this will drive to a statistical rejection of the model. Some of the other alphas are individually significant as well, most clustered in the upper left corner and two more in the lower right. This result could drive to reject the model, time-series estimation cannot explain very well the portfolio returns,

[8] Remember the previous descriptive analysis on test portfolios (Table 2.1).

Table 2.3: Jensen alpha estimators $\widehat{\alpha}$ for the 25 size-value portfolios

Fama-French model					
	low	**2**	**3**	**4**	**high**
small	**−0.864**	**−0.419**	−0.042	0.124	0.056
2	**−0.233**	−0.034	0.073	0.083	−0.018
3	**−0.148**	0.084	0.022	0.082	−0.087
4	0.074	−0.053	0.093	−0.013	**−0.211**
big	0.077	0.021	0.038	**−0.136**	−0.188
Carhart model					
small	**−0.668**	−0.262	0.028	0.1570	0.153
2	**−0.204**	0.051	0.085	0.0693	0.023
3	−0.057	0.064	0.044	0.0672	0.011
4	0.051	−0.019	0.111	0.0742	−0.099
big	0.104	0.034	0.064	−0.052	−0.099

Alphas (pricing errors) less than −0.1 and statistically different from zero in bold. Estimates correspond to equation 1.2.1

especially those of extreme characteristics.

On the other hand, the Carhart model seems to do better. The spread and the magnitude of the alphas in Table 2.3 are smaller, as well as the number of statistically significance estimates (see Table 2.31). In general terms, this model seems to price better the portfolios that Fama-French had difficulties to price such as 11 and 12. Such difficulty is represented in the two bold inputs on the lower panel of Table 2.3.

Remember we present $\widehat{\alpha}_i$ standard errors and t-statistics in the appendix Tables 2.26 and 2.31, there are three version of them: OLS following equation (1.2.6) and GMM following equation (1.2.5) taking 0 and 12 Newey-West lags in the spectral density matrix.[9] The various standard errors are broadly the same, note that GMM standard errors, which accounts for bias, is not always larger than OLS.

[9] See Newey and West [84].

Table 2.4: OLS betas in the Fama-French model

$\widehat{b_i}$ **beta on *rmrf***					
	low	**2**	**3**	**4**	**high**
small	1.307	1.203	1.048	0.966	0.970
2	1.087	1.044	0.979	0.979	1.059
3	1.168	1.013	1.032	0.956	1.149
4	1.059	1.050	1.012	1.062	1.242
big	1.034	0.960	0.977	1.063	1.167
$\widehat{s_i}$ **beta on *smb***					
small	1.364	1.522	1.218	1.234	1.398
2	1.049	0.978	0.866	0.806	0.898
3	0.784	0.506	0.400	0.455	0.520
4	0.274	0.223	0.211	0.188	0.351
big	−0.159	−0.197	−0.211	−0.160	−0.117
$\widehat{h_i}$ **beta on *hml***					
small	0.514	0.338	0.460	0.618	0.935
2	−0.275	0.181	0.399	0.553	0.840
3	−0.192	0.089	0.369	0.524	0.905
4	−0.365	0.150	0.323	0.623	1.023
big	−0.240	−0.002	0.314	0.689	1.057

Most of these betas are statistically different from zero.

We already show the portfolios with large alphas, now it is time to look at their t-statistics. Again, we present 3 different versions since there are 3 different standard errors calculated. It is noteworthy that 11 and 12 portfolios present a significant alpha, even though the significance level is smaller when estimating Carhart model. It is also interesting that alphas on portfolios 45 and 54 become to zero from Fama-French to Carhart. We will formally test this sense of improvement on Table 2.7.

Betas

Now let's move to the betas estimates following equation (1.2.1). The estimates for the Fama-French model are in Table 2.4, and in Table 2.5 for the Carhart model.

Results for the Fama-French model in Table 2.4 show that the market betas $\widehat{b}_i$ are all about one for most portfolios. The *smb* betas $\widehat{s}_i$ vary up and down, while the *hml* betas $\widehat{h}_i$ vary side to side. Therefore, this results imply that the *characteristic* size and book-to-market are in fact associated with *behavior*, covariance with *smb* and *hml* portfolio. This is also one of the reason of getting high R^2 values, as we show in Table 2.6.

We present the standard errors values for Fama-French betas in the appendix, Tables 2.27, 2.28 and 2.29. In the case of $se(\widehat{b}_i)$, OLS and GMM formulas can be quite different. For example, GMM standard errors can be two times as large as plain OLS standard errors. In every case we reject the null hypothesis that the estimate $\widehat{b}_i$ is equal to zero using the t-statistic. In the case of $se\left(\widehat{s}_i\right)$, again, GMM standard errors can be three times large as plain OLS standard errors, even for simple estimation like OLS time-series regressions. The estimate $\widehat{s}_i$ is also statistically different from zero in every case[10]. And, for completeness, the *hml* standard errors $se(\widehat{h}_i)$ are shown in Table 2.29, GMM are often 2-3 times as big. In this last case, the estimator is statistically equal to zero in only three of the 25 portfolios when considering autocorrelation and heteroskedasticity. Hence, must of the estimators are statistically different from zero.

In all cases, there is not much difference between zero and 12 Newey-West lags, as there is not a lot of autocorrelation in stock returns. Clearly, heteroskedasticity is the problem. In this hugely long dataset with high R^2 the betas are all well estimated

[10]In fact, there is an exception: the estimator for the *big-high* portfolio is the only one statistically equal to zero when using GMM standard errors.

Table 2.5: OLS betas in the Carhart model

	$\widehat{b}_i$					$\widehat{s}_i$				
	Low	**2**	**3**	**4**	**High**	**Low**	**2**	**3**	**4**	**High**
Small	1.2	1.1	1.0	0.9	0.9	1.3	1.5	1.2	1.2	1.3
2	1.0	1.0	0.9	0.9	1.0	1.0	0.9	0.8	0.8	0.8
3	1.1	1.0	1.0	0.9	1.1	0.7	0.5	0.3	0.4	0.5
4	1.0	1.0	1.0	1.0	1.2	0.2	0.2	0.2	0.1	0.3
Big	1.0	0.9	0.9	1.0	1.1	−0.1	−0.1	−0.2	−0.1	**−0.1**
	$\widehat{h}_i$					$\widehat{u}_i$				
Small	**0.4**	0.2	0.4	0.6	0.8	0.1	0.1	0.0	**0.0**	0.0
2	−0.2	0.1	0.3	0.5	0.8	**0.0**	0.0	**0.0**	**0.0**	0.0
3	−0.2	0.0	0.3	0.5	0.8	0.0	**0.0**	**0.0**	**0.0**	0.0
4	−0.3	0.1	0.3	0.5	0.9	**0.0**	0.0	**0.0**	0.0	0.0
Big	−0.2	−0.0	0.3	0.6	1.0	0.0	**0.0**	**0.0**	0.0	0.0

Values statistically equal to zero in bold.

(standard errors are low) but that is often not the case.

The Carhart model shows $\widehat{b}_i$ values (see Table 2.5) even closer than one as we compare them with Fama-French model. The fourth factor is statistically equal to zero in 11 out of 25 portfolios (see bold values in Table 2.5) according to a t-statistic, but it is interesting that it is actually different from zero for those portfolios that Fama-French cannot price well. To put it in another way, the momentum factor seems to do a good job on pricing the portfolios located on top-left and bottom-right clusters. For that reason, we should expect that tests will show smaller pricing errors compared with the Fama-French model.

Table 2.6: Goodness of fit, time-series estimation

	Fama-French					**Carhart**				
	R^2									
	low	**2**	**3**	**4**	**high**	**low**	**2**	**3**	**4**	**high**
small	**0.65**	0.84	0.88	0.93	0.93	**0.66**	0.84	0.88	0.93	0.93
2	0.89	0.93	0.93	0.95	0.95	0.89	0.93	0.93	0.95	0.95
3	0.93	0.92	0.92	0.92	0.94	0.93	0.92	0.92	0.92	0.94
4	0.92	0.92	0.91	0.92	0.92	0.92	0.92	0.91	0.92	0.92
big	0.95	0.92	0.90	0.92	0.86	0.95	0.92	0.90	0.92	0.86
	adjusted R^2									
small	**0.65**	0.84	0.88	0.93	0.93	**0.66**	0.84	0.88	0.93	0.93
2	0.89	0.93	0.93	0.95	0.95	0.89	0.93	0.93	0.95	0.95
3	0.93	0.92	0.92	0.92	0.94	0.93	0.92	0.92	0.92	0.94
4	0.92	0.92	0.91	0.92	0.92	0.92	0.92	0.91	0.92	0.92
big	0.95	0.92	0.90	0.92	0.86	0.95	0.92	0.90	0.92	0.86

In bold the worst priced portfolio small-low.

Goodness of fit

Both, R^2 and adjusted R^2 are used in order to analyze the goodness of fit of equation (1.2.1).

These values are in general high as we already expected from our previous analysis. In fact, when running a CAPM on the same period, we obtained a mean R^2 of 0.7, while the mean of Fama-French model is 0.9 and 0.91 for Carhart.

These results are also consistent with previous work in the subject. According to Cochrane [25], most factor models have fairly high R^2, so $\sigma^2(\varepsilon) < \sigma^2(f)$. Common CAPM values of $R^2 = 1 - \frac{\sigma^2(\varepsilon)}{\sigma^2(f)}$ for large portfolios are about 0.6 - 0.7; and multifactor models have R^2 often over 0.9. Typical numbers of assets $N = 10$ to 50 make the first term vanish compared to the second term.

Other authors such as Petkova [87] show that using the commonly employed 25 size and book-to-market ranked portfolios as test assets, there is not much statistical evidence to establish that other models like the five-factor intertemporal capital asset pricing model (ICAPM) outperforms even the simple unconditional CAPM in terms of cross-sectional R^2. However, the advantage of the Fama and French three-factor model over the CAPM is statistically significant for this metric.

Nevertheless, we know that R^2 is not the only way to measure the goodness of fit[11], these results are intended to give a rough idea rather than exhaustive test suite. In the next subsection we will use the formal statistical measure for evaluating the time-series regressions detailed before in section 1.2.3.

Tests statistic

There is one and only one implication for the data: *all regression intercepts* $\widehat{\alpha}_i$ *should be zero*[12].

First column refers to (1.2.7) test, second and third to (1.2.9) and the last one to (1.2.10). The critical values for both models are the same because degrees of freedom do not depend on the number of factors. On the other hand, in the case of the GRS test, the degrees of freedom does depend on the number of factors but at least for the third decimal place both critical values are the same. The null hypothesis is that alphas are *jointly* equal to zero. As we can see, the tests all dramatically reject. We also expected this from the t-statistic on individual alphas above, specially for the Fama-French model. Joint tests can only be worse. As the GMM standard errors of alpha are a bit smaller, the GMM χ^2 values are a little bit larger, leading to more

[11] See Kan, Robotti and Shanken [65] for a discussion about use of R^2 for the judgement about the empirical success of a beta pricing models.

[12] In time-series regression, intercepts are equal to the pricing errors.

Table 2.7: Tests statistics, time-series estimation

	χ^2 (eq. 1.2.7 and 1.2.9)			F (eq. 1.2.10)
	Normality	**0 lags**	**12 lags**	**GRS**
Fama-French model	74.9	81.8	105.1	2.898
Carhart model	54.0	56.7	82.4	2.089
5% critical value	37.6525			1.519
1% critical value	44.3141			1.795

H_0: Pricing errors are equal to zero.

dramatic rejections.

Table 2.7 support the idea that adding momentum is not a bad idea, since the Carhart's statistics are smaller than Fama-French's.

For completeness, we present estimates and standard errors of the factor risk premia λ in Table 2.8. For a time-series regression, these are just the means of the factors as we indicated on (1.2.2). As we can see, there is so little autocorrelation that the 12 lag and i.i.d. standard errors are about the same. All of the risk premia are statistically significant. The units are percent per month, so they are economically large as well.

It is clear that all factor risk premia are statistically significant in Table 2.8. However there are differences among factors, the *smb* factor has the biggest p-value while momentum has the lowest. In other words, momentum seems to be the most significant in statistical terms, even more than the market.

It is appropriate to show a plot of actual versus predicted means returns for the Fama-French and Carhart model in Figures 2.3 and 2.4 respectively. We put the actual on the y axis so that this is like $E\left(R^e\right)$ versus β that we would plot for a single factor model as CAPM. It is clear to see the decent fit, as well as the problems with

Table 2.8: Market, size, value and momentum factors' risk premia

	rmrf	***smb***	***hml***	***umd***
Estimate equation (1.2.2)	**0.709**	**0.278**	**0.463**	**0.693**
Standard error (normality)	0.182	0.116	0.124	0.159
Standard error (GMM, 12 lags)	0.184	0.122	0.140	0.144
t-statistic (normality)	3. 898	2. 384	3. 730	4. 352
t-statistic (GMM, 12 lags)	3. 850	2. 271	3. 293	4. 808

Values statistically different from zero in bold.

the 11 and 12 portfolios[13]. Naturally, the distance between each portfolio and the 45 degree line represent the Jensen's alpha (see [60]) that we shown before in Table 2.3.

Figure 2.3 serves to clarify even further our previous analysis, the 11 portfolio is hard to price for the Fama-French model and the goodness of fit is not bad at all. Therefore, we reject the hypothesis that pricing errors are zero. One way to summarize these kind of plots is collecting their root mean square errors $\sqrt{\frac{1}{N}\alpha'\alpha}$ (rmse hereafter), as well as the mean absolute value (mav hereafter). The results are 0.2171 and 0.1310 respectively.

Now, let's take a look of the same representation for Carhart model in Figure 2.4. Probably one of the most notorious differences among Figures 2.3 and 2.4 is portfolio 11 valuation, the Fama-French predicted value is close to 1.6 while in Carhart model is close to 1.4. This is good since this portfolio seems to be the hardest to price, and a better prediction of it represents an improvement of the model. Actually, we have a better fit with respect to Fama-French since rmse is equal to 0.1669.

The goodness of fit is related to the standard deviation of the expected returns' portfolios. In other words, portfolios with higher expected return dispersion will be

[13]The labels are size, then book-to-market, so 11 = *small-low*.

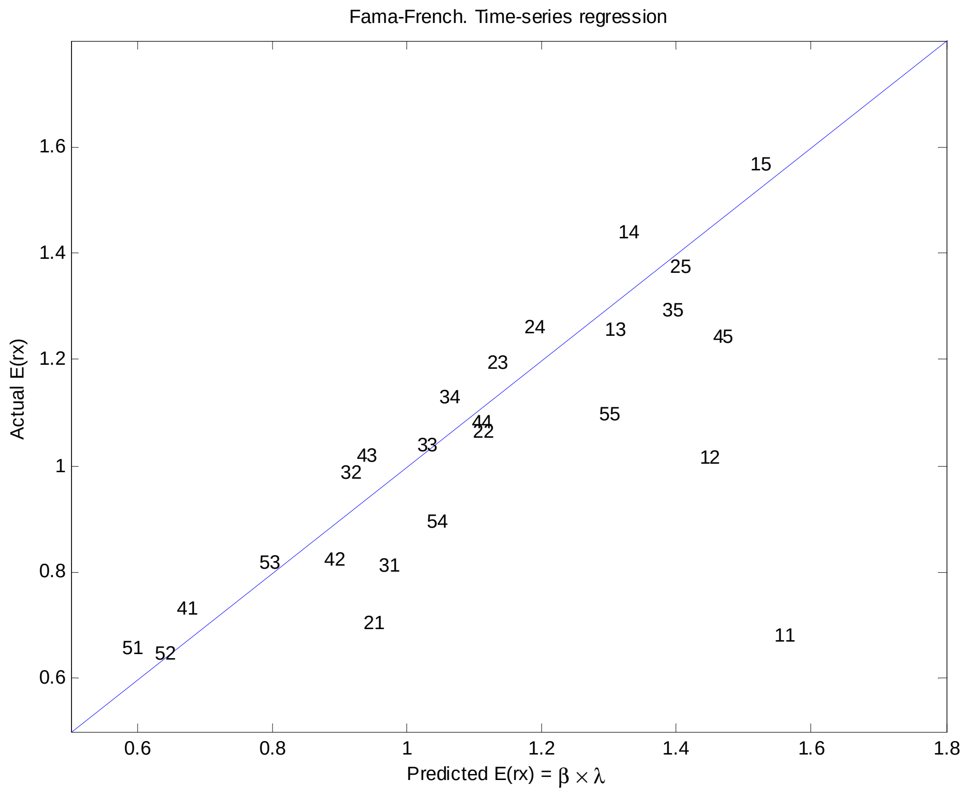

Figure 2.3: Time series Fama-French model actual versus predicted returns

harder to explain independently of the model. In the next chapter we will show how portfolios formation with low dispersion like the 10 size portfolios are associated with high goodness of fit even when estimating the CAPM model.

In the next section, we will perform a similar analysis based on cross-sectional estimation and evaluation.

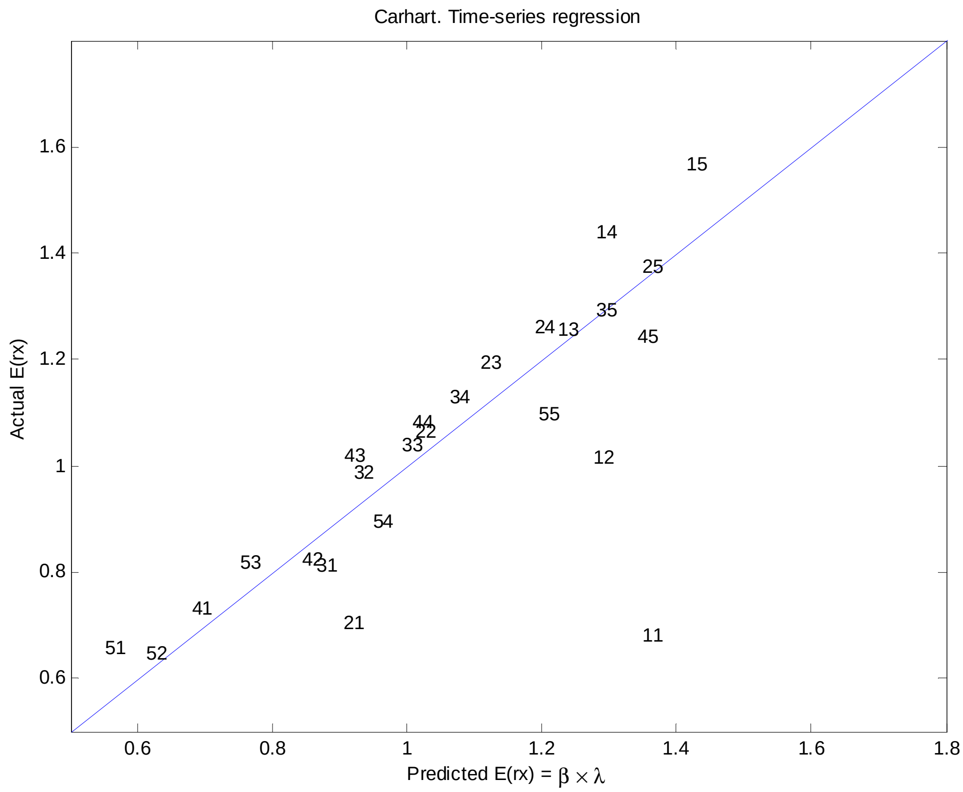

Figure 2.4: Time series Carhart model actual versus predicted returns

2.3.2 Cross-sectional

Recall that betas and intercepts are now the right hand variables taken from the time-series regression, so we do not have to estimate them again in this section 2.3.2. The difference of course is the factor risk premia in (1.3.1). For a detailed exposition of the econometrics used hereafter, see section 1.3.

We present the results for the cross-sectional estimates in Tables 2.9 and 2.10 for the Fama-French and Carhart models respectively. In general, the lambdas are in fact sensibles to the model estimated. Note most of the lambdas are close to the factor

Table 2.9: Cross-sectional estimation, Fama-French model

Equation	Description	const.	*rmrf*	*smb*	*hml*
Ordinary Least Squares (OLS)					
(1.3.2)	λ without a constant		**0.6964**	0.1715	**0.4544**
(1.2.2)	factor mean		**0.7096**	**0.2780**	**0.4637**
(1.3.2)	λ with a constant	**1.96**	**−1.1713**	0.2065	**0.4904**
(1.3.4)	se, i.i.d., no Shanken		0.1842	0.1228	0.1278
(1.3.12)	se, i.i.d., Shanken		0.1843	0.1229	0.1279
(1.2.5)	se, GMM, 0 lags		0.1833	0.1175	0.1224
(1.2.5)	se, GMM, 12 lags		0.1806	0.1166	0.1441
Generalized Least Squares (GLS)					
(1.3.7)	λ without a constant		**0.7364**	**0.2698**	**0.4499**
(1.2.2)	factor mean		**0.7096**	**0.2780**	**0.4637**
(1.3.7)	λ with a constant	**1.583**	**−0.8071**	**0.2804**	**0.4557**
(1.3.9)	se, i.i.d., no Shanken		0.1827	0.1179	0.1264
(1.3.13)	se, i.i.d., Shanken		0.1827	0.1179	0.1264
(1.2.5)	se, GMM, 0 lags		0.1808	0.1161	0.1237
(1.2.5)	se, GMM, 12 lags		0.1816	0.1223	0.1447

Values statistically different from zero in bold.

means, but they are not exactly the same as factor means. Also, as expected, OLS cross-sectional regressions (upper panels in Tables 2.9 and 2.10) allows some *alpha* on the factors in order to better fit the other portfolios.

Ordinary and generalized least squares

Allowing a constant in the cross-sectional regressions changes *smb* and *hml* λ a little in both models, but has a dramatic effect on the market premium. The market is basically a constant here, as there is very little spread in market betas[14]. Thus, the market and the constant are close to be collinear and the regression has trouble picking

[14]This is also shown in Tables 2.4 and 2.5, where $\widehat{b}_i$ is close to 1 for every portfolio.

between them. As we will see later, the cross-sectional regressions with constant do much better in the actual versus predicted plots, but as we already know, they are not believable. In fact, adding a constant leads to a misleading economically interpretation such as a negative market premium.

One interesting result of this section is that the standard errors with and without the Shanken correction are almost numerically identical except for the momentum factor in the last column of Table 2.5. Recall the only difference between equations 1.3.4 and 1.3.12 is the multiplicative $(1+\lambda'\Sigma_f^{-1}\lambda)$ correction to the first term, and that is a small correction. In this case, the effects of autocorrelation and heteroskedasticity are small, so GMM versus Shanken turns out to make little difference at least on the market, size and value factors. It is also interesting to see that in OLS cross-sectional regressions, *smb* (when estimated) is not statistically significant no matter which method do we pick in order to calculate standard errors.

Note that even in GLS estimation in both Tables 2.9 and 2.10, the constant and market premium are so collinear that adding a constant makes a big difference. Also, the lambdas are closer to the factor means. If the test portfolios spanned the factors, that is, if we could recover *rmrf*, *smb*, *hml* and/or *umd* as combinations of the test portfolios, then the lambdas would come out exactly equal to the factor means, and the cross sectional and time-series regressions would yield the same results. Here, we can *almost* get the factors back from the test portfolios but not quite, so it is only very close[15].

Shanken correction makes no difference out to four decimal places. In fact, $\frac{\Sigma_f}{T}$

[15] Actually, if we could get the test portfolios back exactly, the GLS regression would have failed as Σ would have been singular. For a detailed discussion about the use of spanned and unspanned factors at estimating risk premiums see Hou and Kimmel [52].

is the dominant term, so a small multiplicative $\lambda'\Sigma_f^{-1}\lambda$ term makes little difference. Shanken versus GMM also makes little difference in this not very autocorrelated or heteroskedastic dataset.

Momentum factor, as well as the market factor, are basically a constant; so if we add a constant in the model, the regression has trouble picking between them and constant becomes strongly significant. To see this, refer to the Table 2.5, market beta (upper left panel) is close to 1 and momentum beta (lower right panel) is close to zero for every portfolio. Table 2.5 show how standard errors on momentum (last column) are very sensible to changes in the estimation procedure. Here, contrary to the rest of the factors, effects of autocorrelation and heteroskedasticity are not as small at all.

An interesting read on both Tables 2.9 and 2.10 is to verify that GLS standard errors are in fact smaller than OLS without exception.

Actual versus predicted plots

Here we present plots of actual returns versus model predictions, i.e., $\beta\lambda$. Figure 2.5 refer to the Fama-French model and Figure 2.6 to the Carhart model. As we did before, we report the root mean square errors $\sqrt{\frac{1}{N}\alpha'\alpha}$ (rmse hereafter), as well as the mean absolute value (mav hereafter) in Table 2.11.

It is clear that in all cases, a free constant allows the model to explain the returns better for both models (see Figures 2.5 and 2.6), though by allowing a substantial alpha and leading to economically misleading estimates. Table 2.11 show that error measures are between 23 to 97 per cent lower when imposing a free constant. The more evident case is OLS Fama-French while the less evident is OLS Carhart model.

Table 2.10: Cross-sectional estimation, Carhart model

Ordinary Least Squares (OLS)						
Equation	**Description**	**const.**	***rmrf***	***smb***	***hml***	***umd***
(1.3.2)	λ without a constant		**0.775**	**0.274**	**0.600**	**3.279**
(1.2.2)	factor mean		**0.709**	**0.278**	**0.463**	0.693
(1.3.2)	λ with a constant	**1.362**	**−0.564**	0.238	**0.540**	1.402
(1.3.4)	se, i.i.d., no Shanken		0.185	0.120	0.129	0.671
(1.3.12)	se, i.i.d., Shanken		0.187	0.123	0.133	0.900
(1.2.5)	se, GMM, 0 lags		0.186	0.128	0.131	1.342
(1.2.5)	se, GMM, 12 lags		0.192	0.126	0.143	1.526
Generalized Least Squares (GLS)						
(1.3.2)	λ without a constant		**0.788**	**0.265**	**0.497**	**1.886**
(1.2.2)	factor mean		**0.709**	**0.278**	**0.463**	0.693
(1.3.2)	λ with a constant	**1.148**	**−0.360**	**0.275**	**0.475**	0.855
(1.3.4)	se, i.i.d., no Shanken		0.182	0.117	0.126	0.387
(1.3.12)	se, i.i.d., Shanken		0.183	0.118	0.127	0.438
(1.2.5)	se, GMM, 0 lags		0.184	0.115	0.128	0.760
(1.2.5)	se, GMM, 12 lags		0.189	0.121	0.149	0.964

Values statistically different from zero in bold.

Table 2.11: Cross-sectional error measures

Procedure	**rmse**	**mav**
Fama-French (Figure 2.5)		
OLS without constant	0.1938	0.1401
OLS with constant	0.0979	0.0715
GLS without constant	0.2233	0.1318
GLS with constant	0.1176	0.0733
Carhart (Figure 2.6)		
OLS without constant	0.1143	0.0860
OLS with constant	0.0892	0.0699
GLS without constant	0.1394	0.0868
GLS with constant	0.1068	0.0648

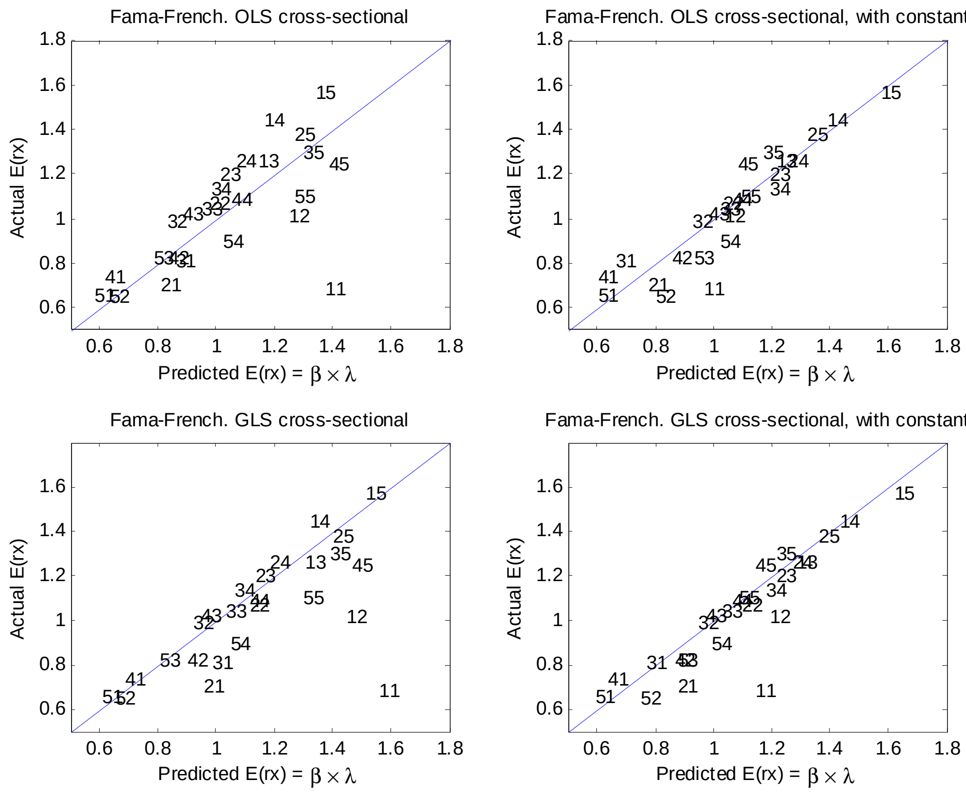

Figure 2.5: Cross-section Fama-French model actual versus predicted returns

On the other hand, it is clear that Carhart model is associated with lower error measures in all cases, even though we add a fourth factor. Anyway, we do a formal test in order to clarify this subject in Table 2.12, it should be not a surprise that they lead to a rejection of Fama-French model when confronted to Carhart.

Another part of the story is the much better fit of the 11 and 12 portfolios. The 11 portfolio in Figure 2.5 had a large beta $\widehat{b}_i = 1.307$. By fitting a huge negative market risk premium (-1.1713 for OLS and -0.8071 for GLS) but making up for it with a huge positive constant for the other assets, (1.9660 for OLS and 1.5832 for GLS), the regression can fit this observation better, but not very credible.

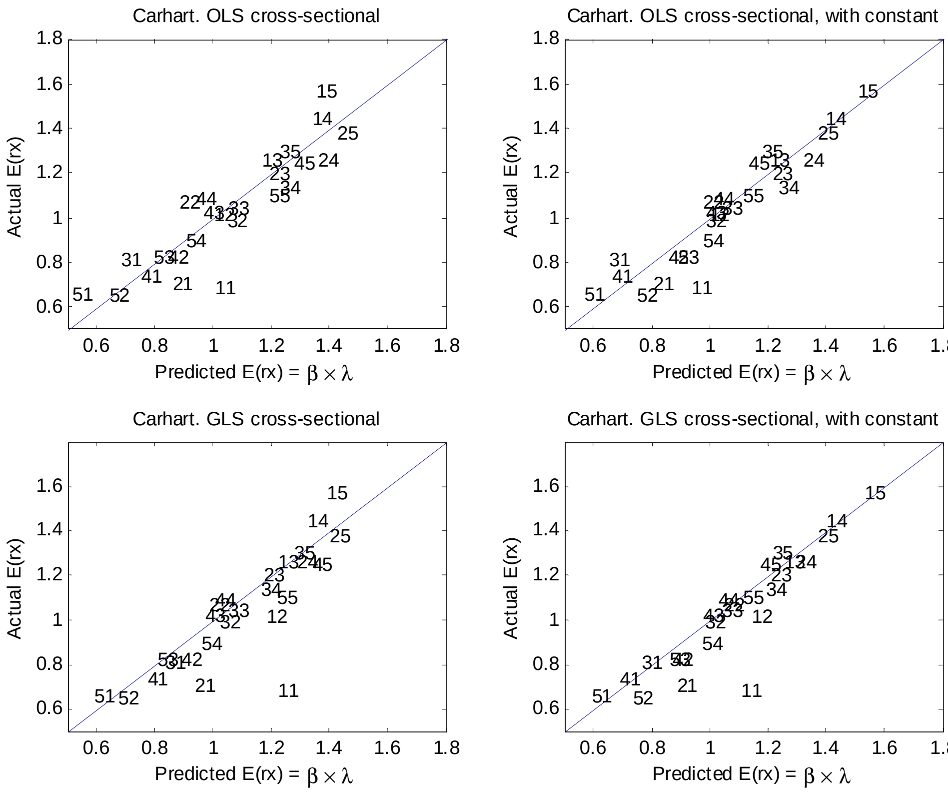

Figure 2.6: Cross-section Carhart model actual versus predicted returns

Tests statistics

It is not convenient to use a measure like a cross-sectional R^2 (average returns on predicted average returns) because can be a dangerous statistic. The cross-sectional R^2 rises automatically as we add factors, besides it depends a lot on the estimation method. R^2 is only well-defined for an OLS cross-sectional regression of average returns on betas with a free intercept. For any other estimation technique, various

Table 2.12: Cross sectional estimation, test statistics

Fama-French model*				
	fixed beta	**0 lags**	**12 lags**	**Shanken**
OLS χ^2	72.8960	78.4661	93.6301	70.9358
p-value $\times$ 1.0e-006	0.2299	0.0289	0.0001	0.4708
GLS χ^2	72.8960	78.2119	94.9172	70.7715
p-value $\times$ 1.0e-006	0.2299	0.0318	0.00001	0.4999
Carhart model**				
OLS χ^2	**26.0161**	**14.1000**	**13.1512**	**14.7668**
p-value	0.2058	0.8653	0.9032	0.8345
GLS χ^2	**38.0466**	**28.4584**	**15.4205**	**23.8316**
p-value	0.0127	0.1276	0.8013	0.3013

* Fama-French: 22 df, 5% and 1% critical values: 33.9244 and 40.2894.
** Carhart: 21 df, 5% and 1% critical values: 32.6706 and 38.9322.

ways of computing R^2 can give wildly different results. These criticisms are of course solved by statistical measures; test statistics based on $\alpha' V^{-1} \alpha$ which pricing errors are invariant to portfolio formation and take account of degrees of freedom. These tests are those presented in equations (1.3.6), (1.3.16) and (1.3.11).

Table 2.12 illustrate that the Fama-French model is rejected while Carhart is not[16]. Ferson and Harvey [37] also reject that a conditional version of the Fama-French three-factor model captures all the return predictability for these test portfolios. Now we can say with certain that adding momentum does help the model to price the average returns of portfolios at least in this 71 years sample using a cross-sectional estimation. Our next task in this work is to repeat this exercise using Fama-MacBeth procedure and GMM and see if these results are consistent with changes on the estimation

[16] Small p-values suggest that the null hypothesis is unlikely to be true. The smaller it is, the more convincing is the rejection of the null hypothesis. It indicates the strength of evidence for say, rejecting the null hypothesis H_0, rather than simply concluding reject H_0 or do not reject H_0.

method.

It is interesting to check again that Shanken correction does not make much difference for this kind of tests for the Fama-French model, on the other hand does make difference when testing the Carhart model. This finding is consistent with Tables 2.9 and 2.10.

2.3.3 Fama-MacBeth

The main results of this subsection 2.3.3 are summarized in Tables 2.13 and 2.14, which show the estimators and the evaluation results. Our results are consistent to those on Lettau and Ludvigson [71], who follow a similar approach that allow factor risk premia and betas to vary over time, and find statistically insignificant *smb* factor over the 1965-1998 period.

In this section, we show that, when the right-hand variables do not vary over time, Fama-MacBeth is numerically equivalent to pooled time-series, cross-section OLS with standard errors corrected for cross-sectional correlation, and also to a single cross-sectional regression on time-series averages with standard errors corrected for cross-sectional correlations.

The Fama-MacBeth estimates and standard errors are exactly the same as OLS cross-sectional regression estimates (1.3.2) and non-Shanken corrected i.i.d. standard errors on (1.3.4). It turns our that the Fama-MacBeth procedure is another way of calculating the standard errors, corrected for cross-sectional correlation. Furthermore, GLS Fama-MacBeth is exactly the same as GLS cross-sectional regression (1.3.2). Also see Tables 2.9 and 2.10 from cross-sectional regressions.

Again, as in the cross-sectional regressions, we reject the Fama-French model and

Table 2.13: Fama-MacBeth estimators and standard errors

OLS (Equation 1.4.1)							
	Fama-French			**Carhart**			
	rmrf	***smb***	***hml***	***rmrf***	***smb***	***hml***	***umd***
λ_{FMB}	**0.696**	0.171	**0.454**	**0.775**	**0.274**	**0.600**	**3.279**
FMB, se	0.184	0.122	0.127	0.185	0.120	0.129	0.671
GMM, FMB, se	0.181	0.120	0.150	0.183	0.121	0.155	0.646
GLS (Equation 1.4.1)							
λ_{FMB}	**0.736**	**0.269**	**0.449**	**0.788**	**0.265**	**0.497**	**1.886**
FMB, se	0.182	0.117	0.126	0.183	0.117	0.126	0.387

Values statistically different from zero in bold.

Table 2.14: Test statistics, Fama-MacBeth

			Critical values	
Model	**Fama-MacBeth χ^2 statistic**	**p-value**	**5%**	**1%**
Fama-French	67.1507	0.0002	33.9244	40.2894
Carhart	**36.4203**	0.0196	32.6706	38.9322

do not reject the Carhart model, at least with $\alpha = 0.05$, following equation 1.4.2.

2.4 Empirical results: Generalized Method of Moments / Stochastic Discount Factor

Remember first-stage estimates imposes no serial correlation and regression errors independent of right-hand variables. Second-stage estimators are calculated using two prespecified weighting matrixes as we detailed in sections 1.5.2 and 1.5.3: the second moment $E_T\left(R^e f'\right)$ and the covariance $E_T(R^e \widetilde{f'})$; and finally using the optimal weighting matrix estimated by picking b and S simultaneously (see section 1.5.4). When using prespecified weighting matrix we consider zero (correcting only for conditional heteroskedasticity) and twelve lags on S (correcting for conditional heteroskedasticity and for high-order autocorrelation). Hence, second-stage GMM/SDF estimates, as well as standard errors, depend on which spectral density weighting matrix is used as weighting matrix.

2.4.1 Second moment matrix as the weighting matrix in second-stage estimators: $\mathbf{GMM_A}$

Hansen and Jagannathan [49] propose the use of the second-moment matrix of the payoffs. They point out that this matrix may be of interest because the minimized GMM loss function can be interpreted as the distance between the estimated SDF and the SDF that prices all assets (the true one). Another good characteristic of this weighting matrix is that it is invariant to the initial choice of assets.

Table 2.15: Fama-French: GMM_A estimates.

Estimates of b	***rmrf***	***smb***	***hml***
First-stage	**0.0202**	0.0006	**0.0283**
0 lag se	0.0072	0.0101	0.0091
t-statistic	2.7965	0.0584	3.1254
12 lag se	0.0082	0.0100	0.0117
t-statistic	2.4433	0.0590	2.4153
Second-stage, 0 lag	**0.0192**	0.0111	**0.0258**
se	0.0068	0.0090	0.0088
t-statistic	2.8285	1.2370	2.9461
Second-stage, 12 lag	**0.0219**	0.0107	**0.0313**
se	0.0071	0.0069	0.0079
t-statistic	3.0976	1.5365	3.9467

Values statistically different from zero in bold.

We present the results of estimating b in first- and second-stage according to equation (1.5.2), and its standard errors as indicated on equation (1.5.3). In particular, the results for Fama-French and Carhart model are in Tables 2.15 and 2.16 respectively. Also, the corresponding error measures are on Table 2.21.

Estimates b give some relative sense of how important each factor is in the discount factor. We can see that the estimates are reasonable stable across first- and second-stage at least in the Fama-French model on Table 2.15. The b t-statistic also gives a clue about which factor could be dropped, probably the *smb* factor can be dropped with little effect on the pricing of the portfolios, since it has low t-statistic, especially in the first-stage. In section 2.5 we evaluate this alternative, and the results are so interesting that we continue the analysis in the next chapter.

Note the contrast between the b t-statistic and the λ t-statistic. In the time-series

Table 2.16: Carhart: GMM_A estimates.

Estimates of b	***rmrf***	***smb***	***hml***	***umd***
First-stage	**0.0517**	0.0214	**0.1116**	**0.1849**
0 lag se	0.0130	0.0212	0.0281	0.0660
t-statistic	3.9905	1.0092	3.9709	2.8039
12 lag se	0.0126	0.0171	0.0332	0.0747
t-statistic	4.0996	1.2502	3.3625	2.4751
Second-stage, 0 lag	**0.0463**	0.0127	**0.0797**	**0.1567**
se	0.0092	0.0141	0.0172	0.0298
t-statistic	5.0570	0.8972	4.6365	5.2554
Second-stage, 12 lag	**0.0526**	**0.0275**	**0.0656**	**0.1734**
se	0.0084	0.0126	0.0176	0.0250
t-statistic	6.2991	2.1925	3.7320	6.9233

Values statistically different from zero in bold.

regression (on Table 2.8), we found that *smb* was *marginally* significant. Thus, this factor *is priced* though it does not *help to price* other securities[17].

On the other hand, Table 2.16 show an interesting result in favor of Carhart model compared with Fama-French. In particular, momentum is not only highly significant, it also increases the significance level of market, size and value factors, in such way that makes *smb* different from zero on second stage estimate with 12 lags.

Table 2.17 show that the Fama-French model is rejected at a huge significance level following equation 1.5.5. There is not much difference according to how it

[17] b_j asks whether factor j *helps to price* assets given the other factors. b_j gives the *multiple* regression coefficient of m on f_j given other factors.

λ_j asks wether factor j *is priced*, or whether its factor-mimicking portfolio carries a positive risk premium. λ_j gives the *single* regression coefficient of m on f_j.

Therefore, when factors are correlated, one should test $b_j = 0$ to see whether to include a factor j given other factors rather than test $\lambda_j = 0$.

Table 2.17: Test of overidentification, $\mathrm{GMM_A}$.

	First-stage			
	Fama-French*		**Carhart****	
	0 lags	**12 lags**	**0 lags**	**12 lags**
χ^2 statistic	79.3192	95.2363	**19.7046**	**21.0869**
p-value	0.0	0.0	0.5400	0.4536
	Second-stage			
χ^2 statistic	79.3192	95.2363	**19.7046**	**21.0869**
p-value	0.0	0.0	0.5400	0.4536

* Fama-French: 22 df, 5% and 1% critical values: 33.9244 and 40.2894.
** Carhart: 21 df, 5% and 1% critical values: 32.6706 and 38.9322.

is calculated. On the other hand, as expected, Carhart model is not rejected, the pricing errors are jointly equal to zero in all cases. Note that the values are the same for first and second-stage, this interesting result was already pointed out in section 1.5.2.

Now, we proceed to make the usual actual versus predicted plots in Figures 2.7 and 2.8 for the Fama-French and Carhart model respectively. Here we did it by plotting $E\left(R^e\right)$ versus $E\left(R^e f'\right) b$. The first-stage $\mathrm{GMM_A}$ on the upper panels looks just about like the OLS cross-sectional regression, and the second-stage (lower panels) looks about like GLS cross-sectional regressions. The second-stage using 12 lags shows a pattern which apparently happens often here (see Cochrane [25, ch. 15] for a discussion). Second-stage $\mathrm{GMM_A}$ is paying attention to portfolios $S^{-\frac{1}{2}}R^e$, and in doing so is not doing a very good job on R^e itself. In other words, second-stage favor the estimator efficiency by allowing some pricing error.

Note that Fama-French model, estimated via $\mathrm{GMM_A}$ still have difficulties at pricing portfolios 11 and 12, so we may think it is a matter of model and not of method.

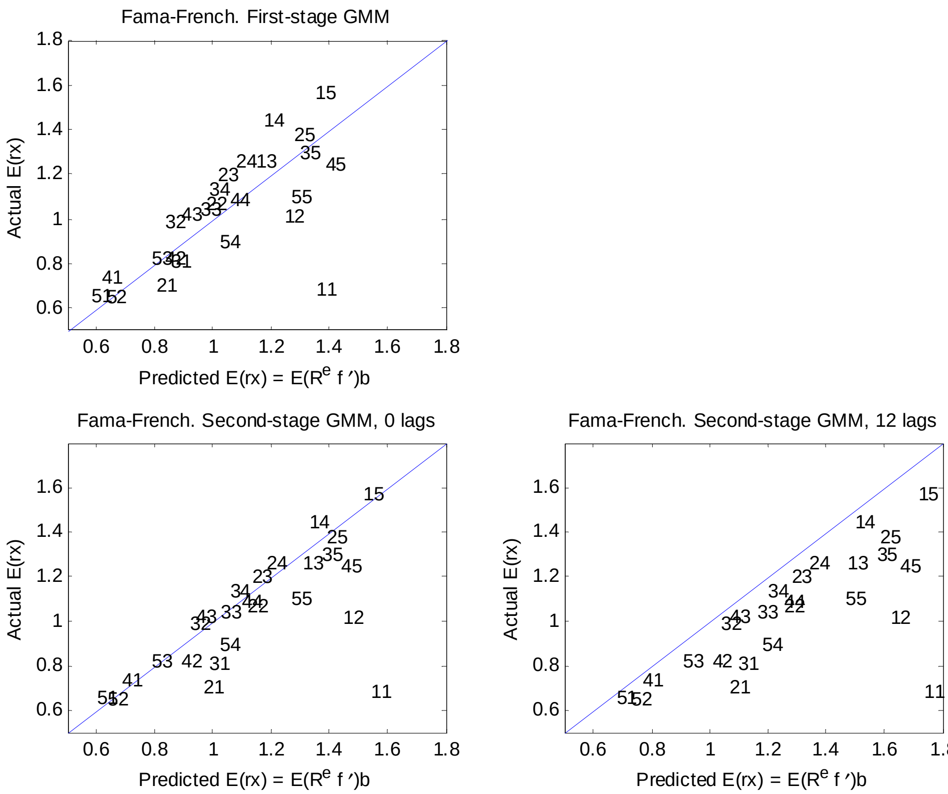

Figure 2.7: $\mathrm{GMM_A}$ actual versus predicted returns, Fama-French model

In particular, portfolio 11 is the most poorly priced, suggesting that the missing element in Fama-French model and in less extent in Carhart model may be an account of the cost of short-selling or the thinness of the market, see D'Avolio [27], and Lamont and Thaler [70].

It is important to notice that using 0 lags versus 12 lags made little difference to the standard errors, as we expect since there is little autocorrelation in these returns. But it makes a big difference to the second-stage estimate. So, *little variations in S matrices make little difference to standard errors, where S enters in the numerator,* *but can make big differences to second-stage estimates, which use S^{-1} to weight moments*. The inverse is much more sensitive to small perturbations. This is another reason to be very careful when doing second-stage or efficient GMM, and to think carefully about prespecified weighting matrices, or other ways of making sure that weighting matrix is not focusing on garbage.

The Fama-French GMM_A plots in Figure 2.7 are similar, but slightly different from the OLS and GLS cross-sectional regressions. That is because weighting by Σ^{-1} is not the same as weighting by S^{-1}, though it is close. The differences are not conclusive since the they are not that big. However, we develop the comparison between GMM and cross-sectional methods in the next chapter, in which we compare a number of estimators[18].

As shown in Figure 2.8 for the Carhart model, first-stage GMM_A considerably improves valuation of 11 and 12 portfolios. Actually this is one of the best fit plots that we get, with a rmse value of 0.0896 for the first-stage. Now, using 0 and 12 lags in the second-stage, does makes much more difference compared with first-stage. As we can see in the lower panels, inadequate weighting matrix can lead to large pricing errors, in particular about 3 times more than first-stage.

It is clear that second-stage estimators lead to bigger rmse and mav errors than Fama-French, see Table 2.21. Even though, the previous statistical test do not reject the Carhart model in all cases including second-stage. The reason is that S matrix in equation (1.5.2) is too big for weighting moments and reduces $\widehat{b}_2$, which reduces the predicted value $E\left(R^e f'\right) b$, so the model underestimates the actual expected returns. We should be cautious in the sense that J_T test may improve because a big S matrix

[18]The main result of chapter 3 about the GMM and cross-sectional methods in comparing pricing errors is that, in general, GMM lead to lower pricing errors than cross-sectional.

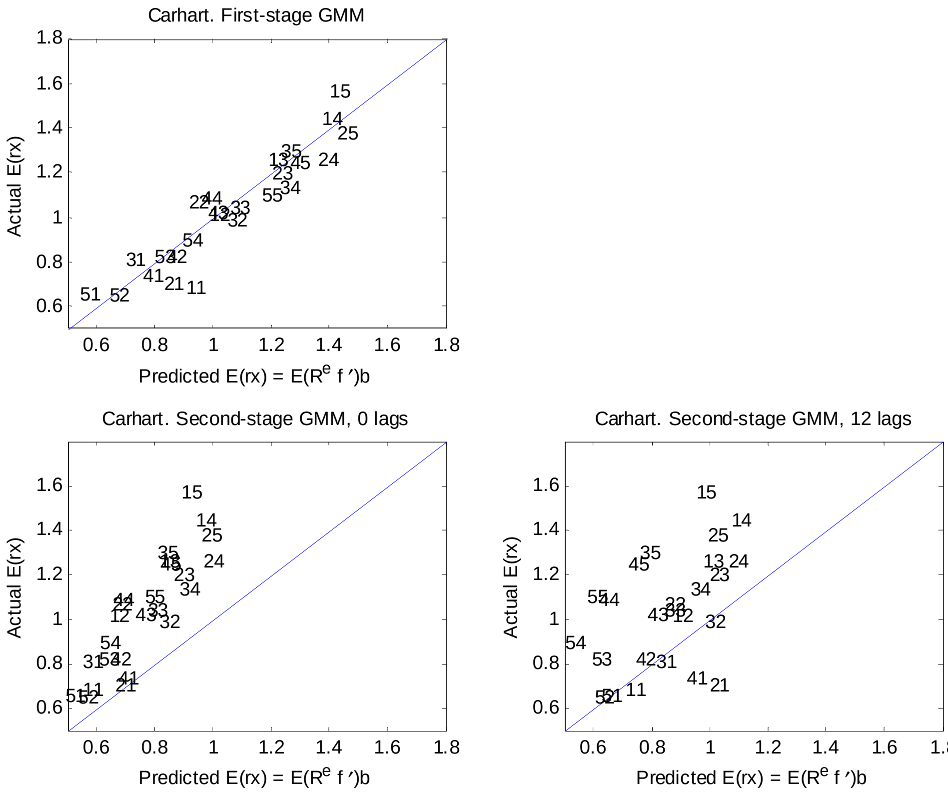

Figure 2.8: GMM_A actual versus predicted returns, Carhart model

rather than a low pricing errors.

2.4.2 Covariance as the weighting matrix in second-stage estimators: GMM_B

Table 2.18 shows the result of Fama-French model estimators (equation 1.5.6) and its standard errors (equation 1.5.7). The first-stage GMM_B estimate from a regression of average excess returns on covariances yield almost exactly the same results as the first-stage from excess returns on second moments GMM_A ($\widehat{b} = 0.0202; 0.0006; 0.0283$).

Table 2.18: Fama-French: GMM_B estimates

	b Estimates			$E(f)$ Estimates		
	rmrf	***smb***	***hml***	***rmrf***	***smb***	***hml***
First-stage $\left(\widehat{b}_1\right)$	**0.0208**	0.0006	**0.0287**	**0.7096**	**0.2780**	**0.4637**
se, with no correction						
0 lags	0.0075	0.0103	0.0094			
t-statistic	2.7863	0.0599	3.0626			
12 lags	0.0086	0.0102	0.0122			
t-statistic	2.4257	0.0606	2.3586			
correct se						
0 lags	0.0075	0.0103	0.0094	0.1820	0.1166	0.1243
t-statistic	2.7685	0.0599	3.0511	3. 898 9	2. 384 2	3. 730 5
12 lags	0.0086	0.0102	0.0122	0.1843	0.1224	0.1408
t-statistic	2.4200	0.0606	2.3514	3. 850 2	2. 271 2	3. 293 3
Second-stage						
0 lag estimator	**0.0176**	0.0108	**0.0248**	**0.5657**	**0.2638**	**0.3472**
se	0.0071	0.0092	0.0091	0.1697	0.1064	0.1148
t-statistic	2.4943	1.1701	2.7314	3.3345	2.4803	3.0240
12 lag estimator	0.0200	0.0104	0.0303	0.4398	0.2254	0.1997
se	0.0074	0.0071	0.0082	0.1425	0.0805	0.0951
t-statistic	2.7122	1.4560	3.6747	3.0873	2.8013	2.1009

Values statistically different from zero in bold.

Following the econometric discussion in section 1.5.3, in Table 2.18, we refer to the standard error with no correction as ignoring that Ef is actually estimated, that is, using $d = cov\,(R^e f')$. Corrected standard errors are those estimated by (1.5.7). Because there are little autocorrelation, both versions leads to very similar results. However, even though the empirical results show no big differences, we have to emphasize that we do have to correct the standard errors as usual.

Note that varying from second moments (Table 2.15) to covariance matrix as weighting matrix (Table 2.18) does not have a big impact on estimates and standard errors. There is no reason to expect a big difference, specially when estimating a single portfolio formation like the 25 Fama-French portfolios. On the next chapters, we extend the comparison between $\mathrm{GMM_A}$ and $\mathrm{GMM_B}$ in order to collect more evidence about the implications of both formulations.

It is important to compare these results of Table 2.19 with the second moment matrix of returns and factors (Table 2.16). First, independently of the weighting matrix, *smb* is statistically equal to zero in most cases. Second, independently of the standard error used t-statistics are practically the same at least for the first-stage $\mathrm{GMM_B}$ estimates.

Now go on the second-stage $\mathrm{GMM_B}$ estimates. Estimates are estimated from the minimization of (1.5.8). Remember that the mean factor $E_T\,(f)$ is 0.7096; 0.2780; 0.4637 and now the estimated values $E\,(f)$ are different in order to do better on the other moments. Standard errors are calculated using (1.5.3).

Table 2.20 show that one more time, we cannot reject Carhart model, and Fama-French show the usual rejection according to equation 1.5.5.

Table 2.19: Carhart: GMM_B estimates

	b Estimates				$E(f)$ Estimates			
	rmrf	***smb***	***hml***	***umd***	***rmrf***	***smb***	***hml***	***umd***
First-stage $\left(\widehat{b}_1\right)$	**0.062**	0.025	**0.133**	**0.216**	**0.709**	**0.278**	**0.463**	**0.693**
se, with no correction								
0 lags	0.018	0.024	0.038	0.087				
t-statistic	3.475	1.019	3.501	2.484				
12 lags	0.018	0.020	0.046	0.103				
t-statistic	3.400	1.228	2.880	2.093				
correct se								
0 lags	0.018	0.024	0.039	0.088	0.182	0.116	0.124	0.159
t-statistic	3.388	1.023	3.416	2.437	3. 898	2. 384	3. 730	4. 352
12 lags	0.018	0.020	0.047	0.105	0.184	0.122	0.140	0.144
t-statistic	3.301	1.226	2.824	2.046	3. 850	2. 271	3. 293	4. 808
Second-stage								
0 lag estimator	**0.051**	0.015	**0.085**	**0.159**	**0.686**	**0.315**	**0.475**	**0.733**
se	0.012	0.016	0.023	0.043	0.162	0.101	0.110	0.139
t-statistic	4.052	0.919	3.609	3.675	4.223	3.116	4.286	5.255
12 lag estimator	**0.059**	**0.032**	**0.071**	**0.184**	**0.686**	**0.275**	**0.524**	**0.719**
se	0.011	0.015	0.023	0.037	0.151	0.097	0.107	0.096
t-statistic	5.137	2.123	3.090	4.935	4.539	2.838	4.867	7.482

Values statistically different from zero in bold.

Table 2.20: Test of overidentification, GMM_B

	First-stage			
	Fama-French*		**Carhart****	
	0 lags	**12 lags**	**0 lags**	**12 lags**
χ^2 statistic	78.4661	93.6301	**19.8431**	**22.6294**
p-value	0.0	0.0	0.5312	0.3641
	Second-stage			
χ^2 statistic	78.6093	93.5595	**19.7514**	**22.4447**
p-value	0.0	0.0	0.5371	0.3743

* Fama-French: 22 df, 5% and 1% critical values: 33.9244 and 40.2894.
** Carhart: 21 df, 5% and 1% critical values: 32.6706 and 38.9322.

Table 2.21: GMM_A and GMM_B error measures

	rmse		**mav**	
Procedure	GMM_A	GMM_B	GMM_A	GMM_B
Fama-French model				
First-stage	0.1885	0.1831	0.1363	0.1251
Second-stage 0	0.2175	0.1995	0.1273	0.1408
Second-stage 12	0.3128	0.1984	0.2231	0.1375
Carhart model				
First-stage	0.0896	0.1061	0.0705	0.0741
Second-stage 0	0.3320	0.3050	0.2971	0.2690
Second-stage 12	0.3085	0.3065	0.2529	0.2703

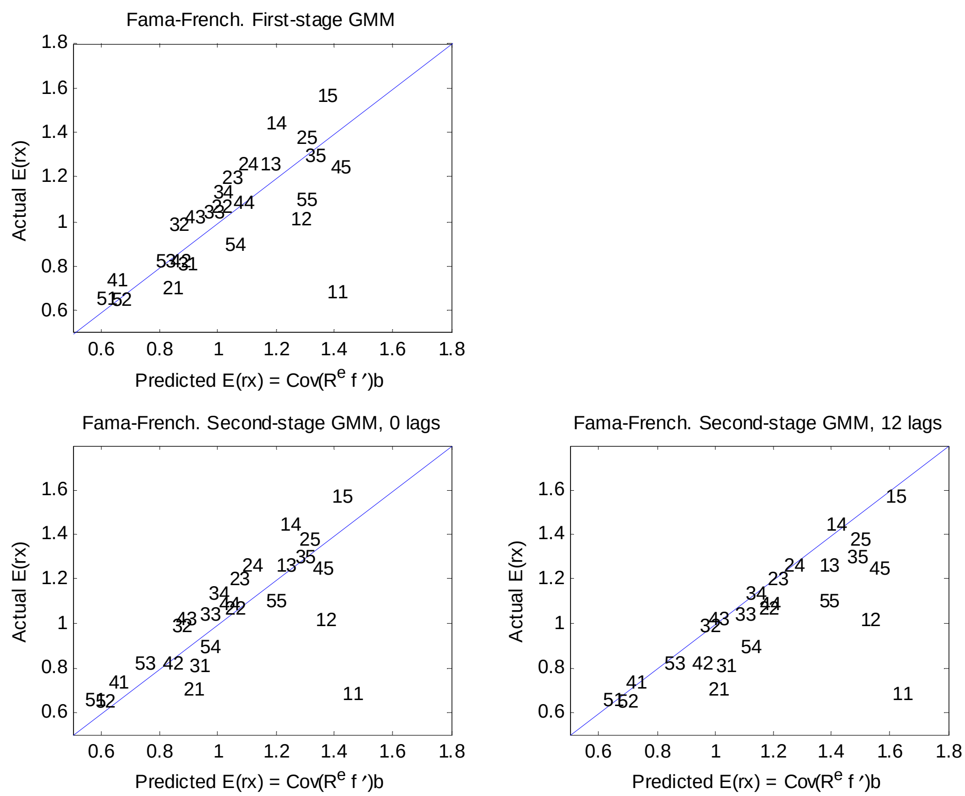

Figure 2.9: $\mathrm{GMM_B}$ actual versus predicted returns, Fama-French model

In general, Figures 2.9 and 2.10 look very much like the plots based on second moments above. It is interesting to highlight that the second-stage $\mathrm{GMM_B}$ with 12 lags looks decidedly better, so there is some indication that the covariance based estimate is a little more stable in the second-stage where we apply strong weightings. Even though, we must keep in mind that if we choose a too long value of lags, together with lack of autocorrelation, the performance of the estimate and test deteriorates. The optimum value of lags depends on how much persistence or low-frequency movements there is in a particular application versus accuracy of the estimate.

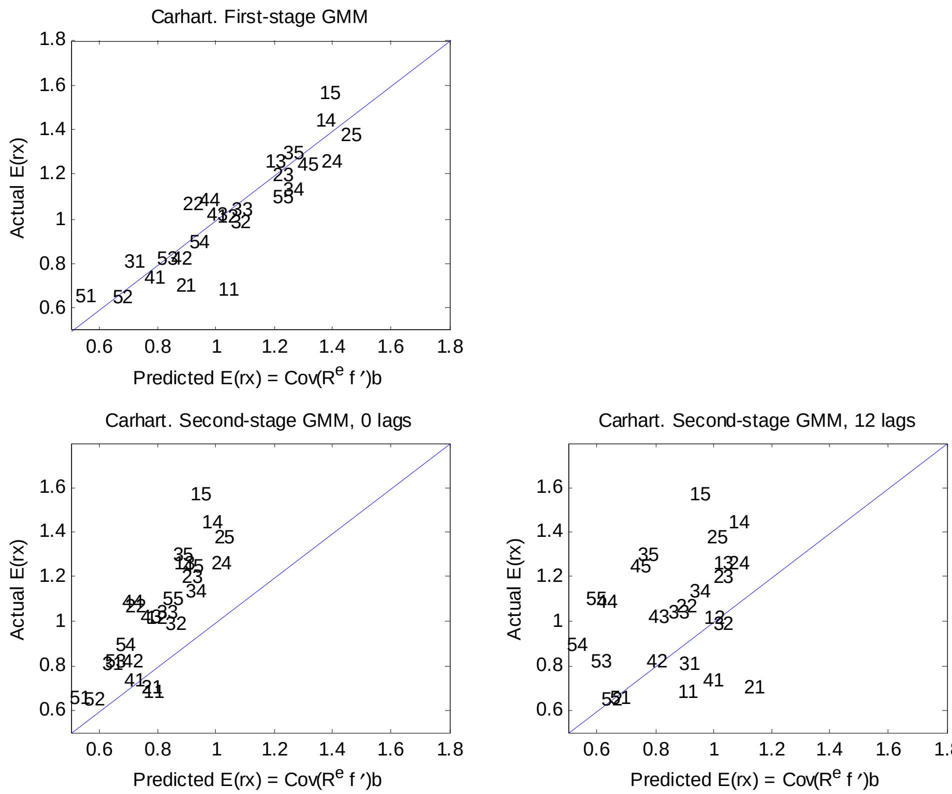

Figure 2.10: $\mathrm{GMM_B}$ actual versus predicted returns, Carhart model

Based on error measures (rmse and mav), according to Table 2.21, $\mathrm{GMM_A}$ second moment estimators presented on section 2.4.1 leads to a marginal better fit than $\mathrm{GMM_B}$ covariances estimators at least for second-stage estimators. Once again, we get a big S matrix which leads to a false sense of improvement in J_T test in the way that pricing errors are actually bigger than Fama-French. Far from being a J_T comparison among models, we just want to clarify why we get a better statistic with bigger pricing errors.

Table 2.22: GMM_C estimators, Fama-French model

	b Estimator			$E(f)$ Estimator		
	rmrf	***smb***	***hml***	***rmrf***	***smb***	***hml***
Estimator	**0.0708**	−0.0110	**0.1007**	**0.8384**	**0.2983**	**0.4623**
se	0.0137	0.0192	0.0226	0.1821	0.1166	0.1243
t-statistic	5.1535	−0.5743	4.4580	4.6046	2.5590	3.7189

Values statistically different from zero in bold.

2.4.3 Efficient GMM: GMM_C

It seems interesting to try second-stage (efficient GMM_C) with the estimated spectral density matrix, we did it by picking b and S simultaneously as explained in section 1.5.4. The estimates and tests are done according to (1.5.9) and the corresponding Tables are 2.22 and 2.23.

Clearly, estimators are very different from others. There is no sense of improvement, so probably at estimating S and b at the same time, we are not making any progress on lowering the pricing errors as explained before. Also, GMM_C standard errors in Fama-French model are two times bigger than in GMM_B Cochrane [25] specification (section 2.4.2), where we take $E_T(R^e\widetilde{f}\prime)$ as the weighting matrix in second-stage estimators.

Next, we review the case of the Carhart model on GMM_C (Table 2.23). It seems that in this apparently bad specification, Carhart model is doing better than Fama-French. Even though there is no sense of improvement by estimating b and S simultaneously. We already knew this possibility from the methodology in section 1.5.4. Note that GMM_C standard errors are more than five times bigger than in GMM_B Cochrane [25] specification, see Tables 2.18 and 2.19.

Table 2.23: GMM_C estimators, Carhart model

	b Estimator				$E(f)$ Estimator			
	rmrf	***smb***	***hml***	***umd***	***rmrf***	***smb***	***hml***	***umd***
Estimator	0.0569	0.0659	0.0997	0.4862	**0.6231**	0.2239	**0.4225**	**0.8680**
se	0.0445	0.0794	0.1057	0.3088	0.1820	0.1166	0.1243	0.1594
t-statistic	1.2798	0.8309	0.9433	1.5745	3.4231	1.9208	3.3978	5.4442

Values statistically different from zero in bold.

The χ^2 test on Table 2.24 give the usual rejections for Fama-French model, but now with a higher p-value. The calculations are following equation 1.5.5.

As expected, we reach a new minimum J_T value on GMM_C specification because it has more free parameters than GMM_B specification, where we have a fixed S. This leads to a lower TJ_T test statistic and a higher p-value, suggesting that the null hypothesis is more likely to be true. This new TJ_T value is not low enough in order to not reject Fama-French model, on the other hand, Carhart model is not rejected (again) by the data.

The previous are unusual results. First, GMM_C blows up standard errors of estimates, see Tables 2.22 and 2.23. Second, higher p-values on tests of overidentifying restrictions. Usual plots will help on clarifying what is going on.

First-stage GMM_A and GMM_B plots are taken from Figures 2.7 and 2.9, we include them again in order to facilitate their comparisons. The second-stage plots on the lower panels (GMM_C) are the new plots.

The fist impression is that error measures are extremely big. Now look at the Carhart model plot in Figure 2.12 before concluding on GMM_C results.

Remember we have $J_T \equiv \min_{\{b,Ef\}} \left[g_T(b,Ef)' \, S^{-1}(b,Ef) \, g_T(b,Ef) \right]$. The models *improve* TJ_T statistic because it blows up the estimates of S, rather than by making

Table 2.24: Tests of overidentification, GMM_{C}

Cochrane [25]: $J_T \equiv \min_{\{b,Ef\}} \left[g_T(b,Ef)' S^{-1} g_T(b,Ef)\right]$		
	Fama-French*	**Carhart****
J_T	0.0923	0.0263
$TJ_T \sim \chi^2$ statistic	78.6093	**22.4447**
p-value	0.0	0.3743
Efficient: $J_T \equiv \min_{\{b,Ef\}} \left[g_T(b,Ef)' S^{-1}(b,Ef)\, g_T(b,Ef)\right]$		
	Fama-French*	**Carhart****
J_T	0.0781	0.0233
$TJ_T \sim \chi^2$ statistic	66.5440	**19.8527**
p-value	$2.2871e-006$	0.5306

* Fama-French: 22 df, 5% and 1% critical values: 33.9244 and 40.2894.
** Carhart: 21 df, 5% and 1% critical values: 32.6706 and 38.9322.

any progress in lowering the pricing errors. Thus, results on Table 2.24 represent a false sense of improvement. Estimators place too much weight on portfolios that are economically uninteresting.

The objective of using the efficient matrix, given by the spectral density of the sample moments computed in the first stage of GMM, is to maximize the asymptotic information in the sample about a model. The danger of using such a matrix is that it may blow up, as we can see, standard errors rather than improve pricing errors. The efficient matrix will focus on linear combinations of returns that have low variance. Therefore, it may ignore the value premium and the size effect if they are hard to price in terms of variability. With a prespecified weighting matrix, we are giving up asymptotic efficiency but still obtaining consistent and more robust estimations. Nevertheless, we compute the spectral density matrix in order to have the correct variances of the estimates and the moments.

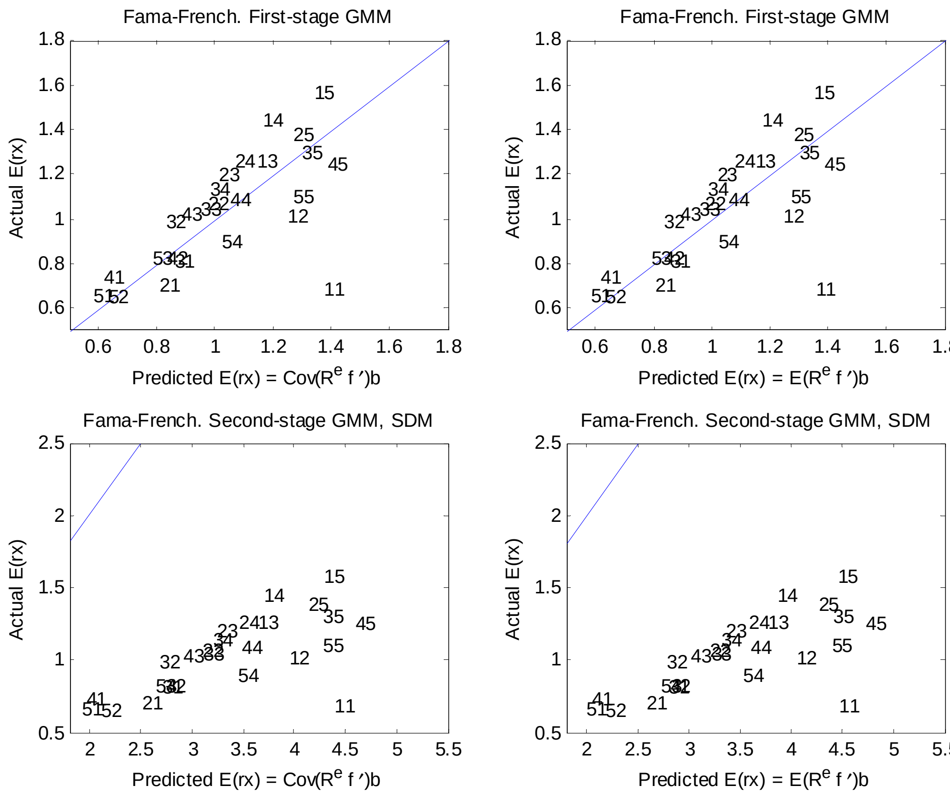

Figure 2.11: GMM_C actual versus predicted returns, Fama-French model

2.5 Can we drop *smb* factor?

It is known that size effect disappeared around the mid 1980's (see Banz [6]), and our results confirm in some way this popular belief. In Figure 2.1, we present a plot of average returns versus market beta for the 25 Fama-French portfolios and it was clear that size effect was not so evident, or at least not as much as the value premium. Cochrane [25] advocates that it is likely that we can get similar results with book-to-market effects alone. Furthermore, most of the empirical results from section 2.4 (GMM/SDF) suggest that we could drop *smb* factor. Hence, we have theoretical

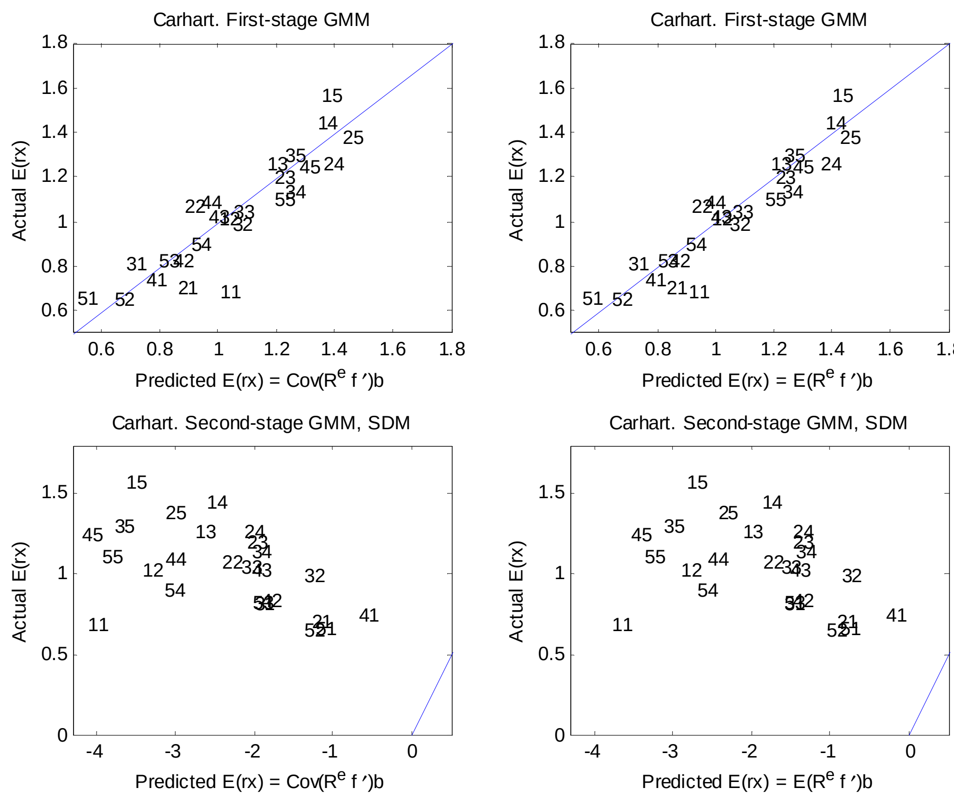

Figure 2.12: GMM_C actual versus predicted returns, Carhart model

and empirical reasons in order to think that a reasonable specification that could be superior to Fama-French and Carhart model is a linear three factor model: *rmrf*, *hml* and *umd*, that is Carhart model without size factor.

In the next chapter we will compare the performance of the Carhart model without size factor, and we will denote it by RUH model, because of the included factors (r for excess market return, u for momentum and h for value factors).

Let us briefly review the results from estimating this proposed model in Table 2.25. When testing the model according to the cross-sectional estimation via OLS and GLS

Table 2.25: Pricing errors by method and model

	Root mean square error (rmse)		
Estimation method	Fama-French *rmrf, smb, hml*	Proposed model *rmrf, hml, umd*	Carhart *rmrf, smb, hml, umd*
Time-series	0.2171	0.1799	0.1669
Cross-sectional - OLS without constant	0.1938	**0.1489**	**0.1143**
Cross-sectional - GLS without constant	0.2233	**0.1622**	**0.1394**
$\mathrm{GMM_A}$: First-stage	0.1885	**0.1250**	**0.0896**
$\mathrm{GMM_B}$: First-stage	0.1831	**0.1407**	**0.1061**
$\mathrm{GMM_C}$	2.354	**4.571**	**3.696**

Pricing errors statistically equal to zero in bold.

procedure, we cannot reject the null hypothesis that pricing errors are equal to zero, even though rmse is bigger than Carhart model but smaller than Fama-French. All estimators are statistically different from zero when we use a prespecified weighting matrix on GMM procedure, and pricing errors are equal to zero except when we impose a 12 lag on second-stage estimators.

In general, statistical test and error measures are slightly worse than Carhart model, but decidedly better than Fama-French. One could argue that in someway the new proposed model outperform Carhart model since it has three instead of four factors. Furthermore, their pricing errors are statistically equal to zero like the four factor model.

Here we summarize the root mean square errors produced by different estimation methods. Note that the proposed model and Carhart model are not rejected by their respective test statistics under the same estimation criteria. Even though the magnitude is bigger on the proposed three factor model.

2.6 Conclusions

For Fama-French and Carhart asset pricing models, we apply apparently different approaches for estimation and evaluation such as times-series, cross-sectional, Fama-MacBeth and GMM; these methods are actually the ones used in empirical practice. However, at the end, all these approaches do the same thing: they pick free parameters of the model to make it fit best, which usually means to minimize pricing errors; and they evaluate the model by examining how big those pricing errors are.

We find that GMM first-stage estimates perform better in terms of rmse, mav and statistical tests than second-stage estimates, because it only focus on minimizing the pricing errors. Also, efficient GMM estimates (GMM_C) perform worst than the second-stage estimators (GMM_A and GMM_B) since an efficient weighting matrix focus on economically uninteresting moments. On the other hand, Carhart model is able to price the 25 test portfolios while Fama-French model fails on specific ones. Finally, according to our findings, we propose a slightly different specification that works somewhat better than Fama-French and Carhart models on explaining cross-sectional returns.

Extensions should go over the economic interpretation of the *smb*, *hml* and *umd* factors. Among the many competing explanations behind the success of these models is the one based on time-varying investment opportunities. Specifically, Fama and French [30] suggest that *hml* and *smb* might proxy for state variables that describe time variation in the investment opportunity set. This is done by relating the Fama-French factors to macroeconomic variables and business cycle fluctuations. Liew and Vassalou [74], for instance, show that *hml* and *smb* help forecast future rates of economic growth, and both Lettau and Ludvigson [71] and Vassalou [99] show that accounting for macroeconomic risk reduces the information content of *hml* and *smb*. On the other hand, authors such as Petkova [87] argue that changes in financial investment opportunities are not necessarily exclusively related to news about future macro variables; furthermore, Campbell [16] points out that the factors in the model should be related to innovations in state variables that forecast future investment opportunities.

■

2.7 Appendix

2.7.1 Time-series

Table 2.26: Fama-French and Carhart standard errors se($\widehat{\alpha}_i$)

	Fama-French					**Carhart**				
OLS										
	low	**2**	**3**	**4**	**high**	**low**	**2**	**3**	**4**	**high**
small	0.253	0.146	0.105	0.078	0.088	0.259	0.149	0.108	0.081	0.090
2	0.091	0.068	0.064	0.055	0.065	0.093	0.069	0.066	0.057	0.067
3	0.070	0.064	0.063	0.062	0.072	0.071	0.066	0.065	0.064	0.073
4	0.058	0.061	0.065	0.069	0.085	0.060	0.063	0.067	0.070	0.087
big	0.041	0.050	0.061	0.064	0.107	0.042	0.051	0.063	0.065	0.110
GMM										
small	0.203	0.118	0.093	0.072	0.076	0.228	0.157	0.098	0.071	0.087
2	0.080	0.063	0.057	0.051	0.063	0.091	0.065	0.058	0.052	0.067
3	0.064	0.062	0.062	0.060	0.068	0.067	0.064	0.062	0.064	0.070
4	0.056	0.058	0.060	0.066	0.082	0.058	0.060	0.068	0.072	0.081
big	0.041	0.050	0.061	0.063	0.109	0.044	0.052	0.064	0.061	0.116
GMM 12 Newey-West lags										
small	0.177	0.121	0.082	0.075	0.070	0.173	0.109	0.079	0.068	0.079
2	0.078	0.074	0.055	0.052	0.061	0.076	0.067	0.060	0.048	0.065
3	0.061	0.057	0.059	0.055	0.066	0.060	0.059	0.056	0.059	0.068
4	0.069	0.073	0.062	0.066	0.087	0.065	0.064	0.061	0.076	0.081
big	0.045	0.052	0.067	0.055	0.118	0.043	0.050	0.061	0.048	0.122

Table 2.27: Fama-French and Carhart standard errors se($\widehat{b}_i$)

	Fama-French					Carhart				
OLS										
	low	**2**	**3**	**4**	**high**	**low**	**2**	**3**	**4**	**high**
small	0.051	0.029	0.021	0.016	0.018	0.052	0.030	0.022	0.016	0.018
2	0.018	0.014	0.013	0.011	0.013	0.019	0.014	0.013	0.011	0.013
3	0.014	0.013	0.013	0.013	0.014	0.014	0.013	0.013	0.013	0.015
4	0.012	0.012	0.013	0.014	0.017	0.012	0.013	0.013	0.014	0.017
big	0.008	0.010	0.012	0.013	0.022	0.008	0.010	0.013	0.013	0.022
GMM										
small	0.170	0.097	0.052	0.025	0.031	0.143	0.072	0.050	0.024	0.031
2	0.039	0.019	0.017	0.016	0.018	0.037	0.018	0.017	0.017	0.019
3	0.023	0.021	0.030	0.016	0.024	0.019	0.019	0.029	0.016	0.020
4	0.019	0.027	0.025	0.029	0.026	0.017	0.027	0.020	0.023	0.023
big	0.010	0.020	0.021	0.024	0.046	0.010	0.019	0.021	0.023	0.041
GMM 12 Newey-West lags										
small	0.113	0.072	0.055	0.020	0.038	0.094	0.060	0.054	0.020	0.040
2	0.040	0.022	0.021	0.018	0.025	0.036	0.019	0.023	0.019	0.026
3	0.031	0.026	0.031	0.016	0.027	0.026	0.023	0.031	0.014	0.021
4	0.013	0.028	0.019	0.027	0.029	0.012	0.027	0.018	0.020	0.023
big	0.011	0.025	0.031	0.027	0.034	0.011	0.025	0.029	0.025	0.030

Table 2.28: Fama-French and Carhart standard errors se($\widehat{s}_i$)

	Fama-French					**Carhart**				
OLS										
	low	**2**	**3**	**4**	**high**	**low**	**2**	**3**	**4**	**high**
small	0.078	0.045	0.033	0.024	0.027	0.078	0.045	0.033	0.024	0.027
2	0.028	0.021	0.020	0.017	0.020	0.028	0.021	0.020	0.017	0.020
3	0.022	0.020	0.019	0.019	0.022	0.021	0.020	0.019	0.019	0.022
4	0.018	0.019	0.020	0.021	0.026	0.018	0.019	0.020	0.021	0.026
big	0.013	0.015	0.019	0.020	0.033	0.013	0.015	0.019	0.020	0.033
GMM										
small	0.208	0.117	0.099	0.110	0.092	0.213	0.115	0.098	0.112	0.091
2	0.065	0.081	0.090	0.053	0.049	0.063	0.082	0.091	0.053	0.048
3	0.049	0.038	0.036	0.051	0.046	0.050	0.038	0.036	0.052	0.043
4	0.050	0.033	0.047	0.040	0.046	0.051	0.033	0.046	0.038	0.044
big	0.025	0.026	0.033	0.036	**0.079**	0.024	0.027	0.033	0.036	0.077
GMM 12 Newey-West lags										
small	0.270	0.142	0.092	0.122	0.113	0.275	0.141	0.092	0.123	0.110
2	0.061	0.086	0.097	0.062	0.042	0.057	0.084	0.098	0.064	0.041
3	0.046	0.058	0.059	0.070	0.050	0.048	0.061	0.058	0.071	0.046
4	0.058	0.050	0.073	0.030	0.045	0.058	0.049	0.073	0.030	0.040
big	0.031	0.042	0.033	0.042	**0.092**	0.030	0.042	0.033	0.041	0.093

Table 2.29: Fama-French and Carhart standard errors $\text{se}(\widehat{h}_i)$

	Fama-French					**Carhart**				
					OLS					
	low	**2**	**3**	**4**	**high**	**low**	**2**	**3**	**4**	**high**
small	0.070	0.040	0.029	0.022	0.024	0.074	0.043	0.031	0.023	0.026
2	0.025	0.019	0.018	0.015	0.018	0.027	0.020	0.019	0.016	0.019
3	0.019	0.018	0.017	0.017	0.020	0.020	0.019	0.019	0.018	0.021
4	0.016	0.017	0.018	0.019	0.024	0.017	0.018	0.019	0.020	0.025
big	0.011	**0.014**	0.017	0.018	0.030	0.012	0.015	0.018	0.019	0.031
					GMM					
small	**0.318**	**0.173**	0.078	0.038	0.068	0.250	0.116	0.069	0.041	0.056
2	0.065	0.036	0.036	0.029	0.032	0.051	0.038	0.038	0.028	0.033
3	0.045	0.037	0.033	0.031	0.048	0.041	0.032	0.032	0.031	0.037
4	0.039	0.034	0.052	0.053	0.040	0.033	0.034	0.039	0.041	0.036
big	0.016	**0.026**	0.029	0.029	0.080	0.016	0.026	0.026	0.029	0.066
					GMM 12 Newey-West lags					
small	**0.276**	0.070	0.062	0.039	0.069	0.092	0.096	0.040	0.023	0.046
2	0.069	0.056	0.058	0.043	0.028	0.057	0.027	0.030	0.025	0.017
3	0.071	**0.058**	0.056	0.055	0.058	0.033	0.039	0.035	0.033	0.031
4	0.033	0.057	0.063	0.075	0.058	0.022	0.032	0.041	0.042	0.028
big	0.024	**0.047**	0.046	0.037	0.068	0.018	0.027	0.032	0.021	0.031

Table 2.30: Carhart standard errors se($\widehat{u}_i$)

OLS					
	low	**2**	**3**	**4**	**high**
small	0.060	0.034	0.025	0.019	0.021
2	0.022	0.016	0.015	0.013	0.015
3	0.016	0.015	0.015	0.015	0.017
4	0.014	0.015	0.016	0.016	0.020
big	0.010	0.012	0.014	0.015	0.025
GMM					
small	0.186	0.163	0.050	0.024	0.042
2	0.057	0.021	0.022	0.018	0.021
3	0.025	0.031	0.028	0.025	0.027
4	0.029	0.022	0.047	0.037	0.025
big	0.013	0.020	0.027	0.018	0.047
GMM 12 Newey-West lags					
small	0.092	0.096	0.040	0.023	0.046
2	0.057	0.027	0.030	0.025	0.017
3	0.033	0.039	0.035	0.033	0.031
4	0.022	0.032	0.041	0.042	0.028
big	0.018	0.027	0.032	0.021	0.031

Table 2.31: Fama-French and Carhart OLS $t(\widehat{\alpha}_i)$

	Fama-French model					**Carhart model**				
	OLS									
	low	**2**	**3**	**4**	**high**	**low**	**2**	**3**	**4**	**high**
small	**−3.4**	**−2.8**	−0.4	1.5	0.6	**−2.5**	−1.7	0.2	1.9	1.7
2	**−2.5**	−0.4	1.1	1.5	−0.2	**−2.1**	0.7	1.2	1.2	0.3
3	**−2.1**	1.3	0.3	1.3	−1.2	−0.8	0.9	0.6	1.0	0.1
4	1.2	−0.8	1.4	−0.1	**−2.4**	0.8	−0.3	1.6	1.0	−1.1
big	1.8	0.4	0.6	**−2.1**	−1.7	**2.4**	0.6	1.0	−0.8	−0.9
	GMM									
small	**−4.2**	**−3.5**	−0.4	1.7	0.7	**−2.9**	−1.6	0.2	2.2	1.7
2	**−2.9**	−0.5	1.2	1.6	−0.2	**−2.2**	0.7	1.4	1.3	0.3
3	**−2.3**	1.3	0.3	1.3	−1.2	−0.8	1.0	0.7	1.0	0.1
4	1.3	−0.9	1.5	−0.2	**−2.5**	0.8	−0.3	1.6	1.0	−1.2
big	1.9	0.4	0.6	**−2.1**	−1.7	**2.3**	0.6	1.0	−0.9	−0.8
	GMM 12 Newey-West lags									
small	**−4.8**	**−3.4**	−0.5	1.6	0.7	**−3.8**	**−2.4**	0.3	**2.3**	1.9
2	**−2.9**	−0.4	1.3	1.5	−0.2	**−2.7**	0.7	1.4	1.4	0.3
3	**−2.4**	1.4	0.3	1.4	−1.3	−0.9	1.0	0.8	1.1	0.1
4	1.0	−0.7	1.4	−0.2	**−2.4**	0.7	−0.3	1.8	0.9	−1.2
big	1.7	0.4	0.5	**−2.4**	−1.5	**2.4**	0.6	1.0	−1.1	−0.8

Values statistically different from zero in bold.

2.7.2 Cross-sectional

Derivation: $\widehat{\lambda}$ in generalized least squares procedure

Let $A = I + \beta'\Sigma^{-1}\beta\Sigma_f^{-1}$.

Hence,

$$\begin{aligned}
\widehat{\lambda} &= \left[\beta'\left(\beta\Sigma_f^{-1}\beta' + \Sigma\right)^{-1}\beta\right]^{-1} A^{-1}A\beta'\left(\beta\Sigma_f^{-1}\beta' + \Sigma\right)^{-1} E_T\left(R^e\right), \\
\widehat{\lambda} &= \left[A\beta'\left(\beta\Sigma_f^{-1}\beta' + \Sigma\right)^{-1}\beta\right]^{-1} A\beta'\left(\beta\Sigma_f^{-1}\beta' + \Sigma\right)^{-1} E_T\left(R^e\right).
\end{aligned}$$

Now

$$\begin{aligned}
A\beta' &= \left(I + \beta'\Sigma^{-1}\beta\Sigma_f^{-1}\right)\beta' \\
A\beta' &= \beta'\left(I + \Sigma^{-1}\beta\Sigma_f^{-1}\beta'\right) \\
A\beta' &= \beta'\Sigma^{-1}\left(\Sigma + \beta\Sigma_f^{-1}\beta'\right).
\end{aligned}$$

So, $\widehat{\lambda} = \left(\beta'\Sigma^{-1}\beta\right)^{-1}\beta'\Sigma^{-1}E_T\left(R^e\right)$.

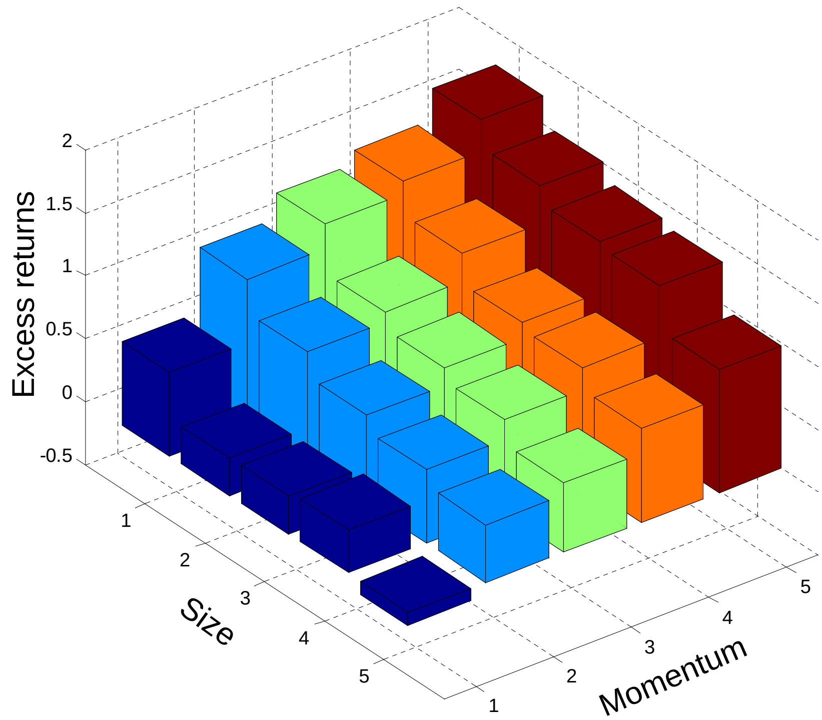

Figure 2.13: Average returns for 25 size-momentum portfolios.

Chapter 3

Evaluating alternative methods for testing asset pricing models with historical data

§ †

We follow the correct Jagannathan and Wang [58] framework for comparing the estimates and specification tests of the classical Beta and stochastic discount factor/generalized method of moments (SDF/GMM) methods. We extend previous studies by considering single and multifactor models, and by taking into account some of the prescriptions for improving empirical tests suggested by Lewellen, Nagel and Shanken [72]. For this purpose, we use a broad cross-section N and a multiple-length T of US test portfolios for the CAPM, Fama-French and an alternative three-factor model based on market, value and momentum called RUH. Our results reveal

§A paper-version of this chapter is currently being written jointly with Gonzalo Rubio (Universidad Cardenal Herrera CEU).

†An earlier version of this work was presented at the Brown Bag Seminar (University of the Basque Country, June 2007); XV AEFIN Finance Forum (Universitat de les Illes Balears, November 2007); and at the Advanced Finance Research Seminar II (Manchester Business School, March 2008). I would like to thank Rosa Rodríguez (Universidad Carlos III de Madrid) for helpful comments as discussant, and to an anonymous Journal of Empirical Finance referee. Martín Lozano gratefully acknowledges financial assistance from the Consolidate Research Team 9/UPV-00038.321-15094 of the University of the Basque Country (Euskal Herriko Unibertsitatea).

that SDF/GMM first-stage estimators lead to lower pricing errors than OLS, while SDF/GMM second-stage estimators display higher pricing errors than the classical Beta GLS method. While Jagannathan and Wang [58], and Cochrane [25] conclude that there are no differences when estimating and testing by the Beta and SDF/GMM methods for the CAPM, we show that their conclusion can not be extensible for multifactor models. The Beta method (OLS and GLS) seem to dominate the SDF/GMM (first and second-stage) procedure in terms of estimators' properties. These results are consistent across benchmark portfolios and sample periods.

3.1 Introduction

Explaining cross-sectional differences in asset expected returns is one of the great challenges of modern finance [1]. Although early empirical tests have largely found substantial empirical support for the traditional Capital Asset Pricing Model [2](CAPM) posterior well known papers by Shanken [91], MacKinley and Richardson [80], and Fama and French [29] among many others show that the usual proxies for the market portfolios are not mean-variance efficient.

The new evidence tends to find that not only the market but other aggregate risk factors seem to be important in describing the cross-sectional variation of average returns. In a very important contribution, Fama and French [30] introduce a three-factor model by adding a market capitalization (size) and a book-to-market (value) factor to the CAPM excess market factor return. Furthermore, Carhart [18] proposes

[1]See Daniel and Titman [26] for a discussion.

[2]For a detailed theoretical exposition of the CAPM see Sharpe [95] and Lintner [75]. See Black [8] for the zero-beta CAPM.

a four-factor model by appending the three Fama and French factors with a momentum factor after the study by Jegadeesh and Titman [59] on returns to momentum strategies[3].

In this paper, we argue that another plausible possibility is a three factor model combining market, momentum and value (RUH hereafter). In fact, this model outperforms other specifications in numerous tests on different samples and portfolios that we use in this paper. In any case, of course, the three models analyzed here (CAPM, Fama-French and RUH) all link excess stock returns to the returns of orthogonal factor-mimicking portfolios (or simply factors). Moreover, it should be recognized that the existing empirical studies that analyze the single-factor or multifactor models are far from decisive on the added value of specific multiple factors. In fact, Lewellen, Nagel and Shanken [72] provide an interesting empirical exercise showing how asset pricing tests are often highly misleading. They demonstrate that if the set of test assets has returns with a strong factor structure, like size or book-to-market sorted portfolios, almost any proposed factor weakly correlated with the Fama-French factors is likely to produce betas that line up with average returns generating a high cross-sectional R^2.

These risk-return models have been extensively tested in the finance literature by the regression based *traditional method* or Beta method, in which a cross-sectional regression model is proposed for average stock returns, and the theoretical implications

[3]The main difference among the various models lies in the way they determine the important factors. There are, broadly speaking, two main approaches to the issue of factor selection. Some models specify factors based on equilibrium arguments. The most important factor of this type is the return on the market portfolio, which is based on the Capital Asset Pricing Model first derived by Sharpe [95]. Other models specify factors based on economic intuition. Examples of such factors are term premium, default premium, the growth rate of industrial production, and inflation as suggested by Chen, Roll and Ross [21], and the size and book-to-market factors as proposed by Fama and French [29]. See Skoulakis [96] for a discussion

are tested as hypothesis on the parameters of the regression model.

However, it is well known that linear asset pricing models such as CAPM or Fama-French, and many others, including nonlinear specifications, can be unified in a stochastic discount factor (SDF) framework. This involves estimating the asset pricing model using its SDF representation and, in most cases, the generalized method of moments (GMM). The SDF method has become extremely popular in the recent finance literature. As Kan and Zhou [68] argue, although both the Beta and the SDF methods are used by many researchers in many different contexts, usually only one of them is used in a given application. It is therefore important to know which of the two methods may be better in some well defined statistical sense. In addition, as suggested by Jagannathan and Wang [58], the comparison can be so important that it might change the course of our empirical research on asset pricing models. For example, if the traditional method performs better in linear models, it is natural to speculate that it can also perform better in situations that involve nonlinear models. This is relevant because many nonlinear SDF cases are often linearized, like the famous papers by Campbell [16] or Cochrane [24] to cite just a few, and we could therefore study them by the traditional Beta method.

The comparison of the two methods (Beta and SDF) is not an easy matter even for linear models, since the parameters of interest are different under the two setups. The Beta method is formulated to analyze the factor risk premia, and these are the primary parameter of interest. In contrast, the SDF representation is formulated to analyze the parameters that enter into the imposed stochastic discount factor. The first formal comparison between the two methods is performed by Kan and Zhou [68]. They argue that the SDF is inferior to the traditional maximum likelihood approach, even in a simple test of the CAPM, as long as returns are identical and independent normally distributed random variables. Jagannathan and Wang [58], in a very influential paper, show that Kan and Zhou's conclusions in [68] are incorrect. These authors fail to explicitly incorporate the transformation between the risk premium parameters in the two methods, and they ignore the information about the mean and the variance of the factor while estimating the risk premium. Once this is done, Jagannathan and Wang [58] analytically show that the SDF method is as asymptotically efficient as the Beta method. Moreover, they also demonstrate that the SDF method has the same power as the Beta method.[4] Cochrane [25] also show that using 10 size-sorted portfolios, a given sample period and the simple CAPM case, the two methods produce basically the same standard errors, t-statistics, and statistics that the pricing errors are jointly zero.

The Jagannathan and Wang [58] CAPM empirical results are based on a set of simulations. In particular, they assume that the returns of 10 size-sorted portfolios and the market factor are drawn from a multivariate normal distribution, considering four different time horizons. In this paper, we follow the correct [58] framework for comparing the estimates and specification tests of the classical Beta and SDF methods, using historical data instead of simulations.

Furthermore, and contrary to Cochrane [25], we test not only the single factor model but also the Fama-French and the RUH models with a diverse number of test assets and time periods in order to address the tight factor structure problem advocated by Lewellen, Nagel and Shanken [72]. In fact, the contribution of this paper is the performance of a comprehensive extension of the analysis reported by

[4] Recent works such as Grauer and Janmaat [42] examine power tests for competing Beta pricing models.

Cochrane [25]. From our point of view this covers an important gap in the empirical asset pricing literature. The closest paper is probably due to Shanken and Zhou [94]. However, although the objective of the paper is similar, they report empirical results based only on simulations rather than on real data sets.

Specifically, we use six families of N test portfolios: 5 and 10 formed on ME (size); 5 and 10 formed on BE/ME (value); 6, 25 and 100 formed by the intersections of ME and BE/ME (Fama-French portfolios); 6 and 25 formed on ME and MOM (size and momentum); 5, 17 and 30 industry portfolios; and an extended test assets case in which we simultaneously combine the 25 FF portfolios and 17 industry portfolios. In this way, we can be confident that our results are not driven by the factor structure argument of Lewellen, Nagel and Shanken [72]. We also conduct our analysis using 6 time horizons T of US tests portfolios: 60, 120, 240, 360, 480 (all of them to cover the post-1963 data) and 948 monthly observations (the longest time-series used in this paper which goes from January 1927 to December 2005). The chosen time horizons are similar to those on related works such as Shanken and Zhou [94] and Grauer and Janmaat [42], just to mention a few.

Moreover, we calculate three kinds of the Beta model estimators: OLS, GLS and WLS. And five SDF estimators: first and second-stage returns on second moments $\mathrm{GMM_A}$, following Hansen and Jagannathan [49]; first and second-stage returns on covariances $\mathrm{GMM_B}$, following Cochrane [25]; and the continuous updating estimate $\mathrm{GMM_C}$ following Hansen, Heaton and Yaron [47]. We are therefore interested on evaluating how (and if) the Jagannathan and Wang [58] and Cochrane [25] results change in this richer framework.

Our results provide new evidence about finite-sample setups in which SDF/GMM formulation lead to almost the same results as the Beta method and also others in which there are significant discrepancies. These differences emerge even in linear models. Therefore, it may not be necessary to have a hard set up such as highly nonlinearity in order to anticipate differences between the two methods. In particular, our evidence reveals that SDF/GMM first-stage estimators lead to lower pricing errors than OLS, while SDF/GMM second-stage estimators display higher pricing errors than the classical Beta GLS method. Moreover, the Beta method (OLS and GLS) seem to dominate the SDF/GMM (first and second-stage) procedure in terms of estimators' properties. These results are consistent across benchmark portfolios and sample periods.

This paper is organized as follows. Section 3.2 briefly reviews the econometrics of estimating and evaluating asset pricing models. A full description of the data employed in the paper is presented in Section 3.3. Section 3.4 discusses the empirical results and a detailed analysis of different comparisons, while Section 3.5 concludes.

3.2 Description of the Beta and SDF/GMM methods

There is a large literature on econometric techniques to estimate and evaluate asset pricing models. As pointed out by Cochrane [25], each technique looks for answers on the same questions: how to estimate parameters, how to calculate standard errors of the estimated parameters, how to calculate standard errors of the pricing errors, and how to test the model. For a full description of the econometrics used in this paper, see sections 1.3 and 1.5. We now briefly describe the Beta and the SDF procedures.

3.2.1 The Beta method

We want to fit the following simple regression model $E\left(R^{ei}\right) = \beta_i'\lambda + \alpha_i$. As we explain in section 1.3, the idea is of course to learn why average returns vary across assets.

In this chapter, this is done by running an OLS, WLS and GLS cross-sectional regressions of average returns on the betas. Since betas are estimated in a time-series regression, we correct asymptotic standard errors by applying the Shanken [92] multiplicative correction (see section 1.3.3). We finally test whether all pricing errors are jointly zero with the asymptotic OLS, WLS and GLS test of pricing errors.

The GLS regression should give more precise estimates of the parameters and improve their efficiency. However, as Cochrane [25] points out, the variance-covariance matrix may be hard to estimate and invert when the cross-section N is large. This suggests that we may prefer the robustness of OLS over the (asymptotic) advantages of GLS. In any case, it is always true that the GLS regression pays more attention to the statistically more informative test assets, as we show on previous chapter.

3.2.2 The SDF method

The use of the SDF method for econometric evaluation of asset pricing models has become common in the recent empirical literature. As we explain in section 1.5, the first order pricing equation from the intertemporal optimization of the representative agent can be written as

$$p_t = E_t(m_{t+1}x_{t+1}),$$
$$m_{t+1} = f(\text{data, parameters}),$$

where p_t is the price of any stock, m_{t+1} is the SDF which is the intertemporal marginal rate of substitution of consumption, x_{t+1} is the future payoff of the stock and E is the conditional expectation operator. An asset pricing model identifies a particular SDF (a proxy for the marginal rate of substitution of aggregate consumption) that is a function of observable variables and the model parameters. The SDF method involves estimating the asset pricing model using its SDF representation and the GMM procedure.

The development of the GMM by Hansen [46] has had a major impact on empirical research in finance because it allows for conditional heteroskedasticity, serial correlation and non-normal distributions. See Jagannathan, Skoulakis and Wang [54] for an excellent review of GMM methodology in finance. In this section 3.2.2, we review the estimation and testing of linear discount factor models expressed as,

$$p = E(mx),$$
$$m = b'f,$$

This pricing expressions lead naturally to the GMM when testing asset pricing models, where the pricing errors are precisely the moments used in the estimation.

First and second-stage GMM estimators

The idea is to choose b to make the pricing errors $g_T(b)$ as small as possible, by minimizing the quadratic form in equation 1.5.1.

When imposing $W = I$, GMM treats all test assets symmetrically, and we just minimize the sum of squared pricing errors. The result of making such simplification is what we call first-stage estimators. This estimator is consistent and asymptotically normal.

The second-stage estimate makes a formal statistical choice of the weighting matrix W. Since returns are correlated, the usual procedure chooses the variance-covariance matrix, so that the matrix pays more attention to linear combinations of moments for which the available data is more informative. Hansen [46] shows formally that the choice $W = S^{-1}$, where $S \equiv \sum_{j=-\infty}^{\infty} E\left(u_t u'_{t-j}\right)$, is the statistically optimal weighting matrix, meaning that it produces estimates with lowest asymptotic variance.

Hansen and Jagannathan [49]: $\mathbf{GMM_A}$ estimators

Another example of prespecified economically interesting weighting matrix is the second moment matrix of returns and factors, advocated by Hansen and Jagannathan [49]. They also introduce the Hansen-Jagannathan distance, which measure specification errors of SDF models by least squares distances between an SDF model and the set of admissible SDFs that can correctly price a set of test assets. In recent works such as Li, Xu and Zhang [73] the Hansen-Jagannathan distance is used in order to evaluate asset pricing models.

Hence, for this subsection we will refer to S as the second moment matrix $E\left(xx'\right) =$

$cov\,(x) + E\,(x)\,E\,(x)'$. The complete derivation of $\mathrm{GMM_A}$ estimates are in section 1.5.2, in particular in equation 1.5.2, and their respective standard errors are represented in equation 1.5.3.

The $\mathrm{GMM_A}$ estimate is a cross-sectional regression of mean excess return on second moment matrix with factors. The model test is a quadratic form in the vector of pricing errors, see equation 1.5.5. Note that there are two ways to get a small value of the J test statistic. First and desirable, we can generate small pricing errors with a high degree of precision or, and this is not desirable, we can generate large pricing errors with even higher standard errors of those errors. Thus, in this paper we would care not only on specifications test results but also on the pricing errors in order to avoid this trap.

Cochrane [25]: $\mathbf{GMM_B}$ estimators

Alternatively, we can run a cross-sectional regression of mean excess returns on covariances by choosing the normalization $a = 1 + b'E\,(f)$ rather than $a = 1$. Then, the model is $m = 1 - b'\,[f - E\,(f)]$ with mean $E\,(m) = 1$. The pricing errors are

$$g_T(b) = E_T\,(mR^e) = E_T\,(R^e) - E_T(R^e\widetilde{f'})b,$$

where we denote $\widetilde{f'} \equiv f - E\,(f)$. We have

$$d = \tfrac{\partial g_T(b)}{\partial b'} = E(R^e\widetilde{f'}),$$

which now denotes the covariance matrix of returns and factors. We must bear in mind that the mean of the factor is estimated in $\mathrm{GMM_B}$ (as well as b), and the distribution theory should recognize sampling variation induced by this fact as we usually do in the cross-sectional regressions. Second-stage estimators comes from the minimization of equation 1.5.8.

Hansen, Heaton and Yaron [47]: $\mathbf{GMM_C}$ (continuous updating) estimators

Another possibility is estimating the spectral density matrix or, in other words, use the optimal weighting matrix instead of taking the prespecified weighting matrix on the second-stage estimators, as we advocate in $\mathrm{GMM_A}$ and $\mathrm{GMM_B}$. The iterated $\mathrm{GMM_C}$ estimator using the optimal weighting matrix may present two related problems as we empirically show on the previous chapter. First, if the variance-covariance matrix for the iterated $\mathrm{GMM_C}$ estimator is poorly measured, then the estimator will put too much weight on moments that spuriously appear to be measured precisely. Moreover, the iterated estimator may place too much weight on test assets that are economically uninteresting, in the sense of being extreme short and long positions in some of the stocks.

Furthermore, the fact that the S matrix changes with the model, may improve the J_T statistic because it blows up the estimate of S, rather than by lowering the pricing errors. As Cochrane [25] emphasizes we should not compare formally J_T tests across models. This is one of the reasons why it is recommended to use a common weighting matrix for comparing models like those discussed above. There are several alternatives to the second-stage procedure. We will use the continuous updating estimator which states that it is not true that S must be held fixed as one searches for b. Instead, one can use a new $S(b)$ for each value of b, and estimate b by using equation 1.5.9. The estimates produced by this simultaneous search will not be numerically the same in a finite sample as the two-step or iterated estimates.

To wrap up, we will have three econometric specifications in the Beta method: OLS, GLS and WLS. And five in the SDF/GMM method: First and second stages of GMM_A, GMM_B, and the continuous updating GMM_C. When estimating, we collect the central parameter (λ for Beta method and b for SDF method), standard errors and bias from the factor mean. On the other hand, when testing, we collect the pricing error and the p-value of the model specification test.

3.3 Data

Three single (size, value, and industry), two double-sorted (size-value and size-momentum), and one combined (size-value plus industry) test portfolios are taken from the data library of Kenneth French because of familiarity and availability to the general readership.

In sum, we take five types of N test portfolios: 5 and 10 formed on ME; 5 and 10 formed on BE/ME; 6, 25 and 100 formed by the intersections of ME and BE/ME (Fama-French portfolios); 6 and 25 formed on ME and MOM; and 5, 17 and 30 industry portfolios. Lewellen, Nagel and Shanken [72] suggest that one could expand the set of test portfolios to price all of them at the same time. In this paper, besides the previous five types of test assets, we use an extended set formed by 25 Fama-French portfolios plus 17 industry portfolios, resulting in a total of 42 test portfolios. Note that we take at least two different values of N within each test assets in order to provide the robustness checks. Other recent works such as Li, Xu and Zhang [73] also recommend considering other portfolio formations, since models may tend to have small pricing errors for the traditional test assets by construction.

From our point of view this wide set of test portfolios offers new insights in the empirical asset pricing literature since, as pointed out by Lewellen, Nagel and Shanken [72], empirical tests frequently focus on certain assets as the 10 size-sorted or 25 Fama-French portfolios. In fact, to the best of our knowledge, the closest empirical study which use a wide set of test portfolios is Shanken and Zhou [94] although their objective was somewhat different as ours.

Aggregate risk factors are also taken from Kenneth French homepage. In the CAPM, the only relevant factor is the excess market return. The Fama-French model introduces two additional factors, the *small minus big* (SMB) portfolio and the *high minus low* (HML) portfolio. The RUH model, introduced in the previous chapter, contains the excess market return, the *up minus down* (UMD) portfolio and the HML factor. These additional factors are intended to capture common non-market risk factors that are related to size, value and momentum.

We conduct our analysis using six values for the time length monthly observations T: 60 (January 2001-December 2005), 240 (January 1986-December 2005); 360 (January 1976-December 2005); 480 (January 1966-December 2005) and 948 (January 1927-December 2005). The choice for a monthly interval reflects the trade-off between the sampling error of a sufficiently large sample, and a realistic evaluation horizon. Increasing the return interval (e.g. yearly) would lead to a small data set, while decreasing it (e.g. daily) to an unrealistically short evaluation horizon. Further, the use of high-frequency data introduces well known microstructure problems which may distort the empirical results. Therefore, we adhere to the common approach of using monthly returns.

Taking into account the three models, thirteen test portfolios, six time periods and eight econometric specifications, we end up with 1636 and 2730 observations of λ and

b estimators with their corresponding standard errors and bias from the factor mean. On the other hand, we have 702 Beta and 1170 SDF observations of pricing errors with their corresponding p-values of the model specification test J_T[5]. This amount of results is a considerable expansion to similar previous works like Cochrane [25] or Shanken and Zhou [94], and thus we are clearly able to broaden the comparisons between the two methods.

It is worthwhile to emphasize that our results comes from historical data while results on Jagannathan and Wang [58], and Kan and Zhou [66, 68] are based on simulations, and those were calibrated using the 10 size-sorted portfolios.

3.4 Empirical results

In this section, we first analyze the simple and well documented case of the evaluation of the CAPM, using monthly data from January 1927 to December 2005, and ten size-sorted portfolios for the US. Then, we move towards a more complex and interesting setup in which we first analyze specification tests and then the estimators' properties.

3.4.1 A classical and simple setup: CAPM with size-sorted portfolios

In this initial testing we follow the framework presented by Cochrane [25]. Indeed, we will show that the SDF estimates, standard errors, and χ^2 statistics are very close to time-series and cross-sectional regression estimates. It is important to carry out this

[5]Note that the number of outcomes from the Beta method is always less than from the SDF procedure. This is because we have three specifications for Beta, and five for the SDF. However, most of the comparisons conducted are based on similar numbers of observations.

analysis to better understand further and more complex settings in which differences among methods will come out.

We argue that the results of this section 3.4.1 may lead to conclude that both methodologies are quite similar. There are numerous examples in which researchers refer to these similarities based on works such as Jagannathan and Wang [58]. For example, see Wang and Zhang [102], Jagannathan, Skoulakis and Wang [55], Vassalou, Li and Xing [100], Cochrane [23, 25], Smith and Wickens [97], Nieto and Rodríguez [85], Balvers and Huang [5], Brandt and Chapman [10], Cai and Hong [15], and Ferson [35], just to mention a few.

However, such similarities are driven by using a set of test portfolios with low dispersion. As far as the right hand variables have low standard deviation, virtually any statistical method would lead to similar results when trying to explain them. This is well explained in Lewellen, Nagel, and Shanken [72], they show that it is difficult to differentiate models that have been developed to explain the cross-sectional returns of the 25 size and book-to-market portfolios using traditional methods, because these models tend to have small pricing errors for the test assets by construction.

The time-series approach implies that all alpha estimates should be zero in equation 1.2.1. The time-series framework estimates the factor risk premium from the sample mean of the factor ignoring any information of the other assets. In other words, this specification sends the expected return-beta line through the market return. The OLS cross-sectional regression from the Beta method minimizes the sum of squared pricing errors. This implies that there is the possibility of some market pricing error in order to obtain a better fit in other test assets. The GLS cross-sectional regression from the Beta method weights pricing errors by the residual variance-covariance matrix. Thus, when the market factor is a return and is included in the test portfolios, it turns out that the GLS procedure reduces to the time-series regression. On the other hand, if the market portfolio is not one of the test assets, as in most empirical analysis, the GLS cross-sectional regression is not identical to the time-series regression.

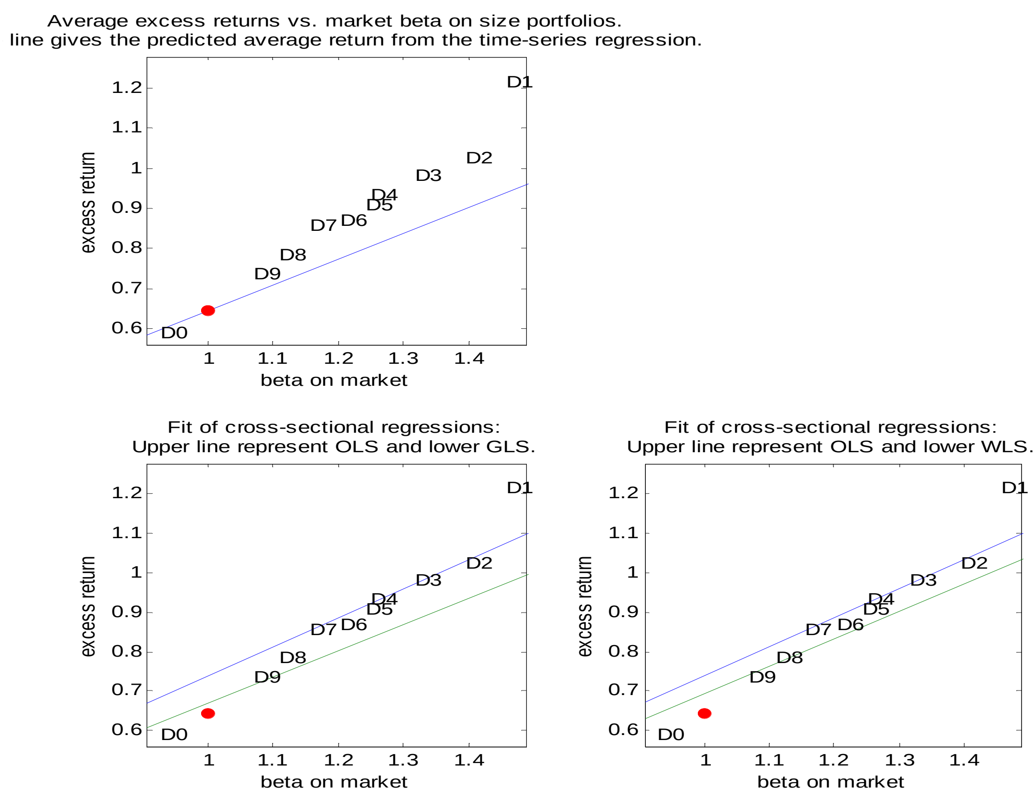

Figure 3.1: CAPM in size portfolios, time-series and Beta method comparison.

Figure 3.1 displays the results from the time-series regression, and the OLS cross-sectional regression compared to a full GLS regression and a WLS case. In none of these cases, the market return (red circle) is taken as one of the test portfolios, so the time-series and GLS regressions are not identical. D1 contains the smallest firms, while D0 is formed from the largest stocks. The positive D1 pricing error is the size effect anomaly.

Figure 3.1 illustrates the difference between the time-series and cross-sectional regressions in evaluating the CAPM on monthly size-sorted portfolios. Since the time-series approach estimates the factor risk premium from the sample average of the factor, the regression draws the expected return-beta line across assets by making it fit precisely on two points, the market return (red circle) and the risk-free rate. On the other hand, in the cross-sectional regressions, the market portfolio is priced with error to reduce the pricing errors of other test assets.

It is interesting to point out that the GLS is practically the same as the time-series case. This makes sense, since the GLS cross-sectional procedure, when the market portfolio is included as a test asset, pays especial attention to fit the market line since the market return has no residual variance. Recall that the GLS cross-sectional expressions weight the various test portfolios by the inverse of the variance-covariance matrix of residuals. As long as the size-sorted portfolios span the market portfolio, the GLS cross-sectional regression and the time-series approach will generate very similar results. Finally, as expected given this reasoning, the WLS regression does not generate results as close as the time-series methodology as the ones found under the GLS procedure.

Figure 3.2 illustrates the $\mathrm{GMM_A}$ estimates with the same data as in Figure 3.1. The horizontal axis is the second moment of returns and factors rather than beta, and the vertical axis is the excess return as in Figure 3.1. The expressions used to

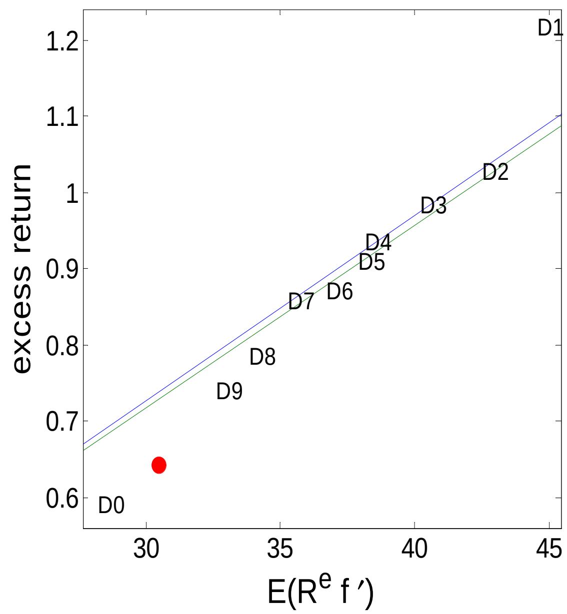

Figure 3.2: CAPM in size portfolios, $\mathrm{GMM_A}$ estimate.

obtain the results come from section 1.5.2.

The first-stage estimate (Figure 3.2, upper line, $\mathrm{GMM}^1_{\mathrm{A}}$) is an OLS cross-sectional regression of average returns on second moments. As expected, it generates pricing errors basically equal to those of the cross-sectional OLS in Figure 3.1. In particular, the OLS pricing error is 0.057 with a p-value of 0.92, while the $\mathrm{GMM}^1_{\mathrm{A}}$ error is 0.056 with a p-value of 0.89. On the other hand, the second-stage estimate (lower line, $\mathrm{GMM}^2_{\mathrm{A}}$) minimizes pricing errors weighted by the spectral density matrix which is of course different from the variance-covariance matrix of residuals. This explains why the line is quite far away from the market portfolio. In any case, both approaches generate very similar results with large pricing for the smallest and largest firms.

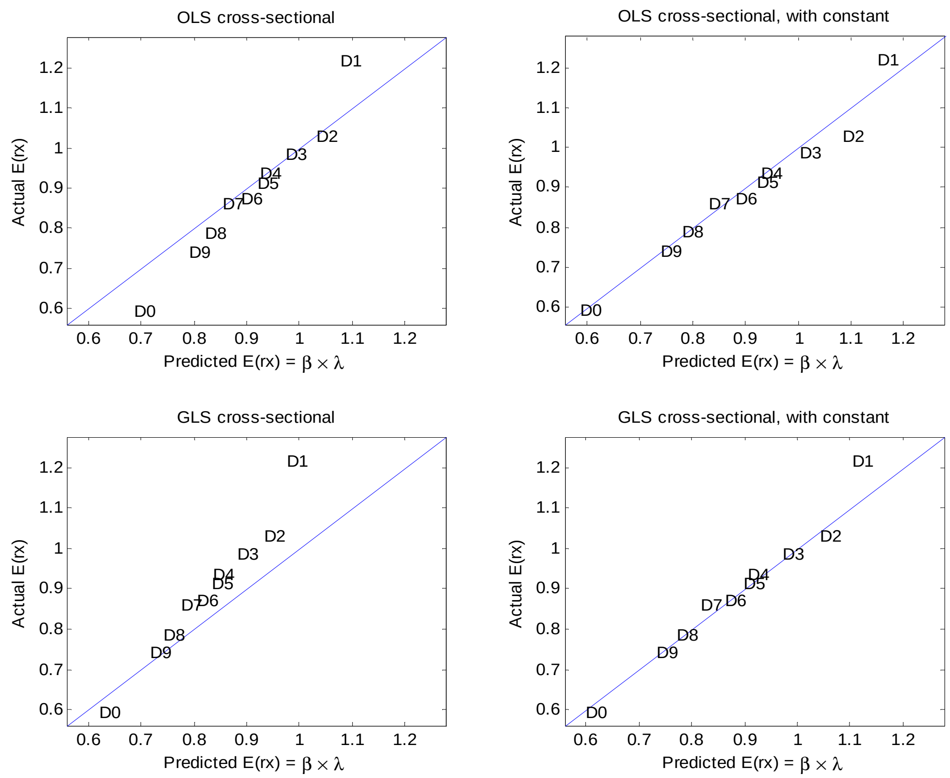

Figure 3.3: CAPM in size portfolios. Actual versus predicted returns. Beta method.

Figures 3.3 and 3.4 show the scatter plots of actual versus predicted expected returns. The vertical distance from a point to the sloped line reveals the residual; this is to say, the difference of the actual and the predicted value. Note that in both figures, the CAPM does a relatively good job at explaining the 0.6 per cent monthly return spread of the size-sorted portfolios. This was expected since the pricing errors of both methods are very similar.

Figure 3.3 makes it clear that a free constant (right panels) allows the model to explain better the cross-sectional variation of expected returns, even for the smallest and biggest portfolios. However, they obtain a considerable alpha, leading to economically misleading estimates. Thus, the relevant plots are the ones without constant (left panels).

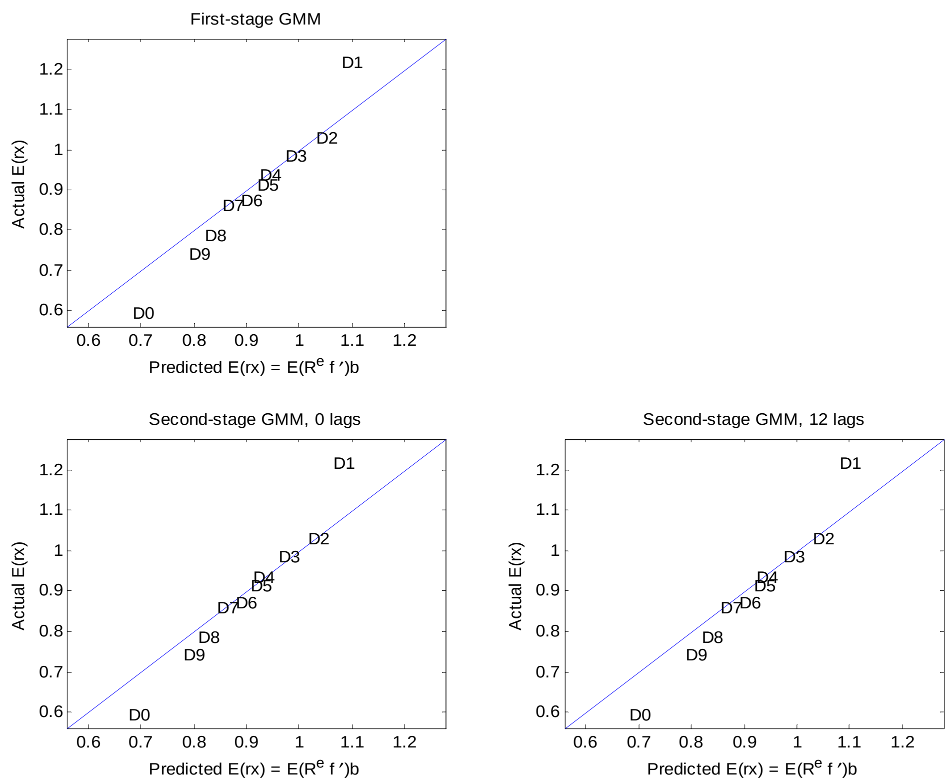

Figure 3.4: CAPM in size portfolios. Actual versus predicted returns. $\mathrm{GMM_A}$ method.

Figure 3.3 and 3.4 also illustrate the performance of GLS and second-stage $\mathrm{GMM}^2_{\mathrm{A}}$ methods. Their pricing errors are 0.103 and 0.058 respectively; again in this case, the

SDF leads to lower pricing errors than the Beta method. Nevertheless, the p-values are the same as the p-values from OLS and first-stage $\text{GMM}_{\text{A}}^{1}$ methods because the higher pricing errors (0.103 and 0.058 vs. 0.057 and 0.056) are compensated with the weighting matrix Σ^{-1} and S^{-1} in the test statistic.

In general, as in Cochrane [25], these results about pricing errors and p-values do not suggest any strong reason to prefer any particular method. The high and similar p-values of the Beta (0.92) and SDF (0.89) methods imply that pricing errors are statistically equal to zero in all cases. Furthermore, OLS, and GMM_{A} (first and second-stages), have strikingly similar pricing errors, all of them between 0.056 and 0.058. The GLS method lead to twice as big pricing errors than the other methods, nevertheless GLS is not intended to minimize the pricing errors as OLS does.

We now turn from the pricing error and specification tests to the estimators' comparison, following what Jagannathan and Wang [58] called the correct framework for making comparisons between alternative estimates. They argue that b estimates are not directly comparable to the risk premium estimates, thus we have to calculate:

$$b = \frac{E\left(R^{em}\right)}{E\left(R^{em2}\right)}.$$

According to our dataset, we have $E\left(R^{em}\right) = 0.6434\%$ and $\sigma\left(R^{em}\right) = 5.4816\%$, so $100 \times b = 100(0.6434)/\left(0.6434^2 + 5.4816^2\right) = 2.1121$. In the original work of Cochrane [25] this value is 2.17, and In Jagannathan and Wang [58], they set larger risk premiums. So, the market risk premium in the Beta method is $\lambda = E_T\left(R^{em}\right) = 0.64$ and $b = 2.11$ for the GMM_{A} method. Their corresponding OLS, GLS, first- and second-stage GMM_{A} estimators are $\widehat{\lambda}_{\text{OLS}} = \underset{15\%(0.19)}{0.74^{***}}$, $\widehat{\lambda}_{\text{GLS}} = \underset{4\%(0.18)}{0.67^{***}}$, $\widehat{b}_1 = \underset{15\%(0.59)}{2.43^{***}}$, $\widehat{b}_2 = \underset{13\%(0.58)}{2.39^{***}}$, where the percent values in brackets represent biases relative to the realized risk premium, while in parenthesis we report their corresponding standard errors. Note that, as in Jagannathan and Wang [58], we do not set the values of $E\left(R^{em}\right)$ and λ to be equal, therefore, we do not impose the restriction that the factor is the return on a portfolio of tradable assets.

These parameters are the slopes of the lines in the Figures 3.1 and 3.2, the market price of risk λ in the Beta model and the relationship between mean returns and second moments b in the SDF method. By including their percentage biases we aim to follow Lewellen, Nagel and Shanken [72] recommendation about *taking the magnitude of the cross-sectional slopes seriously.* Thus, we are not only interested on evaluating the model's goodness of fit, but also on whether the estimated slopes are reasonable close to $\lambda = E_T\left(R^{em}\right) = 0.64$ and $b = 2.11$. It is worthwhile to highlight that in this case the GLS and the second-stage do improve efficiency by giving more precise estimates (measured by the percentage bias relative to the risk premium): from 15% to 4% in the case of Beta method and from 15% to 13% in the case of SDF.

Note that this simple and classic setup is associated with the fact that Σ and S are not hard to estimate and invert. Thus, the methods are expected to do very well, as in fact do. Unfortunately, however, we will show in the next subsection 3.4.2 that this well behaved result is extremely sensible to changes in time length T and/or cross-section N. If we use the double-sorted portfolios (e.g. 25 size-momentum), instead of 10 size portfolios, the 0.6% monthly return spread turns out to be three times higher. In this alternative setting the CAPM is easily rejected.

A fair comparison within methods of the last four estimates should be $\hat{\lambda}_{OLS}$ versus $\hat{b}_1$ (both estimates are weighted by I) and $\hat{\lambda}_{GLS}$ versus $\hat{b}_2$ (weighting by Σ^{-1} is not the same as weighting by S^{-1}, though it is sufficiently close). Looking at the percentage biases, we do not see large differences; they all go from 4% to 15%. On the other hand, standard errors are somewhat higher for the SDF method, even though, this difference is not significant in the sense that the estimates are still statistically different from zero at the highest significance level (by looking at the t-statistics). In general, the numerical values of b are higher than λ. Then, standard errors will also be higher; this is the reason of stressing the importance of not only comparing the standard errors within methods, but also to take into account the t-statistics.

In sum, the results from testing and estimating the CAPM in this classic setup, illustrate that the differences between the Beta and the SDF method are almost irrelevant. However, the key point is that we will show how the differences become significant in a more complex setup.

3.4.2 A full comparison: pricing errors and specification tests

Let us begin comparing the pricing errors of Beta (OLS and GLS) versus GMM_A first and second-stage methods for CAPM, Fama-French and RUH models and the test portfolios described on section 3.3; that is 360 pricing error observations for each method. We show each estimation on the appendix, specifically in sections 3.6.8, 3.6.9 and 3.6.10 for CAPM, Fama-French and RUH models respectively. Each of these sections contains several panels which correspond to the different test portfolios, in this section 3.4.2 we will show frequency tables which summarize the results.

Here, we are concerned with two key issues. The first one regarding which method leads to lower pricing errors, while the second one analyzes how well does that method performs in testing the three models. Our findings reveal that Hansen and Jagganathan [49] first-stage $\mathrm{GMM_A}$ produce the lowest pricing errors; and that multi-factor models, particularly the RUH specification is more likely to successfully price the test portfolios returns.

Which method leads to lower pricing errors?

By construction, we know that OLS and first-stage GMM should lead to smaller pricing errors than GLS and second-stage GMM respectively. We also know that OLS and first-stage GMM are in general more robust but have less asymptotical statistical advantages than GLS and second-stage GMM. So, in this subsection we are interested in analyzing which method leads to lower pricing errors by confronting first OLS and first-stage $\mathrm{GMM_A^1}$ and, secondly, by making comparisons between GLS and the second-stage $\mathrm{GMM_A^2}$.

Regarding the first issue, Shanken and Zhou [94] show that first-stage does it better in Beta representation models. However, they do not analyze SDF models. Furthermore, in Shanken and Zhou [94] there is no conclusive answer for the second stage estimators in this sense. On the other hand, there are other type of works which focus exclusively on SDF models such as Farnsworth, Ferson, Jackson and Todd [33], who find that measures of performance are not highly sensitive to the SDF representation. Thus, we consider that analyzing which method leads to lower pricing errors by confronting first OLS and first-stage $\mathrm{GMM_A^1}$ represent an issue that has been not fully explored yet. Our results show that, in terms of achieving lower pricing errors, first-stage $\mathrm{GMM_A^1}$ does it better than OLS, and GLS outperform $\mathrm{GMM_A^2}$.

Table 3.1 summarize the results of comparing Beta versus GMM/SDF pricing

Table 3.1: Comparison of pricing errors

The numbers denote the frequency in which Beta method (OLS or GLS) leads to higher pricing errors compared with SDF/GMMA method (first and second stage). For example, 0.83 means that 83 percent of the times, the Beta method present a higher pricing error relative to the SDF/GMMA method. We employ four time-periods, T=240, 360, 480, and 948. Size includes 5 and 10 size-sorted portfolios; Value has 5 and 10 BE/ME-sorted assets; Size and Value are the 6, 25 and 100 Fama-French portfolios; Size and Momentum includes 6 and 25 portfolios formed on ME and MOM; Industry incorporates 5, 17 and 30 industry-sorted assets, and Size, Value and Industry has the 25 Fama-French extended with 17 industry portfolios. Pricing errors are those from the estimation of CAPM, Fama-French and RUH models.

Portfolio Classification	OLS vs. $\mathrm{GMM}^1_{\mathrm{A}}$	GLS vs. $\mathrm{GMM}^2_{\mathrm{A}}$
Size (ME)	0.83	0.33
Value (BE/ME)	0.71	0.42
Size and Value (ME & BE/ME)	1.00	0.33
Size and Momentum (ME & MOM)	0.83	0.21
Industry	0.89	0.17
Size, Value and Industry	1.00	0.11

errors, using four time periods from T=240 to T=948. Our results suggest that first-stage $\mathrm{GMM}^1_{\mathrm{A}}$ dominates OLS, while GLS dominates second-stage $\mathrm{GMM}^2_{\mathrm{A}}$ at minimizing pricing errors, and this result is consistent across test portfolios and sample periods.[6]

The OLS and first-stage $\mathrm{GMM}^1_{\mathrm{A}}$ estimators are intended to minimize the root mean square errors since there is no weighting matrix, but first-stage estimators do it better than OLS. This result is consistent with the findings reported by Shanken and Zhou [94] on Beta models. Second column in Table 3.1 tells us a different story. In

[6] The tables used for constructing Table 3.1 are in sections 3.6.8, 3.6.9 and 3.6.10 for CAPM, Fama-French and RUH models respectively.

this case, Beta method does it better in achieving lower pricing errors than second-stage GMM^2_{A} estimators, even though these methods do not have the pure objective of minimizing the sum of square errors, as first-stage GMM^1_{A} and OLS do.

Testing the alternative pricing model specifications

According to Table 3.1, first-stage GMM^1_{A} does good job at minimizing squared pricing errors. Hence, we can now look at their corresponding test statistics with more confidence. A new question that arises is how different are the specification tests results from first-stage GMM^1_{A} when applied to three models, four values of T, and six families of test portfolios N. In order to summarize the results, let us focus on the p-values and group them in quartiles by test portfolios and by models in Table 3.2.

Our test results are summarized in Table 3.2. Each panel represents a particular asset pricing model: CAPM (upper), Fama-French (middle) and RUH (lower). Columns from left to right in each panel are families of N tests portfolios formed by ME, BE/ME, ME & BE/ME, Industry, and the extended set of ME & BE/ME (or Fama-French portfolios) plus Industry respectively. It should be recalled that each column is formed by at least two different values of N and four time-lengths. Rows are the probability intervals of not rejecting the null. Therefore, a column with larger proportion of the last interval 76-100 implies that the null will not be rejected in most cases for that model and test portfolio.

Section 3.4.1 above shows that, under an ad-hoc setup, we were unable of rejecting the CAPM. This result is certainly not robust as we already know from many other existing studies. Indeed, Table 3.2 shows the CAPM is rejected for all test portfolios

Table 3.2: Probability intervals of not rejecting the null hypothesis $\mathrm{GMM}_{\mathrm{A}}^{1}$, by tests assets.

The numbers denote the frequency in which each p-value falls into each probability interval. The probability intervals are given in the first column for the three alternative panels which correspond to the CAPM, the three-factor Fama-French model, and the three-factor model with the excess market return, the momentum factor and the value factor (RUH). High frequencies in the first interval (0-25) and low frequencies in the last interval (76-100) means that the model has a bad performance in terms of pricing errors obtained under the $\mathrm{GMM}_{\mathrm{A}}^{1}$ method. We employ four time-periods, $T = 240$, 360, 480, and 948. Size includes 5 and 10 size-sorted portfolios; Value has 5 and 10 BE/ME-sorted assets; Size and Value are the 6, 25 and 100 Fama-French portfolios; Size and Momentum includes 6 and 25 portfolios formed on ME and MOM; Industry incorporates 5, 17 and 30 industry-sorted assets, and Size, Value and Industry has the 25 Fama-French extended with 17 industry portfolios.

1 - p-value	ME	BE/ME	ME&BE/ME	ME&MOM	Industry	25 ME&BE/ME + 17 Industry
			Panel A: CAPM			
0-25	0.42	0.42	1.00	1.00	0.25	1.00
26-50	0.17	0.25	0.00	0.00	0.42	0.00
51-75	0.25	0.17	0.00	0.00	0.25	0.00
76-100	0.17	0.17	0.00	0.00	0.08	0.00
			Panel B: Fama-French			
0-25	0.33	0.08	1.00	1.00	0.08	1.00
26-50	0.25	0.00	0.00	0.00	0.42	0.00
51-75	0.17	0.25	0.00	0.00	0.25	0.00
76-100	0.25	0.67	0.00	0.00	0.25	0.00
			Panel C: RUH			
0-25	0.08	0.00	0.50	1.00	0.00	1.00
26-50	0.00	0.08	0.08	0.00	0.17	0.00
51-75	0.25	0.17	0.17	0.00	0.42	0.00
76-100	0.67	0.75	0.25	0.00	0.42	0.00

more often than any other competing model. In particular, for the ME & BE/ME (or Fama-French portfolios), the size and momentum, and the combined Fama-French portfolios plus Industry, the rejection of the CAPM is absolute in the sense that all their tests have p-values between 0 and 25 percent. It should now be clear that the results on section 3.4.1 are not representative at all, because they are only 17% of our observations when testing the single-factor model with size-sorted portfolios.

The Fama-French model is originally suggested to explain the pricing errors of the CAPM. It is therefore interesting to compare this model with the CAPM. In particular, Table 3.2 provides evidence favorable to multifactor models since the probability of not rejecting the null substantially improves in the Fama-French and RUH specifications. Most notably, the number of cases in which the probability that the pricing errors are zero on the highest p-value interval using the BE/ME portfolios goes from 17% for the CAPM to 67% and 75% for the Fama-French and RUH models. As in the CAPM, the Fama-French model is not able to price the same families of portfolios. Although the pricing errors are in fact lower, they are not enough for getting at least a portion of the 26-50 interval. Only the RUH specification is capable of successfully pricing the Fama-French portfolios in 25% of the cases, and also reduces to 50% the cases in which the model can not price these portfolios at all.

Generally speaking, this evidence suggests that multiple factors (as the Fama-French model), and more precisely multiple and adequate factors (as the RUH specification) help rather than hurt for all these test portfolios' valuations. The RUH model performance is particular remarkable since there is no track of this specification in previous literature. The probability of not rejecting the null is higher for every test portfolio and model tested under the RUH specification.

Table 3.2 also suggest that the specification testing results depend on the test portfolios employed. It seems that portfolios' characteristics are driving the rejection of the null hypothesis. In particular, the dispersion of average returns across portfolios seems to be positively correlated with the pricing success of any given model. Returns on portfolios formed by ME has a cross-sectional standard deviation of 12%, 16% for BE/ME and Industry portfolios, 22% for the combined Fama-French plus Industry, 27% for ME & BE/ME, and 45% for ME & MOM. It turns out that the hardest test portfolio to value is precisely the ones with higher dispersion. It is also true that double-sorted and combined portfolios are harder to price because they have significant higher dispersion than single-sorted portfolios.

The industry classification is a special case because is not motivated by known patterns in historical return series. Additionally, these portfolios are independent of financial ratios such as BE/ME. Hence, it is surprising to observe that their results are not particularly different from the other test portfolios. One may argue that their relatively low dispersion on average returns explain these similarities.

One may always think that the results reported in Table 3.2 are being drive by some particular characteristics of the historical data on a given time-period. Thus, we perform a robustness check by changing the perspective view over the same results.

In Table 3.3, we now control for the sample period rather than controlling for the test portfolios as we did on Table 3.2. Then, the columns in Table 3.3 now represent time periods (instead of portfolios), and panels and intervals are exactly the same as in Table 3.2. Recall the full results are on sections 3.6.8, 3.6.9 and 3.6.10 for CAPM, Fama-French and RUH models respectively.

This strategy allows us to analyze whether the evidence in favor of multifactor

Table 3.3: Probability intervals of not rejecting the null hypothesis $\mathrm{GMM}_{\mathrm{A}}^{1}$, by time length.

The numbers denote the frequency in which each p-value falls into each probability interval. The probability intervals are given in the first column for the four alternative time periods (T=240, 360, 480, and 948) and the three panels which correspond to the CAPM, the three-factor Fama-French model, and the three-factor model with the excess market return, the momentum factor and the value factor (RUH). High frequencies in the first interval (0-25) and low frequencies in the last interval (76-100) means that the model has a bad performance in terms of pricing errors obtained under the $\mathrm{GMM}_{\mathrm{A}}^{1}$ method.

1 - p-value	T = 240	T =360	T = 480	T = 948
Panel A: CAPM				
0-25	0.44	0.56	0.56	0.56
26-50	0.33	0.22	0.11	0.22
51-75	0.11	0.22	0.33	0.00
76-100	0.11	0.00	0.00	0.22
Panel B: Fama-French				
0-25	0.44	0.33	0.33	0.44
26-50	0.22	0.11	0.11	0.11
51-75	0.00	0.11	0.22	0.33
76-100	0.33	0.44	0.33	0.11
Panel C: RUH				
0-25	0.22	0.22	0.11	0.44
26-50	0.00	0.00	0.11	0.11
51-75	0.22	0.22	0.33	0.11
76-100	0.56	0.56	0.44	0.33

models reported before is independent of the time length. In this regard, it should be pointed out that some studies suggest that the choice of the sample period and size segment are important for judging the empirical validity of the CAPM.

For example, Loughran [77] shows that a substantial portion of the BE/ME effect is driven by the low returns of small newly-listed growth stocks. Further, Ang and Chen [3] estimate conditional factor models and show that the BE/ME effect disappears in the pre-1963 period. However, these studies do not consider the combined effect of the benchmark set and the sample period as we do. Also, these studies focus on the CAPM and do not directly address the issue of the added value of multiple factors. Contrary to these studies, our evidence suggest that the choice of the sample period is not relevant for evaluating the empirical validity of the CAPM and multifactor models, at least in tests of overidentifying restrictions. Even though, we find that the pricing errors diminish for larger time horizons. Our results are also consistent with those on Shanken and Zhou [94].

In general, Tables 3.2 and 3.3 show that RUH model outperforms Fama-French and CAPM models. Again, we can see that adding momentum and value to the CAPM does help explaining the cross-sectional variation of asset returns, and it does for every T value at different magnitudes. On the other hand, there is no clear pattern between the length of the sample period and the rejection of the null hypothesis. While for Fama-French, and especially for RUH, the possibility of rejecting the model is lower when T is different from 948, the opposite seems to hold true for the CAPM. To conclude, multiple factors improve the empirical fit relative to the single-factor CAPM for all test portfolios and sample periods.

3.4.3 A full comparison: estimators' properties

We now turn from tests and pricing errors to evaluate the estimators' properties. Remember we are dealing with historical data. We understand that an estimator has desirable properties if it has the following three conditions (*i*) it has low standard error, (*ii*) it is statistically different from zero, and (*iii*) it has low bias (measured as the percentage error) relative to the observable factor. In this subsection, tables summarize the results on sections 3.6.1 to 3.6.7.

We are now concerned with the following two questions: First, which method leads to better estimators' properties within methods (OLS versus GLS, and first versus second-stage GMM_A)? And second, which method leads to better estimators' properties intra methods (OLS versus first-stage GMM^1_A, and GLS versus second-stage GMM^2_A)? We will answer these questions by aggregating by models, test portfolios and sample periods. Then, we are actually comparing 2184 estimators with their corresponding standard errors, t-statistics and percentage bias. Our results suggest that, in general, the Beta method leads to better properties than SDF.

Due to the large number of λ and b estimates, it is useful to classify their properties into three categories, in a similar way as we did before. For this purpose, we define the category A in the following Tables for those estimators who are less than 50 percent biased from the observable risk factor and are statistically different from zero. The category B corresponds to those estimators with biases between 51 percent and 100 percent whether or not they are statistically different from zero. Finally, the category C is for estimates with 101 percent to 1000 percent biases whether or not they are statistically different from zero. Few observations with a bias even higher than 1000 percent are dropped out from the analysis; it is worthwhile to mention that 95 percent of these dropped values correspond to SDF estimators.

Naturally, the category A represents the best properties. By restricting to be statistically different to zero we guarantee that the standard error is relatively small, and the bias condition assures that the estimate is reasonable as Lewellen, Nagel and Shanken [72] emphasize. In the category B, we are not interested on the size of the standard error; the only condition is the bias interval. Thus, the unreasonable estimates will fall into this category. The category C represents obviously the worst properties because their bias is extremely high; these estimators become not only unreasonable but also unreliable.

The full results of Tables 3.4, 3.5 and 3.6 are on sections 3.6.1 to 3.6.7.

Table 3.4 shows that GLS leads to better properties than OLS, except for the industry portfolios in which the category A goes from 48% in OLS to 41% in GLS. The rest of portfolios increase the proportion of category A between 4 to 10 percentage points. Thus, in general, GLS is actually doing its job at providing better properties by giving up some pricing errors. Furthermore, the category C is actually smaller for the GLS, strengthening the fact that GLS has better properties than OLS. This is true for all portfolios except again for the industry classification in which the red area goes from 27% in OLS to 32% in GLS. The rest show a decrease from 0 to 26 percentage points.

On the other hand, the results on Table 3.4 regarding the second-stage $\mathrm{GMM}_{\mathrm{A}}^{2}$ relative to the first-stage $\mathrm{GMM}_{\mathrm{A}}^{1}$ estimators are less clear in achieving better properties. Thus, the second-stage estimators obtain better properties, except for the double-sorted portfolios in which the category A decreases from 27% (Fama-French) and 40% (size-momentum) in first-stage to 18% and 17% respectively in second-stage.

Table 3.4: Properties of Estimators

The numbers denote the frequency in which each estimator falls into each category. We consider all three models simultaneously, the CAPM, the Fama-French-three factor model, and the three-factor model with excess market return, the momentum factor and the value factor (RUH). Category A is for those estimators which are less than 50 percent biased from the realized factor, and are statistically different from zero; Category B corresponds to estimators with biases between 51 and 100 percent, and Category C for estimates with 101 to 1000 percent biases. We employ six time-periods, T=60, 120, 240, 360, 480, and 948. Size includes 5 and 10 size-sorted portfolios; Value has 5 and 10 BE/ME-sorted assets; Size and Value are the 6, 25 and 100 Fama-French portfolios; Size and Momentum includes 6 and 25 portfolios formed on ME and MOM; Industry incorporates 5, 17 and 30 industry-sorted assets, and Size, Value and Industry has the 25 Fama-French extended with 17 industry portfolios.

1 - p-value	ME	BE/ME	ME&BE/ME	ME&MOM	Industry	25 ME&BE/ME + 17 Industry
Panel A: Beta method OLS						
C	0.30	0.16	0.14	0.30	0.27	0.12
B	0.28	0.35	0.13	0.20	0.25	0.21
A	0.42	0.49	0.72	0.51	0.48	0.67
Panel B: Beta method GLS						
C	0.18	0.13	0.09	0.04	0.32	0.12
B	0.35	0.35	0.09	0.27	0.27	0.12
A	0.47	0.53	0.82	0.69	0.41	0.76
Panel C: SDF/GMM method $\mathrm{GMM}_{\mathrm{A}}^{1}$						
C	0.41	0.24	0.50	0.32	0.32	0.50
B	0.39	0.44	0.23	0.28	0.40	0.21
A	0.20	0.33	0.27	0.40	0.28	0.29
Panel D: SDF/GMM method $\mathrm{GMM}_{\mathrm{A}}^{2}$						
C	0.40	0.23	0.71	0.59	0.31	0.80
B	0.35	0.43	0.11	0.23	0.40	0.12
A	0.26	0.35	0.18	0.17	0.29	0.08

The combined portfolios also show worst properties in second-stage estimators. The rest have a modest improvement between 1 and 6 percentage points. The category C in second-stage is lower than in first-stage, except again for the double sorted portfolios which goes from 50% (Fama-French) and 32% (size-momentum) to 71% and 59% in second-stage. The rest of them have a tiny improvement of 1 percentage point each.

To summarize, GLS and second-stage $\mathrm{GMM}_{\mathrm{A}}^{2}$ estimators tend to present better properties than OLS and first-stage $\mathrm{GMM}_{\mathrm{A}}^{1}$. But clearly the difference is most notably when comparing the Beta rather than the SDF method. In this comparison, it is interesting to point out the role of the double-sorted portfolios. Note that the Beta method can achieve better estimators' properties even in portfolios with high dispersion, while the SDF method cannot. This is consistent with the idea that GMM has difficulties in small samples. In our case this difficulty is associated with the higher dispersion in the portfolios' expected returns. In other words, GMM seems to have difficulties in pricing assets when changing from single-sorted to double and combined-sorted portfolios.

Regarding the second question, which method leads to better estimators' properties intra methods (OLS versus first-stage $\mathrm{GMM}_{\mathrm{A}}^{1}$, and GLS versus second-stage $\mathrm{GMM}_{\mathrm{A}}^{2}$)? The Beta method dominates SDF in terms of estimators' properties. The category A consistently becomes larger from first-stage $\mathrm{GMM}_{\mathrm{A}}^{1}$ to OLS and from second-stage $\mathrm{GMM}_{\mathrm{A}}^{2}$ to GLS. In particular, the increases go from 11 (size-momentum portfolios) to 45 (Fama-French portfolios) percentage points. On the other hand, there is also a substantial decrease of the category C between 10 (value portfolios) to 62 (Fama-French portfolios) percentage points; in this case the only exception is the industry portfolios which slightly increases the red area from 31% (second-stage) to 32% (GLS).

Jagannathan and Wang [58] show that asymptotically no method dominates in terms of estimators' properties, but their findings was supported on the analysis of the single factor model under artificial data. So far however, we report evidence showing that the Beta method dominates the SDF framework. Now, we would like to split the data on Table 3.4 in order to analyze differences among the single and multifactor models. Once we do that, we could further compare our results with those on Jagannathan and Wang [58].

Our next task in Table 3.5 is therefore to compare the estimators' properties across models. Given that the Beta method dominates SDF, then, the next question is whether this evidence is consistent for each model and for all methods. For this purpose, we use an even broader set of estimators than before: we now calculate three kinds of estimators in a Beta formulation: OLS, GLS and WLS; and five estimators in a SDF formulation: returns on second moments GMM_A; returns on covariances GMM_B; and the continuous updating GMM_C. These estimations were obtained by taking our full sample on T and N. We finally end up with about 1636 lambda and 2730 b estimators.

The evidence is summarized in Table 3.5. Each panel represents the alternative factor models, while columns are λ and b estimates' properties. In each column we report the frequency of each category by factor (market, size, value and momentum). We drop some estimators with more than 1,000 percent bias from the factor mean. As before, it is worthwhile to point out that 95 percent of the 151 total dropped values correspond to the SDF. This already suggests which method is more likely to deliver

Table 3.5: Properties of Estimators. Beta Method: OLS, WLS, and GLS; SDF/GMM Method: First- and Second-Stage GMM_A, First- and Second-Stage GMM_B, and GMM_C. Full-Time Period Data.

The numbers denote the frequency in which each estimator falls into each category. We consider all three models, the CAPM, the FF-three factor model, and the three-factor model with excess market return, the momentum factor and the value factor (RUH). Category A is for those estimators which are less than 50 percent biased from the realized factor, and are statistically different from zero; Category B corresponds to estimators with biases between 51 and 100 percent, and Category C for estimates with 101 to 1000 percent biases. We employ six time-periods, T = 60, 120, 240, 360, 480, and 948, and all N portfolios: Size includes 5 and 10 size-sorted portfolios; Value has 5 and 10 BE/ME-sorted assets; Size and Value are the 6, 25 and 100 Fama-French portfolios; Size and Momentum includes 6 and 25 portfolios formed on ME and MOM; Industry incorporates 5, 17 and 30 industry-sorted assets, and Size, Value and Industry has the 25 Fama-French extended with 17 industry portfolios. The estimators are obtained using OLS, WLS, and GLS for the Beta method, and First- and Second-Stage GMM_A, First- and Second-Stage GMM_B, and GMM_C for the SDF/GMM method.

	Risk Premia (first column: λ; second column: b coefficient)			
Category	Market	Size	Value	Momentum
	Panel A: CAPM			
C	0.08 ; 0.17	-	-	-
B	0.09 ; 0.10	-	-	-
A	0.83 ; 0.73	-	-	-
	Panel B: Fama-French			
C	0.01 ; 0.28	0.33 ; 0.50	0.24 ; 0.44	-
B	0.20 ; 0.48	0.36 ; 0.30	0.35 ; 0.37	-
A	0.79 ; 0.24	0.31 ; 0.20	0.41 ; 0.19	-
	Panel C: RUH			
C	0.04 ; 0.54	-	0.47 ; 0.62	0.28 ; 0.50
B	0.15 ; 0.33	-	0.28 ; 0.25	0.30 ; 0.37
A	0.81 ; 0.12	-	0.25 ; 0.13	0.42 ; 0.12

an estimator with worst properties.

Let us focus on the first panel (CAPM) of Table 3.5. As in Jagannathan and Wang [58] and Cochrane [25], we find that the λ and b associated with the market factor have almost identical properties in both methods, even in a more complex setup than the simple CAPM with 10 size-sorted portfolios. In particular, we find that the probability of having good properties is 83% and 73% for the Beta and SDF respectively. This result is remarkable because it is actually what Jagannathan and Wang [58] show in their empirical results with simulated data, and we find exactly the same results using historical data.

Now, let us analyze the λ and b associated with the market factor in the rest of the models. We argue that we are actually assessing a gap in the previous literature by extending the current well known result in the first panel of Table 3.5 to the second and third panels under multifactor models. The evidence is represented by the first two pairs of columns for the Fama-French and RUH specifications. The results are stunning; the λ from the Beta method has much better properties than the b from SDF method. In particular, the category A is 79% versus 24% for the Fama-French model and 81% versus 12% for the RUH. It seems clear then that we can not say that Beta and SDF methods lead to the same estimators' properties.

In any case, why does the λ and b estimates associated with the market factor have so similar properties in the CAPM and so different in multifactor models? One plausible answer is that b gives a multiple regression coefficient of m on the factor given the other factors; while λ gives the single regression coefficient. In the CAPM model of course, there is no difference between single or multiple regression coefficient since there is only one factor. So, λ and b behave in a very similar way as long as both report single regression coefficients, but things change when adding more factors.

The size factor from Fama-French model and the momentum factor from RUH model in Tables 3.5 and 3.6 are priced and help pricing assets given the other factors in a similar magnitude. That is, even though we find that Beta method lead to better estimators' properties than SDF, the difference is not as large as when comparing the market factor in the Fama-French and RUH models.

The results regarding the λ and b associated with the value factor in Tables 3.5 and 3.6 are practically the same among multifactor models. On the other, once again, the Beta method does achieve better properties than the SDF procedure. In the Fama-French model, the category A is almost twice for λ (43% versus 23%), and the RUH specification shows a similar pattern (35% versus 16%).

As Amsler and Schmidt [2] and Shanken and Zhou [94], we find that when T is small (say 60 or 120), these estimators can be very volatile across different test portfolios. Thus, in Table 3.6 we exclude time lengths equal to 60, 120 and 240. When taking away the three smaller sample periods, we are actually dropping off the estimators with higher bias and standard errors, and then our conclusions about Figure 3.5 strengthen.

It is notably how well the methods can achieve desirable estimators' properties, especially when estimating the CAPM, independently of the test portfolio. Moreover, the differences between Beta and SDF are still very low: 10 percentage points in Table 3.5 and 9 percentage points in Table 3.6.

The properties regarding the multifactor models get better when dropping out the smallest time-periods. The category A is now consistently larger and the category C smaller than in Table 3.5. Note that Fama-French model outperforms the others

Table 3.6: Properties of Estimators. Beta Method: OLS, WLS, and GLS; SDF/GMM Method: First- and Second-Stage GMM_A, First- and Second-Stage GMM_B, and GMM_C. Reduced-Time Period Data.

The numbers denote the frequency in which each estimator falls into each category. We consider all three models, the CAPM, the FF-three factor model, and the three-factor model with excess market return, the momentum factor and the value factor (RUH). Category A is for those estimators which are less than 50 percent biased from the realized factor, and are statistically different from zero; Category B corresponds to estimators with biases between 51 and 100 percent, and Category C for estimates with 101 to 1000 percent biases. We employ three time-periods, T=360, 480, and 948, and all N portfolios: Size includes 5 and 10 size-sorted portfolios; Value has 5 and 10 BE/ME-sorted assets; Size and Value are the 6, 25 and 100 Fama-French portfolios; Size and Momentum includes 6 and 25 portfolios formed on ME and MOM; Industry incorporates 5, 17 and 30 industry-sorted assets, and Size, Value and Industry has the 25 Fama-French extended with 17 industry portfolios. The estimators are obtained using OLS, WLS, and GLS for the Beta method, and First- and Second-Stage GMM_A, First- and Second-Stage GMM_B, and GMM_C for the SDF/GMM method.

	Risk Premia (first column: λ; second column: b coefficient)			
Category	Market	Size	Value	Momentum
	Panel A: CAPM			
C	0.00 ; 0.05	-	-	-
B	0.00 ; 0.04	-	-	-
A	1.00 ; 0.91	-	-	-
	Panel B: Fama-French			
C	0.00 ; 0.12	0.31 ; 0.42	0.29 ; 0.35	-
B	0.02 ; 0.51	0.31 ; 0.34	0.28 ; 0.42	-
A	0.98 ; 0.37	0.38 ; 0.24	0.43 ; 0.23	-
	Panel C: RUH			
C	0.00 ; 0.35	-	0.34 ; 0.53	0.27 ; 0.41
B	0.01 ; 0.46	-	0.31 ; 0.31	0.21 ; 0.42
A	0.99 ; 0.19	-	0.35 ; 0.16	0.52 ; 0.17

in terms of estimators' properties, while the RUH model does better than the others in achieving lower pricing errors. Also, λ and b for the size factor are the estimators with worst properties in the Fama-French model, while λ and b of the momentum factor are the estimators with worst properties in the RUH model.

Finally, our evidence shows that the generality of SDF method comes at a cost of slightly misleading standard errors, especially in the second-stage (see Table 3.4 again) and the continuous updating estimators. In particular, standard errors are always larger for the SDF method. Also, using a long sample period always helps improving the estimators' properties (see Tables 3.5 and 3.6), lower the pricing errors, but not necessarily improve the tests results (see Table 3.3).

3.5 Conclusions

Our objective is to contribute to the current knowledge about the differences between Beta and SDF methods when estimating factor pricing models. It is well known that no differences arise in simple setups as we show in section 3.4.1, but there is a gap of empirical evidence regarding the conditions and consequences of using more complex setups. It is also well known that GMM has difficulties in finite samples, when dealing with extreme nonlinearities; nonetheless we find that their difficulties can arise even in linear models.

We find that Hansen and Jagannathan [49] first-stage $\mathrm{GMM}^1_{\mathrm{A}}$ achieve lower pricing errors than OLS Beta method for all test portfolios and time-lengths. On the other hand, their specifications tests show evidence in favor of multifactor models such as RUH because the likeliness of not rejecting the null is greater than in Fama-French and CAPM models. We also find that double-sorted portfolios are hardest to price compared with single-sorted portfolios, and this difference is correlated with the higher dispersion on the test portfolios' average returns.

Theory indicates that GLS should lead to better estimators' properties than OLS, and this should also apply to second and first-stage GMM estimators. Our results suggest that the Beta method actually do better than the SDF method. Moreover, when pricing double-sorted portfolios, the properties on second-stage are actually worse than in first-stage.

We are capable to reproduce Jagannathan and Wang [58] results for the CAPM even in a finite-sample framework, which reinforce the strength of the fact that there is no difference between Beta and SDF methods when comparing λ and b properties under the simple CAPM. Our main and original contribution relies on extending the comparison for the Fama-French and the RUH specifications. Our results imply that differences between the performance of the methods arises in more complex setups such as the ones suggested by Lewellen, Nagel and Shanken [72]. In particular λ from the Beta method has better properties in multifactor models such as Fama-French and RUH than b from the SDF method across tests portfolios and sample periods.

We are also capable to reproduce most of the results on Shanken and Zhou [94], which analyze models under the Beta representation. Our contribution here is to extend the analysis to SDF methods as well. In particular, we demonstrate that first-stage GMM estimators are in general superior than Beta estimators in order to achieve lower pricing errors. One practical implication for this finding is that a good asset pricing model for evaluating performance of managed funds should have small pricing errors. Therefore, fund managers are more likely to accurately evaluate mutual funds and hedge funds, by implementing SDF rather than Beta methods.[7]

[7]See Wang and Zhang [102] for a brief discussion, and for the implications of using SDF asset

Our results confirm that multi-factor models perform best in our model comparison tests than the single-factor model. This result has been documented in similar and very recent works such as in Kan and Robotti and Shanken [65]; in particular, they find that the CAPM and the unconditional consumption CAPM are frequently dominated by other models like the intertemporal CAPM of Petkova [87], the conditional CAPM of Jagannathan and Wang [56], and the Fama and French [30] models.

Further work may address the implications for using simulated data (e.g. from multivariate normal and t distributions). This benefits the analysis in terms of providing size and power tests, and also may go deeper into the analysis of estimators' properties.

■

pricing models over contingent claims. Also, see Ferson and Siegel [39] for an interesting hedge fund example.

3.6 Appendix of Tables

3.6.1 Market risk premium in CAPM model

The following five panels report the estimate values, the bias from the risk premium in percent values, the standard error in parenthesis, and t-statistics. Each panel corresponds to a one set of test portfolios. T is 60 (January 2001 - December 2005), 120, 240, 360, 480 and 948 (January 1927 - December 2005) monthly observations. The two-pass (Fama-MacBeth) cross-sectional estimate is the slope coefficient λ in $E(R^e) = \beta\lambda$ calculated by OLS, GLS and WLS. Next, we turn from beta representation to a discount factor formulation for GMM approach. The $\mathrm{GMM_A}$ and $\mathrm{GMM_B}$ first and second-stage estimates are the parameters $100 \times b$ in $E(R^e) = E(R^e f')\, b$ (returns on second moments, following Hansen and Jagannathan [49]) and $E(R^e) = E\left(R^e \widetilde{f}'\right) b$ (returns on covariances, following Cochrane [25]) respectively. The $\mathrm{GMM_C}$ is the continuous updating estimate (following Hansen, Heaton and Yaron [47]).

Panel A. N Portfolios formed on ME

	T = 60	**T = 120**	**T = 240**	**T = 360**	**T = 480**	**T = 948**
$\lambda = Ef$	0.13	0.56	0.64	0.62	0.44	0.64
$b = \frac{\lambda}{E(R^{em2})}$	0.66	2.52	3.15	3.11	2.11	2.11
			N = 5			
Beta: $\widehat{\lambda}_{\text{OLS}}$	0.68 415%(0.61)	0.75** 34%(0.47)	0.70*** 8%(0.31)	0.77*** 24%(0.25)	0.56*** 29%(0.22)	0.73*** 14%(0.19)
Beta: $\widehat{\lambda}_{\text{GLS}}$	0.11 −19%(0.57)	0.62* 11%(0.43)	0.69*** 8%(0.29)	0.67*** 8%(0.23)	0.47*** 9%(0.21)	0.67*** 4%(0.18)
Beta: $\widehat{\lambda}_{\text{WLS}}$	0.14 5%(0.58)	0.63* 13%(0.43)	0.70*** 9%(0.29)	0.69*** 11%(0.24)	0.49*** 13%(0.21)	0.67*** 5%(0.18)
GMM$_\text{A}$: $\widehat{b}_1$	3.46 422%(3.26)	3.40* 35%(2.31)	3.43** 9%(1.77)	3.84*** 23%(1.41)	2.71*** 29%(1.13)	2.40*** 14%(0.59)
GMM$_\text{A}$: $\widehat{b}_2$	1.18 78%(3.12)	3.25** 29%(2.11)	3.71*** 17%(1.62)	3.29*** 6%(1.30)	2.31*** 10%(1.05)	2.30*** 9%(0.59)
GMM$_\text{B}$: $\widehat{b}_1$	3.47 424%(3.35)	3.46* 37%(2.44)	3.50** 11%(1.89)	3.93*** 26%(1.51)	2.74*** 30%(1.17)	2.44*** 15%(0.62)
GMM$_\text{B}$: $\widehat{b}_2$	0.74 11%(3.15)	2.95* 17%(2.23)	3.64** 15%(1.73)	3.30*** 6%(1.38)	2.31** 9%(1.09)	2.33*** 10%(0.61)
GMM$_\text{C}$: $\widehat{b}$	2.03 206%(3.29)	2.92* 16%(2.39)	3.70** 17%(1.91)	3.33*** 7%(1.47)	2.32** 10%(1.15)	2.33*** 10%(0.62)
			N = 10			
Beta: $\widehat{\lambda}_{\text{OLS}}$	0.74* 456%(0.62)	0.78** 40%(0.47)	0.71*** 11%(0.31)	0.78*** 26%(0.25)	0.57*** 31%(0.22)	0.74*** 15%(0.19)
Beta: $\widehat{\lambda}_{\text{GLS}}$	0.12 −7%(0.57)	0.62** 12%(0.43)	0.70*** 9%(0.29)	0.68*** 9%(0.23)	0.48*** 10%(0.21)	0.67*** 4%(0.18)
Beta: $\widehat{\lambda}_{\text{WLS}}$	0.39 195%(0.58)	0.71** 27%(0.43)	0.73*** 14%(0.29)	0.73*** 18%(0.23)	0.53*** 21%(0.21)	0.69*** 8%(0.18)
GMM$_\text{A}$: $\widehat{b}_1$	3.74* 464%(3.30)	3.53** 40%(2.34)	3.50*** 11%(1.79)	3.91*** 26%(1.42)	2.76*** 31%(1.13)	2.43*** 15%(0.59)
GMM$_\text{A}$: $\widehat{b}_2$	1.53 131%(3.13)	3.80** 51%(2.09)	3.83*** 21%(1.61)	3.43*** 10%(1.29)	2.38*** 13%(1.05)	2.39*** 13%(0.58)
GMM$_\text{B}$: $\widehat{b}_1$	3.75* 466%(3.40)	3.60** 43%(2.48)	3.58*** 14%(1.91)	4.01*** 29%(1.53)	2.79*** 32%(1.18)	2.46*** 17%(0.62)
GMM$_\text{B}$: $\widehat{b}_2$	0.27 −60%(3.14)	3.05* 21%(2.22)	3.46*** 10%(1.72)	3.28*** 5%(1.38)	2.33*** 10%(1.09)	2.41*** 14%(0.61)
GMM$_\text{C}$: $\widehat{b}$	−1.19 −280%(3.04)	3.93** 56%(2.54)	4.10*** 30%(1.98)	3.49*** 12%(1.49)	2.37*** 12%(1.16)	2.42*** 15%(0.62)

Panel B. N Portfolios formed on BE/ME

	T = 60	**T = 120**	**T = 240**	**T = 360**	**T = 480**	**T = 948**
$\lambda = Ef$	0.13	0.56	0.64	0.62	0.44	0.64
$b = \frac{\lambda}{E(R^{em2})}$	0.66	2.52	3.15	3.11	2.11	2.11
			N = 5			
Beta: $\widehat{\lambda}_{\mathrm{OLS}}$	0.39 194%(0.60)	0.82** 48%(0.45)	0.83*** 29%(0.30)	0.80*** 29%(0.24)	0.60*** 37%(0.21)	0.73*** 13%(0.18)
Beta: $\widehat{\lambda}_{\mathrm{GLS}}$	0.06 −58%(0.58)	0.58* 5%(0.43)	0.70*** 9%(0.29)	0.67*** 7%(0.24)	0.49*** 11%(0.21)	0.66*** 3%(0.18)
Beta: $\widehat{\lambda}_{\mathrm{WLS}}$	0.18 35%(0.58)	0.66* 19%(0.43)	0.76*** 18%(0.29)	0.72*** 16%(0.24)	0.53*** 21%(0.21)	0.68*** 5%(0.18)
$\mathrm{GMM_A}$: $\widehat{b}_1$	1.98 199%(3.11)	3.74** 48%(2.23)	4.05*** 28%(1.74)	4.01*** 29%(1.34)	2.89*** 37%(1.08)	2.38*** 13%(0.59)
$\mathrm{GMM_A}$: $\widehat{b}_2$	0.54 −19%(3.03)	2.72* 8%(2.13)	3.64*** 15%(1.67)	3.34*** 7%(1.30)	2.59*** 23%(1.06)	2.35*** 11%(0.58)
$\mathrm{GMM_B}$: $\widehat{b}_1$	1.98 199%(3.15)	3.81** 51%(2.38)	4.16*** 32%(1.88)	4.10*** 32%(1.45)	2.92*** 38%(1.12)	2.42*** 15%(0.61)
$\mathrm{GMM_B}$: $\widehat{b}_2$	0.43 −35%(3.04)	2.60 3%(2.25)	3.58** 13%(1.79)	3.17*** 2%(1.39)	2.48*** 18%(1.10)	2.37*** 12%(0.60)
$\mathrm{GMM_C}$: $\widehat{b}$	2.55 285%(3.25)	3.07* 22%(2.32)	3.91** 24%(1.86)	3.61*** 16%(1.42)	2.53*** 20%(1.11)	2.39*** 13%(0.61)
			N = 10			
Beta: $\widehat{\lambda}_{\mathrm{OLS}}$	0.42 219%(0.60)	0.84*** 51%(0.45)	0.83*** 30%(0.30)	0.82*** 31%(0.24)	0.61*** 39%(0.21)	0.72*** 13%(0.18)
Beta: $\widehat{\lambda}_{\mathrm{GLS}}$	0.06 −52%(0.57)	0.59** 6%(0.43)	0.70*** 10%(0.29)	0.66*** 7%(0.24)	0.49*** 11%(0.21)	0.67*** 4%(0.18)
Beta: $\widehat{\lambda}_{\mathrm{WLS}}$	0.31 132%(0.58)	0.76** 36%(0.43)	0.79*** 23%(0.29)	0.76*** 22%(0.24)	0.56*** 27%(0.21)	0.69*** 8%(0.18)
$\mathrm{GMM_A}$: $\widehat{b}_1$	2.15 224%(3.15)	3.82** 51%(2.25)	4.10*** 30%(1.76)	4.07*** 31%(1.35)	2.93*** 39%(1.08)	2.38*** 12%(0.58)
$\mathrm{GMM_A}$: $\widehat{b}_2$	0.68 2%(3.03)	2.80* 11%(2.13)	3.79*** 20%(1.65)	3.46*** 11%(1.29)	2.58*** 22%(1.06)	2.22*** 5%(0.57)
$\mathrm{GMM_B}$: $\widehat{b}_1$	2.15 224%(3.19)	3.89** 54%(2.40)	4.20*** 33%(1.90)	4.17*** 34%(1.46)	2.96*** 40%(1.13)	2.41*** 14%(0.61)
$\mathrm{GMM_B}$: $\widehat{b}_2$	0.52 −21%(3.04)	2.58* 2%(2.25)	3.69*** 17%(1.78)	3.27*** 5%(1.38)	2.47*** 17%(1.10)	2.19*** 4%(0.59)
$\mathrm{GMM_C}$: $\widehat{b}$	3.11 370%(3.34)	2.88* 14%(2.32)	3.96*** 26%(1.88)	3.68*** 18%(1.43)	2.52*** 19%(1.11)	2.26*** 7%(0.61)

Panel C. N Portfolios formed by the intersections of ME and BE/ME

	T = 60	**T = 120**	**T = 240**	**T = 360**	**T = 480**	**T = 948**
$\lambda = Ef$	0.13	0.56	0.64	0.62	0.44	0.64
$b = \frac{\lambda}{E\left(R^{em2}\right)}$	0.66	2.52	3.15	3.11	2.11	2.11
			N = 6			
Beta: $\widehat{\lambda}_{\mathrm{OLS}}$	0.55 315%(0.61)	0.74** 33%(0.46)	0.72*** 12%(0.30)	0.79*** 27%(0.24)	0.60*** 37%(0.22)	0.75*** 17%(0.19)
Beta: $\widehat{\lambda}_{\mathrm{GLS}}$	0.05 −59%(0.57)	0.56* 1%(0.43)	0.63*** −2%(0.29)	0.61*** −2%(0.23)	0.45*** 2%(0.21)	0.65*** 1%(0.18)
Beta: $\widehat{\lambda}_{\mathrm{WLS}}$	0.15 15%(0.58)	0.65** 17%(0.43)	0.73*** 14%(0.29)	0.72*** 16%(0.24)	0.53*** 22%(0.21)	0.69*** 7%(0.18)
$\mathrm{GMM_A}$: $\widehat{b}_1$	2.80 322%(3.21)	3.39** 35%(2.26)	3.57*** 13%(1.77)	3.97*** 28%(1.40)	2.89*** 37%(1.12)	2.46*** 17%(0.59)
$\mathrm{GMM_A}$: $\widehat{b}_2$	0.77 16%(3.07)	3.42** 35%(2.11)	2.98** −6%(1.64)	2.92*** −6%(1.31)	2.34*** 11%(1.06)	2.21*** 4%(0.58)
$\mathrm{GMM_B}$: $\widehat{b}_1$	2.80 322%(3.27)	3.44* 36%(2.39)	3.63** 15%(1.88)	4.06*** 30%(1.50)	2.92*** 39%(1.16)	2.50*** 18%(0.62)
$\mathrm{GMM_B}$: $\widehat{b}_2$	0.41 −38%(3.09)	2.24 −11%(2.22)	2.07* −34%(1.73)	2.12** −32%(1.39)	1.90** −10%(1.10)	2.09*** −1%(0.60)
$\mathrm{GMM_C}$: $\widehat{b}$	4.82* 627%(3.63)	6.32*** 151%(2.84)	5.13*** 63%(2.13)	4.37*** 40%(1.57)	2.75*** 30%(1.16)	2.26*** 7%(0.61)
			N = 25			
Beta: $\widehat{\lambda}_{\mathrm{OLS}}$	0.71* 436%(0.62)	0.81** 45%(0.47)	0.75*** 17%(0.31)	0.83*** 34%(0.25)	0.62*** 42%(0.22)	0.74*** 14%(0.19)
Beta: $\widehat{\lambda}_{\mathrm{GLS}}$	0.11 −15%(0.57)	0.67** 21%(0.43)	0.69*** 8%(0.29)	0.65*** 4%(0.23)	0.46*** 6%(0.21)	0.67*** 4%(0.18)
Beta: $\widehat{\lambda}_{\mathrm{WLS}}$	0.54 307%(0.58)	0.85*** 53%(0.43)	0.81*** 27%(0.29)	0.84*** 35%(0.23)	0.62*** 41%(0.21)	0.73*** 14%(0.18)
$\mathrm{GMM_A}$: $\widehat{b}_1$	3.61* 445%(3.32)	3.70** 47%(2.35)	3.71*** 17%(1.82)	4.17*** 34%(1.44)	3.00*** 42%(1.14)	2.42*** 14%(0.60)
$\mathrm{GMM_A}$: $\widehat{b}_2$	4.72** 613%(2.99)	7.13*** 183%(2.02)	6.76*** 114%(1.56)	4.93*** 58%(1.25)	3.48*** 65%(1.04)	2.37*** 12%(0.55)
$\mathrm{GMM_B}$: $\widehat{b}_1$	3.62 446%(3.42)	3.75** 49%(2.50)	3.78*** 20%(1.94)	4.26*** 37%(1.56)	3.03*** 44%(1.19)	2.45*** 16%(0.62)
$\mathrm{GMM_B}$: $\widehat{b}_2$	1.27 91%(3.06)	3.18** 26%(2.20)	3.03*** −4%(1.70)	2.75*** −12%(1.35)	2.36*** 12%(1.09)	2.08*** −2%(0.57)
$\mathrm{GMM_C}$: $\widehat{b}$	×	×	×	7.72*** 148%(2.00)	3.40*** 61%(1.23)	2.38*** 13%(0.63)
			N = 100			
Beta: $\widehat{\lambda}_{\mathrm{OLS}}$		0.83*** 49%(0.47)	0.76*** 19%(0.31)	0.85*** 37%(0.25)	0.63*** 45%(0.22)	0.80*** 24%(0.19)
Beta: $\widehat{\lambda}_{\mathrm{GLS}}$		0.65** 16%(0.42)	0.68*** 7%(0.29)	0.65*** 5%(0.23)	0.47*** 7%(0.21)	0.69*** 7%(0.18)
Beta: $\widehat{\lambda}_{\mathrm{WLS}}$		0.91*** 63%(0.43)	0.82*** 28%(0.29)	0.87*** 40%(0.23)	0.63*** 45%(0.21)	0.76*** 18%(0.18)
$\mathrm{GMM_A}$: $\widehat{b}_1$		3.80** 51%(2.36)	3.77*** 20%(1.83)	4.26*** 37%(1.45)	3.06*** 45%(1.15)	2.62*** 24%(0.60)
$\mathrm{GMM_A}$: $\widehat{b}_2$		×	12.75*** 304%(1.41)	8.69*** 179%(1.16)	5.16*** 144%(1.01)	2.92*** 38%(0.44)
$\mathrm{GMM_B}$: $\widehat{b}_1$		3.85** 53%(2.52)	3.84*** 22%(1.96)	4.35*** 40%(1.57)	3.09*** 46%(1.20)	2.66*** 26%(0.63)
$\mathrm{GMM_B}$: $\widehat{b}_2$		2.75** 9%(2.07)	3.45*** 9%(1.62)	3.35*** 8%(1.29)	2.43*** 15%(1.06)	2.51*** 19%(0.47)
$\mathrm{GMM_C}$: $\widehat{b}$		4.36** 73%(2.67)	−0.78 −125%(1.47)	×	15.97*** 657%(2.43)	2.52*** 19%(0.63)

Panel D. N Portfolios formed on ME and MOM

	T = 60	**T = 120**	**T = 240**	**T = 360**	**T = 480**	**T = 948**
$\lambda = Ef$	0.13	0.56	0.64	0.62	0.44	0.64
$b = \frac{\lambda}{E(R^{em2})}$	0.66	2.52	3.15	3.11	2.11	2.11
			N = 6			
Beta: $\widehat{\lambda}_{\text{OLS}}$	0.51 283%(0.60)	0.65* 17%(0.46)	0.66*** 2%(0.30)	0.73*** 17%(0.24)	0.52*** 18%(0.22)	0.65*** 2%(0.19)
Beta: $\widehat{\lambda}_{\text{GLS}}$	0.13 −4%(0.57)	0.57* 3%(0.43)	0.66*** 2%(0.29)	0.64*** 3%(0.24)	0.46*** 6%(0.21)	0.70*** 9%(0.18)
Beta: $\widehat{\lambda}_{\text{WLS}}$	0.29 117%(0.58)	0.74** 33%(0.44)	0.75*** 17%(0.30)	0.75*** 21%(0.24)	0.53*** 22%(0.21)	0.69*** 8%(0.18)
GMM_A: $\widehat{b}_1$	2.58 289%(3.16)	2.99* 19%(2.23)	3.26** 3%(1.72)	3.67*** 18%(1.37)	2.51*** 19%(1.10)	2.16*** 2%(0.59)
GMM_A: $\widehat{b}_2$	1.63 145%(3.06)	3.38** 34%(2.10)	4.54*** 44%(1.57)	4.30*** 38%(1.27)	2.84*** 35%(1.05)	1.96*** −7%(0.58)
GMM_B: $\widehat{b}_1$	2.58 289%(3.21)	3.02* 20%(2.34)	3.31** 5%(1.82)	3.73*** 20%(1.46)	2.52*** 19%(1.14)	2.18*** 3%(0.61)
GMM_B: $\widehat{b}_2$	1.07 62%(3.09)	2.57* 2%(2.21)	3.11** −2%(1.68)	2.85*** −8%(1.36)	2.04** −3%(1.08)	1.65*** −22%(0.60)
GMM_C: $\widehat{b}$	8.88*** 1241%(4.14)	3.29* 30%(2.39)	3.80*** 20%(1.88)	6.06*** 95%(1.71)	3.08*** 46%(1.17)	1.54*** −27%(0.60)
			N = 25			
Beta: $\widehat{\lambda}_{\text{OLS}}$	0.60 353%(0.61)	0.68** 23%(0.47)	0.67*** 5%(0.31)	0.76*** 23%(0.25)	0.54*** 24%(0.22)	0.69*** 7%(0.19)
Beta: $\widehat{\lambda}_{\text{GLS}}$	0.18 39%(0.57)	0.51* −8%(0.43)	0.66*** 3%(0.29)	0.65*** 4%(0.24)	0.49*** 13%(0.21)	0.68*** 6%(0.18)
Beta: $\widehat{\lambda}_{\text{WLS}}$	0.65* 395%(0.58)	0.83*** 49%(0.43)	0.78*** 22%(0.29)	0.83*** 33%(0.23)	0.59*** 36%(0.21)	0.73*** 14%(0.18)
GMM_A: $\widehat{b}_1$	3.05 361%(3.20)	3.14** 25%(2.28)	3.36*** 6%(1.76)	3.84*** 23%(1.40)	2.63*** 25%(1.12)	2.27*** 8%(0.59)
GMM_A: $\widehat{b}_2$	7.21*** 988%(2.91)	5.71*** 126%(2.05)	6.81*** 116%(1.52)	6.47*** 108%(1.23)	3.92*** 86%(1.04)	2.58*** 22%(0.56)
GMM_B: $\widehat{b}_1$	3.05 361%(3.27)	3.16* 25%(2.40)	3.40*** 8%(1.86)	3.90*** 25%(1.50)	2.64*** 25%(1.16)	2.29*** 9%(0.62)
GMM_B: $\widehat{b}_2$	2.58 289%(3.03)	2.73* 8%(2.19)	3.20*** 1%(1.65)	3.12*** 0%(1.34)	2.34*** 11%(1.08)	2.06*** −2%(0.58)
GMM_C: $\widehat{b}$	−4.48 −776%(3.06)	21.42*** 749%(5.69)	32.10*** 917%(6.98)	28.50*** 815%(4.75)	7.83*** 271%(1.54)	2.40*** 13%(0.62)

Panel E. N Industry Portfolios

	T = 60	**T = 120**	**T = 240**	**T = 360**	**T = 480**	**T = 948**
$\lambda = Ef$	0.13	0.56	0.64	0.62	0.44	0.64
$b = \frac{\lambda}{E\left(R^{em2}\right)}$	0.66	2.52	3.15	3.11	2.11	2.11
			N = 5			
Beta: $\widehat{\lambda}_{\text{OLS}}$	0.02 −83%(0.58)	0.60* 8%(0.43)	0.72*** 13%(0.29)	0.71*** 14%(0.24)	0.52*** 19%(0.21)	0.71*** 11%(0.18)
Beta: $\widehat{\lambda}_{\text{GLS}}$	0.13 −5%(0.57)	0.61* 10%(0.43)	0.70*** 9%(0.29)	0.69*** 11%(0.23)	0.50*** 14%(0.21)	0.69*** 8%(0.18)
Beta: $\widehat{\lambda}_{\text{WLS}}$	0.18 32%(0.59)	0.67* 19%(0.44)	0.75*** 17%(0.30)	0.73*** 18%(0.24)	0.52*** 20%(0.21)	0.69*** 8%(0.18)
GMM_{A}: $\widehat{b}_1$	0.12 −82%(2.94)	2.75* 9%(2.08)	3.58*** 13%(1.65)	3.56*** 14%(1.30)	2.51*** 19%(1.06)	2.34*** 11%(0.60)
GMM_{A}: $\widehat{b}_2$	0.71 8%(2.93)	2.88* 14%(2.07)	3.41** 8%(1.64)	3.42*** 10%(1.29)	2.44*** 16%(1.05)	2.19*** 4%(0.59)
GMM_{B}: $\widehat{b}_1$	0.12 −83%(2.94)	2.78* 10%(2.17)	3.65** 16%(1.77)	3.64*** 17%(1.39)	2.54*** 20%(1.10)	2.37*** 12%(0.63)
GMM_{B}: $\widehat{b}_2$	0.59 −12%(2.93)	2.79* 11%(2.17)	3.35** 6%(1.75)	3.44*** 11%(1.38)	2.45*** 16%(1.09)	2.20*** 4%(0.62)
GMM_{C}: $\widehat{b}$	1.78 169%(3.04)	3.04* 20%(2.20)	3.63** 15%(1.77)	3.56*** 14%(1.39)	2.48*** 17%(1.09)	2.24*** 6%(0.62)
			N = 17			
Beta: $\widehat{\lambda}_{\text{OLS}}$	0.39 195%(0.60)	0.63** 14%(0.45)	0.68*** 6%(0.30)	0.65*** 5%(0.24)	0.49*** 12%(0.21)	0.68*** 6%(0.18)
Beta: $\widehat{\lambda}_{\text{GLS}}$	0.16 23%(0.57)	0.58** 5%(0.43)	0.70*** 9%(0.29)	0.70*** 13%(0.23)	0.51*** 18%(0.21)	0.72*** 11%(0.18)
Beta: $\widehat{\lambda}_{\text{WLS}}$	0.30 124%(0.58)	0.55* −2%(0.43)	0.65*** 2%(0.29)	0.66*** 5%(0.24)	0.47*** 9%(0.21)	0.69*** 7%(0.18)
GMM_{A}: $\widehat{b}_1$	1.99 201%(3.13)	2.90* 15%(2.19)	3.37*** 7%(1.73)	3.29*** 6%(1.35)	2.36*** 12%(1.08)	2.24*** 6%(0.59)
GMM_{A}: $\widehat{b}_2$	1.63 146%(3.00)	3.35** 33%(2.06)	3.86*** 22%(1.59)	3.66*** 17%(1.27)	2.65*** 25%(1.04)	2.16*** 2%(0.58)
GMM_{B}: $\widehat{b}_1$	1.99 200%(3.17)	2.93* 16%(2.29)	3.43*** 9%(1.84)	3.35*** 8%(1.43)	2.38*** 13%(1.12)	2.27*** 7%(0.62)
GMM_{B}: $\widehat{b}_2$	0.90 36%(3.03)	2.56* 1%(2.16)	3.50*** 11%(1.70)	3.40*** 9%(1.35)	2.50*** 18%(1.08)	2.12*** 0%(0.60)
GMM_{C}: $\widehat{b}$	−1.02 −254%(2.98)	1.95 −23%(2.20)	3.73*** 18%(1.88)	3.64*** 17%(1.45)	2.53*** 20%(1.12)	2.19*** 4%(0.62)
			N = 30			
Beta: $\widehat{\lambda}_{\text{OLS}}$	0.42 219%(0.59)	0.72** 29%(0.45)	0.73*** 13%(0.30)	0.70*** 13%(0.24)	0.53*** 20%(0.21)	0.71*** 10%(0.18)
Beta: $\widehat{\lambda}_{\text{GLS}}$	0.09 −34%(0.57)	0.63** 13%(0.43)	0.73*** 13%(0.29)	0.73*** 18%(0.24)	0.53*** 22%(0.21)	0.74*** 14%(0.18)
Beta: $\widehat{\lambda}_{\text{WLS}}$	0.38 190%(0.58)	0.67** 20%(0.43)	0.71*** 11%(0.29)	0.71*** 14%(0.24)	0.51*** 17%(0.21)	0.70*** 8%(0.18)
GMM_{A}: $\widehat{b}_1$	2.17 227%(3.12)	3.30** 31%(2.19)	3.59*** 14%(1.75)	3.53*** 13%(1.36)	2.54*** 21%(1.09)	2.33*** 10%(0.60)
GMM_{A}: $\widehat{b}_2$	2.62 296%(3.00)	4.79*** 90%(2.03)	4.58*** 45%(1.57)	4.31*** 38%(1.26)	2.84*** 34%(1.04)	2.09*** −1%(0.58)
GMM_{B}: $\widehat{b}_1$	2.15 225%(3.16)	3.32** 32%(2.31)	3.66*** 16%(1.87)	3.60*** 16%(1.45)	2.57*** 22%(1.13)	2.36*** 12%(0.63)
GMM_{B}: $\widehat{b}_2$	0.90 36%(3.03)	3.15** 25%(2.17)	3.77*** 20%(1.70)	3.78*** 21%(1.35)	2.59*** 23%(1.08)	1.97*** −7%(0.60)
GMM_{C}: $\widehat{b}$	1.78 169%(3.16)	2.74* 9%(2.26)	4.31*** 37%(1.96)	3.90*** 25%(1.47)	2.75*** 30%(1.14)	2.02*** −5%(0.62)

3.6.2 Market risk premium in Fama-French model

The following five panels report the estimate values, the bias from the risk premium in percent values, the standard error in parenthesis, and t-statistics. Each panel corresponds to a one set of test portfolios. T is 60 (January 2001 - December 2005), 120, 240, 360, 480 and 948 (January 1927 - December 2005) monthly observations. The two-pass (Fama-MacBeth) cross-sectional estimate is the slope coefficient λ in $E(R^e) = \beta\lambda$ calculated by OLS, GLS and WLS. Next, we turn from beta representation to a discount factor formulation for GMM approach. The $\mathrm{GMM_A}$ and $\mathrm{GMM_B}$ first and second-stage estimates are the parameters $100 \times b$ in $E(R^e) = E(R^e f')\, b$ (returns on second moments, following Hansen and Jagannathan [49]) and $E(R^e) = E\left(R^e \widetilde{f}'\right) b$ (returns on covariances, following Cochrane [25]) respectively. The $\mathrm{GMM_C}$ is the continuous updating estimate (following Hansen, Heaton and Yaron [47]).

Panel A. N Portfolios formed on ME

	T = 60	T = 120	T = 240	T = 360	T = 480	T = 948
$\lambda = Ef$	0.13	0.56	0.64	0.62	0.44	0.64
$b = \frac{\lambda}{E(R^{em2})}$	0.66	2.52	3.15	3.11	2.11	2.11
			N = 5			
Beta: $\widehat{\lambda}_{\mathrm{OLS}}$	0.06 −57%(0.58)	0.66* 19%(0.43)	0.69** 7%(0.29)	0.66** 6%(0.23)	0.46** 6%(0.21)	0.68*** 6%(0.18)
Beta: $\widehat{\lambda}_{\mathrm{GLS}}$	0.10 −27%(0.57)	0.62* 11%(0.43)	0.67** 4%(0.29)	0.65** 5%(0.23)	0.46** 5%(0.21)	0.67*** 4%(0.18)
Beta: $\widehat{\lambda}_{\mathrm{WLS}}$	0.08 −38%(0.57)	0.62* 12%(0.43)	0.67** 5%(0.29)	0.66** 5%(0.23)	0.46** 5%(0.21)	0.67*** 4%(0.18)
$\mathrm{GMM_A}$: $\widehat{b}_1$	×	−0.33 −113%(4.98)	4.65 47%(4.18)	4.87* 56%(3.29)	3.97* 88%(2.74)	2.09** −1%(0.76)
$\mathrm{GMM_A}$: $\widehat{b}_2$	×	5.51 119%(4.16)	7.78** 147%(3.57)	5.84** 88%(2.87)	4.12* 95%(2.40)	1.97** −7%(0.73)
$\mathrm{GMM_B}$: $\widehat{b}_1$	×	−0.94 −137%(4.95)	4.39 39%(4.84)	5.15 65%(3.97)	4.19 98%(3.16)	2.12** 0%(0.77)
$\mathrm{GMM_B}$: $\widehat{b}_2$	×	4.39 74%(4.36)	8.00** 154%(4.22)	6.32* 103%(3.48)	4.32* 105%(2.75)	2.00** −5%(0.74)
$\mathrm{GMM_C}$: $\widehat{b}$	×	7.30 189%(5.44)	9.41* 198%(6.00)	6.49* 108%(4.16)	4.48* 112%(3.20)	2.01** −5%(0.76)
			N = 10			
Beta: $\widehat{\lambda}_{\mathrm{OLS}}$	0.07 −45%(0.58)	0.66** 19%(0.43)	0.68*** 6%(0.29)	0.66*** 7%(0.23)	0.47*** 8%(0.21)	0.67*** 4%(0.18)
Beta: $\widehat{\lambda}_{\mathrm{GLS}}$	0.12 −12%(0.57)	0.60* 8%(0.43)	0.69*** 7%(0.29)	0.67*** 8%(0.23)	0.47*** 8%(0.21)	0.67*** 4%(0.18)
Beta: $\widehat{\lambda}_{\mathrm{WLS}}$	0.08 −36%(0.57)	0.59* 7%(0.43)	0.68*** 5%(0.29)	0.66*** 6%(0.23)	0.47*** 7%(0.21)	0.67*** 3%(0.18)
$\mathrm{GMM_A}$: $\widehat{b}_1$	×	1.88 −25%(3.82)	5.60*** 78%(2.60)	4.84*** 56%(2.17)	3.40*** 61%(1.75)	1.88*** −11%(0.71)
$\mathrm{GMM_A}$: $\widehat{b}_2$	×	7.58*** 201%(3.05)	6.58*** 109%(2.30)	4.82*** 55%(1.97)	3.14*** 49%(1.60)	1.96*** −7%(0.68)
$\mathrm{GMM_B}$: $\widehat{b}_1$	6.64 903%(6.26)	0.97 −62%(4.14)	5.52** 75%(3.05)	4.99*** 60%(2.56)	3.47** 65%(1.93)	1.92*** −9%(0.73)
$\mathrm{GMM_B}$: $\widehat{b}_2$	×	6.26** 148%(3.39)	5.79*** 84%(2.67)	4.38*** 41%(2.29)	2.99** 42%(1.76)	1.99*** −6%(0.70)
$\mathrm{GMM_C}$: $\widehat{b}$	×	10.41*** 313%(5.49)	7.92*** 151%(3.46)	5.08*** 63%(2.59)	3.29** 56%(1.93)	2.01*** −5%(0.73)

Panel B. N Portfolios formed on BE/ME

	T = 60	**T = 120**	**T = 240**	**T = 360**	**T = 480**	**T = 948**
$\lambda = Ef$	0.13	0.56	0.64	0.62	0.44	0.64
$b = \frac{\lambda}{E(R^{em2})}$	0.66	2.52	3.15	3.11	2.11	2.11
			N = 5			
Beta: $\widehat{\lambda}_{\text{OLS}}$	0.08 −43%(0.58)	0.50 −10%(0.49)	0.69** 7%(0.31)	0.69** 10%(0.24)	0.48** 11%(0.21)	0.66*** 2%(0.18)
Beta: $\widehat{\lambda}_{\text{GLS}}$	0.07 −44%(0.58)	0.55 −1%(0.47)	0.66** 3%(0.31)	0.69** 10%(0.24)	0.48** 10%(0.21)	0.66*** 2%(0.18)
Beta: $\widehat{\lambda}_{\text{WLS}}$	0.07 −47%(0.58)	0.50 −9%(0.52)	0.69** 7%(0.31)	0.68** 10%(0.24)	0.49** 12%(0.21)	0.66*** 2%(0.18)
GMM_A: $\widehat{b}_1$	0.84 27%(3.82)	4.99** 98%(2.52)	4.99** 58%(2.00)	4.59** 47%(1.96)	3.13** 49%(1.62)	2.05** −3%(0.79)
GMM_A: $\widehat{b}_2$	1.05 59%(3.79)	4.89** 94%(2.47)	5.37** 70%(1.90)	4.49** 44%(1.82)	3.54** 68%(1.55)	1.97** −7%(0.78)
GMM_B: $\widehat{b}_1$	0.99 50%(4.45)	5.21* 106%(2.78)	5.25** 66%(2.23)	4.96** 59%(2.17)	3.29** 56%(1.72)	2.09** −1%(0.81)
GMM_B: $\widehat{b}_2$	1.19 80%(4.42)	5.12* 103%(2.74)	5.62** 78%(2.14)	4.85** 56%(2.03)	3.70** 75%(1.65)	2.01** −5%(0.80)
GMM_C: $\widehat{b}$	1.19 79%(4.47)	5.17* 105%(2.85)	5.63** 78%(2.28)	4.85** 56%(2.18)	3.71** 76%(1.72)	2.01** −5%(0.81)
			N = 10			
Beta: $\widehat{\lambda}_{\text{OLS}}$	0.09 −33%(0.58)	0.56* 1%(0.46)	0.68*** 5%(0.30)	0.65*** 5%(0.24)	0.48*** 10%(0.21)	0.66*** 2%(0.18)
Beta: $\widehat{\lambda}_{\text{GLS}}$	0.07 −44%(0.58)	0.56* 0%(0.45)	0.65*** 2%(0.30)	0.65*** 5%(0.24)	0.48*** 10%(0.21)	0.66*** 3%(0.18)
Beta: $\widehat{\lambda}_{\text{WLS}}$	0.09 −34%(0.58)	0.57* 2%(0.46)	0.67*** 4%(0.30)	0.65*** 5%(0.24)	0.48*** 10%(0.21)	0.66*** 3%(0.18)
GMM_A: $\widehat{b}_1$	0.60 −10%(3.75)	4.86** 93%(2.58)	5.03*** 60%(1.92)	5.06*** 63%(1.71)	3.22*** 53%(1.54)	2.18*** 3%(0.75)
GMM_A: $\widehat{b}_2$	−0.63 −195%(3.40)	5.26*** 108%(2.49)	5.60*** 77%(1.83)	5.27*** 69%(1.60)	3.47*** 65%(1.45)	2.10*** −1%(0.73)
GMM_B: $\widehat{b}_1$	0.79 20%(4.37)	5.12** 103%(2.85)	5.28*** 67%(2.16)	5.43*** 74%(1.92)	3.37*** 60%(1.64)	2.22*** 5%(0.78)
GMM_B: $\widehat{b}_2$	−0.85 −228%(3.92)	5.31*** 111%(2.79)	5.79*** 84%(2.08)	5.59*** 80%(1.81)	3.62*** 71%(1.54)	2.10*** −1%(0.76)
GMM_C: $\widehat{b}$	−1.06 −261%(4.42)	5.57** 121%(2.95)	5.91*** 87%(2.16)	5.65*** 82%(1.93)	3.60*** 71%(1.64)	2.19*** 4%(0.78)

Panel C. N Portfolios formed by the intersections of ME and BE/ME

	T = 60	**T = 120**	**T = 240**	**T = 360**	**T = 480**	**T = 948**
$\lambda = Ef$	0.13	0.56	0.64	0.62	0.44	0.64
$b = \frac{\lambda}{E\left(R^{em2}\right)}$	0.66	2.52	3.15	3.11	2.11	2.11
			N = 6			
Beta: $\widehat{\lambda}_{\text{OLS}}$	-0.05 $-138\%(0.58)$	0.46 $-17\%(0.43)$	0.61^{**} $-5\%(0.29)$	0.60^{***} $-4\%(0.23)$	0.43^{**} $-1\%(0.21)$	0.62^{***} $-4\%(0.18)$
Beta: $\widehat{\lambda}_{\text{GLS}}$	0.06 $-57\%(0.57)$	0.57^{*} $3\%(0.43)$	0.64^{**} $-1\%(0.29)$	0.62^{***} $-1\%(0.23)$	0.45^{**} $3\%(0.21)$	0.65^{***} $1\%(0.18)$
Beta: $\widehat{\lambda}_{\text{WLS}}$	0.06 $-54\%(0.57)$	0.58^{*} $5\%(0.43)$	0.66^{**} $3\%(0.29)$	0.63^{***} $1\%(0.23)$	0.46^{**} $5\%(0.21)$	0.65^{***} $1\%(0.18)$
GMM_{A}: $\widehat{b}_1$	1.06 $60\%(3.53)$	5.19^{**} $106\%(2.55)$	5.36^{***} $70\%(1.94)$	5.20^{***} $67\%(1.52)$	3.59^{***} $70\%(1.16)$	1.48^{**} $-30\%(0.65)$
GMM_{A}: $\widehat{b}_2$	1.62 $144\%(3.30)$	6.37^{***} $153\%(2.50)$	5.00^{***} $59\%(1.92)$	5.15^{***} $66\%(1.51)$	3.69^{***} $75\%(1.15)$	1.29^{**} $-39\%(0.65)$
GMM_{B}: $\widehat{b}_1$	1.27 $91\%(4.22)$	5.59^{**} $122\%(3.01)$	5.69^{***} $80\%(2.24)$	5.65^{***} $81\%(1.79)$	3.79^{***} $80\%(1.27)$	1.51^{**} $-28\%(0.67)$
GMM_{B}: $\widehat{b}_2$	0.96 $45\%(3.97)$	4.81^{*} $91\%(2.96)$	3.87^{**} $23\%(2.19)$	4.57^{***} $47\%(1.77)$	3.53^{***} $67\%(1.27)$	1.26^{**} $-40\%(0.67)$
GMM_{C}: $\widehat{b}$	0.85 $29\%(4.49)$	18.01^{***} $614\%(6.05)$	6.77^{***} $114\%(2.33)$	5.90^{***} $90\%(1.84)$	3.86^{***} $83\%(1.30)$	1.41^{**} $-33\%(0.67)$
			N = 25			
Beta: $\widehat{\lambda}_{\text{OLS}}$	-0.05 $-135\%(0.58)$	0.38 $-31\%(0.43)$	0.58^{***} $-9\%(0.29)$	0.58^{***} $-6\%(0.24)$	0.41^{***} $-7\%(0.21)$	0.63^{***} $-2\%(0.18)$
Beta: $\widehat{\lambda}_{\text{GLS}}$	0.11 $-14\%(0.57)$	0.68^{**} $22\%(0.43)$	0.70^{***} $9\%(0.29)$	0.65^{***} $4\%(0.23)$	0.46^{***} $6\%(0.21)$	0.67^{***} $3\%(0.18)$
Beta: $\widehat{\lambda}_{\text{WLS}}$	0.07 $-50\%(0.57)$	0.55^{*} $-2\%(0.43)$	0.65^{***} $2\%(0.29)$	0.62^{***} $-1\%(0.23)$	0.44^{***} $0\%(0.21)$	0.66^{***} $2\%(0.18)$
GMM_{A}: $\widehat{b}_1$	0.79 $19\%(3.59)$	4.90^{***} $94\%(2.60)$	5.29^{***} $68\%(1.98)$	5.15^{***} $65\%(1.55)$	3.54^{***} $68\%(1.19)$	1.69^{***} $-20\%(0.67)$
GMM_{A}: $\widehat{b}_2$	5.00^{***} $655\%(2.72)$	11.67^{***} $363\%(2.28)$	10.45^{***} $231\%(1.75)$	7.35^{***} $136\%(1.39)$	4.94^{***} $134\%(1.12)$	1.57^{***} $-26\%(0.64)$
GMM_{B}: $\widehat{b}_1$	1.04 $57\%(4.29)$	5.31^{***} $111\%(3.06)$	5.61^{***} $78\%(2.29)$	5.60^{***} $80\%(1.83)$	3.74^{***} $77\%(1.31)$	1.74^{***} $-18\%(0.69)$
GMM_{B}: $\widehat{b}_2$	2.41 $264\%(3.39)$	6.11^{***} $142\%(2.85)$	5.37^{***} $70\%(2.11)$	5.15^{***} $65\%(1.67)$	4.14^{***} $96\%(1.24)$	1.41^{***} $-33\%(0.67)$
GMM_{C}: $\widehat{b}$	$\times$	$\times$	32.13^{***} $918\%(6.36)$	13.60^{***} $337\%(2.69)$	6.27^{***} $197\%(1.50)$	4.64^{***} $120\%(1.02)$
			N = 100			
Beta: $\widehat{\lambda}_{\text{OLS}}$		0.34 $-38\%(0.44)$	0.56^{***} $-12\%(0.29)$	0.58^{***} $-8\%(0.24)$	0.40^{***} $-9\%(0.21)$	0.65^{***} $2\%(0.18)$
Beta: $\widehat{\lambda}_{\text{GLS}}$		0.65^{**} $16\%(0.42)$	0.68^{***} $7\%(0.29)$	0.65^{***} $5\%(0.23)$	0.47^{***} $7\%(0.21)$	0.69^{***} $7\%(0.18)$
Beta: $\widehat{\lambda}_{\text{WLS}}$		0.50^{*} $-11\%(0.43)$	0.63^{***} $-3\%(0.29)$	0.60^{***} $-3\%(0.23)$	0.42^{***} $-4\%(0.21)$	0.67^{***} $4\%(0.18)$
GMM_{A}: $\widehat{b}_1$		4.68^{***} $85\%(2.65)$	5.12^{***} $62\%(2.00)$	5.03^{***} $62\%(1.58)$	3.38^{***} $60\%(1.21)$	1.62^{***} $-23\%(0.70)$
GMM_{A}: $\widehat{b}_2$		21.07^{***} $735\%(1.51)$	15.65^{***} $396\%(1.49)$	10.20^{***} $228\%(1.26)$	5.78^{***} $174\%(1.07)$	1.63^{***} $-23\%(0.59)$
GMM_{B}: $\widehat{b}_1$		5.06^{**} $101\%(3.10)$	5.41^{***} $72\%(2.30)$	5.47^{***} $76\%(1.85)$	3.58^{***} $69\%(1.32)$	1.66^{***} $-21\%(0.72)$
GMM_{B}: $\widehat{b}_2$		5.84^{***} $132\%(2.15)$	4.87^{***} $54\%(1.87)$	4.88^{***} $57\%(1.53)$	3.47^{***} $64\%(1.19)$	1.44^{***} $-32\%(0.61)$
GMM_{C}: $\widehat{b}$		-7.34 $-391\%(3.03)$	$\times$	2.14 $-31\%(2.27)$	12.44^{***} $490\%(2.32)$	2.79^{***} $32\%(0.77)$

Panel D. N Portfolios formed on ME MOM

	T = 60	**T = 120**	**T = 240**	**T = 360**	**T = 480**	**T = 948**
$\lambda = Ef$	0.13	0.56	0.64	0.62	0.44	0.64
$b = \frac{\lambda}{E(R^{em2})}$	0.66	2.52	3.15	3.11	2.11	2.11
			N = 6			
Beta: $\widehat{\lambda}_{\text{OLS}}$	0.17 29%(0.58)	0.50 −11%(0.44)	0.71*** 11%(0.29)	0.69*** 11%(0.24)	0.56*** 28%(0.21)	0.84*** 31%(0.18)
Beta: $\widehat{\lambda}_{\text{GLS}}$	0.12 −11%(0.58)	0.55* −2%(0.43)	0.64** 0%(0.29)	0.62*** −1%(0.24)	0.45** 2%(0.21)	0.72*** 11%(0.18)
Beta: $\widehat{\lambda}_{\text{WLS}}$	0.16 18%(0.58)	0.66* 19%(0.45)	0.80*** 25%(0.30)	0.76*** 22%(0.24)	0.58*** 33%(0.21)	0.85*** 32%(0.18)
GMM_A: $\widehat{b}_1$	1.58 139%(4.05)	2.63 4%(2.92)	2.75* −13%(2.17)	2.94** −5%(1.72)	−0.63 −130%(1.77)	3.88*** 84%(1.05)
GMM_A: $\widehat{b}_2$	−0.91 −237%(3.54)	5.64** 124%(2.66)	6.84*** 117%(2.03)	6.38*** 105%(1.61)	4.01*** 90%(1.44)	4.15*** 96%(1.02)
GMM_B: $\widehat{b}_1$	1.82 174%(4.84)	1.55 −39%(3.18)	1.04 −67%(2.27)	0.75 −76%(1.96)	−1.90 −190%(2.05)	3.83*** 81%(1.07)
GMM_B: $\widehat{b}_2$	−1.91 −388%(4.07)	4.34* 72%(2.85)	3.93** 25%(2.05)	3.67** 18%(1.72)	3.21** 52%(1.57)	3.96*** 87%(1.04)
GMM_C: $\widehat{b}$	−4.12 −722%(5.24)	11.50** 356%(6.01)	×	×	×	×
			N = 25			
Beta: $\widehat{\lambda}_{\text{OLS}}$	0.20 54%(0.58)	0.28 −50%(0.46)	0.57*** −12%(0.30)	0.57*** −8%(0.24)	0.46*** 6%(0.21)	0.71*** 10%(0.18)
Beta: $\widehat{\lambda}_{\text{GLS}}$	0.19 40%(0.57)	0.50* −10%(0.43)	0.64*** 0%(0.29)	0.62*** −1%(0.24)	0.47*** 8%(0.21)	0.67*** 4%(0.18)
Beta: $\widehat{\lambda}_{\text{WLS}}$	0.19 44%(0.58)	0.55* −1%(0.44)	0.75*** 16%(0.29)	0.70*** 13%(0.24)	0.56*** 28%(0.21)	0.83*** 29%(0.18)
GMM_A: $\widehat{b}_1$	2.44 268%(4.06)	3.60* 43%(2.83)	4.23*** 34%(2.14)	4.26*** 37%(1.67)	1.62* −23%(1.31)	2.41*** 14%(0.97)
GMM_A: $\widehat{b}_2$	4.19** 532%(2.78)	6.99*** 177%(2.43)	11.09*** 251%(1.87)	9.91*** 218%(1.49)	4.48*** 112%(1.21)	2.43*** 15%(0.84)
GMM_B: $\widehat{b}_1$	2.86 332%(5.01)	3.31 31%(3.17)	3.56** 13%(2.28)	3.44*** 10%(1.82)	0.75 −64%(1.40)	2.45*** 16%(0.99)
GMM_B: $\widehat{b}_2$	1.20 81%(3.47)	2.59 3%(2.78)	4.71*** 49%(2.15)	4.80*** 54%(1.72)	2.46*** 17%(1.28)	2.00*** −5%(0.85)
GMM_C: $\widehat{b}$	5.79 774%(12.95)	19.04*** 655%(5.75)	25.14*** 697%(9.37)	×	×	×

Panel E. N Industry Portfolios

	T = 60	**T = 120**	**T = 240**	**T = 360**	**T = 480**	**T = 948**
$\lambda = Ef$	0.13	0.56	0.64	0.62	0.44	0.64
$b = \frac{\lambda}{E(R^{em2})}$	0.66	2.52	3.15	3.11	2.11	2.11
N = 5						
Beta: $\widehat{\lambda}_{\text{OLS}}$	0.10 −24%(0.58)	0.56 1%(0.43)	0.63** −2%(0.29)	0.66** 6%(0.24)	0.49** 12%(0.21)	0.71*** 10%(0.18)
Beta: $\widehat{\lambda}_{\text{GLS}}$	0.12 −13%(0.57)	0.60 7%(0.43)	0.65** 1%(0.29)	0.65** 5%(0.24)	0.48** 10%(0.21)	0.69*** 7%(0.18)
Beta: $\widehat{\lambda}_{\text{WLS}}$	0.11 −19%(0.58)	0.58 4%(0.44)	0.64** −1%(0.30)	0.66** 6%(0.24)	0.49** 12%(0.21)	0.70*** 8%(0.18)
GMM_{A}: $\widehat{b}_1$	1.81 173%(3.81)	4.34* 72%(2.67)	4.15** 32%(1.91)	4.92*** 58%(1.65)	3.99** 89%(1.66)	4.42*** 109%(1.43)
GMM_{A}: $\widehat{b}_2$	2.34 253%(3.64)	4.78* 90%(2.64)	4.43** 40%(1.87)	4.79*** 54%(1.61)	3.62** 71%(1.52)	3.64*** 72%(1.17)
GMM_{B}: $\widehat{b}_1$	2.15 225%(4.57)	4.41* 75%(2.99)	4.21** 33%(2.07)	4.99** 60%(1.76)	4.01** 90%(1.71)	4.43*** 110%(1.43)
GMM_{B}: $\widehat{b}_2$	2.71 309%(4.38)	4.87* 93%(2.96)	4.43** 40%(2.04)	4.86** 56%(1.73)	3.63** 72%(1.57)	3.65*** 73%(1.18)
GMM_{C}: $\widehat{b}$	2.76 316%(4.66)	4.97* 97%(3.07)	4.48** 42%(2.08)	4.87** 56%(1.75)	3.66** 74%(1.68)	3.77** 78%(1.36)
N = 17						
Beta: $\widehat{\lambda}_{\text{OLS}}$	0.06 −52%(0.58)	0.55* −1%(0.44)	0.66*** 3%(0.29)	0.65*** 5%(0.24)	0.50*** 15%(0.21)	0.72*** 12%(0.18)
Beta: $\widehat{\lambda}_{\text{GLS}}$	0.16 19%(0.57)	0.56* 1%(0.43)	0.68*** 7%(0.29)	0.70*** 12%(0.23)	0.51*** 17%(0.21)	0.72*** 12%(0.18)
Beta: $\widehat{\lambda}_{\text{WLS}}$	0.10 −23%(0.58)	0.49* −12%(0.43)	0.64*** 0%(0.29)	0.66*** 7%(0.24)	0.49*** 12%(0.21)	0.72*** 11%(0.18)
GMM_{A}: $\widehat{b}_1$	1.99 200%(3.79)	3.97 58%(2.59)	4.35*** 38%(1.94)	4.83*** 55%(1.57)	3.61*** 71%(1.28)	3.25*** 54%(0.75)
GMM_{A}: $\widehat{b}_2$	3.94* 495%(3.12)	4.82*** 91%(2.44)	4.46*** 41%(1.75)	4.71*** 51%(1.39)	3.41*** 61%(1.19)	3.08*** 46%(0.72)
GMM_{B}: $\widehat{b}_1$	2.54 284%(4.52)	4.02** 59%(2.83)	4.43*** 41%(2.12)	4.88*** 57%(1.71)	3.62*** 72%(1.34)	3.29*** 56%(0.78)
GMM_{B}: $\widehat{b}_2$	1.91 189%(3.80)	3.55* 41%(2.67)	4.22*** 34%(1.92)	4.40*** 41%(1.52)	3.19*** 51%(1.25)	3.06*** 45%(0.74)
GMM_{C}: $\widehat{b}$	15.13** 2184%(9.72)	4.45** 76%(3.04)	4.54*** 44%(2.14)	4.59*** 47%(1.69)	3.42*** 62%(1.34)	3.18*** 50%(0.77)
N = 30						
Beta: $\widehat{\lambda}_{\text{OLS}}$	0.06 −55%(0.58)	0.57** 2%(0.43)	0.69*** 8%(0.29)	0.73*** 17%(0.24)	0.56*** 29%(0.21)	0.77*** 20%(0.18)
Beta: $\widehat{\lambda}_{\text{GLS}}$	0.07 −44%(0.57)	0.62** 11%(0.43)	0.72*** 11%(0.29)	0.73*** 17%(0.24)	0.53*** 22%(0.21)	0.74*** 15%(0.18)
Beta: $\widehat{\lambda}_{\text{WLS}}$	0.12 −10%(0.58)	0.59** 6%(0.43)	0.69*** 8%(0.29)	0.73*** 17%(0.24)	0.55*** 25%(0.21)	0.75*** 16%(0.18)
GMM_{A}: $\widehat{b}_1$	1.72 160%(3.29)	4.69*** 86%(2.66)	4.68*** 48%(2.01)	5.06*** 63%(1.59)	3.85*** 82%(1.27)	3.28*** 55%(0.75)
GMM_{A}: $\widehat{b}_2$	3.80** 474%(2.57)	6.21*** 146%(2.33)	4.75*** 51%(1.73)	5.14*** 65%(1.38)	3.52*** 67%(1.15)	3.16*** 50%(0.72)
GMM_{B}: $\widehat{b}_1$	2.48 274%(4.07)	4.79** 90%(2.99)	4.77*** 51%(2.21)	5.11*** 64%(1.74)	3.86*** 83%(1.34)	3.33*** 57%(0.78)
GMM_{B}: $\widehat{b}_2$	0.67 1%(3.32)	3.52** 40%(2.67)	3.82*** 21%(1.91)	4.47*** 44%(1.51)	3.17*** 50%(1.21)	3.05*** 45%(0.74)
GMM_{C}: $\widehat{b}$	−0.70 −206%(4.51)	4.95** 96%(3.49)	4.59*** 45%(2.21)	4.63*** 49%(1.73)	3.15*** 49%(1.33)	3.33*** 58%(0.78)

3.6.3 Size risk premium in Fama-French model

The following five panels report the estimate values, the bias from the risk premium in percent values, the standard error in parenthesis, and t-statistics. Each panel corresponds to a one set of test portfolios. T is 60 (January 2001 - December 2005), 120, 240, 360, 480 and 948 (January 1927 - December 2005) monthly observations. The two-pass (Fama-MacBeth) cross-sectional estimate is the slope coefficient λ in $E\left(R^{e}\right) = \beta\lambda$ calculated by OLS, GLS and WLS. Next, we turn from beta representation to a discount factor formulation for GMM approach. The $\mathrm{GMM_A}$ and $\mathrm{GMM_B}$ first and second-stage estimates are the parameters $100 \times b$ in $E\left(R^{e}\right) = E\left(R^{e} f'\right) b$ (returns on second moments, following Hansen and Jagannathan [49]) and $E\left(R^{e}\right) = E\left(R^{e} \widetilde{f}'\right) b$ (returns on covariances, following Cochrane [25]) respectively. The $\mathrm{GMM_C}$ is the continuous updating estimate (following Hansen, Heaton and Yaron [47]).

Panel A. N Portfolios formed on ME

	T = 60	T = 120	T = 240	T = 360	T = 480	T = 948
$\lambda = Ef$	0.80	0.29	0.06	0.29	0.25	0.25
$b = \frac{\lambda}{E(R^{em2})}$	8.14	1.50	0.49	2.77	2.24	2.16
			N = 5			
Beta: $\widehat{\lambda}_{\text{OLS}}$	0.44 −44%(0.49)	0.62 117%(0.48)	0.06 −2%(0.33)	0.23 −21%(0.24)	0.16 −36%(0.22)	0.27 10%(0.22)
Beta: $\widehat{\lambda}_{\text{GLS}}$	0.49 −39%(0.45)	0.35 23%(0.46)	−0.14 −341%(0.31)	0.17 −41%(0.22)	0.14 −41%(0.20)	0.23 −8%(0.19)
Beta: $\widehat{\lambda}_{\text{WLS}}$	0.37 −53%(0.55)	0.55 93%(0.45)	−0.02 −132%(0.29)	0.21 −28%(0.21)	0.15 −38%(0.19)	0.28 15%(0.22)
GMM_A: $\widehat{b}_1$	5.30 −35%(7.12)	−0.07 −105%(3.68)	0.84 74%(3.21)	2.71 −2%(2.05)	1.77 −21%(1.49)	1.33 −39%(2.11)
GMM_A: $\widehat{b}_2$	6.73 −17%(6.65)	4.59* 205%(3.04)	2.54 424%(2.99)	3.02* 9%(1.99)	1.67 −26%(1.48)	0.80 −63%(1.82)
GMM_B: $\widehat{b}_1$	7.99 −2%(8.95)	−0.53 −135%(3.71)	0.55 14%(3.45)	2.88 4%(2.35)	1.88 −16%(1.61)	1.35 −37%(2.13)
GMM_B: $\widehat{b}_2$	8.52 5%(8.32)	3.96 163%(3.17)	2.54 425%(3.20)	3.28* 18%(2.26)	1.76 −21%(1.60)	0.83 −62%(1.84)
GMM_C: $\widehat{b}$	7.07 −13%(10.38)	0.66 −56%(4.07)	2.11 336%(4.32)	3.21 16%(2.51)	1.75 −22%(1.64)	0.78 −64%(2.12)
			N = 10			
Beta: $\widehat{\lambda}_{\text{OLS}}$	0.54* −32%(0.46)	0.52* 80%(0.45)	0.02 −68%(0.26)	0.24* −17%(0.20)	0.20* −20%(0.17)	0.14 −42%(0.14)
Beta: $\widehat{\lambda}_{\text{GLS}}$	0.74** −7%(0.42)	0.17 −41%(0.42)	−0.03 −144%(0.24)	0.26* −11%(0.18)	0.22* −9%(0.16)	0.20** −18%(0.13)
Beta: $\widehat{\lambda}_{\text{WLS}}$	0.41 −48%(0.44)	0.32 10%(0.42)	−0.04 −167%(0.24)	0.22* −23%(0.18)	0.19* −22%(0.16)	0.19* −22%(0.14)
GMM_A: $\widehat{b}_1$	5.14 −37%(7.00)	1.38 −9%(2.98)	1.68 247%(2.46)	2.78** 0%(1.84)	1.77* −21%(1.45)	0.01 −99%(1.23)
GMM_A: $\widehat{b}_2$	6.13* −25%(4.97)	4.85** 222%(2.59)	1.77 265%(2.37)	2.85** 3%(1.83)	1.86* −17%(1.44)	0.40 −82%(1.11)
GMM_B: $\widehat{b}_1$	8.74* 7%(7.73)	0.57 −62%(3.23)	1.43 195%(2.61)	2.86* 3%(2.02)	1.83* −18%(1.53)	0.02 −99%(1.26)
GMM_B: $\widehat{b}_2$	9.05** 11%(6.05)	3.67* 144%(2.74)	1.22 151%(2.53)	2.62* −5%(2.01)	1.86* −17%(1.51)	0.41 −81%(1.14)
GMM_C: $\widehat{b}$	−12.02 −248%(28.74)	7.02** 366%(4.26)	2.04 321%(3.01)	3.04** 10%(2.05)	1.92* −14%(1.52)	0.35 −84%(1.25)

Panel B. N Portfolios formed on BE/ME

	T = 60	**T = 120**	**T = 240**	**T = 360**	**T = 480**	**T = 948**
$\lambda = Ef$	0.80	0.29	0.06	0.29	0.25	0.25
$b = \frac{\lambda}{E(R^{em2})}$	8.14	1.50	0.49	2.77	2.24	2.16
			N = 5			
Beta: $\widehat{\lambda}_{\text{OLS}}$	0.85 6%(0.79)	−0.91 −418%(2.49)	−0.08 −241%(1.48)	0.52 79%(0.85)	0.33 35%(0.53)	0.06 −75%(0.33)
Beta: $\widehat{\lambda}_{\text{GLS}}$	0.73 −9%(0.71)	−0.32 −211%(2.11)	−0.43 −811%(1.30)	0.55 92%(0.78)	0.22 −12%(0.51)	0.08 −69%(0.33)
Beta: $\widehat{\lambda}_{\text{WLS}}$	0.83 5%(0.83)	−0.89 −410%(3.00)	−0.13 −312%(1.44)	0.52 81%(0.69)	0.32 30%(0.43)	0.05 −78%(0.27)
GMM_A: $\widehat{b}_1$	9.60 18%(7.10)	−3.12 −307%(14.43)	0.13 −73%(12.39)	5.51 99%(7.84)	3.06 37%(4.71)	−0.74 −134%(3.02)
GMM_A: $\widehat{b}_2$	8.95 10%(6.77)	1.03 −32%(12.28)	−2.08 −528%(11.07)	5.80 110%(7.07)	2.18 −3%(4.55)	−0.44 −120%(2.99)
GMM_B: $\widehat{b}_1$	11.07 36%(8.98)	−3.94 −362%(15.35)	0.08 −84%(13.01)	5.95 115%(8.59)	3.19 43%(4.94)	−0.75 −135%(3.07)
GMM_B: $\widehat{b}_2$	10.24 26%(8.60)	0.89 −41%(12.83)	−2.24 −562%(11.62)	6.27 126%(7.75)	2.26 1%(4.78)	−0.46 −121%(3.04)
GMM_C: $\widehat{b}$	10.46 28%(9.04)	−0.18 −112%(15.02)	0.07 −85%(13.08)	6.28 127%(8.59)	2.32 4%(4.96)	−0.46 −122%(3.06)
			N = 10			
Beta: $\widehat{\lambda}_{\text{OLS}}$	0.87 10%(0.80)	−0.38 −233%(1.90)	−0.23 −479%(0.97)	0.14 −52%(0.60)	0.28 12%(0.46)	0.00 −99%(0.27)
Beta: $\widehat{\lambda}_{\text{GLS}}$	0.57 −28%(0.64)	−0.35 −223%(1.53)	×	0.06 −79%(0.54)	0.21 −14%(0.43)	−0.04 −118%(0.27)
Beta: $\widehat{\lambda}_{\text{WLS}}$	0.83 5%(0.75)	−0.40 −239%(1.71)	−0.38 −732%(0.85)	0.12 −58%(0.49)	0.29 18%(0.36)	0.03 −89%(0.22)
GMM_A: $\widehat{b}_1$	10.11* 24%(7.55)	−0.31 −120%(11.54)	−1.13 −332%(8.45)	1.97 −29%(5.69)	2.51 12%(4.16)	−1.30 −160%(2.60)
GMM_A: $\widehat{b}_2$	9.31** 14%(5.76)	0.57 −62%(9.03)	−4.02 −929%(7.31)	1.65 −41%(5.08)	2.18 −3%(3.80)	−1.75 −181%(2.52)
GMM_B: $\widehat{b}_1$	11.31* 39%(9.42)	−0.85 −157%(12.21)	−1.21 −350%(8.85)	2.08 −25%(6.12)	2.62 17%(4.36)	−1.32 −161%(2.64)
GMM_B: $\widehat{b}_2$	9.70* 19%(7.37)	0.03 −98%(9.54)	−4.29 −984%(7.65)	1.61 −41%(5.47)	2.26 1%(3.99)	−1.81 −184%(2.57)
GMM_C: $\widehat{b}$	11.42* 40%(9.53)	0.57 −62%(12.23)	−4.26 −978%(9.40)	1.65 −40%(6.16)	2.45 10%(4.37)	−1.96 −191%(2.67)

Panel C. N Portfolios formed by the intersections of ME and BE/ME

	T = 60	**T = 120**	**T = 240**	**T = 360**	**T = 480**	**T = 948**
$\lambda = Ef$	0.80	0.29	0.06	0.29	0.25	0.25
$b = \frac{\lambda}{E\left(R^{em2}\right)}$	8.14	1.50	0.49	2.77	2.24	2.16
				N = 6		
Beta: $\widehat{\lambda}_{\mathrm{OLS}}$	0.80^{**} 1%(0.39)	0.23 −18%(0.40)	0.02 −69%(0.23)	0.26^{*} −11%(0.17)	0.22^{*} −11%(0.15)	0.23^{**} −8%(0.11)
Beta: $\widehat{\lambda}_{\mathrm{GLS}}$	0.80^{**} 0%(0.39)	0.29 0%(0.40)	0.06 0%(0.23)	0.29^{**} 0%(0.17)	0.25^{*} 0%(0.15)	0.25^{**} 0%(0.11)
Beta: $\widehat{\lambda}_{\mathrm{WLS}}$	0.79^{**} −1%(0.40)	0.23 −20%(0.40)	0.02 −75%(0.23)	0.26^{*} −9%(0.17)	0.22^{*} −9%(0.15)	0.22^{**} −9%(0.11)
$\mathrm{GMM_A}$: $\widehat{b}_1$	9.47^{***} 16%(3.93)	3.97^{*} 163%(2.46)	2.00 312%(2.14)	3.73^{**} 35%(1.79)	2.39^{**} 7%(1.45)	0.93 −57%(0.92)
$\mathrm{GMM_A}$: $\widehat{b}_2$	11.44^{***} 41%(3.68)	4.64^{**} 208%(2.45)	1.39 186%(2.14)	3.80^{**} 37%(1.79)	2.81^{**} 25%(1.44)	1.27^{*} −41%(0.92)
$\mathrm{GMM_B}$: $\widehat{b}_1$	11.11^{**} 36%(5.57)	4.22^{*} 180%(2.80)	2.08 329%(2.31)	4.02^{**} 45%(2.01)	2.51^{*} 12%(1.55)	0.95 −56%(0.94)
$\mathrm{GMM_B}$: $\widehat{b}_2$	11.68^{**} 43%(5.33)	3.56^{*} 136%(2.79)	1.01 107%(2.29)	3.41^{**} 23%(2.00)	2.72^{**} 22%(1.54)	1.26^{*} −42%(0.94)
$\mathrm{GMM_C}$: $\widehat{b}$	13.75^{**} 69%(6.04)	×	−0.29 −160%(2.66)	4.22^{**} 52%(2.07)	3.09^{**} 38%(1.57)	1.22^{*} −44%(0.93)
				N = 25		
Beta: $\widehat{\lambda}_{\mathrm{OLS}}$	0.91^{***} 14%(0.40)	0.41 42%(0.40)	0.08 39%(0.23)	0.28^{**} −2%(0.17)	0.24^{**} −3%(0.15)	0.11 −54%(0.12)
Beta: $\widehat{\lambda}_{\mathrm{GLS}}$	0.85^{***} 7%(0.39)	0.26 −8%(0.40)	0.07 9%(0.23)	0.29^{**} 1%(0.17)	0.25^{**} 1%(0.15)	0.24^{***} −2%(0.11)
Beta: $\widehat{\lambda}_{\mathrm{WLS}}$	0.86^{***} 8%(0.40)	0.34 17%(0.40)	0.09 47%(0.23)	0.31^{***} 7%(0.17)	0.25^{**} 3%(0.15)	0.23^{***} −8%(0.11)
$\mathrm{GMM_A}$: $\widehat{b}_1$	10.71^{***} 32%(3.93)	5.22^{***} 247%(2.47)	2.69^{*} 455%(2.17)	4.11^{***} 48%(1.81)	2.73^{***} 22%(1.48)	−0.20 −109%(0.97)
$\mathrm{GMM_A}$: $\widehat{b}_2$	19.05^{***} 134%(3.04)	9.45^{***} 528%(2.23)	4.60^{***} 848%(2.00)	5.40^{***} 95%(1.71)	3.26^{***} 46%(1.41)	1.11^{*} −49%(0.88)
$\mathrm{GMM_B}$: $\widehat{b}_1$	12.38^{***} 52%(5.81)	5.52^{***} 266%(2.89)	2.79^{*} 475%(2.35)	4.40^{***} 59%(2.04)	2.85^{***} 28%(1.58)	−0.20 −109%(0.99)
$\mathrm{GMM_B}$: $\widehat{b}_2$	13.06^{***} 60%(4.94)	5.10^{***} 239%(2.69)	2.32^{*} 379%(2.18)	3.81^{***} 38%(1.94)	2.79^{***} 25%(1.52)	1.06^{*} −51%(0.90)
$\mathrm{GMM_C}$: $\widehat{b}$	70.04^{***} 760%(25.20)	−4.19 −378%(9.79)	×	−6.67 −341%(3.96)	1.18 −47%(1.94)	1.02 −53%(1.57)
				N = 100		
Beta: $\widehat{\lambda}_{\mathrm{OLS}}$		0.51^{*} 77%(0.41)	0.14 139%(0.23)	0.32^{***} 12%(0.18)	0.27^{***} 9%(0.16)	0.19^{**} −22%(0.12)
Beta: $\widehat{\lambda}_{\mathrm{GLS}}$		0.32 13%(0.40)	0.10 62%(0.23)	0.32^{***} 12%(0.17)	0.27^{***} 11%(0.15)	0.23^{***} −5%(0.11)
Beta: $\widehat{\lambda}_{\mathrm{WLS}}$		0.43^{*} 50%(0.40)	0.14 132%(0.23)	0.34^{***} 17%(0.17)	0.29^{***} 16%(0.15)	0.24^{***} −2%(0.11)
$\mathrm{GMM_A}$: $\widehat{b}_1$		5.93^{***} 294%(2.46)	3.23^{**} 565%(2.17)	4.52^{***} 63%(1.82)	3.01^{***} 35%(1.49)	0.50 −77%(0.93)
$\mathrm{GMM_A}$: $\widehat{b}_2$		×	×	7.97^{***} 188%(1.55)	4.54^{***} 103%(1.34)	1.29^{***} −40%(0.71)
$\mathrm{GMM_B}$: $\widehat{b}_1$		6.18^{***} 310%(2.92)	3.30^{**} 580%(2.36)	4.80^{***} 73%(2.06)	3.12^{***} 39%(1.59)	0.51 −76%(0.95)
$\mathrm{GMM_B}$: $\widehat{b}_2$		6.70^{***} 345%(2.28)	2.89^{**} 495%(1.99)	4.48^{***} 62%(1.77)	3.09^{***} 38%(1.44)	1.22^{***} −43%(0.72)
$\mathrm{GMM_C}$: $\widehat{b}$		−1.03 −169%(3.57)	×	5.97^{***} 116%(3.43)	9.83^{***} 340%(2.39)	0.72 −66%(0.99)

Panel D. N Portfolios formed on ME MOM

	T = 60	**T = 120**	**T = 240**	**T = 360**	**T = 480**	**T = 948**
$\lambda = Ef$	0.80	0.29	0.06	0.29	0.25	0.25
$b = \frac{\lambda}{E(R^{em2})}$	8.14	1.50	0.49	2.77	2.24	2.16
			N = 6			
Beta: $\widehat{\lambda}_{\text{OLS}}$	0.90^{**} 13%(0.44)	0.71^{**} 147%(0.43)	0.24 305%(0.24)	0.47^{***} 62%(0.18)	0.48^{***} 94%(0.17)	0.71^{***} 188%(0.14)
Beta: $\widehat{\lambda}_{\text{GLS}}$	1.05^{***} 32%(0.41)	0.55^{*} 91%(0.41)	0.25 319%(0.24)	0.47^{***} 62%(0.18)	0.40^{***} 64%(0.16)	0.43^{***} 73%(0.12)
Beta: $\widehat{\lambda}_{\text{WLS}}$	1.03^{***} 30%(0.43)	0.80^{**} 180%(0.43)	0.44^{**} 637%(0.25)	0.63^{***} 120%(0.19)	0.61^{***} 149%(0.17)	0.72^{***} 194%(0.13)
GMM_A: $\widehat{b}_1$	9.89^{**} 21%(4.38)	4.51^{**} 200%(2.59)	0.48 0%(2.36)	3.50^{**} 26%(1.84)	1.64 −27%(1.76)	5.85^{**} 171%(2.49)
GMM_A: $\widehat{b}_2$	13.60^{***} 67%(3.98)	7.03^{***} 367%(2.47)	3.95^{**} 714%(2.28)	7.34^{***} 165%(1.76)	4.03^{***} 80%(1.70)	0.40 −81%(1.90)
GMM_B: $\widehat{b}_1$	11.58^{**} 42%(5.91)	3.32 121%(2.86)	−1.35 −378%(2.42)	1.71 −38%(2.03)	1.09 −51%(1.93)	5.77^{***} 167%(2.42)
GMM_B: $\widehat{b}_2$	13.10^{**} 61%(5.77)	5.74^{**} 281%(2.76)	1.66 243%(2.37)	4.21^{**} 52%(1.99)	2.19 −2%(1.88)	0.58 −73%(1.87)
GMM_C: $\widehat{b}$	20.06^{***} 146%(6.43)	−2.18 −245%(4.59)	×	×	×	×
			N = 25			
Beta: $\widehat{\lambda}_{\text{OLS}}$	0.92^{***} 15%(0.46)	0.82^{***} 185%(0.44)	0.25 322%(0.25)	0.45^{***} 56%(0.18)	0.40^{***} 64%(0.16)	0.87^{***} 252%(0.15)
Beta: $\widehat{\lambda}_{\text{GLS}}$	0.94^{***} 19%(0.40)	0.44^{*} 54%(0.41)	0.19 215%(0.23)	0.38^{***} 32%(0.18)	0.32^{***} 30%(0.16)	0.34^{***} 38%(0.12)
Beta: $\widehat{\lambda}_{\text{WLS}}$	0.95^{***} 19%(0.41)	0.80^{***} 178%(0.41)	0.38^{**} 544%(0.24)	0.59^{***} 106%(0.18)	0.54^{***} 118%(0.16)	0.73^{***} 195%(0.12)
GMM_A: $\widehat{b}_1$	10.18^{***} 25%(4.61)	7.27^{***} 383%(2.65)	3.30^{**} 581%(2.47)	5.42^{***} 95%(1.93)	2.91^{***} 30%(1.50)	7.40^{***} 243%(2.24)
GMM_A: $\widehat{b}_2$	19.70^{***} 142%(3.27)	12.93^{***} 759%(2.08)	×	10.28^{***} 271%(1.71)	4.80^{***} 115%(1.43)	0.59 −73%(1.43)
GMM_B: $\widehat{b}_1$	12.00^{***} 47%(6.32)	6.93^{***} 361%(3.08)	2.39 394%(2.54)	4.57^{***} 65%(2.06)	2.51^{**} 12%(1.59)	7.47^{***} 246%(2.26)
GMM_B: $\widehat{b}_2$	16.53^{***} 103%(5.00)	7.19^{***} 378%(2.68)	3.33^{**} 587%(2.30)	6.41^{***} 131%(1.91)	2.79^{***} 25%(1.50)	0.37 −83%(1.44)
GMM_C: $\widehat{b}$	63.97^{***} 686%(17.11)	−6.74 −548%(5.27)	×	×	×	×

Panel E. N Industry Portfolios

	T = 60	**T = 120**	**T = 240**	**T = 360**	**T = 480**	**T = 948**
$\lambda = Ef$	0.80	0.29	0.06	0.29	0.25	0.25
$b = \frac{\lambda}{E\left(R^{em2}\right)}$	8.14	1.50	0.49	2.77	2.24	2.16
			N = 5			
Beta: $\widehat{\lambda}_{\text{OLS}}$	0.82 3%(0.76)	−0.61 −314%(1.09)	×	−0.65 −325%(0.62)	−0.55 −322%(0.50)	−0.81 −430%(0.58)
Beta: $\widehat{\lambda}_{\text{GLS}}$	0.80 1%(0.76)	−0.14 −149%(0.96)	×	−0.59 −305%(0.60)	−0.43 −276%(0.45)	−0.47 −292%(0.49)
Beta: $\widehat{\lambda}_{\text{WLS}}$	0.75 −6%(0.73)	−0.12 −141%(0.89)	×	−0.62 −315%(0.57)	−0.44 −277%(0.42)	−0.45 −282%(0.43)
GMM_A: $\widehat{b}_1$	9.36 15%(7.42)	−2.73 −281%(7.95)	×	−7.46 −369%(7.40)	−6.43 −387%(5.25)	−9.34 −533%(6.10)
GMM_A: $\widehat{b}_2$	9.19 13%(7.42)	1.11 −26%(6.81)	×	−6.64 −340%(7.05)	−5.05 −326%(4.58)	−5.72 −365%(5.17)
GMM_B: $\widehat{b}_1$	10.99 35%(9.83)	−3.06 −303%(8.17)	×	−7.57 −373%(7.33)	−6.47 −389%(5.25)	−9.40 −536%(6.09)
GMM_B: $\widehat{b}_2$	10.40 28%(9.80)	0.83 −45%(7.04)	×	−6.76 −344%(7.00)	−5.08 −327%(4.58)	−5.77 −367%(5.16)
GMM_C: $\widehat{b}$	11.24 38%(10.04)	0.45 −70%(8.02)	×	−6.80 −345%(7.24)	−5.16 −331%(5.14)	−6.34 −394%(5.83)
			N = 17			
Beta: $\widehat{\lambda}_{\text{OLS}}$	0.58* −27%(0.53)	−0.38 −231%(0.67)	×	−0.59 −304%(0.33)	−0.35 −241%(0.26)	−0.23 −193%(0.18)
Beta: $\widehat{\lambda}_{\text{GLS}}$	0.50 −38%(0.47)	−0.83 −388%(0.56)	×	−0.38 −233%(0.27)	−0.24 −197%(0.23)	−0.18 −171%(0.17)
Beta: $\widehat{\lambda}_{\text{WLS}}$	0.59* −26%(0.48)	−0.68 −339%(0.54)	×	−0.64 −324%(0.26)	−0.40 −263%(0.23)	−0.27 −208%(0.17)
GMM_A: $\widehat{b}_1$	7.88** −3%(5.65)	−1.26 −184%(4.31)	×	−6.51 −335%(3.72)	−4.35 −295%(2.66)	−3.63 −268%(1.78)
GMM_A: $\widehat{b}_2$	12.40*** 52%(4.55)	−5.11 −439%(3.64)	×	−5.20 −288%(3.07)	−3.90 −275%(2.32)	−3.06 −242%(1.64)
GMM_B: $\widehat{b}_1$	8.11* 0%(7.28)	−1.79 −219%(4.35)	×	−6.90 −349%(3.74)	−4.51 −302%(2.67)	−3.67 −270%(1.80)
GMM_B: $\widehat{b}_2$	6.45 −21%(6.09)	−5.85 −489%(3.63)	×	−5.60 −302%(3.09)	−4.01 −280%(2.33)	−3.09 −243%(1.65)
GMM_C: $\widehat{b}$	−7.06 −187%(16.68)	−8.09 −638%(4.63)	×	−6.56 −337%(3.76)	−4.51 −302%(2.69)	−3.35 −255%(1.79)
			N = 30			
Beta: $\widehat{\lambda}_{\text{OLS}}$	0.76** −5%(0.55)	0.10 −65%(0.63)	−0.45 −851%(0.39)	−0.56 −295%(0.30)	−0.30 −223%(0.22)	−0.06 −126%(0.16)
Beta: $\widehat{\lambda}_{\text{GLS}}$	0.81*** 2%(0.45)	−0.54 −289%(0.48)	−0.37 −713%(0.30)	−0.26 −190%(0.22)	−0.17 −168%(0.19)	−0.15 −161%(0.15)
Beta: $\widehat{\lambda}_{\text{WLS}}$	0.91*** 15%(0.48)	−0.43 −249%(0.50)	−0.53 −992%(0.29)	−0.51 −277%(0.23)	−0.28 −216%(0.20)	−0.18 −172%(0.15)
GMM_A: $\widehat{b}_1$	10.30** 26%(6.34)	3.12 107%(4.04)	−3.67 −858%(3.77)	−6.43 −332%(3.41)	−3.94 −276%(2.26)	−2.04 −194%(1.61)
GMM_A: $\widehat{b}_2$	17.25*** 112%(3.73)	0.93 −38%(3.15)	−4.07 −940%(2.89)	−3.74 −235%(2.51)	−2.88 −229%(1.93)	−2.95 −237%(1.44)
GMM_B: $\widehat{b}_1$	10.38* 28%(8.78)	1.97 31%(4.20)	−4.04 −932%(3.88)	−6.86 −348%(3.45)	−4.18 −287%(2.29)	−2.09 −197%(1.62)
GMM_B: $\widehat{b}_2$	10.23*** 26%(5.72)	−2.16 −244%(3.29)	×	−4.23 −253%(2.54)	−3.06 −237%(1.95)	−2.96 −237%(1.45)
GMM_C: $\widehat{b}$	12.88* 58%(10.00)	−9.35 −721%(5.34)	×	−5.50 −299%(3.43)	−3.47 −255%(2.28)	−3.78 −275%(1.74)

3.6.4 Value risk premium in Fama-French model

The following five panels report the estimate values, the bias from the risk premium in percent values, the standard error in parenthesis, and t-statistics. Each panel corresponds to a one set of test portfolios. T is 60 (January 2001 - December 2005), 120, 240, 360, 480 and 948 (January 1927 - December 2005) monthly observations. The two-pass (Fama-MacBeth) cross-sectional estimate is the slope coefficient λ in $E\left(R^e\right) = \beta\lambda$ calculated by OLS, GLS and WLS. Next, we turn from beta representation to a discount factor formulation for GMM approach. The $\mathrm{GMM_A}$ and $\mathrm{GMM_B}$ first and second-stage estimates are the parameters $100 \times b$ in $E\left(R^e\right) = E\left(R^e f'\right) b$ (returns on second moments, following Hansen and Jagannathan [49]) and $E\left(R^e\right) = E\left(R^e \widetilde{f}'\right) b$ (returns on covariances, following Cochrane [25]) respectively. The $\mathrm{GMM_C}$ is the continuous updating estimate (following Hansen, Heaton and Yaron [47]).

Panel A. N Portfolios formed on ME

	T = 60	**T = 120**	**T = 240**	**T = 360**	**T = 480**	**T = 948**
$\lambda = Ef$	0.78	0.46	0.34	0.43	0.43	0.41
$b = \frac{\lambda}{E(R^{em2})}$	7.50	2.90	3.26	4.44	4.70	3.10
			N = 5			
Beta: $\widehat{\lambda}_{\text{OLS}}$	2.10^{**} 168%(1.01)	-1.22 -366%(1.05)	-0.04 -112%(1.07)	0.23 -46%(0.92)	0.46 7%(0.84)	-0.00 -101%(0.58)
Beta: $\widehat{\lambda}_{\text{GLS}}$	1.73^{**} 121%(0.89)	-0.24 -152%(0.92)	0.80 135%(0.97)	0.52 22%(0.81)	0.52 20%(0.71)	0.16 -62%(0.47)
Beta: $\widehat{\lambda}_{\text{WLS}}$	1.99^{*} 154%(1.25)	-0.93 -303%(0.77)	0.26 -25%(0.76)	0.32 -26%(0.66)	0.48 11%(0.62)	-0.00 -101%(0.59)
GMM_A: $\widehat{b}_1$	25.84^{**} 245%(9.23)	-7.37 -354%(11.56)	4.22 29%(13.10)	6.72 51%(10.95)	8.02 71%(10.02)	-0.82 -126%(4.42)
GMM_A: $\widehat{b}_2$	23.09^{**} 208%(8.79)	7.38 155%(9.41)	14.27 337%(11.10)	10.01 125%(9.48)	8.48 80%(8.70)	0.43 -86%(3.64)
GMM_B: $\widehat{b}_1$	28.95^{**} 286%(14.41)	-8.78 -403%(11.11)	2.86 -12%(14.43)	6.98 57%(12.44)	8.41 79%(11.17)	-0.85 -127%(4.48)
GMM_B: $\widehat{b}_2$	28.69^{**} 283%(13.30)	4.67 61%(9.58)	14.10 332%(12.40)	10.72 141%(10.84)	8.81 88%(9.66)	0.41 -87%(3.70)
GMM_C: $\widehat{b}$	35.79^{**} 377%(17.20)	11.57 300%(11.91)	18.33 462%(17.64)	11.25 153%(13.02)	9.44 101%(11.31)	0.44 -86%(4.44)
			N = 10			
Beta: $\widehat{\lambda}_{\text{OLS}}$	1.76^{***} 125%(0.74)	-0.81 -276%(0.85)	0.22 -35%(0.54)	0.19 -57%(0.51)	0.26 -40%(0.43)	0.36 -12%(0.33)
Beta: $\widehat{\lambda}_{\text{GLS}}$	0.94^{**} 20%(0.63)	0.44 -4%(0.62)	0.30 -12%(0.43)	0.04 -90%(0.43)	0.14 -66%(0.38)	0.23 -45%(0.25)
Beta: $\widehat{\lambda}_{\text{WLS}}$	1.86^{***} 137%(0.72)	-0.07 -115%(0.53)	0.36 6%(0.37)	0.25 -41%(0.35)	0.28 -34%(0.32)	0.26 -37%(0.28)
GMM_A: $\widehat{b}_1$	28.93^{***} 286%(7.83)	-1.59 -155%(8.16)	7.62^{*} 134%(6.50)	6.64 49%(6.18)	5.72 22%(5.42)	2.11 -32%(2.25)
GMM_A: $\widehat{b}_2$	27.50^{***} 267%(7.46)	10.84^{**} 274%(5.86)	9.70^{**} 197%(5.34)	6.45^{*} 45%(5.43)	4.73 1%(4.90)	1.42 -54%(1.78)
GMM_B: $\widehat{b}_1$	24.61^{***} 228%(11.01)	-4.25 -247%(8.80)	6.57 101%(7.41)	6.38 44%(6.96)	5.65 20%(5.83)	2.14 -31%(2.31)
GMM_B: $\widehat{b}_2$	20.63^{***} 175%(10.16)	7.63^{*} 164%(6.45)	7.67^{*} 135%(5.98)	4.84 9%(6.02)	4.03 -14%(5.25)	1.43 -54%(1.83)
GMM_C: $\widehat{b}$	$\times$	20.42^{**} 605%(10.91)	12.91^{**} 295%(8.31)	7.10 60%(7.03)	5.19 10%(5.83)	1.47 -53%(2.32)

Panel B. N Portfolios formed on BE/ME

	T = 60	T = 120	T = 240	T = 360	T = 480	T = 948
$\lambda = Ef$	0.78	0.46	0.34	0.43	0.43	0.41
$b = \frac{\lambda}{E(R^{em2})}$	7.50	2.90	3.26	4.44	4.70	3.10
			N = 5			
Beta: $\widehat{\lambda}_{OLS}$	0.56 $-28\%(0.48)$	0.52 $13\%(0.84)$	0.16 $-54\%(0.44)$	0.19 $-55\%(0.27)$	0.24 $-43\%(0.20)$	0.31^{**} $-24\%(0.15)$
Beta: $\widehat{\lambda}_{GLS}$	0.60 $-23\%(0.46)$	0.34 $-27\%(0.73)$	0.25 $-28\%(0.40)$	0.18 $-57\%(0.26)$	0.28^{*} $-35\%(0.19)$	0.31^{**} $-25\%(0.15)$
Beta: $\widehat{\lambda}_{WLS}$	0.58 $-25\%(0.50)$	0.51 $12\%(0.97)$	0.16 $-53\%(0.42)$	0.19 $-55\%(0.24)$	0.24 $-44\%(0.18)$	0.31^{**} $-23\%(0.14)$
GMM_A: $\widehat{b}_1$	7.71^{*} $3\%(4.91)$	4.49 $55\%(4.22)$	4.88^{*} $50\%(3.08)$	7.00^{***} $58\%(1.89)$	5.56^{***} $18\%(1.74)$	1.75^{*} $-44\%(1.12)$
GMM_A: $\widehat{b}_2$	8.30^{*} $11\%(4.57)$	5.41 $87\%(4.05)$	4.69^{*} $44\%(3.00)$	7.00^{***} $58\%(1.89)$	5.81^{***} $24\%(1.71)$	1.66^{*} $-47\%(1.11)$
GMM_B: $\widehat{b}_1$	8.94^{*} $19\%(6.19)$	4.52 $56\%(4.75)$	5.12^{*} $57\%(3.37)$	7.56^{***} $70\%(2.22)$	5.82^{***} $24\%(1.90)$	1.78^{*} $-43\%(1.14)$
GMM_B: $\widehat{b}_2$	9.55^{*} $27\%(5.86)$	5.60 $93\%(4.53)$	4.90^{*} $50\%(3.28)$	7.56^{***} $70\%(2.22)$	6.07^{***} $29\%(1.87)$	1.69^{*} $-46\%(1.13)$
GMM_C: $\widehat{b}$	9.64^{*} $29\%(6.24)$	5.44 $88\%(4.75)$	5.33^{*} $63\%(3.39)$	7.57^{***} $70\%(2.23)$	6.11^{***} $30\%(1.90)$	1.70^{*} $-45\%(1.14)$
			N = 10			
Beta: $\widehat{\lambda}_{OLS}$	0.51 $-35\%(0.48)$	0.34 $-26\%(0.73)$	0.18 $-46\%(0.35)$	0.29^{*} $-33\%(0.23)$	0.25^{*} $-42\%(0.19)$	0.30^{***} $-27\%(0.14)$
Beta: $\widehat{\lambda}_{GLS}$	0.57^{*} $-26\%(0.44)$	0.35 $-24\%(0.61)$	0.28 $-19\%(0.32)$	0.30^{*} $-31\%(0.22)$	0.27^{**} $-37\%(0.18)$	0.33^{***} $-19\%(0.14)$
Beta: $\widehat{\lambda}_{WLS}$	0.52 $-34\%(0.48)$	0.33 $-27\%(0.66)$	0.21 $-39\%(0.31)$	0.29^{*} $-33\%(0.21)$	0.25^{**} $-42\%(0.17)$	0.30^{***} $-28\%(0.13)$
GMM_A: $\widehat{b}_1$	7.13^{*} $-5\%(5.20)$	4.89^{*} $69\%(3.53)$	4.58^{**} $40\%(2.69)$	6.98^{***} $57\%(1.91)$	5.51^{***} $17\%(1.75)$	1.67^{**} $-46\%(1.12)$
GMM_A: $\widehat{b}_2$	7.89^{**} $5\%(4.35)$	5.59^{**} $93\%(3.37)$	4.32^{**} $32\%(2.63)$	7.01^{***} $58\%(1.89)$	5.66^{***} $21\%(1.72)$	1.74^{**} $-44\%(1.10)$
GMM_B: $\widehat{b}_1$	8.29^{*} $11\%(6.41)$	5.04^{*} $74\%(3.98)$	4.79^{**} $47\%(2.94)$	7.48^{***} $68\%(2.21)$	5.76^{***} $23\%(1.90)$	1.70^{**} $-45\%(1.14)$
GMM_B: $\widehat{b}_2$	8.49^{**} $13\%(5.55)$	5.49^{**} $90\%(3.79)$	4.44^{**} $36\%(2.86)$	7.43^{***} $67\%(2.19)$	5.89^{***} $25\%(1.87)$	1.74^{**} $-44\%(1.12)$
GMM_C: $\widehat{b}$	9.26^{*} $24\%(6.59)$	5.83^{**} $101\%(4.05)$	4.49^{**} $38\%(3.05)$	7.50^{***} $69\%(2.22)$	5.91^{***} $26\%(1.91)$	1.81^{**} $-42\%(1.14)$

Panel C. N Portfolios formed by the intersections of ME and BE/ME

	T = 60	**T = 120**	**T = 240**	**T = 360**	**T = 480**	**T = 948**
$\lambda = Ef$	0.78	0.46	0.34	0.43	0.43	0.41
$b = \frac{\lambda}{E\left(R^{em2}\right)}$	7.50	2.90	3.26	4.44	4.70	3.10
			N = 6			
Beta: $\widehat{\lambda}_{\mathrm{OLS}}$	0.82^{**} 4%(0.40)	0.49^{*} 7%(0.36)	0.36^{**} 5%(0.21)	0.45^{***} 5%(0.16)	0.43^{***} 1%(0.14)	0.40^{***} −1%(0.12)
Beta: $\widehat{\lambda}_{\mathrm{GLS}}$	0.78^{**} 0%(0.40)	0.46^{*} 0%(0.36)	0.34^{**} 0%(0.21)	0.43^{***} 0%(0.16)	0.43^{***} 0%(0.14)	0.41^{***} 0%(0.12)
Beta: $\widehat{\lambda}_{\mathrm{WLS}}$	0.78^{**} −1%(0.42)	0.45^{*} −1%(0.37)	0.36^{**} 5%(0.22)	0.45^{***} 6%(0.17)	0.45^{***} 4%(0.14)	0.39^{***} −4%(0.12)
$\mathrm{GMM_A}$: $\widehat{b}_1$	10.02^{***} 34%(4.07)	8.60^{***} 197%(2.90)	7.92^{***} 143%(2.30)	9.31^{***} 110%(1.77)	7.66^{***} 63%(1.48)	2.48^{***} −20%(0.83)
$\mathrm{GMM_A}$: $\widehat{b}_2$	11.92^{***} 59%(3.89)	10.91^{***} 277%(2.85)	8.66^{***} 165%(2.29)	9.90^{***} 123%(1.76)	8.45^{***} 80%(1.46)	2.21^{***} −29%(0.82)
$\mathrm{GMM_B}$: $\widehat{b}_1$	11.74^{**} 57%(5.78)	9.16^{***} 216%(3.61)	8.31^{***} 155%(2.70)	10.03^{***} 126%(2.19)	8.03^{***} 71%(1.70)	2.53^{***} −19%(0.86)
$\mathrm{GMM_B}$: $\widehat{b}_2$	12.27^{**} 64%(5.60)	8.77^{***} 203%(3.60)	7.29^{***} 123%(2.68)	9.31^{***} 110%(2.18)	8.21^{***} 75%(1.69)	2.19^{***} −30%(0.85)
$\mathrm{GMM_C}$: $\widehat{b}$	13.78^{**} 84%(6.31)	27.46^{***} 848%(8.33)	10.54^{***} 223%(2.90)	11.04^{***} 149%(2.27)	9.03^{***} 92%(1.74)	2.26^{***} −27%(0.85)
			N = 25			
Beta: $\widehat{\lambda}_{\mathrm{OLS}}$	0.81^{***} 4%(0.42)	0.51^{**} 12%(0.37)	0.37^{***} 9%(0.21)	0.47^{***} 9%(0.17)	0.47^{***} 10%(0.14)	0.44^{***} 7%(0.12)
Beta: $\widehat{\lambda}_{\mathrm{GLS}}$	0.76^{***} −2%(0.40)	0.45^{*} −3%(0.36)	0.34^{**} −2%(0.21)	0.43^{***} 1%(0.16)	0.44^{***} 2%(0.14)	0.39^{***} −6%(0.12)
Beta: $\widehat{\lambda}_{\mathrm{WLS}}$	0.67^{**} −15%(0.41)	0.34 −27%(0.37)	0.27^{*} −21%(0.21)	0.41^{***} −4%(0.17)	0.44^{***} 2%(0.14)	0.37^{***} −10%(0.12)
$\mathrm{GMM_A}$: $\widehat{b}_1$	10.21^{***} 36%(4.06)	9.29^{***} 221%(2.93)	8.37^{***} 156%(2.39)	9.63^{***} 117%(1.85)	8.18^{***} 74%(1.53)	2.79^{***} −10%(0.88)
$\mathrm{GMM_A}$: $\widehat{b}_2$	19.87^{***} 165%(2.95)	19.17^{***} 562%(2.36)	15.91^{***} 387%(2.04)	13.21^{***} 198%(1.67)	10.55^{***} 125%(1.44)	2.07^{***} −33%(0.85)
$\mathrm{GMM_B}$: $\widehat{b}_1$	11.80^{***} 57%(5.85)	9.85^{***} 240%(3.73)	8.73^{***} 168%(2.82)	10.34^{***} 133%(2.30)	8.57^{***} 82%(1.77)	2.81^{***} −9%(0.91)
$\mathrm{GMM_B}$: $\widehat{b}_2$	15.91^{***} 112%(4.72)	11.09^{***} 283%(3.39)	9.62^{***} 195%(2.58)	10.28^{***} 131%(2.13)	9.11^{***} 94%(1.70)	1.97^{***} −36%(0.87)
$\mathrm{GMM_C}$: $\widehat{b}$	70.51^{***} 841%(27.38)	×	×	22.05^{***} 397%(4.01)	14.01^{***} 198%(2.19)	8.15^{***} 163%(1.53)
			N = 100			
Beta: $\widehat{\lambda}_{\mathrm{OLS}}$		0.50^{**} 10%(0.38)	0.35^{**} 4%(0.22)	0.46^{***} 6%(0.17)	0.45^{***} 5%(0.14)	0.46^{***} 11%(0.13)
Beta: $\widehat{\lambda}_{\mathrm{GLS}}$		0.38^{*} −16%(0.36)	0.26^{*} −24%(0.21)	0.39^{***} −9%(0.16)	0.42^{***} −3%(0.14)	0.42^{***} 2%(0.12)
Beta: $\widehat{\lambda}_{\mathrm{WLS}}$		0.36 −22%(0.36)	0.29^{**} −16%(0.21)	0.44^{***} 2%(0.17)	0.43^{***} 1%(0.14)	0.38^{***} −8%(0.12)
$\mathrm{GMM_A}$: $\widehat{b}_1$		9.57^{***} 230%(3.01)	8.41^{***} 158%(2.44)	9.65^{***} 117%(1.89)	7.99^{***} 70%(1.56)	2.88^{***} −7%(0.86)
$\mathrm{GMM_A}$: $\widehat{b}_2$		×	21.71^{***} 565%(1.77)	17.14^{***} 286%(1.46)	12.97^{***} 176%(1.35)	2.40^{***} −23%(0.80)
$\mathrm{GMM_B}$: $\widehat{b}_1$		9.99^{***} 245%(3.84)	8.67^{***} 166%(2.88)	10.28^{***} 131%(2.35)	8.28^{***} 76%(1.79)	2.91^{***} −6%(0.89)
$\mathrm{GMM_B}$: $\widehat{b}_2$		11.09^{***} 283%(2.58)	8.17^{***} 150%(2.33)	9.76^{***} 120%(1.92)	8.24^{***} 75%(1.61)	2.28^{***} −27%(0.82)
$\mathrm{GMM_C}$: $\widehat{b}$		−15.19 −624%(4.23)	×	26.00^{***} 485%(4.01)	22.74^{***} 384%(3.17)	2.60^{***} −16%(0.95)

Panel D. N Portfolios formed on ME MOM

	T = 60	**T = 120**	**T = 240**	**T = 360**	**T = 480**	**T = 948**
$\lambda = Ef$	0.78	0.46	0.34	0.43	0.43	0.41
$b = \frac{\lambda}{E(R^{em2})}$	7.50	2.90	3.26	4.44	4.70	3.10
			N = 6			
Beta: $\widehat{\lambda}_{\text{OLS}}$	0.63 $-19\%(0.73)$	-0.48 $-204\%(0.54)$	-0.81 $-336\%(0.37)$	-0.83 $-294\%(0.35)$	-1.33 $-410\%(0.38)$	-1.76 $-530\%(0.34)$
Beta: $\widehat{\lambda}_{\text{GLS}}$	0.53 $-32\%(0.56)$	0.31 $-33\%(0.44)$	0.13 $-61\%(0.29)$	0.29 $-33\%(0.25)$	0.18 $-59\%(0.22)$	-0.32 $-178\%(0.22)$
Beta: $\widehat{\lambda}_{\text{WLS}}$	0.55 $-30\%(0.53)$	-0.39 $-184\%(0.50)$	-0.67 $-296\%(0.32)$	-0.59 $-238\%(0.28)$	-0.86 $-301\%(0.24)$	-1.36 $-431\%(0.20)$
GMM_A: $\widehat{b}_1$	8.90* 19%(6.81)	3.46 19%(4.29)	-1.08 $-133\%(4.33)$	1.52 $-66\%(3.61)$	-10.37 $-321\%(4.77)$	-15.40 $-596\%(5.27)$
GMM_A: $\widehat{b}_2$	7.57* 1%(5.23)	9.34*** 222%(3.79)	7.39** 126%(4.00)	10.93*** 146%(3.13)	4.98* 6%(3.45)	1.86 $-40\%(3.07)$
GMM_B: $\widehat{b}_1$	10.28 37%(8.78)	-0.21 $-107\%(5.06)$	-7.74 $-337\%(4.82)$	-7.69 $-273\%(4.58)$	-15.75 $-435\%(5.75)$	-15.30 $-593\%(5.09)$
GMM_B: $\widehat{b}_2$	6.20 $-17\%(6.68)$	6.65* 130%(4.23)	0.72 $-78\%(4.21)$	3.14 $-29\%(3.69)$	2.63 $-44\%(3.97)$	0.95 $-69\%(3.03)$
GMM_C: $\widehat{b}$	8.16 9%(8.84)	20.95** 624%(11.14)	×	×	×	×
			N = 25			
Beta: $\widehat{\lambda}_{\text{OLS}}$	0.84* 7%(0.66)	0.19 $-59\%(0.41)$	-0.05 $-116\%(0.27)$	-0.01 $-102\%(0.24)$	-0.42 $-198\%(0.23)$	-1.27 $-410\%(0.27)$
Beta: $\widehat{\lambda}_{\text{GLS}}$	0.60* $-24\%(0.46)$	0.31 $-32\%(0.40)$	0.40** 17%(0.26)	0.59*** 38%(0.22)	0.34*** $-22\%(0.19)$	0.21* $-49\%(0.18)$
Beta: $\widehat{\lambda}_{\text{WLS}}$	0.89*** 14%(0.45)	-0.06 $-112\%(0.40)$	-0.32 $-193\%(0.24)$	-0.21 $-148\%(0.20)$	-0.46 $-208\%(0.17)$	-1.10 $-369\%(0.15)$
GMM_A: $\widehat{b}_1$	11.42*** 52%(6.10)	8.72*** 201%(3.55)	6.38** 95%(3.76)	8.58*** 93%(2.86)	1.04 $-78\%(2.51)$	-10.85 $-450\%(3.54)$
GMM_A: $\widehat{b}_2$	17.20*** 129%(3.15)	15.93*** 450%(2.83)	18.77*** 475%(3.02)	21.37*** 381%(2.30)	9.00*** 92%(2.15)	3.32** 7%(2.12)
GMM_B: $\widehat{b}_1$	13.22** 76%(8.39)	7.23** 150%(4.26)	3.02 $-8\%(3.95)$	3.91* $-12\%(3.22)$	-3.43 $-173\%(2.89)$	-11.23 $-462\%(3.60)$
GMM_B: $\widehat{b}_2$	15.88*** 112%(4.57)	7.07*** 144%(3.68)	6.92*** 112%(3.59)	10.09*** 127%(2.88)	3.05* $-35\%(2.39)$	2.59* $-17\%(2.15)$
GMM_C: $\widehat{b}$	44.15*** 489%(19.71)	5.15 78%(7.48)	×	×	×	×

Panel E. N Industry Portfolios

	T = 60	**T = 120**	**T = 240**	**T = 360**	**T = 480**	**T = 948**
$\lambda = Ef$	0.78	0.46	0.34	0.43	0.43	0.41
$b = \frac{\lambda}{E\left(R^{em2}\right)}$	7.50	2.90	3.26	4.44	4.70	3.10
			N = 5			
Beta: $\widehat{\lambda}_{\text{OLS}}$	0.79^{*} 1%(0.52)	0.22 −51%(0.42)	0.06 −82%(0.27)	0.04 −90%(0.23)	0.01 −97%(0.21)	−0.08 −121%(0.18)
Beta: $\widehat{\lambda}_{\text{GLS}}$	0.79^{*} 1%(0.52)	0.21 −55%(0.42)	0.06 −84%(0.26)	0.04 −90%(0.23)	0% −100%(0.21)	−0.10 −125%(0.17)
Beta: $\widehat{\lambda}_{\text{WLS}}$	0.79^{*} 1%(0.48)	0.23 −49%(0.41)	0.08 −76%(0.26)	0.04 −91%(0.23)	0.01 −98%(0.20)	−0.09 −121%(0.17)
GMM_{A}: $\widehat{b}_1$	10.18^{**} 36%(5.36)	2.72 −6%(5.92)	−1.26 −139%(5.36)	1.04 −77%(4.08)	0.70 −85%(2.96)	−1.27 −141%(1.77)
GMM_{A}: $\widehat{b}_2$	10.73^{**} 43%(5.12)	5.07 75%(5.42)	0.12 −96%(5.23)	1.18 −73%(4.05)	0.62 −87%(2.95)	−1.61 −152%(1.75)
GMM_{B}: $\widehat{b}_1$	12.02^{*} 60%(7.38)	2.60 −10%(6.34)	−1.35 −141%(5.43)	1.04 −76%(4.18)	0.69 −85%(2.98)	−1.28 −141%(1.77)
GMM_{B}: $\widehat{b}_2$	12.42^{*} 66%(7.11)	4.97 72%(5.86)	−0.12 −104%(5.32)	1.18 −73%(4.14)	0.61 −87%(2.98)	−1.62 −152%(1.74)
GMM_{C}: $\widehat{b}$	12.70^{*} 69%(7.52)	4.87 68%(6.14)	−0.35 −111%(5.27)	1.19 −73%(4.10)	0.65 −86%(2.88)	−1.56 −150%(1.62)
			N = 17			
Beta: $\widehat{\lambda}_{\text{OLS}}$	0.81^{**} 4%(0.52)	0.12 −73%(0.42)	0.04 −88%(0.26)	0.04 −92%(0.22)	−0.02 −106%(0.18)	−0.03 −108%(0.16)
Beta: $\widehat{\lambda}_{\text{GLS}}$	0.35 −55%(0.45)	0.14 −70%(0.39)	−0.03 −108%(0.23)	−0.07 −117%(0.20)	−0.12 −127%(0.17)	−0.06 −114%(0.15)
Beta: $\widehat{\lambda}_{\text{WLS}}$	0.75^{**} −5%(0.46)	0.21 −54%(0.39)	0.07 −81%(0.24)	0.02 −96%(0.20)	0.03 −94%(0.17)	0.01 −99%(0.16)
GMM_{A}: $\widehat{b}_1$	10.57^{***} 41%(5.52)	2.80 −3%(4.35)	0.92 −72%(3.99)	1.48 −67%(3.62)	0.85 −82%(2.61)	−0.97 −131%(1.30)
GMM_{A}: $\widehat{b}_2$	14.85^{***} 98%(4.11)	1.80 −38%(3.90)	−0.56 −117%(3.36)	0.57 −87%(2.96)	−0.41 −109%(2.35)	−1.29 −142%(1.23)
GMM_{B}: $\widehat{b}_1$	11.92^{**} 59%(7.32)	2.40 −17%(4.59)	0.75 −77%(4.11)	1.15 −74%(3.71)	0.63 −87%(2.63)	−1.01 −133%(1.30)
GMM_{B}: $\widehat{b}_2$	9.90^{**} 32%(5.72)	−0.12 −104%(4.06)	−1.03 −131%(3.45)	−0.19 −104%(3.02)	−0.85 −118%(2.36)	−1.32 −143%(1.23)
GMM_{C}: $\widehat{b}$	56.16^{***} 649%(19.29)	−0.79 −127%(5.38)	−1.32 −140%(4.38)	−0.99 −122%(3.63)	−1.14 −124%(2.62)	−1.33 −143%(1.31)
			N = 30			
Beta: $\widehat{\lambda}_{\text{OLS}}$	0.90^{**} 15%(0.53)	0.19 −59%(0.41)	0.06 −82%(0.25)	−0.02 −105%(0.21)	−0.04 −110%(0.18)	−0.19 −147%(0.18)
Beta: $\widehat{\lambda}_{\text{GLS}}$	0.46^{*} −41%(0.43)	0.08 −82%(0.38)	−0.06 −119%(0.23)	−0.08 −118%(0.19)	−0.10 −123%(0.16)	−0.10 −125%(0.15)
Beta: $\widehat{\lambda}_{\text{WLS}}$	0.62^{**} −20%(0.44)	0.10 −79%(0.39)	−0.02 −107%(0.23)	−0.09 −120%(0.19)	−0.08 −118%(0.17)	−0.13 −131%(0.16)
GMM_{A}: $\widehat{b}_1$	11.59^{***} 55%(4.92)	6.27^{**} 116%(3.82)	2.25 −31%(3.69)	1.09 −76%(3.43)	1.04 −78%(2.56)	−2.28 −174%(1.50)
GMM_{A}: $\widehat{b}_2$	16.22^{***} 116%(3.09)	7.15^{***} 147%(3.29)	0.67 −79%(3.04)	1.39 −69%(2.70)	0.56 −88%(2.20)	−2.09 −167%(1.22)
GMM_{B}: $\widehat{b}_1$	13.25^{***} 77%(7.12)	5.42^{*} 87%(4.28)	1.94 −41%(3.85)	0.70 −84%(3.51)	0.68 −86%(2.58)	−2.39 −177%(1.50)
GMM_{B}: $\widehat{b}_2$	12.01^{***} 60%(4.68)	2.18 −25%(3.73)	−0.57 −117%(3.15)	0.11 −98%(2.77)	−0.17 −104%(2.23)	−2.14 −169%(1.23)
GMM_{C}: $\widehat{b}$	15.88^{***} 112%(8.86)	−12.40 −528%(6.42)	0.00 −100%(3.92)	−0.40 −109%(3.43)	−1.17 −125%(2.54)	−2.72 −188%(1.59)

3.6.5 Market risk premium in RUH model

The following five panels report the estimate values, the bias from the risk premium in percent values, the standard error in parenthesis, and t-statistics. Each panel corresponds to a one set of test portfolios. T is 60 (January 2001 - December 2005), 120, 240, 360, 480 and 948 (January 1927 - December 2005) monthly observations. The two-pass (Fama-MacBeth) cross-sectional estimate is the slope coefficient λ in $E\left(R^{e}\right) = \beta\lambda$ calculated by OLS, GLS and WLS. Next, we turn from beta representation to a discount factor formulation for GMM approach. The $\mathrm{GMM_A}$ and $\mathrm{GMM_B}$ first and second-stage estimates are the parameters $100 \times b$ in $E\left(R^{e}\right) = E\left(R^{e} f'\right) b$ (returns on second moments, following Hansen and Jagannathan [49]) and $E\left(R^{e}\right) = E\left(R^{e} \widetilde{f}'\right) b$ (returns on covariances, following Cochrane [25]) respectively. The $\mathrm{GMM_C}$ is the continuous updating estimate (following Hansen, Heaton and Yaron [47]).

Panel A. N Portfolios formed on ME

	T = 60	T = 120	T = 240	T = 360	T = 480	T = 948
$\lambda = Ef$	0.13	0.56	0.64	0.62	0.44	0.64
$b = \frac{\lambda}{E(R^{em2})}$	0.66	2.52	3.15	3.11	2.11	2.11
			N = 5			
Beta: $\widehat{\lambda}_{\text{OLS}}$	0.11 −20%(0.58)	0.61 9%(0.44)	0.70** 8%(0.30)	0.70*** 13%(0.24)	0.59** 34%(0.24)	0.67*** 4%(0.18)
Beta: $\widehat{\lambda}_{\text{GLS}}$	0.07 −44%(0.57)	0.68* 22%(0.43)	0.70** 10%(0.29)	0.68** 9%(0.23)	0.48** 10%(0.21)	0.67*** 4%(0.18)
Beta: $\widehat{\lambda}_{\text{WLS}}$	0.04 −68%(0.59)	0.64* 16%(0.43)	0.70** 9%(0.29)	0.69** 11%(0.24)	0.52** 19%(0.25)	0.67*** 4%(0.18)
GMM_A: $\widehat{b}_1$	5.45 723%(11.51)	8.18* 224%(5.63)	7.06** 124%(3.30)	5.98* 92%(4.10)	7.78 269%(7.16)	3.32* 57%(1.85)
GMM_A: $\widehat{b}_2$	1.33 101%(8.76)	9.51* 277%(5.06)	7.04** 123%(3.24)	5.18* 66%(3.45)	4.65 120%(4.94)	3.25** 54%(1.58)
GMM_B: $\widehat{b}_1$	7.11 973%(17.01)	8.49 237%(8.08)	8.45* 168%(4.73)	7.01 125%(5.45)	5.01 137%(5.33)	3.77* 79%(2.57)
GMM_B: $\widehat{b}_2$	1.19 79%(12.39)	11.15* 342%(7.02)	8.28* 163%(4.62)	6.00 93%(4.84)	4.99 137%(5.29)	3.67* 74%(2.18)
GMM_C: $\widehat{b}$	2.31 249%(14.95)	13.02 416%(10.39)	8.68* 175%(4.73)	6.33 103%(4.93)	×	3.68* 74%(2.53)
			N = 10			
Beta: $\widehat{\lambda}_{\text{OLS}}$	0.18 34%(0.59)	0.60* 8%(0.44)	0.69*** 8%(0.30)	0.71*** 15%(0.24)	0.59*** 34%(0.23)	0.67*** 4%(0.18)
Beta: $\widehat{\lambda}_{\text{GLS}}$	0.11 −20%(0.57)	0.67** 20%(0.43)	0.71*** 10%(0.29)	0.68*** 10%(0.23)	0.48*** 10%(0.21)	0.66*** 3%(0.18)
Beta: $\widehat{\lambda}_{\text{WLS}}$	0.06 −55%(0.59)	0.64** 15%(0.43)	0.70*** 9%(0.29)	0.70*** 12%(0.24)	0.53*** 22%(0.21)	0.66*** 3%(0.18)
GMM_A: $\widehat{b}_1$	2.36 256%(8.20)	7.65*** 204%(3.80)	6.24*** 98%(2.58)	4.42** 42%(2.37)	4.16*** 97%(2.15)	2.12*** 0%(0.99)
GMM_A: $\widehat{b}_2$	×	7.98*** 216%(2.88)	6.02*** 91%(2.26)	4.25*** 36%(1.99)	3.52** 67%(2.01)	2.24*** 6%(0.87)
GMM_B: $\widehat{b}_1$	1.02 54%(11.07)	7.96** 215%(5.22)	6.91*** 119%(3.29)	4.93** 58%(2.88)	2.90 38%(2.68)	2.12** 0%(1.12)
GMM_B: $\widehat{b}_2$	×	8.88*** 252%(3.87)	6.15*** 95%(2.89)	4.51** 45%(2.52)	3.60** 71%(2.23)	2.24*** 6%(0.98)
GMM_C: $\widehat{b}$	×	9.74** 286%(5.69)	7.34*** 133%(3.37)	4.94** 59%(2.77)	4.57** 116%(2.73)	2.26*** 7%(1.12)

Panel B. N Portfolios formed on BE/ME

	T = 60	**T = 120**	**T = 240**	**T = 360**	**T = 480**	**T = 948**
$\lambda = Ef$	0.13	0.56	0.64	0.62	0.44	0.64
$b = \frac{\lambda}{E(R^{em2})}$	0.66	2.52	3.15	3.11	2.11	2.11
			N = 5			
Beta: $\widehat{\lambda}_{\text{OLS}}$	0.07 $-44\%(0.58)$	0.81 45%(0.85)	0.64^{**} 0%(0.31)	0.64^{**} 3%(0.24)	0.45^{**} 3%(0.21)	0.67^{***} 4%(0.18)
Beta: $\widehat{\lambda}_{\text{GLS}}$	0.06 $-53\%(0.58)$	0.78 39%(0.74)	0.65^{**} 1%(0.31)	0.66^{**} 5%(0.24)	0.47^{**} 7%(0.21)	0.67^{***} 5%(0.18)
Beta: $\widehat{\lambda}_{\text{WLS}}$	0.06 $-55\%(0.58)$	0.80 43%(0.86)	0.64^{**} 0%(0.31)	0.65^{**} 4%(0.24)	0.46^{**} 4%(0.21)	0.67^{***} 4%(0.18)
GMM_A: $\widehat{b}_1$	$\times$	16.97 573%(27.03)	1.43 $-55\%(14.56)$	5.38^{***} 73%(1.54)	3.05 45%(2.24)	2.69^{*} 28%(1.44)
GMM_A: $\widehat{b}_2$	$\times$	17.05 576%(25.32)	2.87 $-9\%(11.66)$	5.47^{***} 76%(1.52)	3.97^{**} 88%(1.92)	2.83^{**} 34%(1.40)
GMM_B: $\widehat{b}_1$	$\times$	26.50 951%(72.71)	0.56 $-82\%(14.81)$	5.73^{**} 84%(2.10)	2.60 23%(2.97)	2.82^{*} 33%(1.72)
GMM_B: $\widehat{b}_2$	$\times$	26.83 964%(70.17)	1.88 $-40\%(12.12)$	6.02^{***} 93%(2.04)	3.96^{*} 88%(2.43)	3.00^{*} 42%(1.68)
GMM_C: $\widehat{b}$	$\times$	$\times$	0.27 $-91\%(15.54)$	5.98^{**} 92%(2.20)	3.43 63%(2.71)	3.11^{*} 47%(1.76)
			N = 10			
Beta: $\widehat{\lambda}_{\text{OLS}}$	0.09 $-32\%(0.58)$	0.59^{*} 6%(0.43)	0.67^{***} 4%(0.29)	0.63^{***} 2%(0.24)	0.45^{***} 3%(0.21)	0.69^{***} 8%(0.18)
Beta: $\widehat{\lambda}_{\text{GLS}}$	0.06 $-55\%(0.57)$	0.58^{*} 3%(0.43)	0.68^{***} 5%(0.29)	0.64^{***} 2%(0.24)	0.47^{***} 6%(0.21)	0.69^{***} 8%(0.18)
Beta: $\widehat{\lambda}_{\text{WLS}}$	0.07 $-48\%(0.58)$	0.59^{*} 6%(0.43)	0.67^{***} 5%(0.29)	0.63^{***} 1%(0.24)	0.45^{***} 4%(0.21)	0.69^{***} 8%(0.18)
GMM_A: $\widehat{b}_1$	6.39 864%(10.06)	5.94^{**} 135%(3.58)	3.84^{**} 22%(2.40)	5.31^{***} 71%(1.46)	3.21^{***} 52%(1.37)	3.86^{***} 83%(1.44)
GMM_A: $\widehat{b}_2$	2.77 318%(7.14)	6.90^{***} 173%(3.46)	4.06^{**} 29%(2.30)	5.29^{***} 70%(1.42)	3.44^{***} 63%(1.32)	3.75^{***} 77%(1.24)
GMM_B: $\widehat{b}_1$	6.91 942%(11.02)	5.95^{*} 136%(4.39)	3.64^{*} 15%(2.64)	5.50^{***} 77%(1.75)	3.05^{***} 45%(1.55)	3.89^{***} 84%(1.79)
GMM_B: $\widehat{b}_2$	2.22 235%(7.83)	6.59^{**} 161%(4.31)	3.82^{**} 21%(2.51)	5.33^{***} 71%(1.67)	3.36^{***} 59%(1.47)	3.47^{***} 64%(1.56)
GMM_C: $\widehat{b}$	-2.40 $-463\%(8.81)$	6.69^{**} 165%(4.51)	3.68^{*} 17%(2.66)	5.36^{***} 72%(1.72)	3.24^{***} 53%(1.55)	12.25^{**} 480%(7.28)

Panel C. N Portfolios formed by the intersections of ME and BE/ME

	T = 60	**T = 120**	**T = 240**	**T = 360**	**T = 480**	**T = 948**
$\lambda = Ef$	0.13	0.56	0.64	0.62	0.44	0.64
$b = \frac{\lambda}{E\left(R^{em2}\right)}$	0.66	2.52	3.15	3.11	2.11	2.11
			N = 6			
Beta: $\widehat{\lambda}_{\mathrm{OLS}}$	0.40 203%(0.59)	0.68^{*} 22%(0.49)	0.70^{**} 8%(0.32)	0.66^{***} 6%(0.24)	0.67^{***} 54%(0.28)	0.70^{***} 8%(0.18)
Beta: $\widehat{\lambda}_{\mathrm{GLS}}$	0.11 −16%(0.58)	0.72^{**} 29%(0.43)	0.76^{***} 18%(0.29)	0.72^{***} 15%(0.24)	0.56^{***} 29%(0.22)	0.72^{***} 13%(0.18)
Beta: $\widehat{\lambda}_{\mathrm{WLS}}$	0.07 −50%(0.58)	0.72^{*} 29%(0.44)	0.74^{***} 16%(0.30)	0.71^{***} 14%(0.24)	0.55^{***} 27%(0.22)	0.67^{***} 4%(0.18)
$\mathrm{GMM_A}$: $\widehat{b}_1$	2.68 304%(9.14)	12.28^{***} 387%(5.05)	8.18^{**} 159%(3.73)	5.81^{***} 87%(2.41)	8.46^{**} 301%(4.45)	3.66^{***} 73%(1.33)
$\mathrm{GMM_A}$: $\widehat{b}_2$	×	13.01^{***} 416%(2.92)	8.99^{***} 185%(3.07)	6.50^{***} 109%(2.22)	7.50^{***} 255%(3.10)	5.12^{***} 142%(1.00)
$\mathrm{GMM_B}$: $\widehat{b}_1$	×	16.58^{**} 557%(8.76)	10.94^{**} 247%(6.50)	8.01^{**} 157%(4.11)	12.13^{**} 475%(6.80)	3.00^{**} 42%(1.83)
$\mathrm{GMM_B}$: $\widehat{b}_2$	×	17.62^{***} 599%(5.45)	12.17^{**} 286%(5.24)	9.31^{***} 199%(3.58)	10.52^{**} 398%(5.85)	5.63^{***} 167%(1.26)
$\mathrm{GMM_C}$: $\widehat{b}$	×	18.36^{**} 628%(9.29)	12.79^{**} 305%(7.38)	9.90^{**} 218%(5.80)	13.61^{**} 545%(8.26)	9.13^{*} 332%(5.82)
			N = 25			
Beta: $\widehat{\lambda}_{\mathrm{OLS}}$	0.50 278%(0.61)	0.79^{**} 42%(0.51)	0.76^{***} 18%(0.34)	0.72^{***} 16%(0.26)	0.65^{***} 49%(0.24)	0.68^{***} 6%(0.18)
Beta: $\widehat{\lambda}_{\mathrm{GLS}}$	0.14 4%(0.57)	0.72^{**} 30%(0.43)	0.76^{***} 18%(0.29)	0.69^{***} 11%(0.24)	0.51^{***} 16%(0.21)	0.70^{***} 8%(0.18)
Beta: $\widehat{\lambda}_{\mathrm{WLS}}$	0.31 136%(0.58)	0.76^{***} 36%(0.43)	0.76^{***} 19%(0.29)	0.74^{***} 19%(0.24)	0.59^{***} 35%(0.21)	0.71^{***} 10%(0.18)
$\mathrm{GMM_A}$: $\widehat{b}_1$	3.04 358%(5.05)	12.23^{***} 385%(4.13)	8.41^{***} 167%(3.41)	6.31^{***} 103%(2.44)	7.48^{***} 254%(2.66)	3.75^{***} 77%(0.91)
$\mathrm{GMM_A}$: $\widehat{b}_2$	×	14.12^{***} 460%(2.36)	10.75^{***} 241%(2.14)	7.71^{***} 148%(1.86)	6.61^{***} 213%(1.75)	4.57^{***} 116%(0.73)
$\mathrm{GMM_B}$: $\widehat{b}_1$	4.35 557%(5.75)	15.29^{***} 506%(6.18)	10.73^{***} 240%(5.16)	8.43^{***} 171%(3.86)	8.72^{***} 313%(2.95)	3.55^{***} 68%(1.09)
$\mathrm{GMM_B}$: $\widehat{b}_2$	7.21^{**} 989%(4.76)	15.86^{***} 529%(3.89)	11.46^{***} 263%(3.22)	8.11^{***} 161%(2.68)	7.12^{***} 237%(2.15)	4.40^{***} 108%(0.89)
$\mathrm{GMM_C}$: $\widehat{b}$	×	25.01^{***} 892%(10.28)	18.43^{***} 484%(8.44)	14.61^{***} 369%(5.92)	9.88^{***} 368%(4.49)	7.42 251%(7.17)
			N = 100			
Beta: $\widehat{\lambda}_{\mathrm{OLS}}$		0.85^{***} 53%(0.49)	0.77^{***} 21%(0.32)	0.78^{***} 25%(0.26)	0.63^{***} 44%(0.23)	0.71^{***} 11%(0.19)
Beta: $\widehat{\lambda}_{\mathrm{GLS}}$		0.65^{**} 17%(0.42)	0.69^{***} 8%(0.29)	0.66^{***} 6%(0.23)	0.48^{***} 10%(0.21)	0.70^{***} 9%(0.18)
Beta: $\widehat{\lambda}_{\mathrm{WLS}}$		0.82^{***} 47%(0.43)	0.78^{***} 21%(0.29)	0.78^{***} 25%(0.23)	0.60^{***} 38%(0.21)	0.72^{***} 11%(0.18)
$\mathrm{GMM_A}$: $\widehat{b}_1$		10.11^{***} 301%(3.14)	7.37^{***} 134%(2.64)	6.38^{***} 105%(2.09)	6.21^{***} 194%(1.68)	3.04^{***} 44%(0.84)
$\mathrm{GMM_A}$: $\widehat{b}_2$		14.22^{***} 464%(1.04)	14.43^{***} 357%(1.25)	11.99^{***} 285%(1.21)	8.35^{***} 295%(1.12)	3.79^{***} 80%(0.61)
$\mathrm{GMM_B}$: $\widehat{b}_1$		11.43^{***} 353%(4.45)	8.42^{***} 167%(3.45)	7.76^{***} 149%(2.81)	6.83^{***} 224%(1.95)	3.00^{***} 42%(0.93)
$\mathrm{GMM_B}$: $\widehat{b}_2$		−3.17 −226%(1.46)	8.48^{***} 169%(1.85)	5.90^{***} 90%(1.71)	4.94^{***} 134%(1.42)	3.24^{***} 54%(0.70)
$\mathrm{GMM_C}$: $\widehat{b}$		×	×	30.10^{***} 867%(7.56)	18.72 787%(17.99)	5.17^{***} 145%(2.31)

Panel D. N Portfolios formed on ME MOM

	T = 60	**T = 120**	**T = 240**	**T = 360**	**T = 480**	**T = 948**
$\lambda = Ef$	0.13	0.56	0.64	0.62	0.44	0.64
$b = \frac{\lambda}{E(R^{em2})}$	0.66	2.52	3.15	3.11	2.11	2.11
				N = 6		
Beta: $\widehat{\lambda}_{\text{OLS}}$	0.39 195%(0.59)	0.82** 48%(0.45)	0.78*** 22%(0.30)	0.80*** 29%(0.24)	0.61*** 39%(0.22)	0.69*** 8%(0.18)
Beta: $\widehat{\lambda}_{\text{GLS}}$	0.13 −1%(0.57)	0.57* 2%(0.43)	0.65** 1%(0.29)	0.63*** 1%(0.24)	0.45** 4%(0.21)	0.66*** 3%(0.18)
Beta: $\widehat{\lambda}_{\text{WLS}}$	0.15 13%(0.58)	0.70* 26%(0.43)	0.76*** 19%(0.29)	0.74*** 18%(0.24)	0.54*** 23%(0.21)	0.66*** 3%(0.18)
GMM$_{\text{A}}$: $\widehat{b}_1$	5.06 663%(5.68)	5.42** 115%(2.84)	5.61*** 78%(2.27)	5.09*** 63%(1.76)	4.79*** 127%(1.48)	2.88*** 36%(0.70)
GMM$_{\text{A}}$: $\widehat{b}_2$	0.80 21%(5.38)	6.64*** 163%(2.70)	6.67*** 112%(2.19)	6.03*** 94%(1.71)	5.60*** 165%(1.43)	2.79*** 32%(0.67)
GMM$_{\text{B}}$: $\widehat{b}_1$	6.99 955%(7.79)	5.50* 118%(3.38)	5.77** 83%(2.72)	4.82** 55%(2.07)	4.47*** 112%(1.70)	3.22*** 53%(0.80)
GMM$_{\text{B}}$: $\widehat{b}_2$	0.44 −34%(6.71)	5.85** 132%(3.21)	4.34** 38%(2.62)	3.55** 14%(2.00)	3.84** 82%(1.66)	2.92*** 38%(0.78)
GMM$_{\text{C}}$: $\widehat{b}$	×	10.90** 332%(4.65)	8.89*** 182%(3.45)	24.64*** 692%(6.48)	17.02*** 706%(3.68)	3.04*** 44%(0.82)
				N = 25		
Beta: $\widehat{\lambda}_{\text{OLS}}$	0.50 281%(0.60)	0.91*** 63%(0.47)	0.82*** 28%(0.31)	0.84*** 35%(0.25)	0.63*** 45%(0.22)	0.63*** −1%(0.18)
Beta: $\widehat{\lambda}_{\text{GLS}}$	0.20 50%(0.57)	0.53* −5%(0.43)	0.65*** 2%(0.29)	0.63*** 2%(0.24)	0.49*** 12%(0.21)	0.63*** −2%(0.18)
Beta: $\widehat{\lambda}_{\text{WLS}}$	0.36 174%(0.58)	0.81*** 45%(0.43)	0.78*** 21%(0.29)	0.79*** 27%(0.23)	0.58*** 32%(0.21)	0.62*** −3%(0.18)
GMM$_{\text{A}}$: $\widehat{b}_1$	5.77* 770%(5.08)	7.30*** 189%(2.82)	6.22*** 97%(2.42)	6.23*** 100%(1.85)	5.44*** 158%(1.50)	2.66*** 26%(0.73)
GMM$_{\text{A}}$: $\widehat{b}_2$	×	11.86*** 370%(2.38)	12.35*** 292%(1.80)	11.10*** 257%(1.44)	8.88*** 321%(1.22)	3.27*** 55%(0.66)
GMM$_{\text{B}}$: $\widehat{b}_1$	×	7.65*** 203%(3.55)	6.63*** 110%(2.97)	6.63*** 113%(2.30)	5.78*** 174%(1.82)	3.03*** 43%(0.86)
GMM$_{\text{B}}$: $\widehat{b}_2$	6.88** 938%(4.34)	6.72*** 166%(3.15)	7.04*** 123%(2.39)	7.43*** 139%(1.92)	6.64*** 215%(1.56)	3.38*** 60%(0.79)
GMM$_{\text{C}}$: $\widehat{b}$	×	23.59*** 835%(7.06)	31.92*** 912%(8.12)	×	×	4.21*** 99%(0.93)

Panel E. N Industry Portfolios

	T = 60	**T = 120**	**T = 240**	**T = 360**	**T = 480**	**T = 948**
$\lambda = Ef$	0.13	0.56	0.64	0.62	0.44	0.64
$b = \frac{\lambda}{E\left(R^{em2}\right)}$	0.66	2.52	3.15	3.11	2.11	2.11
			N = 5			
Beta: $\widehat{\lambda}_{\mathrm{OLS}}$	0.03 −80%(0.58)	0.76^{*} 37%(0.50)	0.79^{**} 23%(0.30)	0.76^{***} 21%(0.25)	0.55^{**} 25%(0.21)	0.76^{***} 17%(0.18)
Beta: $\widehat{\lambda}_{\mathrm{GLS}}$	0.06 −58%(0.58)	0.80^{*} 43%(0.49)	0.79^{**} 24%(0.30)	0.72^{***} 15%(0.24)	0.52^{**} 19%(0.21)	0.72^{***} 11%(0.18)
Beta: $\widehat{\lambda}_{\mathrm{WLS}}$	0.05 −61%(0.59)	0.79^{*} 42%(0.51)	0.79^{**} 23%(0.31)	0.72^{***} 16%(0.25)	0.53^{**} 20%(0.21)	0.72^{***} 12%(0.18)
GMM$_{\mathrm{A}}$: $\widehat{b}_1$	×	8.21 226%(7.01)	5.71^{**} 81%(2.77)	4.43^{**} 42%(1.73)	3.57^{**} 69%(1.55)	4.34^{***} 105%(1.20)
GMM$_{\mathrm{A}}$: $\widehat{b}_2$	−4.43 −768%(11.31)	8.97 256%(6.67)	5.81^{**} 84%(2.72)	4.36^{**} 40%(1.71)	3.16^{**} 50%(1.46)	3.75^{***} 77%(1.07)
GMM$_{\mathrm{B}}$: $\widehat{b}_1$	×	9.57 280%(10.42)	6.56^{*} 108%(3.79)	4.94^{**} 59%(2.12)	3.86^{**} 83%(1.87)	4.67^{***} 121%(1.51)
GMM$_{\mathrm{B}}$: $\widehat{b}_2$	−4.30 −749%(13.76)	10.47 315%(10.09)	6.62^{*} 110%(3.70)	4.69^{**} 51%(2.08)	3.31^{**} 57%(1.72)	3.93^{***} 86%(1.35)
GMM$_{\mathrm{C}}$: $\widehat{b}$	×	11.48 355%(12.33)	6.72^{*} 113%(3.79)	4.74^{**} 52%(1.98)	3.39^{**} 60%(1.74)	4.09^{**} 94%(1.45)
			N = 17			
Beta: $\widehat{\lambda}_{\mathrm{OLS}}$	0.25 90%(0.59)	0.64^{**} 15%(0.45)	0.73^{***} 14%(0.30)	0.69^{***} 11%(0.24)	0.53^{***} 21%(0.21)	0.72^{***} 12%(0.18)
Beta: $\widehat{\lambda}_{\mathrm{GLS}}$	0.20 50%(0.57)	0.62^{**} 12%(0.43)	0.72^{***} 13%(0.29)	0.72^{***} 16%(0.24)	0.52^{***} 20%(0.21)	0.73^{***} 13%(0.18)
Beta: $\widehat{\lambda}_{\mathrm{WLS}}$	0.22 67%(0.58)	0.56^{*} 0%(0.44)	0.68^{***} 5%(0.29)	0.69^{***} 11%(0.24)	0.50^{***} 15%(0.21)	0.72^{***} 12%(0.18)
GMM$_{\mathrm{A}}$: $\widehat{b}_1$	×	6.89^{***} 173%(3.27)	5.66^{***} 79%(2.48)	4.69^{***} 51%(1.75)	3.85^{***} 83%(1.50)	3.52^{***} 67%(0.85)
GMM$_{\mathrm{A}}$: $\widehat{b}_2$	×	7.84^{***} 211%(2.62)	5.42^{***} 72%(1.93)	4.46^{***} 43%(1.52)	2.89^{***} 37%(1.34)	3.06^{***} 45%(0.80)
GMM$_{\mathrm{B}}$: $\widehat{b}_1$	×	7.25^{**} 188%(4.14)	6.24^{***} 98%(3.08)	5.18^{***} 66%(2.13)	4.03^{***} 91%(1.76)	3.58^{***} 69%(1.00)
GMM$_{\mathrm{B}}$: $\widehat{b}_2$	×	6.46^{***} 156%(3.36)	5.19^{***} 65%(2.38)	4.23^{***} 36%(1.82)	2.42^{**} 15%(1.53)	2.92^{***} 38%(0.92)
GMM$_{\mathrm{C}}$: $\widehat{b}$	×	7.17^{**} 184%(4.17)	5.73^{***} 82%(2.88)	4.43^{***} 42%(1.94)	3.02^{***} 43%(1.58)	3.28^{***} 55%(0.98)
			N = 30			
Beta: $\widehat{\lambda}_{\mathrm{OLS}}$	0.24 80%(0.59)	0.80^{***} 43%(0.45)	0.75^{***} 17%(0.30)	0.70^{***} 13%(0.24)	0.55^{***} 27%(0.21)	0.76^{***} 19%(0.18)
Beta: $\widehat{\lambda}_{\mathrm{GLS}}$	0.09 −34%(0.57)	0.65^{**} 17%(0.43)	0.73^{***} 13%(0.29)	0.75^{***} 20%(0.24)	0.55^{***} 25%(0.21)	0.75^{***} 17%(0.18)
Beta: $\widehat{\lambda}_{\mathrm{WLS}}$	0.27 105%(0.58)	0.70^{**} 26%(0.44)	0.72^{***} 12%(0.29)	0.72^{***} 16%(0.24)	0.54^{***} 23%(0.21)	0.74^{***} 15%(0.18)
GMM$_{\mathrm{A}}$: $\widehat{b}_1$	6.66^{*} 905%(5.56)	9.26^{***} 267%(3.52)	5.66^{***} 80%(2.54)	4.62^{***} 48%(1.74)	4.06^{***} 93%(1.52)	4.21^{***} 99%(0.90)
GMM$_{\mathrm{A}}$: $\widehat{b}_2$	×	9.49^{***} 276%(2.08)	6.19^{***} 96%(1.86)	5.58^{***} 79%(1.46)	4.02^{***} 90%(1.34)	3.55^{***} 68%(0.79)
GMM$_{\mathrm{B}}$: $\widehat{b}_1$	×	10.50^{***} 316%(4.90)	6.20^{***} 97%(3.11)	4.97^{***} 60%(2.06)	4.20^{***} 99%(1.77)	4.26^{***} 101%(1.06)
GMM$_{\mathrm{B}}$: $\widehat{b}_2$	6.36^{**} 860%(4.66)	8.59^{***} 241%(2.91)	5.00^{***} 59%(2.29)	4.99^{***} 60%(1.75)	3.32^{***} 57%(1.54)	3.09^{***} 46%(0.93)
GMM$_{\mathrm{C}}$: $\widehat{b}$	4.51 581%(6.78)	11.58^{***} 359%(5.63)	6.16^{***} 95%(2.98)	6.00^{***} 93%(2.34)	3.84^{***} 82%(1.71)	4.01^{***} 90%(1.06)

3.6.6 Momentum risk premium in RUH model

The following five panels report the estimate values, the bias from the risk premium in percent values, the standard error in parenthesis, and t-statistics. Each panel corresponds to a one set of test portfolios. T is 60 (January 2001 - December 2005), 120, 240, 360, 480 and 948 (January 1927 - December 2005) monthly observations. The two-pass (Fama-MacBeth) cross-sectional estimate is the slope coefficient λ in $E(R^e) = \beta\lambda$ calculated by OLS, GLS and WLS. Next, we turn from beta representation to a discount factor formulation for GMM approach. The $\mathrm{GMM_A}$ and $\mathrm{GMM_B}$ first and second-stage estimates are the parameters $100 \times b$ in $E(R^e) = E(R^e f')\, b$ (returns on second moments, following Hansen and Jagannathan [49]) and $E(R^e) = E\left(R^e \widetilde{f}'\right) b$ (returns on covariances, following Cochrane [25]) respectively. The $\mathrm{GMM_C}$ is the continuous updating estimate (following Hansen, Heaton and Yaron [47]).

Panel A. N Portfolios formed on ME

	T = 60	**T = 120**	**T = 240**	**T = 360**	**T = 480**	**T = 948**
$\lambda = Ef$	0.34	0.86	0.84	0.89	0.85	0.76
$b = \frac{\lambda}{E(R^{em2})}$	0.86	2.56	3.99	4.85	4.75	3.37
			N = 5			
Beta: $\widehat{\lambda}_{\text{OLS}}$	0.59 72%(2.86)	3.60 316%(3.06)	2.74 225%(2.00)	2.94 229%(2.32)	3.16 274%(4.09)	1.22 59%(1.91)
Beta: $\widehat{\lambda}_{\text{GLS}}$	0.11 −68%(2.00)	4.83* 458%(2.88)	2.41 187%(1.87)	2.00 123%(1.77)	1.64 95%(3.05)	1.16 52%(1.63)
Beta: $\widehat{\lambda}_{\text{WLS}}$	0.68 99%(3.98)	3.04 251%(3.45)	2.44 190%(3.00)	3.45* 285%(2.33)	5.67 571%(11.98)	1.14 49%(2.28)
GMM_A: $\widehat{b}_1$	×	12.12* 374%(6.45)	12.56* 215%(6.82)	15.04* 210%(10.29)	28.93 509%(51.16)	8.69 158%(9.94)
GMM_A: $\widehat{b}_2$	×	13.48** 427%(6.13)	11.77* 195%(6.73)	11.19 131%(8.65)	8.02 69%(30.16)	8.23 144%(8.18)
GMM_B: $\widehat{b}_1$	×	12.69 396%(9.57)	14.84* 272%(9.11)	17.66 264%(13.84)	19.19 304%(33.14)	9.88 193%(12.57)
GMM_B: $\widehat{b}_2$	×	16.29** 537%(8.31)	13.86* 248%(9.02)	12.76 163%(11.87)	6.40 35%(26.32)	9.28 175%(10.32)
GMM_C: $\widehat{b}$	×	18.17* 610%(11.91)	14.74* 270%(9.10)	14.67 202%(12.34)	×	9.33 177%(12.38)
			N = 10			
Beta: $\widehat{\lambda}_{\text{OLS}}$	−0.85 −350%(2.02)	3.54** 310%(1.94)	2.03** 141%(1.39)	2.22** 148%(1.37)	1.49 76%(1.67)	−0.13 −118%(0.75)
Beta: $\widehat{\lambda}_{\text{GLS}}$	0.61 77%(1.28)	3.75*** 334%(1.53)	1.67** 98%(0.98)	1.68*** 87%(0.79)	0.54 −36%(0.99)	−0.05 −106%(0.61)
Beta: $\widehat{\lambda}_{\text{WLS}}$	0.12 −65%(2.30)	3.55*** 311%(1.82)	2.06* 145%(1.46)	2.65*** 197%(1.10)	1.62 91%(2.43)	0.35 −54%(1.06)
GMM_A: $\widehat{b}_1$	×	11.92*** 366%(4.55)	10.65*** 167%(5.62)	11.69** 141%(7.10)	14.85 213%(15.86)	1.69 −50%(4.39)
GMM_A: $\widehat{b}_2$	−6.93 −902%(4.31)	11.37*** 345%(3.48)	8.42*** 111%(3.70)	8.89*** 83%(3.77)	4.45 −6%(5.96)	1.68 −50%(3.47)
GMM_B: $\widehat{b}_1$	×	12.40*** 385%(5.70)	11.00** 176%(6.28)	12.94** 167%(8.63)	8.80 85%(9.95)	1.20 −64%(4.68)
GMM_B: $\widehat{b}_2$	−1.32 −253%(6.33)	12.99*** 408%(4.53)	8.54** 114%(4.62)	9.40*** 94%(4.86)	3.61 −24%(6.17)	1.26 −63%(3.69)
GMM_C: $\widehat{b}$	×	14.33*** 460%(6.00)	11.28** 183%(6.32)	11.48** 137%(8.04)	9.26 95%(9.57)	1.24 −63%(4.68)

Panel B. N Portfolios formed on BE/ME

	$\mathbf{T = 60}$	$\mathbf{T = 120}$	$\mathbf{T = 240}$	$\mathbf{T = 360}$	$\mathbf{T = 480}$	$\mathbf{T = 948}$
$\lambda = Ef$	0.34	0.86	0.84	0.89	0.85	0.76
$b = \frac{\lambda}{E(R^{em2})}$	0.86	2.56	3.99	4.85	4.75	3.37
			N = 5			
Beta: $\widehat{\lambda}_{\text{OLS}}$	3.45 908%(3.85)	$\times$	−4.24 −605%(10.56)	−0.06 −106%(2.03)	−1.61 −290%(2.47)	0.33 −56%(1.14)
Beta: $\widehat{\lambda}_{\text{GLS}}$	1.80 427%(3.02)	9.43 990%(26.81)	−4.01 −577%(8.98)	0.54 −39%(1.86)	−0.55 −165%(1.95)	0.50 −35%(1.10)
Beta: $\widehat{\lambda}_{\text{WLS}}$	3.41 898%(3.83)	$\times$	−4.40 −623%(10.14)	0.07 −92%(1.98)	−1.61 −290%(2.57)	0.40 −47%(1.28)
GMM_A: $\widehat{b}_1$	$\times$	25.46 895%(56.97)	−21.17 −631%(83.59)	1.22 −75%(10.87)	−6.66 −240%(15.76)	3.62 7%(5.94)
GMM_A: $\widehat{b}_2$	$\times$	25.86 911%(53.47)	−13.23 −432%(66.87)	3.84 −21%(10.09)	−0.05 −101%(13.14)	4.20 25%(5.82)
GMM_B: $\widehat{b}_1$	$\times$	$\times$	−20.99 −627%(64.08)	0.30 −94%(12.40)	−8.92 −288%(16.69)	3.66 8%(6.61)
GMM_B: $\widehat{b}_2$	$\times$	$\times$	−15.51 −489%(52.51)	3.77 −22%(11.33)	−1.18 −125%(13.39)	4.38 30%(6.45)
GMM_C: $\widehat{b}$	$\times$	$\times$	−22.69 −669%(67.34)	3.49 −28%(12.30)	−4.69 −199%(15.07)	4.82 43%(6.80)
			N = 10			
Beta: $\widehat{\lambda}_{\text{OLS}}$	1.26 267%(1.86)	0.09 −89%(1.63)	−1.52 −281%(1.52)	−0.62 −169%(1.06)	−1.12 −233%(1.01)	1.08 42%(1.07)
Beta: $\widehat{\lambda}_{\text{GLS}}$	0.07 −79%(1.49)	0.24 −73%(1.62)	−1.41 −268%(1.35)	−0.83 −193%(0.97)	−0.91 −208%(0.94)	1.07 39%(0.96)
Beta: $\widehat{\lambda}_{\text{WLS}}$	0.85 148%(1.58)	0.12 −86%(1.84)	−1.39 −266%(1.54)	−0.77 −187%(1.06)	−1.02 −221%(1.11)	0.93 22%(1.05)
GMM_A: $\widehat{b}_1$	3.77 336%(12.43)	2.31 −10%(5.30)	−6.63 −266%(8.60)	−2.45 −151%(6.31)	−5.21 −210%(6.72)	8.98* 166%(6.63)
GMM_A: $\widehat{b}_2$	−0.28 −132%(8.48)	3.07 20%(5.23)	−6.41 −261%(8.01)	−3.90 −180%(5.78)	−4.46 −194%(6.53)	9.76** 189%(5.38)
GMM_B: $\widehat{b}_1$	4.08 372%(13.46)	1.60 −38%(5.92)	−7.02 −276%(7.98)	−2.96 −161%(6.43)	−5.96 −226%(6.63)	8.11 140%(7.39)
GMM_B: $\widehat{b}_2$	−0.60 −170%(9.19)	2.15 −16%(5.87)	−6.85 −272%(7.41)	−4.48 −192%(5.81)	−4.83 −202%(6.29)	7.43* 120%(6.25)
GMM_C: $\widehat{b}$	−7.13 −925%(9.84)	1.79 −30%(6.02)	−7.78 −295%(8.10)	−4.83 −200%(6.70)	−5.78 −222%(6.63)	$\times$

Panel C. N Portfolios formed by the intersections of ME and BE/ME

	T = 60	**T = 120**	**T = 240**	**T = 360**	**T = 480**	**T = 948**
$\lambda = Ef$	0.34	0.86	0.84	0.89	0.85	0.76
$b = \frac{\lambda}{E\left(R^{em2}\right)}$	0.86	2.56	3.99	4.85	4.75	3.37
				N = 6		
Beta: $\widehat{\lambda}_{\mathrm{OLS}}$	0.84 147%(1.88)	6.06^{*} 601%(3.85)	5.05^{**} 501%(2.95)	4.97^{**} 456%(2.75)	8.27^{**} 878%(4.26)	0.44 −43%(1.31)
Beta: $\widehat{\lambda}_{\mathrm{GLS}}$	2.76^{**} 706%(1.54)	5.57^{***} 544%(1.89)	5.51^{***} 555%(1.65)	7.13^{***} 697%(2.64)	7.34^{**} 768%(3.66)	2.48^{***} 224%(0.88)
Beta: $\widehat{\lambda}_{\mathrm{WLS}}$	0.17 −51%(2.65)	5.83^{**} 575%(3.13)	4.75^{**} 465%(2.81)	4.45^{***} 398%(1.59)	4.15 391%(3.64)	-0.26 −134%(1.32)
GMM$_{\mathrm{A}}$: $\widehat{b}_1$	-3.36 −489%(12.47)	16.41^{**} 541%(7.84)	20.18^{**} 406%(10.62)	21.99^{**} 353%(11.88)	39.56^{*} 733%(25.18)	7.90^{*} 134%(6.02)
GMM$_{\mathrm{A}}$: $\widehat{b}_2$	×	16.22^{***} 534%(4.13)	21.84^{***} 448%(6.23)	29.05^{***} 499%(7.41)	31.98^{**} 574%(15.34)	16.34^{***} 384%(3.76)
GMM$_{\mathrm{B}}$: $\widehat{b}_1$	3.60 317%(11.99)	22.09^{**} 764%(13.11)	26.73^{*} 571%(18.54)	29.56^{*} 509%(21.13)	51.11^{*} 977%(32.97)	4.53 34%(7.33)
GMM$_{\mathrm{B}}$: $\widehat{b}_2$	×	21.94^{***} 758%(7.46)	29.54^{***} 641%(11.48)	41.13^{***} 748%(14.31)	43.19^{*} 810%(29.90)	16.47^{***} 388%(4.21)
GMM$_{\mathrm{C}}$: $\widehat{b}$	×	23.17^{**} 806%(13.73)	31.67^{*} 695%(21.34)	49.00^{*} 910%(33.80)	×	×
				N = 25		
Beta: $\widehat{\lambda}_{\mathrm{OLS}}$	-0.23 −167%(1.19)	4.88^{***} 464%(2.02)	4.22^{***} 402%(1.46)	4.20^{***} 370%(1.74)	3.88^{***} 359%(1.09)	0.91^{**} 19%(0.63)
Beta: $\widehat{\lambda}_{\mathrm{GLS}}$	1.34^{**} 292%(0.93)	2.81^{***} 225%(0.83)	3.77^{***} 349%(0.68)	3.17^{***} 255%(0.68)	2.85^{***} 237%(0.68)	1.27^{***} 66%(0.37)
Beta: $\widehat{\lambda}_{\mathrm{WLS}}$	-0.01 −103%(1.02)	4.00^{***} 362%(1.39)	3.58^{***} 326%(1.22)	3.64^{***} 307%(0.86)	2.22^{***} 163%(1.09)	0.71^{***} −8%(0.41)
GMM$_{\mathrm{A}}$: $\widehat{b}_1$	-3.74 −532%(4.95)	15.04^{***} 488%(4.62)	18.48^{***} 364%(5.70)	20.19^{***} 316%(9.00)	25.55^{***} 438%(9.27)	9.07^{***} 169%(3.58)
GMM$_{\mathrm{A}}$: $\widehat{b}_2$	-1.94 −325%(3.66)	14.60^{***} 471%(1.96)	19.31^{***} 384%(2.53)	18.20^{***} 275%(3.06)	16.96^{***} 257%(3.85)	11.42^{***} 238%(2.04)
GMM$_{\mathrm{B}}$: $\widehat{b}_1$	-2.28 −364%(5.83)	18.19^{***} 611%(6.42)	22.55^{***} 466%(8.66)	25.20^{***} 419%(14.34)	24.65^{***} 419%(8.27)	7.46^{***} 121%(3.92)
GMM$_{\mathrm{B}}$: $\widehat{b}_2$	4.45 415%(4.34)	17.25^{***} 575%(3.21)	22.83^{***} 473%(4.08)	20.23^{***} 317%(4.74)	17.53^{***} 269%(4.90)	9.85^{***} 192%(2.37)
GMM$_{\mathrm{C}}$: $\widehat{b}$	×	×	41.63^{***} 944%(14.61)	40.14^{***} 727%(21.53)	44.70^{***} 842%(14.31)	×
				N = 100		
Beta: $\widehat{\lambda}_{\mathrm{OLS}}$		2.61^{***} 202%(0.87)	2.04^{***} 143%(0.55)	2.43^{***} 172%(0.68)	1.91^{***} 126%(0.42)	0.43^{*} −43%(0.40)
Beta: $\widehat{\lambda}_{\mathrm{GLS}}$		0.67^{*} −23%(0.53)	1.34^{***} 60%(0.35)	0.81^{***} −10%(0.30)	1.00^{***} 18%(0.30)	0.63^{***} −18%(0.24)
Beta: $\widehat{\lambda}_{\mathrm{WLS}}$		2.35^{***} 172%(0.66)	1.71^{***} 103%(0.42)	2.18^{***} 144%(0.36)	1.59^{***} 89%(0.37)	0.41^{***} −46%(0.24)
GMM$_{\mathrm{A}}$: $\widehat{b}_1$		9.77^{***} 282%(2.31)	10.90^{***} 173%(2.54)	13.76^{***} 184%(3.06)	13.81^{***} 191%(2.45)	5.74^{***} 70%(2.12)
GMM$_{\mathrm{A}}$: $\widehat{b}_2$		15.65^{***} 512%(0.48)	13.44^{***} 237%(0.95)	13.06^{***} 169%(1.27)	14.02^{***} 195%(1.40)	7.72^{***} 129%(1.10)
GMM$_{\mathrm{B}}$: $\widehat{b}_1$		10.42^{***} 307%(3.66)	11.38^{***} 185%(3.61)	14.85^{***} 206%(4.11)	12.63^{***} 166%(3.00)	4.86^{***} 44%(2.21)
GMM$_{\mathrm{B}}$: $\widehat{b}_2$		-3.75 −247%(0.96)	8.60^{***} 116%(1.51)	4.13^{***} −15%(1.75)	7.63^{***} 61%(1.74)	5.31^{***} 57%(1.24)
GMM$_{\mathrm{C}}$: $\widehat{b}$		×	×	24.09^{***} 397%(11.76)	×	22.49^{***} 566%(11.01)

Panel D. N Portfolios formed on ME MOM

	T = 60	**T = 120**	**T = 240**	**T = 360**	**T = 480**	**T = 948**
$\lambda = Ef$	0.34	0.86	0.84	0.89	0.85	0.76
$b = \frac{\lambda}{E(R^{em2})}$	0.86	2.56	3.99	4.85	4.75	3.37
			N = 6			
Beta: $\widehat{\lambda}_{\text{OLS}}$	0.14 −58%(0.81)	0.88** 1%(0.52)	0.81*** −4%(0.29)	0.87*** −3%(0.22)	0.82*** −3%(0.19)	0.75*** −2%(0.15)
Beta: $\widehat{\lambda}_{\text{GLS}}$	0.34 0%(0.80)	0.86** 0%(0.52)	0.84*** 0%(0.29)	0.89*** 0%(0.22)	0.85*** 0%(0.19)	0.76*** 0%(0.15)
Beta: $\widehat{\lambda}_{\text{WLS}}$	0.03 −92%(0.85)	0.58 −33%(0.55)	0.50** −40%(0.31)	0.55*** −38%(0.23)	0.54*** −36%(0.20)	0.58*** −23%(0.16)
GMM_A: $\widehat{b}_1$	×	3.60** 41%(1.90)	4.46*** 12%(1.78)	4.90*** 1%(1.51)	5.32*** 12%(1.37)	6.56*** 94%(1.01)
GMM_A: $\widehat{b}_2$	−1.34 −255%(3.80)	4.93*** 93%(1.78)	5.21*** 31%(1.73)	6.06*** 25%(1.47)	6.37*** 34%(1.33)	6.57*** 95%(0.97)
GMM_B: $\widehat{b}_1$	×	3.74** 46%(2.23)	4.74** 19%(2.12)	5.10*** 5%(1.79)	5.48*** 15%(1.62)	7.20*** 113%(1.28)
GMM_B: $\widehat{b}_2$	0.12 −86%(4.91)	4.45** 74%(2.12)	3.66** −8%(2.06)	4.22*** −13%(1.74)	5.00*** 5%(1.57)	6.69*** 98%(1.22)
GMM_C: $\widehat{b}$	×	0.37 −86%(2.67)	7.75*** 95%(2.72)	−0.37 −108%(4.07)	2.39 −50%(3.18)	7.74*** 129%(1.34)
			N = 25			
Beta: $\widehat{\lambda}_{\text{OLS}}$	0.17 −50%(0.82)	1.02*** 18%(0.53)	0.92*** 9%(0.30)	0.98*** 9%(0.23)	0.96*** 13%(0.19)	0.89*** 17%(0.16)
Beta: $\widehat{\lambda}_{\text{GLS}}$	0.39 13%(0.80)	0.82** −5%(0.52)	0.86*** 2%(0.29)	0.93*** 4%(0.22)	0.92*** 9%(0.19)	0.84*** 10%(0.16)
Beta: $\widehat{\lambda}_{\text{WLS}}$	0.04 −88%(0.83)	0.76** −13%(0.54)	0.69*** −18%(0.30)	0.74*** −17%(0.23)	0.74*** −12%(0.20)	0.79*** 4%(0.16)
GMM_A: $\widehat{b}_1$	−6.63 −868%(4.04)	4.40*** 72%(1.99)	5.03*** 26%(1.89)	5.60*** 15%(1.63)	6.18*** 30%(1.45)	7.76*** 130%(1.01)
GMM_A: $\widehat{b}_2$	−1.94 −325%(2.43)	8.42*** 229%(1.57)	10.00*** 151%(1.45)	9.75*** 101%(1.31)	9.48*** 100%(1.19)	7.99*** 137%(0.91)
GMM_B: $\widehat{b}_1$	−6.94 −904%(5.44)	4.68*** 83%(2.45)	5.45*** 37%(2.32)	6.14*** 27%(2.06)	6.74*** 42%(1.83)	8.75*** 159%(1.34)
GMM_B: $\widehat{b}_2$	6.72*** 678%(3.07)	5.09*** 99%(2.07)	5.95*** 49%(1.92)	6.44*** 33%(1.73)	7.00*** 47%(1.56)	8.09*** 140%(1.20)
GMM_C: $\widehat{b}$	×	2.28 −11%(3.01)	10.22** 156%(7.21)	16.01*** 230%(8.72)	14.69*** 209%(5.07)	10.50*** 211%(1.48)

Panel E. N Industry Portfolios

	T = 60	**T = 120**	**T = 240**	**T = 360**	**T = 480**	**T = 948**
$\lambda = Ef$	0.34	0.86	0.84	0.89	0.85	0.76
$b = \frac{\lambda}{E(R^{em2})}$	0.86	2.56	3.99	4.85	4.75	3.37
			N = 5			
Beta: $\widehat{\lambda}_{\text{OLS}}$	−2.21 −746%(3.50)	2.62 203%(4.18)	2.04 142%(1.57)	1.31 47%(1.30)	1.11 31%(0.94)	1.29* 69%(0.83)
Beta: $\widehat{\lambda}_{\text{GLS}}$	−1.79 −625%(3.24)	3.13 262%(4.10)	2.08* 148%(1.44)	0.86 −4%(1.12)	0.75 −11%(0.81)	0.83 8%(0.71)
Beta: $\widehat{\lambda}_{\text{WLS}}$	−1.56 −556%(2.86)	3.11 260%(4.14)	2.02* 141%(1.40)	0.81 −9%(1.10)	0.73 −14%(0.78)	0.79 4%(0.63)
GMM_A: $\widehat{b}_1$	×	8.63 238%(14.80)	9.49 138%(8.17)	7.07 46%(7.59)	6.57 38%(5.83)	7.81* 131%(4.75)
GMM_A: $\widehat{b}_2$	×	10.06 293%(14.15)	9.10 128%(7.30)	4.37 −10%(6.57)	4.27 −10%(4.97)	5.57 65%(4.35)
GMM_B: $\widehat{b}_1$	×	9.99 291%(19.65)	10.89 173%(10.56)	7.74 60%(8.97)	7.01 48%(6.69)	8.15* 141%(5.48)
GMM_B: $\widehat{b}_2$	×	11.48 349%(19.06)	10.32 159%(9.42)	4.49 −7%(7.62)	4.35 −8%(5.64)	5.60 66%(5.02)
GMM_C: $\widehat{b}$	×	13.45 426%(23.34)	10.67 168%(10.49)	5.13 6%(8.07)	4.77 1%(6.09)	6.33 88%(5.18)
			N = 17			
Beta: $\widehat{\lambda}_{\text{OLS}}$	1.56* 355%(1.34)	1.27 46%(1.20)	1.41*** 68%(0.80)	1.25** 40%(0.71)	1.04** 23%(0.69)	0.63* −17%(0.51)
Beta: $\widehat{\lambda}_{\text{GLS}}$	2.74*** 702%(1.10)	1.00 15%(0.95)	0.74* −12%(0.64)	0.69* −23%(0.54)	0.30 −65%(0.53)	0.22 −72%(0.42)
Beta: $\widehat{\lambda}_{\text{WLS}}$	1.41* 314%(1.12)	1.05 22%(1.01)	1.18*** 40%(0.63)	1.21*** 36%(0.56)	0.88** 4%(0.53)	0.56* −27%(0.44)
GMM_A: $\widehat{b}_1$	5.88 580%(10.02)	5.51** 115%(3.83)	7.43*** 86%(3.95)	7.48*** 54%(3.91)	6.89** 45%(3.95)	5.27*** 56%(2.92)
GMM_A: $\widehat{b}_2$	×	4.24** 66%(2.93)	3.66* −8%(3.03)	3.34* −31%(2.91)	1.72 −64%(3.06)	2.58 −23%(2.52)
GMM_B: $\widehat{b}_1$	7.31 746%(10.89)	5.41* 111%(4.55)	7.82** 96%(4.75)	7.53** 55%(4.69)	6.69** 41%(4.56)	4.71** 40%(3.19)
GMM_B: $\widehat{b}_2$	×	2.85 12%(3.44)	3.03 −24%(3.53)	2.39 −51%(3.33)	0.31 −94%(3.34)	1.81 −46%(2.70)
GMM_C: $\widehat{b}$	×	4.88* 91%(4.53)	5.37* 35%(4.39)	4.57 −6%(4.28)	2.91 −39%(4.17)	3.29 −3%(3.09)
			N = 30			
Beta: $\widehat{\lambda}_{\text{OLS}}$	1.12* 228%(1.04)	2.64*** 206%(1.38)	1.38*** 65%(0.76)	0.89** −1%(0.61)	1.02** 21%(0.62)	0.97*** 26%(0.49)
Beta: $\widehat{\lambda}_{\text{GLS}}$	0.97* 185%(0.86)	0.61 −29%(0.70)	−0.01 −101%(0.49)	0.40 −55%(0.43)	0.31 −63%(0.44)	0.23 −70%(0.39)
Beta: $\widehat{\lambda}_{\text{WLS}}$	0.65 92%(0.94)	1.53*** 77%(0.88)	0.89** 6%(0.53)	0.98*** 9%(0.45)	0.83*** −2%(0.43)	0.57** −25%(0.38)
GMM_A: $\widehat{b}_1$	2.25 160%(5.93)	9.37*** 266%(3.90)	7.44*** 87%(3.64)	5.76*** 19%(3.36)	7.13*** 50%(3.71)	7.24*** 115%(3.18)
GMM_A: $\widehat{b}_2$	7.10*** 722%(3.73)	5.78*** 126%(1.93)	3.72** −7%(2.45)	5.41*** 11%(2.42)	4.63*** −2%(2.68)	3.94** 17%(2.60)
GMM_B: $\widehat{b}_1$	3.49 304%(6.82)	10.30*** 303%(5.24)	7.66*** 92%(4.34)	5.39** 11%(3.85)	6.63** 40%(4.19)	6.31*** 87%(3.42)
GMM_B: $\widehat{b}_2$	5.07* 487%(4.38)	4.80*** 88%(2.51)	1.77 −56%(2.85)	3.51* −28%(2.80)	2.47 −48%(3.03)	1.78 −47%(2.81)
GMM_C: $\widehat{b}$	1.44 67%(6.63)	8.71** 241%(5.83)	4.79* 20%(4.08)	8.32*** 72%(4.43)	5.18* 9%(4.04)	6.12*** 81%(3.33)

3.6.7 Value risk premium in RUH model

The following five panels report the estimate values, the bias from the risk premium in percent values, the standard error in parenthesis, and t-statistics. Each panel corresponds to a one set of test portfolios. T is 60 (January 2001 - December 2005), 120, 240, 360, 480 and 948 (January 1927 - December 2005) monthly observations. The two-pass (Fama-MacBeth) cross-sectional estimate is the slope coefficient λ in $E(R^e) = \beta\lambda$ calculated by OLS, GLS and WLS. Next, we turn from beta representation to a discount factor formulation for GMM approach. The $\mathrm{GMM_A}$ and $\mathrm{GMM_B}$ first and second-stage estimates are the parameters $100 \times b$ in $E(R^e) = E(R^e f')\, b$ (returns on second moments, following Hansen and Jagannathan [49]) and $E(R^e) = E\left(R^e \widetilde{f}'\right) b$ (returns on covariances, following Cochrane [25]) respectively. The $\mathrm{GMM_C}$ is the continuous updating estimate (following Hansen, Heaton and Yaron [47]).

Panel A. N Portfolios formed on ME

	T = 60	T = 120	T = 240	T = 360	T = 480	T = 948
$\lambda = Ef$	0.78	0.46	0.34	0.43	0.43	0.41
$b = \frac{\lambda}{E(R^{em2})}$	7.50	2.90	3.26	4.44	4.70	3.10
			N = 5			
Beta: $\widehat{\lambda}_{\text{OLS}}$	3.58^{*} 358%(2.26)	−0.30 −165%(0.89)	0.28 −17%(0.78)	0.31 −27%(0.90)	−0.27 −164%(1.22)	0.70 71%(0.55)
Beta: $\widehat{\lambda}_{\text{GLS}}$	2.05^{**} 163%(0.91)	−0.29 −163%(0.96)	0.23 −32%(0.72)	0.24 −43%(0.80)	0.44 3%(0.76)	0.67^{*} 64%(0.48)
Beta: $\widehat{\lambda}_{\text{WLS}}$	3.78^{**} 383%(1.31)	−0.16 −136%(1.29)	0.29 −16%(0.88)	0.33 −22%(1.57)	1.18 175%(2.16)	0.69^{*} 68%(0.41)
GMM_A: $\widehat{b}_1$	42.12^{*} 462%(23.25)	5.30 83%(10.78)	8.71 167%(10.16)	10.21 130%(13.62)	14.25 203%(18.46)	8.17 163%(7.45)
GMM_A: $\widehat{b}_2$	34.16^{***} 356%(11.19)	6.87 137%(9.95)	8.51 161%(10.11)	8.18 84%(11.75)	8.70 85%(15.83)	7.76 150%(6.26)
GMM_B: $\widehat{b}_1$	59.19 690%(47.46)	4.54 57%(13.09)	10.29 215%(12.95)	11.04 149%(16.59)	3.25 −31%(18.45)	9.28 199%(9.63)
GMM_B: $\widehat{b}_2$	42.24^{**} 463%(21.80)	7.29 152%(11.94)	9.74 199%(12.83)	9.11 105%(15.21)	9.45 101%(14.85)	8.76 182%(8.05)
GMM_C: $\widehat{b}$	47.51 534%(40.64)	10.17 251%(17.09)	10.55 223%(12.94)	9.90 123%(14.94)	×	8.80 183%(9.49)
			N = 10			
Beta: $\widehat{\lambda}_{\text{OLS}}$	2.85^{***} 265%(1.48)	−0.39 −185%(0.79)	0.09 −73%(0.61)	−0.11 −126%(0.55)	−0.46 −208%(0.80)	0.42^{*} 3%(0.37)
Beta: $\widehat{\lambda}_{\text{GLS}}$	1.16^{**} 49%(0.64)	−0.23 −150%(0.66)	0.05 −84%(0.42)	−0.10 −123%(0.43)	0.14 −67%(0.39)	0.35^{*} −14%(0.27)
Beta: $\widehat{\lambda}_{\text{WLS}}$	3.55^{***} 354%(1.00)	−0.09 −119%(0.66)	0.15 −56%(0.40)	−0.10 −124%(0.50)	0.04 −91%(0.50)	0.54^{***} 32%(0.26)
GMM_A: $\widehat{b}_1$	42.62^{***} 469%(19.54)	4.32 49%(8.37)	6.64 103%(7.50)	4.45 0%(7.08)	2.73 −42%(8.60)	3.48 12%(3.99)
GMM_A: $\widehat{b}_2$	30.76^{***} 310%(10.56)	5.05^{*} 75%(4.47)	6.53^{**} 100%(4.31)	4.70 6%(5.09)	5.28 12%(5.52)	3.10 0%(3.05)
GMM_B: $\widehat{b}_1$	53.80^{**} 618%(35.46)	3.60 24%(9.91)	6.91 112%(8.71)	4.33 −3%(8.30)	−1.90 −140%(11.03)	3.16 2%(4.37)
GMM_B: $\widehat{b}_2$	2.72 −64%(17.38)	5.24 81%(5.48)	6.34^{*} 94%(5.19)	4.50 1%(6.11)	4.82 2%(5.89)	2.82 −9%(3.30)
GMM_C: $\widehat{b}$	×	5.60 94%(10.91)	7.29 123%(8.87)	5.30 19%(7.93)	6.97 48%(11.11)	2.79 −10%(4.36)

Panel B. N Portfolios formed on BE/ME

	T = 60	**T = 120**	**T = 240**	**T = 360**	**T = 480**	**T = 948**
$\lambda = Ef$	0.78	0.46	0.34	0.43	0.43	0.41
$b = \frac{\lambda}{E(R^{em2})}$	7.50	2.90	3.26	4.44	4.70	3.10
			N = 5			
Beta: $\widehat{\lambda}_{\mathrm{OLS}}$	1.08^{*} 38%(0.58)	0.76 66%(1.39)	−0.02 −106%(0.53)	0.34^{**} −20%(0.18)	0.29^{*} −32%(0.16)	0.31^{**} −25%(0.13)
Beta: $\widehat{\lambda}_{\mathrm{GLS}}$	0.86^{*} 10%(0.50)	0.68 48%(1.12)	−0.01 −103%(0.48)	0.34^{**} −21%(0.18)	0.32^{**} −26%(0.16)	0.32^{**} −22%(0.13)
Beta: $\widehat{\lambda}_{\mathrm{WLS}}$	1.10^{**} 41%(0.54)	0.75 63%(1.45)	−0.03 −108%(0.50)	0.33^{*} −23%(0.18)	0.28^{*} −34%(0.16)	0.32^{**} −23%(0.13)
$\mathrm{GMM_A}$: $\widehat{b}_1$	1.53 −80%(13.62)	16.04 454%(24.33)	−1.53 −147%(27.31)	7.20^{**} 62%(2.72)	4.32 −8%(4.39)	3.24 4%(2.92)
$\mathrm{GMM_A}$: $\widehat{b}_2$	2.80 −63%(13.33)	16.24 461%(22.76)	1.34 −59%(21.29)	7.61^{**} 71%(2.63)	6.21^{*} 32%(3.69)	3.56 15%(2.85)
$\mathrm{GMM_B}$: $\widehat{b}_1$	1.60 −79%(14.88)	25.08 766%(66.94)	−2.36 −172%(24.85)	7.53^{**} 70%(3.87)	3.47 −26%(5.65)	3.35 8%(3.36)
$\mathrm{GMM_B}$: $\widehat{b}_2$	3.22 −57%(14.59)	25.61 784%(64.56)	−0.01 −100%(19.85)	8.35^{**} 88%(3.66)	6.15 31%(4.54)	3.75 21%(3.27)
$\mathrm{GMM_C}$: $\widehat{b}$	−1.47 −120%(17.06)	28.89 897%(77.99)	−2.62 −180%(26.09)	8.18^{**} 84%(3.77)	5.09 8%(5.09)	3.95 27%(3.43)
			N = 10			
Beta: $\widehat{\lambda}_{\mathrm{OLS}}$	0.81^{**} 3%(0.46)	0.30 −34%(0.41)	0.09 −74%(0.24)	0.35^{***} −19%(0.18)	0.31^{***} −28%(0.15)	0.32^{***} −23%(0.13)
Beta: $\widehat{\lambda}_{\mathrm{GLS}}$	0.65^{**} −17%(0.43)	0.34 −25%(0.41)	0.09 −73%(0.24)	0.32^{**} −25%(0.18)	0.31^{***} −29%(0.15)	0.34^{***} −18%(0.13)
Beta: $\widehat{\lambda}_{\mathrm{WLS}}$	0.83^{**} 6%(0.44)	0.30 −34%(0.39)	0.09 −74%(0.23)	0.33^{**} −23%(0.18)	0.30^{***} −30%(0.15)	0.32^{***} −22%(0.13)
$\mathrm{GMM_A}$: $\widehat{b}_1$	7.31 −2%(9.12)	5.87^{**} 103%(3.71)	2.75 −16%(3.95)	6.68^{***} 50%(2.27)	4.71^{***} 0%(2.36)	5.67^{***} 83%(2.97)
$\mathrm{GMM_A}$: $\widehat{b}_2$	11.22^{**} 50%(6.78)	6.91^{***} 139%(3.61)	3.23 −1%(3.61)	6.36^{***} 43%(2.17)	5.03^{***} 7%(2.27)	5.74^{***} 85%(2.54)
$\mathrm{GMM_B}$: $\widehat{b}_1$	7.82 4%(10.19)	5.94^{*} 105%(4.51)	2.50 −23%(4.15)	6.86^{***} 55%(2.72)	4.42^{**} −6%(2.68)	5.38^{**} 73%(3.51)
$\mathrm{GMM_B}$: $\widehat{b}_2$	11.09^{*} 48%(7.89)	6.65^{**} 130%(4.44)	2.96 −9%(3.77)	6.35^{***} 43%(2.55)	4.91^{***} 4%(2.52)	4.88^{**} 57%(3.08)
$\mathrm{GMM_C}$: $\widehat{b}$	15.20^{**} 103%(9.62)	6.72^{**} 132%(4.60)	2.77 −15%(4.23)	6.36^{***} 43%(2.78)	4.70^{**} 0%(2.67)	20.62^{*} 565%(14.61)

Panel C. N Portfolios formed by the intersections of ME and BE/ME

	T = 60	**T = 120**	**T = 240**	**T = 360**	**T = 480**	**T = 948**
$\lambda = Ef$	0.78	0.46	0.34	0.43	0.43	0.41
$b = \frac{\lambda}{E\left(R^{em2}\right)}$	7.50	2.90	3.26	4.44	4.70	3.10
			N = 6			
Beta: $\widehat{\lambda}_{\mathrm{OLS}}$	0.66^{*} −16%(0.40)	0.47 2%(0.39)	0.38^{**} 12%(0.23)	0.45^{***} 6%(0.17)	0.35^{**} −18%(0.17)	0.40^{***} −3%(0.12)
Beta: $\widehat{\lambda}_{\mathrm{GLS}}$	0.78^{**} 0%(0.40)	0.46^{*} 0%(0.36)	0.34^{**} 0%(0.21)	0.43^{***} 0%(0.16)	0.43^{***} 0%(0.14)	0.41^{***} 0%(0.12)
Beta: $\widehat{\lambda}_{\mathrm{WLS}}$	0.69^{*} −12%(0.45)	0.37 −20%(0.45)	0.27 −22%(0.28)	0.29^{*} −32%(0.20)	0.31^{**} −28%(0.18)	0.28^{**} −31%(0.13)
$\mathrm{GMM_A}$: $\widehat{b}_1$	13.32 78%(11.32)	11.45^{***} 295%(3.70)	10.77^{***} 230%(3.50)	11.16^{***} 151%(3.13)	14.05^{***} 199%(5.51)	5.64^{**} 82%(2.71)
$\mathrm{GMM_A}$: $\widehat{b}_2$	−2.27 −130%(7.81)	12.59^{***} 335%(3.13)	11.62^{***} 256%(3.27)	11.65^{***} 162%(2.82)	12.74^{***} 171%(5.05)	9.15^{***} 195%(1.75)
$\mathrm{GMM_B}$: $\widehat{b}_1$	7.57 1%(10.83)	15.45^{**} 433%(6.68)	14.37^{***} 340%(6.11)	15.22^{***} 243%(5.96)	19.95^{**} 325%(10.00)	4.43^{*} 43%(3.48)
$\mathrm{GMM_B}$: $\widehat{b}_2$	−1.45 −119%(8.45)	17.04^{***} 489%(5.63)	15.79^{***} 384%(5.71)	16.73^{***} 277%(5.03)	17.85^{**} 280%(9.68)	9.92^{***} 220%(2.16)
$\mathrm{GMM_C}$: $\widehat{b}$	−4.91 −166%(22.43)	17.78^{***} 514%(7.14)	16.45^{**} 404%(7.00)	17.87^{**} 302%(8.81)	23.25^{**} 395%(12.34)	16.37^{*} 427%(12.15)
			N = 25			
Beta: $\widehat{\lambda}_{\mathrm{OLS}}$	0.60^{**} −24%(0.42)	0.41 −10%(0.42)	0.40^{**} 17%(0.25)	0.51^{***} 19%(0.17)	0.41^{***} −4%(0.16)	0.46^{***} 13%(0.12)
Beta: $\widehat{\lambda}_{\mathrm{GLS}}$	0.78^{***} −1%(0.40)	0.50^{**} 9%(0.36)	0.40^{***} 16%(0.21)	0.46^{***} 8%(0.16)	0.46^{***} 7%(0.14)	0.41^{***} −1%(0.12)
Beta: $\widehat{\lambda}_{\mathrm{WLS}}$	0.70^{**} −11%(0.42)	0.38 −18%(0.39)	0.32^{**} −6%(0.23)	0.40^{***} −7%(0.18)	0.34^{***} −22%(0.15)	0.37^{***} −9%(0.13)
$\mathrm{GMM_A}$: $\widehat{b}_1$	13.64^{***} 82%(6.23)	11.26^{***} 289%(3.57)	11.01^{***} 238%(3.69)	11.86^{***} 167%(2.95)	12.39^{***} 164%(3.16)	6.71^{***} 116%(1.77)
$\mathrm{GMM_A}$: $\widehat{b}_2$	23.79^{***} 217%(4.21)	17.65^{***} 509%(2.36)	17.25^{***} 429%(2.42)	15.43^{***} 247%(2.36)	12.05^{***} 156%(2.68)	7.47^{***} 141%(1.23)
$\mathrm{GMM_B}$: $\widehat{b}_1$	12.00^{**} 60%(7.44)	13.95^{***} 382%(5.69)	13.89^{***} 326%(5.40)	15.39^{***} 246%(4.85)	14.18^{***} 202%(3.79)	6.27^{***} 102%(2.11)
$\mathrm{GMM_B}$: $\widehat{b}_2$	8.22^{**} 10%(5.74)	20.57^{***} 610%(4.05)	20.34^{***} 523%(3.78)	17.19^{***} 287%(3.66)	13.50^{***} 187%(3.30)	7.14^{***} 130%(1.47)
$\mathrm{GMM_C}$: $\widehat{b}$	×	×	33.61^{***} 930%(10.11)	30.74^{***} 592%(7.92)	14.99^{***} 219%(6.18)	10.90 251%(16.23)
			N = 100			
Beta: $\widehat{\lambda}_{\mathrm{OLS}}$		0.24 −48%(0.42)	0.29^{*} −15%(0.25)	0.41^{***} −5%(0.18)	0.37^{***} −13%(0.16)	0.52^{***} 26%(0.12)
Beta: $\widehat{\lambda}_{\mathrm{GLS}}$		0.38^{*} −16%(0.36)	0.26^{*} −24%(0.21)	0.39^{***} −8%(0.16)	0.42^{***} −2%(0.14)	0.43^{***} 4%(0.12)
Beta: $\widehat{\lambda}_{\mathrm{WLS}}$		0.20 −56%(0.37)	0.20 −42%(0.21)	0.33^{***} −23%(0.17)	0.33^{***} −23%(0.14)	0.44^{***} 6%(0.12)
$\mathrm{GMM_A}$: $\widehat{b}_1$		8.81^{***} 204%(3.63)	8.89^{***} 172%(3.12)	10.36^{***} 133%(2.58)	9.85^{***} 110%(2.36)	5.77^{***} 86%(1.29)
$\mathrm{GMM_A}$: $\widehat{b}_2$		20.66^{***} 613%(0.92)	19.07^{***} 484%(1.35)	17.69^{***} 298%(1.54)	15.75^{***} 235%(1.54)	6.80^{***} 119%(0.94)
$\mathrm{GMM_B}$: $\widehat{b}_1$		9.76^{***} 237%(5.04)	9.92^{***} 204%(4.08)	12.05^{***} 171%(3.52)	10.59^{***} 125%(2.90)	5.52^{***} 78%(1.52)
$\mathrm{GMM_B}$: $\widehat{b}_2$		−6.96 −340%(1.53)	9.90^{***} 203%(2.09)	7.22^{***} 63%(2.24)	8.30^{***} 77%(2.02)	5.65^{***} 82%(1.09)
$\mathrm{GMM_C}$: $\widehat{b}$		×	28.64 778%(28.25)	×	44.24^{*} 841%(39.12)	8.24^{***} 165%(3.60)

Panel D. N Portfolios formed on ME MOM

	T = 60	**T = 120**	**T = 240**	**T = 360**	**T = 480**	**T = 948**
$\lambda = Ef$	0.78	0.46	0.34	0.43	0.43	0.41
$b = \frac{\lambda}{E(R^{em2})}$	7.50	2.90	3.26	4.44	4.70	3.10
			N = 6			
Beta: $\widehat{\lambda}_{\text{OLS}}$	2.66^{**} 240%(1.43)	-0.39 -185%(0.61)	-0.05 -115%(0.41)	-0.17 -139%(0.35)	0.03 -92%(0.27)	0.72^{**} 75%(0.37)
Beta: $\widehat{\lambda}_{\text{GLS}}$	0.75^{*} -4%(0.60)	0.26 -43%(0.44)	0.14 -58%(0.29)	0.30 -30%(0.25)	0.37^{*} -15%(0.23)	0.71^{***} 72%(0.24)
Beta: $\widehat{\lambda}_{\text{WLS}}$	2.74^{**} 251%(1.29)	-0.42 -190%(0.51)	-0.27 -179%(0.33)	-0.36 -185%(0.31)	-0.25 -158%(0.33)	0.73^{***} 78%(0.27)
GMM_A: $\widehat{b}_1$	40.21^{*} 436%(26.57)	1.78 -39%(4.71)	4.54 39%(4.67)	4.33 -2%(4.22)	6.41^{**} 36%(3.41)	7.63^{***} 146%(2.59)
GMM_A: $\widehat{b}_2$	11.90 59%(11.67)	4.85^{*} 67%(3.35)	4.98^{*} 53%(3.54)	7.36^{**} 66%(3.13)	8.89^{***} 89%(2.95)	7.48^{***} 141%(1.90)
GMM_B: $\widehat{b}_1$	45.73 510%(37.55)	1.35 -53%(5.19)	4.01 23%(5.16)	2.36 -47%(4.73)	4.11 -13%(3.88)	8.21^{***} 164%(3.01)
GMM_B: $\widehat{b}_2$	5.22 -30%(15.18)	3.66 26%(3.76)	1.23 -62%(3.94)	2.06 -54%(3.53)	4.27^{*} -9%(3.35)	7.42^{***} 139%(2.23)
GMM_C: $\widehat{b}$	6.14 -18%(35.33)	19.69^{**} 580%(9.67)	5.01 53%(6.02)	$\times$	$\times$	9.73^{***} 213%(3.15)
			N = 25			
Beta: $\widehat{\lambda}_{\text{OLS}}$	2.03^{***} 160%(0.77)	-0.12 -126%(0.48)	0.05 -85%(0.32)	0.17 -60%(0.26)	0.29^{**} -32%(0.21)	1.02^{***} 149%(0.29)
Beta: $\widehat{\lambda}_{\text{GLS}}$	0.66^{**} -15%(0.46)	0.31 -34%(0.40)	0.40^{**} 18%(0.26)	0.60^{***} 41%(0.22)	0.46^{***} 7%(0.19)	0.87^{***} 113%(0.20)
Beta: $\widehat{\lambda}_{\text{WLS}}$	1.97^{***} 152%(0.54)	0.01 -98%(0.39)	0.01 -97%(0.24)	0.10 -77%(0.20)	0.24^{*} -45%(0.18)	0.92^{***} 125%(0.18)
GMM_A: $\widehat{b}_1$	30.91^{***} 312%(11.62)	4.78^{*} 65%(3.89)	5.74^{**} 76%(3.68)	7.73^{***} 74%(3.20)	8.29^{***} 76%(2.77)	10.15^{***} 227%(2.05)
GMM_A: $\widehat{b}_2$	32.21^{***} 330%(5.35)	11.90^{***} 311%(2.61)	15.73^{***} 382%(2.53)	18.37^{***} 314%(2.14)	15.31^{***} 226%(2.12)	9.58^{***} 209%(1.53)
GMM_B: $\widehat{b}_1$	34.88^{***} 365%(17.75)	4.56 57%(4.62)	5.67^{*} 74%(4.32)	7.35^{***} 65%(3.93)	8.05^{***} 71%(3.38)	11.39^{***} 267%(2.48)
GMM_B: $\widehat{b}_2$	-11.79 -257%(8.42)	6.16^{***} 113%(3.36)	8.28^{***} 154%(3.27)	1.00^{***} 175%(2.92)	10.53^{***} 124%(2.72)	9.77^{***} 215%(1.87)
GMM_C: $\widehat{b}$	$\times$	10.47^{**} 262%(6.36)	$\times$	$\times$	$\times$	11.56^{***} 273%(2.71)

Panel E. N Industry Portfolios

	T = 60	**T = 120**	**T = 240**	**T = 360**	**T = 480**	**T = 948**
$\lambda = Ef$	0.78	0.46	0.34	0.43	0.43	0.41
$b = \frac{\lambda}{E(R^{em2})}$	7.50	2.90	3.26	4.44	4.70	3.10
			N = 5			
Beta: $\widehat{\lambda}_{\mathrm{OLS}}$	1.61* 106%(1.06)	−0.00 −101%(0.52)	−0.12 −136%(0.28)	−0.10 −124%(0.27)	−0.09 −120%(0.21)	−0.16 −140%(0.18)
Beta: $\widehat{\lambda}_{\mathrm{GLS}}$	1.49* 91%(0.98)	−0.03 −107%(0.53)	−0.11 −131%(0.28)	−0.06 −113%(0.26)	−0.08 −119%(0.20)	−0.14 −135%(0.17)
Beta: $\widehat{\lambda}_{\mathrm{WLS}}$	1.42* 82%(0.92)	−0.04 −109%(0.53)	−0.11 −133%(0.30)	−0.03 −107%(0.27)	−0.06 −113%(0.21)	−0.13 −131%(0.17)
$\mathrm{GMM_A}$: $\widehat{b}_1$	32.59 335%(23.76)	5.95 105%(5.60)	3.99 22%(3.58)	3.19 −28%(2.75)	2.47 −47%(2.87)	1.51 −51%(2.62)
$\mathrm{GMM_A}$: $\widehat{b}_2$	29.38 292%(22.68)	6.53 126%(5.34)	4.02 23%(3.45)	3.12 −30%(2.75)	1.83 −61%(2.75)	0.50 −84%(2.45)
$\mathrm{GMM_B}$: $\widehat{b}_1$	38.92 419%(33.84)	6.95 140%(8.05)	4.59 41%(4.52)	3.55 −20%(3.20)	2.64 −44%(3.22)	1.45 −53%(2.85)
$\mathrm{GMM_B}$: $\widehat{b}_2$	31.96 326%(31.95)	7.62 163%(7.80)	4.57 40%(4.33)	3.33 −25%(3.19)	1.87 −60%(3.06)	0.38 −88%(2.67)
$\mathrm{GMM_C}$: $\widehat{b}$	×	8.40 190%(9.69)	4.67 43%(4.49)	3.45 −22%(3.03)	2.01 −57%(3.02)	0.66 −79%(2.68)
			N = 17			
Beta: $\widehat{\lambda}_{\mathrm{OLS}}$	0.66* −15%(0.50)	0.13 −73%(0.43)	0.01 −97%(0.26)	0.05 −88%(0.22)	0.00 −100%(0.19)	0.01 −97%(0.17)
Beta: $\widehat{\lambda}_{\mathrm{GLS}}$	0.33 −58%(0.47)	0.13 −71%(0.39)	−0.03 −109%(0.23)	−0.03 −108%(0.20)	−0.10 −122%(0.17)	−0.04 −109%(0.16)
Beta: $\widehat{\lambda}_{\mathrm{WLS}}$	0.65* −16%(0.49)	0.24 −49%(0.39)	0.11 −67%(0.24)	0.09 −79%(0.20)	0.07 −83%(0.17)	0.03 −91%(0.16)
$\mathrm{GMM_A}$: $\widehat{b}_1$	5.27 −30%(9.06)	5.67** 96%(3.64)	4.83** 48%(3.05)	4.93*** 11%(2.59)	3.70** −21%(2.64)	1.80 −42%(2.15)
$\mathrm{GMM_A}$: $\widehat{b}_2$	−3.91 −152%(6.47)	6.74*** 133%(2.59)	3.33** 2%(2.33)	3.17** −29%(2.21)	0.91 −81%(2.31)	0.32 −90%(1.85)
$\mathrm{GMM_B}$: $\widehat{b}_1$	4.02 −46%(9.64)	5.93** 105%(4.40)	5.28** 62%(3.64)	5.31** 20%(3.06)	3.66* −22%(2.97)	1.37 −56%(2.30)
$\mathrm{GMM_B}$: $\widehat{b}_2$	−5.93 −179%(6.78)	5.50** 90%(3.29)	2.96* −9%(2.74)	2.62 −41%(2.56)	0.02 −100%(2.51)	−0.21 −107%(1.95)
$\mathrm{GMM_C}$: $\widehat{b}$	−4.15 −155%(14.15)	6.33** 118%(4.43)	3.69* 13%(3.41)	2.77 −38%(2.85)	1.05 −78%(2.75)	0.36 −88%(2.25)
			N = 30			
Beta: $\widehat{\lambda}_{\mathrm{OLS}}$	0.88** 13%(0.52)	0.12 −73%(0.43)	−0.02 −106%(0.25)	0.01 −97%(0.21)	0.01 −99%(0.18)	−0.14 −134%(0.19)
Beta: $\widehat{\lambda}_{\mathrm{GLS}}$	0.41 −48%(0.43)	0.10 −77%(0.38)	−0.05 −116%(0.23)	−0.03 −108%(0.19)	−0.06 −114%(0.17)	−0.08 −119%(0.16)
Beta: $\widehat{\lambda}_{\mathrm{WLS}}$	0.54* −31%(0.45)	0.06 −87%(0.39)	−0.04 −111%(0.23)	−0.03 −106%(0.19)	−0.01 −102%(0.17)	−0.12 −130%(0.16)
$\mathrm{GMM_A}$: $\widehat{b}_1$	10.79*** 44%(5.93)	7.47*** 158%(4.05)	4.59** 41%(3.12)	4.34** −2%(2.57)	4.02** −14%(2.61)	1.61 −48%(2.42)
$\mathrm{GMM_A}$: $\widehat{b}_2$	14.45*** 93%(3.76)	10.04*** 247%(2.12)	5.35*** 64%(2.19)	5.81*** 31%(2.05)	4.22*** −10%(2.14)	0.96 −69%(1.85)
$\mathrm{GMM_B}$: $\widehat{b}_1$	9.95** 33%(6.94)	8.39** 190%(5.33)	4.93** 51%(3.69)	4.42** 0%(2.96)	3.83* −19%(2.93)	0.83 −73%(2.53)
$\mathrm{GMM_B}$: $\widehat{b}_2$	8.42*** 12%(4.89)	9.10*** 214%(2.96)	3.88** 19%(2.62)	4.57*** 3%(2.43)	2.71* −42%(2.43)	−0.47 −115%(1.99)
$\mathrm{GMM_C}$: $\widehat{b}$	13.58*** 81%(7.53)	13.36*** 361%(5.91)	6.08*** 86%(3.53)	7.17*** 61%(3.33)	3.27* −30%(2.84)	1.84 −41%(2.46)

3.6.8 Specification tests in CAPM model

The following five panels report the root mean square error and %p-value of the model specification tests. Each panel corresponds to a one set of test portfolios. T is 60 (January 2001 - December 2005), 120, 240, 360, 480 and 948 (January 1927 - December 2005) monthly observations. The time-series test is the Gibbons-Ross-Shanken [41] (GRS) F test. The two-pass (Fama-MacBeth) cross-sectional test are asymptotic χ^2 tests of the hypothesis that all pricing errors are zero under the null that the model is true by dividing them by their variance-covariance matrix; for these three tests, we use the well known Shanken correction, that is why %p-value for OLS and GLS are not exactly the same. Next, we turn from beta representation to a discount factor formulation for GMM approach; thus, the rest are χ^2 tests based on Hansen [46] tests for the overindentifying restrictions (or J tests). The GMM_A formulation is $E(R^e) = E(R^e f')\, b$ (returns on second moments, following Hansen and Jagannathan [49]), in this case the statistic for first and second-stage turns out to be the same (see section 1.5.2 for a discussion), then %p-value for second-stage is represented by ✓. GMM_B formulation is $E(R^e) = E\left(R^e \widetilde{f}'\right) b$ (returns on covariances, following Cochrane [25]). The GMM_C is the continuous updating estimate (following Hansen, Heaton and Yaron [47]).

Panel A. N Portfolios formed on ME

Method	**T = 60**	**T = 120**	**T = 240**	**T = 360**	**T = 480**	**T = 948**
			N = 5			
TS: GRS	0.753 3%	0.275 6%	0.076 2%	0.177 1%	0.153 1%	0.116 16%
Beta: $OLS_{Shanken}$	0.466 1%	0.189 18%	0.051 39%	0.084 62%	0.069 63%	0.045 79%
Beta: $GLS_{Shanken}$	0.774 1%	0.231 18%	0.051 39%	0.136 62%	0.117 63%	0.089 79%
Beta: $WLS_{Shanken}$	0.747 0%	0.227 55%	0.051 91%	0.119 15%	0.102 15%	0.084 27%
GMM_A: Fist-stage	0.463 1%	0.185 15%	0.050 36%	0.082 58%	0.068 61%	0.045 77%
GMM_A: Second-stage	0.671 ✓	0.188 ✓	0.077 ✓	0.143 ✓	0.113 ✓	0.058 ✓
GMM_B: Fist-stage	0.425 1%	0.172 15%	0.047 35%	0.077 59%	0.063 62%	0.041 77%
GMM_B: Second-stage	0.742 1%	0.221 15%	0.058 35%	0.156 59%	0.119 62%	0.061 77%
GMM_C	0.561 1%	0.226 15%	0.065 35%	0.150 59%	0.118 62%	0.061 77%
			N = 10			
TS: GRS	0.806 0%	0.307 5%	0.107 1%	0.192 0%	0.163 1%	0.131 42%
Beta: $OLS_{Shanken}$	0.480 0%	0.204 9%	0.078 12%	0.085 32%	0.068 59%	0.057 92%
Beta: $GLS_{Shanken}$	0.814 0%	0.259 8%	0.080 12%	0.143 32%	0.122 59%	0.103 92%
Beta: $WLS_{Shanken}$	0.607 0%	0.216 58%	0.080 90%	0.100 36%	0.084 48%	0.079 67%
GMM_A: Fist-stage	0.478 0%	0.200 8%	0.077 13%	0.083 31%	0.068 57%	0.056 89%
GMM_A: Second-stage	0.668 ✓	0.209 ✓	0.103 ✓	0.132 ✓	0.110 ✓	0.058 ✓
SDF: $GMM^B_{1,0}$	0.458 0%	0.194 8%	0.075 13%	0.081 31%	0.065 57%	0.055 88%
SDF: $GMM^B_{2,0}$	0.877 0%	0.239 8%	0.083 13%	0.173 31%	0.125 57%	0.061 88%
SDF: GMM^C	1.151 0%	0.216 8%	0.132 13%	0.138 31%	0.118 57%	0.060 89%

Panel B. N Portfolios formed on BE/ME

Methods	**T = 60**	**T = 120**	**T = 240**	**T = 360**	**T = 480**	**T = 948**
			N = 5			
TS: GRS	0.411 27%	0.322 61%	0.230 29%	0.278 2%	0.244 1%	0.121 40%
Beta: $\text{OLS}_{\text{Shanken}}$	0.340 22%	0.232 47%	0.158 34%	0.221 2%	0.191 2%	0.080 50%
Beta: $\text{GLS}_{\text{Shanken}}$	0.453 22%	0.307 46%	0.194 34%	0.255 2%	0.218 2%	0.105 50%
Beta: $\text{WLS}_{\text{Shanken}}$	0.389 0%	0.268 5%	0.169 5%	0.233 0%	0.201 0%	0.096 10%
GMM_A: Fist-stage	0.340 29%	0.227 53%	0.154 39%	0.216 2%	0.188 1%	0.078 45%
GMM_A: Second-stage	0.424 ✓	0.296 ✓	0.172 ✓	0.250 ✓	0.197 ✓	0.079 ✓
GMM_B: Fist-stage	0.311 29%	0.212 54%	0.144 40%	0.202 2%	0.174 1%	0.073 45%
GMM_B: Second-stage	0.436 29%	0.319 54%	0.189 40%	0.279 2%	0.209 1%	0.082 45%
GMM_C	0.357 27%	0.269 51%	0.165 39%	0.240 2%	0.206 1%	0.081 45%
			N = 10			
TS: GRS	0.424 71%	0.325 86%	0.232 69%	0.285 16%	0.248 12%	0.122 12%
Beta: $\text{OLS}_{\text{Shanken}}$	0.334 61%	0.222 79%	0.152 78%	0.221 15%	0.189 18%	0.083 15%
Beta: $\text{GLS}_{\text{Shanken}}$	0.465 60%	0.303 79%	0.192 78%	0.262 14%	0.221 18%	0.104 14%
Beta: $\text{WLS}_{\text{Shanken}}$	0.349 2%	0.233 27%	0.156 21%	0.227 0%	0.195 0%	0.090 20%
GMM_A: Fist-stage	0.333 72%	0.217 79%	0.148 79%	0.216 16%	0.186 16%	0.082 13%
GMM_A: Second-stage	0.423 ✓	0.287 ✓	0.158 ✓	0.245 ✓	0.199 ✓	0.098 ✓
GMM_B: Fist-stage	0.318 72%	0.211 80%	0.144 79%	0.211 17%	0.180 16%	0.079 13%
GMM_B: Second-stage	0.441 72%	0.325 80%	0.178 79%	0.276 17%	0.212 16%	0.111 13%
GMM_C	0.377 70%	0.288 79%	0.159 79%	0.240 16%	0.208 16%	0.097 13%

Panel C. N Portfolios formed by the intersections of ME and BE/ME

Method	T = 60	T = 120	T = 240	T = 360	T = 480	T = 948
			N = 6			
TS: GRS	0.713 2%	0.463 0%	0.323 0%	0.375 0%	0.326 0%	0.200 0%
Beta: $OLS_{Shanken}$	0.573 1%	0.428 0%	0.314 0%	0.333 0%	0.279 0%	0.155 0%
Beta: $GLS_{Shanken}$	0.763 1%	0.461 0%	0.326 0%	0.380 0%	0.321 0%	0.194 0%
Beta: $WLS_{Shanken}$	0.701 0%	0.437 2%	0.314 1%	0.341 0%	0.287 0%	0.171 1%
GMM_A: Fist-stage	0.571 2%	0.420 0%	0.307 0%	0.324 0%	0.275 0%	0.153 0%
GMM_A: Second-stage	0.701 ✓	0.420 ✓	0.328 ✓	0.389 ✓	0.300 ✓	0.178 ✓
GMM_B: Fist-stage	0.530 2%	0.396 0%	0.290 0%	0.308 0%	0.258 0%	0.144 0%
GMM_B: Second-stage	0.746 2%	0.494 0%	0.438 0%	0.508 0%	0.354 0%	0.212 0%
GMM_C	0.708 1%	0.723 0%	0.425 0%	0.341 0%	0.283 0%	0.178 0%
			N = 25			
TS: GRS	0.854 11%	0.545 0%	0.370 0%	0.413 0%	0.361 0%	0.257 0%
Beta: $OLS_{Shanken}$	0.599 0%	0.486 0%	0.354 0%	0.352 0%	0.303 0%	0.231 0%
Beta: $GLS_{Shanken}$	0.869 0%	0.503 0%	0.358 0%	0.401 0%	0.347 0%	0.246 0%
Beta: $WLS_{Shanken}$	0.626 0%	0.487 1%	0.359 0%	0.352 0%	0.303 0%	0.231 0%
GMM_A: Fist-stage	0.597 0%	0.476 0%	0.345 0%	0.343 0%	0.299 0%	0.227 0%
GMM_A: Second-stage	0.640 ✓	0.880 ✓	0.712 ✓	0.377 ✓	0.317 ✓	0.228 ✓
GMM_B: Fist-stage	0.588 0%	0.476 0%	0.347 0%	0.345 0%	0.297 0%	0.226 0%
GMM_B: Second-stage	0.774 0%	0.501 0%	0.384 0%	0.466 0%	0.337 0%	0.269 0%
GMM_C	3.452 0%	4.891 0%	7.741 0%	0.744 0%	0.314 0%	0.233 0%
			N = 100			
TS: GRS		0.605 35%	0.402 0%	0.440 0%	0.385 0%	0.306 0%
Beta: $OLS_{Shanken}$		0.543 0%	0.384 0%	0.372 0%	0.322 0%	0.245 0%
Beta: $GLS_{Shanken}$		0.571 0%	0.392 0%	0.425 0%	0.368 0%	0.276 0%
Beta: $WLS_{Shanken}$		0.548 0%	0.389 0%	0.372 0%	0.322 0%	0.249 0%
GMM_A: Fist-stage		0.531 0%	0.374 0%	0.362 0%	0.318 0%	0.241 0%
GMM_A: Second-stage		5.258 ✓	1.873 ✓	0.989 ✓	0.568 ✓	0.265 ✓
GMM_B: Fist-stage		0.540 0%	0.382 0%	0.370 0%	0.320 0%	0.244 0%
GMM_B: Second-stage		0.594 0%	0.392 0%	0.425 0%	0.354 0%	0.251 0%
GMM_C		0.562 0%	1.004 0%	11.143 0%	2.695 0%	0.250 0%

Panel D. N Portfolios formed on ME MOM

Methods	T = 60	T = 120	T = 240	T = 360	T = 480	T = 948
			N = 6			
TS: GRS	0.733 3%	0.500 2%	0.421 0%	0.434 0%	0.411 0%	0.410 0%
Beta: $OLS_{Shanken}$	0.583 0%	0.490 0%	0.420 0%	0.420 0%	0.402 0%	0.410 0%
Beta: $GLS_{Shanken}$	0.737 0%	0.497 0%	0.420 0%	0.430 0%	0.406 0%	0.414 0%
Beta: $WLS_{Shanken}$	0.638 0%	0.499 5%	0.432 0%	0.420 0%	0.402 0%	0.413 0%
GMM_A: Fist-stage	0.581 1%	0.482 0%	0.412 0%	0.410 0%	0.398 0%	0.404 0%
GMM_A: Second-stage	0.622 ✓	0.490 ✓	0.492 ✓	0.431 ✓	0.405 ✓	0.411 ✓
GMM_B: Fist-stage	0.539 1%	0.454 0%	0.389 0%	0.389 0%	0.372 0%	0.380 0%
GMM_B: Second-stage	0.681 1%	0.500 0%	0.422 0%	0.457 0%	0.416 0%	0.451 0%
GMM_C	1.551 2%	0.496 0%	0.433 0%	0.628 0%	0.422 0%	0.469 0%
			N = 25			
TS: GRS	0.936 9%	0.603 1%	0.472 0%	0.491 0%	0.462 0%	0.485 0%
Beta: $OLS_{Shanken}$	0.741 0%	0.589 0%	0.470 0%	0.468 0%	0.447 0%	0.481 0%
Beta: $GLS_{Shanken}$	0.899 0%	0.614 0%	0.470 0%	0.484 0%	0.450 0%	0.481 0%
Beta: $WLS_{Shanken}$	0.744 0%	0.608 0%	0.484 0%	0.473 0%	0.451 0%	0.484 0%
GMM_A: Fist-stage	0.738 0%	0.578 0%	0.460 0%	0.457 0%	0.442 0%	0.474 0%
GMM_A: Second-stage	1.244 ✓	0.828 ✓	0.863 ✓	0.720 ✓	0.531 ✓	0.488 ✓
GMM_B: Fist-stage	0.727 0%	0.577 0%	0.461 0%	0.459 0%	0.439 0%	0.472 0%
GMM_B: Second-stage	0.753 0%	0.597 0%	0.472 0%	0.496 0%	0.453 0%	0.489 0%
GMM_C	1.997 0%	3.943 0%	5.573 0%	4.733 0%	1.225 0%	0.484 0%

Panel E. N Industry Portfolios

Method	**T = 60**	**T = 120**	**T = 240**	**T = 360**	**T = 480**	**T = 948**
			N = 5			
TS: GRS	0.436 35%	0.295 60%	0.227 23%	0.164 15%	0.123 11%	0.122 1%
Beta: $\text{OLS}_{\text{Shanken}}$	0.422 17%	0.292 56%	0.212 37%	0.139 61%	0.093 75%	0.103 23%
Beta: $\text{GLS}_{\text{Shanken}}$	0.434 17%	0.292 56%	0.214 37%	0.141 61%	0.095 75%	0.104 23%
Beta: $\text{WLS}_{\text{Shanken}}$	0.448 6%	0.298 24%	0.214 17%	0.141 32%	0.093 62%	0.104 20%
GMM_A: Fist-stage	0.422 17%	0.287 58%	0.208 41%	0.136 65%	0.092 76%	0.101 21%
GMM_A: Second-stage	0.438 ✓	0.289 ✓	0.210 ✓	0.139 ✓	0.093 ✓	0.110 ✓
GMM_B: Fist-stage	0.385 17%	0.266 58%	0.194 41%	0.127 65%	0.085 76%	0.094 22%
GMM_B: Second-stage	0.432 17%	0.292 58%	0.220 41%	0.144 65%	0.095 76%	0.115 22%
GMM_C	0.534 19%	0.297 58%	0.213 41%	0.140 64%	0.094 76%	0.109 21%
			N = 17			
TS: GRS	0.639 51%	0.360 63%	0.255 52%	0.232 10%	0.176 4%	0.124 1%
Beta: $\text{OLS}_{\text{Shanken}}$	0.581 12%	0.353 41%	0.252 68%	0.230 45%	0.169 40%	0.117 45%
Beta: $\text{GLS}_{\text{Shanken}}$	0.627 12%	0.356 41%	0.253 68%	0.234 46%	0.171 40%	0.123 45%
Beta: $\text{WLS}_{\text{Shanken}}$	0.589 8%	0.362 44%	0.254 21%	0.230 4%	0.170 16%	0.118 24%
GMM_A: Fist-stage	0.580 10%	0.348 32%	0.247 63%	0.225 42%	0.167 35%	0.116 41%
GMM_A: Second-stage	0.584 ✓	0.359 ✓	0.265 ✓	0.236 ✓	0.178 ✓	0.118 ✓
GMM_B: Fist-stage	0.565 10%	0.343 32%	0.245 63%	0.223 42%	0.164 35%	0.114 42%
GMM_B: Second-stage	0.621 10%	0.361 32%	0.253 63%	0.230 42%	0.171 35%	0.126 42%
GMM_C	0.844 12%	0.402 33%	0.259 63%	0.236 42%	0.172 35%	0.119 41%
			N = 30			
TS: GRS	0.808 73%	0.593 71%	0.312 58%	0.248 19%	0.222 19%	0.189 1%
Beta: $\text{OLS}_{\text{Shanken}}$	0.754 2%	0.575 36%	0.301 63%	0.234 53%	0.202 63%	0.177 19%
Beta: $\text{GLS}_{\text{Shanken}}$	0.825 2%	0.581 36%	0.301 63%	0.236 54%	0.202 63%	0.179 19%
Beta: $\text{WLS}_{\text{Shanken}}$	0.755 15%	0.577 66%	0.301 54%	0.234 23%	0.203 37%	0.177 4%
GMM_A: Fist-stage	0.752 2%	0.565 30%	0.294 57%	0.229 44%	0.200 60%	0.174 12%
GMM_A: Second-stage	0.757 ✓	0.637 ✓	0.352 ✓	0.276 ✓	0.209 ✓	0.189 ✓
GMM_B: Fist-stage	0.742 2%	0.566 27%	0.296 55%	0.230 42%	0.199 59%	0.174 12%
GMM_B: Second-stage	0.793 2%	0.577 27%	0.302 55%	0.237 42%	0.202 59%	0.212 12%
GMM_C	0.762 2%	0.586 28%	0.326 56%	0.241 42%	0.206 59%	0.205 12%

3.6.9 Specification tests in Fama-French model

The following five panels report the root mean square error and $\%p$-value of the model specification tests. Each panel corresponds to a one set of test portfolios. T is 60 (January 2001 - December 2005), 120, 240, 360, 480 and 948 (January 1927 - December 2005) monthly observations. The time-series test is the Gibbons-Ross-Shanken [41] (GRS) F test. The two-pass (Fama-MacBeth) cross-sectional test are asymptotic χ^2 tests of the hypothesis that all pricing errors are zero under the null that the model is true by dividing them by their variance-covariance matrix; for these three tests, we use the well known Shanken correction, that is why $\%p$-value for OLS and GLS are not exactly the same. Next, we turn from beta representation to a discount factor formulation for GMM approach; thus, the rest are χ^2 tests based on Hansen [46] tests for the overindentifying restrictions (or J tests). The GMM_A formulation is $E\left(R^e\right) = E\left(R^e f'\right) b$ (returns on second moments, following Hansen and Jagannathan [49]), in this case the statistic for first and second-stage turns out to be the same (see section 1.5.2 for a discussion), then $\%p$-value for second-stage is represented by $\checkmark$. GMM_B formulation is $E\left(R^e\right) = E\left(R^e \widetilde{f}'\right) b$ (returns on covariances, following Cochrane [25]). The GMM_C is the continuous updating estimate (following Hansen, Heaton and Yaron [47]).

Panel A. N Portfolios formed on ME

Method	**T = 60**	**T = 120**	**T = 240**	**T = 360**	**T = 480**	**T = 948**
				N = 5		
TS: GRS	0.139 52%	0.178 5%	0.065 1%	0.050 0%	0.046 0%	0.086 0%
Beta: $OLS_{Shanken}$	0.121 42%	0.106 7%	0.050 25%	0.015 82%	0.010 80%	0.012 69%
Beta: $GLS_{Shanken}$	0.125 37%	0.135 5%	0.064 29%	0.019 83%	0.010 80%	0.015 69%
Beta: $WLS_{Shanken}$	0.134 47%	0.110 11%	0.054 28%	0.016 83%	0.010 90%	0.014 65%
GMM_A: Fist-stage	0.092 49%	0.110 7%	0.048 32%	0.014 84%	0.009 82%	0.012 65%
GMM_A: Second-stage	0.323 ✓	0.286 ✓	0.065 ✓	0.020 ✓	0.012 ✓	0.014 ✓
GMM_B: Fist-stage	0.096 42%	0.084 4%	0.040 26%	0.012 83%	0.008 82%	0.009 65%
GMM_B: Second-stage	0.369 42%	0.228 4%	0.064 26%	0.021 83%	0.013 82%	0.014 65%
GMM_C	0.399 47%	0.605 8%	0.142 45%	0.021 85%	0.016 83%	0.015 65%
				N = 10		
TS: GRS	0.277 4%	0.193 8%	0.079 1%	0.054 0%	0.046 3%	0.086 0%
Beta: $OLS_{Shanken}$	0.269 0%	0.151 8%	0.073 10%	0.036 29%	0.025 60%	0.023 91%
Beta: $GLS_{Shanken}$	0.275 0%	0.199 9%	0.075 11%	0.037 28%	0.026 59%	0.025 91%
Beta: $WLS_{Shanken}$	0.276 16%	0.169 6%	0.076 44%	0.037 85%	0.025 94%	0.025 94%
GMM_A: Fist-stage	0.203 9%	0.152 11%	0.069 12%	0.034 28%	0.024 57%	0.022 88%
GMM_A: Second-stage	1.471 ✓	0.380 ✓	0.090 ✓	0.036 ✓	0.026 ✓	0.030 ✓
GMM_B: Fist-stage	0.236 0%	0.133 4%	0.064 10%	0.032 28%	0.022 58%	0.020 88%
GMM_B: Second-stage	1.012 0%	0.259 5%	0.094 10%	0.053 28%	0.027 58%	0.028 88%
GMM_C	3.730 23%	0.273 20%	0.093 12%	0.041 28%	0.026 57%	0.031 88%

Panel B. N Portfolios formed on BE/ME

Method	**T = 60**	**T = 120**	**T = 240**	**T = 360**	**T = 480**	**T = 948**
			N = 5			
TS: GRS	0.162 73%	0.151 8%	0.095 6%	0.072 6%	0.072 3%	0.072 1%
Beta: $OLS_{Shanken}$	0.030 88%	0.044 71%	0.025 79%	0.005 98%	0.025 61%	0.014 69%
Beta: $GLS_{Shanken}$	0.032 87%	0.046 70%	0.027 80%	0.006 98%	0.027 61%	0.014 69%
Beta: $WLS_{Shanken}$	0.030 88%	0.044 75%	0.025 78%	0.005 98%	0.025 60%	0.014 78%
GMM_A: Fist-stage	0.025 88%	0.042 73%	0.024 80%	0.005 98%	0.024 63%	0.014 69%
GMM_A: Second-stage	0.031 ✓	0.044 ✓	0.051 ✓	0.012 ✓	0.043 ✓	0.018 ✓
GMM_B: Fist-stage	0.023 88%	0.035 73%	0.020 80%	0.004 98%	0.020 63%	0.011 69%
GMM_B: Second-stage	0.038 88%	0.047 73%	0.050 80%	0.013 98%	0.043 63%	0.019 69%
GMM_C	0.037 88%	0.046 72%	0.065 79%	0.013 98%	0.046 63%	0.018 69%
			N = 10			
TS: GRS	0.201 65%	0.174 18%	0.108 8%	0.080 11%	0.077 9%	0.110 0%
Beta: $OLS_{Shanken}$	0.076 93%	0.067 93%	0.039 98%	0.028 96%	0.030 98%	0.050 16%
Beta: $GLS_{Shanken}$	0.085 93%	0.067 93%	0.043 98%	0.029 96%	0.031 98%	0.052 16%
Beta: $WLS_{Shanken}$	0.077 98%	0.067 96%	0.041 99%	0.028 98%	0.031 97%	0.051 29%
GMM_A: Fist-stage	0.066 94%	0.064 92%	0.037 98%	0.026 96%	0.029 98%	0.050 12%
GMM_A: Second-stage	0.282 ✓	0.090 ✓	0.081 ✓	0.039 ✓	0.043 ✓	0.077 ✓
GMM_B: Fist-stage	0.067 94%	0.059 92%	0.034 98%	0.025 96%	0.027 98%	0.044 13%
GMM_B: Second-stage	0.375 94%	0.073 92%	0.073 98%	0.034 96%	0.042 98%	0.089 13%
GMM_C	0.368 95%	0.095 92%	0.089 98%	0.040 96%	0.045 98%	0.070 12%

Panel C. N Portfolios formed by the intersections of ME and BE/ME

Method	T = 60	T = 120	T = 240	T = 360	T = 480	T = 948
			N = 6			
TS: GRS	0.225 48%	0.240 0%	0.173 0%	0.144 0%	0.117 0%	0.095 0%
Beta: $OLS_{Shanken}$	0.143 23%	0.212 0%	0.165 0%	0.139 0%	0.115 0%	0.087 0%
Beta: $GLS_{Shanken}$	0.173 22%	0.247 0%	0.172 0%	0.143 0%	0.119 0%	0.097 0%
Beta: $WLS_{Shanken}$	0.173 7%	0.239 0%	0.174 0%	0.144 0%	0.120 0%	0.092 0%
GMM_A: Fist-stage	0.121 13%	0.195 0%	0.155 0%	0.128 0%	0.108 0%	0.085 0%
GMM_A: Second-stage	0.252 ✓	0.248 ✓	0.231 ✓	0.134 ✓	0.114 ✓	0.101 ✓
GMM_B: Fist-stage	0.117 11%	0.173 0%	0.135 0%	0.114 0%	0.094 0%	0.071 0%
GMM_B: Second-stage	0.151 11%	0.287 0%	0.400 0%	0.253 0%	0.124 0%	0.122 0%
GMM_C	0.180 11%	3.985 0%	0.288 0%	0.144 0%	0.121 0%	0.093 0%
			N = 25			
TS: GRS	0.291 36%	0.297 0%	0.213 0%	0.172 0%	0.144 0%	0.220 0%
Beta: $OLS_{Shanken}$	0.263 0%	0.275 0%	0.209 0%	0.168 0%	0.141 0%	0.193 0%
Beta: $GLS_{Shanken}$	0.290 0%	0.354 0%	0.231 0%	0.180 0%	0.152 0%	0.225 0%
Beta: $WLS_{Shanken}$	0.273 10%	0.292 0%	0.218 0%	0.173 0%	0.145 0%	0.214 0%
GMM_A: Fist-stage	0.220 3%	0.252 0%	0.196 0%	0.154 0%	0.133 0%	0.189 0%
GMM_A: Second-stage	1.362 ✓	1.149 ✓	0.791 ✓	0.408 ✓	0.295 ✓	0.210 ✓
GMM_B: Fist-stage	0.249 0%	0.260 0%	0.198 0%	0.159 0%	0.133 0%	0.182 0%
GMM_B: Second-stage	0.332 0%	0.298 0%	0.263 0%	0.220 0%	0.151 0%	0.206 0%
GMM_C	5.863 3%	4.203 0%	2.647 0%	0.832 0%	0.278 0%	1.822 0%
			N = 100			
TS: GRS		0.380 55%	0.269 0%	0.210 0%	0.189 0%	0.200 0%
Beta: $OLS_{Shanken}$		0.354 0%	0.263 0%	0.208 0%	0.187 0%	0.198 0%
Beta: $GLS_{Shanken}$		0.404 0%	0.275 0%	0.217 0%	0.198 0%	0.205 0%
Beta: $WLS_{Shanken}$		0.365 0%	0.268 0%	0.210 0%	0.190 0%	0.202 1%
GMM_A: Fist-stage		0.324 0%	0.247 0%	0.190 0%	0.177 0%	0.193 0%
GMM_A: Second-stage		3.177 ✓	1.743 ✓	0.989 ✓	0.516 ✓	0.210 ✓
GMM_B: Fist-stage		0.349 0%	0.260 0%	0.205 0%	0.185 0%	0.195 0%
GMM_B: Second-stage		0.377 0%	0.287 0%	0.242 0%	0.189 0%	0.205 0%
GMM_C		1.990 0%	6.682 0%	1.824 0%	1.822 0%	0.446 0%

Panel D. N Portfolios formed on ME and MOM

Method	**T = 60**	**T = 120**	**T = 240**	**T = 360**	**T = 480**	**T = 948**
			N = 6			
TS: GRS	0.184 20%	0.501 3%	0.459 0%	0.434 0%	0.444 0%	0.463 0%
Beta: $OLS_{Shanken}$	0.144 11%	0.411 1%	0.408 0%	0.388 0%	0.349 0%	0.218 0%
Beta: $GLS_{Shanken}$	0.161 11%	0.482 1%	0.454 0%	0.434 0%	0.434 0%	0.352 0%
Beta: $WLS_{Shanken}$	0.166 70%	0.479 1%	0.463 0%	0.432 0%	0.400 0%	0.259 0%
GMM_A: Fist-stage	0.123 6%	0.399 1%	0.409 0%	0.386 0%	0.370 0%	0.224 0%
GMM_A: Second-stage	0.221 ✓	0.535 ✓	0.659 ✓	0.582 ✓	0.591 ✓	1.031 ✓
GMM_B: Fist-stage	0.118 6%	0.336 0%	0.333 0%	0.317 0%	0.285 0%	0.178 0%
GMM_B: Second-stage	0.421 5%	0.452 0%	0.461 0%	0.410 0%	0.451 0%	0.892 0%
GMM_C	0.571 7%	1.320 6%	×	264.398 0%	×	×
			N = 25			
TS: GRS	0.352 28%	0.568 3%	0.489 0%	0.461 0%	0.472 0%	0.512 0%
Beta: $OLS_{Shanken}$	0.312 0%	0.481 0%	0.463 0%	0.442 0%	0.430 0%	0.352 0%
Beta: $GLS_{Shanken}$	0.327 0%	0.527 0%	0.507 0%	0.481 0%	0.478 0%	0.489 0%
Beta: $WLS_{Shanken}$	0.313 14%	0.532 0%	0.502 0%	0.472 0%	0.460 0%	0.372 0%
GMM_A: Fist-stage	0.258 3%	0.445 0%	0.443 0%	0.414 0%	0.427 0%	0.356 0%
GMM_A: Second-stage	1.108 ✓	0.848 ✓	1.014 ✓	0.940 ✓	0.617 ✓	0.695 ✓
GMM_B: Fist-stage	0.295 0%	0.454 0%	0.437 0%	0.418 0%	0.406 0%	0.332 0%
GMM_B: Second-stage	0.416 0%	0.501 0%	0.466 0%	0.448 0%	0.444 0%	0.541 0%
GMM_C	3.244 11%	2.489 1%	3.349 2%	4.324 0%	×	×

Panel E. N Industry Portfolios

Method	**T = 60**	**T = 120**	**T = 240**	**T = 360**	**T = 480**	**T = 948**
				N = 5		
TS: GRS	0.082 99%	0.212 63%	0.222 9%	0.197 6%	0.214 0%	0.178 0%
Beta: $\text{OLS}_{\text{Shanken}}$	0.075 74%	0.110 55%	0.067 48%	0.018 94%	0.026 86%	0.052 40%
Beta: $\text{GLS}_{\text{Shanken}}$	0.076 74%	0.127 55%	0.070 47%	0.022 94%	0.033 86%	0.061 38%
Beta: $\text{WLS}_{\text{Shanken}}$	0.077 76%	0.131 52%	0.074 41%	0.019 92%	0.030 81%	0.060 28%
GMM_{A}: Fist-stage	0.063 74%	0.107 57%	0.066 45%	0.018 93%	0.025 86%	0.052 45%
GMM_{A}: Second-stage	0.088 ✓	0.128 ✓	0.074 ✓	0.023 ✓	0.035 ✓	0.064 ✓
GMM_{B}: Fist-stage	0.059 74%	0.087 55%	0.053 45%	0.014 93%	0.020 86%	0.041 45%
GMM_{B}: Second-stage	0.096 74%	0.134 55%	0.073 45%	0.024 93%	0.035 86%	0.066 45%
GMM_{C}	0.112 74%	0.128 55%	0.072 44%	0.023 93%	0.034 86%	0.062 41%
				N = 17		
TS: GRS	0.340 55%	0.427 13%	0.302 3%	0.301 0%	0.251 0%	0.159 0%
Beta: $\text{OLS}_{\text{Shanken}}$	0.315 21%	0.286 52%	0.164 94%	0.145 65%	0.134 45%	0.088 71%
Beta: $\text{GLS}_{\text{Shanken}}$	0.389 15%	0.304 55%	0.170 94%	0.159 63%	0.138 44%	0.089 71%
Beta: $\text{WLS}_{\text{Shanken}}$	0.317 91%	0.301 81%	0.172 97%	0.146 74%	0.135 57%	0.089 71%
GMM_{A}: Fist-stage	0.270 10%	0.277 41%	0.159 92%	0.143 58%	0.133 40%	0.087 73%
GMM_{A}: Second-stage	0.597 ✓	0.326 ✓	0.170 ✓	0.163 ✓	0.140 ✓	0.093 ✓
GMM_{B}: Fist-stage	0.290 4%	0.263 44%	0.151 93%	0.134 58%	0.124 41%	0.081 73%
GMM_{B}: Second-stage	0.355 4%	0.322 44%	0.174 93%	0.156 58%	0.139 41%	0.100 73%
GMM_{C}	1.765 22%	0.360 46%	0.182 94%	0.165 60%	0.143 42%	0.094 72%
				N = 30		
TS: GRS	0.453 80%	0.550 19%	0.321 8%	0.300 0%	0.262 0%	0.221 0%
Beta: $\text{OLS}_{\text{Shanken}}$	0.447 14%	0.516 40%	0.249 71%	0.165 68%	0.175 68%	0.157 37%
Beta: $\text{GLS}_{\text{Shanken}}$	0.513 9%	0.546 40%	0.254 70%	0.179 64%	0.181 66%	0.163 37%
Beta: $\text{WLS}_{\text{Shanken}}$	0.470 97%	0.540 79%	0.255 93%	0.166 88%	0.177 70%	0.163 32%
GMM_{A}: Fist-stage	0.372 23%	0.486 20%	0.241 61%	0.162 57%	0.174 59%	0.156 35%
GMM_{A}: Second-stage	0.777 ✓	0.532 ✓	0.257 ✓	0.203 ✓	0.177 ✓	0.185 ✓
GMM_{B}: Fist-stage	0.426 2%	0.492 22%	0.238 63%	0.157 55%	0.167 60%	0.150 36%
GMM_{B}: Second-stage	0.570 2%	0.562 23%	0.269 63%	0.179 55%	0.184 61%	0.210 36%
GMM_{C}	0.830 11%	1.212 37%	0.260 60%	0.174 53%	0.184 62%	0.222 39%

3.6.10 Specification tests in RUH model

The following five panels report the root mean square error and %p-value of the model specification tests. Each panel corresponds to a one set of test portfolios. T is 60 (January 2001 - December 2005), 120, 240, 360, 480 and 948 (January 1927 - December 2005) monthly observations. The time-series test is the Gibbons-Ross-Shanken [41] (GRS) F test. The two-pass (Fama-MacBeth) cross-sectional test are asymptotic χ^2 tests of the hypothesis that all pricing errors are zero under the null that the model is true by dividing them by their variance-covariance matrix; for these three tests, we use the well known Shanken correction, that is why %p-value for OLS and GLS are not exactly the same. Next, we turn from beta representation to a discount factor formulation for GMM approach; thus, the rest are χ^2 tests based on Hansen [46] tests for the overindentifying restrictions (or J tests). The GMM_{A} formulation is $E\left(R^e\right) = E\left(R^e f'\right) b$ (returns on second moments, following Hansen and Jagannathan [49]), in this case the statistic for first and second-stage turns out to be the same (see section 1.5.2 for a discussion), then %p-value for second-stage is represented by $\checkmark$. GMM_{B} formulation is $E\left(R^e\right) = E\left(R^e \widetilde{f}'\right) b$ (returns on covariances, following Cochrane [25]). The GMM_{C} is the continuous updating estimate (following Hansen, Heaton and Yaron [47]).

Panel A. N Portfolios formed on ME

Method	**T = 60**	**T = 120**	**T = 240**	**T = 360**	**T = 480**	**T = 948**
				N = 5		
TS: GRS	0.605 16%	0.368 16%	0.115 9%	0.178 3%	0.155 1%	0.084 8%
Beta: $OLS_{Shanken}$	0.104 48%	0.073 75%	0.021 73%	0.026 73%	0.059 62%	0.002 99%
Beta: $GLS_{Shanken}$	0.354 27%	0.150 79%	0.024 72%	0.068 68%	0.118 54%	0.005 99%
Beta: $WLS_{Shanken}$	0.113 83%	0.090 86%	0.023 98%	0.032 91%	0.103 48%	0.002 100%
GMM_A: Fist-stage	0.072 62%	0.061 77%	0.017 73%	0.020 77%	0.044 83%	0.002 99%
GMM_A: Second-stage	0.128 ✓	0.092 ✓	0.025 ✓	0.098 ✓	0.169 ✓	0.005 ✓
GMM_B: Fist-stage	0.082 67%	0.058 73%	0.017 73%	0.020 78%	0.046 71%	0.001 99%
GMM_B: Second-stage	0.160 67%	0.127 73%	0.030 73%	0.119 78%	0.194 71%	0.005 99%
GMM_C	0.143 52%	0.119 79%	0.033 73%	0.088 73%	557.638 65%	0.005 99%
				N = 10		
TS: GRS	0.660 1%	0.394 31%	0.135 8%	0.188 4%	0.163 3%	0.108 11%
Beta: $OLS_{Shanken}$	0.219 7%	0.105 90%	0.057 44%	0.034 78%	0.062 56%	0.023 90%
Beta: $GLS_{Shanken}$	0.621 0%	0.122 91%	0.058 40%	0.064 73%	0.123 48%	0.040 90%
Beta: $WLS_{Shanken}$	0.230 94%	0.123 100%	0.058 100%	0.038 100%	0.084 47%	0.026 100%
GMM_A: Fist-stage	0.158 7%	0.089 94%	0.049 55%	0.029 86%	0.055 79%	0.022 88%
GMM_A: Second-stage	1.656 ✓	0.093 ✓	0.058 ✓	0.107 ✓	0.195 ✓	0.026 ✓
GMM_B: Fist-stage	0.192 2%	0.092 93%	0.050 55%	0.030 87%	0.054 68%	0.020 88%
GMM_B: Second-stage	1.534 9%	0.107 93%	0.098 55%	0.136 87%	0.184 68%	0.025 88%
GMM_C	17.160 8%	0.135 94%	0.079 51%	0.088 83%	0.191 52%	0.027 88%

Panel B. N Portfolios formed on BE/ME

Method	**T = 60**	**T = 120**	**T = 240**	**T = 360**	**T = 480**	**T = 948**
			N = 5			
TS: GRS	0.129 69%	0.052 98%	0.056 77%	0.041 84%	0.052 53%	0.040 31%
Beta: $\text{OLS}_{\text{Shanken}}$	0.076 62%	0.017 99%	0.014 94%	0.026 76%	0.028 61%	0.012 88%
Beta: $\text{GLS}_{\text{Shanken}}$	0.093 53%	0.020 98%	0.015 94%	0.028 77%	0.031 58%	0.012 88%
Beta: $\text{WLS}_{\text{Shanken}}$	0.077 69%	0.018 99%	0.014 96%	0.026 70%	0.028 60%	0.012 87%
GMM_A: Fist-stage	0.070 73%	0.011 99%	0.017 94%	0.024 77%	0.029 68%	0.011 88%
GMM_A: Second-stage	0.168 ✓	0.025 ✓	0.023 ✓	0.026 ✓	0.038 ✓	0.016 ✓
GMM_B: Fist-stage	0.060 71%	0.014 99%	0.011 94%	0.021 76%	0.022 68%	0.009 88%
GMM_B: Second-stage	0.197 71%	0.040 99%	0.020 94%	0.028 76%	0.037 68%	0.016 88%
GMM_C	0.265 76%	0.044 99%	0.018 94%	0.028 76%	0.037 61%	0.016 88%
			N = 10			
TS: GRS	0.156 89%	0.081 99%	0.078 80%	0.056 80%	0.062 70%	0.065 12%
Beta: $\text{OLS}_{\text{Shanken}}$	0.122 86%	0.066 94%	0.028 100%	0.028 98%	0.030 99%	0.046 34%
Beta: $\text{GLS}_{\text{Shanken}}$	0.140 84%	0.067 94%	0.028 99%	0.030 98%	0.032 99%	0.047 34%
Beta: $\text{WLS}_{\text{Shanken}}$	0.127 83%	0.066 97%	0.028 100%	0.030 99%	0.031 98%	0.047 47%
GMM_A: Fist-stage	0.112 83%	0.061 93%	0.028 100%	0.027 98%	0.031 100%	0.041 29%
GMM_A: Second-stage	0.185 ✓	0.107 ✓	0.037 ✓	0.028 ✓	0.039 ✓	0.132 ✓
GMM_B: Fist-stage	0.107 81%	0.058 93%	0.024 100%	0.025 98%	0.027 100%	0.040 27%
GMM_B: Second-stage	0.250 81%	0.080 93%	0.032 100%	0.030 98%	0.037 100%	0.115 27%
GMM_C	0.227 83%	0.108 93%	0.036 100%	0.030 98%	0.038 99%	1.286 45%

Panel C. N Portfolios formed by the intersections of ME and BE/ME

Method	**T = 60**	**T = 120**	**T = 240**	**T = 360**	**T = 480**	**T = 948**
				N = 6		
TS: GRS	0.493 7%	0.321 1%	0.165 0%	0.210 0%	0.191 0%	0.121 0%
Beta: $OLS_{Shanken}$	0.438 8%	0.060 82%	0.054 86%	0.078 44%	0.115 49%	0.096 1%
Beta: $GLS_{Shanken}$	0.582 14%	0.074 81%	0.084 87%	0.154 64%	0.149 41%	0.133 5%
Beta: $WLS_{Shanken}$	0.539 0%	0.081 98%	0.084 90%	0.107 69%	0.144 14%	0.114 3%
GMM_A: Fist-stage	0.400 12%	0.044 89%	0.040 92%	0.056 73%	0.069 84%	0.087 4%
GMM_A: Second-stage	0.495 ✓	0.105 ✓	0.100 ✓	0.211 ✓	0.086 ✓	0.207 ✓
GMM_B: Fist-stage	0.358 16%	0.049 89%	0.044 91%	0.064 67%	0.094 77%	0.078 1%
GMM_B: Second-stage	0.611 15%	0.137 89%	0.119 91%	0.272 68%	0.142 77%	0.159 1%
GMM_C	1.336 60%	0.146 89%	0.153 93%	0.358 87%	0.150 78%	1.306 50%
				N = 25		
TS: GRS	0.620 21%	0.434 1%	0.223 0%	0.252 0%	0.227 0%	0.186 0%
Beta: $OLS_{Shanken}$	0.484 0%	0.230 27%	0.133 10%	0.134 14%	0.153 16%	0.176 0%
Beta: $GLS_{Shanken}$	0.686 0%	0.295 3%	0.134 5%	0.149 2%	0.201 3%	0.182 0%
Beta: $WLS_{Shanken}$	0.520 0%	0.239 98%	0.139 100%	0.141 92%	0.165 6%	0.180 3%
GMM_A: Fist-stage	0.440 0%	0.174 39%	0.101 17%	0.098 10%	0.112 53%	0.157 0%
GMM_A: Second-stage	0.794 ✓	0.527 ✓	0.295 ✓	0.197 ✓	0.139 ✓	0.176 ✓
GMM_B: Fist-stage	0.457 0%	0.217 4%	0.125 2%	0.126 6%	0.145 17%	0.167 0%
GMM_B: Second-stage	0.710 0%	0.735 4%	0.424 2%	0.258 5%	0.232 16%	0.191 0%
GMM_C	5.850 4%	1.618 58%	0.636 44%	0.526 19%	0.343 27%	8.484 18%
				N = 100		
TS: GRS		0.503 58%	0.274 0%	0.290 0%	0.263 0%	0.237 3%
Beta: $OLS_{Shanken}$		0.345 0%	0.223 0%	0.202 0%	0.193 0%	0.184 1%
Beta: $GLS_{Shanken}$		0.462 0%	0.243 0%	0.270 0%	0.236 0%	0.198 1%
Beta: $WLS_{Shanken}$		0.348 98%	0.229 95%	0.205 55%	0.196 4%	0.188 51%
GMM_A: Fist-stage		0.285 0%	0.187 0%	0.162 0%	0.160 0%	0.169 1%
GMM_A: Second-stage		0.928 ✓	0.886 ✓	0.823 ✓	0.329 ✓	0.210 ✓
GMM_B: Fist-stage		0.340 0%	0.220 0%	0.199 0%	0.190 0%	0.181 0%
GMM_B: Second-stage		0.872 0%	0.244 0%	0.290 0%	0.289 0%	0.188 0%
GMM_C		21.013 0%	6.322 0%	1.999 0%	5.203 1%	1.535 6%

Panel D. N Portfolios formed on ME MOM

Method	**T = 60**	**T = 120**	**T = 240**	**T = 360**	**T = 480**	**T = 948**
			N = 6			
TS: GRS	0.661 9%	0.420 17%	0.277 0%	0.304 0%	0.267 0%	0.159 0%
Beta: $\text{OLS}_{\text{Shanken}}$	0.355 8%	0.221 2%	0.216 0%	0.225 0%	0.199 0%	0.102 0%
Beta: $\text{GLS}_{\text{Shanken}}$	0.666 0%	0.385 2%	0.263 0%	0.298 0%	0.255 0%	0.110 0%
Beta: $\text{WLS}_{\text{Shanken}}$	0.427 11%	0.266 13%	0.256 1%	0.265 0%	0.239 0%	0.123 3%
GMM_{A}: Fist-stage	0.260 11%	0.207 3%	0.197 0%	0.205 0%	0.182 0%	0.092 0%
GMM_{A}: Second-stage	1.091 ✓	0.305 ✓	0.263 ✓	0.222 ✓	0.194 ✓	0.104 ✓
GMM_{B}: Fist-stage	0.290 22%	0.181 2%	0.176 0%	0.183 0%	0.162 0%	0.084 0%
GMM_{B}: Second-stage	1.101 22%	0.337 2%	0.234 0%	0.326 0%	0.240 0%	0.146 0%
GMM_{C}	4.611 13%	1.043 6%	0.505 0%	1.145 0%	0.775 0%	0.108 0%
			N = 25			
TS: GRS	0.801 18%	0.505 5%	0.306 0%	0.330 0%	0.286 0%	0.213
Beta: $\text{OLS}_{\text{Shanken}}$	0.418 1%	0.336 0%	0.241 0%	0.246 0%	0.205 0%	0.100 0%
Beta: $\text{GLS}_{\text{Shanken}}$	0.780 0%	0.507 0%	0.304 0%	0.325 0%	0.252 0%	0.112 0%
Beta: $\text{WLS}_{\text{Shanken}}$	0.446 14%	0.359 7%	0.259 2%	0.266 0%	0.226 0%	0.112 18%
GMM_{A}: Fist-stage	0.327 0%	0.303 0%	0.217 0%	0.217 0%	0.183 0%	0.088 0%
GMM_{A}: Second-stage	0.805 ✓	0.555 ✓	0.627 ✓	0.533 ✓	0.420 ✓	0.167 ✓
GMM_{B}: Fist-stage	0.395 0%	0.317 0%	0.228 0%	0.232 0%	0.194 0%	0.095 0%
GMM_{B}: Second-stage	0.708 0%	0.545 0%	0.277 0%	0.295 0%	0.214 0%	0.129 0%
GMM_{C}	8.612 5%	3.148 0%	1.614 1%	1.831 0%	1.033 0%	0.221 0%

Panel E. N Industry Portfolios

Method	**T = 60**	**T = 120**	**T = 240**	**T = 360**	**T = 480**	**T = 948**
			N = 5			
TS: GRS	0.223 85%	0.190 49%	0.185 8%	0.193 2%	0.189 1%	0.178 0%
Beta: $\text{OLS}_{\text{Shanken}}$	0.111 75%	0.042 92%	0.023 93%	0.045 76%	0.044 72%	0.053 45%
Beta: $\text{GLS}_{\text{Shanken}}$	0.117 73%	0.048 93%	0.025 93%	0.057 75%	0.053 71%	0.062 43%
Beta: $\text{WLS}_{\text{Shanken}}$	0.121 74%	0.047 91%	0.023 93%	0.055 67%	0.052 63%	0.060 33%
GMM_{A}: Fist-stage	0.091 68%	0.036 93%	0.020 94%	0.040 77%	0.040 73%	0.048 48%
GMM_{A}: Second-stage	0.105 ✓	0.046 ✓	0.033 ✓	0.053 ✓	0.052 ✓	0.060 ✓
GMM_{B}: Fist-stage	0.088 67%	0.033 93%	0.018 94%	0.036 78%	0.034 74%	0.042 49%
GMM_{B}: Second-stage	0.130 68%	0.050 93%	0.036 94%	0.060 78%	0.057 74%	0.065 49%
GMM_{C}	0.777 80%	0.069 94%	0.038 94%	0.056 76%	0.055 72%	0.068 46%
			N = 17			
TS: GRS	0.392 26%	0.290 64%	0.200 12%	0.198 1%	0.181 0%	0.139 0%
Beta: $\text{OLS}_{\text{Shanken}}$	0.334 78%	0.241 62%	0.152 84%	0.160 63%	0.135 37%	0.102 49%
Beta: $\text{GLS}_{\text{Shanken}}$	0.424 89%	0.243 60%	0.172 79%	0.181 57%	0.154 30%	0.107 46%
Beta: $\text{WLS}_{\text{Shanken}}$	0.337 81%	0.248 89%	0.157 90%	0.161 55%	0.137 40%	0.102 30%
GMM_{A}: Fist-stage	0.311 88%	0.215 56%	0.135 87%	0.142 52%	0.123 29%	0.095 51%
GMM_{A}: Second-stage	0.478 ✓	0.332 ✓	0.212 ✓	0.176 ✓	0.153 ✓	0.103 ✓
GMM_{B}: Fist-stage	0.308 88%	0.222 56%	0.140 90%	0.148 63%	0.124 41%	0.094 55%
GMM_{B}: Second-stage	0.435 88%	0.274 56%	0.204 90%	0.187 61%	0.167 40%	0.106 55%
GMM_{C}	0.609 93%	0.246 51%	0.177 83%	0.176 52%	0.149 28%	0.103 48%
			N = 30			
TS: GRS	0.541 50%	0.435 69%	0.248 8%	0.232 1%	0.214 0%	0.203 0%
Beta: $\text{OLS}_{\text{Shanken}}$	0.506 7%	0.345 79%	0.206 73%	0.190 63%	0.170 67%	0.146 30%
Beta: $\text{GLS}_{\text{Shanken}}$	0.640 3%	0.439 48%	0.268 59%	0.207 58%	0.184 59%	0.157 22%
Beta: $\text{WLS}_{\text{Shanken}}$	0.542 52%	0.381 99%	0.215 89%	0.190 70%	0.172 63%	0.149 18%
GMM_{A}: Fist-stage	0.458 12%	0.289 79%	0.183 69%	0.172 65%	0.155 64%	0.134 32%
GMM_{A}: Second-stage	0.617 ✓	0.424 ✓	0.262 ✓	0.215 ✓	0.165 ✓	0.159 ✓
GMM_{B}: Fist-stage	0.483 1%	0.329 72%	0.196 74%	0.181 60%	0.162 69%	0.139 37%
GMM_{B}: Second-stage	0.714 1%	0.446 73%	0.246 75%	0.193 61%	0.184 69%	0.160 37%
GMM_{C}	0.723 5%	0.494 77%	0.225 61%	0.210 63%	0.174 60%	0.150 26%

Chapter 4

The efficiency of the SDF and Beta methods at evaluating multi-factor asset-pricing models

[§] [†]

The classical beta method and the stochastic discount factor (SDF) method may be considered competing paradigms for empirical work in asset pricing. The two methods are equally efficient at estimating risk premiums in the context of the single-factor

[§]A paper-version of this chapter – in Southwestern Finance Association 48th Annual Meeting Proceedings – is currently being written jointly with Stuart Hyde and Ian Garrett (Manchester Business School).

[†]An earlier version of this work was presented (or accepted for presentation) at the 2008 Annual Doctoral Conference at Manchester Business School; Internal Seminar (University Carlos III de Madrid, April 2008); Jornada de Investigación en Finanzas (Universidad de Castilla-La Mancha, October 2008); XVI AEFIN Finance Forum (ESADE Business School, Barcelona, November 2008); International Conference on Finance (National Taiwan University, December 2008); Southwestern Finance Association 48th Annual Meeting; (Oklahoma City, February 2009); Royal Economic Society Conference (London, April 2009); Symposium in Statistics and Econometrics (Lausanne, April 2009); Eastern Finance Association Annual Meeting (Washington D.C., May 2009); Asian Finance Association Conference (Brisbane, July 2009); and at the Symposium in Economics and Finance (Geneva, July 2009). I would like to thank Genaro Sucarrat (Universidad Carlos III de Madrid), Chien-Ting Lin (University of Adelaide), and Sheng Guo (Florida International University) for helpful comments as discussants. Martín Lozano gratefully acknowledges financial assistance from the Delegation of the Basque Government in México (Graduate Student Mobility Grant 2007) and especially to the European Union Marie Curie Program.

model. We show this does not hold for multi-factor models. Inference is consistently more reliable in the Beta method for the estimates in models which include size, value and momentum factors. However, our evidence also illustrates that the SDF method is generally more efficient at estimating sample pricing errors. Finally, the specification test in the Beta method tends to under-reject in finite samples while the SDF method has approximately the correct size. Our Monte Carlo simulation results are consistent whether we use a normal or empirical distribution, or different sets and sizes of tests portfolios.

In this chapter we focus our analysis on simulated data, which it is necessary to develop a finite sample analysis of the estimator efficiency. On the previous chapter, we focus our analysis on several datasets of historical data. Another difference is that in previous chapter we evaluate several estimators in the Beta method, while in this chapter we focus in OLS estimators in order to give more importance to the analysis of the estimator efficiency. Finally, previous chapter analyzes the model performance, while in this chapter we focus our analysis on the method performance.

4.1 Introduction

Empirical finance widely adopts either the classical *Beta* method or the *stochastic discount factor* (SDF) method for the evaluation of asset-pricing models. The Beta method involves estimating the beta representation where the expected return on an asset is a linear function of its factor betas. This approach is widely implemented in the finance literature (see Kan, Robotti and Shanken [65]) using the two-stage cross-sectional regression methodology advocated by Black, Jensen and Scholes [9] and Fama and MacBeth [32]. In the SDF representation, the value of an asset equals the expected value of the product of the asset's payoff and the SDF.[1] This approach estimates the asset pricing model using its SDF representation and the generalized method of moments (GMM).

Typically, it is common for researchers to select one approach over the other (Beta or SDF) and consequently certain specific areas of the literature appear to favor one method over the other. In fact, it is a common trend to compare procedures within methodologies. For example, Jagannathan and Wang [57] compare the asymptotic efficiency of the two-stage cross-sectional regression method and the FamaMacBeth procedure; Shanken and Zhou [94] analyze the finite sample properties and empirical performance of the Fama-MacBeth, maximum likelihood, and generalized method of moments for Beta pricing models[2]; other related examples can be found in Farnsworth, Ferson, Jackson and Todd [33], Velu and Zhou [101], Kan and Robotti [64, 63], Chen and Kan [22], and Amsler and Schmidt [2], just to mention a few. However, only recently have there been attempts to evaluate the two approaches.

In particular, Kan and Zhou [68, 66] and Jagannathan and Wang [58] were the first who evaluate and compare the two methods by examining the efficiency of the SDF approach relative to the Beta method in the framework of a single factor model. The lack of previous studies about the comparison of the Beta and the SDF methods may respond to the fact that there is no a direct one-to-one mapping between the estimators from both methodologies. Here, we intend to contribute to the knowledge about their differences in terms of efficiency.

[1] This was first pointed out by Ross [90] and Dybvig and Ingersoll [28] who derive the SDF representation for the CAPM.

[2] In fact, Shanken [92] provided the first comprehensive analysis of the statistical properties of the classical two-pass estimator on beta models under the assumption that returns and factors exhibit conditional homoscedasticity.

Kan and Zhou [68] made the first formal comparison of both methods in a standardized single-factor model, where the factor mean and variance are known in advance and the factor can be normalized to have zero mean and unit variance.[3] Under this specific assumption, the factor risk premium from the Beta method numerically coincides with the linear coefficient associated with the factor in the pricing kernel of the SDF method but they find that the SDF method is less efficient than the Beta method when both are estimated using GMM. However, factors used in empirical work generally do not have zero mean and unit variance thus the estimates will not be identical and comparison of their estimation efficiency becomes more difficult. Jagannathan and Wang [58] and Cochrane [23] discuss these issues with nonstandardized factors.

Jagannathan and Wang [58] show that under an alternative framework, which augments each method by additional moment conditions, the SDF method is as efficient as the beta method. They note that while the risk premium in the SDF method is not equal to the risk premium in the beta method, they are related by a one-to-one transformation. Explicitly accounting for this transformation they show in the context of the market risk premium from a single-factor model that the Beta method does not dominate the SDF method. Cochrane [23] reaches a similar conclusion.

However much empirical research in asset pricing employs multi-factor models, e.g. adopting the Fama-French three factor model [30, 31] or the Carhart four factor model [18] rather than solely relying on inference from the single-factor CAPM. We extend the empirical work of Jagannathan and Wang [58] to cover this gap and ask whether the estimation efficiency of the SDF method is still similar to the Beta method when

[3] As pointed out by Cochrane [23], this is unusual, but not incorrect, since any mean-variance efficient portfolio can serve as reference return.

one employs more than one factor. To compare the methods we examine the estimation of the risk premiums, the sample pricing errors and the associated specification tests. Our key results show that the finite-sample efficiency of the methods depend on (i) the number of factors included in the model, (ii) the GMM moment restrictions imposed in each method and (iii) the degree of non-normality of the adopted factors. Therefore, the estimation efficiency of the methods may differ.

When evaluating the methods with the single factor CAPM, we find both methods lead to the same results for estimating the market risk premium, reinforcing the previous findings of Jagannathan and Wang [58], Cochrane [25] and the previous chapter as well. However, our findings indicate that this does not hold for the multi-factor models, in which inference is consistently more reliable in the Beta method. This suggests that choosing the single factor CAPM for evaluating the efficiency of risk premiums constitutes a fairly weak scenario. Indeed, we are unable to see any significant difference, even if we allow for different sample sizes, alternative numbers of return portfolios or different factors' distributions.

On the other hand, the relative advantage of the Beta method at estimating risk premiums does not apply to the estimation of sample pricing errors, where invariably the efficiency of the SDF method is superior, even in the smallest sample considered. This result is expected given the previous result on last chapter. Consequently, the specification test in the Beta method generally under-rejects in finite samples whereas the SDF method over-rejects but has roughly the correct size[4].

[4]One added value of performing simulation analysis is that we can perform size and power tests, which were not possible to compute in last chapter. This represent a significant contribution to the literature since there are very few evidence about size and power tests applied on Beta and SDF methods. Recent works such as Grauer and Janmaat [42] examine power tests for competing Beta pricing models.

Our results are consistent whether we assume that returns and factors are drawn from a multivariate normal distribution or from the empirical distribution estimated by the bootstrap method. They are also consistent to different sets and sizes of test portfolios such as the 10 single-sorted size, the 25 double-sorted size/book-to-market (Fama-French) portfolios and the 30 industry-sorted portfolios.

The main implication of our results is that if we are interested on making inference on a multi-factor model estimator(s), we should prefer the Beta method over the SDF method. Conversely, if we are primarily interested on making inference on the sampling pricing error or Jensen's alpha, the SDF method should be preferred. This argument primarily relies on the lower simulated standard error of the random draws of the estimator and sample pricing errors. Hence, there is no method that fully dominates the other, rather they are complementary. Similar conclusions are reached in related papers, such as Shanken and Zhou [94] who conduct a simulation analysis of several procedures applied to Beta models and claim that no single estimation procedure dominates in all respects.

Even though the purpose of the previous chapter is different from this chapter, we can find some similarities. For instance, we previously find some indication of greater efficiency of the Beta method for estimating risk premiums by comparing the estimate value and the bias from the risk premium in percent values. However, in order to test this hypothesis it is necessary to conduct a finite sample analysis, as we do in this chapter.

Our results contribute to cover an important gap in the empirical asset pricing literature, since the generalized idea about the Beta and SDF methods is that both lead to almost identical results in terms of efficiency. For example, Ferson [35] concludes that when the two methods correctly exploit the same moments they deliver nearly identical results. Cochrane [25] also comes to similar conclusions comparing the efficiency of the estimation of the risk premiums. However, we argue that these conclusions are limited. Once the asset pricing model under consideration includes more factors with greater non-normality, the differences in terms of efficiency clearly emerge.[5]

The main difference with previous chapter is that we now perform a finite sample analysis using simulated data. Furthermore, in previous chapter we focus on the model performance while in this chapter our main interest is the method comparison in terms of efficiency. This final chapter is intended to be a response and an extension of the work of Jagannathan and Wang [58] and Kan and Zhou [66, 68] about the comparison of these two methodologies. Therefore, our results on this chapter have direct implications over statistical inference in empirical works.

The outline of the remainder of the paper is as follows. In section 4.2 we present the methodology, describing both the Beta method and the SDF method, how comparison of the methods is undertaken and details the Monte Carlo simulation procedure. Section 4.3 presents the results while section 4.4 concludes.

4.2 Methodology

To compare the estimators and test statistics derived from both the Beta and SDF methods we use Hansen [46]'s GMM methodology[6]. This approach is common, Kan

[5]This is complementary to the findings of Kan and Zhou [66] who argue that estimation is sensitive to the presence of skewness and kurtosis.

[6]Skoulakis [96] follow the CRS method in a similar comparison.

and Zhou [68, 66] and Jagannathan and Wang [58] also employ GMM to examine both approaches. Although the Beta method can be applied using the common two-stage Fama and MacBeth [32] approach or by using the maximum likelihood procedure. Shanken and Zhou [94] examine the performance of the alternative estimation methods in the context of the Beta method. While the GMM approach reduces to both estimators under the appropriate assumptions, it is less restrictive allowing for conditional heteroskedasticity, serial correlation and non-normality. For more examples of applications of the GMM methodology in finance, see Jagannathan, Skoulakis and Wang [54].

Even though we have presented the Beta and the SDF method before in previous chapters, we have not show how to estimate the Beta method via GMM. Therefore, we will have to rewrite some equations. Furthermore, in previous chapter we present risk premium estimates for Beta and SDF methods separately, having different units. Here, it would be crucial to transform one of them in order to have the same units even though they came from different methodologies. This will facilitate to compare their standard errors, size and power tests. For all these reasons, it is convenient to present in some detail the particularities of both methodologies.

4.2.1 The Beta method

Following [58] notation, we denote r_t as the vector of N stock returns in excess of the risk-free rate and f_t a vector of K economy-wide pervasive risk factors during period t. According to the notation in previous chapters, the mean and the covariance matrix of the factors are denoted by μ, where $\mu = \mathrm{E}[f_t]$, and Σ_f respectively. The standard linear asset-pricing model under the Beta representation was introduced by equation

1.3.1, in this section we will rewrite it in order to maintain the usual notation in related works such as [23, 58, 68, 66]. Thus, the Beta representation is given by

$$\mathrm{E}\left[r_t\right] = \delta\beta \tag{4.2.1}$$

where δ is the vector of factor risk premiums, and β is the matrix of factor loadings which measure the sensitivity of asset returns to the factors, defined as

$$\beta_{N\times K} \equiv \mathrm{E}\left[r_t\left(f_t - \mu\right)'\right]\Sigma_f^{-1} \tag{4.2.2}$$

Equivalently, we can identify β as a parameter in the time-series regression (equation 1.2.1): $r_t = \phi + \beta f_t + \epsilon_t$ where the residual ϵ_t has zero mean and is uncorrelated with the factors f_t. The specification of the asset-pricing model under the Beta representation in equation (4.2.1) imposes the following restriction on the time-series intercept, $\phi = (\delta - \mu)\beta$. By substituting this restriction in the regression equation, we obtain:

$$r_t = \left(\delta - \mu + f_t\right)\beta + \epsilon_t \quad \text{where} \quad \begin{cases} \mathrm{E}\left[\epsilon_t\right] = 0_N \\ \mathrm{E}\left[\epsilon_t f_t'\right] = 0_{N\times K} \end{cases} \tag{4.2.3}$$

Hence, the Beta representation in equation (4.2.1) gives rise to the factor model, equation (4.2.3). The associated moment conditions of the factor model, equation (4.2.3) are:

$$\begin{aligned} &\mathrm{E}\left[r_t - \left(\delta - \mu + f_t\right)\beta\right] = 0_N \\ &\mathrm{E}\left[\left[r_t - \left(\delta - \mu + f_t\right)\beta\right] f_t'\right] = 0_{N\times K} \end{aligned} \tag{4.2.4}$$

However, when the factor is the return on a portfolio of traded assets, as in the single and multi-factor models analyzed in this paper – the CAPM, the Fama-French three factor model, and the Carhart four factor model – it can be verified that the estimate of μ (the sample mean of the factor) is also the estimate of the risk premium

δ.[7] Therefore, given $\delta = \mu$, the moment conditions given in equation (4.2.4) simplify to

$$\begin{aligned} &\mathrm{E}\left[r_t - f_t\beta\right] = 0_N \\ &\mathrm{E}\left[\left(r_t - f_t\beta\right) f_t'\right] = 0_{N\times K} \\ &\mathrm{E}\left[f_t - \mu\right] = 0_K \end{aligned} \tag{4.2.5}$$

where neither δ or μ appear in the first two restrictions of equation (4.2.5) but it is necessary to include the definition of μ to identify the vector of risk premiums δ as a third moment restriction.[8,9] Now, following the usual GMM notation, we define the vector of unknown parameters $\theta^{\mathrm{eq}(4.2.5)} = \left[\mathrm{vec}\left(\beta\right)' \quad \mu'\right]'$, where the vec operator 'vectorizes' the $\beta_{N\times K}$ matrix by stacking its columns, and the observable variables are $x_t = \left[r_t' \quad f_t'\right]'$. Then, the function g in the moment restriction is given by

$$g\left(x_t, \theta^{\mathrm{eq}(4.2.5)}\right)_{(N+NK+K)\times 1} = \begin{pmatrix} r_t - f_t\beta \\ \mathrm{vec}\left[\left(r_t - f_t\beta\right) f_t'\right] \\ f_t - \mu \end{pmatrix} \tag{4.2.6}$$

Now, for any θ, the sample analogue of $\mathrm{E}[g\left(x_t, \theta\right)]$ is equal to

$$g_T\left(\theta\right) = \frac{1}{T}\sum_{t=1}^{T} g\left(x_t, \theta\right) \tag{4.2.7}$$

Then, a natural estimation strategy for θ is to choose the values that make $g_T\left(\theta\right)$

[7]Non-traded factors are economic factors such as consumption growth used in the Consumption CAPM see [12] or industrial production growth and inflation adopted in linear factor models, see Chen, Roll, and Ross [21] for similar analysis as well.

[8]Nevertheless, it is also possible to estimate the last moment restriction of equation (4.2.5) outside the GMM framework by computing $\mu = E\left[f_t\right]$. This is because the number of added moment restrictions in equation (4.2.5) compared with equation (4.2.4) is the same as the number of added unknown parameters. Hence, the efficiency of equation (4.2.4) and equation (4.2.5) remains the same. By following this alternative, we drop the factor-mean moment condition without ignoring that it has to be estimated.

[9]An additional moment condition to estimate the variance Σ_f could also be added to equation (4.2.5). However the variance can also be estimated outside the GMM framework without affecting efficiency.

as close to the zero vector as possible. For that reason we choose θ to solve[10]

$$\min_{\theta} g_T(\theta)' W^{-1} g_T(\theta) \tag{4.2.8}$$

To compute the first-stage GMM estimator θ_1 we consider $W = I$ in equation (4.2.8). The second-stage GMM estimator θ_2 is the solution of equation (4.2.8) when the weighting matrix is the spectral density matrix of $g(x_t, \theta_1)$:

$$S = \sum_{j=-\infty}^{\infty} \mathrm{E}\left[g(x_t, \theta_1)\, g(x_t, \theta_1)'\right] \tag{4.2.9}$$

In order to examine the validity of the pricing model derived from the moment restrictions in equation (4.2.5) we can test whether the vector of N Jensen's alphas, given by $\alpha = \mathrm{E}[r_t] - \delta\beta$ is jointly equal to zero.[11] This can be done using the J-statistic with an asymptotic χ^2 distribution. Given there are $N + NK + K$ equations and $NK + K$ unknown parameters in equation (4.2.6), then the degrees of freedom is N.

The covariance matrix of the pricing errors, $\mathrm{Cov}(g_T)$, is given by

$$\mathrm{Cov}(g_T) = \frac{1}{T}\left[\left(I - \beta(\beta'\beta)^{-1}\beta'\right) S \left(I - \beta(\beta'\beta)^{-1}\beta'\right)\right] \tag{4.2.10}$$

and the test is a quadratic form of the vector of pricing errors. In particular, the Hansen [46] J-statistic is computed as (see also equation 1.5.5)

$$\begin{array}{ll} \text{First-stage:} & g_T(\theta_1)' \,\mathrm{Cov}(g_T)^{-1} g_T(\theta_1) \sim \chi^2_N \\ \text{Second-stage:} & T g_T(\theta_2)' S^{-1} g_T(\theta_2) \sim \chi^2_N \end{array} \tag{4.2.11}$$

Both the first and second-stage statistic in equation (4.2.11) lead to the same numerical value. However, if we weight equations (4.2.10) and (4.2.11) by any other matrix different to S, such as $\mathrm{E}[r_t r_t']$ or $\mathrm{Cov}[r_t]$, this result no longer holds.

[10] Even though we introduce this equation earlier in section 1.5.1, it is important to show it again for illustration purposes.

[11] This approach is known as the restricted test, see MacKinlay and Richardson [80].

4.2.2 The SDF method

To derive the SDF representation from the Beta representation we follow Ferson and Jagannathan [38] and Jagannathan and Wang [58] among others. First, we substitute the expression for β (equation 4.2.2) into equation (4.2.1) and rearrange the terms, to give:

$$\mathrm{E}\left[r_t\right] - \mathrm{E}\left[r_t\delta'\Sigma_f^{-1}f_t - r_t\delta'\Sigma_f^{-1}\mu'\right] = \mathrm{E}\left[r_t\left(1+\delta'\Sigma_f^{-1}\mu - \delta'\Sigma_f^{-1}f_t\right)\right] = 0_N$$

again, if we are considering traded factors, then $\delta = \mu$ so $1+\delta'\Sigma_f^{-1}\mu = 1+\mu'\Sigma_f^{-1}\mu \geq 1$, then divide each side by $1+\delta'\Sigma_f^{-1}\mu$,[12]

$$\mathrm{E}\left[r_t\left(1 - \frac{\delta'\Sigma_f^{-1}}{1+\delta'\Sigma_f^{-1}\mu}f_t\right)\right] = 0_N$$

If we transform the vector of risk premiums δ into a vector of new parameters λ as follows,

$$\lambda = \frac{\delta'\Sigma_f^{-1}}{1+\delta'\Sigma_f^{-1}\mu} \tag{4.2.12}$$

then we obtain the following SDF representation and moment restriction of the linear asset-pricing model,

$$\mathrm{E}\left[r_t\left(1-\lambda f_t\right)\right] = 0_N \tag{4.2.13}$$

where the random variable $m_t \equiv 1 - f_t'\lambda$ is the SDF because $\mathrm{E}[r_t m_t] = 0_N$.[13]

From the moment restrictions, equation (4.2.13), we obtain the vector of N pricing errors defined as $\pi = \mathrm{E}[r_t] - \lambda\mathrm{E}[r_t f_t]$. The analytical solution of equation (4.2.13) is obtained by GMM.[14] Writing the sample pricing errors as

$$g_T\left(\lambda\right) = -\mathrm{E}\left[r_t\right] + \lambda\mathrm{E}\left[r_t f_t\right] \tag{4.2.14}$$

[12]Even when the factors are not traded, it is common to suppose $1+\delta'\Sigma_f^{-1}\mu \neq 0$.

[13]Alternatively, we could derive the Beta representation from the SDF representation by expanding m and rearranging the terms.

[14]This is useful given the need to undertake vast numbers of simulations. Similar simplifications of multi-dimensional optimization problems for Beta models can be found in Shanken and Zhou [94].

define $d = -\frac{\partial g_T(\lambda)}{\partial \lambda'} = \mathrm{E}\left[r_t f_t\right]$, the second-moment matrix of returns and factors. The first-order condition to minimize the quadratic form of the sample pricing errors, equation (4.2.8), is $-d'W\left[\mathrm{E}\left[r_t\right] - \lambda d\right] = 0$, where W is the GMM weighting matrix, equal to the identity matrix in the first-stage estimator and equal to the spectral density matrix S, equation (4.2.9), in the second-stage estimator. Therefore, the GMM estimates of λ are:

$$\begin{aligned} \widehat{\lambda}_1^A &= \left(d'd\right)^{-1} d'\mathrm{E}\left[r_t\right] \\ \widehat{\lambda}_2^A &= \left(d'S^{-1}d\right)^{-1} d'S^{-1}\mathrm{E}\left[r_t\right] \end{aligned} \tag{4.2.15}$$

Specifying the SDF as a linear function of the factors as in equation (4.2.13) has been very popular in the empirical literature. However, Kan and Robotti [64] point out that this is problematic because the specification test statistic is not invariant to an affine transformation of the factors. Therefore, following [64], we also consider an alternative specification that defines the SDF as a linear function of de-meaned factors. We decorate with an A to $\widehat{\lambda}$ to indicate that the estimator comes from the un-meaned specification and with a B to indicate that comes from the de-meaned specification.[15]

The alternative de-meaned version of equation (4.2.13) is defined as:

$$\mathrm{E}\left[r_t\left[1 - \lambda\left(f_t - \mu\right)\right]\right] = 0_N \tag{4.2.16}$$

According to [55] and [58], it is also possible to estimate μ in equation (4.2.16) outside of the GMM estimation by computing $\mu = \mathrm{E}[f_t]$. This is because the number of added moment restrictions is the same as the number of added unknown parameters. Hence, the efficiency of the estimators remains the same. By following this alternative,

[15] Burnside [14] also finds some advantages of the de-meaned version in terms of specification tests. This de-meaned SDF specification can be also found in Cochrane [25], and in Balduzzi and Yao [4].

we can drop the factor-mean moment condition without ignoring that it has to be estimated, and obtain analytical expressions for $\widehat{\lambda}_1^B$ and $\widehat{\lambda}_2^B$.

Naturally, the procedure to solve the moment restrictions in equation (4.2.16) is similar to that for the un-meaned SDF^A method. In particular, we substitute $\mathrm{E}[r_t f_t]$ for $\mathrm{Cov}[r_t f_t]$ in equation (4.2.14), then define $b = -\frac{\partial g_T(\lambda)}{\partial \lambda'}$ as the covariance matrix of returns and factors. Finally, the SDF^B first and second stage GMM estimates are:

$$\begin{aligned} \widehat{\lambda}_1^B &= \left(b'b\right)^{-1} b'\mathrm{E}\left[r_t\right] \\ \widehat{\lambda}_2^B &= \left(b'S^{-1}b\right)^{-1} b'S^{-1}\mathrm{E}\left[r_t\right] \end{aligned} \tag{4.2.17}$$

The specification tests can be conducted by following equations (4.2.7) and (4.2.11), the only difference being that we substitute β by $d = \mathrm{E}[r_t f_t]$ (the second moment matrix of returns and factors) for the SDF^A case, and by $b = \mathrm{Cov}[r_t f_t]$ (the covariance matrix of returns and factors) for the SDF^B case. The degrees of freedom in equation (4.2.11) are specific for the Beta method, in the SDF method the degrees of freedom is equal to $N - K$, since there are N equations and K unknown parameters in both equations (4.2.13) and (4.2.16).

Equations (4.2.10) and (4.2.11) are weighted by equation (4.2.9), since it is statistically optimal. This approach was first suggested by Hansen [46] as it maximizes the asymptotic statistical information in the sample about a model, given the choice of moments. However, there are also alternatives for this weighting matrix which are suitable for model comparisons because they are invariant to the model and their parameters. For instance, Hansen and Jagannathan [49] suggest the use of the second moment matrix of excess returns $W = \mathrm{E}[r_t r_t']$ instead of $W = S$. Also, Burnside [14], Balduzzi and Yao [4], and Kan and Robotti [64] suggest that the SDF^B method should use the covariance matrix of excess returns $W = \mathrm{Cov}[r_t]$. We investigate the implications of using these alternative weighting matrices.

4.2.3 Comparison of the methods

There is a one-to-one mapping between δ (from $\theta^{\text{eq}(4.2.5)}$) and λ (from equations 4.2.15 and 4.2.17), which facilitates the comparison of the two methods.[16] Hence we can derive an estimate of λ not only by the SDF method but also by the Beta method. By the same token we can derive an estimate of δ not only by the Beta method but also by the SDF method. Therefore, for convenience, variables decorated with '$*$' refer to the estimates from the Beta method and with '$\widehat{\ }$' to the estimates from the SDF method. From the previous definition of λ in equation (4.2.12), we have:

$$\lambda = \frac{\delta}{\Sigma_f + \delta\mu'} \qquad \text{or} \qquad \delta = \frac{\Sigma_f \lambda}{1 - \mu'\lambda} \tag{4.2.18}$$

Remember μ and Σ_f represent the mean and the variance of the factor f, while δ and λ represent the risk premium estimators from the Beta (solving 4.2.8) and SDF method (solving 4.2.15 or 4.2.17 respectively. In a similar way, substituting equation (4.2.18) into π, we can find a one-to-one mapping between π from the SDF method and α from the Beta method.

$$\pi = \frac{\Sigma_f}{\Sigma_f + \delta\mu'}\alpha \qquad \text{or} \qquad \alpha = \frac{\Sigma_f + \delta\mu'}{\Sigma_f}\pi \tag{4.2.19}$$

In the first formal attempt to compare both methods, Kan and Zhou [68] assume that the factor has zero mean and unit variance, that is $\mu = 0$ and $\Sigma_f = 1$. In this standardized single factor model, equations (4.2.18) and (4.2.19) imply $\lambda = \delta$ and $\pi = \alpha$. By assuming that the mean and the variance of the factor are predetermined without estimation, they ignore the sampling errors associated with the estimates of μ and Σ_f and conclude that the estimates of the Beta method are more efficient. Jagannathan and Wang [58] and Cochrane [23] explain the effects of standardized factors,

[16] Thanks to Raymond Kan (University of Toronto) for kindly sharing complementary econometric notes on Kan and Zhou [66].

showing that in general, predetermining the factor moments reduces the sampling error of the estimate in the Beta method and not in the SDF method.

However, with the Beta moment restrictions, equation (4.2.5), we only can make inference on δ, not on λ. Yet to compare the methods using equation (4.2.18) requires an estimator of Σ_f. One solution is to add an additional moment condition to equation (4.2.5) to estimate Σ_f. An alternative is to estimate μ and Σ_f outside the GMM estimation. In simulation results not showed here, we find that the efficiency of both alternatives is the same. Hence we elect to estimate Σ_f outside the GMM estimation.

Predetermining the values of μ and Σ_f to be known constants – not necessarily $\mu = 0$ and $\Sigma_f = 1$ – gives an informational advantage to the Beta method in terms of efficiency. Predetermining without estimation implies ignoring the sampling errors associated with μ^* and Σ_f^*, as a consequence λ^* becomes considerably more efficient than if we follow equation (4.2.5). In our simulation analysis, we consider the case where μ and Σ_f must be estimated.

To summarize, the Beta method gives the GMM estimate δ^* while the SDF method gives the GMM estimate $\widehat{\lambda}$. In our Monte Carlo simulation results, we transform the estimate δ^* into an estimate of λ and then compare the variances of the sampling distribution of λ^* and $\widehat{\lambda}$. In the same way, we transform α^* into an estimate of π and then compare the efficiency of π^* and $\widehat{\pi}$.

We also compare the distributions of Hansen's [46] test of overidentification using the J-statistic of the transformed beta J^* and $\widehat{J}$ from the SDF method. The null hypothesis is that all pricing errors are zero. In the size tests we calculate the probability of rejection under the null that the asset pricing model is true, in the power tests we calculate the probability of rejection under the null that the asset pricing model is false. To misspecify the asset pricing model, we attach to the usual methodology of using fixed alternatives[17] such as in Jagannathan and Wang [58] and Kan and Zhou [66] among others.

4.2.4 Monte Carlo simulation

Researchers are faced with data sets of finite, and occasionally rather small, sample sizes. It is therefore imperative to obtain a sense of the small sample performance of the two methods. Since finite-sample analytical results can be obtained only under certain distributional assumptions, it is customary to resort to some simulation technique, which allows us to alter the simulation input and develop an understanding of how sensitive the results are with respect to the various features of the data generating process.

We use Monte Carlo simulation [18] to whether the asymptotic GMM estimators and test statistics have any bias. In particular we are interested in evaluating the standard deviation of λ^*, $\widehat{\lambda}$, π^*, $\widehat{\pi}$ and also the tail of the J-statistic distribution to conduct specification tests. We assume that the factors f_t are drawn either from a multivariate normal or an empirical distribution estimated by the bootstrap method. Using the empirical distribution allows for non-normalities, autocorrelation, heteroskedasticity and non-independence of factors and residuals.

To artificially generate the excess returns we use the factor model, equation (4.2.3)

[17] Hall and Inoue [45] show the limiting distribution theory for the GMM estimator when the estimation is based on a population moment condition which is subject to nonlocal (or fixed) misspecification.

[18] Simulations were executed using the North West Grid computational facilities. See Ahn and Gadarowski [1] for an examination of finite-sample properties of several model tests methods.

where $t = 1, ..., T$. For T, we consider the following four time horizons: 60, 360, 600 and 1000 months. As Shanken and Zhou [94] argue, varying T is useful in order to understand the small-sample properties of the tests and the validity of asymptotic approximations. For instance, we elect to examine a 5 year window since this may show how distorted results from taking a really small sample could potentially be and also it is a commonly adopted horizon when using rolling windows, a 30 year window corresponds approximately to the sample sizes of Fama and French [29, 30] and Jagannathan and Wang [56] while the 600 month sample matches the largest sample examined by Jagannathan and Wang [58]. We also examine 1000 months since this approximates the current size of the largest sample available on the Kenneth French's library [July 1926 to December 2007 – 978 months]. The estimators and specification tests are then calculated based on the T samples of the factors and returns generated from the factor model. We repeat this independently to obtain 10,000 draws of the estimators of λ, π (the pricing errors) and J (the overidentifying restriction statistic).

Previous related empirical studies such as Kan and Zhou [68, 66], Jagannathan and Wang [58] and Cochrane [25] focus on the CAPM model to test the Beta and SDF methods. Our contribution is to evaluate the methods on multi-factor models in order to check for consistency in presence of other more leptokurtic factors commonly used by researchers. Therefore, we evaluate the two methods by estimating and testing the single-factor model (CAPM), the Fama and French [30, 31] three factor model, and the Carhart [18] four factor model (based on the findings of Jegadeesh and Titman [59]). We denote the factors as the excess market return (RMRF), size (SMB), value

(HML) and momentum (UMD).[19]

In order to generate the excess returns from equation (4.2.3) we first need the $N \times K$ matrix β, capturing the sensitivity of returns to the factor(s). This β matrix, (equation 4.2.2), represents the slope coefficients in the OLS regressions of each N-test portfolio and K-factor model. We use three values of N to generate β, these are the value weighted returns of the 10 size-sorted portfolios, the 25 Fama-French portfolios (the intersections of the 5 size and 5 book-to-market portfolios) and the 30 industry portfolios. As Lewellen, Nagel and Shanken [72] suggest, the traditional tests portfolios used in empirical work such as the size and 25 size/value sorted portfolios frequently present a strong factor structure, hence it seems reasonable to adopt other criteria (industry) for sorting.

In summary, we have $K = 1, 3, 4$ and $N = 10, 25, 30$; their combinations give rise to nine β matrices, allowing us to add another criteria for evaluating the method's performance, in this case measured by the efficiency. Finally, the covariance matrix $\mathrm{E}[\epsilon_t \epsilon_t' \mid f_t]$ in equation (4.2.3), is set equal to the sample covariance matrix of the residuals obtained in the N OLS regressions.

In Table 4.1 we report the descriptive statistics of historical observations of factors and test portfolios, these values are used to calibrate the Monte Carlos. As can be seen from the four moments shown, the factors associated with the multi-factor models are quite different from the excess market return factor, in particular the momentum factor is almost three times more leptokurtic than the excess market return. Thus, it is important to consider an alternative to the multivariate normal distribution which captures properties more consistent with the data such as excess kurtosis. Similar

[19] See [30], for a complete description of the Fama-French factors.

Table 4.1: Sample statistics of the factors and portfolios

This table reports sample mean, standard deviation, skewness and kurtosis of the excess return (RMRF), small minus big (SMB), high minus low (HML) and up minus down (UMD) factors; and 10 size, 25 Fama-French and 30 industry test portfolios. The sample statistics are obtained using 978 monthly observations over the period July 1926 to December 2007 from the Kenneth French library.

	Factors K				Portfolios N		
	RMRF	SMB	HML	UMD	10 size	25 FF	30 Industry
Mean μ	0.65	0.23	0.41	0.76	0.883	0.847	0.740
Std. Dev.	5.41	3.35	3.58	4.65	0.163	0.328	0.125
Skewness	0.22	2.22	1.90	−3.04	0.103	−1.359	0.099
Kurtosis	10.97	25.28	18.96	31.66	2.932	6.331	2.295

studies such as Kan and Zhou [67] consider the Student-t distribution however the magnitude of kurtosis is still limited for a t-distribution with a finite fourth moment.[20] In previous simulations not showed here, a Student-t distribution with five degrees of freedom implies a kurtosis of 6 for the RMRF factor, which is still much lower than the empirical value of 11. Therefore, we consider the empirical distribution as the alternative to the multivariate normal.

Figure 4.1 illustrates the difference between the simulated distributions, comparing the cumulative distribution function of 1000 random observations from the multivariate normal and empirical distributions. As expected, the cumulative distribution function of the sample data is approximately identical to the empirical distribution. Hence, by simulating from the empirical distribution, we closely replicate the non-normalities of the factors and portfolios described in Table 4.1.

[20]The asymptotic distribution theory for the GMM requires that returns and factors have finite fourth moments. Hence, there must be more than four degrees of freedom.

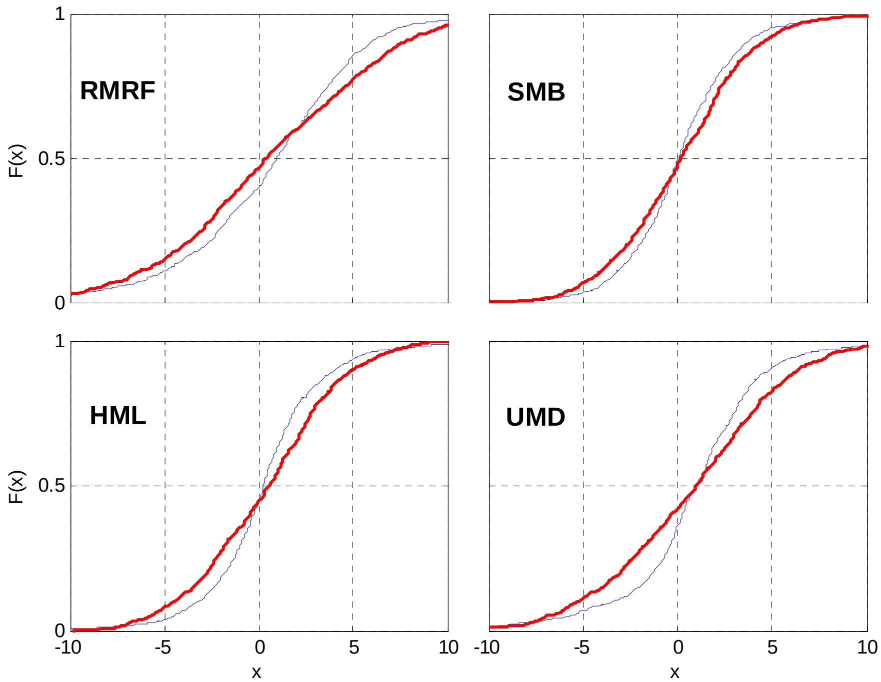

Figure 4.1: Factors CDF plots from empirical distribution in blue (thin line), and from a multivariate normal in red (thick line).

4.3 Results

We fist show the results on the risk premium estimate efficiency in section 4.3.1, and finally the results on pricing errors efficiency , size and power tests in section 4.3.2.

Table 4.2: Relative standard errors of the estimated market risk premium

Independent samples $\{(f_t \epsilon_t')\}_{t=1,...,T}$ are drawn from the empirical distribution to obtain the simulated standard errors. Excess returns on 10 size sorted portfolios are then constructed to satisfy $r_t = f_t\beta + \epsilon_t$ for $t = 1, ..., T$. In each approach, the estimator λ is calculated based on T samples. We repeat this independently to obtain 10,000 draws of the estimator of λ. The simulated standard error is the standard deviation of the random draws of the estimator.

T	$\frac{\sigma\left(\widehat{\lambda}_1^A\right)}{\sigma(\lambda^*)}$	$\frac{\sigma\left(\widehat{\lambda}_2^A\right)}{\sigma(\lambda^*)}$	$\frac{\sigma\left(\widehat{\lambda}_1^B\right)}{\sigma(\lambda^*)}$	$\frac{\sigma\left(\widehat{\lambda}_2^B\right)}{\sigma(\lambda^*)}$	Average
Panel A: CAPM					
60	1.086	1.333	1.186	1.311	1.23
360	1.073	1.052	1.126	1.084	1.08
600	1.081	1.027	1.135	1.067	1.08
1000	1.089	1.017	1.142	1.071	1.08
Panel B: Fama-French					
60	1.056	1.197	1.172	1.246	1.17
360	1.098	1.039	1.147	1.068	1.09
600	1.111	1.049	1.148	1.086	1.10
1000	1.075	1.015	1.106	1.045	1.06
Panel C: Carhart					
60	1.192	1.239	1.432	1.421	1.32
360	1.386	1.339	1.660	1.584	1.49
600	1.450	1.312	1.725	1.562	1.51
1000	1.475	1.344	1.786	1.623	1.56
Average					
	1.18	1.16	1.31	1.26	1.23

4.3.1 Parameter efficiency

Market risk premium

The simulation results for the standard errors of the estimated market risk premium $\sigma(\lambda_{\text{RMRF}})$ are reported in Table 4.2. This table correspond to the 10 size-sorted portfolios. In the appendix, in Tables 4.6 and 4.7, we show the results for the 25 size-value and 30 industry sorted portfolios.

For each estimator of λ_{RMRF} the tables gives the standard deviation of the 10,000 estimated risk premium parameters relative to those obtained from the Beta method, that is $\frac{\sigma(\widehat{\lambda})}{\sigma(\lambda^*)}$. These ratios will facilitate the comparison between methods, models, test portfolios and time lengths.

According to the results on Tables 4.2, 4.6 and 4.7 the Beta method is slightly more efficient than the SDF method, since all values are marginally greater than one. Therefore, there appears to be no significant gain in efficiency when we estimate the parameters by the Beta or the SDF method at least in the case of the market risk premium. This is especially true for both the single and three-factor models (Panels A and B) since the average ratio is no greater than 1.11 for T=1000.

However, the Carhart model (Panel C in Tables 4.2, 4.6 and 4.7) shows a greater difference between the two methods' efficiency. For example, we report an average ratio of 1.57 for T=1000 in Table 4.6. This result suggest that adding a fourth factor such as momentum, tend to distort the estimation of the market risk premium in the Carhart model.

Although we only present the case of the simulations drawn from the empirical distribution, the results using the multivariate normal are qualitatively similar.

Furthermore, they are also consistent across other test portfolios such as the 25 Fama-French and the 30 industry sorted portfolios.[21]

Note also that the efficiency of the method improves as we increase the sample size T for the CAPM model, to the point that the ratio of $\sigma(\widehat{\lambda}_{\mathrm{RMRF}})$ to $\sigma\left(\lambda^*_{\mathrm{RMRF}}\right)$ is close to unity at $T = 1000$. Nevertheless, the multi-factor models may exhibit greater variance as we increase the sample size (see Panels B and C in Tables 4.2, 4.6 and 4.7). This is because it is more likely to have greater dispersion in simulated data when T is sufficiently long.[22]

While Table 4.2 shows that the efficiency of both methods is about the same when estimating the market risk premium even in small samples, we can see that the standard deviation of the second-stage SDF^A is actually the closest to the standard deviation of the Beta method. On the other hand, the more dissimilar standard deviation is with respect to the first-stage SDF^B. By construction, the efficiency of the second-stage GMM estimator is greater than the first-stage for both SDF specifications (see equations 4.2.15 and 4.2.17), however the SDF^A is slightly more efficient than the SDF^B method.[23]

Our results reported in Table 4.2, Panel A are comparable to the results of Jagannathan and Wang [58] and Cochrane [25] for the CAPM with the 10 size-sorted portfolios. Basically, they conclude there are no differences in the standard errors of the estimated λ whichever method is adopted. Our results strongly support this previous finding, showing that there is no substantial efficiency gain from the choice

[21] The results of adopting the multivariate normal distribution, and the actual values of $\sigma\left(\widehat{\lambda}\right)$ and $\sigma\left(\lambda^*\right)$ are available upon request.

[22] The relative standard errors of the estimated pricing errors do diminish as we increase T.

[23] Shanken and Zhou [94] also find that the standard errors of the second-stage estimators are consistently smaller than the standard errors of the fist-stage estimators. Although their study is based on Beta models.

of the method when we estimate the CAPM model to estimate λ_{RMRF}. In particular, the average $\frac{\sigma(\widehat{\lambda})}{\sigma(\lambda^*)}$ ratio is 1.08. Our contribution to this debate is that this is not always the case, as long as we introduce other models in our analysis the differences between the methods emerge.

Even in small samples, the difference is not large, and as we increase T, the magnitude of the standard deviation is almost the same especially for the CAPM and Fama-French models. While the expected returns of the 25 Fama-French test portfolios have higher dispersion than the 10 size-sorted and the 30 industry portfolios (see Table 4.1), these changes in the distribution of N do not alter our conclusions regarding the efficiency of the estimators in either method.

In sum, we do not find significant differences except for the case of the Carhart model (Panel C in Tables 4.2, 4.6 and 4.7). It is interesting to note that the parameter efficiency is apparently more sensible to the model, and number of factors included, rather than to time lengths and portfolio formations.

Size, value and momentum risk premium

Our key result is that the methods are no longer equivalent in terms of estimator efficiency when we compare the rest of the estimators in the multi-factor models. Table 4.3 shows the results for λ_{SMB}, λ_{HML} (estimated in the Fama-French and Carhart models) and λ_{UMD} (estimated in the Carhart model). This table correspond to the 10 size-sorted portfolios. In the appendix, in Tables 4.8 and 4.9, we show the results for the 25 size-value and 30 industry sorted portfolios respectively.

In this case, it is evident that the Beta method is more efficient than the SDF

Table 4.3: Relative standard errors of the estimated size, value and momentum risk premiums

Independent samples $\{(f_t \epsilon_t')\}_{t=1,...,T}$ are drawn from the empirical distribution to obtain the simulated standard errors. Excess returns on 10 size sorted portfolios are then constructed to satisfy $r_t = f_t\beta + \epsilon_t$ for $t = 1, ..., T$. In each approach, the estimator λ is calculated based on T samples. We repeat this independently to obtain 10,000 draws of the estimator of λ. The simulated standard error is the standard deviation of the random draws of the estimator.

	Fama-French					Carhart				
T	$\frac{\sigma(\hat{\lambda}_1^A)}{\sigma(\lambda^*)}$	$\frac{\sigma(\hat{\lambda}_2^A)}{\sigma(\lambda^*)}$	$\frac{\sigma(\hat{\lambda}_1^B)}{\sigma(\lambda^*)}$	$\frac{\sigma(\hat{\lambda}_2^B)}{\sigma(\lambda^*)}$	Average	$\frac{\sigma(\hat{\lambda}_1^A)}{\sigma(\lambda^*)}$	$\frac{\sigma(\hat{\lambda}_2^A)}{\sigma(\lambda^*)}$	$\frac{\sigma(\hat{\lambda}_1^B)}{\sigma(\lambda^*)}$	$\frac{\sigma(\hat{\lambda}_2^B)}{\sigma(\lambda^*)}$	Average
Panel A: λ_{SMB}										
60	1.250	1.362	1.334	1.396	1.34	1.320	1.360	1.474	1.461	1.40
360	1.403	1.323	1.447	1.347	1.38	1.366	1.291	1.472	1.385	1.38
600	1.382	1.284	1.422	1.308	1.35	1.330	1.269	1.446	1.369	1.35
1000	1.333	1.232	1.373	1.262	1.30	1.309	1.226	1.433	1.340	1.33
Panel B: λ_{HML}										
60	2.591	2.263	2.755	2.346	2.49	3.034	2.868	3.368	3.193	3.12
360	2.922	2.264	3.038	2.329	2.64	3.720	3.406	4.233	3.866	3.81
600	2.939	2.267	3.051	2.336	2.65	3.537	3.327	4.126	3.873	3.72
1000	2.933	2.211	3.044	2.277	2.62	3.644	3.377	4.222	3.911	3.79
Panel C: λ_{UMD}										
60	-	-	-	-	-	3.698	3.529	4.345	4.095	3.92
360	-	-	-	-	-	5.697	5.406	6.729	6.333	6.04
600	-	-	-	-	-	5.564	5.205	6.602	6.205	5.89
1000	-	-	-	-	-	6.017	5.517	7.069	6.465	6.27

method at estimating λ, especially in the case of λ_{HML}, and λ_{UMD}. The main implication of our finding is that inference on λ will be in general more accurate if one follows the Beta method than if one follows the SDF method. According to Table 4.3, the value risk premium standard error is 2.6 times bigger when following the SDF method, this may represent the difference between rejecting or not rejecting an hypothesis test. The case of the momentum risk premium is even more evident, the ratio can reach 8.2 for T=1000 (see Table 4.8).

These reported differences considerably accentuates in the case of multi-factor models, therefore presumably the number of factors are relevant at deliver efficient estimators.

It is interesting to highlight the performance of the second-stage SDF estimators since they are, by construction, more efficient than the first-stage: $\sigma(\widehat{\lambda}_1) > \sigma(\widehat{\lambda}_2)$. Our empirical results support this argument, however, they are still far from the efficiency of the OLS Beta estimators: $\sigma(\widehat{\lambda}_1) > \sigma(\widehat{\lambda}_2) > \sigma(\lambda^*)$.

The case of λ_{UMD} is special since the momentum factor has the highest kurtosis relative to the factor mean of all factors examined, see Figure 4.1. Even when pricing the 10 size sorted portfolios – which have lower expected returns variance – and taking the longest sample $T = 1000$, the difference between the standard deviation of $\widehat{\lambda}_{\text{UMD}}$ can be seven times as big as the standard deviation of λ^*_{UMD}. Hence it appears that Table 4.3 suggests the SDF method consistently delivers more inefficient estimators than the Beta method as the factor becomes more non-normal. However, as we explain below, this is not the only and main reason.

Why is the Beta method more efficient?

We argue that there are at least three feasible reasons that explain the differences between the efficiency of the SDF and Beta methods at estimating λ.

1. Factors' non-normalities. The SDF method is not only exposed to the first moments of the returns and factors as in the case of the Beta method, but also to the higher order moments. In particular, the SDF^A estimates depend on $d = \mathrm{E}[r_t f_t]$, and the SDF^B on $b = \mathrm{Cov}[r_t f_t]$, see equations (4.2.15) and (4.2.17). On the other hand, the Beta estimates, δ^*, depend on the first moment of the factor $\delta = \mu = \mathrm{E}[f_t]$, since the asset pricing models we study use traded factors; additionally, δ^* captures the sampling variation of the second moment of the factor, Σ_f, when it is transformed from δ^* to λ^* by equation (4.2.18). Then, as our evidence illustrates in Table 4.3, the more non-normal the factors included in the model the less efficient the estimators in the SDF relative to the Beta method.

2. Numbers of factors. As we show in Table 4.2, the estimation of the single-factor model does not reflect a significant difference between the methods, while Table 4.3 shows that when we evaluate multi-factor models the Beta method clearly outperforms the SDF method. Even though this is one original result, other authors find similar evidence when comparing Beta models such as Shanken and Zhou [94], and Hou and Kimmel [52].[24]

In order to test the influence of factor non-normality and the number of factors, we conduct a further simulation experiment in which single and multi-factor models are loaded with new artificial series, calibrated from either low or high non-normal

[24] They examine theoretical and econometric issues in the estimation of risk pemia in a linear factor model when the model is misspecified. They show that, for a given set of test assets, the risk premium of an unspanned factor is very sensitive to the choice of other factors in the model.

factors. To make our new setup comparable to our previous analysis, we calibrate the low and high non-normal series by considering the historical distribution of the market and momentum factors respectively.[25]

The results indicate the number of factors is in fact more important than the degree of non-normality of the factor in explaining the differences between the efficiency of the Beta over the SDF method. In particular, whether we load a three or four factor model with low or high non-normal factors, we still get consistently more efficient estimators λ in the Beta than in the SDF method. Further, in the case of the single-factor model, no differences emerge independently of the factor's degree of normality.

Therefore, we should expect that SDF method will deliver risk premium estimates with higher variance in models which include more factors such as APT. One practical recommendation is to follow the Beta method at estimating APT risk premiums.

3. GMM Moment restrictions. One may think that if we include the definition of λ, that is $\lambda = \frac{\mu}{\mu^2+\Sigma_f}$ such that $\mathrm{E}[(\mu^2 + \Sigma_f)\,\lambda - \mu] = 0$ as an additional GMM moment restriction in the SDF method, as is usually the case in the Beta method, the puzzle regarding the discrepancy in efficiency of λ^* and $\widehat{\lambda}$ will disappear. This seems reasonable because originally the GMM moment conditions in the SDF method are the definition of the pricing errors, and therefore the efficiency of $\widehat{\lambda}$ may improve if we

[25] An alternative procedure would be to conduct a Box-Cox transformation to the actual factors in order to see whether the main results change once normalized. However, this would require a previous monotonic transformation to get rid of values less than or equal to zero. The historical mean of factors $\mathrm{E}[f]$ is 0.64 and 0.76 for the market and momentum factor, and the minimum value $\min[f]$ is -29.04 and -50.63. Hence, such monotonic transformation will be changing the fundamental relation between the benchmark portfolios and the factors, and the magnitude of the estimates becomes meaningless as well as their variance. Our proposed procedure is free of this problem.

include its definition as an additional restriction. In a Monte Carlo simulation analysis not reported here, we find that the variance of the SDF estimator $\widehat{\lambda}$ diminishes with inclusion of additional moment restrictions. However, the observed decrease is not sufficient to change our main conclusion, not even in the case of the second-stage SDF estimators.[26]

Our key result in Table 4.3 makes an important contribution to the empirical asset pricing literature. Jagannathan and Wang [58], Ferson [35] and Cochrane [25] emphasize that the Beta and SDF methods lead to almost identical results in terms of efficiency. While this holds for the single-factor model we show it does not apply to multi-factor models. Once we include other factors with greater non-normalities, the differences clearly emerge. Kan and Zhou [66] also indicate that estimation in the SDF method is significantly affected by the presence of skewness and kurtosis in factors.

Jagannathan and Wang [58] argues that not only the finite sample efficiency between the methods is the same, they demonstrate that for the case of the single factor model, the SDF method is asymptotically as efficient as the Beta method. However, our finite sample analysis show that this is not the case for the multi-factor models. This original result can be compared to those on Chen and Kan [22], who find that the finite sample distributions of the estimated risk premia differ significantly from their asymptotic distributions when testing two-pass cross-sectional regressions in a Beta formulation.

Table 4.4: Relative standard errors of the estimated pricing errors

Independent samples $\{(f_t\epsilon_t')\}_{t=1,...,T}$ are drawn from the empirical distribution to obtain the simulated standard errors. Excess returns on 10 size sorted portfolios are then constructed to satisfy $r_t = f_t\beta + \epsilon_t$ for $t = 1, ..., T$. In each approach, the vector of sample pricing errors π is calculated based on T samples. We repeat this independently to obtain 10,000 draws of π. The simulated standard error is the standard deviation of the random draws of the sample pricing errors.

T	$\frac{\sigma(\widehat{\pi}_1^A)}{\sigma(\pi^*)}$	$\frac{\sigma(\widehat{\pi}_2^A)}{\sigma(\pi^*)}$	$\frac{\sigma(\widehat{\pi}_1^B)}{\sigma(\pi^*)}$	$\frac{\sigma(\widehat{\pi}_2^B)}{\sigma(\pi^*)}$	Average
Panel A: CAPM					
60	0.545	1.417	0.562	1.137	0.92
360	0.543	1.020	0.552	1.003	0.78
600	0.542	1.010	0.551	1.007	0.78
1000	0.543	1.000	0.551	1.005	0.78
Panel B: Fama-French					
60	0.314	1.964	0.338	1.657	1.07
360	0.334	0.743	0.345	0.629	0.51
600	0.318	0.637	0.327	0.568	0.46
1000	0.308	0.541	0.315	0.519	0.42
Panel C: Carhart					
60	0.166	0.937	0.182	0.828	0.53
360	0.134	0.316	0.146	0.307	0.23
600	0.143	0.277	0.153	0.275	0.21
1000	0.133	0.235	0.145	0.241	0.19
Average					
	0.34	0.84	0.35	0.76	0.57

4.3.2 Pricing error efficiency and specification tests

In this subsection, we evaluate the model misspecification by examining the sample pricing errors. Our calculations are based under the null hypothesis that the asset pricing model, equation (4.2.1), holds. Contrary to the previous section, in which the Beta method is preferred because it yields more efficient estimators, here we show that the SDF method outperforms the Beta method in achieving more efficient pricing errors π, as measured by $\sigma(\pi)$. The moment conditions in the Beta method, equation (4.2.5), and SDF method, equations (4.2.13) and (4.2.16), are $N + NK + K$ and N respectively. To examine the pricing errors π, we take the first N restrictions of the Beta method and transform α^* into π^* by equation (4.2.19). Hence, we compare the standard deviations of π^* and $\widehat{\pi}$ as we did with λ^* and $\widehat{\lambda}$. Analogously, in Table 4.4 we report the comparisons between $\sigma(\pi^*)$ and $\sigma(\widehat{\pi})$ in the same format as before.

The first and most important distinguishing feature is that the standard deviation of the pricing errors using the SDF method is in general smaller than using the Beta method in most of the examined cases.[27] In particular, the difference is greater with respect to the first-stage SDF pricing errors rather than the second-stage pricing errors. This is to be expected as, in general, the first-stage aims to minimize the pricing errors π while the second-stage weights according to the statistically most

[26]Significantly, researchers hardly ever impose this moment restriction when estimating asset pricing models by the SDF method.

[27]This is highly relevant in efficiency tests studies. As pointed out by Ferson and Siegel [39], testing the efficiency of a given portfolio has long been an important topic in empirical asset pricing. They concisely support this argument as follow: The CAPM of Sharpe [95] implies that a market portfolio should be mean variance efficient. Multiple-beta asset pricing models such as Merton [83] imply that a combination of the factor portfolios is minimum variance efficient (Chamberlain [19]; Grinblatt and Titman [43]). The consumption CAPM implies that a maximum correlation portfolio for consumption is efficient (see Breeden [11]). More generally, any stochastic discount factor model implies that a maximum correlation portfolio for the stochastic discount factor is minimum variance efficient (see Hansen and Richard [50]). Classical efficiency tests are studied by Gibbons [40], Jobson and Korkie [61], Stambaugh [98], MacKinlay [78], Gibbons, Ross and Shanken [41] and others.

informative portfolios, for a more complete discussion see [25, section 12.2].

A noisy SDF parameter $\widehat{\lambda}$ does not necessarily imply a noisy SDF pricing error $\widehat{\pi}$. The N moment conditions of the SDF method coincide with the definition of the pricing errors, equation (4.2.14), hence the GMM delivers $\widehat{\lambda}$ such that it minimizes the expected value of π. On the other hand, the moment conditions in the Beta method include not only the N definitions of the Jensen's alphas, but include the other $NK + K$ restrictions. Thus, since the Beta method has additional restrictions, unrelated to the minimization of the pricing errors, it is anticipated to have lower efficiency, i.e. $\sigma(\pi^*) > \sigma(\widehat{\pi})$.

Table 4.4 indicates some evidence in favor of multi-factor models, since the variance of the pricing errors diminishes as we increase the number of factors (i.e. move from Panel A to B and C). This result is consistent with recent work, see for example Shanken and Zhou [94]. There are no notable differences between the performance of un-meaned SDF^A and de-meaned SDF^B, this is consistent with Farnsworth, Ferson, Jackson and Todd [33], who find that measures of performance are not highly sensitive to the SDF representation. Our results indicate that the first-stage results (SDF^A and SDF^B) are quite similar, whereas in the second-stage, the SDF^B method generally performs better than the SDF^A. Nevertheless, in comparing the pricing errors, and as pointed out by Kan and Robotti [64], the SDF^B specification is more appropriate than SDF^A for model comparison.

It is interesting to note the effect of the variance of the central parameter λ and pricing errors π as we increase the size of the time-series T. In the case of the CAPM, the sample range $360 < T < 1000$ does not impact the variance of $\sigma(\lambda)$ and $\sigma(\pi)$. However for both the Fama-French and Carhart models, $\sigma(\pi)$ shows a decrease as T

increases in the range $360 < T < 1000$ while there is no clear pattern in the relation between T and $\sigma(\lambda)$ for the multi-factor models.

Size and power

To examine the test size in the two methods, we use the Monte Carlo simulations to compute the rejection rates under the null hypothesis that the model holds. We report the test size for three significant levels: 1 percent, 5 percent and 10 percent. To estimate the tails of the sampling distribution of the J-statistics, we perform the Monte Carlos with 10,000 simulations.[28]

To examine the power in the two methods, we perform identical Monte Carlo simulations to compute the rejection rates, but now allow for the possibility of deviations from the model. In other words, we study the power under the null that the model does not hold. There are many ways in which the expected return restriction could be violated. In our case, consistent with the extant literature, we add a nonzero Jensen's alpha to the model for generating excess returns, causing the asset pricing model, equation (4.2.3), to be misspecified.

In Table 4.5 we present the results for the size and power test, for brevity, we show only the results from the 10 size-sorted portfolios for the CAPM and in Tables 4.10 and 4.11 the corresponding values for the Fama-French and Carhart models respectively, the test results with the 25 size/value and 30 industry portfolios are available upon request. Contrary to the case of the analysis of the risk premium

[28] To examine the standard deviation of λ and π, there is no significant difference whether one performs 1,000 or 10,000 simulations since one takes the standard deviation of the results. However, when examining the tail of the J-statistic distribution, it is preferable to perform 10,000 instead of 1,000 simulations. This approach is also adopted in Shanken and Zhou [94] and Kan and Zhou [66] among others.

Table 4.5: CAPM Specification tests on 10 size portfolios

	1%					5%					10%				
T	λ^*	$\widehat{\lambda}_1^A$	$\widehat{\lambda}_2^A$	$\widehat{\lambda}_1^B$	$\widehat{\lambda}_2^B$	λ^*	$\widehat{\lambda}_1^A$	$\widehat{\lambda}_2^A$	$\widehat{\lambda}_1^B$	$\widehat{\lambda}_2^B$	λ^*	$\widehat{\lambda}_1^A$	$\widehat{\lambda}_2^A$	$\widehat{\lambda}_1^B$	$\widehat{\lambda}_2^B$
Panel A: Size test $W = S$															
60	3.62	5.43		6.22		9.89	14.45		15.50		16.01	22.55		23.86	
360	0.68	1.48		1.50		3.86	6.26		6.28		7.55	11.90		12.00	
600	0.64	1.09		1.12		3.17	5.62		5.63		6.94	11.04		11.09	
1000	0.68	1.22		1.22		3.41	5.74		5.78		7.09	10.82		10.85	
Panel B: Size test $W = \mathrm{E}[r_t r_t']$															
60	0.11	0.25	0.72	0.54	1.16	1.78	3.30	4.72	4.75	6.48	4.72	8.06	9.58	10.28	12.32
360	0.44	0.73	0.74	0.95	0.99	2.84	4.52	4.60	5.26	5.41	6.19	9.25	9.47	10.58	10.74
600	0.47	0.80	0.81	0.93	0.94	2.66	4.35	4.41	5.22	5.30	6.10	9.36	9.46	10.60	10.69
1000	0.61	0.97	0.99	1.20	1.22	3.10	4.85	4.89	5.65	5.67	6.56	9.85	9.87	10.95	11.01
Panel C: Size test $W = \mathrm{Cov}[r_t]$															
60	3.62	4.85	5.23	6.99	7.36	9.89	13.20	13.71	16.55	17.14	16.01	20.86	21.60	25.37	26.15
360	0.68	1.25	1.25	1.71	1.72	3.86	5.83	5.89	6.75	6.81	7.55	10.97	11.05	12.58	12.66
600	0.64	1.00	1.00	1.20	1.21	3.17	5.15	5.20	6.09	6.11	6.94	10.43	10.46	11.77	11.88
1000	0.68	1.16	1.16	1.38	1.38	3.41	5.36	5.38	6.11	6.14	7.09	10.42	10.45	11.48	11.51
Panel D: Power test $W = S$															
60	3.86	5.48		6.08		10.64	15.56		16.71		17.25	24.39		25.18	
360	2.21	3.97		3.93		8.07	12.18		12.13		14.42	20.56		20.33	
600	3.33	5.44		5.36		11.51	16.76		16.59		19.07	25.96		25.87	
1000	6.24	9.70		9.61		18.43	24.62		24.56		27.93	36.10		35.87	
Panel E: Power test $W = \mathrm{E}[r_t r_t']$															
60	0.19	0.47	0.98	0.76	1.60	1.92	3.60	4.96	4.98	7.02	4.87	8.31	10.65	11.15	13.60
360	1.35	2.27	2.43	2.89	3.05	6.43	9.54	9.75	11.02	11.22	12.05	16.91	17.24	18.84	19.19
600	2.49	4.18	4.32	4.92	5.04	10.10	14.25	14.41	15.91	16.16	17.43	23.01	23.16	25.29	25.49
1000	5.49	8.24	8.32	9.40	9.45	17.31	22.53	22.64	24.64	24.77	26.67	33.66	33.79	36.12	36.20
Panel F: Power test $W = \mathrm{Cov}[r_t]$															
60	3.86	5.06	5.38	7.06	7.64	10.64	14.14	14.89	18.12	19.05	17.25	22.25	23.10	27.08	27.88
360	2.21	3.50	3.63	4.31	4.42	8.07	11.48	11.57	13.16	13.24	14.42	19.43	19.57	21.69	21.87
600	3.33	5.20	5.22	5.96	6.00	11.51	15.79	15.86	17.69	17.76	19.07	24.89	24.98	27.16	27.28
1000	6.24	9.25	9.28	10.35	10.37	18.43	23.79	23.88	25.77	25.83	27.93	35.00	35.09	37.36	37.43

efficiency, the specifications tests do not yield significant differences when comparing the single and multi-factor cases. Therefore, in general, the conclusions reached by Table 4.5 can be extended to those in Tables 4.10 and 4.11.[29]

The size and power tests represent rejection rates of the J-statistic. Nevertheless, it is well known that this statistic can be computed in many ways, for instance, depending on the choice of the weighting matrix used in the aggregation of the pricing errors. We use three alternatives to give a more robust idea about the performance of the methods. In particular we use the spectral density matrix S, the second moment matrix $\mathrm{E}[r_t r_t']$, and the covariance matrix of returns $\mathrm{Cov}[r_t]$.

To add the pricing error vector $g_T(\theta)$, the J-statistic, equation (4.2.11), is weighted by the covariance of the pricing errors, equation (4.2.10) in the first-stage, which is simultaneously weighted by the spectral density matrix S, equation (4.2.9). Therefore, both the first and second-stage J-statistics are weighted by S in order to aggregate the pricing errors vector $g_T(\theta)$. As we know from Hansen [46], this choice is statistically optimal in the sense that it maximizes the asymptotic statistical information in the sample about a model, given the choice of moments. However, since the S matrix changes across models, it is not convenient to use it for model comparison. In particular, we could not claim a better fit because of a smaller J-statistic since we have different values of S across models.

Hansen and Jagannathan [49] suggest the use of the second moment matrix of excess returns $W = \mathrm{E}[r_t r_t']$ instead of $W = S$. This alternative is more suitable for model comparison because it is invariant to the model and their parameters. It

[29] We find that most tests have better finite sample performances for a smaller number of assets, this is consistent with related works for Beta pricing model comparisons such as in Li, Xu and Zhang [73].

provides an economic measure of the model fit instead of a statistical measure and has the property of being invariant to portfolio formation. For our third alternative we follow Burnside [14], Balduzzi and Yao [4] and Kan and Robotti [64] who suggest that the de-meaned SDF^B method should use the covariance matrix of excess returns W=Cov$[r_t]$.

Since we perform size and power tests for the Beta and SDF methods at three significance levels, using three alternative weighting matrices for the calculation of the J-statistic, three models, three benchmark portfolios, and four time-series sizes, for the sake of brevity, we only report a representative sample of the whole tests.

Tables 4.5, 4.10 and 4.11, illustrate that the Beta method consistently under-rejects in finite-samples greater than $T = 60$. In particular, at $T = 1000$ the size is around 30% below the theoretical value for the CAPM measuring between (Table 4.5) 0.61 – 0.68 at 1%; 3.1 – 3.41 at 5% and 6.56 – 7.09 at 10%. The level of under-rejection is considerably more for the multi-factor models where in some cases the size is only about 1% of the theoretical value. The level of under-rejection is slightly greater in Panel B than in Panels A and C. Interestingly in the case of the Beta method, the J-statistic leads to the same size and power results whether we adopt $W = S$ (Panels A and D) or W =Cov$[r_t]$ (Panels D and F). This equality does not hold for the SDF^A and SDF^B specifications.

In contrast to performance of the Beta method, the size of the SDF method is much closer to the theoretical values regardless of the model. In fact, this is partially explained in Table 4.4 since a more efficient pricing error should lead to better specification tests in general. With respect to the differences in size between the two SDF specifications, we see that there is a marginal increase in size for SDF^B

with respect to SDF^A. However, this should not be considered as a real difference, since the misspecification of SDF^B is lessened by the substraction of the mean of the factor (see equations 4.2.13 and 4.2.16). In general, a less misspecified model will lead to better specification test results. Consequently, we argue that our results do not favor a particular SDF specification since the differences in the specification tests can be explained by the de-meaned SDF being not as misspecified as the un-meaned specification by construction. On the other hand, the differences of the risk premium efficiency in section 4.3.1 have real implications since they are conducted assuming that the model is well specified. Even though the results are sensible to the choice of the weighting matrix, Tables 4.5, 4.10 and 4.11 show that the main changes in the specification tests are due to the method. Nevertheless, it is important to bear in mind the theoretical implications of taking one matrix or another.

A similar pattern is found in the power tests, where the SDF has greater capacity to identify a misspecified model than the Beta method, regardless of the weighting matrix and the number of factors. Our results are comparable to those of Burnside [14] since we find the GMM tests have better power under the SDF^B specification, equation (4.2.16), than under the SDF^A specification, equation (4.2.13). But as we state earlier, this difference is due to the effect of subtracting the mean of the factor on the model misspecification.

Thus, while there is an advantage of the SDF method over the Beta method in terms of achieving more efficient pricing errors. The main implication of this finding is that if we are interested in a good model fit we should prefer the SDF method, however this comes at the cost of getting more inefficient risk premium estimates. At the end, both methods are clearly complementary (not equivalent) and the choice is subject to the purpose of the empirical experiment.

4.4 Conclusions

The extant literature demonstrates that the Beta method and SDF method are equally as efficient in terms of the estimation of risk premiums. We examine whether this equality holds for multi-factor asset pricing models. Specifically we investigate both the Fama-French three factor model and the Carhart four factor model in addition to the single-factor CAPM. We find results consistent with the previous literature for the CAPM. However, in the context of multi-factor models, we find that relative to the SDF method the Beta method is in general more efficient at estimating risk premiums. This relative advantage of the Beta method at estimating the risk premiums does not apply to the estimation of the sample pricing errors; however, where invariably the efficiency of the SDF method is superior.

We consider this is a remarkable finding since there are numerous examples in which researchers refer to works such as Jagannathan and Wang [58] which argue that both methodologies are similar in terms of efficiency. For example, see Wang and Zhang [102], Jagannathan, Skoulakis and Wang [55], Vassalou, Li and Xing [100], Cochrane [23, 25], Smith and Wickens [97], Nieto and Rodríguez [85], Balvers and Huang [5], Brandt and Chapman [10], Cai and Hong [15], and Ferson [35], just to mention a few. Our results suggest that this similarity between the two methods only holds under very specific situations.

Previous chapter show some evidence of this result, however the objectives of chapter 3 and 4 are different. Here, we properly demonstrate the difference of the methods in terms of efficiency in simulated data which allow a finite sample analysis, while in chapter 3 we rely on analysis based on several historical data sets.

Commonly used factors and returns in empirical studies habitually exhibit high kurtosis, and commonly tested models are multi-factor, therefore our results suggest that if we are interested in performing inference on risk premiums, we should prefer the Beta method over the SDF method. Conversely, if we are interested in inference based on the sampling pricing error, the SDF method should be preferred. Hence, there is no method that fully dominates the other, rather, they are complementary and, further, they should not be considered as empirically equivalent.

This work intend to contribute to the understanding of the finite sample properties of the Beta and SDF methods by showing evidence about the magnitude, direction, and parameters which determines the parameters bias.

Further extensions to this work could be to explore what happen when considering non-traded factors. In principle, there is no reason to expect a similar pattern. For example, Kan and Robotti [63] show that the standard errors under correctly specified and potentially misspecified models are similar for traded factors, while they can differ substantially for non-traded factors such as the scaled market return and the lagged state variable CAY.

■

4.5 Appendix

Table 4.6: Relative standard errors of the estimated market risk premium, 25 Fama-French portfolios

Independent samples $\{(f_t\epsilon_t')\}_{t=1,\dots,T}$ are drawn from the empirical distribution to obtain the simulated standard errors. Excess returns on 25 size-value sorted portfolios are then constructed to satisfy $r_t = f_t\beta + \epsilon_t$ for $t = 1, ..., T$. In each approach, the estimator λ is calculated based on T samples. We repeat this independently to obtain 10,000 draws of the estimator of λ. The simulated standard error is the standard deviation of the random draws of the estimator.

T	$\frac{\sigma(\widehat{\lambda}_1^A)}{\sigma(\lambda^*)}$	$\frac{\sigma(\widehat{\lambda}_2^A)}{\sigma(\lambda^*)}$	$\frac{\sigma(\widehat{\lambda}_1^B)}{\sigma(\lambda^*)}$	$\frac{\sigma(\widehat{\lambda}_2^B)}{\sigma(\lambda^*)}$	Average
Panel A: CAPM					
60	1.090	2.090	1.189	1.858	1.557
360	1.074	1.149	1.128	1.138	1.122
600	1.068	1.081	1.122	1.095	1.091
1000	1.070	1.053	1.123	1.088	1.083
Panel B: Fama-French					
60	0.989	1.734	1.064	1.635	1.355
360	1.019	1.094	1.057	1.094	1.066
600	1.012	1.049	1.049	1.061	1.043
1000	1.016	1.016	1.047	1.031	1.027
Panel C: Carhart					
60	1.214	1.538	1.421	1.666	1.459
360	1.571	1.229	1.867	1.362	1.507
600	1.605	1.198	1.914	1.358	1.519
1000	1.683	1.190	2.032	1.381	1.571
Average					
	1.201	1.285	1.334	1.314	1.283

Table 4.7: Relative standard errors of the estimated market risk premium, 30 industry portfolios

Independent samples $\{(f_t \epsilon_t')\}_{t=1,...,T}$ are drawn from the empirical distribution to obtain the simulated standard errors. Excess returns on 30 industry sorted portfolios are then constructed to satisfy $r_t = f_t\beta + \epsilon_t$ for $t = 1, ..., T$. In each approach, the estimator λ is calculated based on T samples. We repeat this independently to obtain 10,000 draws of the estimator of λ. The simulated standard error is the standard deviation of the random draws of the estimator.

T	$\frac{\sigma(\widehat{\lambda}_1^A)}{\sigma(\lambda^*)}$	$\frac{\sigma(\widehat{\lambda}_2^A)}{\sigma(\lambda^*)}$	$\frac{\sigma(\widehat{\lambda}_1^B)}{\sigma(\lambda^*)}$	$\frac{\sigma(\widehat{\lambda}_2^B)}{\sigma(\lambda^*)}$	Average
Panel A: CAPM					
60	1.039	2.455	1.130	2.130	1.688
360	1.032	1.168	1.084	1.158	1.111
600	1.014	1.110	1.068	1.110	1.075
1000	1.036	1.071	1.071	1.089	1.067
Panel B: Fama-French					
60	1.028	2.042	1.063	1.898	1.508
360	1.074	1.185	1.111	1.167	1.134
600	1.061	1.122	1.098	1.122	1.101
1000	1.095	1.111	1.127	1.127	1.115
Panel C: Carhart					
60	1.072	1.714	1.169	1.917	1.468
360	1.260	1.298	1.423	1.413	1.349
600	1.280	1.256	1.463	1.390	1.348
1000	1.290	1.242	1.500	1.403	1.359
Average					
	1.107	1.398	1.192	1.410	1.277

Table 4.8: Relative standard errors of the estimated size, value and momentum risk premiums, 25 Fama-French portfolios

Independent samples $\{(f_t \epsilon_t')\}_{t=1,...,T}$ are drawn from the empirical distribution to obtain the simulated standard errors. Excess returns on 25 size-value sorted portfolios are then constructed to satisfy $r_t = f_t\beta + \epsilon_t$ for $t = 1, ..., T$. In each approach, the estimator λ is calculated based on T samples. We repeat this independently to obtain 10,000 draws of the estimator of λ. The simulated standard error is the standard deviation of the random draws of the estimator.

	Fama-French					Carhart				
T	$\frac{\sigma(\widehat{\lambda}_1^A)}{\sigma(\lambda^*)}$	$\frac{\sigma(\widehat{\lambda}_2^A)}{\sigma(\lambda^*)}$	$\frac{\sigma(\widehat{\lambda}_1^B)}{\sigma(\lambda^*)}$	$\frac{\sigma(\widehat{\lambda}_2^B)}{\sigma(\lambda^*)}$	Average	$\frac{\sigma(\widehat{\lambda}_1^A)}{\sigma(\lambda^*)}$	$\frac{\sigma(\widehat{\lambda}_2^A)}{\sigma(\lambda^*)}$	$\frac{\sigma(\widehat{\lambda}_1^B)}{\sigma(\lambda^*)}$	$\frac{\sigma(\widehat{\lambda}_2^B)}{\sigma(\lambda^*)}$	Average
25Panel A: λ_{SMB}										
60	1.162	1.912	1.220	1.785	1.520	1.116	1.590	1.227	1.630	1.391
360	1.178	1.172	1.209	1.153	1.178	1.110	1.091	1.220	1.152	1.143
600	1.176	1.120	1.208	1.120	1.156	1.070	1.047	1.172	1.117	1.102
1000	1.177	1.083	1.208	1.094	1.141	1.070	1.040	1.180	1.120	1.103
Panel B: λ_{HML}										
60	1.146	1.894	1.193	1.786	1.505	1.641	1.786	1.917	2.003	1.837
360	1.146	1.181	1.194	1.181	1.175	2.340	1.483	2.748	1.687	2.065
600	1.153	1.126	1.189	1.135	1.151	2.452	1.443	2.878	1.661	2.109
1000	1.138	1.092	1.172	1.115	1.129	2.575	1.483	3.080	1.724	2.216
Panel C: λ_{UMD}										
60	-	-	-	-	-	3.564	2.645	4.092	3.234	3.384
360	-	-	-	-	-	6.190	3.240	7.060	3.640	5.033
600	-	-	-	-	-	6.773	3.360	7.720	3.853	5.427
1000	-	-	-	-	-	7.121	3.483	8.241	4.034	5.720

Table 4.9: Relative standard errors of the estimated size, value and momentum risk premiums, 30 industry portfolios

Independent samples $\{(f_t \epsilon_t')\}_{t=1,...,T}$ are drawn from the empirical distribution to obtain the simulated standard errors. Excess returns on 30 industry sorted portfolios are then constructed to satisfy $r_t = f_t \beta + \epsilon_t$ for $t = 1, ..., T$. In each approach, the estimator λ is calculated based on T samples. We repeat this independently to obtain 10,000 draws of the estimator of λ. The simulated standard error is the standard deviation of the random draws of the estimator.

	Fama-French					Carhart				
T	$\frac{\sigma(\widehat{\lambda}_1^A)}{\sigma(\lambda^*)}$	$\frac{\sigma(\widehat{\lambda}_2^A)}{\sigma(\lambda^*)}$	$\frac{\sigma(\widehat{\lambda}_1^B)}{\sigma(\lambda^*)}$	$\frac{\sigma(\widehat{\lambda}_2^B)}{\sigma(\lambda^*)}$	Average	$\frac{\sigma(\widehat{\lambda}_1^A)}{\sigma(\lambda^*)}$	$\frac{\sigma(\widehat{\lambda}_2^A)}{\sigma(\lambda^*)}$	$\frac{\sigma(\widehat{\lambda}_1^B)}{\sigma(\lambda^*)}$	$\frac{\sigma(\widehat{\lambda}_2^B)}{\sigma(\lambda^*)}$	Average
indPanel A: λ_{SMB}										
60	1.431	2.536	1.316	2.313	1.899	1.333	2.011	1.288	2.091	1.681
360	1.503	1.571	1.503	1.528	1.526	1.473	1.503	1.564	1.564	1.526
600	1.524	1.508	1.540	1.500	1.518	1.460	1.437	1.571	1.516	1.496
1000	1.531	1.469	1.551	1.469	1.505	1.469	1.408	1.592	1.500	1.492
Panel B: λ_{HML}										
60	1.437	2.440	1.362	2.260	1.875	1.526	2.124	1.562	2.430	1.911
360	1.568	1.500	1.596	1.479	1.536	1.966	1.797	2.203	1.980	1.986
600	1.563	1.438	1.607	1.446	1.513	2.018	1.779	2.292	1.982	2.018
1000	1.593	1.407	1.640	1.430	1.517	2.080	1.761	2.386	2.011	2.060
Panel C: λ_{UMD}										
60	-	-	-	-	-	2.234	2.645	2.328	3.479	2.672
360	-	-	-	-	-	3.630	3.110	4.060	3.480	3.570
600	-	-	-	-	-	3.895	3.250	4.434	3.671	3.813
1000	-	-	-	-	-	4.121	3.328	4.741	3.828	4.004

Table 4.10: Fama-French specification tests on 10 size portfolios

	1%					5%					10%				
T	λ^*	$\widehat{\lambda}_1^A$	$\widehat{\lambda}_2^A$	$\widehat{\lambda}_1^B$	$\widehat{\lambda}_2^B$	λ^*	$\widehat{\lambda}_1^A$	$\widehat{\lambda}_2^A$	$\widehat{\lambda}_1^B$	$\widehat{\lambda}_2^B$	λ^*	$\widehat{\lambda}_1^A$	$\widehat{\lambda}_2^A$	$\widehat{\lambda}_1^B$	$\widehat{\lambda}_2^B$
Panel A: Size test $W = S$															
60	0.69	2.91		3.77		2.64	9.85		10.95		4.68	16.75		17.98	
360	0.21	1.11		1.08		0.99	5.18		5.41		2.45	10.18		10.25	
600	0.17	0.89		0.91		0.83	4.99		5.04		2.05	10.28		10.29	
1000	0.13	0.92		0.99		0.87	5.31		5.28		2.33	10.22		10.04	
Panel B: Size test $W =\mathrm{E}[r_t r_t']$															
60	0.01	0.38	1.03	1.04	2.54	0.31	3.37	5.43	6.78	10.15	0.94	7.82	10.65	13.59	17.47
360	0.10	0.68	0.73	1.13	1.19	0.66	4.19	4.36	6.01	6.14	1.74	8.67	8.88	11.56	11.83
600	0.10	0.61	0.63	1.00	1.07	0.65	4.08	4.15	5.79	5.91	1.70	8.90	9.00	11.48	11.64
1000	0.10	0.72	0.75	1.15	1.16	0.78	4.52	4.59	5.89	5.94	2.13	8.80	8.88	11.30	11.40
Panel C: Size test $W =\mathrm{Cov}[r_t r_t']$															
60	0.69	3.31	3.95	6.69	7.89	2.64	10.05	11.33	16.67	18.49	4.68	16.23	18.01	25.37	27.76
360	0.21	0.97	1.02	1.62	1.74	0.99	5.15	5.28	7.08	7.18	2.45	9.86	10.00	13.10	13.28
600	0.17	0.77	0.78	1.24	1.25	0.83	4.65	4.74	6.50	6.61	2.05	9.64	9.68	12.43	12.52
1000	0.13	0.85	0.85	1.24	1.25	0.87	4.85	4.89	6.21	6.23	2.33	9.30	9.37	11.82	11.90
Panel D: Power test $W = S$															
60	0.91	3.76		4.60		3.22	10.73		12.36		5.92	17.97		19.65	
360	0.56	2.79		2.72		2.64	9.93		9.79		5.19	17.53		17.33	
600	0.88	3.87		3.78		3.81	13.14		13.06		7.16	22.14		21.69	
1000	1.73	7.08		6.94		6.67	20.46		20.18		11.97	30.86		30.59	
Panel E: Power test $W =\mathrm{E}[r_t r_t']$															
60	0.00	0.50	1.18	1.37	3.01	0.39	3.87	6.17	7.48	11.05	1.16	8.77	12.16	15.07	19.13
360	0.33	2.07	2.18	2.93	3.02	2.02	8.28	8.62	10.52	10.84	4.20	15.45	15.69	18.73	19.11
600	0.69	3.12	3.22	4.05	4.20	3.25	11.44	11.71	14.22	14.40	6.37	19.97	20.14	23.39	23.63
1000	1.35	5.89	5.96	7.58	7.64	6.03	18.47	18.60	21.69	21.84	11.20	28.82	28.96	32.43	32.53
Panel F: Power test $W =\mathrm{Cov}[r_t r_t']$															
60	0.91	3.75	4.61	7.29	8.77	3.22	11.17	12.49	18.31	20.14	5.92	18.32	20.03	27.02	29.32
360	0.56	2.73	2.80	3.78	3.85	2.64	9.65	9.76	12.26	12.53	5.19	17.00	17.28	20.64	20.83
600	0.88	3.74	3.79	4.80	4.84	3.81	12.63	12.76	15.35	15.43	7.16	21.21	21.30	24.58	24.67
1000	1.73	6.46	6.51	8.21	8.24	6.67	19.33	19.41	22.74	22.83	11.97	29.60	29.67	33.26	33.37

Table 4.11: Carhart specification tests on 10 size portfolios

	1%					5%					10%					
T	λ^*	$\widehat{\lambda}_1^A$	$\widehat{\lambda}_2^A$	$\widehat{\lambda}_1^B$	$\widehat{\lambda}_2^B$	λ^*	$\widehat{\lambda}_1^A$	$\widehat{\lambda}_2^A$	$\widehat{\lambda}_1^B$	$\widehat{\lambda}_2^B$	λ^*	$\widehat{\lambda}_1^A$	$\widehat{\lambda}_2^A$	$\widehat{\lambda}_1^B$	$\widehat{\lambda}_2^B$
Panel A: Size test $W = S$																
60	0.27	1.80		2.63		1.00	6.85		8.46		2.15	12.63		14.55	
360	0.07	0.75		0.93		0.38	4.53		5.12		1.08	9.06		9.82	
600	0.04	0.69		0.79		0.36	4.70		4.96		1.01	9.69		9.99	
1000	0.07	0.96		1.07		0.38	4.77		4.94		1.07	9.63		9.78	
Panel B: Size test $W = \mathrm{E}[r_t r_t']$																
60	0.01	0.41	1.24	1.56	3.68	0.11	2.75	5.12	7.98	12.37	0.35	6.70	10.78	15.60	20.58
360	0.06	0.52	0.60	1.32	1.46	0.27	3.76	4.18	6.65	7.05	0.78	7.99	8.49	12.25	12.95
600	0.02	0.57	0.59	1.23	1.31	0.26	3.74	3.88	6.56	6.82	0.83	8.12	8.45	12.47	12.76
1000	0.03	0.72	0.75	1.39	1.42	0.33	4.01	4.07	6.39	6.58	0.96	8.35	8.50	12.18	12.31
Panel C: Size test $W = \mathrm{Cov}[r_t r_t']$																
60	0.27	2.12	3.16	6.44	8.22	1.00	7.48	9.89	16.78	19.62	2.15	13.44	16.28	25.18	28.51
360	0.07	0.70	0.83	1.77	1.98	0.38	4.49	4.79	7.56	7.87	1.08	8.91	9.29	13.79	14.19
600	0.04	0.62	0.67	1.46	1.49	0.36	4.11	4.38	7.22	7.47	1.01	8.84	9.14	13.15	13.39
1000	0.07	0.84	0.85	1.50	1.52	0.38	4.23	4.34	6.71	6.81	1.07	8.57	8.72	12.53	12.63
Panel D: Power test $W = S$																
60	0.28	2.22		2.85		1.20	7.67		8.74		2.57	13.23		14.85	
360	0.25	2.06		2.08		1.38	9.19		8.82		3.01	16.57		16.11	
600	0.58	4.09		3.89		2.48	14.08		13.48		4.81	22.78		22.19	
1000	1.02	8.87		8.33		4.95	22.38		21.57		9.23	33.60		32.88	
Panel E: Power test $W = \mathrm{E}[r_t r_t']$																
60	0.00	0.45	1.38	1.66	3.85	0.09	3.65	6.12	8.80	13.11	0.39	8.01	11.87	16.64	21.78
360	0.13	2.14	2.37	3.02	3.22	1.01	9.02	9.44	11.44	11.89	2.24	16.80	17.35	20.04	20.52
600	0.42	3.97	4.11	5.18	5.33	2.16	13.83	14.19	16.19	16.50	4.21	22.88	23.13	26.10	26.34
1000	0.83	8.29	8.41	10.10	10.19	4.55	22.14	22.32	24.77	24.95	8.68	33.38	33.54	36.52	36.71
Panel F: Power test $W = \mathrm{Cov}[r_t r_t']$																
60	0.28	2.91	4.05	7.39	9.30	1.20	8.79	11.01	17.75	20.86	2.57	15.08	18.05	26.50	29.74
360	0.25	2.79	3.04	3.85	4.00	1.38	10.28	10.65	13.03	13.43	3.01	18.26	18.83	21.76	22.20
600	0.58	4.51	4.63	6.00	6.10	2.48	14.79	14.99	17.29	17.52	4.81	23.79	24.03	27.19	27.43
1000	1.02	8.94	9.03	10.68	10.77	4.95	22.92	22.99	25.55	25.62	9.23	34.01	34.12	37.36	37.48

Concluding Remarks

The interest on learning about the asymptotic and finite sample properties of asset pricing model estimators, like risk premiums and pricing errors, have attracted the attention of researchers for decades. This interest is motivated for an extensive list of theoretical and empirical applications mainly – but not exclusive – in economics and finance areas. Roughly speaking, the study of estimator properties usually involves a trade-off. For example, an econometric method which leads to more efficient estimators may comes at the cost of higher pricing errors and vice versa. This evidence is useful for researchers and practitioners because they could choose a proper procedure in terms of a given application. Consequently, the adequate selection of the econometric procedure lead to more accurate hypothesis tests and other kind of computations.

Even though any asset pricing model can be defined either under the Beta or the SDF representation, the literature has traditionally focus on analyzing one of these two representations. In other words, it is not uncommon to find comparisons of different Beta procedures, and in the other hand, comparisons of different SDF procedures. Only until recently, researchers have established the correct framework for comparing the two methods, opening a rich research field in financial econometrics and asset pricing.

Our original contribution provides new evidence on the comparison between the Beta and SDF methodologies. We argue that current studies which compare both methodologies are conducted under certain conditions that are not sufficient to differentiate them. Once we relax those conditions, we show that differences between the two methods emerge. Specifically, we find evidence that suggest that the Beta method lead to better risk premium estimators while the SDF method lead to better pricing error estimators in terms of efficiency. We evaluate the magnitude of the resulting biases and their possible explanations.

In chapters three and four we show the main pieces of empirical evidence that support our principal and original findings. In brief, we perform extensive combinations of estimations and tests under different approaches which not only confirm previous findings but provide new arguments to the current debate between the differences of both methods. Chapters two and three contributes to the knowledge of the finite sample performance of different Beta and SDF procedures by separate. In addition, our set of results also serves to empirically evaluate not only methods but models too. Particularly, we study single and multi-factor asset pricing models in which factors are tradable assets. Furthermore, chapters two and three introduce an empirical motivated model which outperforms other well known pricing models. Finally, chapter one provides a comprehensive econometric review, based on recognized authors, which is specifically orientated to facilitate the programming.

■

Bibliography

[1] Seung C. Ahn and Christopher Gadarowski, *Small sample properties of the model specification test based on the hansen-jagannathan distance*, Journal of Empirical Finance **11(1)** (2004), 109–132.

[2] Christine E. Amsler and Peter Schmidt, *A monte carlo investigation of the accuracy of multivariate capm tests*, Journal of Financial Economics **14(3)** (1985), 359–375.

[3] Andrew Ang and Joseph S. Chen, *Capm over the long-run: 1926-2001*, Journal of Empirical Finance **14(1)** (2007), 1–40.

[4] Perluigi Balduzzi and Tong Yao, *Testing heterogeneous-agent models: An alternative aggregation approach*, Journal of Monetary Economics **54(2)** (2007), 369–412.

[5] Ronald J. Balvers and Dayong Huang, *Productivity-based asset pricing: Theory and evidence*, Journal of Financial Economics **86(2)** (2007), 405–445.

[6] Rolf W. Banz, *The relationship between return and market value of common stocks*, Journal of Financial Economics **9(1)** (1981), 3–18.

[7] Sanjoy Basu, *Investment performance of common stocks in relationship to their price-earnings ratios: A test of the efficient market hypothesis*, Journal of Finance **32(3)** (1977), 663–682.

[8] Fischer Black, *Capital market equilibrium with restricted borrowing*, Journal of Business **45(3)** (1972), 444–455.

[9] Fischer Black, Michael Jensen, and Myron S. Scholes, *The capital asset pricing model: Some empirical tests*, Studies in Theory of Capital Markets. Jensen, Ed. New York: Preager Publishers. **17** (1972), 79–121.

[10] Michael W. Brandt and David A. Chapman, *Linear approximations and tests of conditional pricing models*, National Bureau of Economic Research Working Papers 12513, (2006).

[11] Douglas T. Breeden, *An intertemporal asset pricing model with stochastic consumption and investment opportunities*, Journal of Financial Economics **7(3)** (1979), 265–296.

[12] Douglas T. Breeden, Michael R. Gibbons, and Robert H. Litzenberger, *Empirical tests of the consumption-oriented capm*, Journal of Finance **44(2)** (1989), 231–262.

[13] Michael J. Brennan, Tarun Chordia, and Avanidhar Subrahmanyam, *Alternative factor specifications, security characteristics and the cross-section of expected stock returns*, Journal of Financial Economics **49(3)** (1998), 345–373.

[14] Craig Burnside, *Empirical asset pricing and statistical power in the presence of weak risk factors*, Duke University and NBER Working Papers with number 13357 (2007).

[15] Zongwu Cai and Yongmiao Hong, *Some recent developments in nonparametric finance*, Advances in Econometrics **Forthcoming** (2009).

[16] John Y. Campbell, *Understanding risk and return*, Journal of Political Economy **104(2)** (1996), 298–345.

[17] John Y. Campbell, Andrew W. Lo, and A. Craig MacKinlay, *The econometrics of financial markets*, Princeton NJ: Princeton University Press, 1997.

[18] Mark M. Carhart, *On persistence in mutual fund performance*, Journal of Finance **52(1)** (1997).

[19] Gary Chamberlain, *Funds, factors and diversification in arbitrage pricing models*, Econometrica **51(5)** (1983), 1305–1323.

[20] David A. Chapman, *Approximating the asset pricing kernel*, Journal of Finance **52(4)** (1997), 1383–1410.

[21] Nai-Fu Chen, Richard R. Roll, and Stephen Alan Ross, *Economic forces and the stock market*, Journal of Business **59(3)** (1986), 383–403.

[22] Robert Chen and Raymond Kan, *Finite sample analysis of two-pass cross-sectional regressions*, Working Paper, University of Toronto, (2003).

[23] John H. Cochrane, *A resurrection of the stochastic discount factor / gmm methodology*, Working Paper, Graduate School of Business, University of Chicago, (2000).

[24] ______, *A cross-sectional test of an investment-based asset pricing model*, Journal of Political Economy **104(3)** (1996), 572–621.

[25] ______, *Asset pricing - revised edition*, Princeton University Press, 2005.

[26] Kent Daniel and Sheridan Titman, *Evidence on the characteristics of cross sectional variation in stock returns*, Journal of Finance **52(1)** (1997), 1–33.

[27] Gene D'Avolio, *The market for borrowing stocks*, Journal of Financial Economics **66(2-3)** (2002), 271–306.

[28] Philip H. Dybvig and Jonathan E. Ingersoll, *Mean variance theory in complete markets*, Journal of Business **55(2)** (1982), 233–251.

[29] Eugene F. Fama and Kenneth R. French, *The cross-section of expected stock returns*, Journal of Finance **47(2)** (1992), 427–465.

[30] ______, *Common risk factors in the returns on stocks and bonds*, Journal of Financial Economics **33(1)** (1993), 3–56.

[31] ______, *Multifactor explanations of asset-pricing anomalies*, Journal of Finance **51(1)** (1996), 55–84.

[32] Eugene F. Fama and James D. MacBeth, *Risk return and equilibrium: empirical tests*, Journal of Financial Political Economy **81(3)** (1973), 607–636.

[33] Heber Farnsworth, Wayne Ferson, David Jackson, and Steven Todd, *Performance evaluation with stochastic discount factors*, Journal of Business **75(3)** (2002), 473–503.

[34] Wayne E. Ferson, *Theory and empirical testing of asset pricing models*, Handbooks in Operations Research and Management Science edited by Jarrow, Maksimovic and Ziemba **9** (1995), 145–200.

[35] ______, *Tests of multifactor pricing models, volatility bounds and portfolio performance*, Handbook of the Economics of Finance. Elsevier Science Publishers, North Holland, 2005.

[36] Wayne E. Ferson and Stephen R. Foerster, *Finite sample properties of the generalized method of moments in tests of conditional asset pricing models*, Journal of Financial Economics **36(1)** (1994), 29–55.

[37] Wayne E. Ferson and Campbell Russell Harvey, *Conditioning variables and the cross section of stock returns*, Journal of Finance **54(4)** (1999), 1325–1360.

[38] Wayne E. Ferson and Ravi Jagannathan, *Econometric evaluation of asset pricing models*, G. S. Maddala and C. R. Rao, Eds. Handbook of Statistics, Vol. 14: Statistical Methods in Finance. Elsevier, 1996.

[39] Wayne E. Ferson and Andrew F. Siegel, *Testing portfolio efficiency with conditioning information*, National Bureau of Economic Research, Inc in its series NBER Working Papers with number 12098. 2007.

[40] Michael R. Gibbons, *Multivariate tests of financial models: A new approach*, Journal of Financial Economics **10(1)** (1982), 3–27.

[41] Michael R. Gibbons, Stephen Alan Ross, and Jay Shanken, *A test of the efficiency of a given portfolio*, Econometrica **57(5)** (1989), 1121–1152.

[42] Robert R. Grauer and Johannus A. Janmaat, *On the power of cross-sectional and multivariate tests of the capm*, Journal of Banking and Finance **33(5)** (2009), 775–787.

[43] Mark Grinblatt and Sheridan Titman, *The relation between mean-variance efficiency and arbitrage pricing*, Journal of Business **60(1)** (1987), 97–112.

[44] Alastair R. Hall, *Generalized method of moments*, Oxford University Press, Oxford UK, 2005.

[45] Alastair R. Hall and Atsushi Inoue, *The large sample behaviour of the generalized method of moments estimator in misspecified models*, Journal of Econometrics **114(2)** (2003), 361–394.

[46] Lars Peter Hansen, *Large sample properties of generalized method of moments estimators*, Econometrica **50(4)** (1982), 1029–1054.

[47] Lars Peter Hansen, John Heaton, and Amir Yaron, *Finite-sample properties of some alternative gmm estimators*, Journal of Business and Economic Statistics **14(3)** (1996), 262–280.

[48] Lars Peter Hansen and Ravi Jagannathan, *Implications of security market data for models of dynamic economies*, Journal of Political Economy **99(2)** (1991), 225–262.

[49] ———, *Assessing specification errors in stochastic discount factor models*, Journal of Finance **55(2)** (1997), 557–590.

[50] Lars Peter Hansen and Scott F. Richard, *The role of conditioning information in deducing testable*, Econometrica **55(3)** (1987), 587–613.

[51] Lars Peter Hansen and Kenneth J. Singleton, *Generalized instrumental variables estimation of nonlinear rational expectations models*, Econometrica **50(5)** (1982), 1269–1288.

[52] Kewei Hou and Robert L. Kimmel, *On the estimation of risk premia in linear factor models*, Working Paper, Ohio State University, (2006).

[53] Jonathan E. Ingersoll, *Theory of financial decision making*, Rowman and Littlefield Publishing, Inc., 1987.

[54] Ravi Jagannathan, Georgios Skoulakis, and Zhenyu Wang, *Generalized method of moments: applications in finance*, Journal of Business and Economics Statistics **20(4)** (2002), 470–481.

[55] ———, *The analysis of the cross section of security returns*, Handbook of Financial Econometrics. Edited By Yacine Ait-Sahalia and Lars Hansen **1** (2008), ch. 14.

[56] Ravi Jagannathan and Zhenyu Wang, *The conditional capm and the cross-section of expected returns*, Journal of Finance **51(1)** (1996), 3–53.

[57] ———, *An asymptotic theory for estimating beta-pricing models using cross-sectional regression*, Journal of Finance **53(4)** (1998), 1285–1309.

[58] ———, *Empirical evaluation of asset-pricing models: A comparison of the sdf and beta methods*, Journal of Finance **57(5)** (2002), 2337–2367.

[59] Narasimhan Jegadeesh and Sheridan Titman, *Returns to buying winners and selling losers: Implications for stock market efficiency*, Journal of Finance **48(1)** (1993), 65–91.

[60] Michael C. Jensen, *The performance of mutual funds in the period 1945-1964*, Journal of Finance **23(2)** (1967), 389–416.

[61] Dave J. Jobson and Robert M. Korkie, *Potential performance and tests of portfolio efficiency*, Journal of Financial Economics **10(4)** (1982).

[62] Christian Julliard and Jonathan A. Parker, *Consumption risk and the cross-section of expected returns*, Journal of Political Economy **113(1)** (2005).

[63] Raymond Kan and Cesare Robotti, *A note on the estimation of asset pricing models using simple regression betas*, Federal Reserve Bank of Atlanta in its series Working Paper with number 2009-12.

[64] ———, *Specification tests of asset pricing models using excess returns*, Journal of Empirical Finance **15(5)** (2008), 816–838.

[65] Raymond Kan, Cesare Robotti, and Jay Shanken, *Pricing model performance and the two-pass cross-sectional regression methodology*, Federal Reserve Bank of Atlanta in its series Working Paper with number 2009-11.

[66] Raymond Kan and Guofu Zhou, *Empirical asset pricing: the beta method versus the stochastic discount factor method*, Working Paper, University of Toronto and Washington University in St. Louis, (2002).

[67] ———, *Modeling non-normality using multivariate t: Implications for asset pricing*, Working Paper, (2003).

[68] ______, *A critique of the stochastic discount factor methodology*, Journal of Finance **54(4)** (1999), 1221–1248.

[69] S. P. Kothari, Jay Shanken, and Richard G. Sloan, *Another look at the cross-section of expected stock returns*, Journal of Finance **50(1)** (1995), 185–224.

[70] Owen A. Lamont and Richard H. Thaler, *Can the market add and subtract? mispricing in tech stock carve-outs*, Journal of Political Economy **111(2)** (2003), 227–268.

[71] Martin Lettau and Sydney C. Ludvigson, *Resurrecting the (c)capm: A cross-sectional test when risk premia are time-varying*, Journal of Political Economy **109(6)** (2001), 1238–1287.

[72] Jonathan Lewellen, Stefan Nagel, and Jay Shanken, *A skeptical appraisal of asset pricing tests*, Journal of Financial Economics **Forthcoming (2009)** (2009).

[73] Haitao Li, Yuewu Xu, and Xiaoyan Zhang, *Evaluating asset pricing models using the second hansen-jagannathan distance*, Journal of Financial Economics **Forthcoming** (2009).

[74] Jimmy Liew and Maria Vassalou, *Can book-to-market, size, and momentum be risk factors that predict economic growth?*, Journal of Financial Economics **57(2)** (2000), 221–245.

[75] John Lintner, *Security prices, risk and maximal gains from diversification*, Journal of Finance **20(4)** (1965), 587–615.

[76] Andrew W. Lo and A. Craig MacKinlay, *Data-snooping biases in tests of financial asset pricing models*, Review of Financial Studies **3(3)** (1990), 431–467.

[77] Tim Loughran, *Book-to-market across firm size, exchange, and seasonality: Is there an effect?*, Journal of Financial and Quantitative Analysis **32(3)** (1997), 249–268.

[78] A. Craig MacKinlay, *On multivariate tests of the capm*, Journal of Financial Economics **18(2)** (1987), 341–371.

[79] ______, *Multifactor models do not explain deviations from the capm*, Journal of Financial Economics **38(1)** (1995), 3–28.

[80] A. Craig MacKinlay and Matthew P. Richardson, *Using generalized method of moments to test mean-variance efficiency*, Journal of Finance **46(2)** (1991), 511–527.

[81] José M. Marín and Gonzalo Rubio, *Economía financiera*, Antoni Bosch, editor, 2001.

[82] Robert C. Merton, *Optimum consumption and portfolio rules in a continuous time model*, Journal of Economic Theory **3(4)** (1971), 373–413.

[83] ______, *An intertemporal capital asset pricing model*, Econometrica **41(5)** (1973), 867–87.

[84] Whitney K. Newey and Kenneth D. West, *A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix*, Econometrica **55(3)** (1987), 703–708.

[85] Belén Nieto and Rosa Rodríguez, *Modelos de valoración de activos condicionales: un panorama comparativo*, Investigaciones Económicas **XXIX (1)** (2005), 33–72.

[86] Masao Ogaki, *Generalized method of moments: econometric applications. handbook of statistics*, North-Holland, Amsterdam **11** (1993), 455–488.

[87] Ralitsa Petkova, *Do the fama-french factors proxy for innovations in predictive variable?*, Journal of Finance **61(2)** (2006), 581–612.

[88] Marc R. Reinganum, *A new empirical perspective on the capm*, Journal of Financial and Quantitative Analysis **16(4)** (1981), 439–462.

[89] R. W. Roll, *A critique of the asset pricing theory's tests part i: On past and potential testability of the theory*, Journal of Financial Economics **4(2)** (1977), 129–176.

[90] Stephen Alan Ross, *A simple approach to the valuation of risky streams*, Journal of Business **51(3)** (1978), 453–475.

[91] Jay Shanken, *Multivariate tests of the zero-beta capm*, Journal of Financial Economics **14(3)** (1985), 327–348.

[92] ______, *On the estimation of beta pricing models*, Review of Financial Studies **5(1)** (1992), 1–33.

[93] ______, *Statistical methods in tests of portfolio efficiency: A synthesis*, In: Maddala, G.S., Rao, C.R. (Eds.), Handbook of Statistics, North-Holland, Amsterdam **14** (1996), 693–711.

[94] Jay Shanken and Guofu Zhou, *Estimating and testing beta pricing models: Alternative method and their performance in simulations*, Journal of Financial Economics **84(1)** (2007), 40–86.

[95] William F. Sharpe, *Capital asset prices: A theory of market equilibrium under conditions of risk*, Journal of Finance **19(3)** (1964), 425–442.

[96] Georgios Skoulakis, *Assessment of asset-pricing models using cross-sectional regressions*, Working Paper, Finance Department, Northwestern University (2005).

[97] Peter Smith and Michael Wickens, *Asset pricing with observable stochastic discount factors*, Journal of Economic Surveys **16(3)** (2002), 397–446.

[98] Robert F. Stambaugh, *On the exclusion of assets from tests of the two-parameter model*, Journal of Financial Economics **10(3)** (1982), 237–268.

[99] Maria Vassalou, *News related to future gdp growth as a risk factor in equity returns*, Journal of Financial Economics **68(1)** (2003), 47–73.

[100] Maria Vassalou, Qing Li, and Yuhang Xing, *Sector investment growth rates and the cross-section of equity returns*, Journal of Business **79(3)** (2006), 1637–1636.

[101] Raja Velu and Guofu Zhou, *Testing multi-beta asset pricing models*, Journal of Empirical Finance **6(3)** (1999), 219–241.

[102] Zhenyu Wang and Xiaoyan Zhang, *Empirical evaluation of asset pricing models: Arbitrage and pricing errors over contingent claims*, Working Paper, Federal Reserve Bank of New York in its series Staff Reports with number 265, (2005).

■■