Journal of Empirical Finance 17 (2010) 54–80

Contents lists available at ScienceDirect

Journal of Empirical Finance

journal homepage: www.elsevier.com/locate/econbase

Asset pricing models and economic risk premia: A decomposition☆

Pierluigi Balduzzi[a], Cesare Robotti[b,*]

[a] Boston College, United States
[b] Research Department, Federal Reserve Bank of Atlanta, 1000 Peachtree St. N.E., Atlanta, Georgia, 30309, United States

ARTICLE INFO

Article history:
Received 27 June 2008
Received in revised form 4 September 2009
Accepted 7 September 2009
Available online 19 September 2009

JEL classification:
G12

Keywords:
Economic factor
Risk premium
Pricing kernel
Maximum-correlation portfolio

ABSTRACT

The risk premia of linear factor models on economic (non-traded) risk factors can be decomposed into: i) the premium on maximum-correlation portfolios mimicking the factors; ii) (minus) the covariance between the non-traded components of the pricing kernel and the factors; and iii) (minus) the mispricing of the maximum-correlation portfolios. For a given set of assets available for investment, the first component is the same across models and is typically estimated with little bias and high precision. We conclude that the premia on maximum-correlation portfolios are appealing alternatives to the risk premia of linear factor models, with the dividend yield being the only economic factor significantly priced.

1. Introduction

The estimates of risk premia associated with economic (non-traded) risk variables are relevant for practitioners, policymakers, and academics. Practitioners are interested in understanding how to price contracts and securities whose payoffs are affected by economic risk factors. Policymakers are interested in learning about the economic risks that agents want to avoid the most. Academics are interested in relating security risk premia to the exposures to systematic sources of risk. Yet, existing estimates of risk premia in linear factor models vary substantially in sign and statistical significance from one study to the other. For example, the premium on inflation is negative and significant in Chen et al. (1986); positive and insignificant in McElroy and Burmeister (1988); negative and marginally significant in Ferson and Harvey (1991); and negative and insignificant in Jagannathan and Wang (1996). Another example is the premium on the slope of the term structure, which is negative and mostly insignificant in Chen et al. (1986); positive and significant in McElroy and Burmeister (1988); positive and marginally significant in Ferson and Harvey (1991); and negative and insignificant in Jagannathan and Wang (1996).

In this paper, we reconsider the evidence on economic risk premia in light of a novel decomposition. We show that an economic risk premium is model dependent and equals the negative of the covariance between a normalized (mean-one) candidate pricing kernel and the risk factor. The economic risk premium can then be broken into three components: i) the expected excess return on the maximum-correlation portfolio tracking the factor (λ^*); ii) (minus) the covariance between the

☆ We are grateful to Wayne Ferson, Raymond Kan, Jay Shanken, two anonymous referees, and seminar participants at Boston College, Concordia University, Ente Einaudi, the Federal Reserve Bank of Atlanta, the University of Brescia, and the University of Minnesota for useful comments. The views expressed here are the authors' and not necessarily those of the Federal Reserve Bank of Atlanta or the Federal Reserve System.

* Corresponding author. Tel.: +1 404 498 8543; fax: +1 404 498 8810.
E-mail address: Cesare.Robotti@atl.frb.org (C. Robotti).

0927-5398/$ – see front matter
doi:10.1016/j.jempfin.2009.09.009

non-traded components of the factor and of the candidate pricing kernel of a given model (δ_n); and iii) (minus) the mispricing that a given model (pricing kernel) assigns to the maximum-correlation mimicking portfolio tracking the factor (δ_m).[1]

The decomposition above highlights the two conceptual issues that arise when considering economic risk premia. First, the non-traded component of an economic risk premium is arbitrary, as it does not depend on security-return information. This point is further illustrated in the case where the pricing kernel and the economic factor are contaminated by common measurement error, unrelated to security returns. In this case, we can make the economic risk premium arbitrarily large (in absolute value) without affecting the pricing properties of the kernel. Second, the mispricing component reflects the inability of the model to price assets and is also arbitrary. In fact, we obtain different estimates of the pricing-kernel parameters, and, hence, of an economic risk premium, depending on the moment conditions imposed in estimation. On the other hand, the risk premia on the maximum-correlation portfolios, delivering the first component of an economic risk premium, are immune from these shortcomings. In addition, these risk premia have special economic significance, as they are the risk premia on the hedging portfolios of Merton (1973).[2]

We also show how one can derive bounds on the non-traded and mispricing components of a risk premium, in the spirit of the volatility bounds of Balduzzi and Kallal (1997) and the good-deal pricing bounds of Cochrane and Saá-Requejo (2000). Specifically, we show that the absolute value of the non-traded component of a risk premium is bounded by the non-traded variability (i.e., the standard deviation) of a pricing kernel times the non-traded variability of a factor. On the other hand, the absolute value of the mispricing component is bounded by the Hansen and Jagannathan (1997) distance times the traded variability of a factor.

Our decomposition holds regardless of whether the underlying kernel prices assets correctly, i.e., whether the kernel is admissible. In fact, the determination of the λ^* and δ_n components of a risk premium does not hinge on the mispricing component δ_m: the premium on the maximum-correlation portfolio depends only on the set of assets available for investment and is the same across models; while the non-traded component does not depend on security-return information. Moreover, looking at possibly misspecified kernels seems sensible, as all models are false and, at best, close approximations to reality. Researchers are often interested in studying potentially misspecified models and sometimes keep the issue of estimating risk premia separate from the issue of testing whether the underlying asset pricing model is literally true or not. Examples of this practice are, among others, Ferson and Harvey (1991), Jagannathan and Wang (1996), and Cochrane (1996). The increasing interest in the statistical properties of economic risk premia estimates under potentially misspecified models is further reflected in several recent papers, such as Hall and Inoue (2003), Hou and Kimmel (2006), Shanken and Zhou (2007), Kan and Robotti (2008, 2009), and Kan et al. (2009).

In the empirical analysis, we study the three components of the risk premia associated with the innovations in the term structure, the dividend yield, consumption growth, the default premium, inflation, and the real rate of interest, where we obtain estimates of these innovations via a VAR(1) of traded and non-traded factors. We consider two linear factor models with non-traded factors: the intertemporal capital asset pricing model (I-CAPM) and the consumption capital asset pricing model (C-CAPM). In addition, we consider two models with traded factors only: the Fama and French (1993) three-factor model and the capital asset pricing model (CAPM). In our setting, what distinguishes models with traded factors is that they are calibrated in such a way that the traded factors are priced correctly, i.e., the risk premia assigned by the models to the traded factors coincide with the sample means of the factors. This restriction was introduced by Shanken (1992) and has been implemented in several studies of beta-pricing models. We also examine the issue of time variation: we allow risk premia to vary over time as linear functions of a vector of instruments. This is consistent, for example, with the intuition of the I-CAPM that some economic variables should be priced risks because they affect the position of the investment-opportunity set.

We find that the estimates of economic risk premia assigned by models with non-traded factors are almost always much larger (in absolute value) than the estimates of the premia on maximum-correlation portfolios. Indeed, the absolute size of the economic risk premia assigned by these models is often completely unrealistic. Moreover, the risk premia estimates tend to exhibit large (absolute) biases and standard errors, and can be sensitive to the choice of moment conditions. The premia on maximum-correlation portfolios themselves tend to be estimated precisely and with little bias. For example, consider the term-structure risk premium assigned by the I-CAPM. We have estimates of −3.0554 and −2.2015, depending on the moment conditions, while the risk premium on the maximum-correlation portfolio is −0.0216. Since we are working with standardized factors, i.e., factors with unit standard deviation, these estimates have the interpretation of Sharpe ratios on portfolios tracking the slope of the term structure. These estimates can be compared with the Sharpe ratio on the U.S. equity index of 0.0954 for the same period. It is hard to believe that a portfolio tracking the slope of the term structure could have a Sharpe ratio that is, in absolute size, 32 times the Sharpe ratio on the market! The biases of the two risk premia estimates are also substantial, 2.4037 and 1.5931, respectively, and so are the standard errors of 1.3096 and 0.8503, respectively. On the other hand, the bias of the premium on the maximum correlation portfolio is only −0.0007 and the standard error is only 0.0139. We document how the large discrepancies between the estimates of economic risk premia and the estimates of premia on maximum-correlation portfolios are mainly due to the non-traded component, and that this component is mainly responsible for the poor properties of the estimates.

[1] Related to the approach of the present paper are Hou and Kimmel (2006) and Balduzzi and Robotti (2008). Hou and Kimmel (2006) derive the asymptotic properties of estimates of the economic risk premia and of the premia on the maximum-correlation portfolios under a normality assumption on factors and returns. Balduzzi and Robotti (2008) investigate the small-sample properties of tests of linear factor models with non-traded factors when the original factors are used, and when the factors are replaced by the excess returns on the maximum-correlation portfolios.

[2] Fama (1996) shows that Merton's investors hold overall portfolios that minimize the return variance for given expected return and covariances with the state variables. Other papers focusing on the maximum-correlation portfolios include Breeden (1979), Breeden et al. (1989), Fama (1996), Lamont (2001), Ferson et al. (2005), and Van den Goorbergh et al. (2005).

It is worth noting that the estimates of the economic risk premia assigned by kernels with non-traded factors not only tend to be larger in absolute value, but also tend to be more significant than the estimates of the premia on maximum-correlation portfolios. In particular, the only maximum-correlation portfolio with a significant (5% level) risk premium is the one tracking the dividend yield with a Sharpe ratio of −0.0629. In contrast, four of the six risk premia assigned by the I-CAPM are significant. For example, while the *p*-value of the premium on the portfolio tracking the inflation rate is 42.5%, the *p*-value of the I-CAPM inflation risk premium is 0.5%.

We then implement the bounds on the mispricing and non-traded components of the economic risk premia, in the spirit of the risk premium bounds of Balduzzi and Kallal (1997) and the good-deal pricing bounds of Cochrane and Saá-Requejo (2000). We show that for our models, the bounds on both components can be wide relative to the risk premia on the maximum-correlation portfolios.

Finally, we explore two extensions of our analysis by modifying the properties of the data. First, we perform a simulation under the null to determine whether the absence of mispricing affects our results. Second, we perform a series of simulations where we change the tradeability of an economic factor. We find that eliminating mispricing in the population does not alter the main result of our analysis: the non-traded component of the economic risk premia is substantial and poorly estimated. On the other hand, as the non-traded variability of a factor decreases, the statistical properties of the risk premia estimates improve markedly.

The results of the empirical analysis are certainly specific to our choice of non-traded factors and our approach to modeling innovations. Yet, our results demonstrate how the estimation of economic risk premia assigned by models with non-traded factors can be challenging. On the other hand, the risk premia on maximum-correlation portfolios are an appealing alternative, as they do not depend on the mispricing of a given asset pricing model and they only reflect security-return information. Moreover, they tend to be estimated well in small samples. Our results shed light on the lack of robustness of risk premia estimates documented in previous studies: as the specifications of the linear factor model differ from one study to the other, the non-traded component of an economic risk premium changes as well. This can lead to substantially different results. Moreover, the non-traded component of the risk premium is estimated imprecisely. Hence, relatively minor changes in the collection of assets considered and/or the sample period can also lead to substantially different results.

This paper is organized as follows: Section 2 illustrates the decomposition of risk premia assigned by linear factor models. Section 3 presents the orthogonality conditions imposed in estimation. Section 4 describes the data. Section 5 presents the empirical results. Section 6 presents extensions of our analysis. Section 7 concludes.

2. Decomposing risk premia

2.1. Economic risk premia

We start with the familiar asset pricing equation

$$E_t(q_{t+1} r_{t+1}) = 0, \tag{1}$$

where q_{t+1} is a mean-one, admissible pricing kernel, r_{t+1} is an $N \times 1$ vector of security returns in excess of the risk-free rate, and $E_t(.)$ denotes a conditional expectation.

Consider now a $K \times 1$ vector of risk factors y_{t+1}. We define the risk premia π_t as

$$\pi_t \equiv -\mathrm{Cov}_t(q_{t+1}, y_{t+1}), \tag{2}$$

where $\mathrm{Cov}_t(.,.)$ denotes a conditional covariance. Note that, if the factors were traded, then there would be exact mimicking portfolios with returns

$$r_{y,t+1} = E_t(r_{y,t+1}) + [y_{t+1} - E_t(y_{t+1})], \tag{3}$$

and the risk premia π_t would coincide with the expected excess returns on the exact mimicking portfolios:

$$E_t(r_{y,t+1}) = \pi_t. \tag{4}$$

On the other hand, if the factors are not traded, i.e., they are *economic* factors, then we need an explicit model to determine π_t. We call the risk premia associated with economic factors *economic* risk premia. Moreover, if the factors are not traded, then markets are incomplete, and we have a multiplicity of admissible kernels q_{t+1}. This also implies a multiplicity of economic risk premia π_t.

Since, in general, we do not know the true asset pricing model, we have to work with a *candidate* pricing kernel x_{t+1}. While the analysis of the following sections focuses on the case of linear factor models, at this stage we do not take a stand on the form of

x_{t+1}. The kernel x_{t+1} could be, for example, the intertemporal marginal rate of substitution of the nonlinear C-CAPM. In addition, the candidate kernel does not necessarily price all the available assets correctly; i.e., thecandidate kernel is not necessarily admissible and

$$E_t(x_{t+1}r_{t+1}) \equiv \alpha_t, \tag{5}$$

where α_t is the $N\times 1$ vector of pricing errors. The risk premia λ_t assigned by the candidate pricing kernel are defined as

$$\lambda_t \equiv -\mathrm{Cov}_t(x_{t+1}, y_{t+1}). \tag{6}$$

These would also be the risk premia on exact mimicking portfolios, according to the candidate kernel, if such exact mimicking portfolios existed.

Note that our definition of a risk premium follows that of Balduzzi and Kallal (1997), where the candidate pricing kernel x_{t+1} is not necessarily perfectly spanned by a linear combination of the risk factors y_{t+1} considered by the econometrician. The main reason for this approach is that a researcher could have strong priors on the validity of a model. Given the chosen model, though, one may be interested in pricing the variability of factors that occupy a different space than the one in which the researcher went looking for a kernel. For example, the researcher may think that the CAPM is the true model, but he may be interested in pricing a contract whose payoffs depend on a different non-traded factor, as in the case of the Chicago Mercantile Exchange CPI futures contract.

2.2. Decomposition

The main goal of this section is to provide an economic interpretation of the economic risk premia λ_t assigned by the candidate kernel x_{t+1}. For this purpose, consider the projection of x_{t+1} onto the span of asset excess returns r_{t+1} augmented with a constant, $x^\star_{t+1}$. Consider also the minimum-variance (MV) admissible kernel,

$$q^\star_{t+1} \equiv 1 - [r_{t+1} - E_t(r_{t+1})]^\top \Sigma_{rrt}^{-1} E_t(r_{t+1}) \tag{7}$$

(see Hansen and Jagannathan, 1991), where Σ_{rrt} is the conditional covariance matrix of asset excess returns. It is straightforward to verify that

$$E_t(q^\star_{t+1} r_{t+1}) = 0. \tag{8}$$

Since $q^\star_{t+1}$ is admissible and is linear in r_{t+1}, it is also the projection of any admissible kernel q_{t+1} onto the augmented span of asset returns. Note that $q^\star_{t+1}$ is perfectly negatively correlated with the return on the tangency portfolio. Indeed, the fact that $q^\star_{t+1}$ is admissible is equivalent to the statement that all asset returns have a beta representation with respect to the return on the tangency portfolio.

Given the definitions above, we can then write

$$x_{t+1} \equiv q^\star_{t+1} + (x_{t+1} - x^\star_{t+1}) + (x^\star_{t+1} - q^\star_{t+1}). \tag{9}$$

The first component is the projection of q_{t+1} onto the augmented span of asset returns. This component prices assets correctly by construction. The second term of the equation above is the non-traded component of the candidate kernel. This second component is orthogonal to asset returns and does not affect pricing. The third component is the difference between the projections of x_{t+1} and q_{t+1} onto the augmented span of asset returns. This component determines the pricing properties of x_{t+1}. In fact, we have

$$E_t[(x^\star_{t+1} - q^\star_{t+1}) r_{t+1}] = E_t(x^\star_{t+1} r_{t+1}) - E_t(q^\star_{t+1} r_{t+1}) = E_t(x^\star_{t+1} r_{t+1}) = E_t(x_{t+1} r_{t+1}) = \alpha_t. \tag{10}$$

Consider now $y^\star_{t+1} \equiv (\gamma^\star_t)^\top r_{t+1}$, the variable part of the projection of y_{t+1} onto the augmented span of excess returns. Based on the decomposition of x_{t+1} in Eq. (9), we can rewrite the risk premia in Eq. (6) as

$$\lambda_t \equiv -\mathrm{Cov}_t(q^\star_{t+1}, y_{t+1}) + \mathrm{Cov}_t[(x^\star_{t+1} - x_{t+1}), y_{t+1}] + \mathrm{Cov}_t[(q^\star_{t+1} - x^\star_{t+1}), y_{t+1}]. \tag{11}$$

It is easy to see that the first term of the r.h.s. of Eq. (11) equals

$$-\mathrm{Cov}_t(q^\star_{t+1}, y_{t+1}) = (\gamma^\star_t)^\top E_t(r_{t+1}) \equiv \lambda^\star_t, \tag{12}$$

which are the expected excess returns on the maximum-correlation portfolios.

Also, note that the third term of the r.h.s. of Eq. (11) equals

$$\begin{aligned}\mathrm{Cov}_t[(q^\star_{t+1}-x^\star_{t+1}),y_{t+1}] &= \mathrm{Cov}_t[(q^\star_{t+1}-x^\star_{t+1}),y^\star_{t+1}]\\ &= -\lambda^\star_t-(\gamma^\star_t)^\top \mathrm{Cov}_t(x^\star_{t+1},r_{t+1})\\ &= -(\gamma^\star_t)^\top E_t(r_{t+1})-(\gamma^\star_t)^\top \mathrm{Cov}_t(x^\star_{t+1},r_{t+1})\\ &= -(\gamma^\star_t)^\top [E_t(r_{t+1})+\mathrm{Cov}_t(x^\star_{t+1},r_{t+1})]\\ &= -(\gamma^\star_t)^\top \alpha_t,\end{aligned} \tag{13}$$

which is (minus) the mispricing, or "alphas," of the maximum-correlation portfolios according to the candidate kernel x_{t+1}.

Hence, we can write

$$\lambda_t = \lambda^\star_t-\mathrm{Cov}_t[(x_{t+1}-x^\star_{t+1}),(y_{t+1}-y^\star_{t+1})]-(\gamma^\star_t)^\top\alpha_t \equiv \lambda^\star_t+\delta_{nt}+\delta_{mt}. \tag{14}$$

In other words, the economic risk premia λ_t equal the premia on the maximum-correlation portfolios ($\lambda^\star_t$), plus two components: the negative of the covariance between the non-traded components of x_{t+1} and of y_{t+1} (δ_{nt}), and the negative of the mispricing of the mimicking portfolios by the candidate kernel x_{t+1} (δ_{mt}). Note that, holding the set of available assets constant, the first component is constant across asset pricing models.

The decomposition above highlights two interpretations of the risk premia λ_t. First, they are the risk premia a factor mimicking portfolio would have, if the factors were spanned, i.e., if the factors were traded, and if the candidate kernel priced all assets correctly. Alternatively, if the factors are not traded, the risk premia λ_t are essentially "fictional" quantities, assigned to the factor partly because of observed asset returns, but also partly because of the covariance between the non-traded component of the factor and the non-traded component of the pricing kernel, and partly because of the mispricing of the corresponding maximum-correlation mimicking portfolio.

2.3. The role of mispricing

It is important to note that the determination of the $\lambda^\star_t$ and δ_{nt} components does not hinge on the mispricing component δ_{mt}: $\lambda^\star_t$ depends only on the properties of security returns, and not on the properties of the kernel; while δ_{nt} depends only on the non-traded variability of the kernel, which does not affect the pricing properties of the kernel.

The point above can be illustrated if we "adjust" a candidate kernel to make it admissible by adding the quantity $q^\star_{t+1}-x^\star_{t+1}$. This quantity affects the pricing properties of the kernel, but does not affect its non-traded variability. In fact, this is the minimum adjustment, in terms of volatility, that makes x_{t+1} admissible: the variance of $q^\star_{t+1}-x^\star_{t+1}$ is the Hansen–Jagannathan distance of Hansen and Jagannathan (1997) (see Kan and Robotti, 2008, for a discussion of the Hansen–Jagannathan, 1997, distance in the case of *excess* returns). We define the adjusted kernel as

$$x^a_{t+1} \equiv x_{t+1}+(q^\star_{t+1}-x^\star_{t+1}). \tag{15}$$

The risk premium λ^a_t assigned by the adjusted kernel x^a_{t+1} equals

$$\begin{aligned}\lambda^a_t &= \lambda_t-\mathrm{Cov}_t[(q^\star_{t+1}-x^\star_{t+1}),y_{t+1}]\\ &= \lambda_t-\mathrm{Cov}_t(q^\star_{t+1},y^\star_{t+1})+\mathrm{Cov}_t(x^\star_{t+1},y^\star_{t+1})\\ &= \lambda_t+(\gamma^\star_t)^\top E_t(r_{t+1})-(\gamma^\star_t)^\top[E_t(r_{t+1})-\alpha_t]\\ &= \lambda_t+(\gamma^\star_t)^\top\alpha_t\\ &= \lambda^\star_t-\mathrm{Cov}_t[(x_{t+1}-x^\star_{t+1}),(y_{t+1}-y^\star_{t+1})]-(\gamma^\star_t)^\top\alpha_t+(\gamma^\star_t)^\top\alpha_t\\ &= \lambda^\star_t+\delta_{nt},\end{aligned} \tag{16}$$

and δ_{mt} disappears, while $\lambda^\star_t$ and δ_{nt} are the same as for the original candidate kernel.

Alternatively, instead of adjusting the pricing kernel, we can adjust the excess returns by subtracting the vector α_t (this is the approach used in the analysis under the null of Section 6.1):

$$r^a_{t+1} \equiv r_{t+1}-\alpha_t. \tag{17}$$

In this case, λ_t^a equals

$$\begin{aligned}\lambda_t^a &= (\gamma_t^\star)^\top E_t(r_{t+1}^a) - \mathrm{Cov}_t[(x_{t+1} - x_{t+1}^\star), (y_{t+1} - y_{t+1}^\star)] \\ &= (\gamma_t^\star)^\top [E_t(r_{t+1}) - \alpha_t] + \delta_{nt} \\ &= \lambda_t.\end{aligned} \tag{18}$$

In other words, if we adjust the excess returns to make them consistent with a candidate pricing kernel, the total risk premium and the non-traded component are unaffected, while the mispricing component becomes part of the traded component.

2.4. Linear factor models

In the case where the asset pricing model is a linear factor model with non-traded factors, the corresponding candidate pricing kernel is

$$x_{t+1} = 1 - [y_{t+1} - E_t(y_{t+1})]^\top b_t, \tag{19}$$

where b_t is a set of time-varying coefficients. Note that if we substitute the linear kernel x_{t+1} in Eq. (6) we obtain

$$\lambda_t = \Sigma_{yyt} b_t, \tag{20}$$

where Σ_{yyt} is the conditional covariance matrix of the factors. Hence, we have

$$b_t = \Sigma_{yyt}^{-1} \lambda_t, \tag{21}$$

which highlights the relation between the coefficients of the pricing kernel and the economic risk premia assigned by x_{t+1}. Moreover, we have

$$E_t(r_{t+1}) = \alpha_t + \beta_t^\top \lambda_t, \tag{22}$$

where $\beta_t = \Sigma_{yyt}^{-1} \Sigma_{yrt}$ and Σ_{yrt} is the conditional cross-covariance matrix between y_{t+1} and r_{t+1}. If the linear factor model holds, $\alpha_t = 0$, and all security risk premia are linear combinations of the economic risk premia.

2.5. λ_t and $\lambda_t^\star$: relative size

As anticipated in the Introduction, one of the results of the empirical analysis is that the estimates of the economic risk premia assigned by linear factor models with non-traded factors are often unrealistically large. Hou and Kimmel (2006) and Balduzzi and Robotti (2008) derive the relationship between λ_t and $\lambda_t^\star$ in a multivariate setting. Here, we focus on the special case of a one-factor model that prices the factor mimicking portfolio *exactly*. In this specialized setting, we can easily characterize the relative size of λ_t and $\lambda_t^\star$. In fact, the result in Eq. (14) simplifies to

$$\lambda_t = \lambda_t^\star + (\sigma_{yt}^2 - \sigma_{y^\star t}^2) b_t = \lambda_t^\star + \frac{\sigma_{yt}^2 - \sigma_{y^\star t}^2}{\sigma_{yt}^2} \lambda_t, \tag{23}$$

where σ_{yt}^2 and $\sigma_{y^\star t}^2$ are the conditional variances of the factor and of the factor mimicking portfolio, respectively. The expression above can be rewritten as

$$\lambda_t = \lambda_t^\star \frac{\sigma_{yt}^2}{\sigma_{y^\star t}^2}. \tag{24}$$

Hence, λ_t is always larger than $\lambda_t^\star$ in absolute value, and the discrepancy between the two measures of the risk premium increases as the R^2 of the projection of the factor onto the augmented span of excess returns decreases.

2.6. Measurement error

Assume that we add to y_{t+1} a mean-zero measurement error, e_{t+1}, that is uncorrelated with factors and asset returns. Consider now the premium assigned to the noisy factors, $y_{t+1}^n = y_{t+1} + e_{t+1}$,

$$\lambda_t^n \equiv -\mathrm{Cov}_t(x_{t+1}, y_{t+1}^n) = -\mathrm{Cov}_t(x_{t+1}, y_{t+1}) - \mathrm{Cov}_t(x_{t+1}, e_{t+1}) = \lambda_t - \mathrm{Cov}_t(x_{t+1}, e_{t+1}). \tag{25}$$

If x_{t+1} depends on e_{t+1} through the noisy factors y^n_{t+1}, then there is the potential for λ^n_t and λ_t to differ considerably, without any real underlying economic reason.

Indeed, for simplicity, consider the case where there is only one factor driving a linear kernel, where the kernel depends on y^n_{t+1}, rather than on y_{t+1}:

$$x^n_{t+1} = 1 - \frac{y^n_{t+1} - E_t(y^n_{t+1})}{\sigma^2_{yt}}\lambda_t. \tag{26}$$

It follows that

$$\lambda^n_t = \left(1 + \frac{\sigma^2_{et}}{\sigma^2_{yt}}\right)\lambda_t, \tag{27}$$

where σ^2_{et} is the conditional variance of the measurement error. As σ^2_{et} increases, λ^n_t increases in absolute value, again without any change in the economic fundamentals.

Moreover, even if we standardize the factor so that λ^n_t has the dimension of a theoretical Sharpe ratio, we have

$$\frac{\lambda^n_t}{\sqrt{\sigma^2_{yt} + \sigma^2_{et}}} = \frac{1 + \sigma^2_{et} / \sigma^2_{yt}}{\sqrt{\sigma^2_{yt} + \sigma^2_{et}}}\lambda_t. \tag{28}$$

It is easy to see that the derivative of $(1 + \sigma^2_{et} / \sigma^2_{yt}) / \sqrt{\sigma^2_{yt} + \sigma^2_{et}}$ w.r.t. σ^2_{et} is positive. Hence, even the theoretical Sharpe ratio on the noisy factors increases (in absolute value) as the variance of the measurement error in the factor increases.

On the contrary, the risk premium on the maximum-correlation portfolio is unaffected by measurement error since

$$-\mathrm{Cov}_t(q^\star_{t+1}, y^n_{t+1}) = -\mathrm{Cov}_t(q^\star_{t+1}, y_{t+1}) = \lambda^\star_t. \tag{29}$$

Moreover, going back to the original decomposition, if we use the pricing kernel x^n_{t+1} to assign risk premia, we obtain

$$\lambda^n_t = \lambda^\star_t - \mathrm{Cov}_t[(x^n_{t+1} - x^\star_{t+1}), (y^n_{t+1} - y^\star_{t+1})] - (\gamma^\star_t)^\top \alpha_t. \tag{30}$$

While the first and third components are unaffected by measurement error, the second component is affected. Obviously, if we standardize λ^n_t by the standard deviation of the noisy factor to obtain a Sharpe ratio, and we increase the variability of the measurement error, the first and third components of the Sharpe ratio converge to zero, while the second component diverges to plus or minus infinity.

2.7. Bounds

To gain a better understanding of the different components of λ_t, we derive bounds in the spirit of the risk premia bounds of Balduzzi and Kallal (1997) and the good-deal pricing bounds of Cochrane and Saá-Requejo (2000). Specifically, by the Cauchy-Schwarz inequality, we have

$$|\delta_{nt}| \equiv |\mathrm{Cov}_t[(x_{t+1} - x^\star_{t+1}), (y_{t+1} - y^\star_{t+1})]| \leq \sqrt{\mathrm{Var}_t(x_{t+1} - x^\star_{t+1})}\sqrt{\mathrm{Var}_t(y_{t+1} - y^\star_{t+1})}. \tag{31}$$

Since $\mathrm{Var}_t(x_{t+1} - x^\star_{t+1}) = \mathrm{Var}_t(x_{t+1}) - \mathrm{Var}_t(x^\star_{t+1})$ and $\mathrm{Var}_t(y_{t+1} - y^\star_{t+1}) = \mathrm{Var}_t(y_{t+1}) - \mathrm{Var}_t(y^\star_{t+1})$, we can rewrite the inequality above as

$$|\delta_{nt}| \leq \sqrt{\mathrm{Var}_t(x_{t+1}) - \mathrm{Var}_t(x^\star_{t+1})}\sqrt{\mathrm{Var}_t(y_{t+1}) - \mathrm{Var}_t(y^\star_{t+1})}. \tag{32}$$

Hence, the variability of a kernel, in excess of the variability of the projection of the kernel onto the augmented span of asset returns, places upper and lower bounds on δ_n. If the candidate kernel is very volatile relative to its projection, then the range of possible values of δ_{nt} and λ_t is wide. If, on the other hand, $\mathrm{Var}_t(x_{t+1}) - \mathrm{Var}_t(x^\star_{t+1}) = 0$, $\delta_{nt} = 0$ and $\lambda_t = \lambda^\star_t + \delta_{mt}$.

Using the approach above, we can also place bounds on the mispricing component of the risk premium, δ_{mt}. In this case, we have

$$|\delta_{mt}| \leq \sqrt{\mathrm{Var}_t(q^\star_{t+1} - x^\star_{t+1})}\sqrt{\mathrm{Var}_t(y^\star_{t+1})}, \tag{33}$$

where $\sqrt{\mathrm{Var}_t(q^\star_{t+1} - x^\star_{t+1})}$ is the conditional Hansen–Jagannathan distance statistic. The bounds in Eqs. (32) and (33) highlight the origin of the departures of λ_t from $\lambda^\star_t$. For given non-traded and traded variability of a factor, the potential for λ_t and $\lambda^\star_t$ to differ increases with a model's non-traded variability and mispricing.

3. Estimation

We focus on linear factor models. Hence, our pricing kernels are linear in the factors. This allows us to identify all parameters in closed form, based on a set of moment conditions that includes the standard asset pricing equations. In the following, we describe the details of the estimation procedure.

3.1. Traded and non-traded factors

In the general case, both traded and non-traded factors drive a candidate kernel, and the factors whose premia we want to estimate may not be the same non-traded factors driving the candidate kernel. Hence, we consider *three* sets of risk factors. The first set of K_1 factors, $y_{1,t+1}$, are traded factors, i.e., specific excess security returns driving a candidate kernel, such as the excess return on the market portfolio. The second set of K_2 factors, $y_{2,t+1}$, are the *innovations* in the non-traded risk factors driving a candidate pricing kernel, such as unexpected consumption growth. The third set of K_3 factors, $y_{3,t+1}$, are the *innovations* in the non-traded risk factors whose risk premia we want to estimate, such as unexpected inflation. In the empirical analysis, $y_{2,t+1}$ and $y_{3,t+1}$ are the mean-zero residuals of a VAR(1) system of nine state variables that include the six non-traded factors and the three traded factors described in Section 4.2. These residuals are standardized to have unit variance.

3.2. Conditioning information and conditional variation

Denote by Z_t the $(J\times 1)$ vector of instruments $(1\ z_t)^\top$, where the z_t are demeaned and standardized to have unit variance. In the interest of parsimony, the z_t in the analysis below represent a subset of the nine state variables used to form innovations in the non-traded factors, i.e., we only consider the lagged realizations of the market excess return, the dividend yield, and the real T-bill rate. We assume that expected returns and expected factors are linear functions of Z_t, while conditional variances and covariances are constant: $\Sigma_{rrt}=\Sigma_{rr}$, $\Sigma_{yyt}=\Sigma_{yy}$, and $\Sigma_{ryt}=\Sigma_{ry}$, where Σ_{ryt} is the conditional covariance matrix between returns and factors. This approach to incorporating conditioning information is used, for instance, in Campbell and Viceira (1996), DeRoon et al. (1998) and DeRoon and Nijman (2001). While the assumption of constant conditional second moments is less than fully realistic, it allows us to derive convenient expressions for the parameter estimates.[3]

3.3. Minimum-variance kernel and $\lambda_t^\star$

The (conditionally) MV kernel $q_{t+1}^\star$ has the form

$$q_{t+1}^\star = 1-[r_{t+1}-E_t(r_{t+1})]^\top b_{rt}, \tag{34}$$

where $E_t(r_{t+1})=\mu_r^\top Z_t$, $b_{rt}=b_r^\top Z_t$, and $\lambda_t^\star=(\lambda^\star)^\top Z_t$.

We have the following sets of moment conditions:

$$E[(r_{t+1}-\mu_r^\top Z_t)Z_t^\top]=0, \tag{35}$$

$$E(r_{t+1}q_{t+1}^\star Z_t^\top)=0, \tag{36}$$

$$E\{[(1-q_{t+1}^\star)y_{3,t+1}-\lambda_t^\star]Z_t^\top\}=0. \tag{37}$$

The first set of moments identifies the parameters of the conditional means of excess returns. The second set of moments requires that $q_{t+1}^\star$ prices excess returns, conditional on the realization of Z_t. The third set of moments identifies the parameters of the risk premia on the maximum-correlation portfolios. The resulting conditional risk premia are the projections of $(1-q_{t+1}^\star)y_{3,t+1}$ onto Z_t. Note that the alternative approach to estimating $\lambda^\star$ is to construct conditional maximum-correlation portfolios and to project their excess returns onto Z_t. This alternative approach leads to the same $\lambda^\star$ estimates.

The system of orthogonality conditions above is exactly identified. Indeed, we can use the assumption of constant conditional second moments and the law of iterated expectations to obtain closed-form expressions for the estimates of b_r and $\lambda^\star$ that satisfy Eqs. (35)–(37) *exactly*. We have

$$\begin{aligned} E(r_{t+1}q_{t+1}^\star Z_t^\top) &= E\{E_t[r_{t+1}-r_{t+1}(r_{t+1}-E_t(r_{t+1}))^\top b_r^\top Z_t]Z_t^\top\} \\ &= E\{(\mu_r^\top Z_t-\Sigma_{rr}b_r^\top Z_t)Z_t^\top\} \\ &= E(\mu_r^\top-\Sigma_{rr}b_r^\top)E(Z_t^\top Z_t^\top) \\ &= 0. \end{aligned} \tag{38}$$

[3] While not reported in the paper, we also implemented an *unconditional* version of our empirical analysis, where the vector Z_t only contains the constant, while the factors $y_{2,t+1}$ and $y_{3,t+1}$ are the same standardized VAR(1) innovations as in the conditional analysis. The unconditional analysis does not require the specification of a VAR(1) to describe the law of motion of the instruments.

Postmultiplying both sides of the equality above by $E(Z_t^\top Z_t^\top)^{-1}$ leads to the estimate

$$\hat{b}_r = \hat{\mu}_r \hat{\Sigma}_{rr}^{-1}, \tag{39}$$

where $\hat{\Sigma}_{rr} = \hat{E}[(r_{t+1} - \hat{\mu}_r^\top Z_t)(r_{t+1} - \hat{\mu}_r^\top Z_t)^\top]$ and

$$\hat{\mu}_r = \hat{E}(Z_t Z_t^\top)^{-1} \hat{E}(Z_t r_{t+1}^\top). \tag{40}$$

Using a similar approach, we have

$$\begin{aligned} E\{[(1-q_{t+1}^\star)y_{3,t+1} - \lambda_t^\star]Z_t^\top\} &= E\{E_t[y_{3,t+1}(r_{t+1} - E_t(r_{t+1}))^\top b_r^\top Z_t - (\lambda^\star)^\top Z_t]Z_t^\top\} \\ &= [\Sigma_{y_3 r} b_r^\top - (\lambda^\star)^\top] E(Z_t Z_t^\top) \\ &= 0, \end{aligned} \tag{41}$$

where $\Sigma_{y_3 r} = E_t(y_{3,t+1} r_{t+1}^\top)$. Postmultiplying both sides of the equality above by $E(Z_t^\top Z_t^\top)^{-1}$ leads to the estimate

$$\hat{\lambda}^\star = \hat{b}_r \hat{\Sigma}_{ry_3}, \tag{42}$$

where $\hat{\Sigma}_{ry_3} = \hat{E}(r_{t+1} y_{3,t+1}^\top)$.

3.4. Candidate kernel and λ_t: Non-traded factors only

We illustrate the estimation procedure for the case where the candidate kernel depends only on the non-traded factors $y_{2,t+1}$. The cases where the candidate kernel depends on traded factors only, and on both traded and non-traded factors are discussed in the Appendix.

The candidate kernel x_{t+1} has the form

$$x_{t+1} = 1 - y_{2,t+1}^\top b_{2t}, \tag{43}$$

where $b_{2t} = b_2^\top Z_t$. Note that $E_t(x_{t+1}) = 1$, as $y_{2,t+1}$ are the mean-zero innovations defined earlier.

In all cases, we assume that the risk premia are time varying and linear in Z_t: $\lambda_t = \lambda^\top Z_t$. These are natural assumptions, since we have assumed that conditional expected excess returns are linear in the instruments, and since we have ruled out time variation in the second conditional moments.

We have the following sets of moment conditions:

$$E(r_{t+1} x_{t+1} Z_t^\top) = 0, \tag{44}$$

$$E\{[(1-x_{t+1})y_{3,t+1} - \lambda_t]Z_t^\top\} = 0. \tag{45}$$

The first set of moments requires that x_{t+1} prices excess returns, conditional on the realization of Z_t. The second set of moments identifies the parameters λ, which correspond to the coefficients of a projection of $(1-x_{t+1})y_{3,t+1}$ onto Z_t. Note that the system above is overidentified. Hence, in order to obtain closed-form expressions for the parameter estimates, we replace Eq. (44) with a linear combination of the original asset pricing equations $E(r_{t+1} x_{t+1} Z_t^\top) = 0$, i.e., we set

$$\Sigma_{y_2 r} W E(r_{t+1} x_{t+1} Z_t^\top) = 0, \tag{46}$$

where $\Sigma_{y_2 r} = E_t(y_{2,t+1} r_{t+1}^\top)$ and W is a positive-definite symmetric weighting matrix. The equation above requires that K_2 linear combinations of the asset pricing equations are set to zero. We choose the matrix $\Sigma_{y_2 r} W$ to construct the linear combination of the asset pricing equations that we set to zero because this leads to estimates of b_2 that have a familiar interpretation: they are the result of running a cross-sectional regression (CSR) of the estimated expected excess returns on the covariances between returns and factors (see, for example, Cochrane, 1996). We consider two choices of W: $W = I$ and $W = \hat{\Sigma}_{rr}^{-1}$. Choosing $W = I$ means that the researcher is interested in ordinary least squares (OLS) CSRs. Jagannathan and Wang (1996) and Lewellen, Nagel and Shanken (2009) argue that OLS CSRs and related goodness-of-fit measures can be economically meaningful if the objective is to model the expected returns for a particular set of assets. On the other hand, setting $W = \hat{\Sigma}_{rr}^{-1}$ means that the researcher is interested in generalized least squares (GLS) CSRs. The reason for considering this second choice of weighting matrix is that one may want to downweigh the orthogonality conditions associated with the returns displaying the higher variability. In addition, Kandel and Stambaugh (1995) argue that GLS CSRs are meaningful if we are interested in determining the relative efficiency of an index.

With the system of orthogonality conditions above, and using the assumption of constant conditional second moments and the law of iterated expectations, we obtain estimates of b_r and $\lambda^\star$ in closed form. We have

$$\begin{aligned}\Sigma_{y_2 r}WE(r_{t+1}x_{t+1}Z_t^\top) &= \Sigma_{y_2 r}WE[E_t(r_{t+1}-r_{t+1}y_{2,t+1}^\top b_2^\top Z_t)Z_t^\top] \\ &= \Sigma_{y_2 r}WE[(\mu_r^\top Z_t-\Sigma_{ry_2}b_2^\top Z_t)Z_t^\top] \\ &= \Sigma_{y_2 r}W(\mu_r^\top-\Sigma_{ry_2}b_2^\top)E(Z_tZ_t^\top) \\ &= 0.\end{aligned} \tag{47}$$

Premultiplying both sides of the equation above by $(\Sigma_{y_2 r}W\Sigma_{ry2})^{-1}$ and postmultiplying both sides by $E(Z_tZ_t^\top)^{-1}$ leads to

$$\hat{b}_2 = \hat{\mu}_r W\hat{\Sigma}_{ry_2}(\hat{\Sigma}_{y_2 r}W\hat{\Sigma}_{ry_2})^{-1}, \tag{48}$$

where $\hat{\Sigma}_{ry2}=\hat{\mathrm{E}}(r_{t+1}y_{2,t+1}^\top)$. We also have

$$\begin{aligned}E\{[(1-x_{t+1})y_{3,t+1}-\lambda_t]Z_t^\top\} &= E[E_t(y_{3,t+1}y_{2,t+1}^\top b_2^\top Z_t-\lambda^\top Z_t)Z_t^\top] \\ &= (\Sigma_{y_3y_2}b_2^\top-\lambda^\top)E(Z_tZ_t^\top) \\ &= 0,\end{aligned} \tag{49}$$

where $\Sigma_{y_3y_2}\ E_t(y_{3,t+1}y_{2,t+1}^\top)$. Postmultiplying both sides of the equation above by $E(Z_tZ_t^\top)^{-1}$ leads to

$$\hat{\lambda} = \hat{b}_2\hat{\Sigma}_{y_2y_3}, \tag{50}$$

where $\hat{\Sigma}_{y_2y_3}=\hat{E}(y_{2,t+1}y_{3,t+1}^\top)$.

As pointed out earlier, $\hat{\mathrm{b}}_{2t}$ corresponds to the vector of coefficients of a cross-sectional regression of the estimated conditional expected excess returns on the estimated conditional covariances of returns with the factors:

$$\hat{b}_{2t} = (\hat{\Sigma}_{y_2 r}W\hat{\Sigma}_{ry_2}^\top)^{-1}\hat{\Sigma}_{y_2 r}W(\hat{\mu}_r^\top Z_t). \tag{51}$$

The projection of x_{t+1} onto the augmented span of excess returns leads to

$$x_{t+1}^\star = 1-[y_{2,t+1}^\star-E_t(y_{2,t+1}^\star)]^\top b_{2t}, \tag{52}$$

where $y_{2,t+1}^\star=(\gamma_2^\star)^\top r_{t+1}$ satisfies the set of moment conditions

$$E[(y_{2,t+1}-y_{2,t+1}^\star)r_{t+1}^\top] = 0. \tag{53}$$

We obtain

$$\hat{\gamma}_2^\star = \hat{\Sigma}_{rr}^{-1}\hat{\Sigma}_{ry_2}. \tag{54}$$

Finally, the non-traded and mispricing components of the risk premia, δ_{nt} and δ_{mt}, are the conditional expectations of $(x_{t+1}^\star-x_{t+1})y_{3,t+1}$ and $(q_{t+1}^\star-x_{t+1}^\star)y_{3,t+1}$, respectively. Hence, it is natural to assume that the δ_{nt} and δ_{mt} components are linear functions of the instruments Z_t: $\delta_{nt}=\delta_n^\top Z_t$ and $\delta_{mt}=\delta_m^\top Z_t$. Specifically, δ_n and δ_m are the parameters of the conditional projections of $(x_{t+1}^\star-x_{t+1})y_{3,t+1}$ and $(q_{t+1}^\star-x_{t+1}^\star)y_{3,t+1}$ onto Z_t. The moment conditions identifying the parameters δ_n and δ_m are

$$E\{[(x_{t+1}^\star-x_{t+1})y_{3,t+1}-\delta_n^\top Z_t]Z_t^\top\} = 0, \tag{55}$$

$$E\{[(q_{t+1}^\star-x_{t+1}^\star)y_{3,t+1}-\delta_m^\top Z_t]Z_t^\top\} = 0. \tag{56}$$

We have

$$\begin{aligned}E\{[(x_{t+1}^\star-x_{t+1})y_{3,t+1}-\delta_n^\top Z_t]Z_t^\top\} &= E\{E_t[y_{3,t+1}(y_{2,t+1}-(y_{2,t+1}^\star-E_t(y_{2,t+1}^\star)))^\top b_2^\top Z_t-\delta_n^\top Z_t]Z_t^\top\} \\ &= E\{E_t[y_{3,t+1}(y_{2,t+1}-(\gamma_2^\star)^\top(r_{t+1}-E_t(r_{t+1})))^\top b_2^\top Z_t-\delta_n^\top Z_t]Z_t^\top\} \\ &= E\{[(\Sigma_{y_3y_2}-\Sigma_{y_3 r}\gamma_2^\star)b_2^\top Z_t-\delta_n^\top Z_t]Z_t^\top\} \\ &= [(\Sigma_{y_3y_2}-\Sigma_{y_3 r}\gamma_2^\star)b_2^\top-\delta_n^\top]E(Z_tZ_t^\top) \\ &= 0.\end{aligned} \tag{57}$$

Postmultiplying both sides of the equation above by $E(Z_t Z_t^\top)^{-1}$, we obtain the following estimate:

$$\hat{\delta}_n = \hat{b}_2[\hat{\Sigma}_{y_2 y_3} - (\hat{\gamma}_2^\star)^\top \hat{\Sigma}_{ry_3}]. \tag{58}$$

Similarly, we have

$$\begin{aligned} E\{[(q_{t+1}^\star - x_{t+1}^\star)y_{3,t+1}]Z_t^\top - \delta_m^\top Z_t Z_t^\top\} &= E\{E_t[y_{3,t+1}((y_{2,t+1}^\star - E_t(y_{2,t+1}^\star))^\top b_2^\top Z_t - (r_{t+1} - E_t(r_{t+1}))^\top b_r^\top Z_t)]Z_t^\top\} - \delta_m^\top E(Z_t Z_t^\top) \\ &= E\{E_t[y_{3,t+1}((r_{t+1} - E_t(r_{t+1}))^\top \gamma_2^\star b_2^\top Z_t - (r_{t+1} - E_t(r_{t+1}))^\top b_r^\top Z_t)]Z_t^\top\} - \delta_m^\top E(Z_t Z_t^\top) \\ &= E[\Sigma_{y_3 r}(\gamma_2^\star b_2^\top - b_r^\top) Z_t Z_t^\top] - \delta_m^\top E(Z_t Z_t^\top) \\ &= [\Sigma_{y_3 r}(\gamma_2^\star b_2^\top - b_r^\top) - \delta_m^\top] E(Z_t Z_t^\top) \\ &= 0. \end{aligned} \tag{59}$$

Postmultiplying both sides of the equation above by $E(Z_t Z_t^\top)^{-1}$, we obtain the following estimate:

$$\hat{\delta}_m = [\hat{b}_2(\hat{\gamma}_2^\star)^\top - \hat{b}_r]\hat{\Sigma}_{ry_3}. \tag{60}$$

Note that when $W = \hat{\Sigma}_{rr}^{-1}$, we have

$$\hat{\delta}_m = [\hat{\mu}_r \hat{\Sigma}_{rr}^{-1} \hat{\Sigma}_{ry_2}(\hat{\Sigma}_{y_2 r} \hat{\Sigma}_{rr}^{-1} \hat{\Sigma}_{ry_2})^{-1} \hat{\Sigma}_{y_2 r} \hat{\Sigma}_{rr}^{-1} - \hat{\mu}_r \hat{\Sigma}_{rr}^{-1}] \hat{\Sigma}_{ry_3}. \tag{61}$$

If we focus on the risk premia on the economic factors driving the candidate kernel, we have $y_{3,t+1} = y_{2,t+1}$, $\hat{\Sigma}_{ry_3} = \hat{\Sigma}_{ry_2}$, and $\hat{\delta}_m = 0$. In other words, the maximum-correlation portfolios mimicking the factors $y_{2,t+1}$ are priced *exactly*.

4. Data

This section illustrates the data used in the empirical analysis. The period considered is March 1959 to December 2002 for test assets and economic variables, and February 1959 to November 2002 for the conditioning variables. The choice of the starting date is dictated by macroeconomic data availability.

4.1. Asset returns

We use decile portfolio returns on NYSE, AMEX, and NASDAQ listed stocks from the Center for Research in Security Prices (CRSP). Ten size stock portfolios are formed according to size deciles on the basis of the market value of equity outstanding at the end of the previous year. If a market capitalization was not available for the previous year, the firm was ranked based on the capitalization on the date with the earliest available price in the current year. The returns are value-weighted averages of the firms' returns, adjusted for dividends. The securities with the smallest capitalizations are placed in portfolio one. The partitions on the CRSP file include all securities, excluding ADRs, which were active on NYSE–AMEX–NASDAQ for that year. In addition to returns on size-sorted equity portfolios, we also use returns on the 25 size and book-to-market sorted portfolios, the "Fama–French" portfolios, available from Ken French's website. Results for this alternative choice of assets are briefly summarized in a series of footnotes and are available upon request.

All rates of return are in excess of the risk-free rate. The risk-free rate proxy is the 1-month T-bill rate from Ibbotson Associates and pertains to a bill with at least 1 month to maturity.

4.2. Economic variables and instruments

We concentrate on a set of six non-traded variables, which have been previously used in tests of linear factor models and/or in studies of stock-return predictability (see, for example, Chen et al., 1986; Burmeister and McElroy, 1988; McElroy and Burmeister, 1988; Ferson and Harvey, 1991, 1999; Kirby, 1998). These variables are either statistically significant in multivariate predictive regressions of means and volatilities or they have special economic significance.

HB3 is the 1-month return on a 3-month T-bill less the 1-month return on a 1-month bill (from CRSP, Fama Treasury Bill Term Structure Files).
DIV denotes the monthly dividend yield on the Standard and Poor's 500 stock index (from Global Insight).
CG denotes the logarithm of the monthly gross growth rate of per capita real consumption of nondurable goods and services, percent per month. Per capita quantities are obtained by using data on total resident population. The series used to construct consumption data are from Global Insight.
PREM represents the yield spread between Baa and Aaa rated bonds (from Global Insight).
INF is the monthly rate of inflation (from Ibbotson Associates).
REALTB denotes the real 1-month T-bill (from Ibbotson Associates).

In addition, we consider three traded factors:

XVW represents the value-weighted NYSE–AMEX-NASDAQ index return in excess of the risk-free rate, percent per month.
SMB (Small Minus Big) represents the average return on three small portfolios (small value, small neutral, and small growth) minus the average return on three big portfolios (big value, big neutral, and big growth), percent per month.
HML (High Minus Low) represents the average return on two value portfolios (small value and big value) minus the average return on two growth portfolios (small growth and big growth), percent per month. The three benchmark factors XVW, SMB, and HML are from Kenneth French's website.

Table 1
λ vs $\lambda^{\star}$.

	$\lambda_0^{\star}$ (λ_0)	$\lambda_{XVW}^{\star}$ (λ_{XVW})	$\lambda_{DIV}^{\star}$ (λ_{DIV})	$\lambda_{REALTB}^{\star}$(λ_{REALTB})
Panel A: I-CAPM				
$\lambda^{\star}$(HB3)	−0.0216	−0.0421	−0.0039	−0.0073
λ(HB3)	−3.0554(OLS)	−2.0962(OLS)	−1.4597(OLS)	−0.6159(OLS)
	−2.2015(GLS)	−1.1737(GLS)	−0.8957(GLS)	−0.2563(GLS)
$Bias_{\lambda^{\star}}$ (HB3)	−0.0007	−0.0010	−0.0007	−0.0005
$Bias_{\lambda}$ (HB3)	2.4037(OLS)	1.0330(OLS)	1.2655(OLS)	0.4246(OLS)
	1.5931(GLS)	0.3764(GLS)	0.7694(GLS)	0.1415(GLS)
$Emp.Std_{\lambda^{\star}}$(HB3)	0.0139	0.0190	0.0124	0.0123
$Emp.Std_{\lambda}$(HB3)	1.3096(OLS)	1.2768(OLS)	0.9535(OLS)	0.6568(OLS)
	0.8503(GLS)	1.1808(GLS)	0.6539(GLS)	0.6010(GLS)
$\lambda^{\star}$(DIV)	−0.0629	−0.0412	−0.0446	−0.0672
λ(DIV)	0.4252(OLS)	0.6479(OLS)	0.5860(OLS)	−0.2608(OLS)
	0.2323(GLS)	0.3179(GLS)	0.5214(GLS)	−0.2784(GLS)
$Bias_{\lambda^{\star}}$(DIV)	−0.0005	0.0008	−0.0136	−0.0004
$Bias_{\lambda}$(DIV)	−0.5402(OLS)	−1.0684(OLS)	−0.4567(OLS)	0.1259(OLS)
	−0.3300(GLS)	−0.5746(GLS)	−0.3363(GLS)	0.0943(GLS)
$Emp.Std_{\lambda^{\star}}$(DIV)	0.0324	0.0328	0.0307	0.0314
$Emp.Std_{\lambda}$(DIV)	0.7950(OLS)	0.7807(OLS)	0.5860(OLS)	0.4046(OLS)
	0.5005(GLS)	0.6886(GLS)	0.3890(GLS)	0.3604(GLS)
$\lambda^{\star}$(CG)	0.0103	0.0274	0.0093	−0.0042
λ(CG)	−0.4768(OLS)	0.2913(OLS)	−0.4412(OLS)	−0.7105(OLS)
	−0.1734(GLS)	0.3738(GLS)	−0.1561(GLS)	−0.4856(GLS)
$Bias_{\lambda^{\star}}$(CG)	0.0001	0.0001	0.0025	0.0003
$Bias_{\lambda}$(CG)	0.2848(OLS)	−0.0312(OLS)	0.2950(OLS)	0.3001(OLS)
	0.0929(GLS)	−0.2068(GLS)	0.0979(GLS)	0.1203(GLS)
$Emp.Std_{\lambda^{\star}}$(CG)	0.0164	0.0209	0.0145	0.0146
$Emp.Std_{\lambda}$(CG)	0.8859(OLS)	0.8954(OLS)	0.6557(OLS)	0.4675(OLS)
	0.5794(GLS)	0.8094(GLS)	0.4514(GLS)	0.4264(GLS)
$\lambda^{\star}$(PREM)	0.0045	−0.0166	0.0066	0.0053
λ(PREM)	−2.1647(OLS)	−5.9491(OLS)	−1.2736(OLS)	0.2820(OLS)
	−1.3807(GLS)	−4.1463(GLS)	−1.0028(GLS)	0.0968(GLS)
$Bias_{\lambda^{\star}}$(PREM)	0.0007	−0.0000	0.0014	−0.0004
$Bias_{\lambda}$(PREM)	2.0732(OLS)	5.2462(OLS)	1.2655(OLS)	−0.3040(OLS)
	1.3712(GLS)	3.3239(GLS)	0.9415(GLS)	−0.1188(GLS)
$Emp.Std_{\lambda^{\star}}$(PREM)	0.0124	0.0176	0.0115	0.0107
$Emp.Std_{\lambda}$(PREM)	1.8206(OLS)	1.6574(OLS)	1.3272(OLS)	0.8711(OLS)
	1.1451(GLS)	1.5768(GLS)	0.8912(GLS)	0.8075(GLS)
$\lambda^{\star}$(INF)	0.0036	0.0247	−0.0185	−0.0011
λ(INF)	−1.5613(OLS)	−0.4678(OLS)	−1.2771(OLS)	0.0117(OLS)
	−1.3807(GLS)	0.2607(GLS)	−0.9505(GLS)	0.1632
$Bias_{\lambda^{\star}}$(INF)	0.0006	0.0002	−0.0033	−0.0004
$Bias_{\lambda}$(INF)	2.0732(OLS)	1.2349(OLS)	0.9755(OLS)	0.2282(OLS)
	1.3712(GLS)	0.3685(GLS)	0.6730(GLS)	0.0528(GLS)
$Emp.Std_{\lambda^{\star}}$(INF)	0.0163	0.0206	0.0142	0.0139
$Emp.Std_{\lambda}$(INF)	1.0386(OLS)	1.0519(OLS)	0.7717(OLS)	0.5411(OLS)
	0.6947(GLS)	0.9673(GLS)	0.5326(GLS)	0.4962(GLS)
$\lambda^{\star}$(REALTB)	−0.0047	−0.0314	0.0139	−0.0009
λ(REALTB)	2.1774(OLS)	0.6017(OLS)	1.3183(OLS)	0.0135(OLS)
	1.6145(GLS)	−0.2562(GLS)	1.0097(GLS)	−0.0892(GLS)
$Bias_{\lambda^{\star}}$(REALTB)	−0.0006	−0.0003	0.0024	0.0004
$Bias_{\lambda}$(REALTB)	−2.4250(OLS)	−1.5365(OLS)	−1.1053(OLS)	−0.2234(OLS)
	−1.6754(GLS)	−0.5130(GLS)	−0.7809(GLS)	−0.0850(GLS)
$Emp.Std_{\lambda^{\star}}$(REALTB)	0.0150	0.0197	0.0133	0.0128
$Emp.Std_{\lambda}$(REALTB)	1.1874(OLS)	1.1852(OLS)	0.8790(OLS)	0.6026(OLS)
	0.7863(GLS)	1.0994(GLS)	0.6042(GLS)	0.5553(GLS)

(continued on next page)

As noted earlier, the factors $y_{2,t+1}$ and $y_{3,t+1}$ used in the analysis are mean-zero residuals from a VAR(1) that includes the nine variables described above. We select a constant and the *lagged* realizations of XVW, DIV, and REALTB as Z_t. The reason for using the lagged values of the economic variables as instruments is that, according to the I-CAPM intuition, the variables that drive asset returns should also be the variables affecting the risk-return trade-off, i.e., they should also be the variables predicting returns (see, for example, Campbell, 1996).

Since the factors are standardized to have unit standard deviation, the risk premia have the interpretation of Sharpe ratios. Note that in obtaining standard errors and *p*-values by bootstrap, we bootstrap the original series, not the standardizedseries. Hence, we do account for the sampling variability in the estimates of the standard deviations of the factors.

Table 1 (*continued*)

	λ_0^* (λ_0)	λ_{XVW}^* (λ_{XVW})	λ_{DIV}^* (λ_{DIV})	λ_{REALTB}^*(λ_{REALTB})
Panel B: C−CAPM				
λ^*(HB3)	−0.0216	−0.0421	−0.0039	−0.0073
λ(HB3)	0.0208(OLS)	0.0365(OLS)	0.0147(OLS)	0.0157(OLS)
	0.0044(GLS)	0.0117(GLS)	0.0040(GLS)	−0.0018(GLS)
Bias$_{\lambda^*}$(HB3)	−0.0007	−0.0010	−0.0007	−0.0005
Bias$_{\lambda}$ (HB3)	0.0007(OLS)	0.0041(OLS)	0.0052(OLS)	0.0026(OLS)
	−0.0028(GLS)	−0.0028(GLS)	−0.0005(GLS)	0.0005(GLS)
Emp.Std$_{\lambda^*}$(HB3)	0.0139	0.0190	0.0124	0.0123
Emp.Std$_{\lambda}$(HB3)	0.0436(OLS)	0.0739(OLS)	0.0395(OLS)	0.0369(OLS)
	0.0124(GLS)	0.0219(GLS)	0.0120(GLS)	0.0106(GLS)
λ^*(DIV)	−0.0629	−0.0412	−0.0446	−0.0672
λ(DIV)	−0.1536(OLS)	−0.2688(OLS)	−0.1081(OLS)	−0.1156(OLS)
	−0.0325(GLS)	−0.0864(GLS)	−0.0294(GLS)	0.0132(GLS)
Bias$_{\lambda^*}$(DIV)	−0.0005	0.0008	−0.0136	−0.0004
Bias$_{\lambda}$(DIV)	−0.0075(OLS)	−0.0083(OLS)	−0.0261(OLS)	−0.0079(OLS)
	0.0060(GLS)	0.0158(GLS)	−0.0015(GLS)	−0.0044(GLS)
Emp.Std$_{\lambda^*}$(DIV)	0.0324	0.0328	0.0307	0.0314
Emp.Std$_{\lambda}$(DIV)	0.0764(OLS)	0.0954(OLS)	0.0727(OLS)	0.0712(OLS)
	0.0423(GLS)	0.0554(GLS)	0.0383(GLS)	0.0374(GLS)
λ^*(CG)	0.0103	0.0274	0.0093	−0.0042
λ(CG)	0.07674(OLS)	1.3432(OLS)	0.5402(OLS)	0.5775(OLS)
	0.1626(GLS)	0.4315(GLS)	0.1467(GLS)	−0.0662(GLS)
Bias$_{\lambda^*}$(CG)	0.0001	0.0001	0.0025	0.0003
Bias$_{\lambda}$(CG)	0.0672(OLS)	0.0989(OLS)	0.1582(OLS)	0.0646(OLS)
	−0.0335(GLS)	−0.0724(GLS)	0.0043(GLS)	0.0141(GLS)
Emp.Std$_{\lambda^*}$(CG)	0.0164	0.0209	0.0145	0.0146
Emp.Std$_{\lambda}$(CG)	0.4465(OLS)	0.6218(OLS)	0.4260(OLS)	0.5774(OLS)
	0.2130(GLS)	0.2838(GLS)	0.1903(GLS)	0.1919(GLS)
λ^*(PREM)	0.0045	−0.0166	0.0066	0.0053
λ(PREM)	0.0083(OLS)	0.0146(OLS)	0.0059(OLS)	0.0063(OLS)
	0.0018(GLS)	0.0047(GLS)	0.0016(GLS)	−0.0007(GLS)
Bias$_{\lambda^*}$(PREM)	0.0007	−0.0000	0.0014	−0.0004
Bias$_{\lambda}$(PREM)	0.0022(OLS)	0.0017(OLS)	0.0023(OLS)	0.0012(OLS)
	−0.0009(GLS)	−0.0023(GLS)	0.0001(GLS)	−0.0000(GLS)
Emp.Std$_{\lambda^*}$(PREM)	0.0124	0.0176	0.0115	0.0107
Emp.Std$_{\lambda}$(PREM)	0.0494(OLS)	0.0815(OLS)	0.0427(OLS)	0.0395(OLS)
	0.0129(GLS)	0.0237(GLS)	0.0126(GLS)	0.0104(GLS)
λ^*(INF)	0.0036	0.0247	−0.0185	−0.0011
λ(INF)	−0.1355(OLS)	−0.2372(OLS)	−0.0954(OLS)	−0.1019(OLS)
	−0.0287(GLS)	−0.0762(GLS)	−0.0259(GLS)	0.0117(GLS)
Bias$_{\lambda^*}$(INF)	0.0006	0.0002	−0.0033	−0.0004
Bias$_{\lambda}$(INF)	−0.0072(OLS)	−0.0104(OLS)	−0.0247(OLS)	−0.0083(OLS)
	0.0073(GLS)	0.0148(GLS)	−0.0002(GLS)	−0.0035(GLS)
Emp.Std$_{\lambda^*}$(INF)	0.0163	0.0206	0.0142	0.0139
Emp.Std$_{\lambda}$(INF)	0.0814(OLS)	0.1179(OLS)	0.0778(OLS)	0.0751(OLS)
	0.0382(GLS)	0.0523(GLS)	0.0349(GLS)	0.0344(GLS)
λ^*(REALTB)	−0.0047	−0.0314	0.0139	−0.0009
λ(REALTB)	0.1340(OLS)	0.2346(OLS)	0.0944(OLS)	0.1009(OLS)
	0.0284(GLS)	0.0754(GLS)	0.0256(GLS)	−0.0116(GLS)
Bias$_{\lambda^*}$(REALTB)	−0.0006	−0.0003	0.0024	0.0004
Bias$_{\lambda}$(REALTB)	0.0079(OLS)	0.0119(OLS)	0.0252(OLS)	0.0090(OLS)
	−0.0073(GLS)	−0.0145(GLS)	0.0003(GLS)	0.0034(GLS)
Emp.Std$_{\lambda^*}$(REALTB)	0.0150	0.0197	0.0133	0.0128
Emp.Std$_{\lambda}$(REALTB)	0.0825(OLS)	0.1207(OLS)	0.0790(OLS)	0.0761(OLS)
	0.0381(GLS)	0.0523(GLS)	0.0348(GLS)	0.0344(GLS)

Table 1 (*continued*)

	$\lambda_0^\star$ (λ_0)	$\lambda_{XVW}^\star$ (λ_{XVW})	$\lambda_{DIV}^\star$ (λ_{DIV})	$\lambda_{REALTB}^\star$(λ_{REALTB})
Panel C: FF3				
λ^*(HB3)	−0.0216	−0.0421	−0.0039	−0.0073
λ(HB3)	−0.0363	−0.0252	−0.0113	−0.0083
$Bias_{\lambda^\star}$(HB3)	−0.0007	−0.0010	−0.0007	−0.0005
$Bias_{\lambda}$(HB3)	0.0028	0.0006	0.0002	−0.0002
$Emp.Std_{\lambda^\star}$(HB3)	0.0139	0.0190	0.0124	0.0123
$Emp.Std_{\lambda}$(HB3)	0.0140	0.0151	0.0107	0.0097
λ^*(DIV)	−0.0629	−0.0412	−0.0446	−0.0672
λ(DIV)	−0.0744	−0.0432	−0.0591	−0.0673
$Bias_{\lambda^\star}$(DIV)	−0.0005	0.0008	−0.0136	−0.0004
$Bias_{\lambda}$(DIV)	0.0118	0.0023	−0.0066	−0.0019
$Emp.Std_{\lambda^\star}$(DIV)	0.0324	0.0328	0.0307	0.0314
$Emp.Std_{\lambda}$(DIV)	0.0227	0.0321	0.0215	0.0314
λ^*(CG)	0.0103	0.0274	0.0093	−0.0042
λ(CG)	0.0257	0.0222	0.0177	0.0155
$Bias_{\lambda^\star}$(CG)	0.0001	0.0001	0.0025	0.0003
$Bias_{\lambda}$(CG)	−0.0035	0.0002	0.0012	0.0004
$Emp.Std_{\lambda^\star}$(CG)	0.0164	0.0209	0.0145	0.0146
$Emp.Std_{\lambda}$(CG)	0.0130	0.0144	0.0090	0.0095
λ^*(PREM)	0.0045	−0.0166	0.0066	0.0053
λ(PREM)	0.0171	−0.0080	−0.0011	0.0076
$Bias_{\lambda^\star}$(PREM)	0.0007	−0.0000	0.0014	−0.0004
$Bias_{\lambda}$(PREM)	0.0006	−0.0001	0.0013	−0.0001
$Emp.Std_{\lambda^\star}$(PREM)	0.0124	0.0176	0.0115	0.0107
$Emp.Std_{\lambda}$(PREM)	0.0140	0.0141	0.0089	0.0077
λ^*(INF)	0.0036	0.0247	−0.0185	−0.0011
Panel C: FF3				
λ(INF)	−0.0198	−0.0238	−0.0185	−0.0142
$Bias_{\lambda^\star}$(INF)	0.0006	0.0002	−0.0033	−0.0004
$Bias_{\lambda}$(INF)	0.0046	0.0010	−0.0006	−0.0011
$Emp.Std_{\lambda^\star}$(INF)	0.0163	0.0206	0.0142	0.0139
$Emp.Std_{\lambda}$(INF)	0.0139	0.0161	0.0099	0.0093
λ^*(REALTB)	−0.0047	−0.0314	0.0139	−0.0009
λ(REALTB)	0.0132	0.0130	0.0112	0.0097
$Bias_{\lambda^\star}$(REALTB)	−0.0006	−0.0003	0.0024	0.0004
$Bias_{\lambda}$(REALTB)	−0.0031	−0.0008	0.0006	0.0011
$Emp.Std_{\lambda^\star}$(REALTB)	0.0150	0.0197	0.0133	0.0128
$Emp.Std_{\lambda}$(REALTB)	0.0124	0.0151	0.0090	0.0077
Panel D: CAPM				
λ^*(HB3)	−0.0216	−0.0421	−0.0039	−0.0073
λ(HB3)	−0.0058	−0.0027	−0.0048	−0.0059
$Bias_{\lambda^\star}$(HB3)	−0.0007	−0.0010	−0.0007	−0.0005
$Bias_{\lambda}$(HB3)	0.0009	0.0004	−0.0007	−0.0002
$Emp.Std_{\lambda^\star}$(HB3)	0.0139	0.0190	0.0124	0.0123
$Emp.Std_{\lambda}$(HB3)	0.0049	0.0041	0.0055	0.0065
λ^*(DIV)	−0.0629	−0.0412	−0.0446	−0.0672
λ(DIV)	−0.0653	−0.0304	−0.0542	−0.0666
$Bias_{\lambda^\star}$(DIV)	−0.0005	0.0008	−0.0136	−0.0004
$Bias_{\lambda}$(DIV)	0.0102	0.0022	−0.0071	−0.0018
$Emp.Std_{\lambda^\star}$(DIV)	0.0324	0.0328	0.0307	0.0314
$Emp.Std_{\lambda}$(DIV)	0.0199	0.0306	0.0213	0.0312
λ^*(CG)	0.0103	0.0274	0.0093	−0.0042
λ(CG)	0.0144	0.0067	0.0120	0.0147
$Bias_{\lambda^\star}$(CG)	0.0001	0.0001	0.0025	0.0003
$Bias_{\lambda}$(CG)	−0.0023	−0.0005	0.0016	0.0004
$Emp.Std_{\lambda^\star}$(CG)	0.0164	0.0209	0.0145	0.0146
$Emp.Std_{\lambda}$(CG)	0.0056	0.0072	0.0061	0.0083
λ^*(PREM)	0.0045	−0.0166	0.0066	0.0053
λ(PREM)	0.0065	0.0030	0.0054	0.0067
$Bias_{\lambda^\star}$(PREM)	0.0007	−0.0000	0.0014	−0.0004
$Bias_{\lambda}$(PREM)	−0.0010	−0.0001	0.0006	0.0001
$Emp.Std_{\lambda^\star}$(PREM)	0.0124	0.0176	0.0115	0.0107
$Emp.Std_{\lambda}$(PREM)	0.0047	0.0043	0.0051	0.0061
λ^*(INF)	0.0036	0.0247	−0.0185	−0.0011
λ(INF)	−0.0135	−0.0063	−0.0112	−0.0137
$Bias_{\lambda^\star}$(*INF*)	0.0006	0.0002	−0.0033	−0.0004
$Bias_{\lambda}$(INF)	0.0021	0.0004	−0.0015	−0.0004
$Emp.Std_{\lambda^\star}$(INF)	0.0163	0.0206	0.0142	0.0139

(continued on next page)

Table 1 (*continued*)

	$\lambda_0^\star$ (λ_0)	$\lambda_{XVW}^\star$ (λ_{XVW})	$\lambda_{DIV}^\star$ (λ_{DIV})	$\lambda_{REALTB}^\star$(λ_{REALTB})
Panel D: CAPM				
Emp.Std$_\lambda$(INF)	0.0054	0.0069	0.0059	0.0080
$\lambda^\star$(REALTB)	−0.0047	−0.0314	0.0139	−0.0009
λ(REALTB)	0.0092	0.0043	0.0076	0.0094
Bias$_{\lambda\star}$(REALTB)	−0.0006	−0.0003	0.0024	0.0004
Bias$_\lambda$(REALTB)	−0.0014	−0.0003	0.0010	0.0003
Emp.Std$_{\lambda\star}$(REALTB)	0.0150	0.0197	0.0133	0.0128
Emp.Std$_\lambda$(REALTB)	0.0047	0.0051	0.0052	0.0067

We report estimates of λ_t and $\lambda_t^\star$ for the six economic factors: term structure (HB3), dividend yield (DIV), consumption growth (CG), default premium (PREM), inflation (INF), and real T-bill rate (REALTB). Time variation is driven by lagged values of the market excess return (XVW), the dividend yield, and the real T-bill rate. We consider four multi-beta models. The first model is the I-CAPM (Panel A), the second model is the C-CAPM (Panel B), the third model is the FF3 model (Panel C), and the fourth model is the CAPM (Panel D). The sample period is March 1959–December 2002. Finite-sample absolute biases (Bias) and empirical standard errors (Emp. Std.) are estimated by bootstrap (see Section 5.1 for details). For each economic factor, we report the average conditional risk premium, $\lambda_0^\star$ (λ_0), and the coefficients relating the risk premium to the three instruments: $\lambda_{XVW}^\star$ (λ_{XVW}), $\lambda_{DIV}^\star$ (λ_{DIV}), and $\lambda_{REALTB}^\star$ (λ_{REALTB}). For the models with non-traded factors, we consider estimates based on the two weighting matrices $W=I$ (OLS) and $W=\hat{\Sigma}_{rr}^{-1}$ (GLS).

The traded factors above are drivers of some of the pricing kernels, but they are not part of the set of test assets. The reason for this choice is that these factors – the return on the market, in particular – can be well-tracked by the returns on the test assets.[4] Hence, they add relatively little information to the test asset returns themselves.

5. Empirical analysis

We consider four linear factor models. The first model is the I-CAPM, with factors: market excess return, dividend yield, real T-bill rate, term structure, default premium, consumption growth, and inflation.[5] The second model is the C-CAPM, where the only factor driving the kernel is consumption growth. The third model is the Fama–French three-factor model (FF3). The fourth model is the CAPM, where the only factor driving the kernel is the excess market return.

We present our main results in three tables and two figures. Table 1 compares the estimates of the economic risk premia, λ, and of the premia on the maximum-correlation portfolios, $\lambda^\star$, across the four different models. Table 2 compares the non-traded components of the risk premia, δ_n, across the different models. Table 3 compares the mispricing components of the risk premia, δ_m, across the different models. Figs. 1 and 2 display the bounds on the mispricing and non-traded components of the premia, respectively.

5.1. Bootstrap procedure

For all parameter estimates, we compute small-sample bias and empirical standard errors by parametric bootstrap. The bootstrap exercise is structured as follows.

First, we estimate a VAR(1) for the nine non-traded and traded factors. Since the instruments are the lagged values of the factors, this gives us the law of motion of the factors and of the instruments predicting portfolio returns. Second, we estimate predictive regressions, by regressing the excess returns on Z_t. Third, we jointly bootstrap the residuals in the VAR(1) and in the predictive regressions.[6] The VAR(1) residuals are fed into the estimated law of motion of the economic variables to generate bootstrap samples for the traded and non-traded factors and for the instruments. Using the bootstrap realizations of the instruments and of the predictive-regression residuals, we generate bootstrap samples for excess returns. The initial values of the factors and instruments are the beginning-of-sample values for the corresponding variables. The exercise is repeated 100,000 times.[7]

5.2. Results

5.2.1. Risk premia: λ and $\lambda^\star$

Table 1 reports the average conditional risk premia estimates (λ_0 and $\lambda_0^\star$) and the estimated coefficients relating the risk premium to the market excess return (λ_{XVW} and $\lambda_{XVW}^\star$), the dividend yield (λ_{DIV} and $\lambda_{DIV}^\star$), and the real T-bill rate (λ_{REALTB} and $\lambda_{REALTB}^\star$).

The first result emerging from Table 1 is that the pricing kernels with non-traded factors (I-CAPM and C-CAPM) lead to estimates of the λ parameters that tend to deviate substantially from the corresponding $\lambda^\star$ values, and tend to be much larger in

[4] The R^2s of regressing XVW, SMB, and HML on the test asset returns are 0.9991, 0.7276, and 0.2011, respectively, for the case of size-sorted portfolios. In the case of the Fama–French portfolios, we have: 0.9894, 0.7629, and 0.8757.

[5] This choice of factors for the I-CAPM is analogous to that of Ferson and Harvey (1991).

[6] We perform an i.i.d. bootstrap of the VAR innovations. We also experimented with a stationary block-bootstrap of the VAR innovations, following Politis and Romano (1994), with almost identical results.

[7] We also performed our empirical analysis for the unconditional case; see Section 6 for details. These results are quantitatively and qualitatively similar to the results for the conditional case: i.e., the properties of the components of the unconditional risk premia are very similar to the properties of the components of the conditional risk premia.

Table 2
Non-traded component.

	$\delta_{n,0}$	$\delta_{n,\mathrm{XVW}}$	$\delta_{n,\mathrm{DIV}}$	$\delta_{n,\mathrm{REALTB}}$
Panel A: I-CAPM				
δ_n(HB3)	−3.0280(OLS)	−2.0480(OLS)	0.6948(OLS)	0.2609(OLS)
	−2.1799(GLS)	−1.1312(GLS)	0.3630(GLS)	0.3456(GLS)
$Bias_{\delta_n}$(HB3)	2.4029(OLS)	1.0417(OLS)	−1.0661(OLS)	−0.0389(OLS)
	1.5930(GLS)	0.3769(GLS)	−0.5789(GLS)	−0.2061(GLS)
$Emp.Std_{\delta_n}$(HB3)	1.3022(OLS)	1.2694(OLS)	0.7737(OLS)	0.8848(OLS)
	0.8455(GLS)	1.1752(GLS)	0.6857(GLS)	0.8029(GLS)
δ_n(DIV)	0.4872(OLS)	−5.9292(OLS)	−0.4911(OLS)	0.6333(OLS)
	0.2950(GLS)	−4.1300(GLS)	0.2368(GLS)	−0.2253(GLS)
$Bias_{\delta_n}$(DIV)	−0.5404(OLS)	5.2448(OLS)	1.2273(OLS)	−1.5283(OLS)
	−0.3381(GLS)	3.3242(GLS)	0.3676(GLS)	−0.5122(GLS)
$Emp.Std_{\delta_n}$(DIV)	0.7882(OLS)	1.6489(OLS)	1.0443(OLS)	1.1779(OLS)
	0.4974(GLS)	1.5713(GLS)	0.9613(GLS)	1.0937(GLS)
δ_n(CG)	−0.4838(OLS)	−1.4518(OLS)	0.6306(OLS)	−0.4465(OLS)
	−0.1837(GLS)	−0.8916(GLS)	0.5679(GLS)	−0.1659(GLS)
$Bias_{\delta_n}$(CG)	0.2857(OLS)	1.2644(OLS)	−0.4435(OLS)	0.2898(OLS)
	0.0947(GLS)	0.7695(GLS)	−0.3302(GLS)	0.0970(GLS)
$Emp.Std_{\delta_n}$(CG)	0.8764(OLS)	0.9473(OLS)	0.5810(OLS)	0.6480(OLS)
	0.5738(GLS)	0.6494(GLS)	0.3864(GLS)	0.4462(GLS)
δ_n(PREM)	−2.1673(OLS)	−1.2795(OLS)	−1.2590(OLS)	1.3059(OLS)
	−1.3852(GLS)	−1.0096(GLS)	−0.9315(GLS)	0.9955(GLS)
$Bias_{\delta_n}$(PREM)	2.0723(OLS)	1.2633(OLS)	0.9804(OLS)	−1.1094(OLS)
	1.3715(GLS)	0.9408(GLS)	0.6747(GLS)	−0.7822(GLS)
$Emp.Std_{\delta_n}$(PREM)	1.8114(OLS)	1.3195(OLS)	0.7655(OLS)	0.8728(OLS)
	1.1401(GLS)	0.8863(GLS)	0.5278(GLS)	0.5996(GLS)
δ_n(INF)	−1.5651(OLS)	−0.6055(OLS)	−0.1952(OLS)	−0.7026(OLS)
	−1.0053(GLS)	−0.2489(GLS)	−0.2114(GLS)	−0.4813(GLS)
$Bias_{\delta_n}$(INF)	1.8969(OLS)	0.4244(OLS)	0.1265(OLS)	0.2985(OLS)
	1.2073(GLS)	0.1420(GLS)	0.0946(GLS)	0.1199(GLS)
$Emp.Std_{\delta_n}$(INF)	1.0311(OLS)	0.6517(OLS)	0.3991(OLS)	0.4611(OLS)
	0.6890(GLS)	0.5968(GLS)	0.3565(GLS)	0.4214(GLS)
δ_n(REALTB)	2.1838(OLS)	0.2766(OLS)	0.0118(OLS)	0.0161(OLS)
	1.6192(GLS)	0.0915(GLS)	0.1643(GLS)	−0.0883(GLS)
$Bias_{\delta_n}$(REALTB)	−2.4205(OLS)	−0.3035(OLS)	0.2276(OLS)	−0.2233(OLS)
	−1.6736(GLS)	−0.1184(GLS)	0.0532(GLS)	−0.0854(GLS)
$Emp.Std_{\delta_n}$(REALTB)	1.1799(OLS)	0.8661(OLS)	0.5353(OLS)	0.5973(OLS)
	0.7809(GLS)	0.8035(GLS)	0.4914(GLS)	0.5508(GLS)
Panel B: C−CAPM				
δ_n(HB3)	0.0341(OLS)	0.0597(OLS)	0.0240(OLS)	0.0257(OLS)
	0.0072(GLS)	0.0192(GLS)	0.0065(GLS)	−0.0029(GLS)
$Bias_{\delta_n}$(HB3)	0.0008(OLS)	0.0038(OLS)	0.0067(OLS)	0.0025(OLS)
	−0.0030(GLS)	−0.0036(GLS)	−0.0004(GLS)	0.0008(GLS)
$Emp.Std_{\delta_n}$(HB3)	0.0468(OLS)	0.0789(OLS)	0.0426(OLS)	0.0401(OLS)
	0.0147(GLS)	0.0248(GLS)	0.0140(GLS)	0.0127(GLS)
δ_n(DIV)	−0.0678(OLS)	−0.1186(OLS)	−0.0477(OLS)	−0.0510(OLS)
	−0.0144(GLS)	−0.0381(GLS)	−0.0130(GLS)	0.0058(GLS)
$Bias_{\delta_n}$(DIV)	−0.0068(OLS)	−0.0096(OLS)	−0.0143(OLS)	−0.0059(OLS)
	0.0032(GLS)	0.0062(GLS)	−0.0004(GLS)	−0.0009(GLS)
$Emp.Std_{\delta_n}$(DIV)	0.0500(OLS)	0.0737(OLS)	0.0450(OLS)	0.0434(OLS)
	0.0202(GLS)	0.0285(GLS)	0.0182(GLS)	0.0181(GLS)
δ_n(CG)	0.7188(OLS)	1.2581(OLS)	0.5060(OLS)	0.5408(OLS)
	0.1523(GLS)	0.4041(GLS)	0.1374(GLS)	−0.0620(GLS)
$Bias_{\delta_n}$(CG)	0.0519(OLS)	0.0746(OLS)	0.1395(OLS)	0.0525(OLS)
	−0.0336(GLS)	−0.0725(GLS)	0.0018(GLS)	0.0138(GLS)
$Emp.Std_{\delta_n}$(CG)	0.4223(OLS)	0.5967(OLS)	0.4027(OLS)	0.3883(OLS)
	0.1977(GLS)	0.2655(GLS)	0.1769(GLS)	0.1782(GLS)
δ_n(PREM)	0.0015(OLS)	0.0026(OLS)	0.0010(OLS)	0.0011(OLS)
	0.0003(GLS)	0.0008(GLS)	0.0003(GLS)	−0.0001(GLS)
$Bias_{\delta_n}$(PREM)	0.0022(OLS)	0.0022(OLS)	0.0015(OLS)	0.0012(OLS)
	−0.0005(GLS)	−0.0007(GLS)	−0.0000(GLS)	−0.0003(GLS)
$Emp.Std_{\delta_n}$(PREM)	0.0463(OLS)	0.0767(OLS)	0.0398(OLS)	0.0370(OLS)
	0.0120(GLS)	0.0220(GLS)	0.0117(GLS)	0.0097(GLS)
δ_n(INF)	−0.1095(OLS)	−0.1917(OLS)	−0.0771(OLS)	−0.0824(OLS)
	−0.0232(GLS)	−0.0616(GLS)	−0.0209(GLS)	0.0095(GLS)
$Bias_{\delta_n}$(INF)	−0.0066(OLS)	−0.0107(OLS)	−0.0206(OLS)	−0.0073(OLS)
	0.0061(GLS)	0.0111(GLS)	0.0001(GLS)	−0.0028(GLS)
$Emp.Std_{\delta_n}$(INF)	0.0728(OLS)	0.1114(OLS)	0.0691(OLS)	0.0664(OLS)
	0.0323(GLS)	0.0460(GLS)	0.0291(GLS)	0.0289(GLS)

(continued on next page)

Table 2 (*continued*)

	$\delta_{n,0}$	$\delta_{n,\mathrm{XVW}}$	$\delta_{n,\mathrm{DIV}}$	$\delta_{n,\mathrm{REALTB}}$
Panel B: C−CAPM				
δ_n(REALTB)	0.1161(OLS)	0.2032(OLS)	0.0817(OLS)	0.0874(OLS)
	0.0246(GLS)	0.0653(GLS)	0.0222(GLS)	−0.0100(GLS)
$Bias_{\delta_n}$(REALTB)	0.0071(OLS)	0.0116(OLS)	0.0220(OLS)	0.0078(OLS)
	−0.0065(GLS)	−0.0117(GLS)	0.0000(GLS)	0.0030(GLS)
$Emp.Std_{\delta_n}$(REALTB)	0.0764(OLS)	0.1168(OLS)	0.0728(OLS)	0.0699(OLS)
	0.0341(GLS)	0.0485(GLS)	0.0308(GLS)	0.0305(GLS)
Panel C: FF3				
δ_n(HB3)	−0.0215	−0.0022	0.0020	−0.0015
$Bias_{\delta_n}$(HB3)	0.0001	0.0004	−0.0004	−0.0003
$Emp.Std_{\delta_n}$(HB3)	0.0107	0.0089	0.0067	0.0053
δ_n(DIV)	−0.0058	0.0003	0.0017	0.0004
$Bias_{\delta_n}$(DIV)	0.0006	−0.0017	−0.0018	−0.0009
$Emp.Std_{\delta_n}$(DIV)	0.0081	0.0053	0.0036	0.0026
δ_n(CG)	0.0063	0.0007	−0.0004	0.0006
$Bias_{\delta_n}$(CG)	−0.0005	0.0017	0.0013	0.0005
$Emp.Std_{\delta_n}$(CG)	0.0091	0.0056	0.0036	0.0026
δ_n(PREM)	0.0126	−0.0014	−0.0022	0.0012
$Bias_{\delta_n}$(PREM)	0.0001	0.0006	0.0005	−0.0000
$Emp.Std_{\delta_n}$(PREM)	0.0122	0.0063	0.0049	0.0041
δ_n(INF)	−0.0092	−0.0165	−0.0067	−0.0009
$Bias_{\delta_n}$(INF)	0.0021	0.0017	0.0013	−0.0007
$Emp.Std_{\delta_n}$(INF)	0.0109	0.0086	0.0057	0.0041
δ_n(REALTB)	0.0088	0.0156	0.0064	0.0010
$Bias_{\delta_n}$(REALTB)	−0.0017	−0.0015	−0.0012	0.0006
$Emp.Std_{\delta_n}$(REALTB)	0.0105	0.0079	0.0054	0.0040
Panel D: CAPM				
δ_n(HB3)	0.0004	0.0002	0.0003	0.0004
$Bias_{\delta_n}$(HB3)	−0.0004	−0.0001	−0.0003	−0.0004
$Emp.Std_{\delta_n}$(HB3)	0.0010	0.0007	0.0010	0.0012
δ_n(DIV)	0.0010	0.0005	0.0008	0.0010
$Bias_{\delta_n}$(DIV)	−0.0014	−0.0005	−0.0012	−0.0013
$Emp.Std_{\delta_n}$(DIV)	0.0011	0.0008	0.0012	0.0015
δ_n(CG)	0.0000	0.0000	0.0000	0.0000
$Bias_{\delta_n}$(CG)	0.0005	0.0002	0.0006	0.0006
$Emp.Std_{\delta_n}$(CG)	0.0008	0.0006	0.0008	0.0010
δ_n(PREM)	0.0001	0.0000	0.0001	0.0001
$Bias_{\delta_n}$(PREM)	0.0001	0.0001	0.0001	0.0001
$Emp.Std_{\delta_n}$(PREM)	0.0011	0.0008	0.0012	0.0014
δ_n(INF)	−0.0003	−0.0001	−0.0002	−0.0003
$Bias_{\delta_n}$(INF)	−0.0001	−0.0001	−0.0002	−0.0002
$Emp.Std_{\delta_n}$(INF)	0.0008	0.0006	0.0009	0.0010
δ_n(REALTB)	0.0003	0.0002	0.0003	0.0003
$Bias_{\delta_n}$(REALTB)	−0.0000	0.0001	0.0001	0.0001
$Emp.Std_{\delta_n}$(REALTB)	0.0008	0.0006	0.0009	0.0010

We report estimates of the non-traded component $\delta_{n,t}$ for the six economic factors: term structure (HB3), dividend yield (DIV), consumption growth (CG), default premium (PREM), inflation (INF), and real T-bill rate (REALTB). Time variation is driven by lagged values of the market excess return (XVW), the dividend yield, and the real T-bill rate. We consider four multi-beta models. The first model is the I-CAPM (Panel A), the second model is the C-CAPM (Panel B), the third model is the FF3 model (Panel C), and the fourth model is the CAPM (Panel D). The sample period is March 1959–December 2002. Finite-sample absolute biases (Bias) and empirical standard errors (Emp. Std.) are estimated by bootstrap (see Section 5.1 for details). For each economic factor, we report the average non-traded component, $\delta_{n,0}$, and the coefficients relating the non-traded component to the three instruments: $\delta_{n,\mathrm{XVW}}$, $\delta_{n,\mathrm{DIV}}$, and $\delta_{n,\mathrm{REALTB}}$. For the models with non-traded factors, we consider estimates based on the two weighting matrices $W=I$ (OLS) and $W=\hat{\Sigma}_{rr}^{-1}$ (GLS).

absolute value. Moreover, depending on the model, the estimates can also vary substantially. On the other hand, the pricing kernels with traded factors (FF3 and CAPM) deliver λ estimates that are much closer to their $\lambda^{\star}$ counterparts. For example, consider the premium on consumption growth, which is of special economic significance. The $\lambda_0^{\star}$ estimate is 0.0103. The λ_0 estimate is −0.4768 for the I-CAPM/OLS and 0.7674 for the C-CAPM/OLS. Hence, not only can the discrepancies between the two sets of estimates be large, but the risk premium can even change sign depending on the model used. For a comparison, the FF3 model and the CAPM deliver λ_0 estimates that are much closer to $\lambda_0^{\star}$, 0.0257 and 0.0144, respectively. A similar pattern holds for the inflation risk premium. The $\lambda_0^{\star}$ estimate is 0.0036. The λ_0 estimates are −1.5613 for the I-CAPM/OLS and −0.1355 for the C-CAPM/OLS; while the FF3 model and the CAPM deliver λ_0 estimates of −0.0198 and −0.0135, respectively.

The coefficients relating the conditional risk premia to the instruments can also differ greatly. For example, in the case of consumption growth, the $\lambda_{\mathrm{DIV}}^{\star}$ estimate is 0.0093, while the λ_{DIV} estimates are −0.4412 for the I-CAPM/OLS and 0.5402 for the C-CAPM/OLS. On the other hand, the FF3 model and the CAPM deliver λ_{DIV} estimates of 0.0177 and 0.0120, respectively. In

Table 3
Mispricing component.

	$\delta_{m,0}$	$\delta_{m,\mathrm{XVW}}$	$\delta_{m,\mathrm{DIV}}$	$\delta_{m,\mathrm{REALTB}}$
Panel A: I-CAPM				
δ_m(HB3)	−0.0057(OLS)	−0.0060(OLS)	−0.0054(OLS)	0.0030(OLS)
	0.0000(GLS)	−0.0003(GLS)	−0.0039(gLS)	0.0009(GLS)
$Bias_{\delta_m}$(HB3)	0.0015(OLS)	−0.0078(OLS)	−0.0031(OLS)	0.0076(OLS)
	0.0007(GLS)	0.0005(GLS)	0.0035(GLS)	−0.0008(GLS)
$Emp.Std_{\delta_m}$(HB3)	0.0123(OLS)	0.0141(OLS)	0.0147(OLS)	0.0153(OLS)
	0.0038(GLS)	0.0027(GLS)	0.0227(GLS)	0.0052(GLS)
δ_m(DIV)	0.0009(OLS)	−0.0032(OLS)	−0.0015(OLS)	−0.0002(OLS)
	0.0002(GLS)	0.0004(GLS)	−0.0008(GLS)	0.0005(GLS)
$Bias_{\delta_m}$(DIV)	0.0007(OLS)	0.0014(OLS)	0.0074(OLS)	−0.0079(OLS)
	0.0086(GLS)	−0.0003(GLS)	0.0007(GLS)	−0.0005(GLS)
$Emp.Std_{\delta_m}$(DIV)	0.0181(OLS)	0.0117(OLS)	0.0132(OLS)	0.0133(OLS)
	0.0313(GLS)	0.0028(GLS)	0.0049(GLS)	0.0036(GLS)
δ_m(CG)	−0.0033(OLS)	−0.0040(OLS)	−0.0014(OLS)	−0.0040(OLS)
	−0.0000(GLS)	−0.0002(GLS)	−0.0019(GLS)	0.0004(GLS)
$Bias_{\delta_m}$(CG)	−0.0010(OLS)	0.0018(OLS)	0.0004(OLS)	0.0026(OLS)
	−0.0020(GLS)	0.0006(GLS)	0.0075(GLS)	−0.0017(GLS)
$Emp.Std_{\delta_m}$(CG)	0.0150(OLS)	0.0093(OLS)	0.0129(OLS)	0.0115(OLS)
	0.0072(GLS)	0.0026(GLS)	0.0216(GLS)	0.0050(GLS)
δ_m(PREM)	−0.0018(OLS)	−0.0007(OLS)	0.0004(OLS)	−0.0015(OLS)
	−0.0000(GLS)	0.0002(GLS)	−0.0004(GLS)	0.0003(GLS)
$Bias_{\delta_m}$(PREM)	0.0001(OLS)	0.0008(OLS)	−0.0016(OLS)	0.0016(OLS)
	−0.0010(GLS)	−0.0007(GLS)	0.0016(GLS)	−0.0011(GLS)
$Emp.Std_{\delta_m}$(PREM)	0.0121(OLS)	0.0091(OLS)	0.0091(OLS)	0.0092(OLS)
	0.0039(GLS)	0.0027(GLS)	0.0047(GLS)	0.0034(GLS)
δ_m(INF)	0.0002(OLS)	−0.0031(OLS)	0.0017(OLS)	−0.0037(OLS)
	0.0068(GLS)	0.0000(GLS)	0.0002(GLS)	−0.0000(GLS)
$Bias_{\delta_m}$(INF)	0.0052(OLS)	0.0007(OLS)	−0.0002(OLS)	0.0013(OLS)
	0.0017(GLS)	−0.0000(GLS)	0.0001(GLS)	−0.0000(GLS)
$Emp.Std_{\delta_m}$(INF)	0.0121(OLS)	0.0066(OLS)	0.0079(OLS)	0.0076(OLS)
	0.0068(GLS)	0.0016(GLS)	0.0139(GLS)	0.0032(GLS)
δ_m(REALTB)	−0.0017(OLS)	0.0001(OLS)	0.0011(OLS)	−0.0017(OLS)
	−0.0000(GLS)	−0.0000(GLS)	0.0000(GLS)	−0.0000(GLS)
$Bias_{\delta_m}$(REALTB)	−0.0039(OLS)	−0.0002(OLS)	0.0010(OLS)	−0.0005(OLS)
	−0.0011(GLS)	−0.0000(GLS)	0.0000(GLS)	−0.0000(GLS)
$Emp.Std_{\delta_m}$(REALTB)	0.0122(OLS)	0.0058(OLS)	0.0063(OLS)	0.0063(OLS)
	0.0049(GLS)	0.0017(GLS)	0.0030(GLS)	0.0022(GLS)
Panel B: C-CAPM				
δ_m(HB3)	0.0083(OLS)	0.0189(OLS)	−0.0054(OLS)	−0.0026(OLS)
	0.0188(GLS)	0.0347(GLS)	0.0014(GLS)	0.0085(GLS)
$Bias_{\delta_m}$(HB3)	0.0007(OLS)	0.0014(OLS)	−0.0008(OLS)	0.0005(OLS)
	0.0009(GLS)	0.0018(GLS)	0.0007(GLS)	0.0002(GLS)
$Emp.Std_{\delta_m}$(HB3)	0.0170(OLS)	0.0246(OLS)	0.0143(OLS)	0.0141(OLS)
	0.0137(GLS)	0.0188(GLS)	0.0120(GLS)	0.0121(OLS)
δ_m(DIV)	−0.0229(OLS)	−0.1090(OLS)	−0.0158(OLS)	0.0026(OLS)
	0.0448(GLS)	−0.0070(GLS)	0.0282(GLS)	0.0746(GLS)
$Bias_{\delta_m}$(DIV)	−0.0004(OLS)	0.0005(OLS)	0.0021(OLS)	−0.0011(OLS)
	0.0034(GLS)	0.0089(GLS)	0.0125(GLS)	−0.0031(GLS)
$Emp.Std_{\delta_m}$(DIV)	0.0245(OLS)	0.0300(OLS)	0.0226(OLS)	0.0224(OLS)
	0.0311(GLS)	0.0350(GLS)	0.0277(GLS)	0.0274(GLS)
δ_m(CG)	0.0383(OLS)	0.0578(OLS)	0.0249(OLS)	0.0408(OLS)
	0(GLS)	0(GLS)	0(GLS)	0(GLS)
$Bias_{\delta_m}$(CG)	0.0152(OLS)	0.0242(OLS)	0.0162(OLS)	0.0117(OLS)
	0(GLS)	0(GLS)	0(GLS)	0(GLS)
$Emp.Std_{\delta_m}$(CG)	0.0323(OLS)	0.0421(OLS)	0.0284(OLS)	0.0282(OLS)
	0(GLS)	0(GLS)	0(GLS)	0(GLS)
δ_m(PREM)	0.0024(OLS)	0.0287(OLS)	−0.0018(OLS)	−0.0001(OLS)
	−0.0031(GLS)	0.0205(GLS)	−0.0053(GLS)	−0.0059(GLS)
$Bias_{\delta_m}$(PREM)	−0.0006(OLS)	−0.0004(OLS)	−0.0005(OLS)	0.0005(OLS)
	−0.0011(GLS)	−0.0014(GLS)	−0.0013(GLS)	0.0007(GLS)
$Emp.Std_{\delta_m}$(PREM)	0.0154(OLS)	0.0215(OLS)	0.0133(OLS)	0.0140(OLS)
	0.0119(GLS)	0.0162(GLS)	0.0108(GLS)	0.0104(GLS)
δ_m(INF)	−0.0296(OLS)	−0.0701(OLS)	0.0003(OLS)	−0.0184(OLS)
	−0.0091(GLS)	−0.0393(GLS)	0.0136(GLS)	0.0034(GLS)
$Bias_{\delta_m}$(INF)	−0.0017(OLS)	−0.0007(OLS)	−0.0009(OLS)	−0.0006(OLS)
	0.0007(GLS)	0.0034(GLS)	0.0028(GLS)	−0.0003(GLS)
$Emp.Std_{\delta_m}$(INF)	0.0217(OLS)	0.0283(OLS)	0.0186(OLS)	0.0187(OLS)
	0.0151(GLS)	0.0199(GLS)	0.0130(GLS)	0.0126(GLS)

(continued on next page)

Table 3 (*continued*)

	$\delta_{m,0}$	$\delta_{m,\mathrm{XVW}}$	$\delta_{m,\mathrm{DIV}}$	$\delta_{m,\mathrm{REALTB}}$
Panel B: C-CAPM				
δ_m(REALTB)	0.0226(OLS)	0.0628(OLS)	−0.0013(OLS)	0.0144(OLS)
	0.0085(GLS)	0.0415(GLS)	−0.0105(GLS)	−0.0006(GLS)
$Bias_{\delta_m}$(REALTB)	0.0020(OLS)	0.0014(OLS)	0.0009(OLS)	0.0007(OLS)
	−0.0002(GLS)	−0.0024(GLS)	−0.0020(GLS)	0.0001(GLS)
$Emp.Std_{\delta_m}$(REALTB)	0.0203(OLS)	0.0274(OLS)	0.0176(OLS)	0.0177(OLS)
	0.0141(GLS)	0.0188(GLS)	0.0124(GLS)	0.0119(GLS)
Panel C: FF3				
δ_m(HB3)	0.0068	0.0192	−0.0094	0.0005
$Bias_{\delta_m}$(HB3)	0.0034	0.0011	0.0014	0.0006
$Emp.Std_{\delta_m}$(HB3)	0.0143	0.0138	0.0113	0.0097
δ_m(DIV)	−0.0056	−0.0024	−0.0161	−0.0006
$Bias_{\delta_m}$(DIV)	0.0115	0.0032	0.0089	−0.0004
$Emp.Std_{\delta_m}$(DIV)	0.0329	0.0137	0.0220	0.0133
δ_m(CG)	0.0091	−0.0059	0.0089	0.0192
$Bias_{\delta_m}$(CG)	−0.0031	−0.0015	−0.0026	−0.0004
$Emp.Std_{\delta_m}$(CG)	0.0172	0.0148	0.0132	0.0120
δ_m(PREM)	0.0001	0.0100	−0.0055	0.0012
$Bias_{\delta_m}$(PREM)	−0.0002	−0.0007	−0.0005	0.0003
$Emp.Std_{\delta_m}$(PREM)	0.0109	0.0123	0.0089	0.0084
δ_m(INF)	−0.0143	−0.0320	0.0067	−0.0121
$Bias_{\delta_m}$(INF)	0.0020	−0.0009	0.0014	−0.0000
$Emp.Std_{\delta_m}$(INF)	0.0162	0.0158	0.0125	0.0114
δ_m(REALTB)	0.0091	0.0288	−0.0091	0.0096
$Bias_{\delta_m}$(REALTB)	−0.0008	0.0009	−0.0006	0.0001
$Emp.Std_{\delta_m}$(REALTB)	0.0144	0.0155	0.0116	0.0108
Panel D: CAPM				
δ_m(HB3)	0.0154	0.0393	−0.0012	0.0011
$Bias_{\delta_m}$(HB3)	0.0021	0.0014	0.0005	0.0007
$Emp.Std_{\delta_m}$(HB3)	0.0130	0.0187	0.0113	0.0102
δ_m(DIV)	−0.0034	0.0104	−0.0105	−0.0004
$Bias_{\delta_m}$(DIV)	0.0120	0.0018	0.0077	0.0001
$Emp.Std_{\delta_m}$(DIV)	0.0308	0.0156	0.0214	0.0133
δ_m(CG)	0.0041	−0.0207	0.0026	0.0188
$Bias_{\delta_m}$(CG)	−0.0029	−0.0008	−0.0015	−0.0005
$Emp.Std_{\delta_m}$(CG)	0.0158	0.0197	0.0129	0.0124
δ_m(PREM)	0.0019	0.0196	−0.0013	0.0013
$Bias_{\delta_m}$(PREM)	−0.0017	−0.0002	−0.0007	0.0004
$Emp.Std_{\delta_m}$(PREM)	0.0112	0.0171	0.0099	0.0088
δ_m(INF)	−0.0168	−0.0309	0.0076	−0.0123
$Bias_{\delta_m}$(INF)	0.0017	0.0003	0.0020	0.0003
$Emp.Std_{\delta_m}$(INF)	0.0155	0.0194	0.0124	0.0115
δ_m(REALTB)	0.0135	0.0356	−0.0065	0.0100
$Bias_{\delta_m}$(REALTB)	−0.0008	−0.0001	−0.0015	−0.0001
$Emp.Std_{\delta_m}$(REALTB)	0.0142	0.0189	0.0117	0.0110

We report estimates of the mispricing component $\delta_{m,t}$ for the six economic factors: term structure (HB3), dividend yield (DIV), consumption growth (CG), default premium (PREM), inflation (INF), and real T-bill rate (REALTB). Time variation is driven by lagged values of the market excess return (XVW), the dividend yield, and the real T-bill rate. We consider four multi-beta models. The first model is the I-CAPM (Panel A), the second model is the C-CAPM (Panel B), the third model is the FF3 model (Panel C), the fourth model is the CAPM (Panel D). The sample period is March 1959–December 2002. Finite-sample absolute biases (Bias) and empirical standard errors (Emp. Std.) are estimated by bootstrap (see Section 5.1 for details). For each economic factor, we report the average non-traded component, $\delta_{m,0}$, and the coefficients relating the non-traded component to the three instruments: $\delta_{m,\mathrm{XVW}}$, $\delta_{m,\mathrm{DIV}}$, and $\delta_{m,\mathrm{REALTB}}$. For the models with non-traded factors, we consider estimates based on the two weighting matrices $W=I$ (OLS) and $W=\hat{\Sigma}_{rr}^{-1}$ (GLS).

the case of inflation, the $\lambda^{\star}_{\mathrm{DIV}}$ estimate is −0.0185, while the λ_{DIV} estimates are −1.2771 for the I-CAPM/OLS and −0.0954 for the C-CAPM/OLS; the FF3 model and the CAPM deliver λ_{DIV} estimates of −0.0185 and −0.0112, respectively.

The second result emerging from Table 1 is that the choice of weighting matrix for models with non-traded factors can make a substantial difference. In the case of the I-CAPM consumption premium, for example, the estimate of the λ_0 parameter changes from −0.4768 to −0.1734 going from the OLS to the GLS specification; the estimate of the λ_{DIV} parameter changes from −0.4412 to −0.1561.

The third result emerging from Table 1 is that the λ estimates assigned by models with non-traded factors can exhibit substantial small-sample biases and tend to be estimated imprecisely. Focusing again on the I-CAPM/OLS risk premium on consumption growth, the bias for the λ_0 estimate is 0.2848, with a standard error of 0.8859. On the other hand, the (absolute) biases for models with traded factors are much smaller: −0.0035 and −0.0023 for the FF3 model and the CAPM, respectively. Also modest are the empirical standard errors: 0.0130 and 0.0056. Similarly substantial are the biases and standard errors for the λ_{DIV}

estimates: 0.2950 and 0.6557 for the I-CAPM/OLS. The corresponding biases and standard errors in the FF3 model and in the CAPM are only 0.0012 and 0.0090, and 0.0016 and 0.0061, respectively. For the case of inflation, the bias for the I-CAPM/OLS λ_0 estimate is 2.0732, with a standard error of 1.0386. Biases for models with traded factors are 0.0046 (FF3) and 0.0021 (CAPM), with standard errors of 0.0139 and 0.0054. The bias for the I-CAPM/OLS λ_{DIV} estimate is 0.9755, with standard error of 0.7717. The corresponding biases and standard errors in the FF3 model and in the CAPM are only −0.0006 and 0.0099, and −0.0015 and 0.0059, respectively.

On the other hand, the $\lambda^\star$ estimates exhibit modest biases and standard errors compared to the λ estimates. In the case of consumption growth, the bias in $\lambda_0^\star$ is negligible and the empirical standard error is only 0.0164. The bias in $\lambda_{DIV}^\star$ is also small, 0.0025, and the empirical standard error is 0.0145. In the case of inflation, the bias in $\lambda_0^\star$ is 0.0006 and the empirical standard error is 0.0163. The bias in $\lambda_{DIV}^\star$ is −0.0033, and the empirical standard error is 0.0142.[8]

Finally, it is worth noting that the largest $\lambda_0^\star$ estimate is for the dividend yield, −0.0629, comparable in magnitude to the average conditional Sharpe ratio on the market index, 0.0954. This result is not surprising. Innovations in the dividend yield are mainly driven, in the opposite direction, by equity market returns. Moreover, the risk premium on a maximum-correlation portfolio is proportional to the covariance between a factor and the excess return on the tangency portfolio. Hence, the negative $\lambda_0^\star$ estimate reflects the negative correlation between the return on the tangency portfolio and the innovation in the dividend yield.

While not reported in the paper, we also computed the *p*-values of a one-sided significance test based on the bootstrap distribution of the coefficients. The difference in inference between the $\lambda^\star$ and λ estimates is striking. Of the 24 $\lambda^\star$ coefficients reported in Table 1, only three are significant at the 5% level. Indeed, the only significant $\lambda_0^\star$ estimate is the one for the dividend yield, with a *p*-value of 2.8%. For the I-CAPM/OLS, on the other hand, there are 16 significant parameter estimates. Interestingly, the difference in inference persists when we consider a model with traded factors like the FF3 model. In this case, there are still nine significant estimates out of 24.[9] It is also worth noting that some of the factors are significantly priced across models. For example, the mean risk premium on inflation is significant at the 5% significance level, or better, in three of the four models: the I-CAPM, the FF3 model, and the CAPM.

5.2.2. Non-traded component: δ_n

Table 2 reports the estimate of the mean non-traded component ($\delta_{n,0}$) and the estimated coefficients relating the non-traded component to the market excess return ($\delta_{n,XVW}$), the dividend yield ($\delta_{n,DIV}$), and the real T-bill rate ($\delta_{n,REALTB}$).

It is immediate that, for models with non-traded factors, this component represents a substantial portion of the total estimated risk premium, and its parameters are estimated with large (absolute) biases and imprecisely. Moreover, the estimates can be significantly affected by the choice of weighting matrix. For example, in the case of consumption growth, $\delta_{n,0}$ is −0.4838 for the I-CAPM/OLS and 0.7188 for the C-CAPM/OLS. The corresponding biases are 0.2857 and 0.0519, while the standard errors are 0.8764 and 0.4223. Going from the OLS to the GLS approach leads to estimates of −0.1837 and 0.1523 for the I-CAPM and C-CAPM, respectively. We find a similar pattern for $\delta_{n,DIV}$. The estimate is 0.6306 for the I-CAPM/OLS and 0.5060 for the C-CAPM/OLS. The corresponding biases are −0.4435 and 0.1395, while the standard errors are 0.5810 and 0.4027. Going from the OLS to the GLS approach leads to estimates of 0.5679 and 0.1374 for the I-CAPM and C-CAPM, respectively. Similar patterns also hold for the parameters of the non-traded component of the inflation risk premium and for the other economic factors considered.[10]

5.2.3. Mispricing component: δ_m

Table 3 reports the estimates of the mean mispricing component ($\delta_{m,0}$) and of the coefficients relating the mispricing component to the market excess return ($\delta_{m,XVW}$), the dividend yield ($\delta_{m,DIV}$), and the real T-bill rate ($\delta_{m,REALTB}$).

Unlike the non-traded component, the δ_m component of all models tends to be small compared to the total estimated risk premium. Biases and standard errors are also modest. Moreover, for models with non-traded factors, the choice of weighting matrix is less relevant. The intuition for this is that for a linear factor model with non-traded factors, where the parameters are estimated by cross-sectional GLS-style regressions, the mispricing of the portfolios mimicking the factors is identically zero. Hence, even when the estimation approach departs from this paradigm, i.e., we employ the identity matrix as the weighting matrix, the mispricing of the mimicking portfolios is very modest.[11]

For example, in the case of consumption growth, $\delta_{m,0}$ is −0.0033 for the I-CAPM/OLS and 0.0383 for the C-CAPM/OLS. The corresponding biases are −0.0010 and 0.0152, while the standard errors are 0.0150 and 0.0323. Going from the OLS to the GLS approach leads to estimates of −0.0000 and 0 for the I-CAPM and C-CAPM, respectively.

A similar pattern holds for the $\delta_{m,DIV}$ coefficient: the estimates are −0.0014 for the I-CAPM/OLS and 0.0249 for the C-CAPM/OLS, with biases of 0.0004 and 0.0162, and standard errors of 0.0129 and 0.0284. Going from the OLS to the GLS approach leads to estimates of −0.0019 and 0 for the I-CAPM and C-CAPM, respectively.

[8] Overall, the evidence using the Fama–French portfolios is similar. While the λ estimates tend to be smaller in absolute value and display somewhat smaller biases, it is still the case that the bias in the λ^* estimates is negligible.

[9] In the case of the Fama–French portfolios, we obtain essentially the same results for the estimates of the λ^* parameters. For the estimates of the λ parameters the results are similar, although the number of significant estimates decreases somewhat, especially in the OLS case.

[10] Results are qualitatively similar for the case of the Fama–French portfolios.

[11] Note, though, that in the case of the I-CAPM, we are imposing the restriction that the market is exactly priced. This means that even in the GLS case, the mispricing components of the risk premia on the non-traded factors driving the kernel are not identically zero.

5.2.4. Bounds

We plot the quantities

$$\pm\sqrt{\mathrm{Var}_t(x_{t+1})-\mathrm{Var}_t(x^\star_{t+1})}\sqrt{\mathrm{Var}_t(y_{t+1})-\mathrm{Var}_t(y^\star_{t+1})} \tag{62}$$

and

$$\pm\sqrt{\mathrm{Var}_t(q^\star_{t+1}-x^\star_{t+1})}\sqrt{\mathrm{Var}_t(y^\star_{t+1})} \tag{63}$$

as functions of $\sqrt{\mathrm{Var}_t(x_{t+1})-\mathrm{Var}_t(x^\star_{t+1})}$ and $\sqrt{\mathrm{Var}_t(q^\star_{t+1}-x^\star_{t+1})}$, respectively. All conditional quantities are evaluated at the unconditional average of the instruments Z_t. We report the two quantities above for two economic factors: innovations in consumption and innovations in the dividend yield. The graphs also report vertical lines corresponding to the conditional standard deviations implied by the four different models. The intersections between the vertical lines and the bounds effectively delimit the range of possible values of the two components of the premia for the different models. Fig. 1 reports bounds on δ_{nt}, while Fig. 2 reports bounds on δ_{mt}. The top panel in each picture presents reference values for the models estimated using $W=I$, while the bottom panel presents reference values for the models estimated using $W=\hat{\Sigma}rr^{-1}$.

Several patterns emerge from the two figures. First, the slopes of the bounds are markedly different for the two factors. The bounds on the non-traded (mispricing) component are wider (narrower) for consumption growth, reflecting the fact that consumption is not tracked by security returns as well as the dividend yield is. The slopes of the δ_n bounds are ± 0.7274 (DIV) and ± 0.9682 (CG), respectively, whereas the slopes of the δ_m bounds are ± 0.6862 (DIV) and ± 0.2503 (CG), respectively.

Second, for the I-CAPM and C-CAPM, the bounds are much wider for the non-traded component than for the mispricing component. This reflects the fact that the variability of the non-traded component of the kernels, $x_{t+1}-x^\star_{t+1}$, is much higher than the variability of $q^\star_{t+1}-x^\star_{t+1}$. For example, in the case of the C-CAPM ($W=I$) and consumption growth, the bound on δ_m is ± 0.0704, whereas the bound on δ_n is ± 0.7193. Hence, a kernel with the same non-traded variability and mispricing as the C-CAPM kernel can generate much larger (in absolute value) estimates for δ_n than for δ_m.

Third, the non-traded variabilities and degrees of mispricing of the four models are also markedly different. For example, for $W=I$, the non-traded variability of the I-CAPM kernel is 4.9421, whereas it is only 0.7430 for the C-CAPM kernel. For consumption growth, these non-traded variabilities translate into bounds of ± 4.7847 and ± 0.7193, respectively. Hence, a kernel with the non-traded variability of the I-CAPM has the potential to generate much larger discrepancies between λ and $\lambda^\star$ than a kernel with the non-traded variablity of the C-CAPM. The ranking is different for the bounds on the mispricing component. In this case, the I-CAPM displays lower mispricing, as measured by $\sqrt{\mathrm{Var}_t(q^\star_{t+1}-x^\star_{t+1})}$, than the C-CAPM: 0.1327 as opposed to 0.2177. For consumption growth, this translates into bounds on the mispricing component of ± 0.0332 and ± 0.0704, respectively.

Fourth, when the kernel parameters are estimated using $W=\hat{\Sigma}_{rr}^{-1}$, both the I-CAPM and C-CAPM kernels display lower volatility, and lower mispricing, than for the $W=I$ choice of weighting matrix. Hence, going from the OLS to the GLS estimates, there is a narrowing of the bounds on both components of the premium. For example, for consumption growth and the I-CAPM, the δ_n bound goes from ± 4.7847 to ± 3.5626, and the δ_m bound goes from ± 0.0332 to ± 0.0261.

Finally, it is worth comparing the magnitude of the bounds on δ_n and δ_m to the magnitude of the estimates $\lambda^\star$. For example, in the case of consumption growth, $\lambda^\star_0$ is only 0.0103. Hence, for our models with non-traded factors, the bounds on the non-traded and mispricing components can be wide relative to the risk premia on the maximum-correlation portfolios.

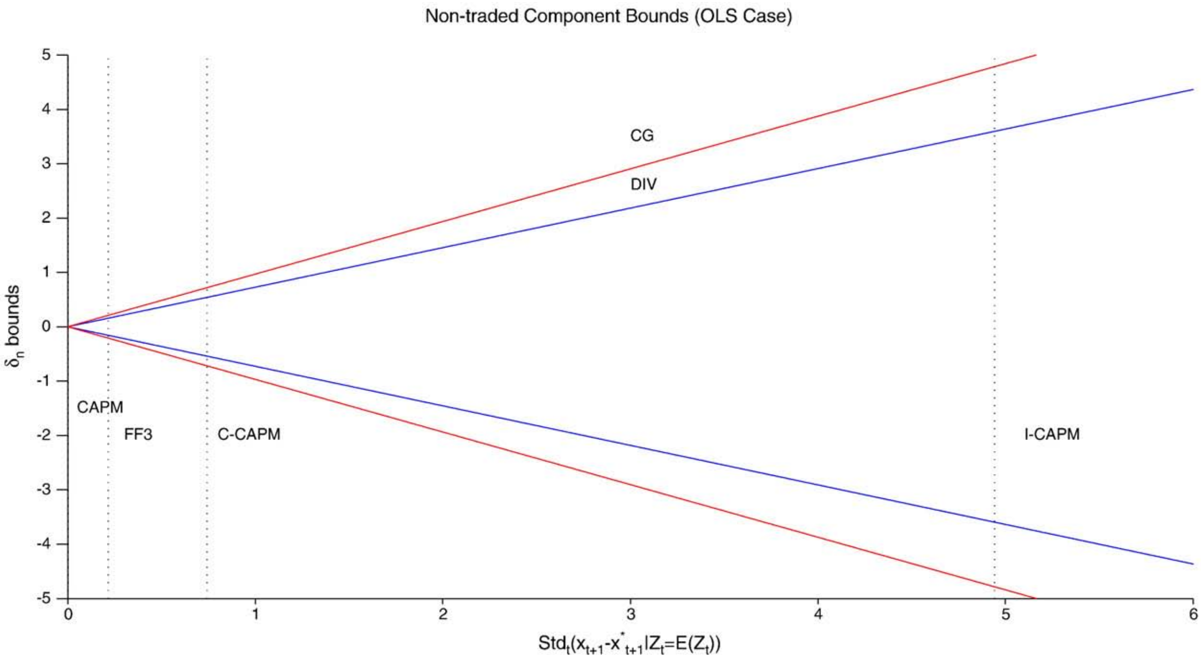
6. Extensions

In this section, we explore two extensions of our analysis where we modify the features of the data. First, we study how our results change in economies where the proposed pricing kernels are admissible, i.e., they price the test assets correctly. In other words, we perform our simulations under the null. Second, focusing on the I-CAPM and C-CAPM, we change the tradeability of the non-traded factors. In the interest of brevity, the tables displaying the detailed results of these exercises are not reported in the paper, but they are available on the authors' websites.

For simplicity, in both extensions we consider unconditional versions of the pricing models; the factors are still defined as mean-zero VAR(1) innovations, but the instrument set, Z_t, predicting returns and driving pricing-kernel parameters and risk premia, only contains the constant. Since we now bootstrap not the *innovations* in returns, but the returns per se, we are concerned that an i.i.d. bootstrap may not capture serial dependence in the data. Hence, we perform a stationary block-bootstrap, following Politis and Romano (1994), and we set the average length of the block to three. In addition, in the bootstrap exercise, we keep the instruments fixed.

6.1. Analysis under the null

We modify the bootstrap exercise by adding an initial step where we estimate the alphas for each of the four models (using $W=I$) and subtract the estimated alphas from the realized excess returns.



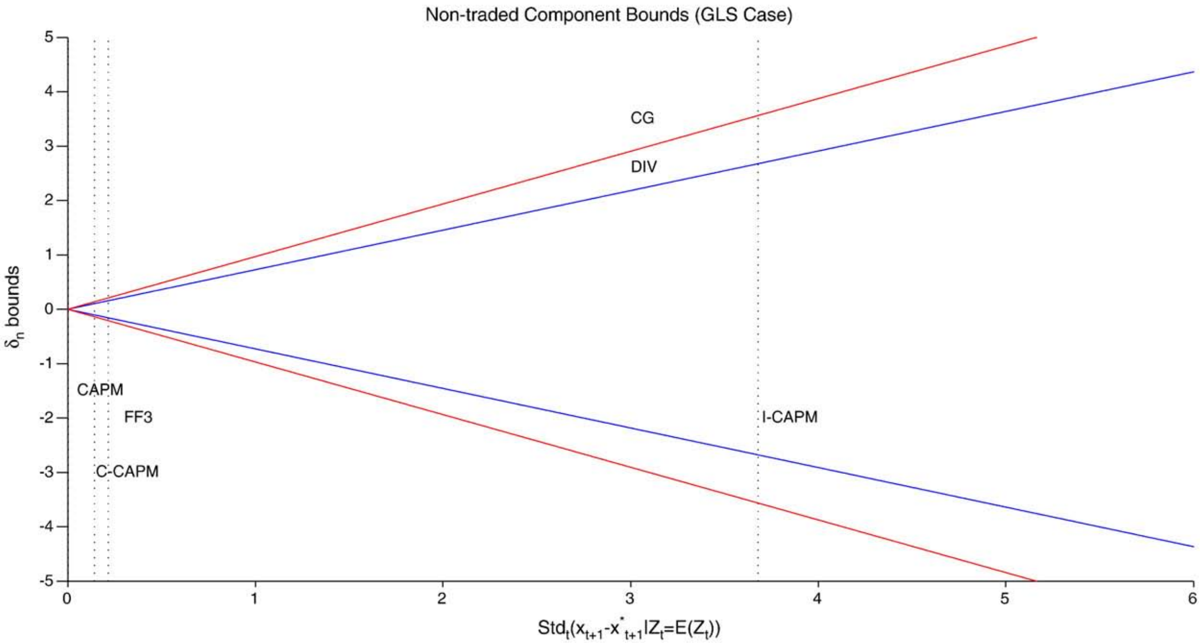


Fig. 1. Non-traded component bounds, δ_{nt}. We plot the bounds on the non-traded component of the risk premium λ_t for consumption growth (CG) and the dividend yield (DIV) as functions of the conditional standard deviation of $(x_{t+1}-x^*_{t+1})$ evaluated at $Z_t = E(Z_t)$. The vertical lines represent the conditional standard deviations of $(x_{t+1}-x^*_{t+1})$ evaluated at $Z_t=E(Z_t)$ for x = I-CAPM, C-CAPM, FF3, and CAPM. When $W=I$ (OLS case), the values on the horizontal axis are 4.9421, 0.7430, 0.2163, and 0.0028 for the I-CAPM, C-CAPM, FF3, and CAPM, respectively. When $W=\hat{\Sigma}_{rr}^{-1}$ (GLS case), the values on the horizontal axis are 3.6798, 0.1421, 0.2163, and 0.0028 for the I-CAPM, C-CAPM, FF3, and CAPM, respectively.

In summary, the results under the null are quantitatively and qualitatively similar to the results under the alternative. For example, consider the case of the C-CAPM and the consumption-growth risk premium. The reference values for λ^* are 0.0481 under the null, and 0.0092 under the alternative. The bias is −0.0006 under the null, and −0.0002 under the alternative. The empirical standard error is 0.0149 under the null, and 0.0165 under the alternative. Moving on to the λ estimates, the reference

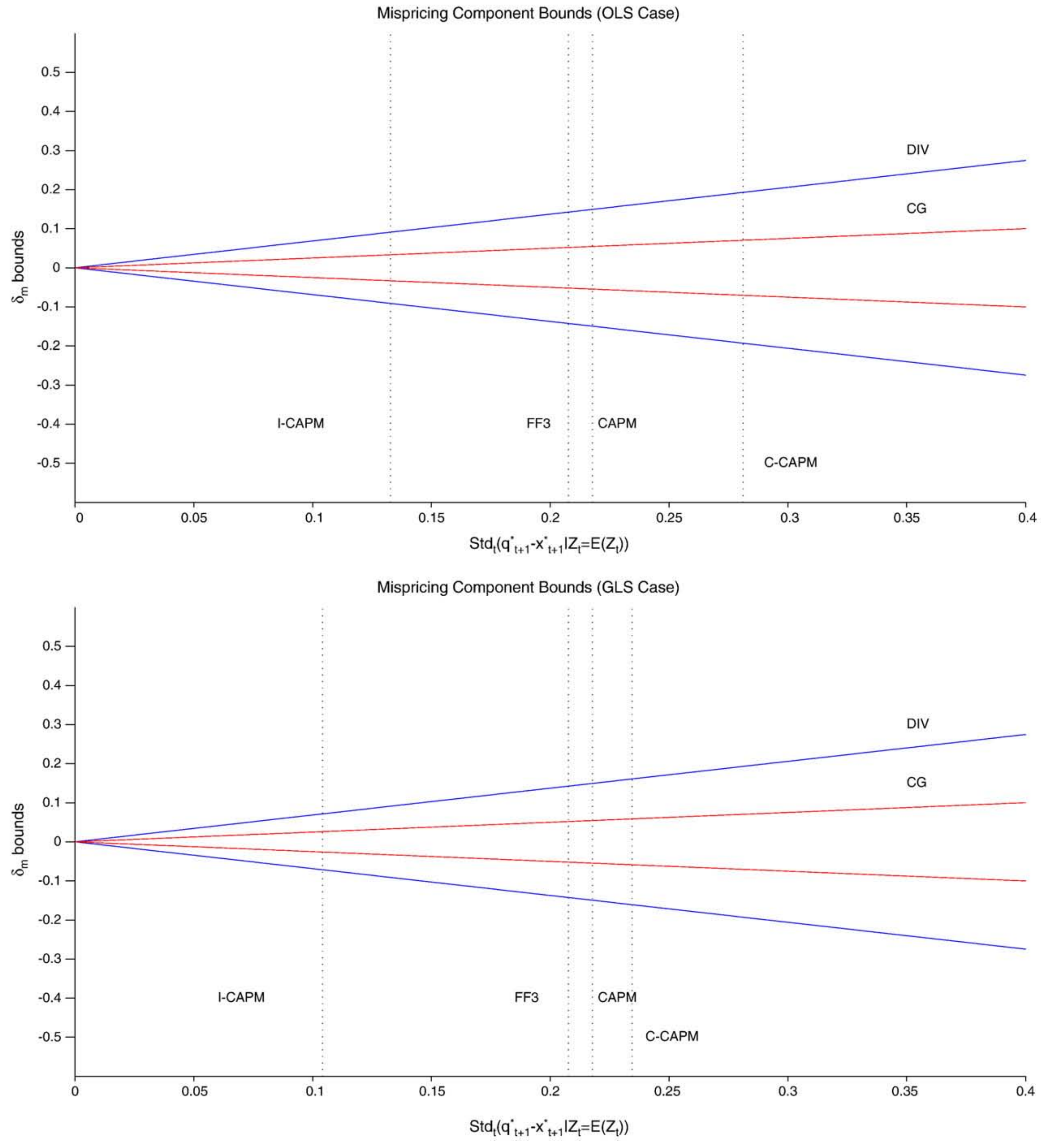


Fig. 2. Mispricing component bounds, δ_{mt}. We plot the bounds on the mispricing component of the risk premium λ_t for consumption growth (CG) and the dividend yield (DIV) as functions of the conditional standard deviation of $(q^*_{t+1} - x^*_{t+1})$ evaluated at $Z_t = E(Z_t)$. The vertical lines represent the conditional standard deviations of $(q^*_{t+1} - x^*_{t+1})$ evaluated at $Z_t = E(Z_t)$ for x = I-CAPM, C-CAPM, FF3, and CAPM. When $W = I$ (OLS case), the values on the horizontal axis are 0.1327, 0.2812, 0.2076, and 0.2177 for the I-CAPM, C-CAPM, FF3, and CAPM, respectively. When $W = \hat{\Sigma}_{rr}^{-1}$ (GLS case), the values on the horizontal axis are 0.1041, 0.2344, 0.2076, and 0.2177 for the I-CAPM, C-CAPM, FF3, and CAPM, respectively.

values for the OLS risk-premium estimate under the null and under the alternative are the same: 0.7674.[12] The bias is 0.0640 under the null, and 0.0566 under the alternative. The empirical standard error is 0.4462 under the null, and 0.4374 under the alternative.

[12] Recall that adjusting excess returns to make them consistent with a candidate pricing kernel leaves the total risk premium unaffected.

6.2. Changing the tradeability of the factors

In the single-factor C-CAPM model, we generate a "modified" consumption-growth factor y^m_{t+1},

$$y^m_{t+1} = \frac{y^\star_{t+1} + c(y_{t+1} - y^\star_{t+1})}{\sqrt{\sigma^2_{y^\star} + c^2\sigma^2_{y-y^\star}}}, \tag{64}$$

where $\sigma^2_{y-y\star}$ is the variance of the non-traded component of y_{t+1}. The modified factor has more or less variability than y_{t+1} depending on whether c is greater or less than one. The properties of returns are left unchanged (the simulation is under the alternative). On the other hand, as a result of working with standardized factors, the population values of both λ and $\lambda^\star$ change (see the Appendix). Indeed, as the tradeability of the factor increases, $\lambda^\star$ increases, while λ decreases (in absolute value). Analogous transformations are applied to obtain modified versions of the factors of the I-CAPM.

We focus on the C-CAPM and the consumption risk premium (results for the I-CAPM and the other factors are qualitatively similar). Consider the case where $c=1/4$. The most striking feature is the improvement in the properties of the λ estimate. In the OLS case, for example, the bias declines from 0.0566 to 0.0078, while the standard error goes from 0.4374 to 0.1207. Similarly, in the GLS case, the bias goes from −0.0314 to −0.0023, while the standard error declines from 0.2233 to 0.0650. When $c=4$, we obtain similar changes in the properties of the estimates, but in the opposite direction.[13]

7. Conclusions

Economic risk premia are the negative of the covariance between a normalized (mean-one) candidate pricing kernel and a set of risk factors. Since the factors are not traded, these premia depend on the specification of an asset pricing model. If the model is a linear factor model in the economic factors, then all security risk premia are linear in the economic risk premia.

We show how the risk premium assigned to a non-traded source of risk can be decomposed into three parts: i) the expected excess cash flow on the maximum-correlation portfolio; ii) (minus) the covariance between the non-traded components of the factor and of the candidate pricing kernel of a given model; and iii) (minus) the mispricing assigned by the candidate pricing kernel x_{t+1} to the maximum-correlation portfolio mimicking the factor.

We estimate the three components for four linear factor models: the intertemporal CAPM (I-CAPM), the consumption CAPM (C-CAPM), the Fama and French (1993) three-factor model (FF3), and the CAPM. We show how models with non-traded factors (I-CAPM and C-CAPM) assign risk premia that deviate substantially from the premia on maximum-correlation portfolios and are often unrealistically large (in absolute value). Moreover, the economic risk premia parameter estimates tend to exhibit large (absolute) biases and standard errors, and can be sensitive to the choice of the weighting matrix. On the other hand, the premia on maximum-correlation portfolios are realistic in magnitude, and tend to be estimated precisely and with little bias. We also show how the discrepancies between the estimates of economic risk premia and the estimates of premia on maximum-correlation portfolios are mainly due to the common non-traded variability of the factors and of the candidate pricing kernels. These patterns hold for both the constant and the time-varying component of the risk premia.

We conclude that the estimates of economic risk premia based on linear factor models with non-traded factors can be unrealistic and unreliable, especially when the factors involved are very "far" from being traded. On the other hand, the risk premia on maximum-correlation portfolios are an appealing alternative as they provide realistic indications of whether a factor is priced, are estimated precisely, and have the theoretical advantage of being invariant to a model's non-traded variability and mispricing. Focusing on these risk premia, we find that the only economic factor to be significantly priced is the dividend yield, with a negative Sharpe ratio that is comparable in absolute magnitude to the market index.

Appendix A. Estimation, candidate kernel and λ_t

We start with the case of a candidate kernel that contains traded and non-traded factors. Then, we specialize our results to the case of a kernel that contains only traded factors.

To impose the restriction that the traded factors $y_{1,t+1}$ are priced exactly by x_{t+1}, it is convenient to consider the residuals of a projection of the non-traded factors $y_{2,t+1}$ onto the *innovations* in the traded factors $y_{1,t+1}$:[14]

$$\epsilon_{2,t+1} \equiv y_{2,t+1} - \gamma^\top[y_{1,t+1} - E_t(y_{1,t+1})], \tag{65}$$

[13] While not reported in the paper (results are available upon request), we performed an empirical exercise where we estimate the consumption risk premium using annual data, following Jagannathan and Wang (2007). The most striking difference between the annual and the monthly analysis is how much closer consumption growth is to being traded. Accordingly, the non-traded component of the consumption risk premium is a much lower portion of the total premium for yearly than for monthly data, and the precision of the δ_n estimates is also much higher in the annual case.

[14] In the empirical analysis, we verified that our findings were robust to the orthogonalization of the traded and non-traded factors driving the candidate kernel. Specifically, we ignored the pricing kernel restriction induced by the factor being traded and obtained results that were quantitatively very close to the ones reported in the tables (results are available upon request).

where $E_t(y_{1,t+1}) = \mu_{y_1}^\top Z_t$. Hence, the candidate kernel x_{t+1} has the form

$$x_{t+1} = 1 - [y_{1,t+1} - E_t(y_{1,t+1})]^\top b_{1t} - \epsilon_{2,t+1}^\top b_{2t}, \tag{66}$$

where $b_{1t} = b_1^\top Z_t$ and $b_{2t} = b_2^\top Z_t$.

The set of moment conditions is given by

$$E(r_{t+1} - \beta_0 - \beta_1^\top y_{1,t+1}) = 0, \tag{67}$$

$$E[(r_{t+1} - \beta_0 - \beta_1^\top y_{1,t+1}) y_{1,t+1}^\top] = 0, \tag{68}$$

$$E[(y_{1,t+1} - \mu_{y_1}^\top Z_t) Z_t^\top] = 0, \tag{69}$$

$$E[\epsilon_{2,t+1}(y_{1,t+1} - \mu_{y_1}^\top Z_t)^\top] = 0, \tag{70}$$

$$E(y_{1,t+1} x_{t+1} Z_t^\top) = 0, \tag{71}$$

$$\Sigma_{\epsilon_2 r} W E[(r_{t+1} - \beta_1^\top y_{1,t+1}) x_{t+1} Z_t^\top] = 0, \tag{72}$$

$$E\{[(1 - x_{t+1}) y_{3,t+1} - \lambda_t] Z_t^\top\} = 0. \tag{73}$$

The first two sets of moments identify the parameters of a projection of r_{t+1} onto $y_{1,t+1}$ and a constant. The vector $r_{t+1} - \beta_1^\top y_{1,t+1}$ represents the component of the excess returns that is not priced by the traded factors $y_{1,t+1}$. This component is what we want to price with the non-traded factors $y_{2,t+1}$. The third set of moments identifies the parameters of the conditional expectation of the traded factors. The fourth set of moments identifies the parameters of the projection of $y_{2,t+1}$ onto the innovations in $y_{1,t+1}$. The fifth set of moments imposes the Shanken (1992) restriction and identifies the parameters b_1. The sixth set of moments sets to zero K_2 linear combinations of the orthogonality conditions $E[(r_{t+1} - \beta_1^\top y_{1,t+1}) x_{t+1} Z_t^\top]$, and identifies the parameters b_2. Finally, the last set of moments identifies the parameters λ.

The parameter estimates are given by

$$\hat{\mu}_{y_1} = [\hat{E}(Z_t Z_t^\top)]^{-1} \hat{E}(Z_t y_{1,t+1}^\top), \tag{74}$$

$$\hat{b}_1 = \hat{\mu}_{y_1} \hat{\Sigma}_{y_1 y_1}^{-1}, \tag{75}$$

$$\hat{b}_2 = (\hat{\mu}_r - \hat{\mu}_{y_1} \hat{\beta}_1) W \hat{\Sigma}_{r\epsilon_2} (\hat{\Sigma}_{\epsilon_2 r} W \hat{\Sigma}_{r\epsilon_2})^{-1}, \tag{76}$$

$$\hat{\lambda} = \hat{b}_1 \hat{\Sigma}_{y_1 y_3} + \hat{b}_2 \hat{\Sigma}_{\epsilon_2 y_3}, \tag{77}$$

where $\hat{\Sigma}_{y_1 y1} = \hat{E}[(y_{1,t+1} - \hat{\mu}_{y_1}^\top Z_t)(y_{1,t+1} - \hat{\mu}_{y_1}^\top Z_t)^\top]$, $\hat{\beta}_1 = \hat{\Sigma}_{y_1 y1}^{-1} \hat{\Sigma}_{y_1 r}$, $\hat{\Sigma}_{r\varepsilon_2} = \hat{E}[(r_{t+1} - \hat{\mu}_r^\top Z_t)\epsilon_{2,t+1}^\top]$, $\hat{\Sigma}_{\epsilon_2 y_3} = \hat{E}(\epsilon_{2,t+1} y_{3,t+1}^\top)$, and $\hat{\Sigma}_{y_1 y_3} = \hat{E}(y_{1,t+1} y_{3,t+1}^\top)$.

From the projection of x_{t+1} onto the span of excess returns, we obtain

$$x_{t+1}^\star = 1 - [y_{1,t+1}^\star - E_t(y_{1,t+1}^\star)]^\top b_{1t} - [\epsilon_{2,t+1}^\star - E_t(\epsilon_{2,t+1}^\star)]^\top b_{2t}, \tag{78}$$

where $y_{1,t+1}^\star = (\gamma_1^\star)^\top r_{t+1}$ satisfies the set of moment conditions

$$E[(y_{1,t+1} - y_{1,t+1}^\star) r_{t+1}^\top] = 0, \tag{79}$$

and $\epsilon_{2,t+1}^\star = (\gamma_2^\star)^\top r_{t+1}$ satisfies the set of moment conditions

$$E[(\epsilon_{2,t+1} - \epsilon_{2,t+1}^\star) r_{t+1}^\top] = 0. \tag{80}$$

Therefore, we have

$$\hat{\gamma}_1^\star = \hat{\Sigma}_{rr}^{-1} \hat{\Sigma}_{ry_1}, \tag{81}$$

$$\hat{\gamma}_2^\star = \hat{\Sigma}_{rr}^{-1} \hat{\Sigma}_{r\varepsilon_2}, \tag{82}$$

where $\hat{\Sigma}_{ry_1} = \hat{E}[(r_{t+1} - \hat{\mu}_r^\top Z_t)(y_{1,t+1} - \hat{\mu}_{y_1}^\top Z_t)^\top]$.

As to the non-traded and mispricing components, the parameter estimates are given by

$$\hat{\delta}_n = \hat{b}_1[\hat{\Sigma}_{y_1y_3} - (\hat{\gamma}_1^\star)^\top \hat{\Sigma}_{ry_3}] + \hat{b}_2[\hat{\Sigma}_{\varepsilon_2 y_3} - (\hat{\gamma}_2^\star)^\top \hat{\Sigma}_{ry_3}], \quad (83)$$

$$\hat{\delta}_m = [\hat{b}_1(\hat{\gamma}_1^\star)^\top + \hat{b}_2(\hat{\gamma}_2^\star)^\top - \hat{b}_r]\hat{\Sigma}_{ry_3}. \quad (84)$$

Note that the model with only the traded factors y_1 is nested within the model with traded and non-traded factors discussed here. As a result, the $\hat{b}_1$ estimates are the same as in Eq. (75), while the $\hat{\lambda}$, $\hat{\delta}_n$, and $\hat{\delta}_m$ estimates are given by

$$\hat{\lambda} = \hat{b}_1 \hat{\Sigma}_{y_1y_3}, \quad (85)$$

$$\hat{\delta}_n = \hat{b}_1[\hat{\Sigma}_{y_1y_3} - (\hat{\gamma}_1^\star)^\top \hat{\Sigma}_{ry_3}], \quad (86)$$

$$\hat{\delta}_m = [\hat{b}_1(\hat{\gamma}_1^\star)^\top - \hat{b}_r]\hat{\Sigma}_{ry_3}. \quad (87)$$

Appendix B. Changing the tradeability of the factors

The projection of y^m_{t+1} onto the augmented span of excess returns is

$$\frac{y^\star_{t+1}}{\sqrt{\sigma^2_{y^\star} + c^2\sigma^2_{y-y^\star}}} \quad (88)$$

and the premium on the maximum-correlation portfolio tracking the modified factor y^m_{t+1} is given by

$$\frac{\lambda^\star}{\sqrt{\sigma^2_{y^\star} + c^2\sigma^2_{y-y^\star}}}, \quad (89)$$

which decreases (in absolute value) as c increases. As to the pricing kernel based on y^m_{t+1}, we have

$$x^m_{t+1} \equiv 1 - b^m y^m_{t+1} = 1 - \frac{b^m}{\sqrt{\sigma^2_{y^\star} + c^2\sigma^2_{y-y^\star}}}[y^\star_{t+1} + c(y_{t+1} - y^\star_{t+1})], \quad (90)$$

where b^m denotes the population coefficient of the modified pricing kernel x^m_{t+1}. Since changing c does not affect the pricing properties of the kernel, we have

$$\frac{b^m}{\sqrt{\sigma^2_{y^\star} + c^2\sigma^2_{y-y^\star}}} = b \quad (91)$$

and

$$b^m = b\sqrt{\sigma^2_{y^\star} + c^2\sigma^2_{y-y^\star}}, \quad (92)$$

where b denotes the population coefficient of the original pricing kernel, $x_{t+1} = 1 - by_{t+1}$. Hence, as c increases, λ also increases (in absolute value). The same arguments above apply to the multivariate case of the I-CAPM.

References

Balduzzi, P., Kallal, H., 1997. Risk premia and variance bounds. Journal of Finance 52, 1913–1949.
Balduzzi, P., Robotti, C., 2008. Mimicking portfolios, economic risk premia, and tests of multi-beta models. Journal of Business and Economic Statistics 26, 354–368.
Breeden, D.T., 1979. An intertemporal asset pricing model with stochastic consumption and investment opportunities. Journal of Financial Economics 7, 265–296.
Breeden, D.T., Gibbons, M.R., Litzenberger, R.H., 1989. Empirical tests of the consumption-oriented CAPM. Journal of Finance 44, 231–262.
Burmeister, E., McElroy, M.B., 1988. Joint estimation of factor sensitivities and risk premia for the arbitrage pricing theory. Journal of Finance 43, 721–733.
Campbell, J.Y., 1996. Understanding risk and return. Journal of Political Economy 104, 298–345.
Campbell, J.Y., Viceira, L.M., 1996. Consumption and portfolio decisions when expected returns are time-varying. Quarterly Journal of Economics 114, 433–495.
Chen, N., Roll, R., Ross, S.A., 1986. Economic forces and the stock market. Journal of Business 59, 383–403.
Cochrane, J.H., 1996. A cross-sectional test of an investment-based asset pricing model. Journal of Political Economy 104, 572–621.
Cochrane, J.H., Saá-Requejo, J., 2000. Beyond arbitrage: good-deal asset price bounds in incomplete markets. Journal of Political Economy 108, 79–119.
DeRoon, F.A., Nijman, Th.E., 2001. Testing for mean-variance spanning: a survey. Journal of Empirical Finance 8, 111–155.
DeRoon, F.A., Nijman, Th.E., Werker, B.J.M., 1998. "Testing for Spanning with Futures Contracts and Nontraded Assets: A General Approach," Working Paper. Tilburg University.
Fama, E.F., 1996. Multifactor portfolio efficiency and multifactor asset pricing. Journal of Financial and Quantitative Analysis 31, 441–465.
Fama, E.F., French, K.R., 1993. Common risk factors in the returns on stocks and bonds. Journal of Financial Economics 33, 3–56.
Ferson, W.E., Harvey, C.R., 1991. The variation of economic risk premiums. Journal of Political Economy 99, 385–415.
Ferson, W.E., Harvey, C.R., 1999. Conditioning variables and the cross-section of stock returns. Journal of Finance 54, 1325–1360.
Ferson, W.E., Siegel, A., Xu, T., 2005. Mimicking portfolios with conditioning information. Journal of Financial and Quantitative Analysis 41, 607–636.
Hall, A.R., Inoue, A., 2003. The large sample behavior of the generalized method of moments estimator in misspecified models. Journal of Econometrics 114, 361–394.

Hansen, L.P., Jagannathan, R., 1991. Implications of security market data for models of dynamic economies. Journal of Political Economy 99, 225–262.
Hansen, L.P., Jagannathan, R., 1997. Assessing specification errors in stochastic discount factor models. Journal of Finance 52, 557–590.
Hou, K., Kimmel, R.L., 2006. "On Estimation of Risk Premia in Linear Factor Models," Working Paper. Princeton University.
Jagannathan, R., Wang, Z., 1996. The conditional CAPM and the cross-section of expected returns. Journal of Finance 51, 3–53.
Jagannathan, R., Wang, Y., 2007. Lazy investors, discretionary consumption, and the cross-section of stock returns. Journal of Finance 62, 1623–1661.
Kan, R., Robotti, C., 2008. Specification tests of asset pricing models using excess returns. Journal of Empirical Finance 15, 816–838.
Kan, R., and C. Robotti, 2009, Model comparison using the Hansen-Jagannathan distance. Review of Financial Studies 22, 3449–3490.
Kan, R., Robotti, C., Shanken, J., 2009. Pricing model performance and the two-pass cross-sectional regression methodology. *NBER* working paper, p. 15047.
Kandel, S., Stambaugh, R.F., 1995. Portfolio inefficiency and the cross-section of expected returns. Journal of Finance 50, 157–184.
Kirby, C., 1998. The restrictions on predictability implied by rational asset pricing models. Review of Financial Studies 11, 343–382.
Lamont, O., 2001. Economic tracking portfolios. Journal of Econometrics 105, 161–184.
Lewellen, J.W., S. Nagel, and J. Shanken, 2009, "A skeptical appraisal of asset-pricing tests," *Journal of Financial Economics*, forthcoming.
McElroy, M., Burmeister, E., 1988. Arbitrage pricing theory as a restricted nonlinear multivariate regression model: ITNLSUR estimates. Journal of Business and Economic Statistics 6, 29–42.
Merton, R., 1973. An intertemporal capital asset pricing model. Econometrica 41, 867–887.
Politis, D.N., Romano, J.P., 1994. The stationary bootstrap. Journal of the American Statistical Association 89, 1303–1313.
Shanken, J., 1992. On the estimation of beta-pricing models. Review of Financial Studies 5, 1–33.
Shanken, J., Zhou, G., 2007. Estimating and testing beta pricing models: alternative methods and their performance in simulations. Journal of Financial Economics 84, 40–86.
Van den Goorbergh, R.W.J., DeRoon, F.A., Werker, B.J.M., 2005. "Economic Hedging Portfolios," Working Paper. Tilburg University.