working paper
1711

Empirical Evaluation of Overspecified Asset Pricing Models

Elena Manresa
Francisco Peñaranda
Enrique Sentana

May 2017

CENTRO DE ESTUDIOS
MONETARIOS Y FINANCIEROS
Casado del Alisal 5, 28014 Madrid, Spain
www.cemfi.es

CEMFI Working Paper No. 1711
May 2017

Empirical Evaluation of Overspecified Asset Pricing Models

Abstract

Asset pricing models with potentially too many risk factors are increasingly common in empirical work. Unfortunately, they can yield misleading statistical inferences. Unlike other studies focusing on the properties of standard estimators and tests, we estimate the sets of SDFs and risk prices compatible with the asset pricing restrictions of a given model. We also propose tests to detect problematic situations with economically meaningless SDFs uncorrelated to the test assets. We confirm the empirical relevance of our proposed estimators and tests with Yogo's (2006) linearized version of the consumption CAPM, and provide Monte Carlo evidence on their reliability in finite samples.

JEL Codes: G12, C52.

Keywords: Continuously updated GMM, factor pricing models, set estimation, stochastic discount factor, underidentification tests.

Elena Manresa
MIT Sloan
emanresa@mit.edu

Francisco Peñaranda
Queens College CUNY
francisco.penaranda@qc.cuny.edu

Enrique Sentana
CEMFI
sentana@cemfi.es

Acknowledgement

We would like to thank Craig Burnside and Motohiro Yogo for kindly making their data available to us. We also thank Cesare Robotti, Enrique Salvador, participants at the Finance Forum (Zaragoza), the Midwest Finance Association Meeting (Chicago), the UK Money and Macro Finance Research Group Workshop on "Empirical Modelling of Financial Markets" (Brunel), the Annual SoFiE Conference (Aarhus), the International Conference on Computational and Financial Econometrics (Sevilla) and the Econometric Society World Congress (Montreal), as well as audiences at CEU Cardenal Herrera (Elche), Fundação Getulio Vargas (Rio), Geneva, Harvard Business School, LSE, McCombs School of Business (Austin) and Queens College CUNY for helpful comments and suggestions. Of course, the usual caveat applies. Financial support from the Spanish Government through grants ECO 2011-25607 (Peñaranda), 2011-26342 and 2014-59262 (Sentana) is gratefully acknowledged. The second author is also very grateful to SanFI for its funding.

1 Introduction

The most popular empirically oriented asset pricing models effectively assume the existence of a common stochastic discount factor (SDF) that is linear in some risk factors, which discounts uncertain payoffs differently across different states of the world. Those factors can be either the returns on some traded securities, non-traded economy wide sources of uncertainty related to macroeconomic variables, or a combination of the two. The empirical success of such models at explaining the so called CAPM anomalies was initially limited, but researchers have progressively entertained a broader and broader set of factors, which has resulted in several success claims. Harvey, Liu and Zhu (2016) contains a comprehensive and up to date list of references, cataloguing 315(!) different factors.

However, several authors have warned that some of those factors, or more generally linear combinations of them, could be uncorrelated with the vector of asset payoffs that they are meant to price, which would result in economically meaningless models (see Burnside (2016), Gospodinov, Khan and Robotti (2015) and the references therein). Further, those papers forcefully argue that such situations can lead to misleading econometric conclusions.

In this context, the purpose of our paper is to study the estimation of risk prices and the testing of the cross-sectional restrictions imposed by overspecified linear factor pricing models. By overspecified models we mean those with at least one non-zero SDF which is uncorrelated with the excess returns on the vector of test assets. We discuss in detail several examples of this situation, which illustrate two important differences between our work and related studies. First, the presence of uncorrelated risk factors is sufficient but not necessary for overspecification. As a result, attempts to find out which factors are uncorrelated on an individual basis fail to provide a complete answer. Second, overspecification is necessary but not sufficient for the model parameters to be underidentified. Therefore, studying parameter identification by means of rank tests does not provide a full answer either.

Our point of departure from the existing literature is that we do not focus exclusively on the properties of the usual estimators and tests. Instead, we use the econometric framework in Arellano, Hansen and Sentana (2012).[1] Thus, we can identify a linear subspace of risk prices compatible with the cross-sectional asset pricing restrictions, a basis of which we can easily parametrize and efficiently estimate using standard GMM methods.

We follow Peñaranda and Sentana (2015) in using single-step procedures, such as the continuously updated GMM estimator (CU-GMM) of Hansen, Heaton and Yaron (1996), to obtain numerically identical test statistics and risk price estimates for SDF and regression methods,

[1] In this sense, our paper can be regarded as a substantial extension of Manresa (2009).

with uncentred or centred moments and symmetric or asymmetric normalizations. GEL methods such as Empirical Likelihood or Exponentially Tilted also share the numerical invariance properties of CU-GMM. However, given that these methods are often more difficult to compute than two-step estimators, and they may sometimes give rise to multiple local minima, we propose simple, intuitive consistent parameter estimators that can be used as sensible initial values, and which will be efficient for elliptically distributed returns and factors. Interestingly, we can also show that these consistent initial values coincide with the GMM estimators recommended by Hansen and Jagannathan (1997), which use the second moment of returns as weighting matrix.

For simplicity of exposition, we initially focus on excess returns, but later on extend our analysis to cover gross returns too. Importantly, we show that single-step GMM procedures yield the same numerical results with both types of payoffs.

In addition to the usual overidentification test, which is informative about the existence of admissible SDFs, we propose simple tests that can diagnose economically meaningless but empirically relevant cases in which the expected values of all SDFs in the identified set are 0, which is equivalent to their being uncorrelated to the test assets. We refer to this situation as complete overspecification, which should not apply to credible empirical models.

In our empirical application, we investigate the potential overspecification of the three-factor extension of the Epstein and Zin (1989) version of the consumption CAPM in Yogo (2006) using quarterly data from the usual Fama and French cross-section of excess returns on size and book-to-market sorted portfolios. Aside from its undisputable influence on the subsequent literature, an important characteristic of this model is that his chosen risk factors were theoretically motivated and not the result of either an extensive search or a reverse engineering process. Nevertheless, the results we obtain with our novel inference procedures indicate that the admissible SDFs in the linearized version of this model lie on a two-dimensional subspace, so there is lack of identification. In addition, we cannot reject the null hypothesis that all those SDFs have zero means, which is tantamount to complete overspecification. Importantly, our simulations show that these empirical findings are not due to lack of power. On the contrary, if anything, our proposed tests tend to overreject for the sample size of this data set.

Our conclusions about the Yogo (2006) model sharpen the analysis in Lewellen, Nagel and Shanken (2010). These authors show that the strong factor covariance structure in the size and book-to-market portfolios implies that it should be possible to find models with macroeconomic factors that can price those test assets. In fact, we find not only one (up to scale) admissible SDF using Yogo's (2006) risk factors, but an entire two-dimensional subspace of SDFs which can price those assets, a situation that standard GMM asymptotic theory cannot cope with.

The rest of the paper is organized as follows. Section 2 introduces linear factor pricing models, and precisely characterizes their potential overspecification. Next, we present our econometric methodology in section 3. Then, we empirically analyze in detail the aforementioned asset pricing model in section 4 and report our simulation evidence in section 5. Finally, we summarize our conclusions and discuss some avenues for further research in section 6. Proofs of formal results and a detailed description of all the possible situations that may arise in models with up to three factors are relegated to appendix A, while appendix B contains the Monte Carlo design.

2 Overspecified Asset Pricing Models

2.1 Stochastic discount factors and moment conditions

Let $\mathbf{r}$ be a given $n \times 1$ vector of excess returns, whose means $E(\mathbf{r})$ we assume are not all equal to zero. Standard arguments such as lack of arbitrage opportunities or the first order conditions of a representative investor imply that

$$E(m\mathbf{r}) = \mathbf{0}$$

for some random variable m called SDF, which discounts uncertain payoffs in such a way that their expected discounted value equals their cost.

The standard approach in empirical finance is to model the SDF as an affine transformation of some $k < n$ observable risk factors $\mathbf{f}$, even though this ignores that m must be positive with probability 1 to avoid arbitrage opportunities, which would require non-linear specifications for m (see Hansen and Jagannathan (1991)). In particular, researchers typically express the pricing equation as

$$E[(a + \mathbf{b}'\mathbf{f})\mathbf{r}] = \mathbf{0} \tag{1}$$

for some coefficients $(a, \mathbf{b})$, which we can refer to as the intercept and slopes of the affine SDF $m = a + \mathbf{b}'\mathbf{f}$.

We can also estimate the SDF mean $c = E(m)$ by adding the moment condition

$$E(a + \mathbf{b}'\mathbf{f} - c) = 0, \tag{2}$$

which exactly identifies c for given $(a, \mathbf{b})$. A non-trivial advantage of this approach is that (1) and (2) are linear in $(a, \mathbf{b}, c)$.[2]

[2] There are two alternative popular approaches to test asset pricing models. One uses $Cov(\mathbf{r}, \mathbf{f})$ instead of $E(\mathbf{r}\mathbf{f}')$ in explaining the cross-section of risk premia, while the other one relies on the regression of $\mathbf{r}$ onto a constant and $\mathbf{f}$. Both require a higher number of parameters to estimate from a higher number of moments, and for that reason we shall not explicitly consider them in this paper. Nevertheless, it is straightforward to extend the results in Peñaranda and Sentana (2015) to our context, so as to prove that all three approaches provide numerically equivalent tests and prices of risk estimates when one uses single-step GMM procedures.

It is pedagogically convenient for our purposes to think about the restrictions the linear factor pricing model above imposes on the parameters $(a, \mathbf{b}, c)$ as we increase the number of assets we consider. For simplicity, we focus on the case of two pricing factors.

When $n = 1$, there is always a two dimensional linear space of admissible solutions, which can be regarded as the dual set to the combination line of expected excess returns and covariances with the risk factors that can be generated by leveraging r_1 up or down.

(Figure 1: One asset)

When $n = 2$, the two dimensional space generated by each asset will generally be different, so their intersection will be a straight line.[3]

(Figure 2: Two assets)

Three assets is the minimum number required to be able to reject the model. The reason is the following. If the asset pricing model does not hold, the three linear subspaces associated to each of the assets will only intersect at the origin. We may then say that there is financial markets segmentation, in the sense that there is no single SDF within the model that can price all the assets.

(Figure 3: Three segmented asset markets)

If on the other hand the asset pricing model holds, the intersection will be a linear subspace of positive dimension. This requires that the three assets are coplanar in the space of expected excess returns and covariances with the risk factors, so that they all lie on the security market plane. When this happens, we may say that there is financial markets integration.

(Figure 4: Three integrated asset markets)

Therefore, when there exist admissible parameter configurations other than the trivial one $(a, \mathbf{b}, c) = (0, \mathbf{0}, 0)$, we can at best identify a direction in $(a, \mathbf{b}, c)$ space, which leaves both the scale and sign of the SDF undetermined. As forcefully argued by Hillier (1990) for single equation IV models, this suggests that we should concentrate our efforts in estimating the identified direction. However, empirical researchers often prefer to estimate points rather than directions, and for that reason they typically focus on some asymmetric scale normalization, such as $(1, \mathbf{b}/a, c/a)$. In this regard, note that $\boldsymbol{\delta} = -\mathbf{b}/a$ can be interpreted as prices of risk

[3] Occasionally, though, the two linear subspaces might coincide. This will happen when the two assets are collinear in the space of expected excess returns and covariances with the risk factors. We will revisit this issue when we discuss Figure 8 in section 2.2.

since we may rewrite (1) as $E(\mathbf{r}) = E(\mathbf{rf}')\boldsymbol{\delta}$. Other normalizations, such as $(a/c, \mathbf{b}/c, 1)$ or $\mathbf{b}'\mathbf{b} + c^2 = 1$ are also popular.

(Figure 5: Normalizations)

Nevertheless, given that any asymmetric normalization is potentially restrictive, we prefer to use invariant estimation methods, such as CU-GMM.

In what follows, we consider models in which the elements of $\mathbf{f}$ are either non-traded (or treated as such) or they are portfolios of $\mathbf{r}$. In those cases, the pricing conditions (1) and (2) contain all the relevant information to estimate and test the asset pricing model. Nevertheless, it would be very easy to extend our analysis to explicitly deal with traded factors whose excess returns do not belong to the linear span of $\mathbf{r}$. In that case, we should add moment conditions such as

$$E[(a + \mathbf{b}'\mathbf{f})\mathbf{f}] = \mathbf{0}$$

to (1) and (2) to complete the asset pricing information that we should consider, as Lewellen, Nagel and Shanken (2010) suggest.

2.2 Admissible SDFs sets

The pricing conditions (1) can be expressed in matrix notation as

$$\begin{bmatrix} E(\mathbf{r}) & E(\mathbf{rf}') \end{bmatrix} \begin{pmatrix} a \\ \mathbf{b} \end{pmatrix} = \mathbf{M}\boldsymbol{\theta} = \mathbf{0}, \tag{3}$$

where $\mathbf{M}$ is an $n \times (k+1)$ matrix of first and second moments of data and $\boldsymbol{\theta}$ a $(k+1) \times 1$ parameter vector.

The highest possible rank of $\mathbf{M}$ is its number of columns $k+1$ because $k < n$. In that case, though, the asset pricing model will not hold because the only value of $\boldsymbol{\theta}$ that satisfies (3) will be the trivial solution $\boldsymbol{\theta} = \mathbf{0}$.

On the other hand, if the rank of $\mathbf{M}$ is k then there will be a one-dimensional subspace of $\boldsymbol{\theta}$'s that satisfy the pricing conditions (3), in which case the solution $\boldsymbol{\theta}$ is unique up to scale, as we explained in the previous section. Not surprisingly, $rank(\mathbf{M}) = k$ coincides with the usual identification condition required for standard GMM inference (see e.g. Hansen (1982) and Newey and McFadden (1994)).

Kan and Zhang (1999) and Burnside (2016) among others have forcefully argued that some empirical asset pricing models effectively rely on factors for which the matrix $Cov(\mathbf{r}, \mathbf{f})$ does not have full column rank. The best known example is a useless factor, which would yield a zero column in the matrix $Cov(\mathbf{r}, \mathbf{f})$. To understand the implications, consider a two-factor model with $Cov(\mathbf{r}, f_2) = \mathbf{0}$, so that the matrix $\mathbf{M}$ becomes

$$[\ E(\mathbf{r}) \quad E(\mathbf{r}f_1) \quad E(\mathbf{r}f_2)\ ] = [\ E(\mathbf{r}) \quad E(\mathbf{r}f_1) \quad E(\mathbf{r})E(f_2)\ ].$$

Given the rank failure of this matrix in those circumstances, we can always find at least a one-dimensional subspace of SDFs whose parameters satisfy (3). Two different situations might occur.

First, if $E(\mathbf{r})$ and $E(\mathbf{r}f_1)$ are linearly independent, then $rank(\mathbf{M}) = 2$, the model parameters will remain econometric identified, and we can still rely on standard GMM inference.[4] However, $E(\mathbf{r})$ and $E(\mathbf{r}f_1)$ linearly independent together with absence of arbitrage opportunities implies that the true SDF must depend at least on an additional genuine risk factor different from f_1 and f_2. As a result, there can be no admissible SDF affine in the two risk factors selected by the empirical researcher with a meaningful economic interpretation that can explain cross-sectional risk premia. Indeed, when $Cov(\mathbf{r}, f_2) = \mathbf{0}$ but $E(\mathbf{r}) \neq \mathbf{0}$, the SDF conditions (1) will trivially hold for any $m \propto [f_2 - E(f_2)]$ because they will all satisfy $E(m) = 0$ and $Cov(\mathbf{r}, m) = \mathbf{0}$, which in turn implies that

$$E(\mathbf{r}m) = E(\mathbf{r})E(m) + Cov(\mathbf{r}, m) = \mathbf{0}.$$

As a result, the admissible SDFs will have $b_1 = 0$ and $c = E(m) = 0$. Thus, this overspecified model is econometrically identified but economically unattractive.

(Figure 6: Valid but unattractive model with a useless factor)

Second, if f_1 were a valid pricing factor, so that $E(\mathbf{r}) = E(\mathbf{r}f_1)\delta_1$, then $rank(\mathbf{M}) = 1$ because

$$[\ E(\mathbf{r}) \quad E(\mathbf{r}f_1) \quad E(\mathbf{r}f_2)\ ] = E(\mathbf{r})[\ 1 \quad 1/\delta_1 \quad E(f_2)\ ].$$

Hence, this overspecified pricing model will be economically meaningful but parametrically underidentified.

(Figure 7: Valid and attractive model with a useless factor)

More generally, there will be rank failures in $Cov(\mathbf{r}, \mathbf{f})$ whenever we can find a valid asset pricing model with fewer factors even though no column of $Cov(\mathbf{r}, \mathbf{f})$ is zero. This will happen in particular if there is only a true pricing factor but we include two different noisy proxies for it, so that their difference will be uncorrelated to the vector of excess returns. As in the

[4] Nevertheless, some asymmetric normalizations may be incompatible with these configurations; see section 4.4 in Peñaranda and Sentana (2015) for further details in the case of a single pricing factor.

previous example, this alternative overspecified two-factor model is economically meaningful but parametrically underidentified even though the two candidate pricing factors are correlated with the returns on the test assets.

As a different example of the same type of overspecified model, assume that both the CAPM and the (linearized) CCAPM hold, in the sense that excess returns on the market and consumption growth can price on their own a cross-section of excess returns, i.e. $E(\mathbf{r}) = E(\mathbf{r}f_1)\delta_1$ and $E(\mathbf{r}) = E(\mathbf{r}f_2)\delta_2$. Then, an SDF that is linear in both factors, as in the (linearized) Epstein-Zin model, implies that $rank[(Cov(\mathbf{r},\mathbf{f})] = 1$. As a consequence, the matrix $\mathbf{M}$ becomes

$$\begin{bmatrix} E(\mathbf{r}) & E(\mathbf{r}f_1) & E(\mathbf{r}f_2) \end{bmatrix} = E(\mathbf{r})\begin{pmatrix} 1 & 1/\delta_1 & 1/\delta_2 \end{pmatrix},$$

which means that we can find a two-dimensional subspace of SDFs whose parameters satisfy $\mathbf{M}\boldsymbol{\theta} = \mathbf{0}$, a basis of which will be given by $m_1 = 1-\delta_1 f_1$ and $m_2 = 1-\delta_2 f_2$, with $c_1 = 1-E(f_1)\delta_1$ and $c_2 = 1-E(f_2)\delta_2$ both different from 0 because f_1 and f_2 are correlated with $\mathbf{r}$.

(Figure 8: Two single factor models)

Finally, there will also be a two-dimensional subspace of SDFs whose parameters satisfy $\mathbf{M}\boldsymbol{\theta} = \mathbf{0}$ when there are two useless factors, i.e. $Cov(\mathbf{r}, f_1) = Cov(\mathbf{r}, f_2) = \mathbf{0}$. Specifically, if

$$\begin{bmatrix} E(\mathbf{r}) & E(\mathbf{r}f_1) & E(\mathbf{r}f_2) \end{bmatrix} = E(\mathbf{r})\begin{bmatrix} 1 & E(f_1) & E(f_2) \end{bmatrix},$$

then any SDF which is a linear combination of $[f_1 - E(f_1)]$ and $[f_2 - E(f_2)]$ will work.

(Figure 9: Two useless factors)

The special feature of this completely overspecified case is that $c = 0$ for all admissible SDFs, so there is not only underidentification but also the absence of any meaningful specification.

3 Econometric methodology

3.1 Set estimation

Given our previous discussion, it is of the utmost importance to use statistical inference tools that can successfully deal with situations in which $rank(\mathbf{M}) \leq k$. Following Arellano, Hansen and Sentana (2012), we begin by specifying the dimension of the subspace of solutions to the pricing conditions (3), which we denote d, so that $rank(\mathbf{M}) = (k+1) - d$. Given that we maintain the hypothesis that $E(\mathbf{r}) \neq \mathbf{0}$, we could in principle consider ranks for $\mathbf{M}$ as low as 1 or, equivalently, any positive integer d up to a maximum value of k.

As we mentioned before, when $d = 1$ we can rely on standard GMM to estimate a unique $\boldsymbol{\theta}$ (up to normalization) and use its associated J test to assess the validity of the asset pricing restrictions. However, when $d \geq 2$, we will have a multidimensional subspace of admissible SDFs even after fixing their scale. Nevertheless, we can efficiently estimate a basis of that subspace by replicating d times the moment conditions (3) as follows:

$$\left.\begin{array}{c} [\ E(\mathbf{r}) \quad E(\mathbf{rf}')\ ]\boldsymbol{\theta}_1 = \mathbf{0}, \\ [\ E(\mathbf{r}) \quad E(\mathbf{rf}')\ ]\boldsymbol{\theta}_2 = \mathbf{0}, \\ \vdots \\ [\ E(\mathbf{r}) \quad E(\mathbf{rf}')\ ]\boldsymbol{\theta}_d = \mathbf{0}, \end{array}\right\} \tag{4}$$

and imposing enough normalizations on $(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, ..., \boldsymbol{\theta}_d)$ to ensure the point identification of a basis of the null space of $\mathbf{M}$.

In this setting, the familiar J test from the work of Sargan (1958) and Hansen (1982) for overidentification of the augmented model becomes a test for "underidentification" of the original model. The rationale is as follows: if we can identify a linear subspace of risk prices without statistical rejection, then the original asset pricing model is not well identified. In contrast, a statistical rejection provides evidence that the prices of risk in the original model are indeed point identified, unless of course the familiar J test continues to reject its overidentifying restrictions.

We can also add moment conditions to estimate $(c_1, c_2, \ldots, c_d)$, which characterize the expected values of the basis SDF's. Specifically, we can combine (4) with the moment conditions

$$\left.\begin{array}{c} [\ 1 \quad E(\mathbf{f}')\ ]\boldsymbol{\theta}_1 - c_1 = 0, \\ [\ 1 \quad E(\mathbf{f}')\ ]\boldsymbol{\theta}_2 - c_2 = 0, \\ \vdots \\ [\ 1 \quad E(\mathbf{f}')\ ]\boldsymbol{\theta}_d - c_d = 0, \end{array}\right\} \tag{5}$$

which are exactly identified for given values of $(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, ..., \boldsymbol{\theta}_d)$.

3.2 Normalizations and starting values

In the presentation of our empirical results, we will use the popular SDF normalization discussed in section 2.1, which fixes the first element of $\boldsymbol{\theta}_i$ to 1, thereby defining the prices of risk as $\boldsymbol{\delta}_i = -\mathbf{b}_i/a_i$. Additionally, we need to impose enough zero restrictions on the prices of risk to achieve identification.[5] Once again, though, the advantage of CU-GMM and other GEL estimators is that our inferences will be numerically invariant to the chosen normalization.

[5] Alternatively, we could make a $d \times d$ block of (a permutation of) the matrix $(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, ..., \boldsymbol{\theta}_d)$ equal to the identity matrix of order d.

Nevertheless, one drawback of these single-step methods is that they involve a non-linear optimization procedure even though the moment conditions are linear in parameters, which may result in multiple local minima. For that reason, we propose to use as starting value a computationally simple intuitive estimator that is always consistent, but which would become efficient when the returns and factors are *i.i.d.* elliptical. This family of distributions includes the multivariate normal and Student t distributions as special cases, which are often assumed in theoretical and empirical finance.

Let us define $(\mathbf{f}_1, \mathbf{f}_2, ..., \mathbf{f}_d)$ as the vectors of factors that enter each one of the SDFs in (4) after imposing the necessary restrictions that guarantee the point identification of the basis of risk prices $(\boldsymbol{\delta}_1, \boldsymbol{\delta}_2, ..., \boldsymbol{\delta}_d)$, so that the corresponding Jacobian matrices $E(\mathbf{r}\mathbf{f}_i')$ have full rank. As a result, we can re-write (4) as

$$E\begin{bmatrix}(1-\mathbf{f}_1'\boldsymbol{\delta}_1)\mathbf{r}\\(1-\mathbf{f}_2'\boldsymbol{\delta}_2)\mathbf{r}\\\vdots\\(1-\mathbf{f}_d'\boldsymbol{\delta}_d)\mathbf{r}\end{bmatrix}=\mathbf{0}, \tag{6}$$

and (5) as

$$E(1-\mathbf{f}_i'\boldsymbol{\delta}_i-c_i)=0, \qquad i=1,2,..,d. \tag{7}$$

Let $\mathbf{r}_t$ and $\mathbf{f}_t$ denote the values of the excess returns on the n assets and the k factors at time t. We can then prove that

Proposition 1 *If $(\mathbf{r}_t, \mathbf{f}_t)$ is an i.i.d. elliptical random vector with bounded fourth moments such that (6) holds, then:*
a) The most efficient GMM estimator of $\boldsymbol{\delta}_i$ $(i=1,\dots,d)$ from the system (6) will be given by

$$\mathring{\boldsymbol{\delta}}_{iT}=\left(\sum_{t=1}^{T}\tilde{\mathbf{r}}_{it}^{+}\tilde{\mathbf{r}}_{it}^{+\prime}\right)^{-1}\sum_{t=1}^{T}\tilde{\mathbf{r}}_{it}^{+}, \tag{8}$$

where $\tilde{\mathbf{r}}_{it}^{+}$ are the relevant elements of the sample factor mimicking portfolios

$$\tilde{\mathbf{r}}_{t}^{+}=\left(\sum_{s=1}^{T}\mathbf{f}_s\mathbf{r}_s'\right)\left(\sum_{s=1}^{T}\mathbf{r}_s\mathbf{r}_s'\right)^{-1}\mathbf{r}_t. \tag{9}$$

b) When we combine the moment conditions (6) with (7), the most efficient GMM estimator of each $\boldsymbol{\delta}_i$ is the same as in a), and the most efficient GMM estimator each c_i is the sample mean of the corresponding SDF.

Intuitively, Proposition 1 states that the optimal GMM estimator in an elliptical setting is such that it prices without error the factor mimicking portfolios in any given sample.

Although the elliptical family is rather broad (see Fang, Kotz and Ng (1990)), it is important to stress that (8) will remain consistent under correct specification even if the assumptions of serial independence and a multivariate elliptical distribution do not hold in practice.

In addition, we can provide a rather different justification for (8). Specifically, we can prove that $\mathring{\boldsymbol{\delta}}_{iT}$ in (8) coincides with the GMM estimator that we would obtain if we used as weighting matrix the second moment of the vector of excess returns $\mathbf{r}$. In other words, $\mathring{\boldsymbol{\delta}}_{iT}$ minimizes the sample counterpart to the Hansen and Jagannathan (1997) distance

$$E\left[\left(1-\mathbf{f}_i'\boldsymbol{\delta}_i\right)\mathbf{r}\right]'\left[E\left(\mathbf{r}\mathbf{r}'\right)\right]^{-1}E\left[\left(1-\mathbf{f}_i'\boldsymbol{\delta}_i\right)\mathbf{r}\right]$$

irrespective of the distribution of returns and the validity of the asset pricing model. The reason is that the f.o.c. of this minimization is

$$E\left(\mathbf{f}_i\mathbf{r}'\right)\left[E\left(\mathbf{r}\mathbf{r}'\right)\right]^{-1}E\left[\left(1-\mathbf{f}_i'\boldsymbol{\delta}_i\right)\mathbf{r}\right]=\mathbf{0},$$

which is equivalent to the exact pricing of the factor mimicking portfolios in Proposition 1.

3.3 Testing restrictions on admissible SDF sets

As we have just seen, our inference framework allows us to estimate the set of SDFs that is compatible with the pricing conditions (1). But we can also use it to test if the elements of this set satisfy some relevant restrictions.

A particularly important null hypothesis that empirical researchers would like to find evidence against is that all SDFs compatible with the data have zero means, a situation we have termed "complete overspecification". In that case, there will be no element in the admissible SDF set that explains the cross-section of expected returns from a meaningful economic perspective, as we illustrated in section 2.2 for $d=1$ and $d=2$ in Figures 6 and 9, respectively. Both those figures show completely overspecified models in which all the SDFs in the corresponding admissible set are uncorrelated with the asset payoffs, which renders them economically uninteresting.

In any given sample, though, the estimated values of the means of the admissible SDFs will not be 0. Given that the SDF means are associated to the parameters $(c_1, c_2, ..., c_d)$ by virtue of (7), a distance metric (DM) test of $H_0 : c_i = 0 \; i = 1, \ldots, d$ will give us a valid test of the null hypothesis of complete overspecification. As is well known, a DM test simply compares the GMM criterion functions (J statistics) with and without those constraints. We can trivially compute the criterion function without the zero mean constraints from the system (6), or equivalently, from the joint system that also considers the exactly identified moment conditions (7). In turn, we can construct the criterion function that imposes the zero mean constraints on all the SDFs from the system

$$E\begin{bmatrix}(1-\mathbf{f}_i'\boldsymbol{\delta}_i)\mathbf{r}\\ 1-\mathbf{f}_i'\boldsymbol{\delta}_i\end{bmatrix}=E\left[(1-\mathbf{f}_i'\boldsymbol{\delta}_i)\mathbf{x}\right]=0, \qquad i=1,2,..,d, \tag{10}$$

where $\mathbf{x}' = (\mathbf{r}', 1)$, which is analogous to (6) for an extended vector of payoffs that includes a fictional unit safe payoff.[6]

Again, normalization-invariant procedures are crucial to avoid obtaining different results for different basis of the admissible SDF set. But given the numerical complications that they may entail, we again propose to use as starting value a computationally simple intuitive estimator that is always consistent, but which would become efficient when the returns and factors are *i.i.d.* elliptical. In fact, we can prove that the optimal estimator of the prices of risk continues to have the same structure as in Proposition 1 if we define the factor mimicking portfolios over the extended payoff space. Specifically:

Proposition 2 *If $(\mathbf{r}_t, \mathbf{f}_t)$ is an i.i.d. elliptical random vector with bounded fourth moments such that (10) holds, then the most efficient GMM estimator of $\boldsymbol{\delta}_i$ $(i = 1, \ldots, d)$ will be given by*

$$\dot{\boldsymbol{\delta}}_{iT} = \left(\sum_{t=1}^{T} \tilde{\mathbf{x}}_{it}^{+} \tilde{\mathbf{x}}_{it}^{+\prime}\right)^{-1} \sum_{t=1}^{T} \tilde{\mathbf{x}}_{it}^{+}, \tag{11}$$

where $\tilde{\mathbf{x}}_{it}^{+}$ are the relevant elements of the sample factor mimicking portfolios

$$\tilde{\mathbf{x}}_{it}^{+} = \left(\sum_{s=1}^{T} \mathbf{f}_s \mathbf{x}_s'\right) \left(\sum_{s=1}^{T} \mathbf{x}_s \mathbf{x}_s'\right)^{-1} \mathbf{x}_t. \tag{12}$$

Another interesting null hypothesis that we may also want to test is whether some particular pricing factor does not appear in any admissible SDF. Formally, the corresponding null hypothesis would be that the entry of b associated to this factor being zero in all the vectors $(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, ..., \boldsymbol{\theta}_d)$.

(Figure 10: An unpriced second factor)

Again, a DM test based on single-step GMM procedures will be ideally suited for testing this restriction on the space of admissible SDFs.[7]

3.4 Comparison to the existing literature

Burnside (2016) and Gospodinov, Kan and Robotti (2015) study the identification of the prices of risk of the linear factor pricing model (1) (or its centred version in (A1), with $m =$

[6] If there really existed an unconditionally safe asset, an SDF that satisfied $E(\mathbf{x}m) = \mathbf{0}$ would allow for arbitrage opportunities in the extended payoff space. Although no such an asset exists in real life, the fact that all the SDFs in the admissible set satisfy those moment conditions signals the problematic economic interpretation of a completely overspecified model.

[7] Given that the moment conditions (5) and (6) are linear in parameters and the restrictions to test are homogenous, the results in Newey and West (1987b) imply that the Wald, Lagrange Multiplier and DM tests would be numerically identical for two-step GMM methods that shared the same weighting matrices. More generally, though, DM tests might be more reliable than Wald tests in non-standard situations with potential identification failures (see Dufour (1997) for closely related results in a likelihood context).

$c + \mathbf{b}'(\mathbf{f} - \boldsymbol{\mu})$, where $\boldsymbol{\mu} = E(\mathbf{f})$ is the vector of risk factor means) by applying the tests proposed by Cragg and Donald (1997) and Kleibergen and Paap (2006) to assess the rank of $E(\mathbf{rf}')$ (or $Cov(\mathbf{r}, \mathbf{f})$), which coincide with the expected Jacobian matrices of the GMM conditions (3) (or their centred counterparts).

As we mentioned in the introduction, though, overspecification is a necessary but not sufficient condition for underidentification. In that regard, we can prove the following result:

Proposition 3 *The CU version of the overidentification test of the original SDF moment conditions (4) and (5) after imposing the d restrictions* $c_1 = \ldots = c_d = 0$ *numerically coincides with the CU version of the test of the null hypothesis* $rank[Cov(\mathbf{r}, \mathbf{f})] = k - d$.

As a result, the DM test of $c_1 = ... = c_d = 0$ we introduced in the previous section can be interpreted as a test of the null hypothesis that $rank[Cov(\mathbf{r}, \mathbf{f})] = k - d$ under the maintained hypothesis that $rank(\mathbf{M}) = (k + 1) - d$. In other words, this DM test implicitly checks whether the difference in the rank of those two matrices is one, or equivalently, whether $E(\mathbf{r})$ cannot be spanned by $Cov(\mathbf{r}, \mathbf{f})$, in which case the only admissible SDFs would be those economically meaningless random variables that exploit the rank failure in $Cov(\mathbf{r}, \mathbf{f})$ in setting to zero the pricing conditions (1).

In contrast, the test of the rank of $Cov(\mathbf{r}, \mathbf{f})$ in Proposition 3 is often uninformative about the existence of economically meaningful SDFs precisely because it does not maintain any hypothesis on the rank of $\mathbf{M}$.[8] For example, both in the model with a valid factor and a useless one depicted in Figure 7, and in the double single factor model described in Figure 8, the matrix $Cov(\mathbf{r}, \mathbf{f})$ has rank 1 instead of 2 while $\mathbf{M}$ has rank 1 instead of 3. As a result, $E(\mathbf{r})$ belongs to the span of $Cov(\mathbf{r}, \mathbf{f})$, which confirms that in those two examples there exist economically meaningful SDFs that correctly price $\mathbf{r}$.

On the other hand, Figure 6 shows another possible situation with $rank[Cov(\mathbf{r}, \mathbf{f})] = 1$ instead of 2 in which $E(\mathbf{r})$ cannot be spanned by $Cov(\mathbf{r}, \mathbf{f})$ because the rank of $\mathbf{M}$ is 2, so that the only admissible SDFs must be uncorrelated to the vector of excess returns.

The other main difference with those papers is that they focus on the implications of those rank failures for standard GMM procedures, which assume point identification, while we propose alternative inference procedures that explicitly handle set identification.

3.5 Gross returns

Let $\mathbf{R}$ denote a vector of gross returns on $N = n + 1$ assets. Without loss of generality, we can understand the vector of excess returns $\mathbf{r}$ that we have used so far as the difference between

[8] The only exception is the extreme case of $Cov(\mathbf{r}, \mathbf{f}) = \mathbf{0}$, which necessarily means $rank(\mathbf{M}) = 1$ when $E(\mathbf{r}) \neq \mathbf{0}$, making it impossible to find meaningful SDFs that can explain $E(\mathbf{r})$, as Figure 9 illustrates.

the gross returns of the last n assets and the first one, R say. In practice, this reference asset could be the real return on US T-bills, whose payoffs are not constant. The relevant pricing equation for $\mathbf{R}$ becomes:

$$E[\mathbf{R}(a + \mathbf{b}'\mathbf{f})] = \boldsymbol{\ell},$$

where $\boldsymbol{\ell}$ is a vector of N ones. Without loss of generality, we can re-write these moment conditions as the combination of (1) with:

$$E\left[R(a + \mathbf{b}'\mathbf{f})\right] = 1. \tag{13}$$

In addition, we can continue to estimate the SDF mean from the moment condition (2).

The addition of the pricing of R in (13) implies that we no longer require an arbitrary normalization of $(a, \mathbf{b}, c)$. As Peñaranda and Sentana (2015) prove in their Appendix A, though, the empirical evidence obtained by single-step methods applied to $\mathbf{R}$ is consistent with the analogous evidence obtained from $\mathbf{r}$ alone. In particular, the overidentification restriction test for the joint system (1) and (13) is numerically identical to the one for (1) alone, and the normalized estimate of $-\boldsymbol{\delta}$ obtained from the moment conditions for excess returns coincides with the ratio of the estimates of $\mathbf{b}$ to a obtained using all the assets. Intuitively, the addition of gross returns allows us to pin down a and the mean of the SDF, c, but otherwise, it simply re-scales this variable.

More formally, we can re-express the SDF as $a(1 - \mathbf{f}'\boldsymbol{\delta})$ and re-write the moment conditions (1) and (13) as $E[(1 - \mathbf{f}'\boldsymbol{\delta})\mathbf{r}] = \mathbf{0}$ and $E[a(1 - \mathbf{f}'\boldsymbol{\delta})R] = 1$, respectively. This last equation does not contain any additional information about the prices of risk $\boldsymbol{\delta}$, it only pins down a, or equivalently c.

Finally, the same comments apply to those situations with $d > 1$. The only difference is that they involve several SDFs of the form $a_i(1 - \mathbf{f}_i'\boldsymbol{\delta}_i)$ for $i = 1, \ldots, d$. But since we add one moment and one parameter for each dimension, the equivalence between the results for excess and gross returns we have just discussed for $d = 1$ continues to hold for any d.

4 Empirical Application

4.1 Original results

To illustrate the practical relevance of our proposed methods, in this section we analyze in detail the linear version of the three-factor model in Yogo (2006) in order to assess the potential overspecification of this popular empirical model.

As is well known, his theoretical model extends the CCAPM by assuming recursive preferences over a consumption bundle of nondurable and durable goods.[9] Therefore, in the linearized

[9] Eichenbaum and Hansen (1990) were the first authors to empirically entertain the idea that it might be

version of his model, the SDF will depend on three factors: the market return, and the consumption growth of nondurables and durables, so that we can write it as:

$$m = a\left(1 - \delta_1 f_1 - \delta_2 f_2 - \delta_3 f_3\right).$$

In practice, the log-growth rate of US real per capita consumption of nondurables and services and durables are identified with f_2 and f_3, respectively. In turn, the return on wealth - proxied by the (log) return on the value-weighted U.S. stock market measured in real terms - is associated with f_1.

We initially evaluate this model with the original data, which corresponds to quarterly excess returns on the Fama-French cross-section of 25 size and book-to-market sorted portfolios from 1951 to 2001 (see Fama and French (1993) for further details).[10] In addition to the insightful nature of Yogo's (2006) theoretically motivated SDF specification, his results became very influential because not only did he fail to reject the asset pricing restrictions but also because he succeeded in aligning the risk premia in the data with the risk premia generated by his model.

(Figure 11: Risk premia from 2S-GMM)

Nevertheless, the theoretical results in Burnside (2016) and Gospodinov, Kan and Robotti (2015) indicate that a high cross-sectional R^2 may spuriously arise in models with useless factors.

As an aside, we find that the results in Figure 11 depend on the estimation method (2-step GMM) and the imposition of some restrictions on the prices of risk.[11] Specifically, if we use instead iterated GMM starting from the 2-step estimates, we encounter a cycle with four different solutions.

(Figure 12: Risk premia from IT-GMM)

Convergence does not improve if we free up the price of risk coefficients: iterated GMM enters yet another cycle of three different solutions.

(Figure 13: Risk premia from IT-GMM, free coefficients)

These discrepancies highlight the advantages of single step methods.

necessary to look at different consumption subaggregates to successfully explain asset risk premia.Yogo's (2006) empirical model goes one important step further by combining their ideas with those in Epstein and Zin (1989).

[10]Note that although the market return is a traded factor, we do not add its pricing condition to (1) because it can effectively be generated as a portfolio of the cross-section of excess returns that we want to price.

[11]In this section, we follow Yogo (2006) in using the moment conditions of the centred SDF approach mentioned in footnote 2, which are given by equation (A1) in the proof of Proposition 2 with the normalization $c = 1$. As Peñaranda and Sentana (2015) argue, this matters for two step and iterated GMM, but not for CU-GMM.

4.2 Overspecification

The sensitivity of the empirical results to the estimation method might be a sign of overspecification. For that reason, we apply our methodology to the same data. Specifically, we use the moment conditions (6) with $d = 1$, 2 and 3 to test for one, two and three-dimensional linear subsets of valid SDFs, respectively. In all cases, we augment those moment conditions with the exactly identified moment conditions (7) to obtain the associated SDF means. As we mentioned in section 3.3, we can then assess whether the model is completely overspecified by testing the joint significance of those means.

We estimate the different subspaces for risk prices and SDFs using single-step GMM methods choosing those normalizations which are arguably easiest to interpret in each context. In the case of $d = 2$, in particular, we present the results for the simple normalization of the prices of risk given by $(\delta_1, \delta_2, 0)$ and $(\delta_1, 0, \delta_3)$.[12] Since the first factor is the market, we can interpret those two SDFs as two variants of the linearized Epstein and Zin (1989) model, one with nondurable consumption and another with durable consumption. In contrast, in the case of $d = 3$ we present the results for the simple normalization $(\delta_1, 0, 0)$, $(0, \delta_2, 0)$ and $(0, 0, \delta_3)$, which effectively imposes that each factor can separately explain risk premia.

Table 1 shows the results of our overspecification analysis of the model. This table displays estimates of the SDF parameters and associated J tests. The criterion function is also reported under the restriction of zero SDF means, which is equivalent to a rank test for $Cov(\mathbf{r}, \mathbf{f})$ from Proposition 3. The p-values of the different J tests are shown in parenthesis.

We complement the J tests with significance tests for the SDF prices of risk. In particular, to the right of the point estimates we report in parenthesis the p-value of the DM test of the null hypothesis of a zero parameter value. The first, second and third blocks of columns refer to SDF sets of dimension 1, 2 and 3, respectively. All the reported results correspond to a weighting matrix a la Newey and West (1987a) with one lag, but we obtain qualitatively similar conclusions when we use a VARHAC procedure also with one lag.[13]

(Table 1: Empirical evaluation of the model 1951-2001)

The results for the one-dimensional set entirely agree with the results in Yogo (2006), who finds that (i) the J test of two-step GMM does not reject his model for these 25 size- and value-sorted portfolios and (ii) durable consumption provides the only non-zero price of risk. In

[12] This normalization is identified as long as $\delta_2 \neq 0$ and $\delta_3 \neq 0$. In that regard, Table 1 below shows that the DM tests that we proposed in section 3.3 reject both $\delta_2 = 0$ and $\delta_3 = 0$.

[13] Den Haan and Levin's (1997) VARHAC procedure assumes that the moment conditions have a finite VAR representation, which they exploit to estimate the required long run covariance matrix.

this regard, the usual overidentification test reported in the first column of Table 1 does not reject the null hypothesis that there exists an SDF affine in the three factors that can price the cross-section of securities (p-value=53.7%).

However, the validity of the asymptotic distribution of this J test crucially depends on the model parameters being point identified. For that reason, we also report the overidentification test for $d = 2$. As explained before, this test assesses whether there is a linear subspace of dimension 2 of admissible SDFs that can price the cross section of risk premia. We obtain a p-value of 13.4%, which suggests that a model with the market together with durable and nondurable consumption as risk factors might be overspecified. In contrast, the overidentification test corresponding to $d = 3$ is rejected.

Therefore, we find evidence that the admissible SDFs of this model lie on a two-dimensional subspace. In addition, the DM test of the null hypothesis that all the admissible SDFs have zero means when $d = 2$ has a p-value of 49.4%. This suggests that the seeming pricing ability of this set of SDFs simply exploits the lack of correlation of its elements with $\mathbf{r}$. In other words, the vector of risk premia does not appear to lie in the span of the covariance matrix of the excess returns and the factors, which suggests the model is completely overspecified.

Our results are in line with Burnside (2016), who finds that the matrix $Cov(\mathbf{r}, \mathbf{f})$ for this combination of test assets and risk factors has rank 1 only. His evidence implies that there are SDFs that price the test assets in an economic unattractive manner. Our results confirm that those SDFs seem to be the only admissible ones that a linearized extension of Epstein and Zin (1989) CCAPM combining durable and nondurable consumption can generate.

4.3 Robustness exercises

One potential concern with our methods is that the number of moments involved may be too large relative to the sample size. For that reason, we assess the reliability of the empirical results in Table 1 in two different ways: using a sample with a longer time span, and also with a smaller cross-section of test assets. In the first case, we use the same data as Burnside (2016), whose sample period is 1949-2012, while in the second one we make use of the 6 size- and value-sorted Fama-French portfolios over the same time span.

Table 2 shows that our main findings are robust to these changes.[14] Specifically, we continue to find that the admissible SDFs lie on a two-dimensional subspace, and that the entire set of admissible SDFs has zero means.

[14] We follow Burnside (2016) in using real excess returns, while Yogo (2006) used nominal excess returns. Given that the effect of inflation is second order for excess returns, the choice between nominal and real returns is inconsequential for our results.

(Table 2: Empirical evaluation of the model 1949-2012)

5 Monte Carlo Evidence

In this section, we assess the finite sample size and power properties of the testing procedures discussed above by means of several extensive Monte Carlo exercises. The exact design of our experiments is described in Appendix B. Given that the number of mean, variance and correlation parameters for returns and factors is large, we have simplified the data generating process (DGP) as much as possible without losing generality, so that in the end we only had to select a handful of parameters whose interpretation is very simple.

We use $n = 6$ and $T = 200$. This number of test assets coincides with the dimension of one of the Fama-French cross-sections in the previous section, while the sample size represents fifty years of quarterly data. Further, we also run simulations with $T = 600$, which corresponds to fifty years of monthly data. In all instances, we simulate $10,000$ samples for each design.

5.1 Numerical details

The main practical difficulty is that we have to rely on numerical optimization methods to maximize the non-linear CU-GMM criterion function even though the moment conditions are linear in the parameters. For that reason, we compute the criterion by means of the auxiliary OLS regressions described in appendix B of Peñaranda and Sentana (2012). We achieve further gains in numerical reliability by using the consistent estimators in Propositions 1 and 2 as starting values.

Given that single-step methods are invariant to different parametrizations of the SDF, we use the uncentered version in (6) because it is the most parsimonious in terms of parameters. Nevertheless, one could exploit the numerical equivalence of the different approaches mentioned in footnote 2, as well as the different normalizations, to check that a global minimum has been reached.

In view of the exactly identified nature of the moment conditions (5), further speed gains can be achieved by minimizing the original moment conditions (6) with respect to $\boldsymbol{\delta}_1, \ldots, \boldsymbol{\delta}_d$ only. Once this is done, the joint criterion function can be minimized with respect to $c_1, \ldots, c_d$ only, keeping $\boldsymbol{\delta}_1, \ldots, \boldsymbol{\delta}_d$ fixed at their CUEs and using the sample means of the estimated SDF basis as consistent starting values.

5.2 Two-dimensional set of admissible SDFs

Table 3 displays the rejection rates of the J and DM tests when there is a two-dimension set of admissible SDFs. In our two factor setting, this means that any of the factors can price the cross-section of returns on its own. Our standard asymptotic theory implies that we expect rejection rates close to size for the J test for $d = 2$. In contrast, the usual J test for $d = 1$ should under-reject because of its generic lack of identification. The only exception arises when $c = 0$, in which case there will be a unique linear combination of the factors that yields an admissible SDF with zero mean, even though the two SDFs that we use in this design have nonzero means. Thus, the J test for $d = 1$ that imposes a zero SDF mean should yield rejection rates close to size too.

Panel A reports the rejection rates when most SDFs in the admissible set have nonzero means, while Panel B shows the corresponding figures when the asset pricing model is completely overspecified. To achieve this, we use two factors that are uncorrelated with the cross-section of returns as the DGP of Panel B. In each panel, we report the rates for 6 tests: the J tests for $d = 2$ and $d = 1$, their variants restricted to have zero SDF means, and the corresponding DM tests.

(Table 3: Rejection rates for a two-dimensional set of admissible SDFs)

The first result we can see in Panel A of Table 3 is that the J test for $d = 2$ performs well, showing only a slight overrejection under the null, and considerable power against $c = 0$. As expected, the J test for $d = 1$ massively under-rejects when we do not impose the restriction that $c = 0$, while it has rejection rates close to size if we do.

On the other hand, Panel B of Table 3 confirms that the J test for $d = 2$ underrejects, the restricted J test performs well, with only a slight overrejection, and the corresponding DM test overrejects. This last overrejection indicates that the fact that this DM test does not reject in our empirical application is not due to lack of power. In that regard, Table B1 in the Appendix shows that this DM test no longer shows any noticeable size distortions for $T = 600$.

5.3 One-dimensional set of admissible SDFs

Table 4 displays the rejection rates of the J and DM tests when the true model contains only one (up to scale) admissible SDF. In that case, we expect that the J test for $d = 1$ yields rejection rates close to size, while the J test for $d = 2$ should now show substantial power.

Once again, Panel A contains the rejection rates when the SDF has a nonzero mean, while Panel B reports the corresponding figures when the model is overspecified. To achieve this, we impose that one the factors is uncorrelated with the cross-section of returns as the DGP of Panel B.

(Table 4: Rejection rates for a one-dimensional set of admissible SDFs)

As expected, Panel A of Table 4 confirms that the J test for $d = 1$ performs well while the J test for $d = 2$ has indeed power. Therefore, our finding an overspecified model in the empirical application cannot be due to lack of power of this second test.

In Panel B of Table 4, the J test for $d = 2$ shows considerable power. Further, the J test for $d = 1$ underrejects, the restricted J test performs well, and the corresponding DM test overrejects. As in the previous section, Table B2 in the Appendix shows that this DM test no longer shows any noticeable size distortions for $T = 600$.

Finally, we also simulated a design where the admissible set of SDFs is empty, as described in Appendix B. In this case, all the tests that we study should reject their respective null hypotheses. Our results, which are available upon request, confirm the power of our proposed procedures such a design.

6 Conclusions

We study the estimation of risk prices and the testing of the cross-sectional restrictions imposed by overspecified linear factor pricing models in which there is at least one non-zero SDF which is uncorrelated with the excess returns of the test assets chosen by the researcher. We provide several examples of this situation, which is necessary but not sufficient for the model parameters to be underidentified. In addition, we also emphasize the distinction between a model with uncorrelated pricing factors, which is necessarily overidentified, from a model with uncorrelated SDFs.

Unlike previous studies, which focus on the non-standard asymptotic properties of the usual estimators and tests, our methods directly estimate the linear subspaces of prices of risk and associated SDFs compatible with the pricing restrictions of the model, which we can easily express in terms of linear moment conditions and efficiently estimate using standard GMM methods. In this regard, a non-trivial advantage of our unusual procedures is that they have standard asymptotic distributions.

We use single-step GMM procedures, and in particular continuously updated GMM, to obtain identical test statistics and risk price estimates for SDF and regression methods, with uncentered or centred moments and symmetric or asymmetric normalizations. Another non-trivial advantage of these methods is that they yield exactly the same conclusions for excess returns and gross returns.

We also propose simple tests to detect economically unattractive but empirically relevant situations in which the expected values of all SDFs in the identified set are 0, which is equivalent to their being uncorrelated to the test assets. In our opinion, researchers could convince readers that their results are meaningful by systematically reporting that they reject the restrictions implicit in these completely overspecified models.

In our empirical application, we investigate the potential overspecification of the three-factor extension of the Epstein and Zin (1989) version of the consumption CAPM model in Yogo (2006), which combines two macroeconomic factors: non-durable and durable consumption, and a stock market factor.

We evaluate the linearized version of this model with the original data, which corresponds to excess returns on the Fama-French cross-section of 25 size and book-to-market sorted portfolios from 1951 to 2001. Our results indicate that the admissible SDFs lie on a two-dimensional subspace. In addition, we cannot reject the null hypothesis that model is completely overspecified. Importantly, our results hold both when we update the sample and when we consider the Fama-French cross-section of 6 portfolios. In addition, our simulations show that these empirical findings are not due to lack of power. On the contrary, if anything, our proposed tests tend to overreject for the sample sizes of our datasets.

Our econometric methodology is *positive* in nature, in the sense that our main objective has been to complement the diagnostics that researchers typically report in support of their preferred linear factor pricing specification so as to increase the empirical credibility of their results. Nevertheless, it might be interesting to combine our procedures with *normative* econometric methods that some researchers use to come up with an acceptable specification. Three recent proposals are Harvey, Liu and Zhu (2016), Bryzgalova (2016) and Kozak, Nagel and Santosh (2017). The application of our proposed diagnostics to models that have been selected after an implicit or explicit specification search raises multiple testing issues that we leave for future research.

Another interesting avenue for further research would be to consider bootstrap versions of our tests to improve their finite sample reliability. Finally, we could also apply our methods to other popular empirical asset pricing models, such as the ones in Jagannathan and Wang (1996) or Lettau and Ludvigson (2001). We are currently pursuing some of these extensions.

References

Arellano, M., L.P. Hansen and E. Sentana (2012): "Underidentification?", *Journal of Econometrics* 170, 256–280.

Bryzgalova, S. (2016): "Spurious factors in linear asset pricing models", Stanford GSB working paper.

Burnside, C. (2016): "Identification and inference in linear stochastic discount factor models with excess returns", *Journal of Financial Econometrics* 14, 295–330.

Cragg, J.G. and S.G. Donald (1997): "Inferring the rank of a matrix", *Journal of Econometrics* 76, 223–250.

Den Haan, W.J. and A.T. Levin (1997): "A practitioner's guide to robust covariance estimation", in G.S. Maddala and C.R. Rao (eds.) *Handbook of Statistics* vol. 15, 299–342, Elsevier.

Dufour, J.-M. (1997): "Some impossibility theorems in econometrics with applications to structural and dynamic models", *Econometrica* 65, 1365–1388.

Eichenbaum, M. and L.P. Hansen (1990): "Estimating models with intertemporal substitution using aggregate time series data", *Journal of Business and Economic Statistics* 8, 53–69.

Epstein, L.G. and S.E. Zin (1989): "Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework", *Journal of Political Economy* 57, 937–969.

Fama, E.F. and K.R. French (1993): "Common risk factors in the returns on stock and bonds", *Journal of Financial Economics* 33, 3–56.

Fang, K.-T., S. Kotz and K.-W. Ng (1990): *Symmetric Multivariate and Related Distributions*, Chapman and Hall.

Gospodinov, N., R. Kan, and C. Robotti (2015): "Spurious inference in reduced-rank asset-pricing models", Federal Reserve Bank of Atlanta working paper.

Hansen, L.P. (1982): "Large sample properties of generalized method of moments estimators", *Econometrica* 50, 1029–1054.

Hansen, L.P., J. Heaton and A. Yaron (1996): "Finite sample properties of some alternative GMM estimators", *Journal of Business and Economic Statistics* 14, 262–280.

Hansen, L.P. and R. Jagannathan (1991): "Implications of security market data for models of dynamic economies", *Journal of Political Economy* 99, 225–262.

Hansen, L.P. and R. Jagannathan (1997): "Assessing specification errors in stochastic discount factor models", *Journal of Finance* 52, 557–590.

Harvey, C. R., Y. Liu, and H. Zhu (2016): "... and the cross-section of expected returns", *Review of Financial Studies* 29, 5–68.

Hillier, G.H. (1990): "On the normalization of structural equations: properties of direct estimators", *Econometrica* 58, 1181–1194.

Householder, A.S. (1964): *The Theory of Matrices in Numerical Analysis*, Blaisdell Publishing Co.

Jagannathan, R. and Z. Wang (1996): "The conditional CAPM and the cross-section of expected returns", *Journal of Finance* 51, 3–53.

Kan, R. and C. Zhang (1999): "GMM tests of stochastic discount factor models with useless factors", *Journal of Financial Economics* 54, 103–127.

Kleibergen, F. and R. Paap (2006): "Generalized reduced rank tests using the singular value decomposition", *Journal of Econometrics* 133, 97–126.

Kozak, S., S. Nagel, and S. Santosh (2017): "Shrinking the cross section", University of Michigan working paper.

Lettau, M., and S. Ludvigson (2001): "Resurrecting the (C)CAPM: a cross-sectional test when risk premia are time-varying", *Journal of Political Economy* 109, 1238–1287.

Lewellen J., S. Nagel and J. Shanken (2010): "A skeptical appraisal of asset-pricing tests, *Journal of Financial Economics* 96, 175–194.

Manresa, E. (2009): "Application of rank tests to linear factor asset pricing models", unpublished Master thesis, CEMFI.

Newey, W.K. and D.L. McFadden (1994): "Large sample estimation and hypothesis testing", in R.F Engle and D.L. McFadden (eds.) *Handbook of Econometrics* vol. IV, 2111–2245, Elsevier.

Newey, W.K. and K.D. West (1987a): "A simple positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix", *Econometrica* 55, 703-8.

Newey, W.K. and K.D. West (1987b): "Hypothesis testing with efficient method of moments estimation", *International Economic Review* 28, 777-787.

Peñaranda, F. and E. Sentana (2012): "Spanning tests in portfolio and stochastic discount factor mean-variance frontiers: a unifying approach", *Journal of Econometrics* 170, 303–324.

Peñaranda, F. and E. Sentana (2015): "A unifying approach to the empirical evaluation of asset pricing models", *Review of Economics and Statistics* 97, 412–435.

Sargan, J.D. (1958): "The estimation of economic relationships using instrumental variables", *Econometrica* 26, 393–415.

Yogo, M. (2006): "A consumption-based explanation of expected stock returns", *Journal of Finance* 61, 539–580.

Appendices

A Proofs and special cases

A.1 Proofs

Proposition 1

We develop most of the proof for the case $d = 2$ to simplify the expressions, but explain the extension to $d > 2$ at the end.

a) When $d = 2$, the moment conditions (6) become

$$E(\mathbf{m} \otimes \mathbf{r}) = E\begin{pmatrix} m_1\mathbf{r} \\ m_2\mathbf{r} \end{pmatrix} = E\begin{pmatrix} (1 - \mathbf{f}_1'\boldsymbol{\delta}_1)\,\mathbf{r} \\ (1 - \mathbf{f}_2'\boldsymbol{\delta}_2)\,\mathbf{r} \end{pmatrix} = \mathbf{0}.$$

We know from Hansen (1982) that the optimal moments correspond to the linear combinations

$$\mathbf{D}'\mathbf{S}^{-1}\frac{1}{T}\sum_{t=1}^{T}\begin{pmatrix} m_{1t}\mathbf{r}_t \\ m_{2t}\mathbf{r}_t \end{pmatrix},$$

where $\mathbf{D}$ is the expected Jacobian and $\mathbf{S}$ the corresponding long-run variance

$$\mathbf{S} = avar\left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\begin{pmatrix} m_{1t}\mathbf{r}_t \\ m_{2t}\mathbf{r}_t \end{pmatrix}\right].$$

In this setting, the expected Jacobian trivially is

$$\mathbf{D} = \begin{pmatrix} \mathbf{D}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_2 \end{pmatrix}, \qquad \mathbf{D}_i = -E\left(\mathbf{r}\mathbf{f}_i'\right).$$

Since we assume that the chosen normalization $(\boldsymbol{\delta}_1, \boldsymbol{\delta}_2)$ is identified, $\mathbf{D}$ has full column rank, which in turn implies that both $\mathbf{D}_1$ and $\mathbf{D}_2$ must have full column rank too.

When $(\mathbf{r}_t, \mathbf{f}_t)$ is an i.i.d. elliptical random vector with bounded fourth moments, we can tediously show that the long-run covariance matrix of the influence functions will be

$$\mathbf{S} = \mathcal{A} \otimes E\left(\mathbf{r}\mathbf{r}'\right) - \mathcal{B} \otimes E(\mathbf{r})\,E(\mathbf{r})',$$

$$\mathcal{A} = (1+\kappa)\,V(\mathbf{m}) + E(\mathbf{m})\,E(\mathbf{m})', \qquad \mathcal{B} = \kappa V(\mathbf{m}) + 2(1-\kappa)\,E(\mathbf{m})\,E(\mathbf{m})',$$

where κ is the coefficient of multivariate excess kurtosis (see Fang, Kotz and Ng (1990)).

To relate the optimal moments to the factor mimicking portfolios

$$\mathbf{r}_i^+ = \mathbf{C}_i\mathbf{r}, \qquad \mathbf{C}_i = E\left(\mathbf{r}\mathbf{f}_i'\right)' E^{-1}\left(\mathbf{r}\mathbf{r}'\right),$$

it is convenient to define the matrix

$$\mathbf{C}' = \begin{pmatrix} \mathbf{C}_1' & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_2' \end{pmatrix},$$

on the basis of which we can compute

$$\mathbf{SC}' = \left[\mathcal{A} \otimes E\left(\mathbf{rr}'\right) - \mathcal{B} \otimes E\left(\mathbf{r}\right) E\left(\mathbf{r}\right)'\right] \begin{pmatrix} \mathbf{C}_1' & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_2' \end{pmatrix}$$

$$= \begin{pmatrix} \mathcal{A}_{11} E\left(\mathbf{rf}_1'\right) & \mathcal{A}_{12} E\left(\mathbf{rf}_2'\right) \\ \mathcal{A}_{12} E\left(\mathbf{rf}_1'\right) & \mathcal{A}_{22} E\left(\mathbf{rf}_2'\right) \end{pmatrix} - \begin{pmatrix} \mathcal{B}_{11} E\left(\mathbf{r}\right) E\left(\mathbf{r}\right)' \mathbf{C}_1' & \mathcal{B}_{12} E\left(\mathbf{r}\right) E\left(\mathbf{r}\right)' \mathbf{C}_2' \\ \mathcal{B}_{12} E\left(\mathbf{r}\right) E\left(\mathbf{r}\right)' \mathbf{C}_1' & \mathcal{B}_{22} E\left(\mathbf{r}\right) E\left(\mathbf{r}\right)' \mathbf{C}_2' \end{pmatrix}.$$

Given that the existence of two valid SDFs implies that $E(\mathbf{r}) = E(\mathbf{rf}_1')\boldsymbol{\delta}_1 = E(\mathbf{rf}_2')\boldsymbol{\delta}_2$, we can write these matrices as

$$\mathbf{SC}' = \begin{pmatrix} \mathcal{A}_{11} E\left(\mathbf{rf}_1'\right) & \mathcal{A}_{12} E\left(\mathbf{rf}_2'\right) \\ \mathcal{A}_{12} E\left(\mathbf{rf}_1'\right) & \mathcal{A}_{22} E\left(\mathbf{rf}_2'\right) \end{pmatrix} - \begin{pmatrix} \mathcal{B}_{11} E\left(\mathbf{rf}_1'\right) \boldsymbol{\delta}_1 \boldsymbol{\delta}_1' \mathbf{G}_1 & \mathcal{B}_{12} E\left(\mathbf{rf}_2'\right) \boldsymbol{\delta}_2 \boldsymbol{\delta}_2' \mathbf{G}_2 \\ \mathcal{B}_{12} E\left(\mathbf{rf}_1'\right) \boldsymbol{\delta}_1 \boldsymbol{\delta}_1' \mathbf{G}_1 & \mathcal{B}_{22} E\left(\mathbf{rf}_2'\right) \boldsymbol{\delta}_2 \boldsymbol{\delta}_2' \mathbf{G}_2 \end{pmatrix},$$

$$\mathbf{G}_i = E\left(\mathbf{rf}_i'\right)' E^{-1}\left(\mathbf{rr}'\right) E\left(\mathbf{rf}_i'\right).$$

In addition, let us define the matrices $\mathbf{Q}_i$ such that $E(\mathbf{rf}_1') = E(\mathbf{rf}_2')\mathbf{Q}_1$ and $E(\mathbf{rf}_2') = E(\mathbf{rf}_1')\mathbf{Q}_2$, which are related by $\mathbf{Q}_2 = \mathbf{Q}_1^{-1}$. The existence of these matrices is guaranteed by the lack of full column rank of $E(\mathbf{rf}')$ together with the full column rank of $E(\mathbf{rf}_1')$ and $E(\mathbf{rf}_2')$. Thus, we can write

$$\mathbf{SC}' = \mathbf{D}\mathbb{Q},$$

$$\mathbb{Q} = -\begin{pmatrix} \mathcal{A}_{11}\mathbf{I}_1 - \mathcal{B}_{11}\boldsymbol{\delta}_1\boldsymbol{\delta}_1'\mathbf{G}_1 & \mathbf{Q}_2\left(\mathcal{A}_{12}\mathbf{I}_1 - \mathcal{B}_{12}\boldsymbol{\delta}_2\boldsymbol{\delta}_2'\mathbf{G}_2\right) \\ \mathbf{Q}_1\left(\mathcal{A}_{12}\mathbf{I}_2 - \mathcal{B}_{12}\boldsymbol{\delta}_1\boldsymbol{\delta}_1'\mathbf{G}_1\right) & \mathcal{A}_{22}\mathbf{I}_2 - \mathcal{B}_{22}\boldsymbol{\delta}_2\boldsymbol{\delta}_2'\mathbf{G}_2 \end{pmatrix}.$$

The assumption that $\mathbf{D}'\mathbf{S}^{-1}$ has full row rank guarantees that the same is true for $\mathbf{C}$, so that $\mathbb{Q}$ will be invertible. Therefore, we have found that

$$\mathbf{D}'\mathbf{S}^{-1} = \mathbb{Q}'^{-1}\mathbf{C}.$$

In other words, the rows of $\mathbf{D}'\mathbf{S}^{-1}$ are spanned by the rows of $\mathbf{C}$, which confirms that the factor mimicking portfolios span the optimal instrumental variables.

As a result, the optimal moments can be expressed as

$$\begin{pmatrix} \mathbf{C}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_2 \end{pmatrix} \frac{1}{T}\sum_{t=1}^{T} \begin{pmatrix} m_{1t}\mathbf{r}_t \\ m_{2t}\mathbf{r}_t \end{pmatrix} = \frac{1}{T}\sum_{t=1}^{T} \begin{pmatrix} \mathbf{r}_{1t}^{+} m_{1t} \\ \mathbf{r}_{2t}^{+} m_{2t} \end{pmatrix} = \mathbf{0},$$

which proves that the optimal estimator of each vector of risk prices simply uses the corresponding factor mimicking portfolios. This estimator is infeasible because we do not know $\mathbf{C}_i$, but under standard regularity conditions we can replace $\mathbf{r}_{it}^{+}$ by its sample counterpart in (9) without affecting the asymptotic distribution.

b) When $d = 2$, the joint system of moments (6) and (7)

$$E(\mathbf{h}) = E\begin{pmatrix} \mathbf{m} \otimes \mathbf{r} \\ \mathbf{m} - \mathbf{c} \end{pmatrix},$$

is composed by

$$E(\mathbf{m} \otimes \mathbf{r}) = E\begin{pmatrix} m_1 \mathbf{r} \\ m_2 \mathbf{r} \end{pmatrix} = E\begin{pmatrix} (1 - \mathbf{f}_1' \boldsymbol{\delta}_1)\, \mathbf{r} \\ (1 - \mathbf{f}_2' \boldsymbol{\delta}_2)\, \mathbf{r} \end{pmatrix} = \mathbf{0},$$
$$E(\mathbf{m} - \mathbf{c}) = E\begin{pmatrix} m_1 - c_1 \\ m_2 - c_2 \end{pmatrix} = E\begin{pmatrix} 1 - \mathbf{f}_1' \boldsymbol{\delta}_1 - c_1 \\ 1 - \mathbf{f}_2' \boldsymbol{\delta}_2 - c_1 \end{pmatrix} = \mathbf{0},$$

with the parameters being

$$\boldsymbol{\theta} = \begin{pmatrix} \boldsymbol{\delta} \\ \mathbf{c} \end{pmatrix}, \qquad \boldsymbol{\delta} = \begin{pmatrix} \boldsymbol{\delta}_1 \\ \boldsymbol{\delta}_2 \end{pmatrix}, \qquad \mathbf{c} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}.$$

The optimal moments correspond to the linear combinations

$$\mathcal{D}' \mathcal{S}^{-1} \frac{1}{T} \sum_{t=1}^{T} \mathbf{h}_t,$$

where $\mathcal{D}$ is the expected Jacobian and $\mathcal{S}$ the corresponding long-run variance

$$\mathcal{S} = avar\left[\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \mathbf{h}_t \right].$$

In this setting, the expected Jacobian can be decomposed as

$$\mathcal{D} = \begin{pmatrix} \mathbf{D} & \mathbf{0} \\ \mathbb{D} & -\mathbf{I}_2 \end{pmatrix},$$

where $\mathbb{D}$ contains the Jacobian of $\mathbf{m} - \mathbf{c}$ with respect to $\boldsymbol{\delta}$, and $\mathbf{I}_2$ is the identity matrix of order 2. The long-run variance for i.i.d. returns and factors can be decomposed as

$$\mathcal{S} = \begin{pmatrix} \mathbf{S} & E(\mathbf{m}\mathbf{m}' \otimes \mathbf{r}) \\ E(\mathbf{m}\mathbf{m}' \otimes \mathbf{r}') & Var(\mathbf{m}) \end{pmatrix}.$$

Once again, we can exploit the structure of the optimal moments to show that the optimal estimator of $\boldsymbol{\delta}$ satisfies the moment conditions

$$\mathbf{D}' \mathbf{S}^{-1} \frac{1}{T} \sum_{t=1}^{T} (\mathbf{m}_t \otimes \mathbf{r}_t) = \mathbf{0}.$$

Hence, the optimal estimator of $\mathbf{c}$ will satisfy the moment conditions

$$\frac{1}{T} \sum_{t=1}^{T} (\mathbf{m}_t - \mathbf{c}) - E\left(\mathbf{m}\mathbf{m}' \otimes \mathbf{r}'\right) \mathbf{S}^{-1} \frac{1}{T} \sum_{t=1}^{T} (\mathbf{m}_t \otimes \mathbf{r}_t) = \mathbf{0}.$$

Obviously, as the additional moments $E(\mathbf{m}-\mathbf{c})=\mathbf{0}$ are exactly identified, the moment conditions that define the optimal estimator of $\boldsymbol{\delta}$ coincide with the conditions in point a), and consequently the same estimator is obtained. The optimal estimator of $\mathbf{c}$ is equal to

$$\frac{1}{T}\sum_{t=1}^{T}\mathbf{m}_t - E\left(\mathbf{m}\mathbf{m}'\otimes\mathbf{r}'\right)\mathbf{S}^{-1}\frac{1}{T}\sum_{t=1}^{T}\left(\mathbf{m}_t\otimes\mathbf{r}_t\right),$$

with $\mathbf{m}_t$ evaluated at the optimal estimator of $\boldsymbol{\delta}$.

When $(\mathbf{r}_t,\mathbf{f}_t)$ is an i.i.d. elliptical random vector with bounded fourth moments, we can show that

$$E\left(\mathbf{m}\mathbf{m}'\otimes\mathbf{r}'\right)=\mathcal{C}\otimes E(\mathbf{r})', \qquad \mathcal{C}=Var(\mathbf{m})-E(\mathbf{m})E(\mathbf{m})'.$$

There are two valid SDFs: $E(\mathbf{r})=E\left(\mathbf{r}\mathbf{f}_1'\right)\boldsymbol{\delta}_1=E\left(\mathbf{r}\mathbf{f}_2'\right)\boldsymbol{\delta}_2$. Thus, we can write

$$E\left(\mathbf{m}\mathbf{m}'\otimes\mathbf{r}'\right)=\begin{pmatrix}\mathcal{C}_{11}E(\mathbf{r})' & \mathcal{C}_{12}E(\mathbf{r})'\\ \mathcal{C}_{12}E(\mathbf{r})' & \mathcal{C}_{22}E(\mathbf{r})'\end{pmatrix}=\begin{pmatrix}\mathcal{C}_{11}\boldsymbol{\delta}_1'E\left(\mathbf{r}\mathbf{f}_1'\right)' & \mathcal{C}_{12}\boldsymbol{\delta}_2'E\left(\mathbf{r}\mathbf{f}_2'\right)'\\ \mathcal{C}_{12}\boldsymbol{\delta}_1'E\left(\mathbf{r}\mathbf{f}_1'\right)' & \mathcal{C}_{22}\boldsymbol{\delta}_2'E\left(\mathbf{r}\mathbf{f}_2'\right)'\end{pmatrix}.$$

Let us focus on the optimal estimator of c_1. We can express it as

$$\begin{aligned}&\frac{1}{T}\sum_{t=1}^{T}m_{1t}-\begin{pmatrix}\mathcal{C}_{11}\boldsymbol{\delta}_1' & \mathcal{C}_{12}\boldsymbol{\delta}_2'\end{pmatrix}\begin{pmatrix}E\left(\mathbf{r}\mathbf{f}_1'\right)' & \mathbf{0}\\ \mathbf{0} & E\left(\mathbf{r}\mathbf{f}_2'\right)'\end{pmatrix}\mathbf{S}^{-1}\frac{1}{T}\sum_{t=1}^{T}\left(\mathbf{m}_t\otimes\mathbf{r}_t\right)\\ &=\frac{1}{T}\sum_{t=1}^{T}m_{1t}+\begin{pmatrix}\mathcal{C}_{11}\boldsymbol{\delta}_1' & \mathcal{C}_{12}\boldsymbol{\delta}_2'\end{pmatrix}\mathbf{D}'\mathbf{S}^{-1}\frac{1}{T}\sum_{t=1}^{T}\left(\mathbf{m}_t\otimes\mathbf{r}_t\right),\end{aligned}$$

where the second term must be zero by definition of the optimal estimator of $\boldsymbol{\delta}$. A similar argument can be applied to the optimal estimator of c_2. Hence, we can conclude that

$$\hat{\mathbf{c}}=\frac{1}{T}\sum_{t=1}^{T}\mathbf{m}_t$$

will be the optimal estimator of the SDF means in an elliptical setting. □

Finally, we can easily extend our proof to $d>2$ because the structures of $\mathbf{D}$, $\mathbf{S}$, and $\mathbf{C}$ are entirely analogous. Specifically, $\mathbf{S}$ will continue to be the same function of $\mathcal{A}$ and $\mathcal{B}$ above, although the dimension of these matrices becomes d instead of 2. In turn, $\mathbf{D}$ and $\mathbf{C}$ will remain block-diagonal, but with d blocks instead of 2 along the diagonal. Lastly, $E\left(\mathbf{m}\mathbf{m}'\otimes\mathbf{r}'\right)$ will continue to be the same function of $\mathcal{C}$ above.

Proposition 2

One again, we develop most of the proof for the case $d=2$ to simplify the expressions, but explaining the extension to $d>2$ at the end.

When $d = 2$, the moment conditions (10) become

$$E(\mathbf{m} \otimes \mathbf{x}) = E\begin{pmatrix} m_1\mathbf{x} \\ m_2\mathbf{x} \end{pmatrix} = E\begin{pmatrix} (1 - \mathbf{f}_1'\boldsymbol{\delta}_1)\,\mathbf{x} \\ (1 - \mathbf{f}_2'\boldsymbol{\delta}_2)\,\mathbf{x} \end{pmatrix} = \mathbf{0}.$$

The optimal moments correspond to the linear combinations

$$\mathbf{D}'\mathbf{S}^{-1}\frac{1}{T}\sum_{t=1}^{T}\begin{pmatrix} m_{1t}\mathbf{x}_t \\ m_{2t}\mathbf{x}_t \end{pmatrix},$$

where $\mathbf{D}$ is the expected Jacobian and $\mathbf{S}$ the corresponding long-run variance. In this setting, the expected Jacobian is block-diagonal with blocks $-E\left(\mathbf{x}\mathbf{f}_i'\right)$.

When $(\mathbf{r}_t, \mathbf{f}_t)$ is an i.i.d. elliptical random vector with bounded fourth moments, and $E(\mathbf{m}) = \mathbf{0}$, we can use our previous results in the proof of Proposition 1 to show that the long-run covariance matrix of the influence functions will be

$$\mathbf{S} = \mathcal{A} \otimes E\left(\mathbf{x}\mathbf{x}'\right) - \mathcal{B} \otimes E(\mathbf{x})\,E(\mathbf{x})',$$
$$\mathcal{A} = (1+\kappa)\,E\left(\mathbf{m}\mathbf{m}'\right), \qquad \mathcal{B} = \kappa E\left(\mathbf{m}\mathbf{m}'\right),$$

where κ is the coefficient of multivariate excess kurtosis.

The structure of $\mathbf{D}$ and $\mathbf{S}$ are similar to the proof of Proposition 1. Therefore, we can follow the same argument to conclude that if we define the factor mimicking portfolios on the extended payoff space as

$$\mathbf{x}_i^+ = \mathbf{C}_i\mathbf{x}, \qquad \mathbf{C}_i = E\left(\mathbf{x}\mathbf{f}_i'\right)' E^{-1}\left(\mathbf{x}\mathbf{x}'\right),$$

then the sample version of the optimal moments can be expressed as

$$\begin{pmatrix} \mathbf{C}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_2 \end{pmatrix}\frac{1}{T}\sum_{t=1}^{T}\begin{pmatrix} m_{1t}\mathbf{x}_t \\ m_{2t}\mathbf{x}_t \end{pmatrix} = \frac{1}{T}\sum_{t=1}^{T}\begin{pmatrix} \mathbf{x}_{1t}^+ m_{1t} \\ \mathbf{x}_{2t}^+ m_{2t} \end{pmatrix}.$$

This expression proves that the optimal estimator of each vector of risk prices simply uses the corresponding factor mimicking portfolios. Once again, this estimator is infeasible because we do not know $\mathbf{C}_i$, but under standard regularity conditions we can replace $\mathbf{x}_{it}^+$ by its sample counterpart in (12) without affecting the asymptotic distribution. □

As in the case of Proposition 1, we can easily extend our proof to $d > 2$ because the structures of $\mathbf{D}$, $\mathbf{S}$, and $\mathbf{C}$ is entirely analogous. Specifically, $\mathbf{S}$ will continue to be the same function of $\mathcal{A}$ and $\mathcal{B}$ above, although the dimension of these matrices becomes d instead of 2. In turn, $\mathbf{D}$ and $\mathbf{C}$ will remain block-diagonal, but with d blocks instead of 2 along the diagonal.

Proposition 3

The proof is simpler if we express the pricing conditions (1) in terms of central moments, which is numerically inconsequential for CU-GMM. Specifically, we can add and subtract $\mathbf{b}'\boldsymbol{\mu}$ from $a + \mathbf{b}'\mathbf{f}$ and define $c = a + \mathbf{b}'\boldsymbol{\mu}$ as the expected value of the affine SDF. This allows us to re-write the pricing conditions as

$$E\left\{\begin{array}{c}[c + \mathbf{b}'(\mathbf{f} - \boldsymbol{\mu})]\,\mathbf{r} \\ \mathbf{f} - \boldsymbol{\mu}\end{array}\right\} = \mathbf{0}. \tag{A1}$$

In this way, the unknown parameters become $(c, \mathbf{b}, \boldsymbol{\mu})$ instead of $(a, \mathbf{b})$, as we have added k extra moments to estimate $\boldsymbol{\mu}$. Empirical researchers often use the implicit normalization $c = 1$, but this would be incompatible with the null hypothesis $H_0 : c = 0$ that we want to test. In contrast, the symmetric normalization $\mathbf{b}'\mathbf{b} + c^2 = 1$ is perfectly compatible with this null hypothesis.

To deal with a $d-$dimensional subspace of admissible SDFs, we need to replicate d times the pricing conditions in (A1). Thus, the centred SDF counterpart to (4) will be based on the moment conditions

$$E\begin{pmatrix}\mathbf{f} - \boldsymbol{\mu} \\ \mathbf{r}m_1 \\ \vdots \\ \mathbf{r}m_d\end{pmatrix} = \mathbf{0}, \qquad m_i = c_i + (\mathbf{f}_i - \boldsymbol{\mu}_i)'\,\mathbf{b}_i, \tag{A2}$$

where $(\mathbf{f}_1, \mathbf{f}_2, ..., \mathbf{f}_d)$ are the vectors of factors that enter each one of the SDFs after imposing the necessary restrictions that guarantee the point identification of the basis $(\mathbf{b}_1, \mathbf{b}_2, ..., \mathbf{b}_d)$.

Let us denote by J the CU-GMM value of the overidentifying restrictions test with free $(c_1, c_2, ..., c_d)$ in (A2). Similarly, let us denote by J_0 the CU-GMM value of the corresponding overidentifying restrictions test after imposing $c_1 = ... = c_d = 0$. In this context, it is straightforward to see that the overidentification test based on J_0 is trivially a rank test on $Cov(\mathbf{r}, \mathbf{f})$ because it is testing the existence of d linear combinations of the columns of this covariance matrix with weights $\mathbf{b}_i$ that are equal to zero. By the invariance properties of single-step GMM methods, it is easy to prove that we would obtain the same value for the overidentification test from the moment conditions (4) and (5). $\square$

Finally, note that our DM test of the null hypothesis $c_1 = ... = c_d = 0$ is based on $J_0 - J$.

A.2 Possible cases with one, two and three factors

We describe all the possible cases for models with one, two or three factors under the maintained assumption that $E(\mathbf{r}) \neq \mathbf{0}$. As a result, we only study cases where the rank of $\mathbf{M}$ is one or higher.

One factor

We cannot have an underidentified single-factor model because the valid SDFs are unique up to scale:

- Identification ($d = 1$): The rank of $\mathbf{M}$ is one.
  - $E(\mathbf{r})$ is not in the span of $Cov(\mathbf{r}, f)$: The rank of $Cov(\mathbf{r}, f)$ is zero, and the model is completely overspecified.
  - $E(\mathbf{r})$ is in the span of $Cov(\mathbf{r}, f)$: The rank of $Cov(\mathbf{r}, f)$ is one.
- Lack of a valid SDF: The rank of $\mathbf{M}$ is two, while the rank of $Cov(\mathbf{r}, f)$ is one.

Two factors

The valid SDFs may belong to a two-dimensional subspace, so we may find underidentified two-factor models:

- Underidentification with $d = 2$: The rank of $\mathbf{M}$ is one.
  - $E(\mathbf{r})$ is not in the span of $Cov(\mathbf{r}, \mathbf{f})$: The rank of $Cov(\mathbf{r}, \mathbf{f})$ is zero, and the model is completely overspecified.
  - $E(\mathbf{r})$ is in the span of $Cov(\mathbf{r}, \mathbf{f})$: The rank of $Cov(\mathbf{r}, \mathbf{f})$ one, so the model is only partially overspecified.
- Identification ($d = 1$): The rank of $\mathbf{M}$ is two.
  - $E(\mathbf{r})$ is not in the span of $Cov(\mathbf{r}, \mathbf{f})$: The rank of $Cov(\mathbf{r}, \mathbf{f})$ is one, and the model is completely overspecified.
  - $E(\mathbf{r})$ is in the span of $Cov(\mathbf{r}, \mathbf{f})$: The rank of $Cov(\mathbf{r}, \mathbf{f})$ is two.
- Lack of a valid SDF: The rank of $\mathbf{M}$ is three, while the rank of $Cov(\mathbf{r}, \mathbf{f})$ is two.

Three factors

The valid SDFs may belong to a three-dimensional subspace:

- Underidentification with $d = 3$: The rank of $\mathbf{M}$ is one.
  - $E(\mathbf{r})$ is not in the span of $Cov(\mathbf{r}, \mathbf{f})$: The rank of $Cov(\mathbf{r}, \mathbf{f})$ is zero, and the model is completely overspecified.

- $E(\mathbf{r})$ is in the span of $Cov(\mathbf{r}, \mathbf{f})$: The rank of $Cov(\mathbf{r}, \mathbf{f})$ is one, so the model is only partially overspecified.

- Underidentification with $d = 2$: The rank of $\mathbf{M}$ is two.
  - $E(\mathbf{r})$ is not in the span of $Cov(\mathbf{r}, \mathbf{f})$: The rank of $Cov(\mathbf{r}, \mathbf{f})$ is one, and the model is completely overspecified.
  - $E(\mathbf{r})$ is in the span of $Cov(\mathbf{r}, \mathbf{f})$: The rank of $Cov(\mathbf{r}, \mathbf{f})$ is two, so the model is only partially overspecified.

- Identification ($d = 1$): The rank of $\mathbf{M}$ is three.
  - $E(\mathbf{r})$ is not in the span of $Cov(\mathbf{r}, \mathbf{f})$: The rank of $Cov(\mathbf{r}, \mathbf{f})$ is two, and the model is completely overspecified.
  - $E(\mathbf{r})$ is in the span of $Cov(\mathbf{r}, \mathbf{f})$: The rank of $Cov(\mathbf{r}, \mathbf{f})$ is three.

- Lack of a valid SDF: The rank of $\mathbf{M}$ is four, while the rank of $Cov(\mathbf{r}, \mathbf{f})$ is three.

B Monte Carlo design

B.1 Data generating process

In this appendix, we extend the design of the single factor Monte Carlo experiment in Peñaranda and Sentana (2015) to a two-factor model. An unrestricted (i.i.d.) Gaussian data generating process (DGP) for $(\mathbf{f}, \mathbf{r})$ is

$$\mathbf{f} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}),$$
$$\mathbf{r} = \boldsymbol{\mu}_r + \mathbf{B}_r(\mathbf{f} - \boldsymbol{\mu}) + \mathbf{u}_r, \qquad \mathbf{u}_r \sim N(\mathbf{0}, \boldsymbol{\Omega}_{rr}),$$

where the $n \times 2$ matrix $\mathbf{B}_r$ is defined by the two beta vectors

$$\mathbf{B}_r = \begin{pmatrix} \boldsymbol{\beta}_1 & \boldsymbol{\beta}_2 \end{pmatrix}.$$

Without loss of generality, we construct the two factors so that their covariance matrix is the identity matrix. In addition, given that we use the simulated data to test that an affine function of $\mathbf{f}$ is orthogonal to $\mathbf{r}$, the only thing that matters is the linear span of $\mathbf{r}$. As a result, we can substantially reduce the number of parameters characterizing the conditional DGP for $\mathbf{r}$ by means of the following steps:

1. a Cholesky transformation of $\mathbf{r}$ which effectively sets the residual variance $\boldsymbol{\Omega}_{rr}$ equal to the identity matrix,

2. a Householder transformation that makes the second to the last entries of the vector of risk premia $\boldsymbol{\mu}_r$ equal to zero (see Householder (1964)),

3. another Householder transformation that makes the third to the last entries of $\boldsymbol{\beta}_1$ equal to zero,

4. a final third Householder transformation that makes the fourth to the last entries of $\boldsymbol{\beta}_2$ equal to zero.

As a result, our simplified DGP for excess returns will be

$$\mathbf{r} = \mu_r \mathbf{e}_1 + (\beta_{11}\mathbf{e}_1 + \beta_{21}\mathbf{e}_2)(f_1 - \mu_1) + (\beta_{12}\mathbf{e}_1 + \beta_{22}\mathbf{e}_2 + \beta_{32}\mathbf{e}_3)(f_2 - \mu_2) + \mathbf{u}_r,$$
$$\mathbf{u}_r \sim N(\mathbf{0}, \mathbf{I}_n),$$

where $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ are the first, second, and third columns of the identity matrix, and

$$\mathbf{f} \sim N(\boldsymbol{\mu}, \mathbf{I}_2).$$

B.2 Model restrictions

We set the values of the three parameters of $\boldsymbol{\mu}$ to 1. In turn, we calibrate the six parameters that define $\mathbf{r}$ as follows. First, we define a Hansen-Jagannathan (HJ) distance for this three-factor model as the minimum with respect to $\boldsymbol{\phi}$ of the quadratic form

$$\boldsymbol{\phi}'\mathbb{M}'Var^{-1}(\mathbf{r})\mathbb{M}\boldsymbol{\phi},$$

where

$$\mathbb{M}\boldsymbol{\phi} = \begin{bmatrix} E(\mathbf{r}) & Cov(\mathbf{r}, \mathbf{f}) \end{bmatrix} \begin{pmatrix} c \\ \mathbf{b} \end{pmatrix}.$$

Note that $\mathbb{M}\boldsymbol{\phi} = \mathbf{M}\boldsymbol{\theta}$ and $rank(\mathbb{M}) = rank(\mathbf{M})$.

The 3×3 weighting matrix

$$\mathbb{W} = \mathbb{M}'Var^{-1}(\mathbf{r})\mathbb{M}$$

$$= \begin{pmatrix} E(\mathbf{r})'Var^{-1}(\mathbf{r})E(\mathbf{r}) & E(\mathbf{r})'Var^{-1}(\mathbf{r})Cov(\mathbf{r}, \mathbf{f}) \\ \cdot & Cov(\mathbf{r}, \mathbf{f})'Var^{-1}(\mathbf{r})Cov(\mathbf{r}, \mathbf{f}) \end{pmatrix} = \begin{pmatrix} \sigma_{00} & \sigma_{01} & \sigma_{02} \\ \cdot & \sigma_{11} & \sigma_{12} \\ \cdot & \cdot & \sigma_{22} \end{pmatrix}$$

can be interpreted as the variance matrix of three noteworthy portfolios. The first one yields the maximum Sharpe ratio

$$r_0 = \mathbf{r}'Var^{-1}(\mathbf{r})E(\mathbf{r}),$$

while the other two are the centred factor mimicking portfolios

$$r_i = \mathbf{r}' Var^{-1}(\mathbf{r})\, Cov(\mathbf{r}, f_i), \qquad i = 1, 2.$$

Note that if we minimize the above quadratic form subject to the symmetric normalization $\boldsymbol{\phi}'\boldsymbol{\phi} = 1$, then this HJ distance will be equal to the minimum eigenvalue of the covariance matrix $\mathbb{W}$.

The first entry of $\mathbb{W}$ is the variance of r_0 or, equivalently, the squared maximum Sharpe ratio. The other two diagonal entries are the variances of (r_1, r_2) or, equivalently, the R^2 of their respective regressions. Finally, the three different off-diagonal elements correspond to the covariances between these three portfolios, which we can pin down by their correlations. In this way, we have six parameters that are easy to interpret and calibrate, from which we can obtain the six parameters that our DGP requires for $\mathbf{r}$.

We start from the free design and progressively add more and more constraints. In addition, we can interpret the constraints that the different models impose as forcing certain linear combinations of (r_0, r_1, r_2) with coefficients (c, b_1, b_2) to have zero variance. We define 3 designs (with some variants) indexed by the dimension of the subspace of prices of risk d.

- Design $d = 0$: The matrix $\mathbb{W}$ has full rank. We need to give values to the six parameters with the interpretations mentioned before. We calibrate their values to the data as explained below. The rest of designs require constraints on the matrix $\mathbb{W}$, which we impose by means of small changes in that matrix.

- Design $d = 1$: The matrix $\mathbb{W}$ has one rank failure defined by a one-dimensional subspace of vectors (c, b_1, b_2). At least one of the factors must enter the SDF to avoid risk neutrality, so we can assume that $b_2 \neq 0$. Thus, we can choose a linear combination $(c^*, b_1^*, -1)$ with zero variance. Equivalently, we can express r_2 as

$$r_2 - \mu_2 = c^*(r_0 - \mu_0) + b_1^*(r_1 - \mu_1),$$

and change the last column of matrix $\mathbb{W}$ to

$$\sigma_{02} = c^*\sigma_{00} + b_1^*\sigma_{01},$$

$$\sigma_{12} = c^*\sigma_{01} + b_1^*\sigma_{11},$$

$$\sigma_{22} = c^{*2}\sigma_{00} + b_1^{*2}\sigma_{11} + 2c^*b_1^*\sigma_{01}.$$

We keep the three parameters that define the covariance matrix of (r_0, r_1) equal to the values they take in design $d = 0$. This design will have two variants: one with nonzero c

in the linear combination (c, b_1, b_2), and a second one with $c^* = 0$. In the former variant, we choose c^* and b_1^* to keep the same σ_{02} and σ_{22} as in the design $d = 0$. In the second variant, we chose $c^* = b_1^* = 0$, which is equivalent to an uncorrelated factor, so that $\sigma_{02} = \sigma_{12} = \sigma_{22} = 0$.

- Design $d = 2$: The matrix $\mathbb{W}$ has two rank failures defined by a two-dimensional subspace of vectors (c, b_1, b_2). We maintain the linear combination $(c^*, b_1^*, -1)$ with zero variance from design $d = 1$, and add a second linear combination $(c^{**}, -1, 0)$ with zero variance. Equivalently, we can express r_1 as

$$r_1 - \mu_1 = c^{**} (r_0 - \mu_0),$$

and modify the matrix $\mathbb{W}$ accordingly

$$\sigma_{01} = c^{**} \sigma_{00},$$

$$\sigma_{11} = c^{**2} \sigma_{00},$$

with $(\sigma_{02}, \sigma_{12}, \sigma_{22})$ satisfying the same equations as in design $d = 1$. We keep σ_{00} equal to the value in design $d = 0$. This design will again have two variants: one with nonzero c in the linear combinations (c, b_1, b_2), and a second one with $c^* = c^{**} = 0$. In the former variant, we choose c^{**} to keep the same σ_{11} as in the design $d = 0$. In the second variant, we have two uncorrelated factors, and hence all entries of $\mathbb{W}$ except σ_{00} are equal to 0.

B.3 Numerical values

- Design $d = 0$: We calibrate the matrix $\mathbb{W}$ from the data. After orthogonalizing the factors (with the market being the first factor), we set the correlations of (r_0, r_1, r_2) and the variances of (r_1, r_2) to values similar to those in the data. In contrast, we lower the variance of r_0, which can be interpreted as the squared maximum Sharpe ratio, so as to ensure a realistic risk-return trade off. Thus, the matrix that we use is

$$\mathbb{W} = \begin{pmatrix} 0.090 & 0.198 & 0.063 \\ \cdot & 0.980 & 0 \\ \cdot & \cdot & 0.100 \end{pmatrix},$$

which is associated to the following DGP

$$\mathbf{r} = 1.433\mathbf{e}_1 + (6.907\mathbf{e}_1 + 1.140\mathbf{e}_2)(f_1 - \mu_1) + (0.049\mathbf{e}_1 - 0.297\mathbf{e}_2 + 0.143\mathbf{e}_3)(f_2 - \mu_2) + \mathbf{u}_r.$$

- Design $d = 1$: We study two variants. In the first one, we impose the existence of an SDF with coefficients
$$\begin{pmatrix} c^* \\ b_1^* \\ b_2^* \end{pmatrix} = \begin{pmatrix} 1.406 \\ -0.319 \\ -1 \end{pmatrix}$$
by means of the matrix
$$\mathbb{W} = \begin{pmatrix} 0.090 & 0.198 & 0.063 \\ \cdot & 0.980 & -0.035 \\ \cdot & \cdot & 0.100 \end{pmatrix},$$
which is associated to the following DGP
$$\mathbf{r} = 1.464\mathbf{e}_1 + (7.158\mathbf{e}_1 + 1.167\mathbf{e}_2)(f_1 - \mu_1) + (-0.229\mathbf{e}_1 - 0.373\mathbf{e}_2)(f_2 - \mu_2) + \mathbf{u}_r.$$
In the second one, we impose the existence of an SDF with coefficients
$$\begin{pmatrix} c^* \\ b_1^* \\ b_2^* \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}$$
by means of the matrix
$$\mathbb{W} = \begin{pmatrix} 0.090 & 0.198 & 0 \\ \cdot & 0.980 & 0 \\ \cdot & \cdot & 0 \end{pmatrix},$$
which is associated to the following DGP
$$\mathbf{r} = 1.432\mathbf{e}_1 + (6.914\mathbf{e}_1 + 1.093\mathbf{e}_2)(f_1 - \mu_1) + \mathbf{u}_r.$$

- Design $d = 2$: We study two variants. In the first one, we impose the existence of two SDFs with coefficients
$$\begin{pmatrix} c^* & c^{**} \\ b_1^* & b_1^{**} \\ b_2^* & b_2^{**} \end{pmatrix} = \begin{pmatrix} 1.406 & 3.300 \\ -0.319 & -1 \\ -1 & 0 \end{pmatrix}$$
by means of the matrix
$$\mathbb{W} = \begin{pmatrix} 0.090 & 0.297 & 0.032 \\ \cdot & 0.980 & 0.104 \\ \cdot & \cdot & 0.011 \end{pmatrix},$$

which is associated to the following DGP

$$\mathbf{r} = 3.182\mathbf{e}_1 + 10.500\mathbf{e}_1 (f_1 - \mu_1) + 1.118\mathbf{e}_1 (f_2 - \mu_2) + \mathbf{u}_r.$$

In the second one, we impose the existence of two SDFs with coefficients

$$\begin{pmatrix} c^* & c^{**} \\ b_1^* & b_1^{**} \\ b_2^* & b_2^{**} \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & -1 \\ -1 & 0 \end{pmatrix}$$

by means of the matrix

$$\mathbb{W} = \begin{pmatrix} 0.090 & 0 & 0 \\ \cdot & 0 & 0 \\ \cdot & \cdot & 0 \end{pmatrix},$$

which is associated to the following DGP

$$\mathbf{r} = 0.300\mathbf{e}_1 + \mathbf{u}_r.$$

Table 1: Empirical evaluation of the model 1951-2001

	One-dimensional Set		Two-dimensional Set			Three-dimensional Set		
Market	0.200	(0.805)	-3.888	0.514	(0.002)	4.793	0	0
Nondur.	24.765	(0.458)	222.902	0	(0.000)	0	115.687	0
Durables	92.229	(0.035)	0	99.333	(0.000)	0	0	121.320
Mean	0.014	(0.790)	-0.099	0.034	(0.494)	0.852	0.421	-0.029
Criterion	20.743	(0.537)		56.687	(0.134)		215.144	(0.000)
Criterion $c = 0$	20.814	(0.592)		58.098	(0.151)			

Note: This table displays estimates of the SDF parameters, as well as the J tests (with free and constrained SDF means) with p-values in parenthesis (). The first, second and third blocks of columns report the results for sets of SDFs of dimension 1, 2 and 3, respectively. We display the estimates for a particular normalization, but our results are numerically invariant to the chosen normalization because we use CU-GMM. The J tests are complemented with significance tests of some SDF parameters. In particular, the p-value of the distance metric test of the null hypothesis of zero parameters is reported in parenthesis to the right of the estimates. Distance metric tests are not reported when the p-value of the J test is lower than 0.01. The payoffs of the test assets correspond to 25 nominal excess returns of size and book-to-market sorted portfolios on a quarterly basis.

Table 2: Empirical evaluation of the model 1949-2012

	One-dimensional Set		Two-dimensional Set			Three-dimensional Set		
Panel A. 25 size and book-to-market sorted portfolios								
Market	0.766	(0.673)	-1.878	0.882	(0.000)	12.882	0	0
Nondur.	-6.452	(0.834)	192.583	0	(0.000)	0	169.191	0
Durables	106.144	(0.024)	0	97.810	(0.000)	0	0	110.143
Mean	0.003	(0.972)	0.052	0.008	(0.757)	0.411	0.065	-0.075
Criterion	18.278	(0.689)		54.818	(0.175)		165.053	(0.000)
Criterion $c = 0$	18.279	(0.742)		55.375	(0.216)			
Panel B. 6 size and book-to-market sorted portfolios								
Market	0.504	(0.542)	-1.873	0.597	(0.053)	4.148	0	0
Nondur.	12.831	(0.837)	208.870	0	(0.000)	0	152.818	0
Durables	92.619	(0.158)	0	100.563	(0.000)	0	0	121.767
Mean	0.021	(0.836)	0.076	0.007	(0.810)	0.875	0.326	-0.173
Criterion	0.526	(0.913)		3.787	(0.876)		47.695	(0.000)
Criterion $c = 0$	0.568	(0.967)		4.207	(0.938)			

Note: This table displays estimates of the SDF parameters, as well as the J tests (with free and constrained SDF means) with p-values in parenthesis (). The first, second and third blocks of columns report the results for sets of SDFs of dimension 1, 2 and 3, respectively. We display the estimates for a particular normalization, but our results are numerically invariant to the chosen normalization because we use CU GMM. The J tests are complemented with significance tests of some SDF parameters. In particular, the p-value of the distance metric test of the null hypothesis of zero parameters is reported in parenthesis to the right of the estimates. Distance metric tests are not reported when the p-value of the J test is lower than 0.01. The payoffs of the test assets correspond to 25 (Panel A) and 6 (Panel B) real excess returns of size and book-to-market sorted portfolios at the quarterly frequency.

Table 3: Rejection rates for a two-dimensional set of admissible SDFs ($T = 200$)

	Nominal size		
	10	5	1
Panel A. Some SDFs have nonzero mean			
J d=2	13.65	7.03	1.61
J d=2, c=0	99.62	99.62	99.62
DM c=0	99.62	99.62	99.62
J d=1	1.03	0.37	0.01
J d=1 c=0	11.78	3.07	1.20
DM c=0	39.97	28.01	10.89
Panel B. All SDFs have zero mean			
J d=2	8.97	4.49	0.71
J d=2, c=0	14.26	7.73	1.72
DM c=0	21.46	13.69	4.34
J d=1	0.75	0.15	0.00
J d=1, c=0	0.89	0.31	0.00
DM c=0	8.03	3.37	0.30

Note: This table displays the rejection rates of CU J tests, their variants restricted to zero SDF means, and the corresponding DM tests, as described in Section 3. The rates are shown in percentage for the asymptotic critical values at 10, 5, and 1%. 10000 samples of 6 excess returns are simulated under the two variants of a two-dimensional set of admissible SDFs in Appendix B, where the parameter values of the DGP are explained. Panel A reports the results for the first variant, when most of these SDFs have nonzero means, and Panel B reports the results for the second variant, when the asset pricing model is completely overspecified.

Table 4: Rejection rates for a one-dimensional set of admissible SDFs ($T = 200$)

	Nominal size		
	10	5	1
Panel A. Some SDFs have nonzero mean			
J d=2	99.17	98.03	92.79
J d=2, c=0	99.99	99.99	99.99
DM c=0	99.97	99.97	99.97
J d=1	10.06	4.99	0.98
J d=1 c=0	97.92	95.39	84.25
DM c=0	99.24	98.71	95.57
Panel B. All SDFs have zero mean			
J d=2	68.55	56.95	33.30
J d=2, c=0	99.85	99.85	99.85
DM c=0	99.84	99.84	99.84
J d=1	6.29	2.66	0.33
J d=1, c=0	11.65	5.96	1.20
DM c=0	20.51	12.92	3.88

Note: This table displays the rejection rates of CU J tests, their variants restricted to zero SDF means, and the corresponding DM tests, as described in Section 3. The rates are shown in percentage for the asymptotic critical values at 10, 5, and 1%. 10000 samples of 6 excess returns are simulated under the two variants of a one-dimensional set of admissible SDFs in Appendix B, where the parameter values of the DGP are explained. Panel A reports the results for the first variant, when the SDF has a nonzero mean, and Panel B reports the results for the second variant, when the asset pricing model is overspecified.

Table B1: Rejection rates for a two-dimensional set of admissible SDFs ($T = 600$)

	Nominal size		
	10	5	1
Panel A. Some SDFs have nonzero mean			
J d=2	10.80	5.55	1.13
J d=2, c=0	99.55	99.55	99.55
DM c=0	99.55	99.55	99.55
J d=1	0.78	0.16	0.00
J d=1 c=0	10.36	5.20	1.23
DM c=0	37.94	26.15	9.82
Panel B. All SDFs have zero mean			
J d=2	10.15	5.03	0.99
J d=2, c=0	11.58	5.98	1.28
DM c=0	13.63	7.58	1.63
J d=1	0.77	0.16	0.00
J d=1, c=0	0.82	0.13	0.00
DM c=0	5.94	2.09	0.19

Note: This table displays the rejection rates of CU J tests, their variants restricted to zero SDF means, and the corresponding DM tests, as described in Section 3. The rates are shown in percentage for the asymptotic critical values at 10, 5, and 1%. 10000 samples of 6 excess returns are simulated under the two variants of a two-dimensional set of admissible SDFs in Appendix B, where the parameter values of the DGP are explained. Panel A reports the results for the first variant, when most of these SDFs have nonzero means, and Panel B reports the results for the second variant, when the asset pricing model is completely overspecified.

Table B2: Rejection rates for a one-dimensional set of admissible SDFs ($T = 600$)

	Nominal size		
	10	5	1
Panel A. Some SDFs have nonzero mean			
J d=2	100	100	100
J d=2, c=0	100	100	100
DM c=0	100	100	100
J d=1	9.94	5.00	1.06
J d=1 c=0	100	100	100
DM c=0	100	100	100
Panel B. All SDFs have zero mean			
J d=2	99.06	97.97	92.69
J d=2, c=0	100.00	100.00	100.00
DM c=0	100.00	100.00	100.00
J d=1	8.87	4.26	0.82
J d=1, c=0	10.50	5.25	1.11
DM c=0	12.79	6.63	1.60

Note: This table displays the rejection rates of CU J tests, their variants restricted to zero SDF means, and the corresponding DM tests, as described in Section 3. The rates are shown in percentage for the asymptotic critical values at 10, 5, and 1%. 10000 samples of 6 excess returns are simulated under the two variants of a one-dimensional set of admissible SDFs in Appendix B, where the parameter values of the DGP are explained. Panel A reports the results for the first variant, when the SDF has a nonzero mean, and Panel B reports the results for the second variant, when the asset pricing model is overspecified.

Figure 1: One asset

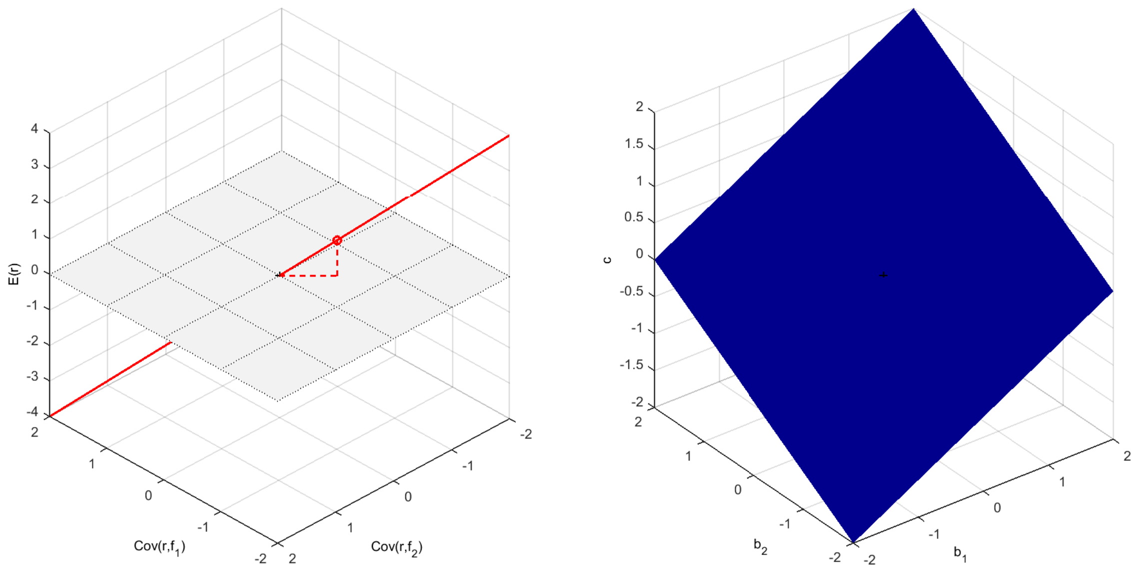


Figure 2: Two assets

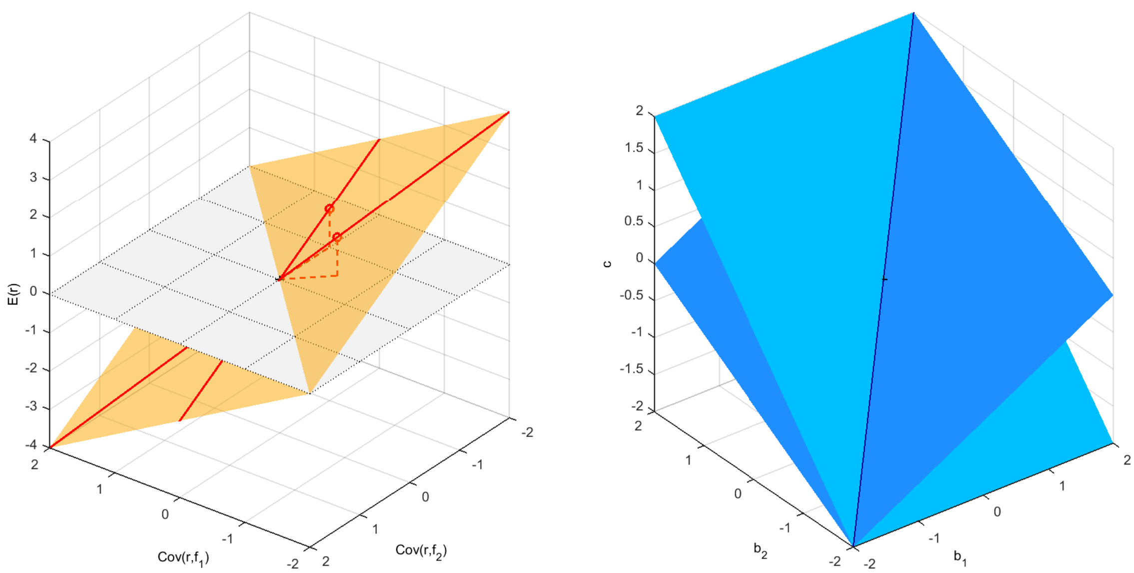


Figure 3: Three segmented asset markets

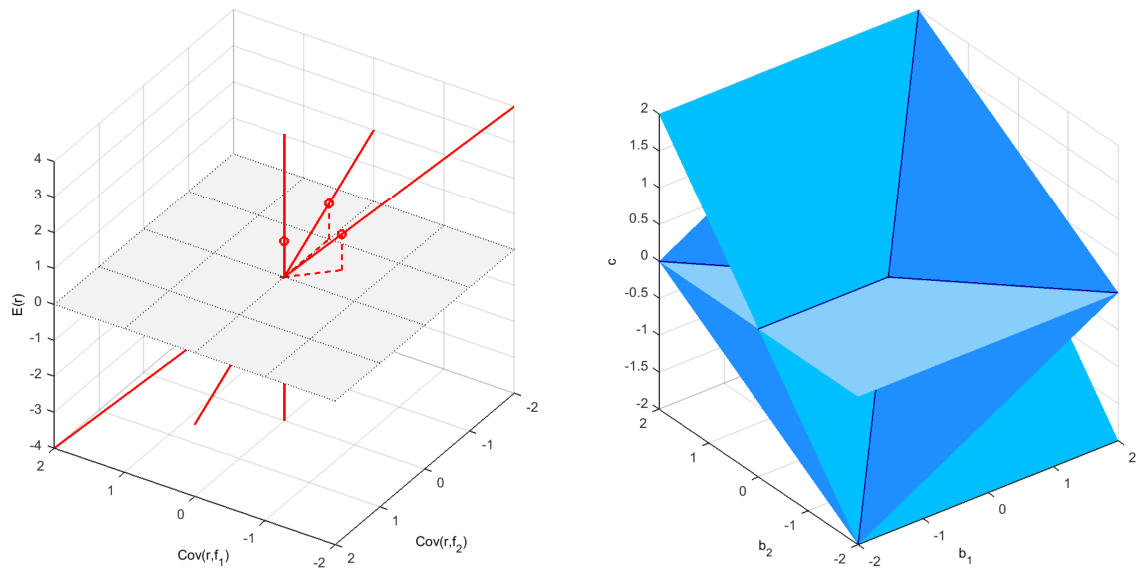


Figure 4: Three integrated asset markets

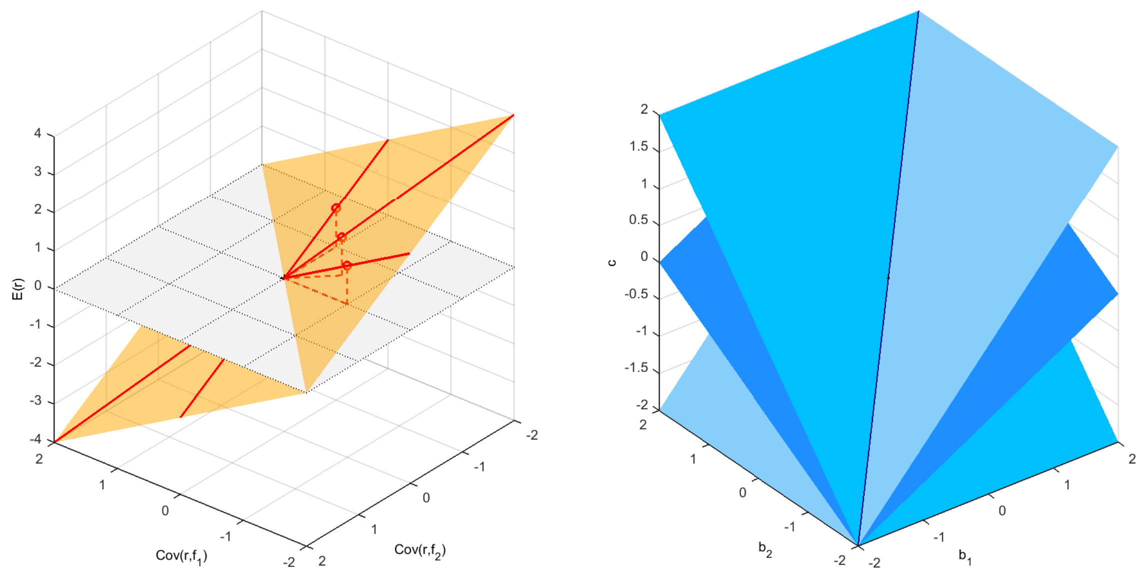


Figure 5: Normalizations

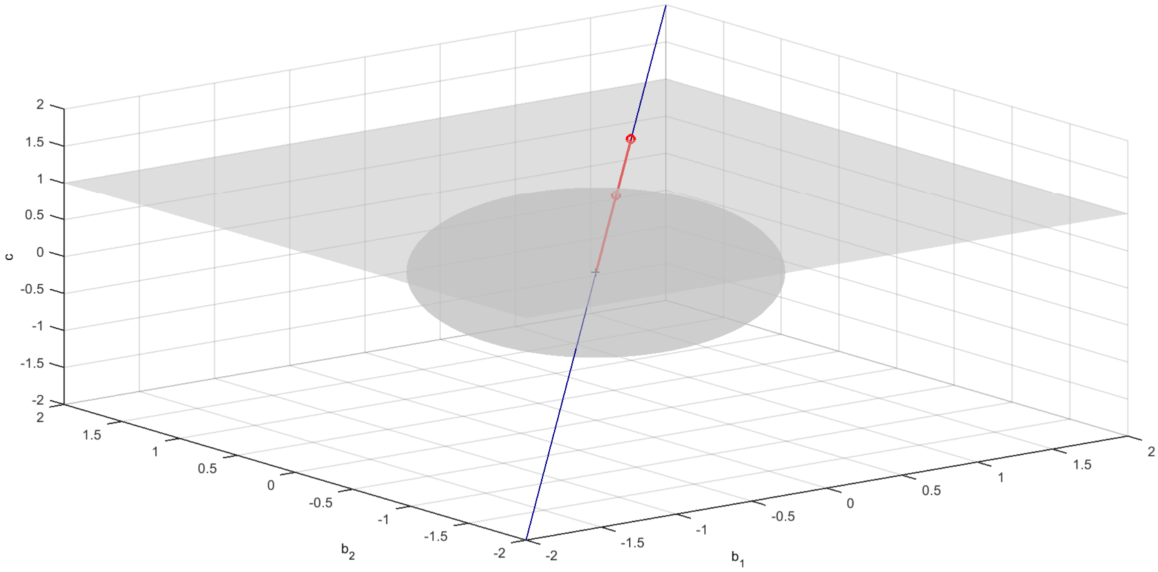
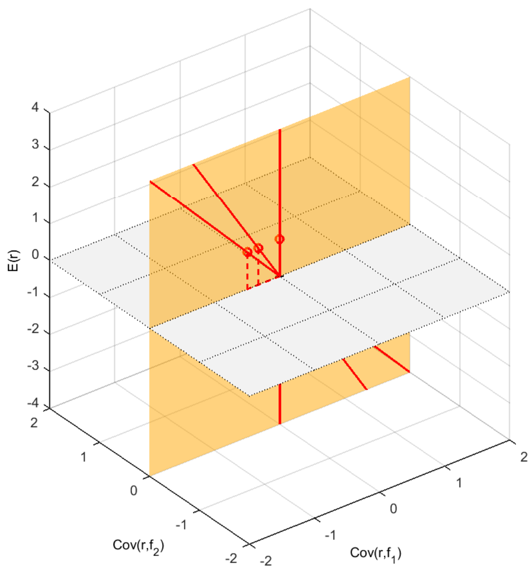


Figure 6: Valid but unattractive model with a useless factor



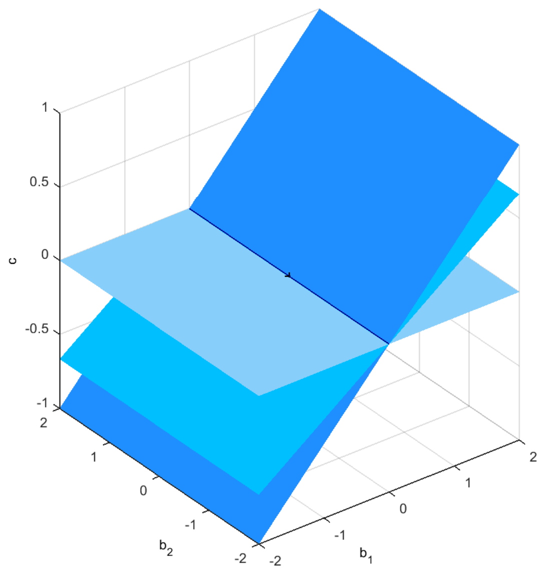


Figure 7: Valid and attractive model with a useless factor

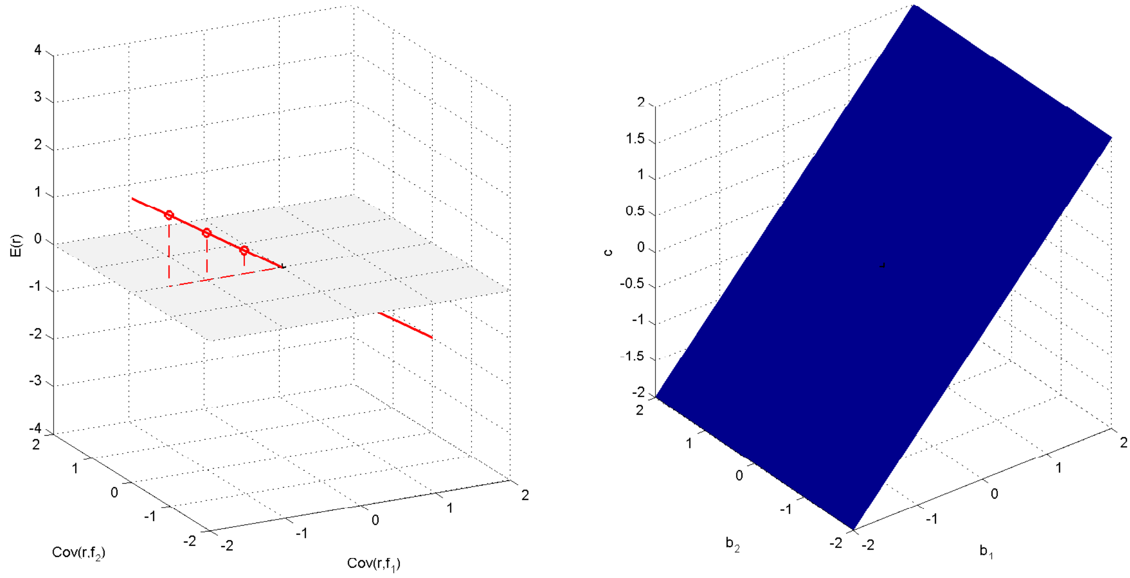


Figure 8: Two single factor models

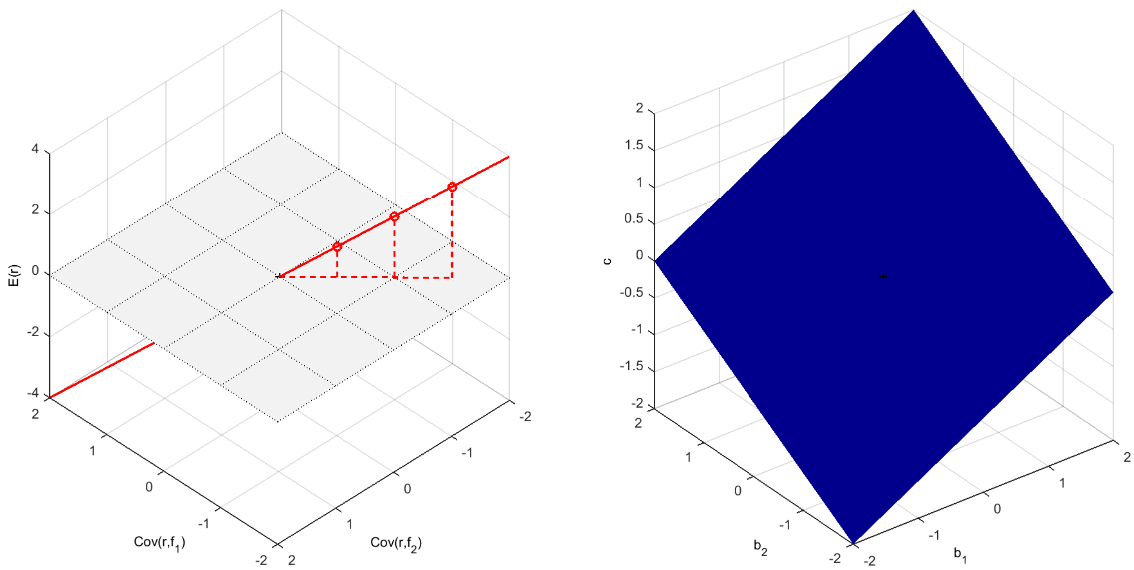


Figure 9: Two useless factors

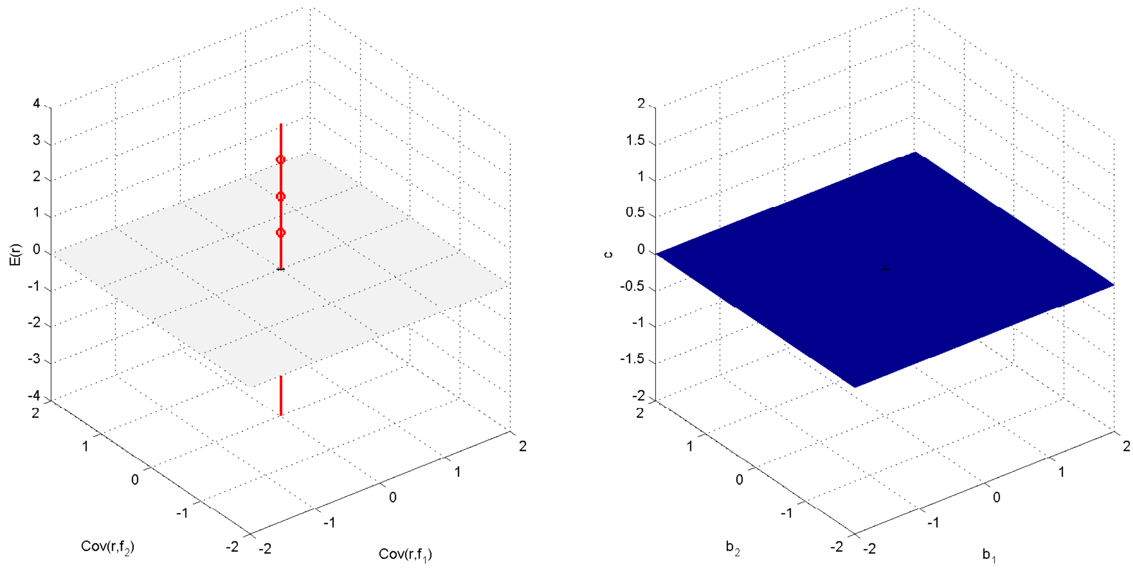


Figure 10: An unpriced second factor

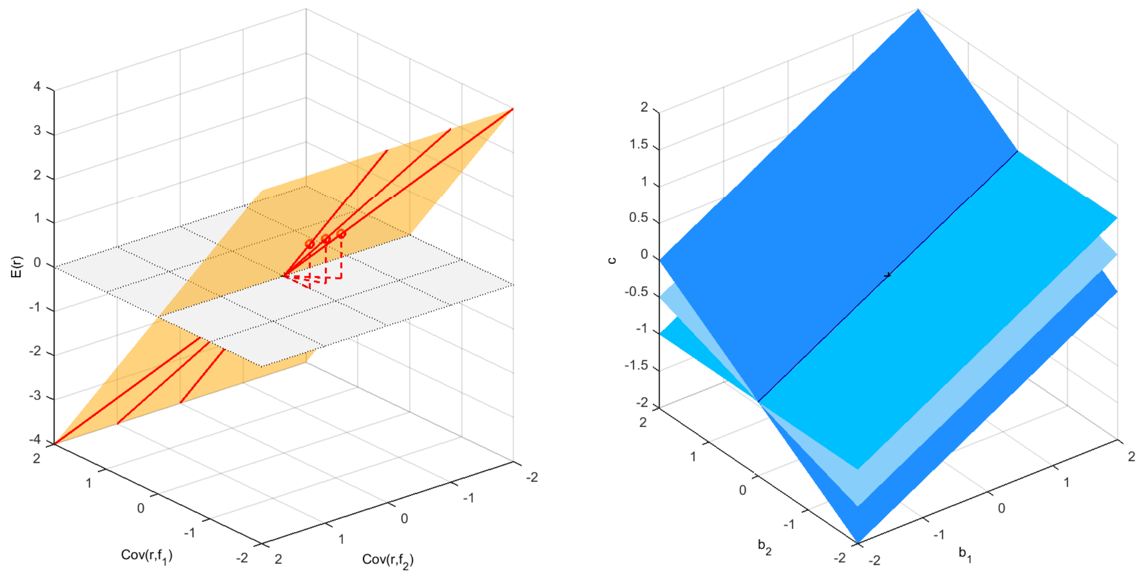


Figure 11: Risk premia from 2S-GMM

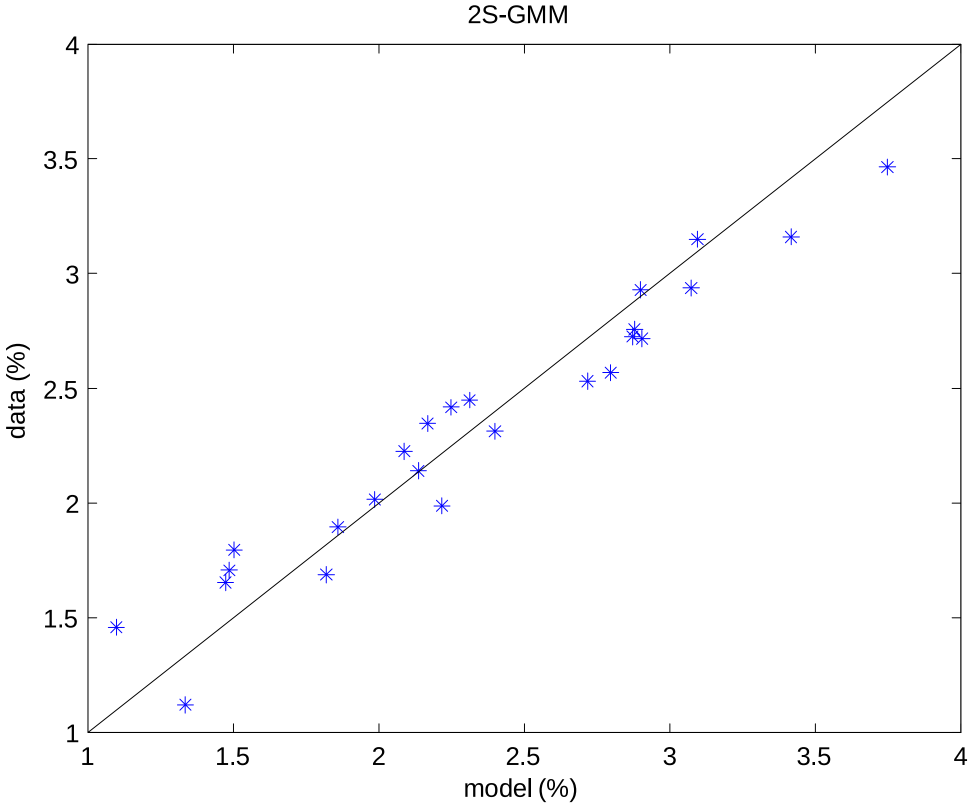


Figure 12: Risk premia from IT-GMM

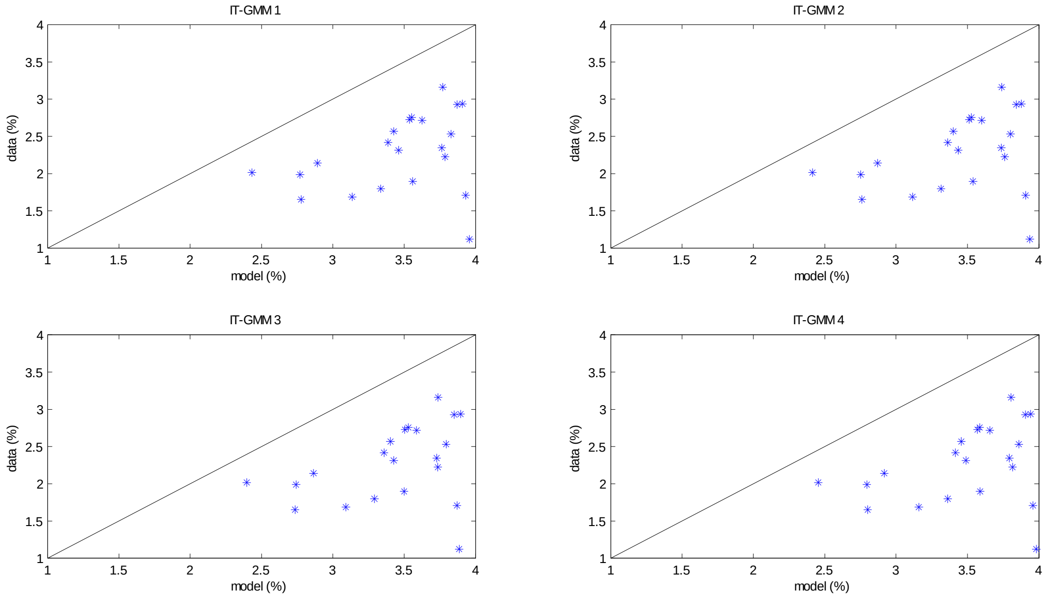


Figure 13: Risk premia from IT-GMM, free coefficients