

NBER WORKING PAPER SERIES

EMPIRICAL ASSET PRICING AND STATISTICAL POWER IN THE PRESENCE
OF WEAK RISK FACTORS

A. Craig Burnside

Working Paper 13357
<http://www.nber.org/papers/w13357>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 2007

I am grateful to the National Science Foundation for financial support (SES-0516697). I thank John Cochrane, Martin Eichenbaum, Bjorn Eraker, Ron Gallant, Lars Hansen, Ravi Jagannathan, Sergio Rebelo, Barbara Rossi, George Tauchen, Mark Watson and Motohiro Yogo, for useful conversations, and seminar participants at Duke, Wharton, and the Board of Governors of the Federal Reserve System for comments. The usual disclaimer applies. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

© 2007 by A. Craig Burnside. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Empirical Asset Pricing and Statistical Power in the Presence of Weak Risk Factors

A. Craig Burnside

NBER Working Paper No. 13357

August 2007

JEL No. C33,F31,G12

ABSTRACT

The risk factors in many consumption-based asset pricing models display statistically weak correlation with the returns being priced. Some GMM-based procedures used to test these models have very low power to reject proposed stochastic discount factors (SDFs) when they are misspecified and the covariance matrix of the asset returns with the risk factors has less than full column rank. Consequently, these estimators provide potentially misleading positive assessments of the SDFs. Working with SDFs specified in terms of demeaned risk factors improves the performance of GMM but the power to reject misspecified SDFs may remain low. Two summary tests for failure of the rank condition have reasonable power, and lead to no Type I errors in Monte Carlo experiments.

A. Craig Burnside

Department of Economics

Duke University

213 Social Sciences Building

Durham, NC 27708-0097

and NBER

burnside@econ.duke.edu

Classical tests of stochastic discount factor (SDF) models are conducted under the maintained hypothesis that the models are true. Confidence regions and standard errors for parameter estimates are derived under two important assumptions: the proposed SDF nests the true specification, and the covariance matrix of the asset returns and risk factors has full column rank. Tests of pricing errors and over-identifying restrictions proceed in the same way. In this paper, I explore the properties of point estimates and the power of inference when these assumptions fail. Using asymptotic theory, as well as small sample simulation-based evidence, I show that some, but not all, standard tests based on the Generalized Method of Moments (GMM) have very low power to reject the proposed SDF when it is false and the rank condition fails to hold. An example of the failure of the rank condition is that the covariance between the vector of asset returns and one of the risk factors is zero. This means that my results are especially relevant for tests of consumption-based asset pricing models, because many consumption-related risk factors proposed in the literature display statistically insignificant covariance with the returns being priced, implying statistical failure of the rank condition.

I consider the conventional, though not universal, case where the model SDF is estimated using data on excess returns. In this case, the mean of the SDF is unidentified but a subset of the model parameters can still be identified by adopting a normalization. Two candidate normalizations described by Cochrane (2005)—one using raw risk factors, the other using demeaned risk factors—are theoretically equivalent when the model is true. In contrast, I show that empirical results obtained using GMM estimators are dramatically different for the two normalizations when consumption-based models are estimated. In particular, the normalization that expresses the SDF in terms of raw risk factors tends to cast the models in a favorable light. Measures of fit are close to 1, and the estimated parameters of the SDF are statistically significant. For the normalization that uses demeaned risk factors, the models appear to fit the data very poorly, and estimated parameters have a much lower degree of statistical significance.

My experiments with Monte Carlo simulation suggest that this lack of robustness does not stem from the small sample properties of the estimators under the null. Rather, I favor an alternative explanation, that the rank condition on the covariance between the returns and the risk factors does not hold and that the proposed consumption-based SDFs are false. In a single factor model, the rank condition fails when the proposed risk factor is uncorrelated

with the returns. This necessarily implies that the model is false or that the excess returns being studied are all mean zero. I discard the second possibility because the returns I study in this paper do not appear to be jointly mean zero. In multifactor models failure of the rank condition can point to a third possibility. The SDF may be true, yet the returns being studied may simply be uninformative about the relevance of one or more, or some linear combination, of the factors. I discard this possibility as well, because it implies that in large samples measures of fit would converge to 1 for any normalization of the model, yet, in practice, these measures differ dramatically across normalizations.

How do failure of the rank condition and falsehood of the proposed SDF explain the lack of robustness across normalizations? Using asymptotic theory I demonstrate that the normalization that uses raw risk factors has the following properties: (i) the estimated parameters of the SDF converge in probability to a non-zero limit such that the estimated SDF is uncorrelated with the returns, (ii) an R^2 measure of model fit converges in probability to 1, (iii) the t -statistic associated with any parameter of the SDF that converges to a non-zero limit (of which there is at least one) will diverge in probability to $\pm\infty$ leading to the conclusion that the associated risk factor helps to price the assets, and (iv) under some additional regularity assumptions the test of the over-identifying restrictions rejects the model with the same probability as the size of the test, that is, as if the model were true. My simulation-based evidence indicates that these properties are borne out in finite samples.

The normalization that uses demeaned risk factors does not share these properties. When the rank condition fails the parameters of this normalization of the SDF are asymptotically unidentified. Under additional assumptions about the data generating processes, I show that the parameter estimates and test statistics have asymptotic distributions, but these results do not lead to general statements about the statistical significance of the parameter estimates and the likelihood of rejecting the model in large samples. Simulation-based evidence suggests that in finite samples a researcher is more likely to conclude that the model has poor fit, reject it based on the test of the pricing errors, and conclude that the spurious factor does not price the assets, when the SDF is expressed in terms of demeaned factors. Nonetheless, even with the demeaned normalization the power to reject the model can be low.

Failure of the rank condition is a possibility of genuine relevance because several of the

risk factors studied in the consumption-based asset pricing literature are only weakly correlated with the returns. Favorable inference regarding the relevance of these risk factors should, therefore, be treated cautiously. Although my results are specifically relevant for empirical work that uses GMM and one of the two normalizations, they have broader relevance, because the first stage of GMM with the demeaned normalization is equivalent to the two-pass regression method pioneered by Black, Jensen and Scholes (1972) and Fama and MacBeth (1973), and widely used in the consumption-based asset pricing literature. In a separate paper, Burnside (2007a), I extend my discussion to further variants of the demeaned specification used by Parker and Julliard (2005) and Yogo (2006).

There is an extensive literature relating to this paper which examines the properties of asset pricing tests in the presence of *spurious* risk factors, that is risk factors that are uncorrelated with the returns. Kan and Zhang (1999a) examine the behavior of GMM estimators in the presence of spurious risk factors. Their results relate to mine for the demeaned specification, although they study risk factors that are mean zero by construction. Kan and Zhang (1999b) study similar issues in the context of the two-pass approach to model evaluation. A more closely related paper to this one is Kan and Robotti (2006). They examine the behavior of the Hansen and Jagannathan (1997) distance measure under the two model normalizations discussed here. They show that results can depend dramatically on the normalization chosen. A contribution of this paper is that I demonstrate the relevance of the theoretical results here and in the prior literature to consumption-based asset pricing.

Going beyond simple failure of the rank condition, or spuriousness of a risk factor, I consider another possibility: that a risk factor has the same covariance with every asset return being explained. I refer to this as *panel spuriousness* and should be considered when factor betas are statistically significant, but have very little spread, as is the case for the quarterly Fama-French portfolios and US consumption growth.¹ A panel spurious factor is not irrelevant, but cannot explain the cross-sectional distribution of the expected returns being studied. I show that in sufficiently large samples both normalizations will lead to rejection of the over-identifying restrictions. However, in finite samples, if the covariance

¹In related work, Lewellen and Nagel (2006) and Lewellen, Nagel and Shanken (2006) criticize empirical estimates of the conditional CAPM (and CCAPM) arguing that the spread in the betas in these models is too small for inference to be reliable. Daniel and Titman (2005) discuss the properties of the Fama-MacBeth procedure when the test assets lie in a low-dimensional subspace of the full payoff space. Their analysis undoubtedly has bearing on GMM-based procedures as well, and one of the sets of data I examine in this paper (the Fama-French 25 returns, described below) is central to their discussion.

between the returns and the factor is small, the performance of the raw-factor normalization mimics the case where the factor is spurious.

My findings suggest that a useful diagnostic, prior to estimation, is to simply estimate factor betas (the slope coefficients in a time series regression of the returns on the risk factors) and perform diagnostic tests on their significance.² A simple test of the rank condition suggested by Cragg and Donald (1997) and Wright (2003) is a useful summary statistic, has reasonable power in finite samples, and leads to no Type I errors in my Monte Carlo experiments.

I also propose a simple post-estimation diagnostic. When a risk factor or linear combination of risk factors is spurious, the probability limit of the mean of the SDF is zero for the normalization that uses raw factors. A test for whether the mean is zero is straightforward, and performs quite well in Monte Carlo simulations. It has reasonable power against false models, especially at the first stage of GMM, and leads to no Type I errors in my experiments.

The paper is organized as follows. Section 1 lays out a standard linear factor model. Section 2 discusses the identification problem and the two normalizations of the SDF. Section 3 discusses the approach to be used in estimating the two model normalizations. Section 4 discusses empirical findings for two sets of data from the finance literature: the Fama and French (1993) returns on 25 portfolios sorted by market capitalization and book-to-market value, and the Lustig and Verdelhan (2007) data on 8 portfolios of returns to long positions in the foreign exchange market, sorted by short-term interest rates. It lays the foundation for the rest of the paper by showing that inference regarding consumption-based models is sharply dependent on the normalization. Section 5 discusses the asymptotic properties of the estimates and diagnostic tests under the maintained assumption that the model is valid, and under the alternative assumption that it is misspecified and a risk factor is spurious. Section 6 performs a series of small-sample Monte Carlo simulation exercises that demonstrate the consequences of failure of the rank condition in samples similar in size to those being studied in the literature. Section 7 concludes.

²Kan and Zhang (1999b) make the same suggestion and point out that while Chen, Roll, and Ross (1986) and Ferson and Harvey (1993) performed such tests, relatively few researchers do so.

1 Model Setup and Moment Conditions

I consider the estimation of a class of linear factor pricing models where the SDF takes the form

$$m_t = a - f_t' b. \quad (1)$$

Here f_t is a $k \times 1$ vector of risk factors, a is a scalar constant and b is $k \times 1$ vector of parameters. If m_t is the true SDF, then standard arguments imply that the price of any asset at time $t - 1$ whose payoff at time t is x_t is $p_{t-1} = E_{t-1}(m_t x_t)$. Therefore the expected price of the asset is $E(p_{t-1}) = E(m_t x_t)$. It follows that the return to an asset, $R_t = x_t/p_{t-1}$, satisfies $E_{t-1}(m_t R_t) = 1$ and $E(m_t R_t) = 1$. Consequently, the difference between the returns on two assets $R_{1t} - R_{2t}$ satisfies $E[m_t(R_{1t} - R_{2t})] = 0$.

Now consider a particular $n \times 1$ vector, R_t^e , whose i th element, R_{it}^e , is the excess return at time t to asset i defined as the difference between the return on asset i and the risk free rate. If m_t given by (1) is the true SDF, then

$$E(R_t^e m_t) = 0. \quad (2)$$

The restriction (2) implies that

$$E(R_t^e) = -\frac{\text{cov}(R_t^e, m_t)}{E(m_t)}. \quad (3)$$

Equation (3) states that the expected return, $E(R_{it}^e)$, is higher for those assets for which $\text{cov}(R_{it}^e, m_t)$ is lower. Variation in $E(R_{it}^e)$ across i implies variation in $\text{cov}(R_{it}^e, m_t)$ across i .

A beta representation of the model, derived from (3), is

$$E(R_t^e) = -\underbrace{\frac{\text{cov}(R_t^e, m_t)}{\text{var}(m_t)}}_{\beta_m} \underbrace{\frac{\text{var}(m_t)}{E(m_t)}}_{\lambda_m}. \quad (4)$$

The term β_m is an $n \times 1$ vector of *SDF betas*, while λ_m is the *market price of risk*, as it measures the expected excess return for an asset whose SDF beta is 1. One such asset is an SDF mimicking portfolio whose excess return is $\lambda_m - [m_t - E(m_t)]$. An alternative beta representation, obtained by substituting (1) into the numerator of (3), is

$$E(R_t^e) = \frac{\text{cov}(R_t^e, f_t' b)}{E(m_t)} = \underbrace{\text{cov}(R_t^e, f_t') \text{var}(f_t)^{-1}}_{\beta_f} \underbrace{\frac{\text{var}(f_t) b}{E(m_t)}}_{\lambda_f}. \quad (5)$$

The term β_f is an $n \times k$ matrix of *factor betas*, while λ_f is a $k \times 1$ vector of *factor risk premia*. Row i in the matrix β_f is the vector of slope coefficients in a time series regression of R_{it}^e on f_t .

2 Normalizations of the SDF

The moment restriction (2) does not separately identify the parameters a and b . This is because the GMM errors, $R_t^e m_t$, for the parameter pair (a, b) are proportional to the GMM errors for the parameter pair (ka, kb) , for any constant k . I consider two approaches to achieving identification, both of which involve picking a particular normalization of the SDF.

To achieve identification, consider rewriting (1) as

$$m_t = am_t^* = a(1 - f_t' b^*) \quad (6)$$

where $b^* = b/a$. Given this way of writing the SDF, the model is estimated by exploiting the moment restrictions:

$$E(R_t^e m_t^*) = 0. \quad (7)$$

I refer to this normalization as the *a-normalization*.

Alternatively (1) can be rewritten as

$$m_t = \xi m_t^\diamond = \xi[1 - (f_t - \mu)' b^\diamond], \quad (8)$$

where μ is the unconditional mean of f_t , ξ is a scalar, $\xi = a - \mu'b$ and $b^\diamond = b/(a - \mu'b)$. Given this way of writing the SDF, the model is estimated by exploiting the moment restrictions:

$$E(f_t) = \mu \quad E(R_t^e m_t^\diamond) = 0. \quad (9)$$

I refer to this normalization as the *ξ -normalization*.

If the original model, (1), is true and (2) holds, the moment restrictions stated in (7) and (9) are also valid. Although b^* and b^\diamond are different parameters (except when $\mu = 0$) the normalizations are equivalent in the sense that the GMM errors $1 - f_t' b^*$ and $1 - (f_t - \mu)' b^\diamond$ are proportional to one another when evaluated at the true parameter values. Also, b^\diamond can be obtained from b^* and μ : $b^\diamond = b^*/(1 - \mu'b^*)$.

3 Estimation and Inference using GMM

3.1 The *a*-Normalization

Using the n moment restrictions given by (7), b^* is estimated using GMM. Define $u_t^*(b^*) = R_t^e m_t^* = R_t^e(1 - f_t' b^*)$ and let $g_T^*(b^*) = \frac{1}{T} \sum_{t=1}^T u_t^* = \bar{R}^e - D_T b^*$ be an $n \times 1$ vector of pricing

errors, where $\bar{R}^e = \frac{1}{T} \sum_{t=1}^T R_t^e$, $D_T = \frac{1}{T} \sum_{t=1}^T R_t^e f_t'$ and T is the sample size. I consider GMM estimators that set $a_T^* g_T^* = 0$, where a_T^* is a $k \times n$ matrix and takes the form $a_T^* = D_T' W_T^*$, where W_T^* is an $n \times n$ positive definite weighting matrix. It follows that the GMM estimator of b^* is

$$\hat{b}^* = (D_T' W_T^* D_T)^{-1} D_T' W_T^* \bar{R}^e. \quad (10)$$

I consider two-stage GMM estimators. In the first stage $W_T^* = I_n$. In the second stage, $W_T^* = (S_T^*)^{-1}$ where $S_T^* = \frac{1}{T} \sum_{t=1}^T \hat{u}_t^* \hat{u}_t^{*'}$, $\hat{u}_t^* = R_t^e (1 - f_t' \hat{b}_1^*)$ and \hat{b}_1^* represents the first stage estimator of b^* .

Let $\delta_T^* = -D_T$. A test of the pricing errors is based on the statistic

$$J^* = T g_T^*(\hat{b}^*) (\hat{V}_g^*)^+ g_T^*(\hat{b}^*), \quad (11)$$

where

$$\hat{V}_g^* = A_T^* S_T^* A_T^{*'} \text{ with } A_T^* = I_n - \delta_T^* (a_T^* \delta_T^*)^{-1} a_T^*, \quad (12)$$

and X^+ indicates the generalized inverse of the matrix X .

Equations (6) and (7) imply that

$$E(R_t^e) = E(R_t^e f_t) b^*. \quad (13)$$

In a finite sample, corresponding to the left-hand side of (13) is the vector of *actual expected returns*, \bar{R}^e . Corresponding to the right-hand side of (13) is a vector of *predicted expected returns* given by $D_T \hat{b}^*$. The model's fit is evaluated using the cross-sectional R^2 :

$$R_*^2 = 1 - \frac{(\bar{R}^e - D_T \hat{b}^*)' (\bar{R}^e - D_T \hat{b}^*)}{(\bar{R}^e - \ddot{R}^e)' (\bar{R}^e - \ddot{R}^e)}, \quad (14)$$

where $\ddot{R}^e = \frac{1}{n} \sum_{i=1}^n \bar{R}_i^e$ is the cross-sectional average of the mean returns in the data.

Equation (13) can be rewritten as

$$E(R_t^e) = \underbrace{E(R_t^e f_t) \text{var}(f_t)^{-1} E(m_t^*)}_{\beta_f} \underbrace{\frac{\text{var}(f_t) b^*}{E(m_t^*)}}_{\lambda_f} \quad (15)$$

Since $m_t^* = m_t/a$ and $b^* = b/a$, it follows that λ_f and β_f are the same objects defined in (5). The mean and covariance matrix of f_t are estimated by GMM using the moment restrictions

$$E(f_t - \mu) = 0 \quad (16)$$

$$E[(f_t - \mu)(f_t - \mu)' - \Sigma_f] = 0. \quad (17)$$

An estimate of λ_f is given by

$$\hat{\lambda}_f^* = \frac{\hat{\Sigma}_f \hat{b}^*}{1 - \hat{\mu}' \hat{b}^*}, \quad (18)$$

where $\hat{\mu} = \bar{f} \equiv \frac{1}{T} \sum_{t=1}^T f_t$ and $\hat{\Sigma}_f = S_f \equiv \frac{1}{T} \sum_{t=1}^T (f_t - \bar{f})(f_t - \bar{f})'$ are, respectively, the sample mean and covariance matrix of f_t . Standard errors for $\hat{\lambda}_f^*$ are obtained by the delta method using the joint distribution of \hat{b}^* , $\hat{\mu}$ and $\hat{\Sigma}_f$. The details are discussed in Appendix A.³

3.2 The ξ -Normalization

Using the $n+k$ moment restrictions given by (9), b^\diamond and μ are estimated using GMM. Define $u_{1t}^\diamond(b^\diamond, \mu) = R_t^e m_t^\diamond = R_t^e [1 - (f_t - \mu)' b^\diamond]$ and let $g_{1T}^\diamond(b^\diamond, \mu) = \frac{1}{T} \sum_{t=1}^T u_{1t}^\diamond = \bar{R}^e - (D_T - \bar{R}^e \mu') b^\diamond$. Define $u_{2t}^\diamond(\mu) = f_t - \mu$ and let $g_{2T}^\diamond(\mu) = \frac{1}{T} \sum_{t=1}^T u_{2t}^\diamond = \bar{f} - \mu$. Define $u_t^\diamond = (u_{1t}^{\diamond'} \ u_{2t}^{\diamond'})'$ and $g_T^\diamond = (g_{1T}^{\diamond'} \ g_{2T}^{\diamond'})'$. I consider GMM estimators that set $a_T^\diamond g_T^\diamond = 0$, where a_T^\diamond is a $2k \times (n+k)$ matrix and takes the form

$$a_T^\diamond = \begin{pmatrix} d_T' W_T^\diamond & 0 \\ 0 & I_k \end{pmatrix}, \quad (19)$$

where $d_T = D_T - \bar{R}^e \bar{f}'$ and W_T^\diamond is an $n \times n$ positive definite weighting matrix. It follows that the GMM estimators of b^\diamond and μ are

$$\hat{b}^\diamond = (d_T' W_T^\diamond d_T)^{-1} d_T' W_T^\diamond \bar{R}^e \quad (20)$$

$$\hat{\mu} = \bar{f}. \quad (21)$$

I consider two-stage GMM estimators. In the first stage $W_T^\diamond = I_n$. In the second stage, $W_T^\diamond = (P_T S_T^\diamond P_T')^{-1}$ where $P_T = (I_n \ \bar{R}^e (\hat{b}^\diamond)')$ and S_T^\diamond is a consistent estimator of $S_0^\diamond = \sum_{j=-\infty}^{+\infty} E(u_t^\diamond u_{t-j}^{\diamond'})$.⁴ Because u_{2t}^\diamond may be serially correlated I use a VARHAC estimator, described in more detail in Appendix A, to compute S_T^\diamond .

Let

$$\delta_T^\diamond = \begin{pmatrix} -d_T & \bar{R}^e \hat{b}^{\diamond'} \\ 0 & -I_k \end{pmatrix}. \quad (22)$$

A test of the pricing errors is based on

$$J^\diamond = T g_T(\hat{b}^\diamond, \hat{\mu}) (\hat{V}_g^\diamond)^+ g_T(\hat{b}^\diamond, \hat{\mu}), \quad (23)$$

³The appendices are available at <http://www.duke.edu/~acb8/research.htm>

⁴Cochrane (2005) suggests using the matrix $(I_n \ 0_{n \times k})$ in place of P_T in the expression for W_T^\diamond . This is less efficient in terms of the covariance matrix of \hat{b}^\diamond , but is asymptotically equivalent in terms of the test of the overidentifying restrictions.

where

$$\hat{V}_g^\diamond = A_T^\diamond S_T^\diamond A_T^{\diamond'} \text{ with } A_T^\diamond = I_{n+k} - \delta_T^\diamond (a_T^\diamond \delta_T^\diamond)^{-1} a_T^\diamond. \quad (24)$$

Equations (8) and (9) imply that

$$E(R_t^e) = E [R_t^e (f_t - \mu)'] b^\diamond. \quad (25)$$

Corresponding to the right-hand side of (25) is a vector of predicted expected returns, $d_T \hat{b}^\diamond$.

The cross-sectional R^2 measure is:

$$R_\diamond^2 = 1 - \frac{(\bar{R}^e - d_T \hat{b}^\diamond)' (\bar{R}^e - d_T \hat{b}^\diamond)}{(\bar{R}^e - \bar{R}^e)' (\bar{R}^e - \bar{R}^e)}. \quad (26)$$

Equation (25) can be rewritten as

$$E(R_t^e) = E \underbrace{[R_t^e (f_t - \mu)']}_{\beta_f} \underbrace{\text{var}(f_t)^{-1} \text{var}(f_t) b^\diamond}_{\lambda_f}. \quad (27)$$

Since $m_t^\diamond = m_t/\xi$ and $b^\diamond = b/\xi$, it follows that λ_f and β_f are the same objects defined in (5).

As in the previous case, the covariance matrix of f_t is estimated by GMM using the moment restriction (17). An estimate of λ_f is given by

$$\hat{\lambda}_f^\diamond = \hat{\Sigma}_f \hat{b}^\diamond, \quad (28)$$

where $\hat{\Sigma}_f = S_f$, as before. Standard errors for $\hat{\lambda}_f^\diamond$ are obtained by the delta method using the joint distribution of \hat{b}^\diamond , $\hat{\mu}$ and $\hat{\Sigma}_f$. The details are discussed in Appendix A.

4 Empirical Findings

Here I consider two sets of data. The first is from Lustig and Verdelhan (2007) (henceforth LV), and consists of eight portfolios of positions in foreign currency, with the U.S. dollar as the home currency, sorted on the basis of the foreign interest rates. The second set of data is widely used in the consumption-based asset pricing literature: the Fama-French set of 25 portfolios of US stocks sorted on size and the book-to-market value ratio (henceforth, FF25). The data are described in more detail in Appendix B.

4.1 Lustig-Verdelhan Foreign Exchange Portfolios

I consider the annual real US dollar excess returns to portfolios of short-term foreign government securities denominated in foreign currency. The sample period is 1953–2002. LV form

these portfolios on the basis of the interest rates on the underlying securities. In particular the real excess returns on a large number of countries' treasury securities are sorted into eight bins in each period according to the nominal interest rates on the securities, from lowest to highest. The returns to holding equally-weighted portfolios of each bin are then calculated.

To explain these returns I use three risk factors individually. These are:

- the log-growth rate of real per capita consumption of nondurables and services, which I will refer to from now on as “consumption growth”
- the log-growth rate of the real stock of consumer durables, which I will refer to from now on as “durables growth”
- the “market premium”; i.e. the difference between the return to a value-weighted portfolio of all US stocks and the risk free rate, denoted $Rm-Rf$, from the Fama-French database.

Apart from these individual risk factors, I also consider three groupings of risk factors from the literature:

- consumption growth and durables growth, which I refer to as the “Consumption Factors,”
- consumption growth, durables growth, and Rm , the market return, as in Yogo (2006) and LV, which I refer to as the “Yogo Factors,”
- $Rm-Rf$, SMB (the return differential between small cap and big cap equities) and HML (the return differential between high book-to-market and low book-to-market equities) from the Fama and French (1993) database. I refer to these as the Fama-French factors.

Tables 1 and 2 and Figures 1 and 2 present the empirical results for these data.

4.1.1 Results with Individual Factors

The top half of Table 1 shows point estimates, measures of fit and test statistics obtained using the a -normalization, while the top half of Table 2 shows results obtained using the ξ -normalization. Figure 1 plots average returns for the individual portfolios against the model-predicted expected returns, based on the second stage GMM estimates. The first

stage estimates for the ξ -normalization are numerically identical to estimates obtained using the two-pass method without a constant in the cross-sectional regression.⁵

Consumption Growth For consumption growth, the results are dramatically different, in qualitative terms, depending on which of the two normalizations is used. In particular, when the a -normalization is used, $\hat{b}^* = 48.8$ and is highly significant, with a t-statistic of 11.3. The estimated factor risk premium is $\hat{\lambda}_f = 4.3$ and is significant at the 5 percent level. The cross-sectional R^2 measure is 0.87. The good fit of the model can be seen in Figure 1(a), which show the model's predicted expected returns plotted against the actual mean returns observed in the data for the eight portfolios. Finally, the model easily passes a test of the over-identifying restrictions.

In contrast, when the ξ -normalization is used, the estimate of b^\diamond is much less precise, though it remains significant at the 5 percent level in the second stage of GMM. The estimated factor risk premium is $\hat{\lambda}_f = 1.5$, which is much smaller than when the a -normalization is used, though it remains significant at the 5 percent level. Figure 1(b) shows that the cross-sectional fit of the model is very poor. The R^2 measure is 0.06 at the second stage of GMM. The over-identifying restrictions are rejected at the 3 percent level of significance.

Durables Growth A similar story holds for durables growth. When the a -normalization is used, the estimate of b^* is positive, large and highly significant. The cross-sectional R^2 measure at the second stage of GMM is 0.81. The good fit of the model can also be seen in Figure 1(c). The model easily passes the test of the over-identifying restrictions. In contrast, when the ξ -normalization is used, the estimate of b^\diamond is much less precise, though, again, it remains significant at the 5 percent level in the second stage of GMM. Figure 1(d) shows that the cross-sectional fit of the model is very poor, with the R^2 measure being -0.13 after the second stage of GMM. The over-identifying restrictions are rejected at the 5 percent level.

The Market Premium The results are also quite striking for the market excess return. Using the a -normalization, the estimate of b^* is positive and significant, the R^2 of the model is 0.45, and the model is not rejected at the 5 percent level. However, when one turns to the

⁵See the appendix of Burnside (2007b) for a detailed discussion of the equivalence between GMM and the two-pass method.

ξ -normalization, the estimate of b° is insignificant, the estimate of λ_f shrinks by a factor of eight, the R^2 of the model is close to zero, and the over-identifying restrictions are strongly rejected.

4.1.2 Results with Groups of Risk Factors

When we turn to groups of risk factors, the most striking statistical sensitivity of the results to the normalization pertains to the R^2 statistics and the tests of the over-identifying restrictions. Estimates of the factor risk premia shrink in magnitude in all three cases, though their significance is not dramatically affected. Results for the a -normalization are in the bottom half of Table 1. Results for the ξ -normalization are in the bottom half of Table 2. Figure 2 plots average returns for the individual portfolios against the model-predicted expected returns.

Consumption Factors Consider, first the model that uses the two consumption factors. This specification corresponds to a linearized version of a consumption-based asset pricing model in which households have expected utility preferences and derive utility from nondurables and services and the stock of durable goods. The apparent fit of the model based on the a -normalization is very good, while the fit of the model based on the ξ -normalization is quite poor. Using the a -normalization the R^2 measure of cross-sectional fit suggests that the model can explain at least 81 percent of the variation in expected returns and the model easily passes the test of the over-identifying restrictions. With the ξ -normalization, however, rather than being close to 1, both R^2 measures are close to zero. Figures 2(a) and 2(b) paint a similar picture. They plot the sample values of \bar{R}^e against the model predicted values $D_T \hat{b}^*$, for the second stage of GMM. For the a -normalization, the model easily passes the test of the over-identifying restrictions. For the ξ -normalization, the over-identifying restrictions are rejected at the 3 percent level.

Yogo Factors Next, consider the linear factor model from Yogo (2006) and LV, which uses consumption growth, durables growth and Rm as the factors. This is a linearized version of a consumption-based asset pricing model in which households have Epstein and Zin (1989) preferences and derive utility from nondurables and services and the stock of durable goods. While the model based on the a -normalization appears to have nearly perfect fit, with an R^2 of between 0.89 and 0.95, the fit of the ξ -normalization is far from the mark, with the R^2

measure being 0.34 at the first stage, and -0.08 at the second stage of GMM. This despite the fact that three risk factors are being used to explain the returns to only eight portfolios. Figures 2(c) and 2(d) paint a similar picture. The apparent fit of the model based on the a -normalization is very good, while the fit of the model based on the ξ -normalization is quite poor. The tests of the over-identifying restrictions are also somewhat different across the two cases. For the a -normalization the model passes with flying colors (the p-value for the test is 0.86), while for the ξ -normalization the p-value on the J -statistic falls to 0.15.

Fama-French Factors Finally, consider the three risk factors proposed by Fama and French (1993). In this case, the results are dramatically different across the two normalizations. The signs of the coefficients on the three risk factors flip in moving from one normalization to the other, and the significance of the $Rm-Rf$ and HML factors is lost when moving to the ξ -normalization. The R^2 measure of fit falls from between 0.27 and 0.56 for the a -normalization to close to zero when the ξ -normalization is used. The model fails the test of the over-identifying restrictions at below the 5 percent level with the ξ -normalization, but easily passes the test for the a -normalization.

4.2 Fama-French 25 Portfolios

I examine the real excess returns to the FF25 portfolios at the quarterly frequency over the period 1949Q1–2005Q4. To explain these returns I use the same risk factors as mentioned above, except that the data are measured at the quarterly frequency. Tables 3 and 4 and Figures 3 and 4 show the results for these data.

4.2.1 Results with Individual Factors

The top halves of Tables 3 and 4 show the results from estimating a linear factor model using each of the scalar risk factors. Figure 3 plots average returns for the individual portfolios against the model-predicted expected returns.

Consumption Growth For consumption growth, the results are dramatically different, in qualitative terms, depending on which of the two normalizations is used. In particular, when the a -normalization is used, $\hat{b}^* = 143$ and is highly significant, with a t -statistic of 14, while $\hat{\lambda}_f = 1.38$ and is highly significant. The cross-sectional R^2 s at the first and second stages of GMM are 0.81 and 0.51, respectively. The good fit of the model can also be seen in

Figure 3(a), which shows the model’s predicted expected returns plotted against the actual mean returns observed in the data for the 25 portfolios. The model passes the test of the over-identifying restrictions, given a p-value of 0.18.

In contrast, when the ξ -normalization is used, \hat{b}^\diamond is a large positive number, but only has a t -statistic of 2, being much less precisely estimated than \hat{b}^* . The factor risk premium is much smaller, $\hat{\lambda}_f = 0.38$, but remains significant at the 5 percent level. The cross-sectional R^2 , however, is negative at both stages of GMM. The model’s bad fit is also seen in Figure 3(b). The model still passes the test of the over-identifying restrictions, but narrowly, with a p-value of 0.051.

Durables Growth For the growth rate of the stock of durables, \hat{b}^* is positive, large and highly significant using the a -normalization, while it is effectively zero using the ξ -normalization. So, the a -normalization leads us to conclude that durables growth helps price the returns, while the ξ -normalization leads us to the opposite conclusion. The fit of the model appears to be quite good using the a -normalization, but not with the ξ -normalization, as can be seen in Figures 3(c) and 3(d). In neither case is λ_f estimated to be significant, nor are the over-identifying restrictions rejected.

The Market Premium The results for the market premium are less striking. The sign and statistical significance of \hat{b}^* and \hat{b}^\diamond are similar. The same is true for the factor risk premia. The R^2 measures are negative in both cases. Figures 3(e) and 3(f) show that, if anything, the fit of the model is somewhat better with the second stage estimates of the ξ -normalization, rather than the a -normalization. The over-identifying restriction are rejected at very low significance levels in both cases.

4.2.2 Results with Groups of Risk Factors

The results are very dramatic when we turn to groups of risk factors. Results for the a -normalization are in the bottom half of Table 3. Results for the ξ -normalization are in the bottom half of Table 4. Figure 4 plots average expected returns for the individual portfolios against the model-predicted expected returns.

Consumption Factors Consider, first a model that uses the two consumption factors. Using the a -normalization, both elements of \hat{b}^* are positive and significant. The factor risk

premia are positive but are not individually significant. The R^2 measure of cross-sectional fit suggests that the model can explain at least 96 percent of the variation in expected returns. The model also easily passes a test of the over-identifying restrictions.

With the ξ -normalization the results are dramatically different. Neither of the b coefficients is significant, while the factor risk premia are both small and insignificant, at the second stage of GMM. Rather than being close to 1, both R^2 measures are negative. Figures 4(a) and 4(b) paint a similar picture. The apparent fit of the model based on the a -normalization is very good, while the fit of the model based on the ξ -normalization is very poor. The model still passes a test of the over-identifying restrictions, but at a much lower level of significance.

Yogo Factors Next, consider the linear factor model from Yogo (2006) that uses consumption growth, durables growth and Rm as the factors. Using the a -normalization the b parameters on consumption growth and durables growth are positive and significant. The factor risk premia are positive but not individually significant. The R^2 measures suggest that the model can explain at least 97 percent of the variation in expected returns. The model also easily passes a test of the over-identifying restrictions.

With the ξ -normalization the results are dramatically different. The b coefficients on consumption and durables growth remain large but no longer have any statistical significance. The factor risk premia are much smaller in magnitude and remain insignificant, except for the market return, which actually becomes significant. Rather than being close to 1, both R^2 measures are negative. Figures 4(c) and 4(d) paint a similar picture. The apparent fit of the model based on the a -normalization is very good, while the fit of the model based on the ξ -normalization is very poor. The model still passes the test of the over-identifying restrictions, but at a slightly smaller level of significance.

The parameters and degree of fit obtained under the a -normalization are qualitatively similar to the results obtained by Yogo (2006). However, he uses the ξ -normalization rather than the a -normalization to estimate the model. In section 4.4, I suggest that Yogo's particular choice of estimator is the main explanation for his positive assessment, despite his use of the ξ -normalization.

Fama-French Factors Finally, consider the three Fama-French factors. Here the results are roughly robust across the two normalizations. The signs and degree of significance of the

coefficients and risk premia are invariant to the choice of normalization. The model fails the J -test using both normalizations. But in both cases the model has quite good fit, at least at the first stage of GMM.

4.3 Accounting for the Findings Using the Two Normalizations

Recall that the two normalizations are equivalent in the following sense. Under the null hypothesis that the true SDF is given by (1), m_t^* and m_t^\diamond are proportional to one another when evaluated at the true parameter values. We have seen several cases, however, where the significance of the b coefficients, the apparent fit of the model, and the estimated risk premia, differ dramatically for the two normalizations. What might explain this apparent lack of equivalence?

One possibility is that there are severe small sample problems in estimating the models. The Monte Carlo simulation experiments in section 6 suggest that this is not the case. The explanation that I focus on, instead, is the possibility that one or more of the risk factors, or some linear combination of the risk factors, is uncorrelated with the returns. I explore this possibility in detail in section 5. Here I provide a quick sketch in the case of a single factor model.

Suppose that $E(R_{it}^e) \neq 0$ for at least one element of R_t^e . Suppose, also, that f_t is a spurious factor, so that $\text{cov}(R_t^e, f_t) = 0$. This means the SDF in (1) cannot be the true SDF because f_t cannot price the elements of R_{it}^e with non-zero means. Consider the moment condition (7). Under the assumptions just given, it can be rewritten

$$E(R_t^e)[1 - E(f_t)b^*] = 0. \quad (29)$$

Notice that if $E(f_t) \neq 0$ the unique solution to (29) is $b^* = 1/E(f_t)$. Now consider the moment conditions (9). Under the assumptions just given they can be rewritten as

$$E(f_t) - \mu = 0 \quad E(R_t^e) \{1 - [E(f_t) - \mu] b^\diamond\} = 0 \quad (30)$$

Since $E(R_{it}^e) \neq 0$ for at least one element of R_t^e , there is no solution for (b^\diamond, μ) . Not surprisingly, as sections 6 and 7 demonstrate, a GMM estimator based on (7) provides misleadingly favorable inference about the SDF. GMM estimators based on (9) encounter their own difficulties, but these do not arise because there is a spurious solution to (9).

There is another, more limited, sense in which a factor can be spurious. Suppose that f_t is *panel spurious* in that the covariance between R_{it}^e and f_t itself does not vary with i . That

is, $\text{cov}(R_t^e, f_t) = c\iota$, where ι is an $n \times 1$ vector of ones, and $c \neq 0$ is some scalar constant. With these assumptions, if the SDF were true then it would follow that $E(R_{it}^e) = cb^\circ$ for $i = 1, \dots, n$. So if $E(R_{it}^e) \neq E(R_{jt}^e)$ for at least one pair of assets, i and j , $i \neq j$, the SDF must be false. Sections 6 and 7 also explore the consequences of panel spuriousness for GMM estimates based on the two normalizations.

When the risk factors are known to be highly correlated with the individual asset returns being studied, and the covariance varies across assets, the spuriousness hypothesis is implausible. To show that spuriousness is a plausible explanation of the results obtained with the LV and FF25 data, I now present evidence on the covariance between the returns and risk factors.

First, I directly estimate the covariances, c_i , between the risk factors and the asset returns using GMM by exploiting the moment restrictions

$$E(f_t - \mu) = 0 \quad (31)$$

$$E[R_{it}^e(f_t - \mu) - c_i] = 0, \quad i = 1, 2, \dots, n. \quad (32)$$

To test for spuriousness I test the null hypothesis that $c_i = 0$ for all i . To test for panel spuriousness I test the null hypothesis that $c_i = c$ for all i .

Second, I estimate the factor betas, $\beta_i = c_i/\sigma_f^2$, by GMM using the following moment restrictions:

$$E(R_{it}^e - a_i - \beta_i f_t) = 0, \quad i = 1, 2, \dots, n. \quad (33)$$

$$E[(R_{it}^e - a_i - \beta_i f_t)f_t] = 0, \quad i = 1, 2, \dots, n. \quad (34)$$

To test for spuriousness I test the null hypothesis that $\beta_i = 0$ for all i . To test for panel spuriousness I test the null hypothesis that $\beta_i = \beta$ for all i .

Table 5 presents the covariances and betas for the LV data. Notably only one portfolio's return, #2, is significantly correlated with consumption growth. The covariance for this portfolio is significant at the 10 percent level, while the factor beta is significant at the 5 percent level. A joint test of the null hypothesis that $c_i = 0$ for all i has a p-value of 0.447, while the p-value associated with the null hypothesis that $\beta_i = 0$ for all i is 0.048. Which test result is more plausible? Using Monte Carlo simulation, Burnside and Eichenbaum (1996) show that GMM-based Wald tests have excessive size in small samples. Therefore Table 5 also presents small-sample p-values for the joint tests computed using Monte Carlo

simulations described in Appendix C. These p-values are 0.618 for the covariance-based test and 0.334 for the beta-based test.

A similar story holds for durables growth. The covariance for Portfolio #7 is statistically significant at the 10 percent level, while the factor beta for Portfolio #7 is significant at the 5 percent level. The joint tests do not reject that $c_i = 0$ and $\beta_i = 0$ for all i at conventional significance levels.

The market return is not significantly correlated with any of the portfolio returns, and the joint tests do not reject that $c_i = 0$ and $\beta_i = 0$ for all i at conventional significance levels.

Table 6 presents the covariances and betas for the FF25 returns. For consumption growth, most of the covariances and betas are significant at the 5 percent level, and all are significant at the 10 percent level. Indeed, when I test the null hypothesis that $\beta_i = 0$ for all i , it is rejected at the 0.001 percent level of significance, although the null hypothesis that $c_i = 0$ for all i is only rejected at the 8.5 percent level of significance. P-values based on Monte-Carlo simulations are larger, but some doubt is cast on the notion that consumption growth is a spurious factor. It is striking, however, that the regression coefficients and covariances are quite similar in value across the returns. The standard deviation of $\hat{\beta}_i$ across i is just 0.76, while the average standard error of $\hat{\beta}_i$ is nearly twice as large at 1.39. This suggests that $\text{cov}(R_{it}^e, f_t)$ does not vary very much across i . Indeed, when I test the null hypothesis that $c_i = c$ for all i , the p-value is 0.29. On the other hand, when I test the null hypothesis that $\beta_i = \beta$ for all i , the asymptotic p-value is 0.009, though the small sample p-value based on Monte Carlo experiments is 0.172. I conclude that it is reasonable to hypothesize that consumption growth is a panel spurious factor for the FF25 returns based on this somewhat mixed evidence.

Turning to durables growth, the evidence suggests that it is spurious. None of the individual regression coefficients and covariances is significant at the 5 percent level. Furthermore, the null hypothesis that $c_i = 0$ for all i is not rejected at conventional significance levels. When I test the null hypothesis that $\beta_i = 0$ for all i , the asymptotic p-value is 0.115, while the small sample p-value based on Monte Carlo experiments is 0.461.

Finally, consider the market excess return. Here there is no evidence of either form of spuriousness. All of the regression coefficients and covariances are significant at the 1 percent level. All joint hypothesis tests lead to overwhelming rejection of both types of spuriousness.

4.4 Robustness

In Appendix D, I explore the robustness of my findings. In particular, I compute long-run covariance matrices using HAC estimators rather than using the VARHAC procedure of den Haan and Levin (2000). Following Cochrane (2005), I use an alternative weighting matrix at the second stage of GMM when estimating the ξ -normalization. I also consider iterated GMM, where the weighting matrix is recomputed multiple times until approximate convergence of the parameter estimates. Mimicking Parker and Julliard (2005) I introduce a common pricing error parameter, akin to the constant sometimes used in the cross-sectional regression of two-pass procedures used to estimate the beta representation of the model. In no case do any of these changes affect my qualitative conclusion that the fit of the estimated models is relatively poor and the statistical significance of the estimated parameters is relatively weak, when using the ξ -normalization.

As a final check on robustness I adopt a variant of the ξ -normalization used by Yogo (2006). In this variant, all information about μ that is available in the moment restrictions is used to estimate it. In the benchmark approach the GMM estimator sets $a_T^\diamond g_T^\diamond = 0$ and a_T^\diamond is designed so that $\hat{\mu} = \bar{f}$. But there is information about μ in the moment condition $E(R_t^e m_t^\diamond) = 0$ because $m_t^\diamond = 1 - (f_t - \mu)' b^\diamond$. Because Yogo's approach takes this information into account, $\hat{\mu}$ will not, in general, be equal to \bar{f} . With this variant of the ξ -normalization I find that the models perform better overall because the estimated means of the factors move substantially to fit the asset pricing conditions. But this depends critically on how the weighting matrix is specified in the first stage of GMM. For the LV data the fit of the models is poor at the second stage of GMM, while parameter estimates are insignificant at the first stage of GMM. For the FF25 data, Yogo's multi-factor model fits considerably better than in my benchmark case. However, again, I find this to be highly dependent on how the weighting matrix is set at the first stage of GMM. If further GMM iterations are performed, this sensitivity to the initial weighting matrix vanishes, and the results are similar to those for my benchmark approach.⁶

5 Large Sample Properties of the GMM Estimators

In this section I explore the asymptotic properties of the GMM procedures described above.

⁶I explore the Parker and Julliard (2005) and Yogo (2006) approaches in more detail in a related paper, Burnside (2007a).

Assumption 1 Let the true SDF be given by (1) and let the true values of the parameters a and b be denoted a_0 and b_0 . Let the true values of b^* , b^\diamond and μ be denoted $b_0^* = b_0/a_0$, $b_0^\diamond = b_0/(a_0 - \mu_0' b_0)$ and $\mu_0 = E(f_t)$.

Assumption 2 Let $E(R_t^e f_t')$ and $\text{cov}(R_t^e, f_t')$ have full column rank.

Assumption 3 Assume that $D_T \xrightarrow{a.s.} D_0 = E(R_t^e f_t')$, $d_T \xrightarrow{a.s.} d_0 = \text{cov}(R_t^e, f_t)$, $W_T^* \xrightarrow{a.s.} W_0^*$ and $W_T^\diamond \xrightarrow{a.s.} W_0^\diamond$, with W_0^* and W_0^\diamond positive definite.

For compactness of notation let $\theta^\diamond = (b^\diamond, \mu)$ and $\theta_0^\diamond = (b_0^\diamond, \mu_0)$. Define

$$\delta_0^* = E \left[\frac{\partial u_t^*(b_0^*)}{\partial b^*} \right] = -E(R_t^e f_t') = -D_0$$

and

$$\delta_0^\diamond = E \left[\frac{\partial u_t^\diamond(\theta_0^\diamond)}{\partial \theta^\diamond} \right] = \begin{pmatrix} -\text{cov}(R_t^e, f_t) & E(R_t^e) b_0^{\diamond'} \\ 0 & -I_k \end{pmatrix} = \begin{pmatrix} -d_0 & E(R_t^e) b_0^{\diamond'} \\ 0 & -I_k \end{pmatrix}$$

Let $S_0^* = E \left[\sum_{j=-\infty}^{+\infty} u_t^*(b_0^*) u_{t-j}^*(b_0^*)' \right]$ and $S_0^\diamond = E \left[\sum_{j=-\infty}^{+\infty} u_t^\diamond(\theta_0^\diamond) u_{t-j}^\diamond(\theta_0^\diamond)' \right]$. Define $a_0^* = D_0' W_0^*$ and

$$a_0^\diamond = \begin{pmatrix} d_0' W_0^\diamond & 0 \\ 0 & I_k \end{pmatrix}.$$

Theorem 1 Suppose assumptions 1–3 are satisfied. Under additional regularity conditions provided in Hansen (1982) $\hat{b}^* \xrightarrow{a.s.} b_0^*$, $\hat{\theta}_0^\diamond \xrightarrow{a.s.} \theta_0^\diamond$, $\sqrt{T}(\hat{b}^* - b_0^*) \xrightarrow{d} N(0, V_b^*)$ and $\sqrt{T}(\hat{\theta}_0^\diamond - \theta_0^\diamond) \xrightarrow{d} N(0, V_\theta^\diamond)$ with $V_b^* = (a_0^* \delta_0^*)^{-1} a_0^* S_0^* a_0^{*\prime} (\delta_0^{*\prime} a_0^*)^{-1}$ and $V_\theta^\diamond = (a_0^\diamond \delta_0^\diamond)^{-1} a_0^\diamond S_0^\diamond a_0^{\diamond'} (\delta_0^{\diamond'} a_0^\diamond)^{-1}$. Also, $R_*^2 \xrightarrow{a.s.} 1$ and $R_\diamond^2 \xrightarrow{a.s.} 1$. The statistics J^* and J^\diamond both converge in distribution to χ^2 random variables with $n - k$ degrees of freedom.

The proof of Theorem 1 is provided in Appendix A. The interpretation of Theorem 1 is that when the model is true (Assumption 1), and when the moment conditions are informative about the risk factors (Assumption 2), both approaches to estimation work well.

Next I turn to a situation where the model remains true, but the returns in the data being studied do not fully shed light on the relevance of the risk factors.

Assumption 2a Let $\text{cov}(R_t^e, f_t')$ have rank $r < k$.

Notice that if we maintain assumption 1, that the model is true, then $E(R_t^e) = \text{cov}(R_t^e, f_t') b_0$ and $E(R_t^e f_t') = \text{cov}(R_t^e, f_t') [I_k + b_0 E(f_t')]$. Hence $E(R_t^e f_t')$ has rank less than or equal to r .

Theorem 2 *Suppose assumptions 1, 2a and 3 are satisfied. It follows that neither b^* nor b^\diamond is asymptotically identified. Nonetheless, $R_*^2 \xrightarrow{a.s.} 1$ and $R_\diamond^2 \xrightarrow{a.s.} 1$.*

The proof of Theorem 2 is provided in Appendix A. The interpretation of Theorem 2 is that even though the model is true (Assumption 1), the moment conditions are insufficiently informative about the risk factors (Assumption 2a), and neither approach to estimation will work well in large samples. Presumably the lack of asymptotic identification would be manifested in large standard errors for \hat{b}^* and \hat{b}^\diamond in finite samples. While the parameters of the SDF are not identified, this is because there are multiple values of b^* and b^\diamond that satisfy the moment conditions asymptotically. As a result the measures of fit limit to 1.

Next I turn to a situation where the model is false.

Assumption 1b Assume that the true SDF is not given by (1) so that, in general, $E(R_t^e) \neq \text{cov}(R_t^e, f_t)b$ for the particular vector f_t being studied.

Assumption 2b Let $E(R_t^e, f_t')$ have rank k but $\text{cov}(R_t^e, f_t')$ have rank $k - 1$.

Assumption 2b is feasible because we are no longer maintaining assumption 1. The matrix $E(R_t^e f_t')$ can have full column rank, even when the matrix $\text{cov}(R_t^e, f_t')$ has less than full column rank.

Theorem 3 *Suppose assumptions 1b, 2b and 3 are satisfied. Under additional regularity conditions provided in Hansen (1982) $\hat{b}^* \xrightarrow{a.s.} b_s^* = x/[x'E(f_t)]$, where x is the unique element of the nullspace of $\text{cov}(R_t^e, f_t)$ whose elements sum to 1, and $R_*^2 \xrightarrow{a.s.} 1$. At least one element of b_s^* is non-zero. In contrast b^\diamond is not asymptotically identified.*

The proof of Theorem 3 is provided in Appendix A. Since the model is not true, the parameter vector b_s^* has no interpretation as the “true” value of b^* . Rather b_s^* is a degenerate value of b^* for which the moment condition (6) holds, even though (1) is not the true SDF. The sum of the elements of b_s^* is $1/[x'E(f_t)]$ which is the inverse of a weighted average of the means of the risk factors.

Before turning to the asymptotic distributions of \hat{b}^* and J^* it is helpful to define the following notation. Let $S_s^* = E \sum_{j=-\infty}^{+\infty} u_t^*(b_s^*)u_{t-j}^*(b_s^*)'$ and $V_s^* = E[u_t^*(b_s^*)u_t^*(b_s^*)']$. In general $S_s^* \neq V_s^*$ under Assumption 1b. Let $A_0^* = I_n - \delta_0^*(a_0^*\delta_0^*)^{-1}a_0^*$ and define $V_g^* = A_0^*V_s^*A_0^{*'} and$

$V_0^* = A_0^* S_s^* A_0^{*'}.$ Finally, diagonalize V_0^* as $V_0^* = P_0 \Lambda_0 P_0'$ where the columns of P_0 are the orthonormal eigenvectors of V_0^* and Λ_0 is a diagonal matrix with the eigenvalues of V_0^* on the diagonal. Let $\tilde{V}_0^* = P_0 \Lambda_0^{1/2}$ so that $\tilde{V}_0^* \tilde{V}_0^{*'} = V_0^*.$

Theorem 4 *Under the assumptions of Theorem 3, $\sqrt{T}(\hat{b}^* - b_s^*) \xrightarrow{d} N(0, V_b^*)$ with $V_b^* = (a_0^* \delta_0^*)^{-1} a_0^* S_s^* a_0^{*'} (\delta_0^{*'} a_0^*)^{-1}.$ If $S_s^* = V_s^*, J^* \xrightarrow{d} \chi_{n-k}^2.$ When $S_s^* \neq V_s^*, J^* \xrightarrow{d} \sum_{i=1}^{n-k} \lambda_i z_i^2$ where z_1, z_2, \dots, z_{n-k} are mutually independent standard normal random variables and $\lambda_1, \lambda_2, \dots, \lambda_{n-k}$ are the non-zero eigenvalues of the matrix $\tilde{V}_0^{*'} (V_g^*)^+ \tilde{V}_0^*.$*

The proof of Theorem 4 is provided in Appendix A. Some examples are helpful in interpreting Theorems 3 and 4.

Single Factor Model With a Spurious Factor In a single factor model $k = 1$ so Assumption 2b implies that $\text{cov}(R_t^e, f_t) = 0$ and, since $E(R_t^e f_t)$ has rank 1, that $D_0 = E(R_t^e f_t) = E(R_t^e) E(f_t) \neq 0.$ In this case $b_s^* = 1/E(f_t).$ A researcher testing the model under the null would compute an inconsistent estimate of V_b^* that would converge, instead, to $(a_0^* \delta_0^*)^{-1} a_0^* V_s^* a_0^{*'} (\delta_0^{*'} a_0^*)^{-1}.$ Nonetheless, because this matrix is finite, the t -statistic for \hat{b}^* would diverge to $+\infty$ if $E(f_t) > 0$ or $-\infty$ if $E(f_t) < 0.$ The predicted expected returns, $D_T \hat{b}^*,$ would converge almost surely to $D_0 b_s^* = E(R_t^e).$ Therefore $R_*^2 \xrightarrow{a.s.} 1.$ Thus, in large samples a researcher testing the model using the a -normalization would conclude that the factor f_t prices the returns (due to the statistical significance of \hat{b}^*) and that the model's fit is perfect. A researcher testing the over-identifying restrictions using, say, a 5 percent critical value from the χ_{n-1}^2 distribution would only reject the model about 5 percent of the time in repeated large samples if $S_s^* = V_s^*.$ It is unclear what would happen in the more general case when $S_s^* \neq V_s^*.$

For the ξ -normalization, Hansen's identification condition requires that there should be a unique b° that solves $(d_0' W_0^\circ d_0) b^\circ = d_0' W_0^\circ E(R_t^e).$ Clearly any value of b° will work when $d_0 = 0.$

Multi-factor Model with a Single Spurious Factor In a multi-factor model $k > 1.$ Without loss of generality, let $\text{cov}(R_t^e, f_{kt}) = 0.$ In this case, the vector x referred to in the statement of Theorem 3 has a 1 as its k th element and zeros elsewhere. So $b_s^* = x/[x' E(f_t)]$ has $1/E(f_{kt})$ as its k th element and zeros elsewhere. Oddly enough, a researcher testing the model using the a -normalization and a large sample of data would conclude that the factor

f_{kt} prices the returns and that the other factors are irrelevant, even though they are the only ones correlated with R_t^e . As in the previous example, $R_*^2 \xrightarrow{a.s.} 1$. A researcher testing the over-identifying restrictions using a 5 percent critical value from the χ_{n-k}^2 distribution would only reject the model about 5 percent of the time in repeated large samples if $S_s^* = V_s^*$.

As in the previous example, for the ξ -normalization, the identification condition is not satisfied because there is no unique b^\diamond that solves $(d_0' W_0^\diamond d_0) b^\diamond = d_0' W_0^\diamond E(R_t^e)$. When the last column is d_0 is zero, but the rest of the matrix has full column rank, it is b_k^\diamond that is unidentified asymptotically. The rest of the parameter vector b^\diamond is identified. If it is also true that $m_t^\diamond = 1 - (f_t - \mu)' b^\diamond$ but $b_k^\diamond = 0$, so that the proposed SDF nests the true SDF, the components of \hat{b}^\diamond other than \hat{b}_k^\diamond will converge in probability to their true values. For more general forms of misspecification these components of \hat{b}^\diamond will converge in probability, although the probability limits will not be interpretable as true values.

Multi-factor Model with Multicollinearity Among the Factors Now consider the more general case where $k > 1$, where $\text{cov}(R_t^e, f_t)$ has no zero columns, but $\text{cov}(R_t^e, f_t)$ has rank $k-1$. In this case, no one factor is spurious, but there is an identification problem akin to multicollinearity. But this identification problem only arises when using the ξ -normalization. Using the a -normalization, $\hat{b}^* \xrightarrow{a.s.} b_s^* = x/[x'E(f_t)]$ and $R_*^2 \xrightarrow{a.s.} 1$. Remarkably, the estimated model puts all its weight on an irrelevant linear combination of the risk factors because

$$\text{cov}(R_t^e, f_t' b_s^*) = d_0 x / [x'E(f_t)] = 0.$$

Any elements of \hat{b}^* that correspond to non-zero elements of b_s^* will be found to be significantly different from zero in large samples. A researcher testing the over-identifying restrictions using a 5 percent critical value from the χ_{n-k}^2 distribution would only reject the model about 5 percent of the time in repeated large samples if $S_s^* = V_s^*$.

Greater Numbers of Spurious Factors Theorems 3 and 4 relied on the assumption that only one linear combination of the risk factors is spurious, in the sense that $\text{rank}[\text{cov}(R_t^e, f_t')] = k - 1$. If $\text{rank}[\text{cov}(R_t^e, f_t')] = k - r$, $r > 1$, then b^* is asymptotically unidentified because there will be many b^* such that $a_0^*[E(R_t^e) - D_0 b^*] = 0$. All these b^* , however, share the property that $E(f_t)' b^* = 1$ and $E(R_t^e) = D_0 b^*$. Presumably the lack of asymptotic identification would be manifested in large standard errors for the individual

elements of \hat{b}^* , but the particular linear combination $E(f_t)' \hat{b}^*$ would have a small standard error, and would be centered around 1. Also, the measure of fit, R_*^2 , would be roughly 1.

More Asymptotic Properties of the ξ -Normalization To this point we have only been able to show lack of asymptotic identification of the ξ -normalization when the rank condition on $\text{cov}(R_t^e, f_t')$ fails. Although b^\diamond is unidentified asymptotically, and does not have a well defined probability limit, it is still possible, under additional regularity conditions, to derive the asymptotic distributions of \hat{b}^\diamond , its associated t -statistic, the model's R^2 and the J -statistic used to test the over-identifying restrictions. Here I consider only the case of a single factor model with the returns and factors being iid and independent of each other. The more general case of a multifactor model with persistent returns or factors can be worked out, but at the cost of algebraic complexity.

Theorem 5 *Make the assumptions of Theorem 3. Let R_t^e and f_t be iid over time and mutually independent, and let $k = 1$. Let $\mu_R = E(R_t^e)$, $\Sigma_R = \text{var}(R_t^e)$, $\mu = E(f_t)$ and $\sigma_f^2 = \text{var}(f_t)$. Define the random variables $X \sim N(0, \sigma_f^2 \Sigma_R)$, $Z = (X' \mu_R)/(X' X)$ and $\tilde{Z} = (X' \Sigma_R^{-1} \mu_R)/(X' \Sigma_R^{-1} X)$. Then $T^{1/2} d_T \xrightarrow{d} X$. It follows that at the first stage of GMM $T^{-1/2} \hat{b}^\diamond \xrightarrow{d} Z$, $t(\hat{b}^\diamond) \xrightarrow{d} (X' \mu_R)/[\sigma_f^2 Z^2 (X' \Sigma_R X)]^{1/2}$ and $R_\diamond^2 \xrightarrow{d} 1 - (\mu_R' M \mu_R)/(\mu_R' M_\iota \mu_R)$, where $M = I_n - X(X' X)^{-1} X'$, $M_\iota = I_n - \iota \iota'/n$ and ι is an $n \times 1$ vector of ones. At the second stage of GMM $T^{-1/2} \hat{b}^\diamond \xrightarrow{d} \tilde{Z}$, $t(\hat{b}^\diamond) \xrightarrow{d} (X' \Sigma_R^{-1} \mu_R)/[\sigma_f^2 Z^2 (X' \Sigma_R^{-1} X)]^{1/2}$ and $R_\diamond^2 \xrightarrow{d} 1 - (\mu_R' \tilde{M}' \tilde{M} \mu_R)/(\mu_R' M_\iota \mu_R)$, where $\tilde{M} = I_n - X(X' \Sigma_R^{-1} X)^{-1} X' \Sigma_R^{-1}$, and $J \xrightarrow{d} (\mu_R' \tilde{M}' \Sigma_R^{-1} \tilde{M} \mu_R)/(\sigma_f^2 Z^2)$.*

The proof of Theorem 5 is provided in Appendix A.⁷ Interpreting Theorem 5 in general terms is difficult because of the dependence of the asymptotic distributions on μ_R and Σ_R . The distribution of \hat{b}^\diamond will spread out as the sample size increases at a rate of $T^{1/2}$. The t -statistic for \hat{b}^\diamond has a well-defined asymptotic distribution at both stages of GMM, so the probability of finding \hat{b}^\diamond to be significantly different from zero will converge to a number generally not equal to zero. Similarly, the J statistic used to test the over-identifying restrictions has a well-defined asymptotic distribution, so the probability of rejecting the model will converge to a number generally not equal to zero nor one. When we explore the small sample properties

⁷Similar asymptotic properties are derived in Kan and Zhang (1999a, 1999b) for the case where f_t is zero mean or, equivalently, has a known mean.

of GMM in section 6 using calibrated examples, we will use the expressions in Theorem 5 to derive the limiting distributions of \hat{b}^\diamond , $t(\hat{b}^\diamond)$, R_\diamond^2 and J as $T \rightarrow \infty$.

Approximate Failure of the Rank Condition The assumption that the rank condition literally does not hold may seem extreme. In single factor models it requires that $\text{cov}(R_t^e, f_t)$ is exactly zero. A standard device in the theory of weak instruments and unit root testing is also useful when it is preferable to assume that $\text{cov}(R_t^e, f_t)$ is small and asymptotically vanishing but not literally zero. Mimicking Hall's (2005) discussion of Staiger and Stock (1997), in the single factor case we might suppose that $R_t^e = \mu_R + c_T(f_t - \mu) + u_t$ where u_t is an $n \times 1$ vector that is uncorrelated with f_t and $c_T = T^{-1/2}c$ where c is a $n \times 1$ vector of constants. Working with this alternative assumption, however, does not change the result stated in Theorem 5 that $T^{1/2}d_T \xrightarrow{d} N(0, \sigma_f^2 \Sigma_R)$ nor does it change the fact that $\bar{R}^e \xrightarrow{p} \mu_R$. As a consequence, the results in Theorem 5 go through unchanged.

Panel Spuriousness We have not yet considered the case where a factor is panel spurious, that is, $\text{cov}(R_t^e, f_t) = c\iota$, where ι is an $n \times 1$ vector of ones, and $c \neq 0$ is some scalar constant. Suppose this is the case and the model being tested is a single factor model.

First, assume that the model is true, or, in other words, maintain Assumptions 1–3. Obviously this implies that $E(R_{it}^e) = cb_0^\diamond$ for all $i = 1, \dots, n$. Then Theorem 1 applies. Both normalizations produce GMM estimates with standard, desirable, asymptotic properties.

Now, assume, instead that the model is false. This means replacing Assumption 1 with Assumption 1b, so that, in general, $E(R_t^e) \neq \text{cov}(R_t^e, f_t)b$. Define $\omega_0^* = [(D_0'W_0^*\iota)/D_0'W_0^*D_0]c$ and $\omega_0^\diamond = \iota'W_0^\diamond E(R^e)/(\iota'W_0^\diamond \iota)$.

Theorem 6 *Suppose assumptions 1b, 2 and 3 are satisfied. Assume, also, that $k = 1$ and $\text{cov}(R_t^e, f_t) = c\iota$. Under additional regularity conditions provided in Hansen (1982) $\hat{b}^* \xrightarrow{a.s.} b_s^* = (1 - \omega_0^*)/E(f_t)$, and the predicted expected returns of the a -normalization converge almost surely to*

$$D_0 b_s^* = \left[E(R^e) + \frac{c}{E(f)} \iota \right] (1 - \omega_0^*).$$

Also $\hat{b}^\diamond \xrightarrow{a.s.} b_s^\diamond = \omega_0^\diamond/c$, and the predicted expected returns of the ξ -normalization converge almost surely to $d_0 b_s^\diamond = \omega_0^\diamond$.

The proof of Theorem 6 is provided in Appendix A. It is clear that in a sufficiently large

sample the model will be rejected using either normalization because the limit of the predicted expected returns, in general, is not $E(R^e)$. For small c , however, it is clear that the a -normalization will imply good fit since for $c \cong 0$, $b_s^* \cong 1/E(f_t)$ and $D_0 b_s^* \cong E(R^e)$. The fit of the ξ -normalization in large samples, in contrast, is independent of c and is very poor. In the first stage of GMM, $W_0^\diamond = I_n$ so that $\omega_0^\diamond = \sum_{i=1}^n E(R_{it}^e)/n$, the cross-sectional mean of the elements of $E(R^e)$. So at the first stage of GMM, $R_\diamond^2 \xrightarrow{a.s.} 0$. At the second stage of GMM, ω_0^\diamond is a weighted average of the elements of $E(R^e)$. So at the second stage of GMM, $\text{plim } R_\diamond^2 \leq 0$.

Inference Regarding Risk Premia I now turn to the issue of estimating and testing for the statistical significance of risk premia. As described earlier, in section 3, the estimates $\hat{\lambda}_f^*$ and $\hat{\lambda}_f^\diamond$ are obtained by adding appropriate moment restrictions and standard errors are obtained by the delta method. When Assumptions 1–3 standard asymptotic theory described in more detail in Appendix A goes through. My main concern, here, is with the properties of risk premia estimated using the a -normalization when the model is false and a linear combination of the risk factors is spurious (under Assumptions 1b, 2b and 3). Recall that the estimate of λ_f based on the a -normalization is given by (18), $\hat{\lambda}_f^* = S_f \hat{b}^*/(1 - \bar{f}'\hat{b}^*)$, where \bar{f} and S_f are, respectively, the sample mean and covariance matrix of f_t . Under assumptions 1b, 2b and 3 $\text{plim } S_f \hat{b}^* = \Sigma_f b_s^* = \Sigma_f x/[x'E(f_t)]$, but $1 - \bar{f}'\hat{b}^* \xrightarrow{a.s.} 0$. Since, in general, $\Sigma_f b_s^* \neq 0$, this implies that in large samples, the risk premia emerging from estimates of the a -normalization will diverge to $\pm\infty$ with probability 1. Would a researcher implementing the a -normalization find these estimates to be statistically significant if he computed the t -statistic for $\hat{\lambda}_f^*$?

Theorem 7 *Under the assumptions of Theorem 3, each element of the vector of squared t -statistics for $\hat{\lambda}_f^*$ converges in distribution to cz^2 where c is a scalar and $z^2 \sim \chi_1^2$. If the lag restrictions imposed in computing the long-run covariance matrix of the GMM errors are valid, $c = 1$.*

The proof of Theorem 7 is provided in Appendix A. The lag restriction imposed in computing the long-run covariance matrix of the GMM errors is that $u_t^*(b^*) = R_t^e(1 - f_t' b^*)$ is orthogonal to lagged information. If this restriction is valid, or if it not imposed on the computation of the long-run covariance matrix, the t -statistics will indicate significance of $\hat{\lambda}_f^*$ in only a small

fraction of repeated large samples when the model is false.

A Simple Diagnostic Test We have seen that when the assumptions 1b, 2b and 3 are satisfied the probability limit of \hat{b}^* is $b_s^* = x/[x'E(f_t)]$. This implies that in the limit, the mean of the fitted SDF, $\hat{m}_t^* = 1 - f_t'\hat{b}^*$, is zero. A simple test for this outcome can be performed by testing the null hypothesis that $\mu'b^* = 1$. Equivalently the mean of the SDF and the standard error of its mean can be calculated. If the null hypothesis that the mean of the SDF is zero cannot be rejected there is good reason to doubt favorable results based on the a -normalization.

Table 7 shows results corresponding to the estimates of the a -normalization that were presented in Tables 1 and 3 for the LV and FF25 data. For the LV data the mean of the estimated SDF is only significantly different from zero for the single factor model that uses the market premium as the risk factor. The other estimated SDFs are all approximately mean zero. These results are not unexpected given that we found the factor betas to be jointly statistically insignificant in section 4.

For the FF25 data the mean of the SDF is significantly different from zero for the single factor models with consumption growth and the market premium as the risk factors, as well as for the multi-factor model that uses the Fama-French factors. This is not unexpected, because we found consumption growth to be significantly correlated with the FF25 returns, and it is well known that the Fama-French factors are also correlated with them. On the other hand, for any model that includes durables growth as one of the factors the mean of the estimated SDF is small and statistically insignificant. This is not surprising given that, as we saw in section 4, durables growth is approximately uncorrelated with the FF25 returns.

Direct Tests of the Rank Condition An alternative to estimating the mean of the SDF is to directly test the rank condition on $\text{cov}(R_t^e, f_t)$. To do this I estimate the $n \times k$ matrix of factor betas, β_f , using GMM and the moment restrictions (33) and (34). Let $B = \text{vec}(\beta_f)$ and let $V(\hat{B})$ be a consistent estimator for the asymptotic covariance matrix of $\sqrt{T}(\hat{B} - B_0)$, where B_0 is the true value of B . Following Cragg and Donald (1997) and Wright (2003), to test the null hypothesis that the rank of β_f is $r < k$ I form the statistic

$$L(r) = \min_{P \in \Omega_r} T[\hat{B} - \text{vec}(P)]'V(\hat{B})^{-1}[\hat{B} - \text{vec}(P)]$$

where Ω_r is the set of all $n \times k$ matrices with rank r . If the true rank of B_0 is r , $L(r) \xrightarrow{d} \chi_{n-r}^2$. For single factor models $k = 1$ and the relevant hypothesis is that $r = 0$. In this case the $L(r)$ statistic is equivalent to testing that all the factor betas are zero.

Table 8 shows results for the rank tests corresponding to the models estimated in Tables 1–4 for the LV and FF25 data. For the single factor models, the test statistics and p-values correspond to the ones shown in Table 6. For the LV data, it is only for the single factor model based on consumption-growth that we can reject the null of insufficient rank, and this rejection is marginal. For all the other models, we cannot reject the null hypothesis that the matrix of factor betas has less than full column rank.

For the FF25 data we can reject the null hypothesis that $\text{cov}(R_t^e, f_t)$ has less than full column rank for the single factor consumption growth model, the single factor CAPM model, and the three factor Fama-French model. But for the durables growth model, the consumption factors model, and the Yogo factors model it is difficult to reject the null that $\text{cov}(R_t^e, f_t)$ has rank $k - 1$. This, presumably, reflects the inclusion of durables growth in these models.

6 Small Sample Properties of the GMM Estimators

To further demonstrate the sensitivity of empirical results to the choice of normalization in the presence of a spurious risk factor, I conduct three sets of Monte Carlo experiments. In the first set of experiments I see how the GMM procedures perform when I confront them with artificial spurious risk factor data in combination with either the LV or FF25 returns. I generate random data for an artificial factor that mimics US consumption data—in terms of mean and standard deviation—yet is either globally spurious or panel spurious. I show that a researcher draws dramatically different conclusions about the model depending on which normalization is used in estimating the model.

In the second set of experiments I generate data from artificial economies. These artificial economies bear some resemblance to the US economy, in that the consumption and returns series share some of the statistical properties of their US counterparts. Here, however, the simulated risk factors price the returns by construction. I generate spurious factors in parallel with the true factors and the returns. I then confront the GMM procedures with either the true risk factors or a set of alternative risk factors that include a spurious risk factor. I find that the two normalizations perform similarly when the researcher estimates the model using the true risk factor(s). However, when one of the risk factors in the proposed SDF is

spurious the researcher draws dramatically different conclusions about the model depending on which normalization is used in estimation.

6.1 Experiments Based on US Data

6.1.1 An Experiment Based on the Lustig-Verdelhan Data

LV's data consist of 50 annual observations on US consumption between 1953 and 2002. Accordingly, with MATLAB's random number generator, I create 10000 samples, each of length $T = 50$, for an artificial risk factor f_t , whose law of motion is

$$f_t = \mu(1 - \rho) + \rho f_{t-1} + \epsilon_t \quad (35)$$

with $\epsilon_t \sim Niid(0, \sigma^2)$. The parameters in (35) are set so that f_t has the same mean (0.0156), standard deviation (0.0147) and serial correlation coefficient (0.292) as the consumption growth series in the LV database. For each sample the factor pricing model is estimated using GMM and each of the two normalizations, $m_t^* = 1 - f_t b^*$ and $m_t^\diamond = \xi[1 - (f_t - \mu)b^\diamond]$. The results are presented in Table 9.

The two normalizations perform very differently in finite samples. As Table 9 indicates, the median value of the R^2 measure of fit is just over 0.7 at the first and second stages of GMM for the a -normalization. In contrast, for the ξ -normalization the median value of the R^2 is less than 0.07 at both stages of GMM. The distribution of \hat{b}^* is almost entirely to the right of 35 at the second GMM stage, and \hat{b}^* is found to be significant at the 5 percent level in 99.9 percent of the samples, indicating that the researcher using the a -normalization would almost always conclude that the spurious factor prices the portfolios. On the other hand, a researcher using the ξ -normalization would find \hat{b}^\diamond centered around 0, and would find it significant at the 5 percent level in only 25 percent of the samples. So he would usually conclude that the spurious factor was unhelpful in pricing the assets. Using the a -normalization the over-identifying restrictions are rejected only 9 percent of the time with size set to 5 percent, while rejections occur in 81 percent of the samples using the ξ -normalization.

There is less contrast between the results for the factor price of risk. For the a -normalization, the distribution of $\hat{\lambda}_f^*$ is mainly to the right of zero, but has very fat tails to the right and left. Using a two-sided test, $\hat{\lambda}_f^*$ is statistically significant at the 5 percent level in 32 percent of the samples, at the second stage of GMM. For the ξ -normalization $\hat{\lambda}_f^\diamond$ is centered around 0 and is only significant in 24 percent of the samples.

The simulations with spurious factors can rationalize our findings with US consumption growth used as the risk factor. As Table 1 indicates, when using the a -normalization we found $\hat{b}^* = 48.6$ and $\hat{\lambda}_f^* = 4.20$ at the first stage of GMM, and $\hat{b}^* = 48.8$ and $\hat{\lambda}_f^* = 4.28$ at the second stage of GMM. These values lie at the 62nd, 57th, 50th and 58th percentiles of the analog empirical distributions in the Monte Carlo experiments. Table 1 reports $R_*^2 = 0.87$ at both stages of GMM, and $J = 5.1$. These values lie, respectively, at the 82nd, 84th and 17th percentiles of the analog empirical distributions in the Monte Carlo experiments. In Table 1, the t -statistics for \hat{b}^* and $\hat{\lambda}_f^*$ are, respectively, 4.67 and 1.05 at the first stage of GMM and 11.3 and 2.61 at the second stage of GMM. The first stage values are at the 63rd and 43rd percentiles of the analog empirical distributions in the Monte Carlo experiments, while the second stage values lie at the 98th and 90th percentiles of their respective empirical distributions. It is only these last two statistics that seem at odds with the notion that consumption growth is a purely spurious factor. This may be due to the simulations not capturing specific features of the US data that matter in constructing the weighting matrix for the second stage of GMM.

A researcher using the a -normalization would obviously improve his chances of rejecting a proposed SDF based on a spurious factor if he tested for the statistical significance of $\hat{\lambda}_f^*$. As Table 9 indicates, however, he would still find $\hat{\lambda}_f^*$ to be significant at the 5 percent level in 32 percent of the repeated samples at the second stage of GMM. He could also test for the statistical significance of Em^* , but as Table 9 indicates, it would be significant at the 5 percent level in 46 percent of the repeated samples at the second stage of GMM. The two statistics are highly correlated with each other, so a joint test would be no more effective.

Testing for the significance of $\hat{\lambda}_f^*$ at the first stage of GMM would be effective, because $\hat{\lambda}_f^*$ is only significant at the 5 percent level in 8 percent of the repeated samples. While this suggests that an effective way of testing for a spurious factor is to test for the significance of $\hat{\lambda}_f^*$, this is only true for single factor models. For multi-factor models, as we will see, rejecting a model only on the basis of the statistical significance of $\hat{\lambda}_f^*$ would lead to over-rejection of the true model.

The most effective test for detecting spurious factors appears to be the direct test for the rank of the β_f matrix. Table 9 shows results from testing the null hypothesis that the rank of β_f is zero. The L statistic exceeds its 5 percent critical value in only 21 percent of the samples. As a result, the test would have a 79 percent success rate in detecting failure of

the rank condition if the size of the test was set at 5 percent.

6.1.2 An Experiment Based on the Fama-French Data

When using the FF25 data, above, I set the sample period to 1949Q1–2005Q4. Accordingly, I create 10000 samples, each of length $T = 228$, for an artificial risk factor f_t . This experiment is more complicated than the previous one, however, because the random factor is designed to be panel spurious so that it mimics the properties of US consumption growth. In each random sample I set $f_t = \mu + c(R_t^e - \bar{R}^e)' \hat{\Sigma}_R^{-1} \iota + \epsilon_t$ where \bar{R}^e and $\hat{\Sigma}_R$ are the sample mean and covariance matrix of the vector of excess returns in the FF25 data, and ϵ_t is $\epsilon_t \sim Niid(0, \sigma^2)$. Across repeated samples $\text{cov}(f_t, R_t) = c\iota$, by construction. I set $\mu = 0.0051$ to match the mean of consumption growth in the US data, $c = 7.49 \times 10^{-5}$ to match the cross-sectional average of the covariance between consumption growth and the FF25 returns, and $\sigma = 0.005$ so that the standard deviation of f_t equals the standard deviation of consumption growth in the data. I do not attempt to replicate the serial correlation properties of US consumption growth in the randomly generated factor. For each sample I estimate the model using both the a and ξ normalizations.

Recall that when the “Consumption growth” model was estimated with actual US data we found \hat{b}^* , $\hat{\lambda}^*$, \hat{b}^\diamond and $\hat{\lambda}^\diamond$ to be statistically significant (see Tables 3 and 4), though only marginally so for the ξ -normalization. The main differences across normalizations were that the R^2 was large and positive for the a -normalization, but negative for the ξ -normalization, at both stages of GMM. The model passed the test of the over-identifying restrictions in both cases but only marginally for the ξ -normalization.

A similar pattern emerges when the model is estimated using panel spurious factors in repeated samples. The results are presented in Table 10. It indicates that \hat{b}^* and $\hat{\lambda}_f^*$ are statistically significant in nearly 100 percent of the repeated samples at the second stage of GMM. In contrast, for the ξ -normalization, \hat{b}^\diamond and $\hat{\lambda}_f^\diamond$ are statistically significant in a little over 60 percent of the repeated samples. Here only the R^2 measure of fit for the ξ -normalization provides a strong test of the model’s ability to explain the returns. While R_*^2 exceeds 0.64 in 50 percent of the samples, R_\diamond^2 is negative in 98 percent of the samples at the second stage of GMM.

6.2 Experiments with an Artificial Model

6.2.1 A Single Factor Model

In this set of experiments I generate artificial data from a single-factor model. The SDF is given by $m_t = a - f_t b$ where $a = 1.42$, $b = 27.7$ and f_t follows the law of motion $f_t \sim Niid(\mu, \sigma^2)$, with $\mu = 0.0156$ and $\sigma = 0.0147$. This ensures that f_t has the same mean and standard deviation as consumption growth in the LV database. Given that the true value of μ is 0.0156 this means that $b^* = 19.5$, $b^\diamond = 28$, $\lambda_f = 0.68$ (expressed in percent) and $Em^* = 0.70$.

I construct an $n \times 1$ (with $n = 8$) vector of artificial excess returns with the law of motion $R_{it}^e = \mu_{Ri} + \eta(m_t - \mu_m) + \Psi \xi_t$, where μ_{Ri} is an $n \times 1$ vector, $\mu_m = a - \mu b$, η is an $n \times 1$ vector, Ψ is an $n \times n$ diagonal matrix, and $\xi_t \sim Niid(0, I_n)$ and is independent of f_t . Given this definition for R_{it}^e , it follows that the variance of R_{it}^e is $\sigma_{Ri}^2 = \eta_i^2 \sigma_m^2 + \psi_i^2$ with $\sigma_m^2 = \sigma^2 b^2$. Let $\rho_i = -\mu_m \mu_{Ri} / (\sigma_m \sigma_{Ri})$, for $i = 1, \dots, n$. Setting $\eta_i = \rho_i \sigma_{Ri} / \sigma_m$ implies that the i th element on the diagonal of Φ is $\psi_i = (1 - \rho_i^2)^{1/2} \sigma_{Ri}$ and that $E(R_{it}^e m_t) = 0$. So that the model returns share some of the characteristics of the LV returns, I set μ_{Ri} and σ_{Ri}^2 equal to their sample equivalents in the LV data, and thereby determine the values of ρ_i , η_i and ψ_i , for $i = 1, \dots, n$.

In these experiments the SDF, m_t , prices the returns. In large samples GMM estimators based on the two normalizations deliver consistent parameter estimates and lead to correct inference about the model. To check small sample performance I simulate 10000 samples of 50, 250 and 1000 observations each from the model, and estimate the model using both normalizations.

The results of the experiments with the true risk factor, f_t , and the a -normalization are summarized in Table 11. The distributions of the parameter estimates are centered near their true values, and in 96 percent of the samples of 50 observations the estimates of b^* are found to be statistically significant at the 5 percent level (90 percent for λ_f^*) at the second stage of GMM. These percentages rise to 100 percent in the larger samples. There is a slight tendency to under-reject the over-identifying restrictions, but size approaches asymptotic size as T increases. The median R^2 in a sample of 50 observations is 0.83, and this rises to 0.96 in samples of 250 observations, and 0.98 in samples of 1000 observations.

The results of the experiments with the true risk factor, f_t , and the ξ -normalization are summarized in Table 12. Once again, the distributions of the parameter estimates are

centered near their true values, and in 73 percent of the samples of 50 observations the estimates of b^\diamond are found to be statistically significant at the 5 percent level (75 percent for λ_f^\diamond) at the second stage of GMM. These percentages rise to 100 percent in the larger samples. Here there is a slight tendency to over-reject the over-identifying restrictions, but size approaches asymptotic size as T increases. The median R^2 in a sample of 50 observations is 0.66, and this rises to 0.92 in samples of 250 observations, and 0.98 in samples of 1000 observations. Overall, the ξ -normalization has a tendency to be less enthusiastic about the true model than the a -normalization, but both normalizations perform well in sufficiently large samples.

To examine the performance of the GMM estimators when the rank condition fails, I generate a globally spurious factor $x_t \sim Niid(\mu, \sigma^2)$, which is orthogonal to R_t^e and f_t , and estimate the model with x_t replacing f_t . The results of the experiments with the spurious risk factor, x_t , are summarized in Table 13 for the a -normalization. Consider, first, samples of 50 observations. With the a -normalization the model is rejected at the 5 percent level in only 2 percent of the samples and the median values of R_*^2 are 0.79 and 0.81 at the two stages of GMM. The estimates of \hat{b}^* are statistically significant at the 5 percent level in 93 and 96 percent of the samples. Thus, a researcher using only the a -normalization would almost always conclude that the spurious factor helps to price the assets. The factor risk premium, $\hat{\lambda}_f^*$, is significant less often, but a researcher would not necessarily reject the model on this basis.⁸ As we will see later, rejecting a model based on lack of significance of $\hat{\lambda}_f^*$ can lead to a lot of Type I errors. The predictions of Theorem 4 are borne out in the simulations. As the sample size increases the distribution of \hat{b}^* tightens and is nearly centered around $1/E(f) = 64.2$. The J statistic rejects the model based on the spurious factor about as often as it would if the model were true. But, as Theorem 7 predicts, $\hat{\lambda}_f^*$ is significant at the 5 percent level in roughly 5 percent of the samples.

In Section 5 I proposed a simple diagnostic test based on an estimate of Em^* given by $1 - \hat{\mu}'\hat{b}^*$. Suppose the researcher rejects the model whenever the estimate of Em^* is not significantly different from zero at the 5 percent level. The results in Table 13 suggest that he would reject the false model 72.7, 89.2, and 94.1 percent of the time in samples of 50, 250 and 1000 observations based on the first stage of GMM. This is a reasonable degree of power. The results in Table 11 suggest that he would make no Type I errors when the model

⁸The standard deviation of $\hat{\lambda}_f^*$ is very large, as Table 12 indicates, but this is due to a small number of extreme observations in the tails of its distribution.

is true.

Turning to the ξ -normalization (Table 14), we can see that in samples of 50 observations the researcher would rarely be misled into thinking that the model based on the spurious factor fits the data. This would not, however, be the result of rejecting the model on the basis of the J statistic. This would only happen 18.3 percent of the time for a test with 5 percent asymptotic size. Here \hat{b}^\diamond and $\hat{\lambda}_f^\diamond$ are rarely statistically significant, and the estimates are centered approximately around zero. Also, the R^2 measure of fit takes on a median value of just 0.10 at the first stage of GMM, and 0.05 at the second stage. Unfortunately, to some extent the ξ -normalization performs less well in large samples. At the first stage of GMM, the frequency with which \hat{b}^\diamond and $\hat{\lambda}_f^\diamond$ are statistically significant at the 5 percent level rises to 50 percent in samples of 1000 observations. For the second stage of GMM, this is only true in about 15 percent of the samples. The distribution of R_\diamond^2 is roughly invariant to the sample size. The test of the over-identifying restrictions becomes more powerful as the sample size increases, although it rejects the model only 56.1 percent of the time in samples of 1000 observations.

The direct test of the rank of β_f proves to be a useful diagnostic. Suppose the researcher rejects the model whenever the test statistic, L , does not exceed the 5 percent critical value of the relevant χ^2 distribution. The results in Table 14 imply that he would reject the false model 65, 90, and 94 percent of the time in samples of 50, 250 and 1000 observations based on the first stage of GMM. This is a reasonable degree of power. The results in Table 12 suggest that he would make no Type I errors when the model is true.

Overall, my results suggest that when the model is true there is little to choose between the normalizations. However, the results with spurious factors suggest that a researcher must interpret his results carefully. The ξ -normalization is less likely to deliver misleadingly favorable inference about the model than the a -normalization, but at the first stage of GMM even the ξ -normalization has low power. The test of $Em^* = 0$, which is based on the a -normalization, and the direct test based on the rank of β_f , both appear to be useful diagnostics for failure of the rank condition, in that they have a reasonable degree of power, and lead to no Type I errors in my simulations.

As a final check on the Monte Carlo simulations with spurious factors, Figures 5 and 6 illustrate the frequency distributions of \hat{b}^* , \hat{b}^\diamond , their respective t -statistics $t(\hat{b}^*)$ and $t(\hat{b}^\diamond)$, R_\star^2 , R_\diamond^2 , J_\star and J_\diamond . For the a -normalization, the distribution of the t -statistic diverges so it is

scaled by a factor of $T^{-1/2}$. For the ξ -normalization, the distribution of \hat{b}^\diamond diverges, so it is scaled by a factor of $T^{-1/2}$. For both normalizations, the distributions across the Monte Carlos experiments converge to their respective probability limits or limiting distributions as $T \rightarrow \infty$.

6.2.2 A Multi-Factor Model

In this set of experiments I generate artificial data from a multi-factor model. The SDF is given by $m_t = a - f_t' b$ where a is a scalar, f_t and b are $k \times 1$ vectors, and f_t follows the law of motion $f_t \sim Niid(\mu, \Sigma_f)$. I set $k = 3$, $a = 1.166$ and $b = (4.50 \quad -0.12 \quad 6.78)'$. I set μ and Σ_f equal to the sample mean and covariance matrix of the *Rm-Rf*, *SMB* and *HML* factors from the Fama-French database. It follows that $b^* = (3.86 \quad -0.10 \quad 5.82)'$, $b^\diamond = (4.51 \quad -0.12 \quad 6.80)'$ and $\lambda_f = (1.92 \quad 0.53 \quad 1.32)'$, expressed in percent. The model for the true SDF mimics the first stage GMM estimates for the Fama-French 3-factor model shown in Table 4.

I generate an $n \times 1$ (with $n = 25$) vector of artificial excess returns $R_t^e = \mu_R + \beta(f_t - \mu) + \Psi \xi_t$ where μ_R is an $n \times 1$ vector, β is an $n \times k$ matrix, Ψ is an $n \times n$ lower triangular matrix, and $\xi_t \sim Niid(0, I_n)$ and is independent of f_t . Given this definition for R_t^e , it follows that the covariance matrix of R_t^e is $\Sigma_R = \beta \Sigma_f \beta' + \Psi \Psi'$. So that the model shares some characteristics with actual data, I set Σ_R equal to its sample equivalent in the FF25 data. I set β equal to the matrix of factor betas for the FF25 returns regressed on *Rm-Rf*, *SMB* and *HML*. I set Ψ equal to the Cholesky decomposition of the covariance matrix of the residuals from those regressions. From the assumptions above we have

$$\begin{aligned} E(R_t^e m_t) &= E\{[\mu_R + \beta(f_t - \mu) + \Psi \xi_t] (a - f_t' b)\} \\ &= (a - \mu' b) [\mu_R - \beta \Sigma_f b / (a - \mu' b)]. \end{aligned} \tag{36}$$

To ensure that the model satisfies the asset pricing restrictions I set $\mu_R = \beta \Sigma_f b / (a - \mu' b)$. This means that the model expected returns correspond to the model-predicted expected returns for the first stage Fama-French model shown in Table 4.

In these experiments the SDF, m_t , prices the returns. In large samples GMM estimators based on the two normalizations deliver consistent parameter estimates and lead to correct inference about the model. To check small sample performance I simulate 10000 samples of 228 observations (the size of quarterly US sample used early) each from the model, and estimate the model using both normalizations.

I perform four sets of experiments. In the first experiment, the true model is tested by making the proposed SDF a linear function of f_t . In the second experiment, a false model with a single relevant factor is tested. Here the proposed SDF is a linear function of just f_{1t} , the first element of f_t . In the third experiment, a false model with a single relevant, but panel spurious, factor is tested. In the fourth experiment, a false multi-factor model is tested. Here the proposed SDF is a linear function of a panel relevant factor, a panel spurious factor, and a purely spurious factor.

Results for the True Model Table 15 shows the results from the first set of experiments with the true model. The results highlight an important fact: even in a sample of 228 observations, the estimated parameters of the true SDF can be statistically significant in a relatively small fraction of the samples. The coefficient on the second factor, which mimics the *SMB* factor, is only significant in around 10 percent of the samples. The factor risk premium, $\hat{\lambda}_f$, for the second factor is also only significant in about 30 to 50 percent of the samples. For the α -normalization, the test of the over-identifying restrictions has size less than asymptotic size, while for the ξ -normalization the two are approximately equal.

Results for a Single Relevant Factor Table 16 shows results for the second set of experiments where a model based on a single relevant factor is estimated. The single factor is $x_{1t} = f_{1t}$, the first factor in the true model. Not surprisingly, the estimates of b^* , b^\diamond , λ_f^* and λ_f^\diamond are statistically significant in nearly every sample. For both normalizations, the low values of the R^2 and the large values of the J -statistic both act as reliable signals that there is a missing factor.

Results for a Single Panel-Spurious Factor Table 17 shows results for the third set of experiments with a panel spurious factor. Here the proposed SDF is a linear function of a factor,

$$x_{2t} = \mu_{x_2} + c\iota'\Sigma_R^{-1}(R_t^e - \mu_R) + u_{2t},$$

where c is a scalar, ι is an $n \times 1$ vector of ones, and $u_{2t} \sim Niid(0, \sigma_{u_2}^2)$ is independent of R_t^e and u_{1t} . I set μ_{x_2} equal to the sample mean of quarterly US consumption growth in the period 1949Q1–2005Q4. I set c equal to the cross-sectional average of the sample covariance between US consumption growth and the FF25 returns over the same period. I set $\sigma_{u_2}^2$ so that the variance of x_{2t} equals the sample variance of US consumption growth over the same

period. As Table 17 indicates, the a -normalization provides very misleading inference about the model. The estimates of b^* , λ_f and Em^* are significant in almost every sample at the second stage of GMM. This is not unexpected. After all, x_{2t} is a relevant factor that is correlated with R_t^e . But the model also has good fit in many samples, with the median R^2 being 0.79 and 0.58 at the first and second stages of GMM, and the J -statistic leading to very few rejections of the over-identifying restrictions.

A very different picture emerges when the model is estimated using the ξ -normalization. Once again the estimates of the model parameters, in this case b^\diamond and λ_f , are often statistically significant due to the relevance of the risk factor. The model is also rarely rejected on the basis of the test of the over-identifying restrictions. But here, the low R^2 values reliably signal that the model has poor fit.

Results for a False Multi-Factor SDF Table 18 shows results for the fourth set of experiments with a false proposed SDF that is a linear function of a relevant factor, a panel spurious factor, and a third, purely spurious factor. In this way, this experiment is meant to mimic the “Yogo Factors” case in Tables 3 and 4. As we saw in section 4, the market premium factor appears to be relevant, but the evidence from the factor betas suggests that consumption growth is panel spurious for the FF25 returns, and durables growth is purely spurious.

The trio of factors, x_t , consists of x_{1t} and x_{2t} (defined above), as well as $x_{3t} = \mu_{x3} + u_{3t}$ where $u_{3t} \sim Niid(0, \sigma_{u3}^2)$ is independent of R_t^e , u_{1t} and u_{2t} . I set σ_{u3}^2 equal to the sample variance of US durables growth over the period 1949Q1–2005Q4.

The results for the a -normalization follow the predictions of Theorem 4. The estimates of b_3^* , the coefficient on the spurious factor, are nearly always found to be statistically significant. In contrast, the estimates of b_1^* , the coefficient on the most relevant factor for the cross-sectional distribution of the returns, is nearly always found to be statistically insignificant. The cross-sectional R^2 measures are also very high and the test of the over-identifying restrictions rarely leads to the model being rejected. Some doubt is cast on the model by the fact that some elements of $\hat{\lambda}_f^*$ are significant in relatively few samples, but as we saw above, this also occurs when the model being tested is the true model. A researcher aware of this fact would be hesitant to reject the model on this basis.

The test for $Em^* = 0$ at the first stage of GMM provides reliable evidence against model, as Em^* is found to be statistically different from zero at the 5 percent level in just 11 percent

of the samples. As Table 15 indicates, using the 5 percent critical value as a threshold for rejection of the model would not lead to Type I errors as the estimate of Em^* is significant in every sample for the true model.

The test of the rank of β_f also provides reliable evidence against model. The statistic L is found to be statistically different from zero at the 5 percent level in just 14 percent of the samples. As Table 15 indicates, using the 5 percent critical value as a threshold for rejection of the model would not lead to Type I errors as L is statistically significant in every sample for the true model.

7 Conclusion

We have seen that standard GMM-based tests based on the a -normalization have very low power to reject proposed SDFs when they are false and the covariance matrix between the returns and the proposed risk factors has less than full column rank. This finding is relevant for tests of consumption-based asset pricing models, because some consumption-related risk factors display little covariance with the returns being priced. Estimates of consumption-based models based on the a -normalization are misleading because parameter estimates are statistically significant and R^2 s are close to 1 despite the lack of significant covariance between the returns and specific risk factors in the models.

The ξ -normalization, which specifies the SDF in terms of demeaned risk factors, generally performs better in small samples, but even then the power to reject a false model can be surprisingly low. In Burnside (2007a) I show that there are related small sample problems with the variants of the ξ -normalization used by Parker and Julliard (2005) and Yogo (2006).

The challenge is to find tests that reliably detect failure of the rank condition. Tests based on the statistical significance of factor risk premia lead to many Type I errors in finite samples. Tests based on the statistical significance of an estimate of Em^* display a reasonable amount of power and lead to very few Type I errors. So do direct tests of the column rank of the matrix of factor betas. Given the problematic nature of the small and large sample properties of GMM when the rank condition fails, these diagnostics are useful tools for researchers working with macroeconomic factor models.

REFERENCES

- Black, Fischer, Michael Jensen and Myron Scholes (1972) “The Capital Asset Pricing Model: Some Empirical Tests,” in Michael Jensen, ed. *Studies in the Theory of Capital Markets*. Praeger: New York.
- Burnside, Craig (2007a) “GMM and the Two-Pass Method in Consumption-Based Asset Pricing: A Critical Review,” mimeo, Duke University.
- Burnside, Craig (2007b) “The Cross-Section of Foreign Currency Risk Premia and Consumption Growth Risk: A Comment,” NBER Working Paper 13129.
- Burnside, Craig and Martin Eichenbaum (1996) “Small Sample Properties of GMM Based Wald Tests,” *Journal of Business and Economic Statistics* 14, 1996, 294–308.
- Chen, Naifu, Richard R. Roll, and Stephen A. Ross (1986) “Economic Forces and the Stock Market,” *Journal of Business* 59, 383–403.
- Cochrane, John H. (2005) *Asset Pricing*, Revised edition. Princeton: Princeton University Press.
- Cragg, John G. and Stephen G. Donald (1997) “Inferring the Rank of a Matrix,” *Journal of Econometrics* 76, 223–50.
- Daniel, Kent and Sheridan Titman (2005) “Testing Factor-Model Explanations of Market Anomalies,” working paper, Northwestern University.
- den Haan, Wouter J. and Andrew T. Levin (2000) “Robust Covariance Matrix Estimation with Data-Dependent VAR Prewhitening Order,” NBER Technical Working Paper No. 255.
- Epstein, Larry G. and Stanley E. Zin (1989) “Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework,” *Econometrica* 57, 937–69.
- Fama, Eugene F. and Kenneth R. French (1993) “Common Risk Factors in the Returns on Stocks and Bonds,” *Journal of Financial Economics* 33, 3–56.
- Fama, Eugene F. and James D. MacBeth (1973) “Risk Return and Equilibrium: Empirical Tests,” *Journal of Political Economy* 71, 607–36.
- Ferson, Wayne E. and Campbell Harvey (1993) “The Risk and Predictability of International Equity Returns,” *Review of Financial Studies* 6, 527–566.
- Hall, Alastair R. (2005) *Generalized Method of Moments*. Oxford: Oxford University Press.
- Hansen, Lars P. (1982) “Large Sample Properties of Generalized Method of Moments Estimators,” *Econometrica* 50, 1029–54.
- Hansen, Lars P. and Ravi Jagannathan (1997) “Assessing Specification Errors in Stochastic Discount Factor Models,” *Journal of Finance* 52, 557–90.

- Kan, Raymond and Cesare Robotti (2006) "Specification Tests of Asset Pricing Models Using Excess Returns," Federal Reserve Bank of Atlanta Working Paper No. 2006-10.
- Kan, Raymond and Chu Zhang (1999a) "GMM Tests of Stochastic Discount Factor Models with Useless Factors," *Journal of Financial Economics* 54, 103–27.
- Kan, Raymond and Chu Zhang (1999b) "Two-Pass Tests of Asset Pricing Models with Useless Factors," *Journal of Finance* 54, 203–35.
- Lewellen, Jonathan and Stefan Nagel (2006) "The Conditional CAPM Does Not Explain Asset-Pricing Anomalies," manuscript, Dartmouth College. Forthcoming, *Journal of Financial Economics*.
- Lewellen, Jonathan, Stefan Nagel and Jay Shanken (2006) "A Skeptical Appraisal of Asset-Pricing Tests," manuscript, Dartmouth College.
- Lustig, Hanno and Adrien Verdelhan (2007) "The Cross-Section of Foreign Currency Risk Premia and Consumption Growth Risk," *American Economic Review*.
- Parker, Jonathan A. and Christian Julliard (2005) "Consumption Risk and the Cross Section of Expected Returns," *Journal of Political Economy* 113, 185–222.
- Staiger, Douglas and James H. Stock and (1997) "Instrumental Variables Regression with Weak Instruments," *Econometrica* 65: 557–86.
- Wright, Jonathan H. (2003) "Detecting Lack of Identification in GMM," *Econometric Theory* 19, 322–30.
- Yogo, Motohiro (2006) "A Consumption-Based Explanation of Expected Stock Returns," *Journal of Finance* 61, 539–80.

TABLE 1: GMM Estimates of Linear Factor Models
Lustig-Verdelhan Data, a -Normalization

Risk factor or factors	First Stage			Second Stage			
	b^*	λ_f	R^2	b^*	λ_f	R^2	J
Consumption growth	48.6 (10.4)	4.20 (4.01)	0.87	48.8 (4.3)	4.28 (1.64)	0.87	5.1 (0.645)
Durables growth	20.2 (3.7)	2.65 (1.63)	0.84	24.3 (1.7)	5.51 (4.41)	0.81	4.8 (0.683)
$Rm-Rf$	5.5 (2.3)	29.1 (18.8)	0.46	6.5 (1.7)	39.1 (17.5)	0.45	10.2 (0.175)
Consumption Factors							
Consumption growth	40.1 (56.0)	3.64 (1.86)	0.87	-5.5 (27.9)	1.86 (1.98)	0.81	2.7 (0.847)
Durables growth	3.7 (28.0)	3.64 (1.95)		26.2 (13.5)	4.81 (3.48)		
Yogo Factors							
Consumption growth	9.6 (30.7)	2.76 (2.06)	0.95	2.3 (21.9)	3.43 (3.41)	0.89	1.9 (0.860)
Durables growth	14.2 (17.1)	2.91 (2.74)		22.0 (11.0)	5.96 (6.21)		
Rm	2.4 (1.7)	33.4 (27.7)		1.1 (0.9)	12.1 (23.2)		
Fama-French Factors							
$Rm-Rf$	6.0 (3.4)	31.3 (30.2)	0.56	7.6 (2.5)	92.8 (91.4)	0.27	5.3 (0.384)
SMB	-5.3 (5.7)	-12.7 (22.1)		-4.5 (4.2)	-17.5 (39.6)		
HML	4.1 (3.9)	12.3 (19.7)		6.5 (2.8)	45.7 (50.7)		

Note: Annual data, 1953–2002. The table reports first and second stage GMM estimates of b^* , from the SDF $m_t = 1 - f_t' b^*$, obtained using the moment restriction $E(R_t^e m_t) = 0$, where R_t^e is an 8×1 vector of excess returns of equally-weighted portfolios of short-term foreign-currency denominated money market securities sorted by their interest differential with the US, and f_t is a scalar or vector of risk factors. The factors are real per household consumption (nondurables & services) growth, real per household durable consumption growth, and the following variables from the Fama-French data: the real value weighted US stock market excess return over the risk free rate ($Rm-Rf$), the gross return to the same portfolio (Rm), and the SMB and HML portfolio excess returns [see Lustig and Verdelhan (2007)]. Estimates of the factor risk premium $\hat{\lambda}_f = S_f \hat{b}^* / (1 - \bar{f}' \hat{b}^*)$ are also reported (in percent), where \bar{f} and S_f are the sample mean and covariance matrix of f_t . GMM-VARHAC standard errors are reported in parentheses for \hat{b}^* and $\hat{\lambda}_f$. The table reports the R^2 measure of fit between the sample mean of R_t^e and the predicted mean returns, given by $D_T \hat{b}^*$, where $D_T = \frac{1}{T} \sum_{t=1}^T R_t^e f_t'$. Tests of the overidentifying restrictions are also reported. The test statistic, J , is asymptotically distributed as a χ_{8-k}^2 , where k is the number of risk factors. The p-value is in parentheses.

TABLE 2: GMM Estimates of Linear Factor Models
Lustig-Verdelhan Data, ξ -Normalization

Risk factor or factors	First Stage				Second Stage			
	μ	b^\diamond	λ_f	R^2	b^\diamond	λ_f	R^2	J
Consumption growth	0.016 (0.003)	45.1 (74.7)	0.95 (1.48)	0.10	72.3 (31.4)	1.52 (0.73)	0.06	15.6 (0.029)
Durables growth	0.034 (0.007)	20.8 (29.2)	0.87 (1.15)	0.16	48.0 (23.4)	2.01 (0.82)	-0.13	14.3 (0.047)
Rm-Rf	0.070 (0.025)	1.8 (3.6)	5.97 (11.8)	0.02	1.4 (2.4)	4.72 (7.76)	0.02	22.0 (0.003)
Consumption Factors								
Consumption growth	0.016 (0.003)	-8.7 (56.9)	0.26 (0.80)	0.16	-10.8 (40.6)	0.74 (0.69)	-0.11	14.1 (0.028)
Durables growth	0.034 (0.007)	23.6 (39.0)	0.82 (1.05)		50.8 (28.6)	1.92 (0.84)		
Yogo Factors								
Consumption growth	0.016 (0.003)	-22.0 (63.6)	0.59 (1.18)	0.34	-4.9 (48.3)	1.26 (1.14)	-0.08	8.2 (0.146)
Durables growth	0.034 (0.007)	45.5 (51.0)	1.10 (1.78)		65.4 (34.7)	2.40 (1.64)		
<i>Rm</i>	0.070 (0.025)	5.2 (3.0)	11.7 (9.42)		3.3 (2.7)	5.02 (7.80)		
Fama-French Factors								
<i>Rm-Rf</i>	0.070 (0.025)	1.5 (4.3)	7.07 (11.4)	0.08	-0.4 (3.5)	1.07 (9.51)	0.05	17.7 (0.003)
<i>SMB</i>	0.024 (0.020)	1.7 (4.6)	4.08 (7.06)		2.2 (3.9)	3.79 (6.04)		
<i>HML</i>	0.057 (0.020)	-2.8 (5.1)	-5.91 (8.78)		-2.5 (4.0)	-4.80 (7.23)		

Note: Annual data, 1953–2002. The table reports first and second stage GMM estimates of μ and b^\diamond , from the SDF $m_t = 1 - (f_t - \mu)'b^\diamond$, obtained using the moment restrictions $E(R_t^e m_t) = 0$, $E(f_t - \mu) = 0$. Since $\hat{\mu}$ is the same for both GMM stages, the estimate is reported once. The variables R_t^e and f_t are defined in the note to Table 1. Estimates of the factor risk premium $\hat{\lambda}_f = S_f \hat{b}^\diamond$ are also reported (in percent), where S_f is the sample covariance matrix of f_t . GMM-VARHAC standard errors are reported in parentheses for $\hat{\mu}$, \hat{b}^\diamond and $\hat{\lambda}_f$. The table reports the R^2 measure of fit between the sample mean of R_t^e and the predicted mean returns, given by $d_T \hat{b}^\diamond$, where $d_T = \frac{1}{T} \sum_{t=1}^T R_t^e (f_t' - \hat{\mu})'$. Tests of the overidentifying restrictions are also reported. The test statistic, J , is asymptotically distributed as a χ_{8-k}^2 , where k is the number of risk factors. The p-value is in parentheses.

TABLE 3: GMM Estimates of Linear Factor Models
Fama-French 25 Data, a -Normalization

Risk factor or factors	First Stage			Second Stage			
	b^*	λ_f	R^2	b^*	λ_f	R^2	J
Consumption growth	126.3 (23.6)	0.94 (0.49)	0.81	142.8 (10.1)	1.38 (0.36)	0.51	30.2 (0.179)
Durables growth	111.1 (18.3)	-2.48 (2.84)	0.89	93.3 (7.2)	11.0 (44.9)	0.43	15.5 (0.906)
$Rm-Rf$	3.3 (0.9)	2.31 (0.57)	-0.55	4.7 (0.8)	3.40 (0.59)	-3.55	65.5 (0.000)
Consumption Factors							
Consumption growth	54.0 (33.8)	2.83 (4.86)	0.98	49.3 (13.8)	1.93 (2.03)	0.96	17.2 (0.800)
Durables growth	64.4 (20.4)	4.24 (8.01)		64.8 (8.3)	3.15 (3.24)		
Yogo Factors							
Consumption growth	41.0 (32.5)	2.10 (3.76)	0.98	38.5 (16.6)	1.47 (1.57)	0.97	17.3 (0.748)
Durables growth	69.9 (18.1)	4.21 (7.19)		68.7 (8.8)	3.02 (2.57)		
Rm	0.2 (0.5)	2.33 (3.42)		0.4 (0.4)	2.96 (3.28)		
Fama-French Factors							
$Rm-Rf$	3.9 (0.9)	1.92 (0.64)	0.75	5.0 (0.8)	2.53 (0.70)	-0.20	47.8 (0.001)
SMB	-0.1 (1.2)	0.52 (0.36)		-0.5 (1.1)	0.58 (0.37)		
HML	5.9 (1.1)	1.35 (0.49)		7.2 (0.9)	1.69 (0.63)		

Note: Quarterly data, 1949–2005. The table reports first and second stage GMM estimates of b^* , from the SDF $m_t = 1 - f_t' b^*$, obtained using the moment restriction $E(R_t^e m_t) = 0$, where R_t^e is a 25×1 vector of excess returns of the Fama-French 25 portfolios of US stocks sorted on size and the book-to-market value ratio, and f_t is a scalar or vector of risk factors. The factors are real per capita consumption (nondurables & services) growth, real per capita durable consumption growth, and the following variables from the Fama-French data: the real value weighted US stock market excess return over the risk free rate ($Rm-Rf$), the gross return to the same portfolio (Rm), and the SMB and HML portfolio excess returns (see Appendix B). Estimates of the factor risk premium $\hat{\lambda}_f = S_f \hat{b}^* / (1 - \bar{f}' \hat{b}^*)$ are also reported (in percent), where \bar{f} and S_f are the sample mean and covariance matrix of f_t . GMM-VARHAC standard errors are reported in parentheses for \hat{b}^* and $\hat{\lambda}_f$. The table reports the R^2 measure of fit between the sample mean of R_t^e and the predicted mean returns, given by $D_T \hat{b}^*$, where $D_T = \frac{1}{T} \sum_{t=1}^T R_t^e f_t'$. Tests of the overidentifying restrictions are also reported. The test statistic, J , is asymptotically distributed as a χ_{8-k}^2 , where k is the number of risk factors. The p-value is in parentheses.

TABLE 4: GMM Estimates of Linear Factor Models
Fama-French 25 Data, ξ -Normalization

Risk factor or factors	First Stage				Second Stage			
	μ	b^\diamond	λ_f	R^2	b^\diamond	λ_f	R^2	J
Consumption growth	0.0051 (0.0005)	335.4 (164.5)	0.89 (0.46)	-0.44	141.9 (70.7)	0.38 (0.17)	-5.93	36.3 (0.051)
Durables growth	0.0104 (0.0012)	-554.0 (534.1)	-1.92 (1.89)	-2.46	-2.2 (121.5)	-0.01 (0.42)	-16.8	2.9 (1.000)
Rm-Rf	0.0197 (0.0054)	3.5 (1.1)	2.30 (0.57)	-0.77	3.1 (1.0)	2.04 (0.54)	-0.97	69.8 (0.000)
Consumption Factors								
Consumption growth	0.0051 (0.0005)	370.3 (166.6)	0.99 (0.46)	-0.42	130.2 (82.3)	0.35 (0.20)	-6.83	28.2 (0.210)
Durables growth	0.0104 (0.0012)	64.9 (165.2)	0.31 (0.56)		6.6 (73.8)	0.05 (0.25)		
Yogo Factors								
Consumption growth	0.0051 (0.0005)	271.6 (177.6)	0.76 (0.45)	-0.33	24.7 (83.0)	0.09 (0.21)	-1.61	20.2 (0.571)
Durables growth	0.0104 (0.0012)	136.4 (136.2)	0.53 (0.48)		40.2 (63.4)	0.14 (0.23)		
Rm	0.0224 (0.0054)	1.5 (2.4)	2.18 (0.73)		2.7 (1.8)	1.76 (0.88)		
Fama-French Factors								
Rm-Rf	0.0197 (0.0054)	4.5 (1.2)	1.92 (0.63)	0.66	4.4 (1.1)	1.75 (0.63)	0.49	53.7 (0.000)
SMB	0.0063 (0.0036)	-0.1 (1.4)	0.53 (0.36)		-0.6 (1.3)	0.36 (0.35)		
HML	0.0119 (0.0036)	6.8 (1.4)	1.32 (0.49)		6.9 (1.3)	1.37 (0.50)		

Note: Quarterly data, 1949–2005. The table reports first and second stage GMM estimates of μ and b^\diamond , from the SDF $m_t = 1 - (f_t - \mu)'b^\diamond$, obtained using the moment restrictions $E(R_t^e m_t) = 0$, $E(f_t - \mu) = 0$. Since $\hat{\mu}$ is the same for both GMM stages, the estimate is reported once. The variables R_t^e and f_t are defined in the note to Table 3. Estimates of the factor risk premium $\hat{\lambda}_f = S_f \hat{b}^\diamond$ are also reported (in percent), where S_f is the sample covariance matrix of f_t . GMM-VARHAC standard errors are reported in parentheses for $\hat{\mu}$, \hat{b}^\diamond and $\hat{\lambda}_f$. The table reports the R^2 measure of fit between the sample mean of R_t^e and the predicted mean returns, given by $d_T \hat{b}^\diamond$, where $d_T = \frac{1}{T} \sum_{t=1}^T R_t^e (f_t' - \hat{\mu})'$. Tests of the overidentifying restrictions are also reported. The test statistic, J , is asymptotically distributed as a χ_{8-k}^2 , where k is the number of risk factors. The p-value is in parentheses.

TABLE 5: Diagnostics for Spurious Factors, Lustig-Verdelhan Data

Portfolio	(a) Covariance with Factor			(b) Factor Beta		
	Consumption Growth	Durables Growth	Market Return	Consumption Growth	Durables Growth	Market Return
1	0.22 (1.66)	1.00 (3.01)	-16.3 (15.0)	0.10 (0.53)	0.24 (0.43)	-0.05 (0.05)
2	1.60 (0.88)	2.04 (1.40)	-7.74 (19.0)	0.76 (0.38)	0.49 (0.31)	-0.02 (0.06)
3	0.55 (1.61)	2.66 (2.79)	1.62 (11.8)	0.26 (0.82)	0.64 (0.53)	0.00 (0.05)
4	0.38 (2.31)	3.72 (3.48)	-29.8 (20.9)	0.18 (1.09)	0.89 (0.54)	-0.09 (0.06)
5	1.34 (1.43)	2.30 (2.31)	4.14 (20.7)	0.63 (0.65)	0.55 (0.54)	0.01 (0.06)
6	0.55 (1.75)	2.90 (2.53)	2.70 (20.2)	0.26 (0.83)	0.69 (0.57)	0.01 (0.06)
7	2.32 (1.72)	5.42 (2.81)	-12.4 (21.1)	1.10 (0.72)	1.30 (0.55)	-0.04 (0.06)
8	0.18 (2.50)	2.82 (2.70)	17.4 (33.1)	0.09 (1.43)	0.68 (1.05)	0.05 (0.10)
Joint-tests						
All = 0	7.87 (0.447) [0.618]	5.69 (0.682) [0.791]	7.52 (0.481) [0.663]	15.6 (0.048) [0.334]	10.0 (0.265) [0.603]	7.6 (0.475) [0.771]
All = constant	5.40 (0.611) [0.731]	2.82 (0.901) [0.932]	7.16 (0.412) [0.572]	11.92 (0.103) [0.369]	7.49 (0.380) [0.629]	7.56 (0.373) [0.664]

Note: Annual data, 1953–2002. The portfolios are equally-weighted groups of short-term foreign-currency denominated money market securities sorted according to their interest differential with the US ($i^* - i$). The risk factors are real per household consumption (nondurables & services) growth, real per household durable consumption growth, and the real value weighted US stock market return from the Fama-French data [see Lustig and Verdelhan (2007)]. Part (a) reports the sample covariance between the excess return to portfolio i and each risk factor. Part (b) reports the slope coefficient from an OLS regression of the excess return to portfolio i on a constant and each risk factor. GMM (VARHAC) standard errors are reported in parentheses. The bottom of the table reports χ^2 test statistics for two tests. The first test is of the hypothesis that the covariances (or betas) with the factor are equal to 0 for all portfolios. The second tests is of the hypothesis that the covariances (or betas) with the factor equal a common value. GMM (VARHAC)-based p-values are reported in parentheses. Small sample p-values computed using a Monte Carlo experiment described in the text are reported in square brackets.

TABLE 6: Diagnostics for Spurious Factors, Fama-French 25 Data

Portfolio	(a) Covariance with Factor			(b) Factor Beta		
	Consumption Growth	Durables Growth	Market Return	Consumption Growth	Durables Growth	Market Return
1,1	1.10 (0.49)	-0.62 (0.61)	100.98 (12.18)	4.14 (1.98)	-1.78 (1.70)	1.55 (0.08)
1,2	1.30 (0.42)	-0.68 (0.66)	88.22 (10.70)	4.91 (1.69)	-1.95 (1.43)	1.36 (0.07)
1,3	0.80 (0.37)	-0.56 (0.63)	74.91 (9.39)	3.03 (1.53)	-1.62 (1.35)	1.15 (0.06)
1,4	0.89 (0.36)	-0.45 (0.64)	71.47 (9.09)	3.36 (1.49)	-1.31 (1.25)	1.10 (0.06)
1,5	0.97 (0.42)	-0.49 (0.72)	75.94 (9.88)	3.68 (1.72)	-1.41 (1.45)	1.17 (0.07)
2,1	0.80 (0.42)	-0.58 (0.64)	95.70 (11.11)	3.01 (1.72)	-1.67 (1.55)	1.47 (0.06)
2,2	0.78 (0.37)	-0.56 (0.60)	80.56 (9.68)	2.96 (1.49)	-1.61 (1.27)	1.24 (0.05)
2,3	0.80 (0.32)	-0.30 (0.59)	70.88 (8.71)	3.04 (1.33)	-0.87 (1.13)	1.09 (0.05)
2,4	0.73 (0.33)	-0.59 (0.61)	68.40 (8.46)	2.77 (1.38)	-1.71 (1.12)	1.05 (0.06)
2,5	0.91 (0.34)	-0.39 (0.72)	72.69 (9.18)	3.44 (1.66)	-1.13 (1.47)	1.12 (0.07)
3,1	0.73 (0.37)	-0.52 (0.50)	88.47 (10.18)	2.75 (1.51)	-1.50 (1.31)	1.36 (0.04)
3,2	0.69 (0.31)	-0.62 (0.57)	73.09 (8.51)	2.60 (1.26)	-1.78 (1.05)	1.12 (0.04)
3,3	0.71 (0.31)	-0.44 (0.54)	65.65 (7.96)	2.67 (1.26)	-1.28 (1.00)	1.01 (0.05)
3,4	0.71 (0.31)	-0.41 (0.61)	64.44 (7.78)	2.67 (1.31)	-1.19 (1.07)	0.99 (0.05)
3,5	0.80 (0.32)	-0.14 (0.73)	67.65 (8.62)	3.04 (1.55)	-0.41 (1.41)	1.04 (0.07)

Table continues on following page

Note: Quarterly data, 1949–2005. The portfolios are the Fama-French 25 portfolios of US stocks sorted on size and the book-to-market value ratio. The risk factors are real per capita consumption (nondurables & services) growth, real per capita durable consumption growth, and the real value weighted US stock market return from the Fama-French data (see Appendix B for more details). Part (a) reports the sample covariance between the excess return to portfolio i and each risk factor. Part (b) reports the slope coefficient from an OLS regression of the excess return to portfolio i on a constant and each risk factor. GMM (VARHAC) standard errors are reported in parentheses.

TABLE 6 (continued): **Diagnostics for Spurious Factors, Fama-French 25 Data**

Portfolio	(a) Covariance with Factor			(b) Factor Beta		
	Consumption Growth	Durables Growth	Market Return	Consumption Growth	Durables Growth	Market Return
4,1	0.58 (0.35)	-0.36 (0.55)	82.24 (9.40)	2.20 (1.38)	-1.04 (1.18)	1.26 (0.04)
4,2	0.56 (0.28)	-0.50 (0.49)	69.14 (8.17)	2.12 (1.16)	-1.43 (1.03)	1.06 (0.04)
4,3	0.53 (0.29)	-0.37 (0.59)	64.00 (7.47)	2.02 (1.19)	-1.07 (1.01)	0.98 (0.04)
4,4	0.64 (0.31)	-0.29 (0.55)	62.69 (7.22)	2.43 (1.27)	-0.83 (1.04)	0.96 (0.05)
4,5	0.91 (0.34)	-0.22 (0.71)	68.85 (8.19)	3.43 (1.59)	-0.62 (1.33)	1.06 (0.06)
5,1	0.63 (0.26)	-0.22 (0.48)	66.63 (8.16)	2.39 (1.05)	-0.63 (0.95)	1.02 (0.03)
5,2	0.43 (0.24)	-0.31 (0.44)	59.36 (6.83)	1.62 (0.97)	-0.88 (0.87)	0.91 (0.03)
5,3	0.44 (0.22)	-0.12 (0.48)	51.60 (6.07)	1.65 (0.93)	-0.34 (0.83)	0.79 (0.03)
5,4	0.50 (0.26)	-0.09 (0.49)	52.94 (6.14)	1.88 (1.07)	-0.27 (1.13)	0.81 (0.04)
5,5	0.78 (0.29)	-0.35 (0.58)	57.95 (6.86)	2.94 (1.23)	-1.00 (1.17)	0.89 (0.05)
Joint-tests						
All = 0	35.2 (0.085) [0.179]	12.3 (0.984) [0.972]	108 (0.000) [0.000]	52.3 (0.001) [0.084]	33.7 (0.115) [0.461]	24079 (0.000) [0.000]
All = constant	27.3 (0.290) [0.436]	11.7 (0.983) [0.963]	66 (0.000) [0.000]	43.4 (0.009) [0.172]	33.4 (0.096) [0.402]	193 (0.000) [0.000]

Note: Quarterly data, 1949–2005. The portfolios are the Fama-French 25 portfolios of US stocks sorted on size and the book-to-market value ratio. The risk factors are real per capita consumption (nondurables & services) growth, real per capita durable consumption growth, and the real value weighted US stock market return from the Fama-French data (see Appendix B for more details). Part (a) reports the sample covariance between the excess return to portfolio i and each risk factor. Part (b) reports the slope coefficient from an OLS regression of the excess return to portfolio i on a constant and each risk factor. GMM (VARHAC) standard errors are reported in parentheses. The bottom of the table reports χ^2 test statistics for two tests. The first test is of the hypothesis that the covariances (or betas) with the factor are equal to 0 for all portfolios. The second tests is of the hypothesis that the covariances (or betas) with the factor equal a common value. GMM (VARHAC)-based p-values are reported in parentheses. Small sample p-values computed using a Monte Carlo experiment described in the text are reported in square brackets.

TABLE 7: Means of Estimated SDFs

	(a) Lustig-Verdelhan Data		(b) Fama-French 25 Data	
	First Stage	Second Stage	First Stage	Second Stage
Consumption growth	0.243 (0.221)	0.240 (0.132)	0.357 (0.115)	0.274 (0.067)
Durables growth	0.319 (0.188)	0.184 (0.163)	-0.156 (0.192)	0.029 (0.116)
<i>Rm-Rf</i>	0.617 (0.181)	0.546 (0.169)	0.935 (0.034)	0.907 (0.040)
Consumption Factors	0.251 (0.161)	0.205 (0.162)	0.056 (0.100)	0.075 (0.077)
Yogo Factors	0.204 (0.139)	0.148 (0.156)	0.060 (0.105)	0.081 (0.082)
Fama-French Factors	0.477 (0.271)	0.213 (0.184)	0.854 (0.040)	0.820 (0.043)

Note: Part (a) of the table displays the sample means of the estimated SDFs corresponding to the parameter estimates in Table 1. Part (b) of the table displays the sample means of the estimated SDFs corresponding to the parameter estimates in Table 3. GMM (VARHAC) standard errors are reported in parentheses.

TABLE 8: Tests of the Rank Condition

	LV Data	FF25 Data
Consumption growth	15.6 (0.048)	52.3 (0.001)
Durables growth	10.0 (0.265)	33.7 (0.115)
<i>Rm-Rf</i>	7.58 (0.475)	24079 (0.000)
Consumption Factors	8.17 (0.318)	23.6 (0.487)
Yogo Factors	4.26 (0.642)	12.9 (0.954)
Fama-French Factors	3.03 (0.805)	7102 (0.000)

Note: This table presents test statistics pertaining to the rank of $\text{cov}(R_t^e, f_t)$. P-values for the statistics are in parentheses. Under the null hypothesis that $\text{rank}[\text{cov}(R_t^e, f_t)] = k - 1$, where k is the number of factors in the candidate SDF, the statistic shown is distributed as a $\chi_{n-(k-1)}^2$ where n is the dimension of R_t^e . For the LV data, R_t^e is an 8×1 vector of excess returns of equally-weighted portfolios of short-term foreign-currency denominated money market securities sorted by their interest differential with the US, and f_t is a scalar or vector of risk factors. The factors are real per household consumption (nondurables & services) growth, real per household durable consumption growth, and the following variables from the Fama-French data: the real value weighted US stock market excess return over the risk free rate (*Rm-Rf*), the gross return to the same portfolio (*Rm*), and the *SMB* and *HML* portfolio excess returns. For the FF data, R_t^e is a 25×1 vector of excess returns of the Fama-French 25 portfolios of US stocks sorted on size and the book-to-market value ratio, and f_t is a scalar or vector of risk factors. The factors are real per capita consumption (nondurables & services) growth, real per capita durable consumption growth, and the following variables from the Fama-French data: the real value weighted US stock market excess return over the risk free rate (*Rm-Rf*), the gross return to the same portfolio (*Rm*), and the *SMB* and *HML* portfolio excess returns.

TABLE 9: Monte-Carlo Experiments with the Lustig-Verdelhan Returns

	GMM Stage	Percentiles					Mean	Std. Dev.	Percent Significant at	
		5	10	50	90	95			10% level	5% level
<i>a</i> -Normalization										
b^*	1	25.1	29.9	44.8	62.1	68.2	45.5	13.2	96.3	93.7
λ_f	1	-7.64	0.83	2.77	10.18	18.00	1.68	213	18.9	8.2
Em^*	1	-0.01	0.06	0.29	0.56	0.64	0.30	0.20	52.0	40.8
R^2	1	0.28	0.40	0.73	0.90	0.93	0.69	0.20	–	–
b^*	2	36.2	38.4	48.7	64.0	69.6	50.3	10.5	100.0	99.9
λ_f	2	-7.66	1.82	3.83	11.98	20.50	3.95	117.28	46.5	31.5
Em^*	2	-0.01	0.05	0.24	0.41	0.46	0.23	0.14	56.7	46.0
R^2	2	0.14	0.32	0.71	0.90	0.93	0.65	0.25	–	–
J	2	3.29	4.16	8.40	13.68	15.24	8.72	3.67	18.8	8.6
ξ -Normalization										
μ	1/2	0.0110	0.0120	0.0155	0.0191	0.0201	0.0156	0.0027	100.0	99.9
b^\diamond	1	-67.0	-49.1	-0.5	47.7	66.6	-0.1	41.2	3.4	1.2
λ_f	1	-1.34	-0.98	-0.01	0.97	1.33	0.00	0.83	3.9	1.5
R^2	1	-0.01	0.00	0.06	0.28	0.36	0.10	0.12	–	–
b^\diamond	2	-82.8	-62.4	0.4	61.2	82.5	0.3	51.6	35.0	25.1
λ_f	2	-1.66	-1.26	0.01	1.26	1.65	0.01	1.03	34.3	23.7
R^2	2	-0.23	-0.13	0.02	0.23	0.32	0.04	0.17	–	–
J	2	8.18	11.05	18.81	21.68	22.61	17.52	4.39	87.3	80.6
L (Rank-test)		3.3	4.3	9.7	20.1	24.5	11.3	7.0	28.7	21.0

Note: The table reports results from 10000 Monte Carlo experiments described in the main text. Each experiment uses a sample of 50 observations. In each experiment R_t^e , $t = 1, \dots, 50$, is the 8×1 vector of excess returns from the Lustig-Verdelhan data (the returns do not vary across experiments). In each experiment a random risk factor, f_t , is generated using MATLAB's random number generator (so the risk factor varies across experiments). The process for f_t is a Gaussian AR(1) with mean, variance and first order serial correlation equal to the sample mean, variance and first order serial correlation of US consumption growth in the Lustig-Verdelhan data. For each experiment the model is estimated using the a -normalization and the ξ -normalization.

TABLE 10: Monte-Carlo Experiments with the Fama-French 25 Returns

	GMM Stage	Percentiles					Mean	Std. Dev.	Percent Significant at	
		5	10	50	90	95			10% level	5% level
<i>a</i> -Normalization										
<i>b</i>	1	101	106	127	157	169	130	21	100.0	100.0
λ_f	1	0.56	0.61	0.94	1.89	2.65	1.15	7.1	70.6	51.3
Em^*	1	0.15	0.21	0.35	0.46	0.48	0.34	0.1	85.3	78.8
R^2	1	0.55	0.62	0.80	0.91	0.93	0.78	0.12	–	–
<i>b</i>	2	115	118	131	147	152	132	11	100.0	100.0
λ_f	2	0.77	0.82	1.04	1.43	1.60	1.09	0.29	99.9	99.6
Em^*	2	0.23	0.26	0.33	0.40	0.42	0.33	0.06	100.0	99.9
R^2	2	-0.39	-0.06	0.64	0.86	0.90	0.50	0.45	–	–
<i>J</i>	2	18.2	20.5	29.1	38.8	41.7	29.4	7.1	28.9	16.4
ξ -Normalization										
μ	1/2	0.0045	0.0047	0.0051	0.0055	0.0056	0.0051	0.0003	100.0	100.0
b^\diamond	1	199	219	335	580	686	371	178	73.9	50.7
λ_f	1	0.53	0.58	0.88	1.51	1.78	0.97	0.45	75.4	53.6
R^2	1	-5.86	-3.20	-0.54	0.25	0.39	-1.30	2.59	–	–
b^\diamond	2	53.0	76.8	146	218	240	146	57.2	71.6	61.5
λ_f	2	0.14	0.20	0.39	0.57	0.62	0.38	0.15	73.8	64.8
R^2	2	-14.8	-13.0	-5.9	-1.3	-0.6	-6.6	4.4	–	–
<i>J</i>	2	9.2	12.2	26.8	43.3	47.5	27.4	12.0	32.0	23.6
<i>L</i> (Rank-test)		30.0	34.2	51.1	74.3	81.5	53.0	16.4	89.6	83.5

Note: The table reports results from 10000 Monte Carlo experiments described in the main text. Each experiment uses a sample of 228 observations. In each experiment R_t^e , $t = 1, \dots, 228$, is the 25×1 vector of excess returns from the Fama-French data (1949–2005). Panel spurious risk factors are generated using MATLAB’s random number generator by setting $f_t = \mu + c(R_t^e - \bar{R}^e)' \hat{\Sigma}_R^{-1} \iota + \epsilon_t$ where \bar{R}^e and $\hat{\Sigma}_R$ are the sample mean and covariance matrix of R_t^e , and ϵ_t is $\epsilon_t \sim Niid(0, \sigma^2)$. Across repeated samples $\text{cov}(R_t^e, f_t) = c\iota$, by construction. So that f_t mimics US consumption growth in the period 1949–2005, I set $\mu = 0.0051$, $c = 7.49 \times 10^{-5}$, and $\sigma = 0.005$. This implies that across repeated samples $E(f_t)$, $\text{var}(f_t)$ and $\text{cov}(R_t^e, f_t)$ match the corresponding sample moments in US data. For each experiment the model is estimated using the *a*-normalization and the ξ -normalization.

TABLE 11: Monte-Carlo Experiments with Artificial Annual Data
 Estimated Model uses the True Risk Factor and the a -Normalization

	GMM Stage	Percentiles					Mean	Std. Dev.	Percent Significant at	
		5	10	50	90	95			10% level	5% level
$T = 50$										
b^*	1	10.2	12.7	20.6	27.0	28.7	20.2	5.7	90.8	86.7
λ_f	1	0.25	0.33	0.63	0.95	1.04	0.64	0.24	85.4	76.9
Em^*	1	0.49	0.53	0.68	0.83	0.87	0.68	0.12	100.0	100.0
R^2	1	0.33	0.52	0.86	0.96	0.97	0.79	0.22	–	–
b^*	2	14.5	17.0	24.0	29.7	31.1	23.6	5.1	97.4	95.7
λ_f	2	0.38	0.47	0.79	1.13	1.23	0.80	0.26	94.4	89.4
Em^*	2	0.45	0.49	0.62	0.77	0.82	0.63	0.11	100.0	100.0
R^2	2	0.15	0.38	0.83	0.95	0.97	0.73	0.28	–	–
J	2	2.4	3.1	6.2	10.5	11.8	6.5	2.9	4.4	1.2
$T = 250$										
b^*	1	15.3	16.4	19.7	22.7	23.6	19.6	2.5	100.0	100.0
λ_f	1	0.42	0.46	0.60	0.73	0.77	0.60	0.11	100.0	100.0
Em^*	1	0.61	0.63	0.69	0.76	0.78	0.69	0.05	100.0	100.0
R^2	1	0.89	0.91	0.96	0.99	0.99	0.95	0.04	–	–
b^*	2	16.5	17.5	20.5	23.3	24.1	20.5	2.3	100.0	100.0
λ_f	2	0.47	0.50	0.63	0.76	0.80	0.63	0.10	100.0	100.0
Em^*	2	0.60	0.62	0.68	0.75	0.76	0.68	0.05	100.0	100.0
R^2	2	0.87	0.90	0.96	0.99	0.99	0.95	0.04	–	–
J	2	2.2	2.8	6.3	11.7	13.5	6.9	3.5	8.9	3.9
$T = 1000$										
b^*	1	17.4	17.9	19.5	21.1	21.5	19.5	1.2	100.0	100.0
λ_f	1	0.50	0.52	0.59	0.66	0.68	0.59	0.05	100.0	100.0
Em^*	1	0.65	0.66	0.70	0.73	0.74	0.70	0.03	100.0	100.0
R^2	1	0.98	0.98	0.99	1.00	1.00	0.99	0.01	–	–
b^*	2	17.8	18.3	19.8	21.2	21.6	19.7	1.1	100.0	100.0
λ_f	2	0.52	0.54	0.60	0.66	0.68	0.60	0.05	100.0	100.0
Em^*	2	0.65	0.66	0.69	0.72	0.73	0.69	0.03	100.0	100.0
R^2	2	0.97	0.98	0.99	1.00	1.00	0.99	0.01	–	–
J	2	2.1	2.9	6.3	11.8	13.8	6.9	3.7	9.4	4.6

Note: The table reports results from 10000 Monte Carlo experiments with sample sizes $T = 50$, $T = 250$ and $T = 1000$. The true risk factor, $f_t \sim Niid(\mu, \sigma^2)$ with μ and σ^2 equal to the sample mean and variance of US consumption growth in the LV data. The SDF is given by $m_t = a - f_t b$, with $a = 1.42$, $b = 27.72$. I generate $R_t^e = \mu_R + \eta(m_t - \mu_m) + \Psi \xi_t$ where μ_R and η are 8×1 vectors, $\mu_m = a - \mu b$, Ψ is an 8×8 diagonal matrix, and $\xi_t \sim Niid(0, I_n)$ and is independent of f_t . The parameters in μ_R , η and Ψ are set so that $E(R_{it}^e m_t) = 0$, and the means and variances of the elements of R_t^e are equal to their sample equivalents in the LV data. The model is estimated by GMM using the a -normalization. The parameters' true values are $b^* = 19.5$, $\lambda_f = 0.678$ (expressed in percent) and $Em^* = 0.70$.

TABLE 12: Monte-Carlo Experiments with Artificial Annual Data
 Estimated Model uses the True Risk Factor and the ξ -Normalization

	GMM Stage	Percentiles					Mean	Std. Dev.	Percent Significant at	
		5	10	50	90	95			10% level	5% level
$T = 50$										
μ	1/2	0.012	0.013	0.016	0.018	0.019	0.016	0.002	100.0	100.0
b^\diamond	1	8.5	12.4	26.7	45.1	51.1	27.9	13.1	73.6	59.4
λ_f	1	0.18	0.26	0.55	0.85	0.94	0.55	0.23	77.0	66.1
R^2	1	0.10	0.27	0.68	0.88	0.91	0.62	0.25	–	–
b^\diamond	2	9.9	13.3	27.6	46.3	52.4	28.9	13.2	82.0	72.5
λ_f	2	0.21	0.28	0.57	0.87	0.96	0.57	0.23	83.1	74.9
R^2	2	0.04	0.22	0.66	0.87	0.90	0.59	0.27	–	–
J	2	2.6	3.3	7.2	13.2	15.1	7.8	3.9	14.5	7.5
L (Rank-test)		406	469	807	1384	1632	885	394	100.0	100.0
$T = 250$										
μ	1/2	0.014	0.014	0.016	0.017	0.017	0.016	0.001	100.0	100.0
b^\diamond	1	19.0	20.8	27.7	35.3	37.7	27.9	5.7	100.0	100.0
λ_f	1	0.41	0.45	0.58	0.71	0.75	0.58	0.11	100.0	100.0
R^2	1	0.795	0.837	0.925	0.970	0.977	0.911	0.060	–	–
b^\diamond	2	19.9	21.5	28.0	35.2	37.5	28.2	5.3	100.0	100.0
λ_f	2	0.43	0.46	0.59	0.71	0.75	0.59	0.10	100.0	100.0
R^2	2	0.78	0.82	0.92	0.97	0.98	0.90	0.07	–	–
J	2	2.2	2.9	6.5	12.3	14.4	7.2	3.8	10.8	5.6
L (Rank-test)		1985	2127	2691	3400	3636	2741	507	100.0	100.0
$T = 1000$										
μ	1/2	0.015	0.015	0.016	0.016	0.016	0.016	0.000	100.0	100.0
b^\diamond	1	23.4	24.4	27.9	31.6	32.7	28.0	2.8	100.0	100.0
λ_f	1	0.50	0.52	0.59	0.66	0.67	0.59	0.05	100.0	100.0
R^2	1	0.952	0.960	0.980	0.992	0.994	0.978	0.014	–	–
b^\diamond	2	23.8	24.7	27.9	31.5	32.4	28.0	2.6	100.0	100.0
λ_f	2	0.51	0.53	0.59	0.65	0.67	0.59	0.05	100.0	100.0
R^2	2	0.95	0.96	0.98	0.99	0.99	0.98	0.02	–	–
J	2	2.1	2.9	6.3	11.9	14.0	7.0	3.7	9.8	4.9
L (Rank-test)		8584	8874	10003	11257	11631	10043	932	100.0	100.0

Note: The table reports results from the Monte Carlo experiments described in Table 11. In this table the results are GMM estimates of the ξ -normalization. The parameters' true values are $b^\diamond = 28$, $\lambda_f = 0.678$ (expressed in percent) and $\mu = 0.0156$.

TABLE 13: Monte-Carlo Experiments with Artificial Annual Data
 Estimated Model uses a Spurious Risk Factor and the α -Normalization

	GMM Stage	Percentiles					Mean	Std. Dev.	Percent Significant at	
		5	10	50	90	95			10% level	5% level
$T = 50$										
b^*	1	27.3	33.3	49.7	69.6	77.6	50.6	15.8	95.7	93.2
λ_f	1	-19.9	-8.2	3.1	14.6	27.9	8.2	499.1	12.5	4.9
Em^*	1	-0.16	-0.06	0.22	0.49	0.59	0.22	0.23	36.6	27.7
R^2	1	0.25	0.42	0.81	0.94	0.96	0.74	0.23	–	–
b^*	2	27.4	31.6	44.7	59.8	65.5	45.4	12.1	97.2	95.6
λ_f	2	0.1	1.0	2.8	8.8	14.7	2.5	164.4	37.6	22.9
Em^*	2	0.00	0.08	0.30	0.52	0.60	0.30	0.19	61.9	52.3
R^2	2	0.22	0.39	0.78	0.93	0.95	0.71	0.24	–	–
J	2	2.4	3.0	6.3	10.8	12.2	6.7	3.0	5.5	1.7
$T = 250$										
b^*	1	47.4	50.0	60.3	74.2	79.1	61.4	9.8	100.0	100.0
λ_f	1	-64.5	-30.3	6.3	35.6	66.0	4.7	757.7	9.6	4.4
Em^*	1	-0.21	-0.14	0.06	0.21	0.25	0.05	0.14	15.0	9.9
R^2	1	0.82	0.87	0.95	0.98	0.99	0.93	0.06	–	–
b^*	2	44.3	46.9	55.6	66.8	70.8	56.3	8.1	100.0	100.0
λ_f	2	-35.9	-14.2	6.5	25.8	45.9	2.8	605.2	21.9	12.2
Em^*	2	-0.09	-0.03	0.13	0.27	0.31	0.12	0.12	32.9	24.3
R^2	2	0.79	0.84	0.94	0.98	0.98	0.92	0.07	–	–
J	2	2.2	2.8	6.1	11.3	13.2	6.7	3.5	7.8	3.7
$T = 1000$										
b^*	1	55.3	57.0	63.3	70.9	73.5	63.7	5.5	100.0	100.0
λ_f	1	-128.1	-63.5	10.2	68.7	133.8	-650.1	95362.6	9.8	4.8
Em^*	1	-0.13	-0.10	0.01	0.11	0.13	0.01	0.08	10.8	6.0
R^2	1	0.96	0.97	0.99	0.99	1.00	0.98	0.01	–	–
b^*	2	54.4	55.9	61.5	68.1	70.4	61.9	4.9	100.0	100.0
λ_f	2	-111.9	-53.5	13.3	62.7	123.9	13.3	756.3	13.9	7.6
Em^*	2	-0.09	-0.06	0.04	0.12	0.15	0.04	0.07	17.6	11.0
R^2	2	0.96	0.96	0.98	0.99	1.00	0.98	0.01	–	–
J	2	2.2	2.9	6.3	11.7	13.7	6.9	3.6	9.0	4.4

Note: The table reports results from the Monte Carlo experiments described in Table 11. Here the asset returns and true risk factor are generated in the same way, but I report results from experiments in which the proposed SDF is $m_t^* = 1 - x_t b^*$ and x_t is a spurious factor. I assume that $x_t \sim Niid(\mu, \sigma^2)$ with μ and σ^2 equal to the sample mean and variance of US consumption growth in the Lustig-Verdelhan data, and x_t is independent of the true risk factor, f_t . The model is estimated by GMM using the α -normalization.

TABLE 14: Monte-Carlo Experiments with Artificial Annual Data
 Estimated Model uses a Spurious Risk Factor and the ξ -Normalization

	GMM Stage	Percentiles					Mean	Std. Dev.	Percent Significant at	
		5	10	50	90	95			10% level	5% level
$T = 50$										
μ	1/2	0.012	0.013	0.016	0.018	0.019	0.016	0.002	100.0	100.0
b^\diamond	1	-101.7	-78.9	-0.1	77.9	100.6	0.0	64.0	16.7	6.2
λ_f	1	-2.05	-1.59	0.00	1.57	2.00	0.00	1.29	19.9	8.6
R^2	1	-0.11	-0.05	0.10	0.52	0.63	0.17	0.24	–	–
b^\diamond	2	-74.1	-55.3	-0.4	56.7	76.4	0.2	47.3	20.7	12.2
λ_f	2	-1.52	-1.13	-0.01	1.15	1.53	0.00	0.95	22.1	13.4
R^2	2	-0.20	-0.12	0.05	0.45	0.57	0.10	0.24	–	–
J	2	3.0	3.9	9.2	16.5	18.7	9.8	4.8	30.5	18.9
L (Rank-test)		3.9	5.2	12.3	26.6	32.4	14.5	9.4	44.7	35.3
$T = 250$										
μ	1/2	0.014	0.014	0.016	0.017	0.017	0.016	0.001	100.0	100.0
b^\diamond	1	-185.7	-152.4	1.5	151.3	186.9	0.4	121.1	49.3	27.6
λ_f	1	-3.88	-3.16	0.03	3.16	3.88	0.01	2.53	50.6	29.6
R^2	1	-0.02	0.00	0.15	0.58	0.68	0.22	0.23	–	–
b^\diamond	2	-112.2	-86.7	-1.0	86.6	113.4	0.1	70.3	20.0	10.3
λ_f	2	-2.34	-1.81	-0.02	1.81	2.36	0.00	1.47	21.0	11.2
R^2	2	-0.06	-0.03	0.08	0.47	0.61	0.16	0.21	–	–
J	2	2.9	4.1	14.1	35.3	40.7	17.3	12.2	56.4	50.1
L (Rank-test)		3.0	3.8	8.0	15.3	18.1	9.0	4.8	16.1	9.5
$T = 1000$										
μ	1/2	0.015	0.015	0.016	0.016	0.016	0.016	0.000	100.0	100.0
b^\diamond	1	-353.4	-291.9	-4.1	296.8	360.7	-0.4	233.8	68.9	49.8
λ_f	1	-7.42	-6.13	-0.09	6.24	7.55	-0.01	4.92	69.3	50.4
R^2	1	-0.01	0.00	0.17	0.59	0.69	0.24	0.23	–	–
b^\diamond	2	-201.7	-157.0	-1.6	158.0	205.6	-0.7	125.9	25.2	15.4
λ_f	2	-4.22	-3.28	-0.03	3.34	4.29	-0.01	2.65	25.5	15.6
R^2	2	-0.03	-0.02	0.09	0.49	0.61	0.17	0.21	–	–
J	2	2.8	4.0	17.1	99.7	121.1	34.1	37.8	60.6	55.6
L (Rank-test)		2.8	3.6	7.6	13.8	16.0	8.2	4.1	11.2	5.8

Note: The table reports results from the Monte Carlo experiments described in Table 11. Here the asset returns and true risk factor are generated in the same way, but I report results from experiments in which the proposed SDF is $m_t^\diamond = 1 - (x_t - \mu)b^\diamond$ and x_t is a spurious factor. I assume that $x_t \sim Niid(\mu, \sigma^2)$ with μ and σ^2 equal to the sample mean and variance of US consumption growth in the Lustig-Verdelhan data, and x_t is independent of the true risk factor, f_t . The model is estimated by GMM using the ξ -normalization.

TABLE 15: Monte-Carlo Experiments with Artificial Quarterly Data, $T = 228$
 Estimated Model Uses the True Risk Factors

	GMM Stage	Percentiles					Mean	Std. Dev.	Percent Significant at	
		5	10	50	90	95			10% level	5% level
<i>a</i> -Normalization										
b_1^*	1	2.49	2.80	3.87	4.91	5.21	3.86	0.83	99.7	99.3
b_2^*	1	-2.28	-1.79	-0.09	1.61	2.10	-0.09	1.32	11.0	5.3
b_3^*	1	3.96	4.37	5.83	7.24	7.63	5.81	1.12	99.9	99.8
λ_{f1}	1	1.02	1.21	1.92	2.61	2.80	1.92	0.54	97.2	94.1
λ_{f2}	1	-0.06	0.07	0.53	1.00	1.14	0.53	0.36	42.5	31.1
λ_{f3}	1	0.70	0.83	1.31	1.79	1.92	1.31	0.37	97.3	94.4
Em^*	1	0.77	0.79	0.85	0.90	0.91	0.85	0.04	100.0	100.0
R^2	1	0.84	0.87	0.94	0.97	0.98	0.93	0.05	–	–
b_1^*	2	2.83	3.17	4.34	5.49	5.81	4.34	0.91	99.9	99.8
b_2^*	2	-2.50	-1.95	-0.06	1.83	2.35	-0.08	1.47	17.2	10.5
b_3^*	2	4.46	4.91	6.53	8.06	8.48	6.51	1.22	100.0	99.9
λ_{f1}	2	1.17	1.40	2.21	3.03	3.26	2.21	0.63	98.2	96.4
λ_{f2}	2	-0.06	0.09	0.62	1.16	1.32	0.62	0.42	52.1	40.4
λ_{f3}	2	0.80	0.95	1.49	2.04	2.20	1.50	0.42	98.3	96.6
Em^*	2	0.75	0.77	0.83	0.88	0.90	0.83	0.05	100.0	100.0
R^2	2	-0.50	-0.10	0.64	0.91	0.94	0.49	0.53	–	–
J	2	13.1	14.7	21.8	30.2	32.9	22.2	6.1	8.6	3.8
ξ -Normalization										
μ_1	1/2	0.0108	0.0127	0.0197	0.0266	0.0285	0.0197	0.0054	97.9	95.6
μ_2	1/2	0.0005	0.0018	0.0063	0.0109	0.0122	0.0063	0.0035	55.7	43.0
μ_3	1/2	0.0059	0.0072	0.0118	0.0164	0.0177	0.0118	0.0036	95.1	90.9
b_1°	1	2.78	3.17	4.55	6.08	6.54	4.60	1.14	99.7	99.2
b_2°	1	-2.72	-2.11	-0.09	1.90	2.49	-0.10	1.56	9.8	4.6
b_3°	1	4.41	4.92	6.84	8.94	9.62	6.90	1.58	99.9	99.8
λ_{f1}	1	1.02	1.21	1.91	2.61	2.80	1.92	0.54	97.3	94.1
λ_{f2}	1	-0.06	0.08	0.53	1.00	1.14	0.53	0.36	42.8	31.3
λ_{f3}	1	0.69	0.83	1.30	1.77	1.91	1.30	0.37	97.1	94.2
R^2	1	0.80	0.83	0.92	0.96	0.96	0.90	0.06	–	–
b_1°	2	2.79	3.15	4.53	6.12	6.58	4.60	1.16	99.8	99.3
b_2°	2	-2.66	-2.08	-0.09	1.92	2.50	-0.10	1.56	12.1	6.2
b_3°	2	4.39	4.89	6.86	9.02	9.69	6.92	1.62	99.9	99.8
λ_{f1}	2	1.00	1.19	1.91	2.63	2.85	1.92	0.56	96.9	93.8
λ_{f2}	2	-0.07	0.07	0.53	1.01	1.15	0.53	0.37	44.2	32.3
λ_{f3}	2	0.69	0.82	1.30	1.79	1.93	1.31	0.38	97.2	94.7
R^2	2	0.45	0.59	0.85	0.93	0.94	0.79	0.18	–	–
J	2	13.5	15.3	23.0	32.5	35.6	23.5	6.8	13.7	7.3
L (Rank-test)		7678	8154	10345	13231	14086	10571	2003	100.0	100.0

Note: The note to Table 15 is on the following page.

Note to Table 15: The table reports results from 10000 Monte Carlo experiments with sample size $T = 228$. The true risk factors, $f_t \sim Niid(\mu, \Sigma_f)$ with μ and Σ_f equal to the sample mean and covariance matrix of the Fama-French factors (1949Q1–2005Q4). The SDF is given by $m_t = a - f_t' b$, with $a = 1.166$, $b = (4.50 \quad -0.12 \quad 6.78)'$. I generate $R_t^e = \mu_R + \beta(f_t - \mu) + \Psi \xi_t$, where μ_R is a 25×1 vector, β is an 25×3 matrix, Ψ is a 25×25 lower triangular matrix, and $\xi_t \sim Niid(0, I_n)$ and is independent of f_t . The elements of β and Ψ so that the model-implied $\text{var}(R_t^e)$ and $\text{cov}(R_t^e, f_t)$ are equal to their sample equivalents for the FF25 portfolios and Fama-French factors (1949Q1–2005Q4). The vector μ_R is set to ensure that $E(R_t^e m_t) = 0$. The model is estimated by GMM using the a - and ξ -normalizations. The parameters' true values are $\mu_1 = 0.0197$, $\mu_2 = 0.0063$, $\mu_3 = 0.0119$, $b_1^* = 3.86$, $b_2^* = -0.10$, $b_3^* = 5.82$, $b_1^\diamond = 4.51$, $b_2^\diamond = -0.12$, $b_3^\diamond = 6.80$, $\lambda_{f1} = 1.92$, $\lambda_{f2} = 0.53$, $\lambda_{f3} = 1.32$. and $Em^* = 0.86$.

TABLE 16: Monte-Carlo Experiments with Artificial Quarterly Data, $T = 228$
 Estimated Model Uses a Single Relevant Factor

	GMM Stage	Percentiles					Mean	Std. Dev.	Percent Significant at	
		5	10	50	90	95			10% level	5% level
<i>a</i> -Normalization										
b^*	1	2.01	2.29	3.33	4.34	4.61	3.32	0.79	99.1	98.2
λ_f	1	1.35	1.56	2.31	3.07	3.28	2.31	0.58	99.1	98.0
Em^*	1	0.87	0.89	0.93	0.97	0.98	0.93	0.03	100.0	100.0
R^2	1	-1.09	-0.89	-0.41	0.01	0.15	-0.43	0.39	–	–
b^*	2	2.27	2.62	3.86	5.03	5.34	3.85	0.94	99.5	99.1
λ_f	2	1.53	1.79	2.71	3.64	3.89	2.71	0.72	99.5	99.0
Em^*	2	0.85	0.87	0.92	0.97	0.97	0.92	0.04	100.0	100.0
R^2	2	-3.18	-2.43	-0.83	-0.17	-0.02	-1.13	1.17	–	–
J	2	28.7	31.5	42.6	54.8	58.1	42.9	9.0	86.2	76.0
ξ -Normalization										
μ	12	0.0108	0.0127	0.0197	0.0266	0.0285	0.0197	0.0054	97.9	95.6
b^\diamond	1	2.04	2.35	3.54	4.84	5.22	3.58	0.97	99.0	97.8
λ_f	1	1.35	1.55	2.29	3.05	3.26	2.29	0.58	99.1	97.9
R^2	1	-1.45	-1.19	-0.60	-0.17	-0.03	-0.65	0.45	–	–
b^\diamond	2	1.64	1.92	3.05	4.31	4.69	3.10	0.94	97.8	95.0
λ_f	2	1.07	1.26	1.98	2.72	2.94	1.99	0.57	98.0	95.4
R^2	2	-2.10	-1.78	-0.99	-0.51	-0.37	-1.08	0.56	–	–
J	2	29.2	32.3	44.0	57.3	61.1	44.5	9.7	88.2	79.1
L (Rank-test)		25259	27305	34312	43574	46898	35047	6588	100.0	100.0

Note: The table reports results from Monte Carlo experiments described in Table 15. The proposed SDF is a single factor model that uses a fully relevant risk factor, x_{1t} , which is equal to the first true risk factor, f_{1t} (described in Table 15). The model is estimated by GMM using the *a*- and ξ -normalizations.

TABLE 17: Monte-Carlo Experiments with Artificial Quarterly Data, $T = 228$
 Estimated Model Uses a Single Factor that is Panel Spurious

	GMM Stage	Percentiles					Mean	Std. Dev.	Percent Significant at	
		5	10	50	90	95			10% level	5% level
<i>a</i> -Normalization										
b^*	1	91.4	99.0	126.4	159.8	171.3	128.4	24.7	100.0	99.9
λ_f	1	0.44	0.52	0.93	2.03	2.91	1.26	18.89	71.2	54.1
Em^*	1	0.14	0.19	0.35	0.50	0.54	0.35	0.12	85.7	79.7
R^2	1	0.46	0.56	0.79	0.91	0.93	0.75	0.18	–	–
b^*	2	98.0	102.8	119.4	137.4	143.1	119.9	13.7	100.0	100.0
λ_f	2	0.53	0.58	0.80	1.17	1.32	0.85	0.26	99.8	99.3
Em^*	2	0.27	0.30	0.39	0.48	0.51	0.39	0.07	99.9	99.8
R^2	2	-1.14	-0.49	0.58	0.85	0.89	0.31	0.87	–	–
J	2	16.5	18.7	26.8	36.2	39.0	27.2	6.8	18.7	9.5
ξ -Normalization										
μ	12	0.0045	0.0047	0.0051	0.0055	0.0056	0.0051	0.0003	100.0	100.0
b^\diamond	1	148.6	183.2	329.2	614.8	742.3	367.6	210.4	75.0	53.9
λ_f	1	0.40	0.48	0.87	1.60	1.92	0.96	0.54	76.0	56.1
R^2	1	-7.88	-4.18	-0.51	0.29	0.43	-1.70	4.23	–	–
b^\diamond	2	16.2	34.3	96.9	163.7	182.9	97.6	51.5	44.7	32.5
λ_f	2	0.04	0.09	0.25	0.43	0.48	0.26	0.14	47.4	36.3
R^2	2	-27.7	-22.0	-8.7	-2.0	-1.1	-10.9	9.2	–	–
J	2	5.4	7.2	17.4	32.1	36.9	18.7	9.8	8.5	5.2
L (Rank-test)		30.1	33.8	51.9	74.8	83.0	53.7	16.9	88.8	83.3

Note: The table reports results from Monte Carlo experiments described in Table 15. The proposed SDF is a single factor model that uses a panel spurious factor. This factor, x_{2t} , is generated according to the law of motion

$$x_{2t} = \mu_{x_2} + c\iota'\Sigma_R^{-1}(R_t^e - \mu_R) + u_{2t}$$

where c is a scalar, ι is an $n \times 1$ vector of ones, μ_R and Σ_R are the mean and covariance matrix of the vector of returns described in Table 15, and $u_{2t} \sim Niid(0, \sigma_{u_2}^2)$ is independent of R_t^e and x_{1t} (described in Table 16). I set μ_{x_2} equal to the sample mean of US consumption growth in the period 1949Q1–2005Q4. I set c equal to the cross-sectional average of the sample covariance between US consumption growth and the FF25 returns over the same period. I set $\sigma_{u_2}^2$ so that the variance of x_{2t} equals the sample variance of US consumption growth over the same period. The factor x_{2t} is panel spurious because $\text{cov}(R_t^e, x_{2t}) = \iota c$, by construction. The model is estimated by GMM using the a - and ξ -normalizations.

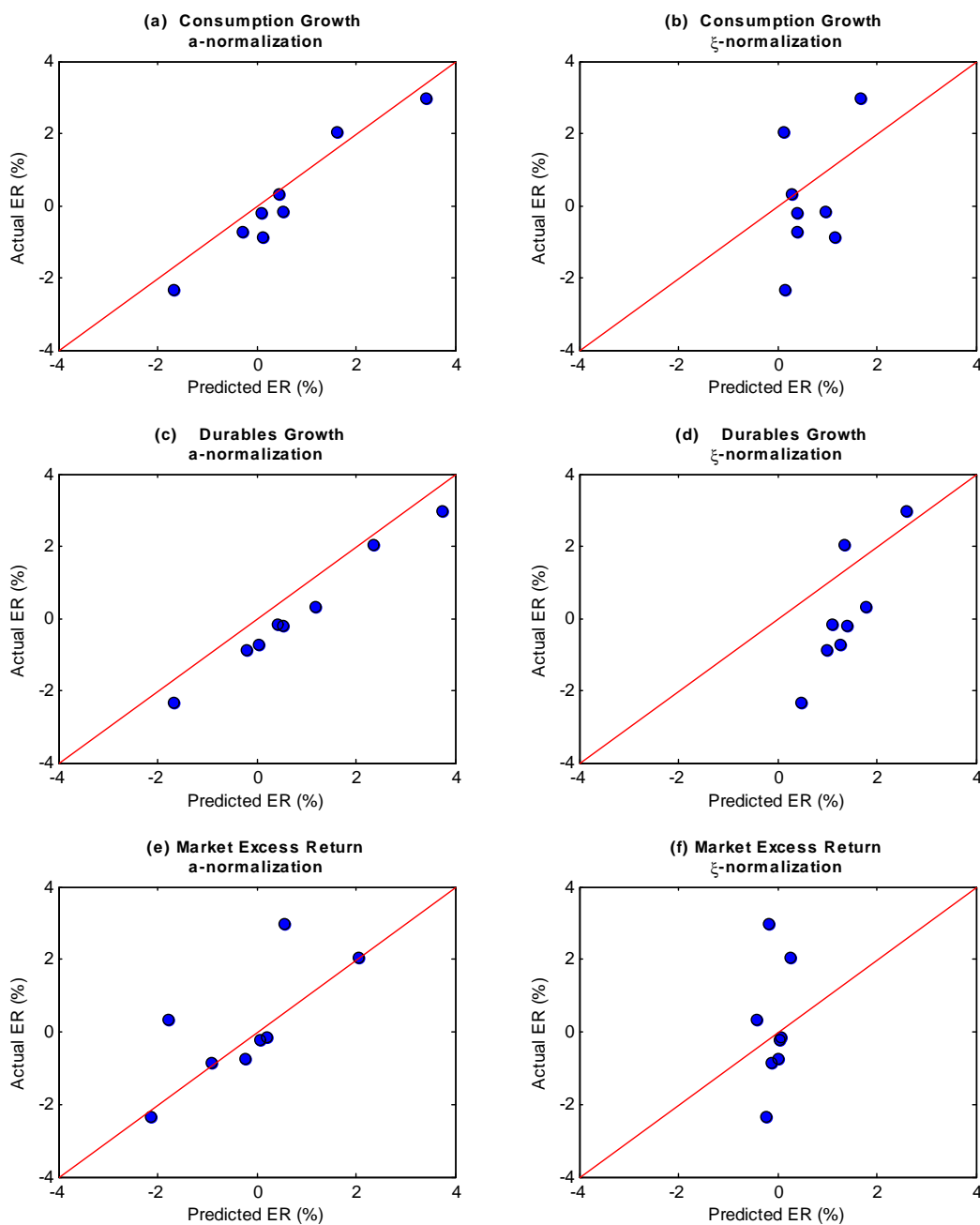
TABLE 18: Monte-Carlo Experiments with Artificial Quarterly Data, $T = 228$
 Estimated Model Has Three Factors, One of Which is Spurious

	GMM Stage	Percentiles					Mean	Std. Dev.	Percent Significant at	
		5	10	50	90	95			10% level	5% level
<i>a</i> -Normalization										
b_1^*	1	-1.22	-0.97	-0.17	0.61	0.82	-0.18	0.63	8.4	3.5
b_2^*	1	-17.0	-4.6	35.8	75.1	87.8	35.5	32.2	36.5	26.3
b_3^*	1	43.7	51.1	73.7	97.2	104.8	73.9	18.8	97.2	95.6
λ_{f1}	1	-2.10	0.00	2.03	4.20	6.18	2.06	31.4	42.5	34.9
λ_{f2}	1	-6.18	-2.61	0.70	3.74	6.59	0.65	34.2	5.8	2.4
λ_{f3}	1	-18.18	-8.48	1.90	9.62	17.78	0.77	152.1	7.1	2.5
Em^*	1	-0.11	-0.07	0.06	0.16	0.20	0.05	0.09	17.5	10.6
R^2	1	0.93	0.94	0.98	0.99	0.99	0.97	0.03	–	–
b_1^*	2	-0.89	-0.70	-0.05	0.58	0.77	-0.06	0.51	12.4	6.9
b_2^*	2	9.0	15.6	37.6	59.4	66.3	37.6	17.6	73.2	64.6
b_3^*	2	49.3	53.1	66.1	79.2	83.2	66.2	10.4	100.0	99.9
λ_{f1}	2	-0.33	0.43	1.94	3.53	4.42	1.81	13.6	49.2	39.5
λ_{f2}	2	0.04	0.26	0.76	2.07	3.26	0.94	17.0	37.0	22.9
λ_{f3}	2	0.67	0.90	1.77	4.85	7.76	1.99	49.3	52.1	37.2
Em^*	2	0.01	0.04	0.12	0.20	0.22	0.12	0.06	62.6	52.4
R^2	2	0.66	0.76	0.93	0.98	0.98	0.89	0.12	–	–
J	2	10.7	12.2	18.6	26.5	28.9	19.1	5.6	2.6	1.0
ξ -Normalization										
μ_1	12	0.0108	0.0127	0.0197	0.0266	0.0285	0.0197	0.0054	97.9	95.6
μ_2	12	0.0045	0.0047	0.0051	0.0055	0.0056	0.0051	0.0003	100.0	100.0
μ_3	12	0.0098	0.0099	0.0104	0.0109	0.0111	0.0104	0.0004	100.0	100.0
b_1^\diamond	1	-3.00	-1.91	1.59	5.19	6.44	1.64	2.92	23.9	16.6
b_2^\diamond	1	-203.6	-101.1	184.3	386.5	446.2	161.2	197.4	43.7	27.2
b_3^\diamond	1	-393.4	-321.2	0.3	318.9	380.9	0.4	245.0	37.3	18.5
λ_{f1}	1	1.21	1.42	2.21	3.00	3.24	2.21	0.61	97.5	94.5
λ_{f2}	1	-0.48	-0.22	0.49	1.00	1.15	0.43	0.49	48.0	32.4
λ_{f3}	1	-1.34	-1.09	0.01	1.09	1.31	0.00	0.84	38.9	20.2
R^2	1	-0.66	-0.45	0.21	0.64	0.72	0.14	0.44	–	–
b_1^\diamond	2	0.32	0.85	2.53	4.24	4.74	2.54	1.34	39.3	26.6
b_2^\diamond	2	-65.3	-39.5	43.9	123.7	148.1	42.8	64.7	11.6	5.9
b_3^\diamond	2	-111.1	-84.4	1.3	83.8	106.3	0.2	66.2	6.7	2.9
λ_{f1}	2	0.90	1.12	1.93	2.74	2.97	1.93	0.63	74.9	62.1
λ_{f2}	2	-0.14	-0.08	0.13	0.34	0.40	0.13	0.16	16.0	8.9
λ_{f3}	2	-0.38	-0.29	0.01	0.28	0.37	0.00	0.23	7.6	3.5
R^2	2	-3.19	-2.44	-0.85	-0.08	0.09	-1.10	1.08	–	–
J	2	6.4	7.5	13.6	26.2	31.6	15.5	8.0	5.6	3.8
L (Rank-test)		14.8	17.2	26.1	37.2	41.4	26.8	8.1	23.4	14.3

Note: The note to Table 18 is on the following page.

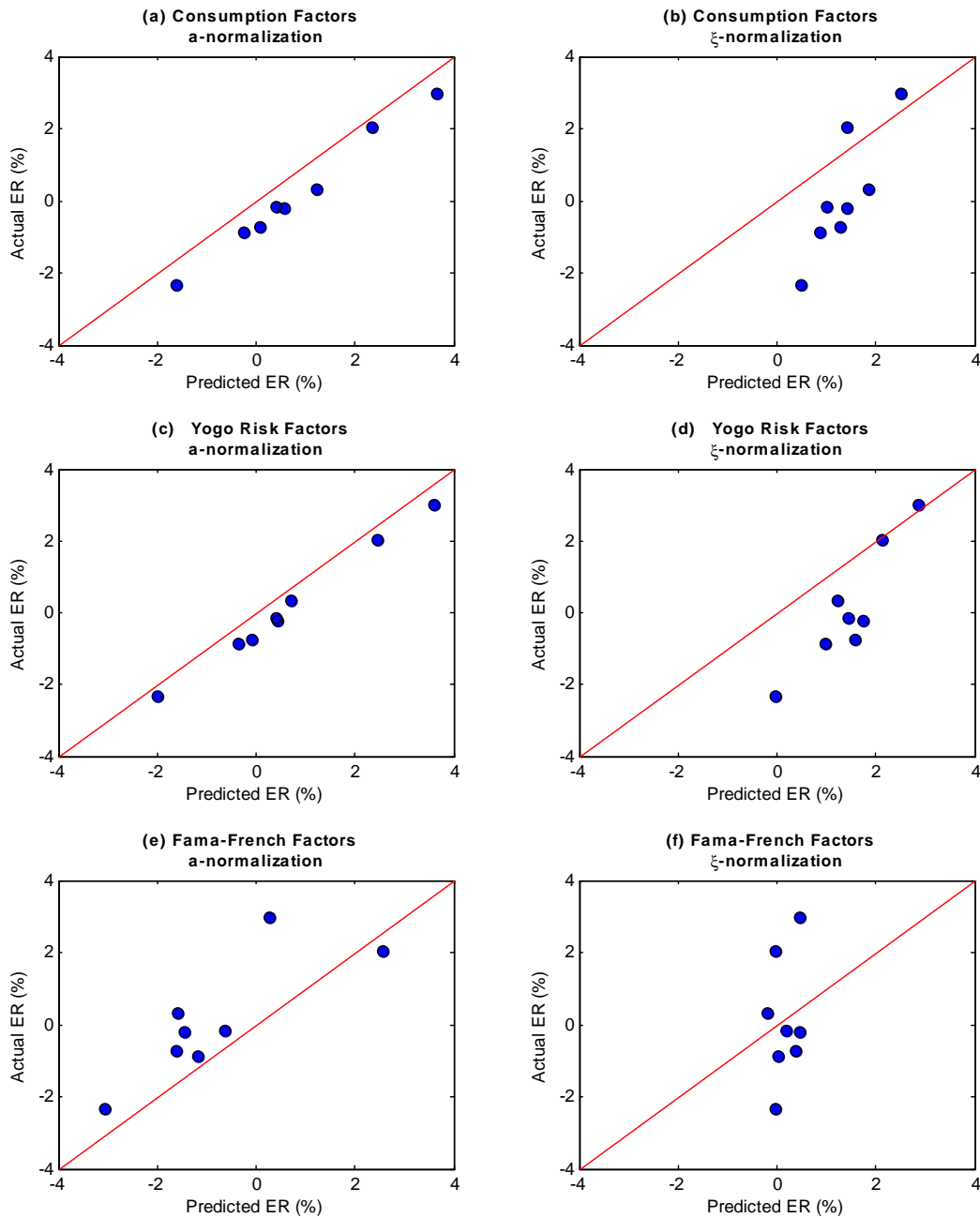
Note to Table 18: The table reports results from Monte Carlo experiments described in Table 15. The proposed SDF is a three factor model that uses a relevant factor, x_{1t} , a panel spurious factor, x_{2t} , and a purely spurious factor, x_{3t} . The factors x_{1t} and x_{2t} are described in Tables 16 and 17, while x_{3t} is generated according to the law of motion $x_{3t} = \mu_{x3} + u_{3t}$, where $u_{3t} \sim Niid(0, \sigma_{u3}^2)$ is independent of R_t^e , u_{1t} and u_{2t} (defined in Tables 16 and 17). I set μ_{x3} equal to the sample mean of durables growth in the period 1949Q1–2005Q4. I set the variance of u_{3t} equal to the sample variance of durables growth in the period 1949Q1–2005Q4. The model is estimated by GMM using the a - and ξ -normalizations.

FIGURE 1
Actual Expected Returns and Predicted Expected Returns for Single Factor Models
 Lustig-Verdelhan Data



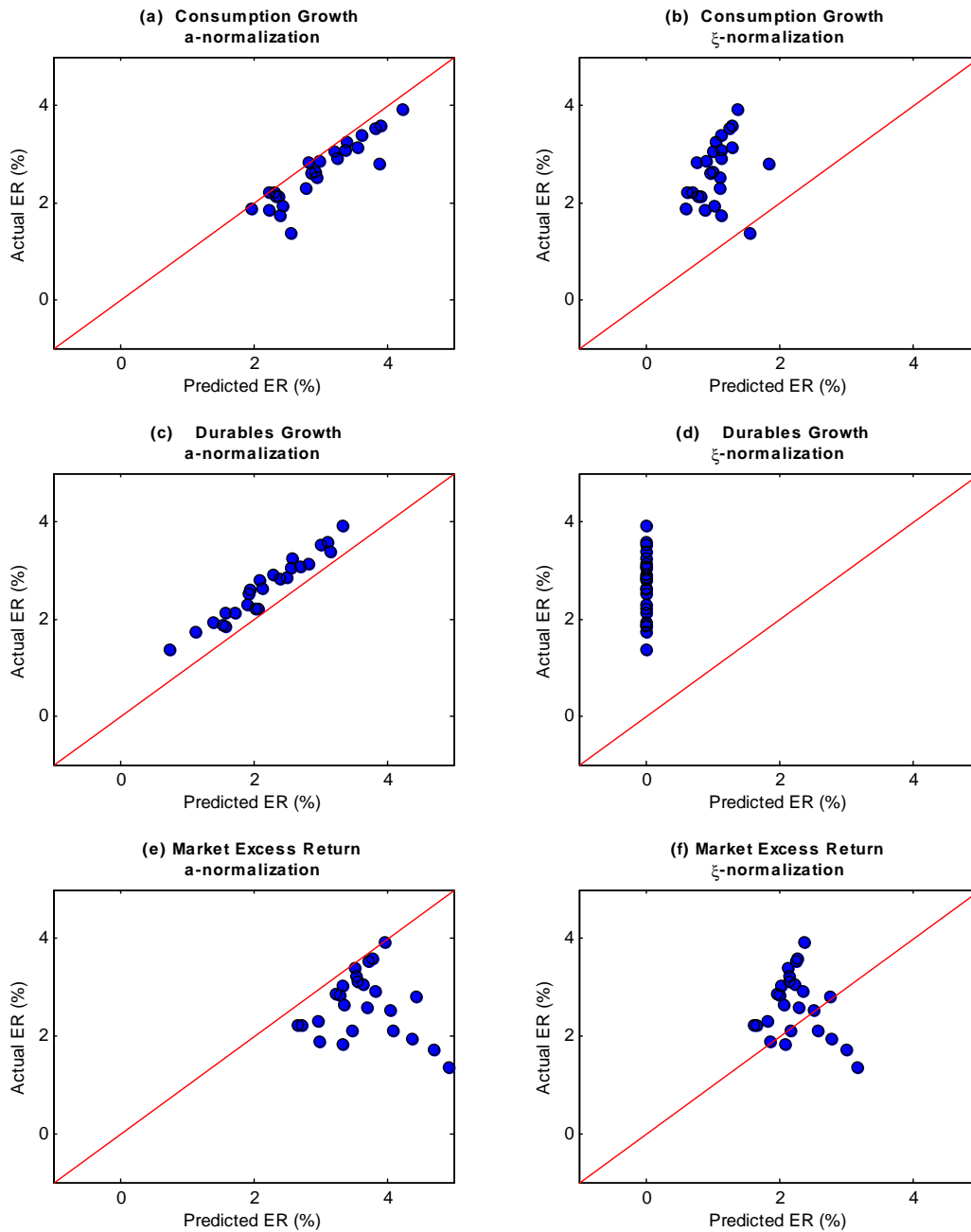
Note: The figures present scatter plots of \bar{R}^e against the model-predicted expected return for the single factor models presented in Tables 1 and 2. For the a -normalization the model predicted expected return is $D_T \hat{b}^*$, while for the ξ -normalization it is $d_T \hat{b}^\xi$.

FIGURE 2
 Actual Expected Returns and Predicted Expected Returns for Multi-Factor Models
 Lustig-Verdelhan Data



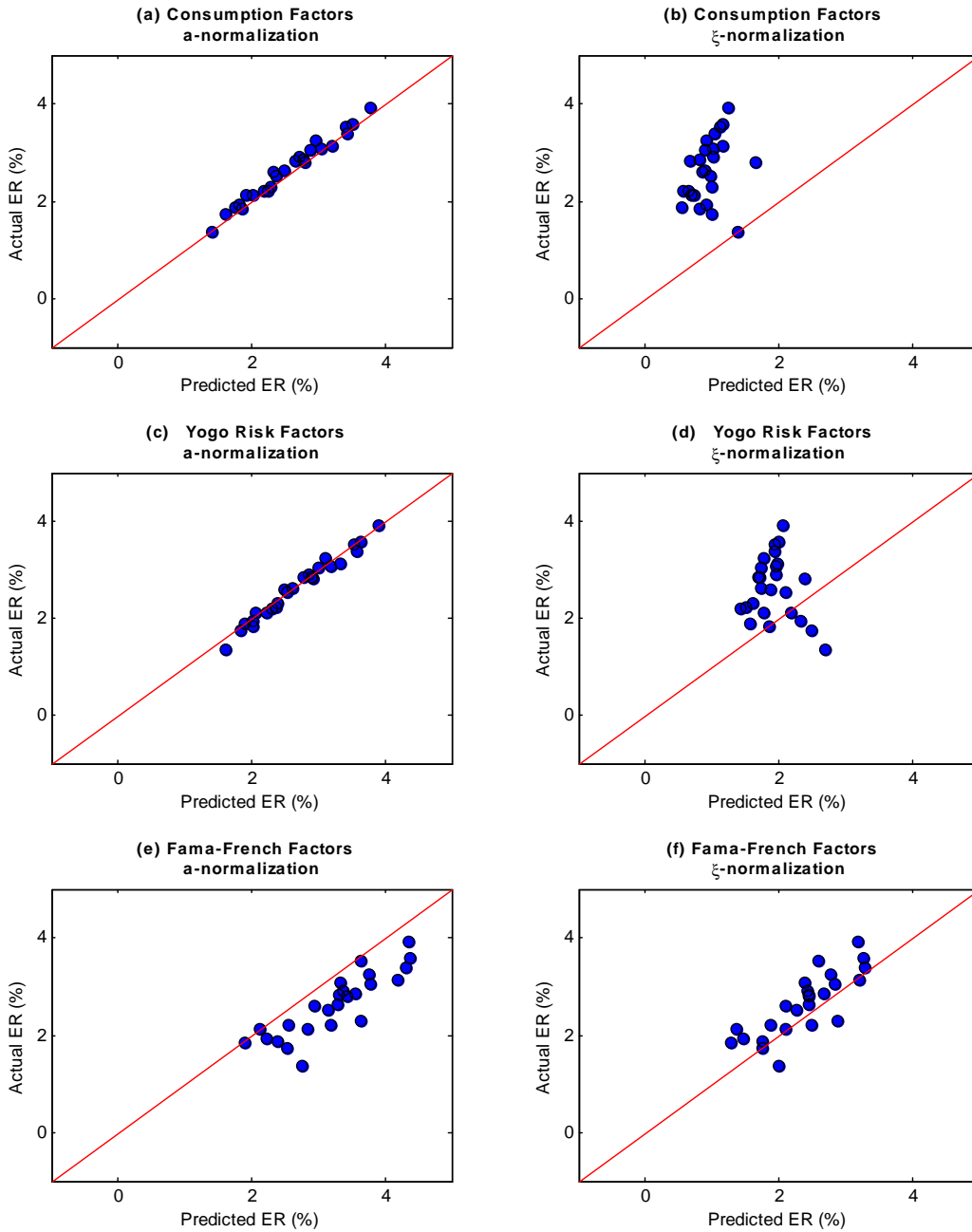
Note: The figures present scatter plots of \bar{R}^e against the model-predicted expected return for the multi-factor models presented in Tables 1 and 2. For the a -normalization the model predicted expected return is $D_T \hat{b}^*$, while for the ξ -normalization it is $d_T \hat{b}^\circ$.

FIGURE 3
 Actual Expected Returns and Predicted Expected Returns for Single Factor Models
 Fama–French 25 Data



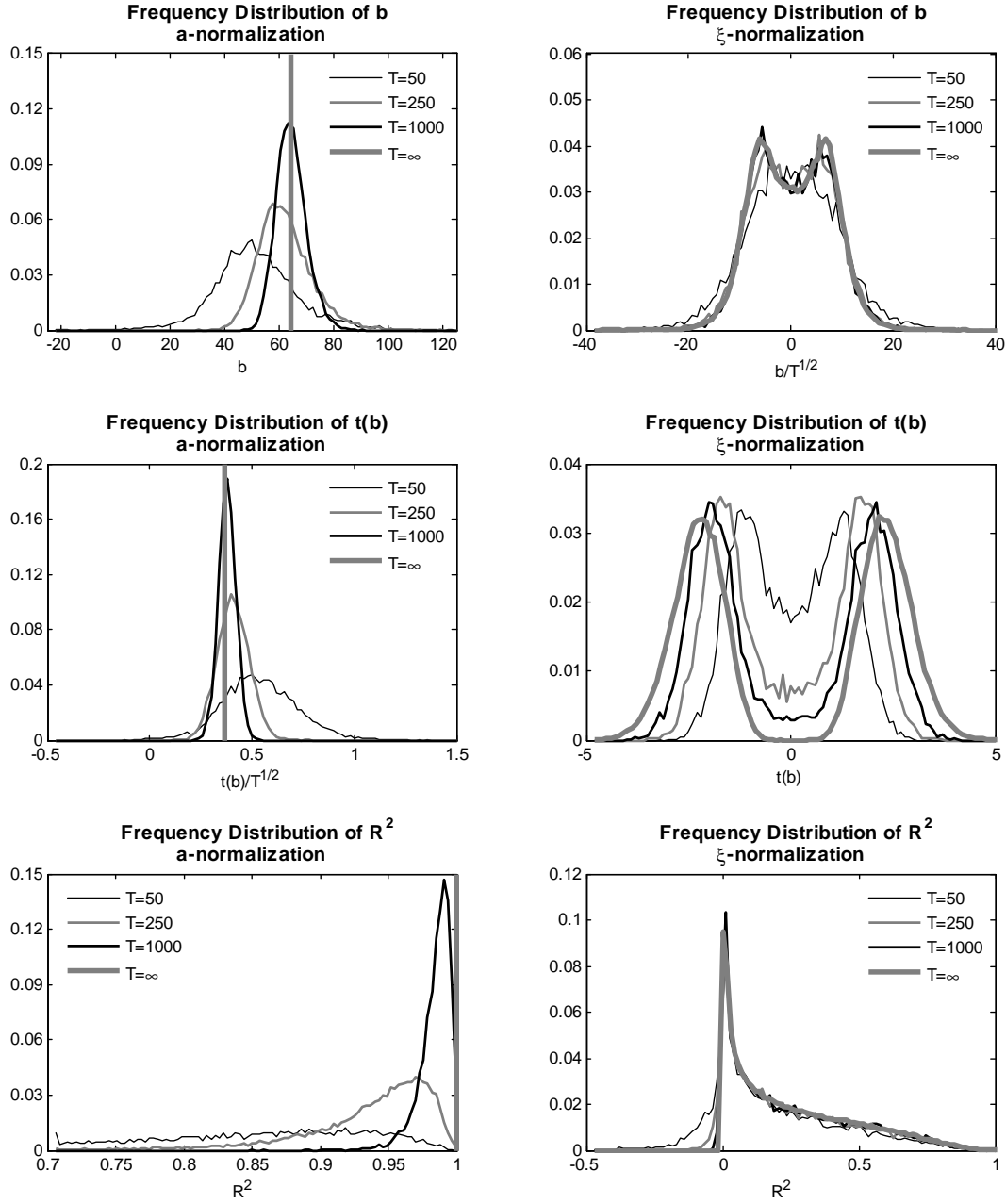
Note: The figures present scatter plots of \bar{R}^e against the model-predicted expected return for the single factor models presented in Tables 3 and 4. For the a -normalization the model predicted expected return is $D_T \hat{b}^*$, while for the ξ -normalization it is $d_T \hat{b}^\circ$.

FIGURE 4
Actual Expected Returns and Predicted Expected Returns for Multi-Factor Models
 Fama-French 25 Data



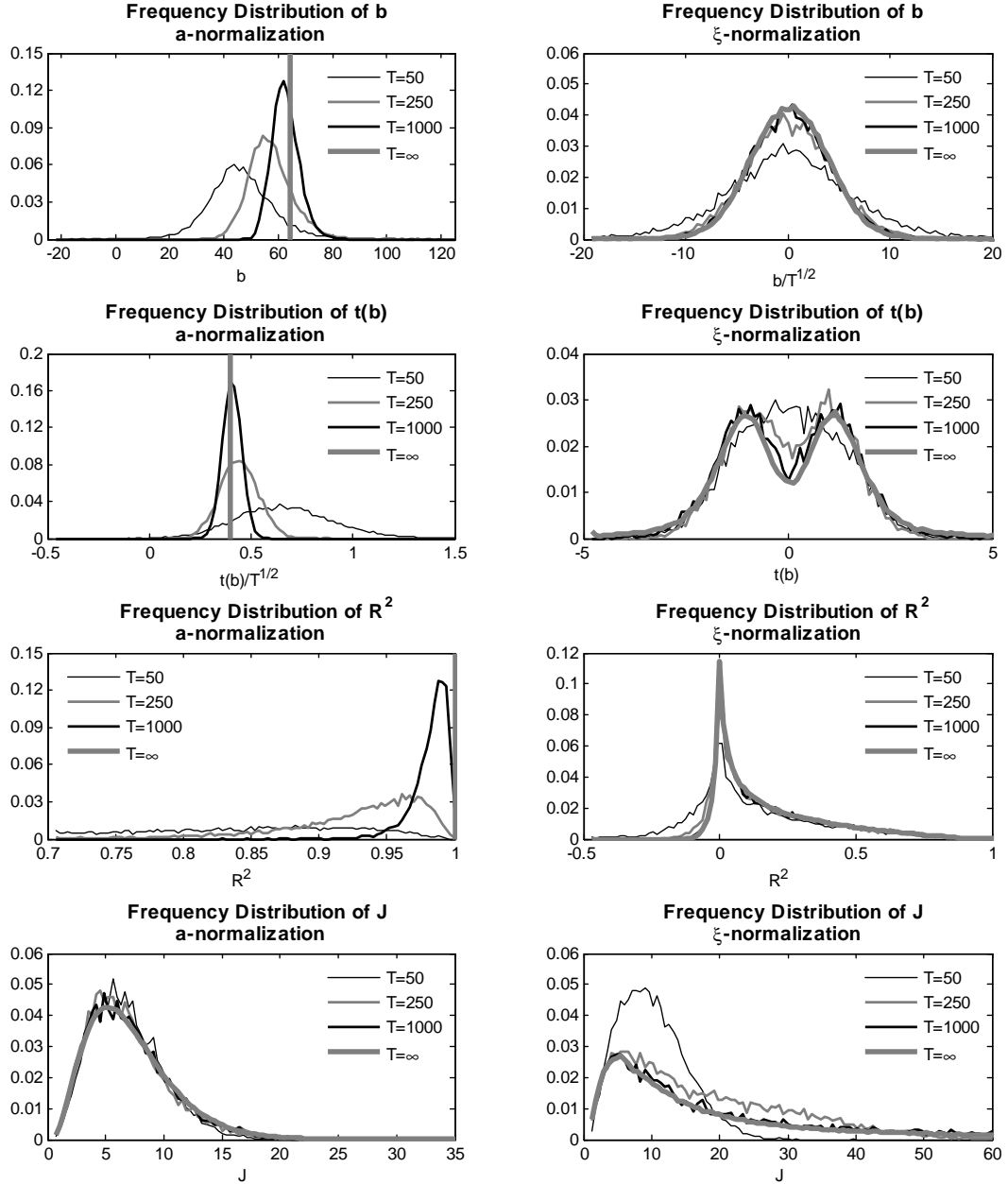
Note: The figures present scatter plots of \bar{R}^e against the model-predicted expected return for the multi-factor models presented in Tables 3 and 4. For the a -normalization the model predicted expected return is $D_T \hat{b}^*$, while for the ξ -normalization it is $d_T \hat{b}^\diamond$.

FIGURE 5
Monte-Carlo Experiments with Artificial Data
 Frequency Distributions of First Stage GMM Estimators



Note: The experiments are described in the notes to Tables 11–14. Statistics are presented for the first stage of GMM when the proposed SDF is constructed from the spurious risk factor. The frequency distributions are computed using 100 points over the indicated range on the x-axis. For $T = \infty$, if the distribution limits to a singleton a vertical red line is used to indicated the probability limit.

FIGURE 6
Monte-Carlo Experiments with Artificial Data
 Frequency Distributions of Second Stage GMM Estimators



Note: The experiments are described in the notes to Tables 11–14. Statistics are presented for the first stage of GMM when the proposed SDF is constructed from the spurious risk factor. The frequency distributions are computed using 100 points over the indicated range on the x-axis. For $T = \infty$, if the distribution limits to a singleton a vertical red line is used to indicate the probability limit.