DIVISION OF THE HUMANITIES AND SOCIAL SCIENCES

CALIFORNIA INSTITUTE OF TECHNOLOGY

PASADENA, CALIFORNIA 91125

An Agnostic and Practically Useful Estimator of the Stochastic Discount Factor

Kuntara Pukthuanthong

University of Missouri

Richard Roll

California Institute of Technology

SOCIAL SCIENCE WORKING PAPER #1405

(Revised) October 2016

An Agnostic and Practically Useful Estimator of the Stochastic Discount Factor

By

Kuntara Pukthuanthong and Richard Roll

October 19, 2016

Abstract

We propose an estimator for the stochastic discount factor (SDF) which is agnostic because it does not require macroeconomic proxies or preference assumptions. It depends only on observed asset returns. Nonetheless, it is immune to the form of the multivariate return distribution, including the distribution's factor structure. Putting our estimator to work, we find that a unique positive SDF prices all U.S. asset classes and satisfies the Hansen/Jagannathan variance bound. In contrast, the Chinese and Indian equity markets do not share the same SDF and hence do not seem to be integrated.

Authors' Coordinates	
Pukthuanthong[1]	Roll
Department of Finance University of Missouri Columbia, MO 65211	Caltech Pasadena, CA 91125 & Emeritus, UCLA
573-884-9785	626-395-3890
pukthuanthongk@missouri.edu	rroll@caltech.edu

Acknowledgement

For many helpful comments and suggestions on previous drafts, we are indebted to Viral Acharya, Tobias Adrian, Serena Agoro-Menyang, Yakov Amihud, Antonio Bernardo, Michael Brennan, Menachem Brenner, Mikhail Chernov, Bhagwan Chowdhry, John Cochrane, Harry DeAngelo, Eric DeBodt, Federico Echenique, Robert Engle, Michael Ewens, Wayne Ferson, Steve Figlewski, Ben Gillen, Lars Hansen, Raymond Kan, Shiki Levy, Francis Longstaff, John McConnell, Hugh McCulloch, Tyler Muir, Andrew Patton, Michael O'Doherty, Stephen Ross, Grigory Vilkov, Kenneth Winston, Rui Yu, Goufu Zhou, and participants in the Caltech and UCLA Brown Bag seminars, the Amundi Paris Dauphine annual workshop and the Stern School Volatility Institute Seminar at New York University.

[1] Corresponding author

We propose a new estimator of the Stochastic Discount Factor (SDF) inspired by a recognition that the SDF appears in a particular mathematical object, an integral equation. The solution to this integral equation makes our proposed estimator novel in several respects. First, it does not depend on macroeconomic proxies or preferences, unlike estimators in the previous literature. It is constructed from observed asset returns only, which is why we call it an "agnostic" estimator. Second, in contrast to typical portfolio applications such as mean/variance analysis, our estimator requires an unconventional condition: the number of assets must exceed the number of time period observations, which allows for a broader application to shorter time samples. Third, although the estimator is a function of observed returns, it does not depend on the distribution of returns. It works for single- and multi-factor data, for any asset class, and for thick or thin tails. Finally, it is immune to the grouping of assets. If N assets share a common SDF, the SDF estimator will be statistically indistinguishable when derived from the N assets as a whole or from subsets of size N/2, N/3, etc.

The final attribute above opens the way for checking whether naturally grouped assets, such as those comprising different asset classes or different countries, share the same SDF and are thus inhabiting a single integrated market. For example, we can (and will) test whether bonds and stocks within the U.S. are traded in segmented markets and have reliably incompatible SDFs or whether, to the contrary, their markets are integrated and their estimated SDFs are the same. This same test can and will be conducted here for other natural asset groups. We can also examine whether the SDF is strictly positive, which implies the absence of arbitrage. [2]

Although our estimator is quite easy to compute, its sampling distribution is complex for reasons that will be fully explained below. Hence, we find it enlightening to conduct a battery of simulations, first to uncover its small sample properties and second to investigate the power of its tests of market integration when asset groups do and do not actually share a common SDF. We find that the estimator is a profligate user of data. To achieve acceptable power, it requires a large

[2] Kan and Zhou (1999) argue that the SDF is more subject to estimation error than other approaches and that it provides inherently weak power in empirical tests. Cochrane (2001b) argues that these problem can be overcome by adding factor moment conditions and conducting joint estimation.

number of assets (up to 1,000) and a smaller but still sizeable number of time periods (more than 200). Such data are available for equity markets in many countries, but other asset classes are not so bountiful and hence some tests could have weak power.

Putting our estimator to practical use in this paper, we investigate market integration over a limited time period, a recent decade, and mainly for US assets, stocks, bonds, real estate, currencies and commodities. Such tests are far from comprehensive, of course, and we cannot draw general conclusions about other decades or countries. For US assets in a recent decade, we find that a unique positive SDF prices them all; i.e., US markets are integrated regardless of the asset class. We also test integration for two dissimilar but adjacent Asian markets, India and China, and statistically reject the proposition that they share the same SDF. This implies that they are segmented, at least during our sample time period.

Before presenting a formal derivation of the estimator, we embark on a brief literature review that places in context with respect to previous work. This is followed by sections containing the derivation, remarks on the sampling distribution, simulations, and finally empirical applications. A summary concludes.

I. Previous SDF Literature

The Stochastic Discount Factor (SDF) has become a dominant paradigm in recent asset pricing literature. For example, Ferson (1995) shows how the main asset pricing results (mean/variance efficiency, multi-beta models) are special cases of the basic SDF relation. Cochrane (2001a) begins with the SDF relation in chapter 1 and expands it into almost all other known models of assets. Exactly the same foundation is established in the first chapter of Singleton (2006) and exploited to study asset price dynamics. Campbell (2014) ordains the SDF as "The Framework of Contemporary Finance," (p. 3) in his essay explaining the 2013 Nobel Prizes awarded to Fama, Hansen, and Shiller. Excellent reviews are provided by Ferson (1995) and Cochrane and Culp (2003).

The empirical success of the SDF approach is less apparent. In many previous empirical applications, the SDF is proxied by a construct that depends typically on aggregate consumption, but occasionally on some other macroeconomic quantity, combined with a risk aversion parameter. For example, Cochrane (1996) employs aggregate consumption changes along with power utility (and a particular level risk aversion) to measure the SDF. Despite giving this specification every empirical benefit of the doubt, Cochrane (2001a, p. 45) admits that it still "…does not do well." A similar imperfect fit between consumption changes, over various horizons, and both equities and bonds, is reported by Singleton (1990).

Lettau and Ludvigson (2000) add in macro variables such as labor income and find that the deviation in wealth from its shared trend with consumption and labor income has strong predictive power for excess stock returns at business cycle frequencies, thereby suggesting that risk premia vary countercyclically. Chapman (1997) adds technology shocks and a battery of conditioning variables, transforming them with orthogonal polynomials, which serve to eliminate the small firm effect but still produce "statistically and economically large pricing errors", (p. 1406.) Da and Yun (2010) employ electricity generation as a proxy for aggregate consumption.[3] Adrian, Crump and Moench (2013) employ an exponential function of a grouping of state variables, which are themselves principal components of Treasury bond returns.

In research published just prior to the hegemony of the SDF paradigm, Long (1990) shows that a "Numeraire" portfolio has many similar properties. Long's Numeraire portfolio η has strictly positive gross returns $(1+R_\eta)$ and exists only if there is no arbitrage within a list of assets from which it is composed. In this case, the expected value of the ratio $(1+R_j)/(1+R_\eta)$ is unity for all assets j on the list, which implies that $1/(1+R_\eta)$ is essentially the same as the modern SDF. Long notes that the Numeraire portfolio is also the growth optimum portfolio. The latter is examined by Roll (1973) who provides an empirical test of whether the expected ratio above is the same for all assets. (He does not find evidence against it.)

Recognizing that aggregate consumption changes are too "smooth" to be well connected with asset prices (Mehra and Prescott [1985)]) and that consumption is likely measured with

[3] See also the variety of specifications discussed by Cochrane and Hansen (1992) in section III, "Other Candidate Discount Factors."

significant error (Rosenberg and Engle [2002]), recent literature avoids aggregate consumption data. In addition to Rosenberg and Engle, such an approach is taken by Aït-Sahalia and Lo (1998, 2000), and Chen and Ludvigson (2009). However, as pointed out by Araujo, Issler, and Fernandes (2005) and Araujo and Issler (2011), the above scholars still find it necessary to impose what might be considered rather ad hoc restrictions on preferences.

Nagel and Singleton (2011) estimate the SDF as a conditionally affine function of a set of priced risk factors in order to evaluate conditional asset pricing models. Korteweg and Nagel (2016) employ an SDF valuation method for assessing venture capital performance. Using the SDF framework, Kozak, Nagel, and Santosh (2016) suggest that behaviorally induced mistakes in asset pricing cannot easily be uncovered by focusing on the covariances of asset returns.

Hansen and Jagannathan (1991) avoid the specification of preferences and are still able to develop their famous bound on the mean and volatility of the SDF, given that SDF is unique. Campbell (1993) surmounts the annoyance with various approximations of nonlinear multiperiod consumption and portfolio-choices. He develops a formula for risk premia that can be tested without using consumption data and suggests a new way to use imperfect data about both market returns and consumption.

Araujo, Issler and Fernandes (2005, hereafter AIF) get around these difficulties by noting that the SDF should be the only serial correlation common feature of the data in the sense of Engle and Kozicki (1993). Then, by exploiting a log transform of returns, they derive a measure of the SDF that does not depend on a macroeconomic variable (notably including the problematic aggregate consumption) and also avoids the imposition of preferences.

Araujo and Issler (2011, hereafter AI) take a similar tack, noting via a logarithmic series expansion that the natural logarithm of the SDF is the only common factor in the log of all returns. Thus, the log SDF can be eliminated by a simple difference in returns. Essentially, the log SDF represents the (single) common APT factor in the sense of Ross (1976).

In both AIF and AI, the SDF measure is a function of average arithmetic and geometric asset returns. AIF compute their measure empirically and report its temporal evolution along with various statistical properties. They also compare it to the time series of riskless returns. AI find that relatively low risk aversion parameters are consistent with their estimated SDFs. They are able to price some stocks successfully, but not stocks with low capitalization levels.

Both AIF and AI essentially assume that the SDF is unique, rejecting that proposition only indirectly in the case of AI with low cap stocks. Our primary goal is to develop tests that offer an opportunity to directly reject SDF uniqueness. Our SDF estimator, which we exploit to develop such tests, does not depend on a factor model or a logarithmic approximation, or any other structural condition. Also, it works regardless of the multivariate distribution of returns, whatever its form, provided that certain lower order moments exist.[4]

II. An Agnostic Estimator for the SDF

This section first shows (in sub-section II.A) how SDFs can be approximated by a transformation of returns, without any additional information about preferences, consumption or other macro-economic data. The following sub-section (II.B) proves that the same SDF estimator arises naturally from minimizing a particular sum of average surprises. This development allows us to infer some useful properties of the SDF estimator. Sub-section II.C provides demonstrations of concept; using simulations, we illustrate a perhaps surprising fact that our proposed estimator works well regardless of the underlying distributions of returns including their factor structure. Finally, sub-section (II.D) proposes a battery of tests of SDF uniqueness using the SDF estimator derived in II.A and II.B.

II.A. Estimating the SDF from Returns Alone

Let $p_{i,t}$ denote the cash value of asset i at time t. When markets are complete, the SDF paradigm implies the existence of a unique m_t, the SDF, such that

$$E_{t-1}(\tilde{m}_t \tilde{p}_{i,t}) = p_{i,t-1} \qquad \forall i,t.^{5} \tag{1}$$

Denoting a gross return between t-1 and t by $R_{i,t} \equiv p_{i,t}/p_{i,t-1}$, equation (1) is the same as

[4] We explain the required moments below.

[5] For a representative agent, m is the discounted future marginal utility of consumption divided by the current marginal utility of consumption. The tilde denotes a random variable as of period t-1.

$$E_{t-1}(\tilde{m}_t \tilde{R}_{i,t}) = 1 \qquad \forall i,t. \text{ [6]} \tag{2}$$

In the financial economics SDF literature where (2) was originally derived, the SDF is invariably considered a function of preferences and consumption, often as held by a representative investor. From a mathematical perspective, however, equation (2) is a special case of an Integral or Volterra equation and, indeed, it is the simplest of these, called a "Fredholm equation of the first type."[7] As such, $\tilde{m}_t$ is simply an unknown mathematical function, not necessarily related to anything economic including the probability distribution of R. This is why it should be obvious that estimators of the SDF need not depend, in any way, on the multivariate form of the distribution of returns.

Integral equations often do not have analytic solutions, so mathematicians and physicists solve them numerically, typically by a "quadrature rule" whereby a system of equations with an equal number of unknowns provides a set of discrete values for the unknown function, which in our application would be some set of observations m_t for, say, t=1,…,T. We are proposing an analogous approach, discretizing as usual but with an over-identified system whose solution is rendered unique by a statistical restriction on the error of estimation.

We begin by noting that the expectation in (2), must correspond to a realization at time t; i.e.,

$$m_t R_{i,t} = E_{t-1}(\tilde{m}_t \tilde{R}_{i,t}) + \varepsilon_{i,t} \tag{3}$$

where $\boldsymbol{\varepsilon}_{\mathbf{i,t}}$ denotes the (complete) surprise in the mR product for asset i in period t. For each time period t, the realization in (3) is determined by whatever state occurs among the many encapsulated

[6] Equation (2) is the only moment condition required by SDF theory. However, the basic SDF relation applies similarly to multiple periods; e.g., $E_t(\tilde{m}_{t+\tau} \tilde{R}_{i,t+\tau}) = 1$ for $\tau>1$ where the gross return spans τ periods and m involves marginal utilities of consumption separated by τ periods. This could provide some interesting features involving a term structure of SDFs but we do not explore that possibility in this paper.

[7] We are grateful to Francis Longstaff for pointing out this isomorphism. See also Polyanin and Manzhirov (1998). It is implied in McCulloch (2003) who shows that the SDF (or "pricing kernel") has finite payoffs even when returns follow stable laws whose second moments are infinite.

in the expectation (2). The surprise is complete if expectations are rational; i.e., if agents can freely change their expectation in response to new information.

Since there is a state realization for each time t, over T time periods, we have, from (3) and (2),

$$\frac{1}{T}\sum_{t=1}^{T} m_t R_{i,t} = \frac{1}{T}\sum_{t=1}^{T}\left[E_{t-1}(\tilde{m}_t \tilde{R}_{i,t})\right] + \frac{1}{T}\sum_{t=1}^{T}\varepsilon_{i,t} = 1 + \frac{1}{T}\sum_{t=1}^{T}\varepsilon_{i,t} \cong 1 \ . \tag{4}$$

where the approximation indicates that the average surprise is not exactly zero in a finite sample, though it should vanish as $T \to \infty$.

The approximation error in (4) equals the time series sample mean of the surprises in the SDF-gross return product, a mean for asset i which we hereafter denote

$$\overline{\varepsilon}_i \equiv \frac{1}{T}\sum_{t=1}^{T}\varepsilon_{i,t} \ .$$

Rational expectations rules out any serial dependence in the surprises,

$$\mathrm{Cov}(\varepsilon_{i,t}, \varepsilon_{i,t-j}) = 0, \ j \neq 0$$

but the surprises could be heteroscedastic. Hence,

$$\mathrm{Var}(\overline{\varepsilon}_i) = \frac{1}{T^2}\sum_{t=1}^{T}\mathrm{Var}_t(\varepsilon_{i,t}) = \frac{1}{T}\overline{\sigma}_i^2$$

where $\overline{\sigma}_i^2$ denotes the mean variance of surprises for asset i over the particular sample period, t=1,…T. Unless the mean variance is growing without bound, the approximation error should disappear as T grows larger.

Now consider a sample of N assets with <u>simultaneous</u> observations over T periods, with N > T. The ensemble of gross returns for the N assets can be expressed as a matrix **R** (hereafter boldface denotes a matrix or vector). There are N columns in **R** and the ith column is $[R_{i,1}:\ldots:R_{i,T}]'$. We also need a column vector **m** ≡ $[m_1:\ldots:m_T]'$ to hold T realized values of the SDF and a N-element column unit vector **1** ≡ [1:…:1]'. The entire SDF ensemble of realizations for all assets and periods can then be written compactly as

$$\mathbf{R'm}/T \cong \mathbf{1}. \tag{5}$$

Pre-multiply (5) by **R**, to obtain

$$(\mathbf{RR'})\mathbf{m}/\mathrm{T} \cong \mathbf{R1}.$$

Since we have chosen N > T, the cross-sectional time-product matrix **RR'** is non-singular unless there are two periods with linearly dependent cross-sectional vectors of returns.[8] Hence, we can usually solve for a time-varying vector of estimated stochastic discount factors as

$$\mathbf{m}/\mathrm{T} \cong (\mathbf{RR'})^{-1}\mathbf{R1}. \tag{6}$$

N.B.: It is very important to emphasize that our solution (6) absolutely requires the number of assets to exceed the number of time periods; i.e., N > T. Many comments on earlier drafts make it clear that this condition, which is unusual and perhaps unprecedented in finance, is hard to grasp. Yet it is essential. It is not possible to uncover a unique vector of SDF realizations if T > N, which is the familiar condition in most other contexts, such as computing non-singular covariance matrices. We MUST have N > T to obtain a unique ***m***. *We hasten to add that this is merely a sample requirement and hence is easy to satisfy; e.g., by reducing T until it falls below N. The condition does not imply anything egregious such as the existence of an arbitrage because we are simply estimating T sample realizations of m, not the entire state space of m in each time period t.*

Hansen and Jagannathan (1991, p. 233) derive an expression that appears similar to (6), but the resemblance is superficial.[9] Their expression involves a covariance matrix of payoffs (or returns). Our **RR**' is not a covariance matrix. They note that their solution involves the first and second moments of the future payoffs and prices. If **RR**' above were diagonal, equation (6) would also involve first and second moments but in this case the (sample) moments would be the cross-sectional mean return in each period divided by the cross-sectional mean of the individual squared returns in that period.

[8] That is, unless the return of every individual asset in a given period is a linear function of the return on that asset in another period, (not that the returns are linearly dependent relative to each other in a given period.)

[9] The Hansen/Jagannathan approach is implemented for performance measurement by Chen and Knez (1996) and is further refined by He, Ng, and Zhang (1999.)

Collecting individual asset sample mean surprises in a column N vector, $\bar{\boldsymbol{\varepsilon}} = (\bar{\varepsilon}_1 : \ldots : \bar{\varepsilon}_N)'$, the approximation error in (6) is equal to

$$(\mathbf{RR}')^{-1}\mathbf{R}\bar{\boldsymbol{\varepsilon}}. \tag{7}$$

This error is not exactly zero because, for each t, there are related components in **R** and $\bar{\boldsymbol{\varepsilon}}$. For very large N and T, these components should become immaterial, but they add sampling error to the estimated SDFs with smaller N and T. We investigate the consequences in the next sub-section after presenting an alternative approach for deriving the same estimator.

II.B. The Minimum Sum of Squared Average Surprises

The exact form of equation (5), (i.e., with no approximation), is

$$\mathbf{R}'\mathbf{m}/\mathrm{T} = \mathbf{1} + \bar{\boldsymbol{\varepsilon}} \tag{8}$$

where $\bar{\boldsymbol{\varepsilon}}$ is the column N vector that contains the average surprises for each asset. A least squares estimator for **m** is available by minimizing the sum of squared average surprises with respect to **m**; i.e.,

$$\min_{\mathbf{m}}[(\bar{\boldsymbol{\varepsilon}}'\bar{\boldsymbol{\varepsilon}}) = (\mathbf{R}'\mathbf{m}/\mathrm{T} - \mathbf{1})'(\mathbf{R}'\mathbf{m}/\mathrm{T} - \mathbf{1}) \ .$$

The first-order condition is

$$\frac{\partial}{\mathbf{m}}(\mathbf{m}'\mathbf{RR}'\mathbf{m}/\mathrm{T}^2 - 2\mathbf{m}'\mathbf{R1}/\mathrm{T}) = 2\mathbf{RR}'\mathbf{m}/\mathrm{T}^2 - 2\mathbf{R1}/\mathrm{T} = \mathbf{0}$$

and the extremum is achieved for the $\hat{\mathbf{m}}$ that satisfies

$$\hat{\mathbf{m}}/\mathrm{T} = (\mathbf{RR}')^{-1}\mathbf{R1} \tag{9}$$

which shows that $\hat{\mathbf{m}}$ is the approximation (6) in section II.A. The second order condition is strictly positive because **RR'** is positive definite (by assumption); hence $\hat{\mathbf{m}}$ provides the minimum sum of squares for the average SDF surprises.

One may legitimately question why the estimator in (6) or (9) should involve a cross-sectional sum of returns (**R1**) in each period. Actually, this is dictated by the mathematical fact that the basic SDF equation (2) has a 1.0 on the right side for every asset. In the robustness section

B.4 in Appendix B, however, we do consider an alternative, a precision weighted sum as opposed to a simpler sum.

Another possible question might occur to some readers in that the estimator seems to use information across all observed time periods even though for any given period t within T, the SDF is a random variable. But the answer is simply that we are estimating the best fit to the entire vector $\hat{\mathbf{m}}$ whose elements have already occurred as realizations of the random variable at each t. There is an associated surprise each t as well and the estimator simply minimizes the sum of squared average surprises.

The least squares estimator in (9) differs from a standard regression estimator in one important respect; since the "dependent" variable here is the T element unit vector, (with every element a constant 1.0), there could be a connection between **R** and $\bar{\varepsilon}$, which would violate the customary spherical regression assumptions. Consequently, the estimator could be biased. There is indeed a linear connection between the **R'm** product and $\bar{\varepsilon}$ but this is slightly different than the source of typical regression bias induced by linear dependence of the disturbances and explanatory variables.

To elucidate this issue, solve (8) for **1** and substitute the result in (9), which simplifies to,

$$\hat{\mathbf{m}} - \mathbf{m} = -\mathrm{T}(\mathbf{RR}')^{-1}\mathbf{R}\bar{\boldsymbol{\varepsilon}} \ .$$

The expected value of this expression is the bias. Expanding $\mathbf{R}\bar{\boldsymbol{\varepsilon}}$ term by term, we observe that most elements are innocuous and close to zero because they involve products such as ($\varepsilon_{j,t}$ $R_{i,t-k}$) for i≠j and k≠0. However, there are a few elements that are unlikely to disappear. For period t, there is

$$R_{1,t}\varepsilon_{1,t} + R_{2,t}\varepsilon_{2,t} + \ldots + R_{N,t}\varepsilon_{N,t} = m_t(R_{1,t}^2 + R_{2,t}^2 + \ldots + R_{N,t}^2) - \sum_{j=1}^{N} R_{j,t}$$

and there are similar terms for other periods. We will study the extent of the resulting bias in the next section using simulation but note already that the bias terms are atypical because the dependence between the explanatory variables (the R's) and the disturbances (the ε's) is not linear.

Despite its possible bias, the estimator in (9) shares some attractive features with OLS regression estimates. In particular, it can be used to define residuals, estimates of the true disturbances, as[10]

$$\hat{\hat{\boldsymbol{\varepsilon}}} = \mathbf{R}'\hat{\mathbf{m}}/\mathrm{T} - \mathbf{1} = \mathbf{R'(RR')^{-1}R1} - \mathbf{1} = -[\mathbf{I\text{-}R'(RR')^{-1}R}]\mathbf{1} \tag{10}$$

The matrix in brackets in (10) is idempotent, so the sum of squared residuals divided by the degrees-of-freedom, N-T, is

$$\frac{\hat{\hat{\boldsymbol{\varepsilon}}}'\hat{\hat{\boldsymbol{\varepsilon}}}}{\mathrm{N\text{-}T}} = \frac{\mathbf{1}'[\mathbf{I}-\mathbf{R}'(\mathbf{RR}')^{-1}\mathbf{R}]\mathbf{1}}{\mathrm{N\text{-}T}} = \frac{\mathrm{N}}{\mathrm{N\text{-}T}} - \frac{\mathbf{1}'[\mathbf{R}'(\mathbf{RR}')^{-1}\mathbf{R}]\mathbf{1}}{\mathrm{N\text{-}T}} \tag{11}$$

For a large enough N, (definitely for N>2T), the mean squared residual in (11) declines with N, holding T constant.[11] Consequently, the quality of our SDF estimator should be better when N is large relative to T; i.e., when there are at least twice as many assets as time periods. The square root of (11) gives the standard error of the estimate,

$$\mathrm{s} \equiv \sqrt{\hat{\hat{\boldsymbol{\varepsilon}}}'\hat{\hat{\boldsymbol{\varepsilon}}}/(\mathrm{N}-\mathrm{T})}\ .$$

The covariance matrix of the estimated SDFs is given by

$$\mathrm{E}[(\hat{\mathbf{m}}-\mathbf{m})(\hat{\mathbf{m}}-\mathbf{m})'\,|\,\mathbf{R},\mathrm{T}] = (\mathbf{RR}')^{-1}\mathbf{R}\mathrm{E}(\mathbf{V}_{\boldsymbol{\Sigma\varepsilon}})\mathbf{R}'(\mathbf{RR}')^{-1} \tag{12}$$

where the (N X N) symmetric matrix $\mathbf{V}_{\boldsymbol{\Sigma\varepsilon}}$ has the following element in the j^{th} row and k^{th} column:

$$(\varepsilon_{j,1}+\varepsilon_{j,2}+\ldots+\varepsilon_{j,T})(\varepsilon_{k,1}+\varepsilon_{k,2}+\ldots+\varepsilon_{k,T})\ .$$

Unlike the analogous covariance matrix of disturbances in standard OLS regressions, the diagonal elements of $\mathbf{V}_{\boldsymbol{\Sigma\varepsilon}}$ are not necessarily equal to each other and the off-diagonal elements

[10] Unlike the true disturbances, the residuals in (10) are orthogonal to **R**.

[11] Proof: The second term on the right side of (11) can be written as $\frac{\mathrm{N}^2}{\mathrm{N}-\mathrm{T}}\overline{\mathbf{R}}'(\mathbf{RR}')^{-1}\overline{\mathbf{R}} \equiv \frac{\mathrm{N}^2}{\mathrm{N}-\mathrm{T}}\boldsymbol{\Psi}$ where $\overline{\mathbf{R}}$ is the T element column vector whose t^{th} element is the cross-sectional mean gross return in period t. The positive quadratic form $\boldsymbol{\Psi}$ does not depend directly on N, so $\frac{\partial}{\mathrm{N}}\left[\frac{\mathrm{N}^2}{\mathrm{N}-\mathrm{T}}\boldsymbol{\Psi}\right] = \left[1-\left(\frac{\mathrm{T}}{\mathrm{N}-\mathrm{T}}\right)^2\right]\boldsymbol{\Psi}$, which is positive for N>2T, at which point both terms in (11) decline with N; QED.

need not have zero expectation. However, we can safely assume that cross-products separated in time, such as $\varepsilon_{j,t}\varepsilon_{k,\tau}$ for $t \neq \tau$, are zero; otherwise, the ε's would not be surprises. This implies that the element in the jth row and kth column of $\mathbf{V}_{\Sigma\varepsilon}$ reduces to $\sum_{t=1}^{T}\varepsilon_{j,t}\varepsilon_{k,t}$. Moreover, if the ε's are not correlated across assets, an arguably dubious condition, this sum has an expected value of zero for j≠k and then E($\mathbf{V}_{\Sigma\varepsilon}$) becomes diagonal and equal to $\mathbf{I}\boldsymbol{\sigma}^2_{\Sigma\boldsymbol{\varepsilon}}$ where I is the identity matrix and $\boldsymbol{\sigma}^2_{\Sigma\boldsymbol{\varepsilon}}$ is the N element column vector whose jth element is $\mathrm{Var}(\sum_{t=1}^{T}\varepsilon_{j,t})$. If the variance of the surprises were the same scalar σ^2 for all assets and time periods, perhaps an even more dubious condition, then (12) simplifies further to

$$E[(\hat{\mathbf{m}}-\mathbf{m})(\hat{\mathbf{m}}-\mathbf{m})'\,|\,\mathbf{R},T] = T\sigma^{\mathbf{2}}(\mathbf{RR}')^{-1} \tag{13}$$

Except for the presence of T, this is the standard regression covariance matrix of the coefficients given IID disturbances.

The square roots of the T diagonal elements of (12) or (13) provide the standard errors of the SDFs period-by-period. We will examine their properties using simulation in the next section. One pertinent property is obvious already, however. For a fixed number of assets, N, the standard errors of estimated SDFs increase with the time series sample size, T. Thus, we anticipate that our estimator will perform better when N-T is large.

II.C. Demonstrations of Concept

We have learned from comments on earlier drafts and in presentations that our proposed estimator is sometimes understood, erroneously, to be a projection on observed sample returns. Such intuition is understandable because the estimator does employ returns; hence, one could easily fall into the mistaken notion that the estimator is akin to a sample mean/variance efficient portfolio, which, of course, is composed differently across various sub-samples of assets.[12]

[12] Such intuition is readily overturned by thinking about the SDF as an unknown function in an integral equation; see section II.A. However, financial economists are not accustomed to thinking in such terms.

But a close examination of our estimator belies such intuition. Instead of a projection on asset returns, it is actually a projection on time periods. As a consequence, it is unaffected by the distributions of returns or even by their identity as long as a unique SDF prices all assets in the cross-section.

To demonstrate this fact, we resort to simulations since they subsume the potential sampling problems discussed in the previous sub-section. We show first that the estimator performs almost perfectly when the sampling noise is small. We then show that the estimator is immune to differences in the distributions of returns and extracts indistinguishable estimates of the SDF even from sub-samples of assets with different factor structures. In this sub-section, we briefly explain the simulations and report the results. Details about all simulations in the paper, including generating equations and parameters, are provided in Appendix A.

Assuming that the SDF unique, we generate "true" SDF realizations with a mean equal to the reciprocal of the gross riskless interest rate, as the SDF paradigm stipulates, and with a given level of time series variation about the mean. (See Appendix A, equation A-1.) We then independently simulate gross returns so that their product with the true SDF averages to unity over a specified sample period; we then add noise to each return observation with a random perturbation, (A-3.) Finally, using the resulting noisy sample returns, we calculate our SDF estimator and compare it with the known "true" SDF.

Our first illustration of concept uses 120 assets and 60 time periods, (a modest degrees-of-freedom according to section I.B), a riskless rate of .4% per period, and a true SDF standard deviation of 4% per period. Initial returns have means of .8% per period (mean gross returns of 1.008) and standard deviations of 8% per period, a material level of return volatility. However, the standard deviation of the perturbation, ϑ in (A-3), is intentionally small, .01% per period.

The final returns (after making sure the means of the SDF-Return product is 1.0 on average) still have substantial volatility. Their average standard deviation is 8.1% over the 120 simulated assets with a minimum (maximum) individual asset standard deviation of 6.17% (11.2%).[13]

[13] The minimum (maximum) individual return is -25.3% (35.6%).

Figure I plots the resulting estimated SDFs against the true SDFs for the 60 time periods. Their difference is trifling. Their correlation is 0.99946 and they are aligned with each other almost perfectly. This illustrates that the theoretical bias discussed in section I.B is empirically trivial when the sampling perturbations are minor.

In reality, of course, returns are correlated with one another and conceivably have heterogeneous factor structures across asset classes. For example, bond returns could be driven by different risk factors than equity returns. Nonetheless, if a unique SDF prices all asset expected returns in the cross-section, the basic SDF equation (1) is valid with the same SDF for all assets.

To consider this situation, we provide a further demonstration of concept by simulating returns that are not only correlated but also have diverse factor structures. In this simulation, we presume that there are two asset classes that share a common factor but that the second asset class is also driven by a second factor that has no influence on the first asset class; (See Appendix A, section A.1.a.)

Figure II, Panels A and B (for two different levels of return perturbation), plots the estimated SDF against the true SDF in the left chart and the SDFs estimated for the two groups against each other in the right chart. As the figure shows, there is sampling variation, but the recovered estimate of the SDF is close to the true SDF and the estimated SDFs from the two divergent (by factor structure) are close to one another.

Finally, we provide another simulation in which two asset groups have completely different factor structures; (Appendix A, Section A.1.b.) There are two factors driving the returns on both groups but the factors themselves are independent of each other across groups. Figure III shows the results. In this illustration, we use the higher level of return perturbation from Figure II.

Again, despite the fact that the factors are entirely different in the two asset groups, there is a strong connection between the true and estimated SDFs and between the SDFs estimated from the two groups. This illustrates our contention that the distributions of returns are inconsequential for our SDF estimator provided that the true SDF is unique and prices all assets regardless of groupings.

Some might find these results quite surprising because our SDF estimator is unaffected by the return distribution. This could be particularly hard to fathom because a competing construct, a sample mean/variance efficient portfolio, also perfectly prices returns in the cross-section, but it obviously depends on the distribution of returns and has a different composition for various groups of assets. But examining carefully the basic SDF equation (2) reveals why our estimator is so robust. Equation (2) is a mathematical object, an integral or "Volterra" equation, that says nothing about the distribution of returns other than the product of each return and the SDF has an expected value of unity. Consequently, every expected return obeys the same cross-sectional linear function of the covariance between the return and the SDF. So long as the first moment of the SDF/return product is finite and the SDF is unique, its estimator needs not be troubled by any other property of the multivariate distribution of returns.

II. D. Using the SDF Estimator to Assess Market Integration

The vector on the right side of (9) is an estimate based on N assets and a sample period of length T, a combination of cross-sectional and time series observations. SDF paradigm also contends that any **other** set of assets within the same integrated market should produce, aside from sampling variation, the same $\hat{\mathbf{m}}$ from concurrent time series observations. Hence, if we denote by $\hat{\mathbf{m}}(k)$ a sample $\hat{\mathbf{m}}$ computed according to (9) (where k indicates a set of K assets, K>T) and then, <u>from the same calendar observations</u>, choose a complement set $j \not\subset k$ with J assets (and J > T), the SDF null hypothesis of market integration can be expressed as

$$H_0\text{: } E[\hat{\mathbf{m}}(k) - \hat{\mathbf{m}}(j)] = \mathbf{0}. \tag{14}$$

Notice that K and J need not be equal, but both must be larger than T.

This test is reminiscent of DeSantis (1993) and Ferson (1995), who suggest comparing SDFs derived from a subset of assets to SDFs derived from all available assets. Testing for the equivalence of pricing operators across two groups of assets is also explored by Chen and Knez (1994)[14] and, in the context of the APT, by Brown and Weinstein (1983).[15]

It is important to emphasize that the philosophy of the above test is standard; i.e., we will never be able to prove that two SDFs are exactly the same and that compared markets are indeed completely integrated, but we do have the possibility of rejecting these implications. If markets are not complete and integrated, an infinite number of stochastic discount factors satisfy equation (1) because $E_{t-1}[(\tilde{m}_t + \tilde{\omega}_t)\tilde{p}_t] = E_{t-1}(\tilde{m}_t \tilde{p}_t)$ whenever ω and p are orthogonal; Cf. Cochrane (2001a, section 4.1). But $\tilde{m}_t + \tilde{\omega}_t$ looks just like the true SDF plus an estimation error. Indeed, if markets are complete, $\tilde{\omega}_t$ **is** an estimation error because $\tilde{m}_t$ is unique. On the contrary, if markets are incomplete $\tilde{\omega}_t$ can differ across groups of assets and hence the null hypothesis in (14) can potentially be rejected.

Many standard tests of equality could be employed for equation (14). For example, the Hotelling (1931) T^2 test could check whether the means of $\hat{\mathbf{m}}(k)$ and $\hat{\mathbf{m}}(j)$ are statistically indistinguishable. The non-parametric Kruskal-Wallis (1952) test (hereafter KW) is designed for this purpose and will reject the null hypothesis if $\hat{\mathbf{m}}(k)$ stochastically dominates $\hat{\mathbf{m}}(j)$ or vice versa. This also provides a test of the equality of medians.

It might be sensible to conduct tests with assets that seem unlikely, *a priori*, to share the same SDF, such as equities in one group and bonds in another (over the same sample period, of course) or perhaps equities in two different countries. This would represent a tougher hurdle for the SDF paradigm but any viable concept should be able to surmount the most severe examination possible.

There is no reason to restrict our attention to just two sets of assets. Every vector computed according to (9), drawn from concurrent calendar observations but with different assets, should be

[14] Chen and Knez (1995) derive a measure of market integration as the minimal amount that two pricing operators differ. They use a similar framework to develop a general approach to portfolio performance measurement in Chen and Knez (1996).

[15] The Arbitrage Pricing Theory due to Ross (1976).

congruent. The Welch (1951) test (hereafter WE) would serve nicely to check whether the means of all such vectors are the same and the KW test can handle multiple comparisons of entire distributions. The Welch test is robust against heterogeneity in the variances of the distributions being compared. On the other hand, the non-parametric Brown/Forsythe (1974) test (hereafter BF) is designed specifically to check for unequal volatilities using absolute deviations.

The KW, WE and BF tests involve necessary conditions for the SDF paradigm. They can detect differences in, respectively, the medians, means and volatilities two estimated SDF vectors, but they are not capable of detecting time-dependent patterns of differences in the individual elements of the two vectors. For example, one vector might be increasing over time and the other decreasing but they could still have the same mean and variance.

The SDF paradigm stipulates not only that the location and volatility in SDFs are the same across groups of assets but also that SDF estimated realizations are the same in every time period. A sufficient condition is that the entire vectors $\hat{\mathbf{m}}(k)$ and $\hat{\mathbf{m}}(j)$ are congruent. Thus, we consider also a test that compares the two vectors element by element, a Hausman (1978) type Chi-Square test (hereafter CH.) [16]

To explain the Hausman type test in our application, let $\hat{\mathbf{m}}_{\mathbf{j,t}}$ and $\hat{\mathbf{m}}_{\mathbf{k,t}}$ denote the estimated SDF observation from asset groups j and k at time t. Under the null SDF hypothesis, they have the same expected value, μ, and a common standard deviation, σ_t. Their correlation is ρ_t. Note that the correlation is not perfect because these are estimates of m, not the true values.

Under the null hypothesis, the variance of $\hat{\mathbf{m}}_{\mathbf{j,t}}\text{-}\hat{\mathbf{m}}_{\mathbf{k,t}}$ is $\mathbf{2\sigma_t^2(1\text{-}\rho_t)}$. Consequently, the standardized variate,

$$z_t = \frac{\hat{m}_{j,t} - \hat{m}_{k,t}}{\sigma_t\sqrt{2(1-\rho_t)}}$$

has mean zero and unit variance. When z is not autocorrelated,

[16] We are indebted to Ben Gillen for suggesting this test.

$$\chi_T^2 = \sum_{t=1}^{T} z_t^2$$

converges asymptotically to a Chi-Square distribution with T degrees of freedom.[17]

The main implementation problem is, of course, that σ_t and ρ_t are unknown parameters that have to be estimated. Ignoring their time variation, this can be accomplished with the usual estimates over the sample of size T. However, since there are two estimated SDF vectors, even with this simplifying assumption there would be two different estimates of σ. The most straightforward and sensible expedient is simply to average the two.

This Chi-Square test is best suited for comparing the SDFs from two groups of assets, but it can be extended to multiple groups if we are willing to assume that the estimation error differences are independent across groups. Given this assumption, the null hypothesis is tested by computing the statistic above for all pairs of groups (each group's SDF being estimated over the same sample of time periods) and then using the Bonferroni correction of the type I error.

For example, suppose we have five groups, which implies ten pairs. Then $\chi_{i,T}^2$ is computed for each pair i by the formula above and compared with the $\alpha/10$ significance level, where α is the usual type I error (e.g., 5%). If <u>none</u> of the pairs have smaller p-values, there is no significant evidence against the null hypothesis. Alternatively, if just one pair has a p-value smaller than $\alpha/10$, the null is rejected.

The Bonferroni adjustment is known to be conservative in the sense that rejection of the null is less likely if there are any issues with the assumptions. In our case, the most likely issue would be dependence in the error differences across group pairs. For this reason and also to examine the asymptotic convergence of the Chi-Square test statistic, we subject it to a battery of simulation experiments.

By implementing all four of the tests just described, we should be able to ascertain whether two or more estimated SDFs have equal means, volatilities, display stochastic dominance or differ

[17] If the SDF estimates are normally distributed and independent across time, the Chi-Square distribution is exact for any sample size.

element by element. Violation of any one of the four tests would be evidence against SDF uniqueness.

Test power is a more difficult issue. As indicated in section I.B, power undoubtedly depends on the relative sizes of the time period, T, and the cross-sections, N. Unless the data are extremely high frequency, one usually has more assets than time periods. But in the present case, unlike with most asset pricing tests, this is an advantage. On the other hand, a large T, but not nearly as large as N, might sometimes confer an advantage because the time series sums of expectation surprises, (the $\varepsilon_{i,t}$'s in (3)) will compromise the accuracy of the SDF estimates for short time series. We investigate this issue in section III using simulated data.

Nothing above requires specification of a macro/preferences-related proxy for the SDF. Even a riskless rate, if there is one, whose gross return R_f satisfies the useful property, $E(m_t) = 1/R_f$, is not necessary. Moreover, tests can be conducted with relatively short time series samples, but still with the caveat that longer samples may be less prone to estimation error.

III. More About the Qualities of Our SDF Estimator

III. A. Comparing the Estimated SDF and the True SDF with an Extended Set of Parameters

To provide further insight about the performance of our SDF estimator, this sub-section offers a series of simulations to compare true SDFs with estimated SDFs. Extending the demonstration of concept discussed in Sub-section I.C above, we provide simulations for a wider set of parameters and sampling variation. The basic setup is identical to that in Sub-section I.C. Technical details are provided by Appendix A.

In all cases, we compare the true and estimated SDFs using two criteria, the simple correlation between $\mathbf{m}$ and $\hat{\mathbf{m}}$ and the Theil (1966) U_2 statistic. The latter is closely related to the mean square prediction error, (MSE). Specifically,

$$\mathrm{MSE} = \sum_{t=1}^{T} (m_t - \hat{m}_t)^2 / T \text{ , and}$$

$$U_2 = MSE / \left(\sum_{t=1}^{T} m_t^2 / T \right) .$$

The correlation is easy to understand but it can be a bit misleading because it fails to measure whether **m** and $\hat{\mathbf{m}}$ are congruent. For example, if $\hat{\mathbf{m}}$ =2**m**, the correlation would be perfect. An advantage of the MSE is that it can be decomposed into three components, one due to a difference in means, another to a difference in volatilities, and third due to a lack of correlation; i.e.,

$$MSE \equiv (\overline{m} - \overline{\hat{m}})^2 + (s_m - s_{\hat{m}})^2 + 2(1-\rho)s_m s_{\hat{m}} \tag{15}$$

where the superior bars indicate means, the s's are standard deviations and ρ is the correlation between **m** and $\hat{\mathbf{m}}$. This decomposition is particularly relevant in our application because we would expect $\hat{\mathbf{m}}$ to have more volatility than **m** due to sampling error and to be imperfectly correlated. However, when the SDF theory is true, the two means should be close to one another.

In simulations with different levels of sampling perturbations, we examine the relative influences of the time series and cross-sectional sample sizes, T and N, respectively, and also the impact of return perturbations, the volatility of the true SDF, and the risk-free rate. With this many parameters, it is hard to summarize results compactly over a continuum of parameter values, so we resort to a hopefully more illuminating expedient. We simply generate the simulated **m** and $\hat{\mathbf{m}}$ with several different choices of the parameters and then present summary linear regressions of the correlations and Theil's U_2 on all the parameters jointly.

Our estimator of the SDF requires N>T, so we let T=30, 60, 90, and 120 and for each T, we set N=240, 360, 480, and 960. These choices are made to roughly match sample sizes and numbers of assets in our later empirical work below. For each N and T, we let the true SDF volatility take the values σ_ξ = .5%, 1%, 1.5% and 2% per month. For each N, T, and σ_ξ, the perturbation volatility σ_ϑ takes on nine values beginning with $\sigma_\xi/5$ and increasing by this increment to terminate at $1.8\sigma_\xi$.[18] Finally, for each choice of the previous parameters, we let the risk-free rate vary as follows: R_F = .1%, .2%, .3%, .4% and .5% per month. This results in 2,880 different parameter combinations. For each parameter combination, we generate completely different true SDFs and returns and hence have independent sets of sample SDFs.

Table I gives the results, panel A for the correlation between $\mathbf{m}$ and $\hat{\mathbf{m}}$, and Panel B for Theil's U_2. In Panel A, we see that the correlation falls with T, rises with N, rises with σ_ξ, the volatility of the true SDF, and falls with σ_ϑ, the perturbation volatility, all with very high levels of significance. Each regression coefficient, of course, indicates the marginal influence holding constant other parameters. For the two volatilities, the directions are intuitively obvious because a greater spread of the true values and a smaller perturbation variance should improve the fit. For N and T, the fit seems related to the degrees-of-freedom, N-T, (remember, N>T). Fewer degrees-of-freedom result in less precise estimation. The riskless rate has no significance whatsoever; this too is hardly surprising because a simple translation of the mean SDF should essentially be immaterial.[19]

The results for Theil's U_2 in Panel B essentially agree with the results for the correlations in Panel A, with opposite signs as expected (since U_2 is larger when the fit is worse), except for the volatility of the true SDF, which has the same sign but less statistical significance. This exception might be explained by the fact that U_2 is scaled by a denominator that relates to the variance of the true SDF. The other three significant variables in panel A are even more significant in Panel B and the overall explanatory power is larger.

We find, after decomposing the MSE into its three components, (equation (15)), virtually no effect at all from the first component, a difference in means between the true and estimated SDFs. On average over the 2,880 combinations of parameters, the mean difference component's fraction of the total MSE has a value of 0.0000 and the largest value is only 0.0012. In contrast,

---

[18] The equations that contain these Greek symbols are displayed in Appendix A.

[19] In unreported results, we verify that this is also true of the mean and variance of the initial returns as generated by equation A-2 in Appendix A.

the averages of the standard deviation difference component and the correlation component are, respectively, 0.2426 and 0.7574 as fractions of the total MSE. (For each parameter set, the three fractional components sum to 1.0 by construction.) The largest and smallest values are, respectively .8346 and 0.000 (1.0000 and 0.1654) for the standard deviation difference component (correlation component.)

Each of the 2,880 parameter combinations uses a different simulated set of "true" SDFs, which results in a corresponding and different set of estimated SDFs. Consequently, we can compare the 2,880 means of true and estimated SDFs. They are very close. The averages over 2,880 sets are 0.9956 and 0.9960 for, respectively, the estimated and true SDF means. The standard deviations of the means across the 2,880 sets are, respectively, 0.2438 and 0.2439. Their correlation is 0.9977. Hence the mean of our estimator is close to the true mean SDF regardless of the parameters.

However, although the means are close, the period-by-period estimated and true SDFs display substantial divergence for some parameter combinations. The average correlation is .189 and the maximum and minimum correlations over the 2,880 parameter combinations are, respectively, 0.951 and -0.547. This makes it very clear that ill-considered parameters degrade the performance of our SDF estimator when there is a large amount of sampling variation.[20]

Panel C of Table I reports determinants of the time series standard deviation of the estimated SDFs. The impact of degrees-of-freedom (essentially N-T) is apparent; Larger N and smaller T reduce sampling error and result in a better-behaved estimated SDF. Holding N and T constant, more volatility in the return perturbation brings, not surprisingly, in a more volatile estimated SDF. The time series volatility of the true SDF, however, has no significant impact and neither does the riskless rate.

The variance of our estimated SDF <u>should</u> increase with T due to the approximation error. This is because the elements in the estimated SDF vector are equal to the right side of (6) multiplied by T. This multiplication converts the average approximation error to the sum of approximation

[20] For Theil's U_2, the mean, maximum and minimum are, respectively, 0.339, 0.787, and 0.0359. Larger values indicate more disagreement between the estimated SDF and the true SDF.

errors, (summed over T periods.) The standard deviation of this sum increases with $\sqrt{T}$. In an unreported alternative regression to Panel C in Table I, using $\sqrt{T}$ instead of T as a regressor, we find that virtually nothing is altered except the coefficient.[21]

In Panel D, of Table I, we finally see something that **is** influenced by the true riskless rate; viz., the implied riskless rate from the reciprocal of the estimated SDF. The t-statistic is 2.42, but the overall explanatory power is meager. Also, both the perturbation volatility and the volatility of the true SDF are marginally significant, which may be explained by Jensen's inequality (since the implied riskless rate is obtained from a reciprocal of an estimated SDF.)

III. B. Test Power

This sub-section provides evidence about the power of our proposed tests by tabulating type II errors under a variety of different simulated conditions. The type II error, often called the "power" of the test, is the probability of correctly rejecting a false null hypothesis. To estimate power, we must set up a simulation so that the true SDFs for different groups of assets are <u>not</u> the same. For two or more sets of assets, we then estimate SDFs and tabulate the rejection frequency of the null hypothesis that all SDF estimates are the same except for sampling error. In a simulation, the rejection frequency is the fraction of replications with test p-values less than the type I error.

The estimated SDF could differ for two distinct reasons. First, even though the basic SDF equation (1) holds for different groups of assets, the stochastic discount factor itself might have different distributions across groups; i.e., different means, volatilities, or other features. Second, the basic SDF equation might be false for one or more groups such that the expectation in (1) is not unity for such groups. We will examine both types of discrepancies in the simulations next.

To examine both types of possible violations indications of market non-integration, we use the four tests explained in Section II.D, both now in simulations and later in the empirical

[21] The t-statistic for $\sqrt{T}$ is 53.8 as opposed to the 53.9 reported for T in Table 1. Everything else is similarly close; e.g., the adjusted R-square is 0.730 as opposed to 0.731.

examinations of actual data. The tests are the Kruskal/Wallis (1952) (KW) non-parametric one-way analysis of variance based on ranks, which rejects a false null hypothesis if one or more sample SDFs is stochastically dominant or has an abnormal median, the Welch (1951) (WE) test for equal means, which allows for unequal variances, the Brown/Forsythe (1974) (BF) test for unequal variances, and the Chi-Square tests that estimated SDF vectors are the same element by element.

The relevant test depends on the nature of the difference among SDFs. For example, if the medians differ but the means and variances are about the same, the KW test should reject the null but the WE and BF test might not. Similarly, if the SDF distributions have similar location on the real line but have disparate volatilities, the BF test should reject but the other tests would not. If the SDF estimates have the same location and volatility but different time patterns, the Chi-Square test should work well.

If one or more asset groups is characterized by departure of the basic SDF expectation (1) from unity, all four tests could conceivably detect it. This suggests that simulations should examine various type of SDF heterogeneity; i.e., different locations or volatilities or both and perhaps differences in higher moments and also failure of the basic SDF equation (1). Obviously, we cannot hope to examine every possible type and size of differences across SDFs, so our simulations are unavoidably limited in scope. However, we will gladly supply the simulation Fortran code to anyone interested in examining power for other parameter choices or types of SDF failures.

To be most relevant for the empirical tests to follow, we perform power calculations for several choices of the most important parameters, which are the number of sample periods, T, the number of assets in each group, N, the means and variances of the true SDFs (which can differ across groups), the number of asset groups, and the volatility of return perturbations. For each choice of parameters, the simulations are replicated 1,000 times and the power is tabulated as the null hypothesis rejection frequency.

III. B. 1. Test Power when the SDF equation is true but the SDF differs across asset groups

In this subsection, we assume that the basic SDF equation (1) is valid for all assets but that the SDF itself differs across asset groups. Our first set of simulations has just two asset groups. Parameter combinations include N=240, 480, 720 and 960. For each N, T=30, 60, 90, and 120. To illustrate differences in the tests, we conduct a simulation with SDFs that differ only in location; i.e., two values for the riskless rate, .1% and 5% per period, but with the same SDF volatility, a standard deviation of 25% per period. A second simulation reduces the volatility to 5% per period. A third simulation has two SDFs with the same mean, R_F=.1%, but different standard deviations, 10% and 25%. In these simulations (and in all that follow), SDFs are generated according to the true SDF model in Appendix A equation (A-1), and returns are generated by (A-3). Note that the underlying SDFs are the same across groups but with differing means and/or volatilities.

Results for the first simulations, with differing SDF means but equal volatilities, test power is reported in columns 3-6 of Table II. Panel A (B) has a perturbation volatility of 1% (2%); see equation (A-3).[22] If a particular test has minimal power, it is not reported. Hence, only the Chi-Square test is reported in column 3 where the SDF volatility is 25%. KW, BF and WE have no power in this case. However, when SDF volatility is reduced to 5%, (columns 4-6) both KW and WE have very good power for all choices of N and T in Panel A and for N > 240 in Panel B. For these simulations with equal volatility, the BF test should not have any power, and it does not.

The CH test exhibits a complex pattern of power. In Panel A, we see that it has perfect power for N=720 and N=960 but for lower N its power declines dramatically with larger T. Evidently, its power is degraded when the degrees of freedom, i.e., N-T, is not sufficient. A similar pattern is observed in Panel B except that the power is uniformly lower and completely absent for lower N and higher T. For lower SDF volatility, columns 4-6, the CH test displays a very similar power pattern as for the higher volatility, column 3. From a power perspective, CH is dominated by KW and WE for lower SDF volatility.

---

[22] The perturbation interacts with other stochastic component to produce estimation error in the SDF, which is considerably more volatile than the perturbation itself.

Clearly, return perturbation volatility has a large deleterious impact on power, but it appears that this can be overcome with a large enough collection of assets and a judicious choice of the time series sample size. It also seems clear that all three of these tests, (KW, WE, and CH), provide valuable information about the uniqueness of the SDF. WE and KW are similar, and in columns 4-6 of Table II, WE has slightly higher power, but WE has the disadvantage of being a parametric test; hence KW might be preferred when one is not sure about the distributions of returns or of the underlying SDFs. CH is also a parametric test but it appears best when SDF volatility is high (column 3).

In the next simulations, the SDFs have the same means, based on riskless rates of 0.1%, but have different volatilities, 10% and 25%. The results are in columns 7-8 of Table II. KW and WE have virtually no power because the locations are the same. The BF test, in contrast, has almost perfect power for $N \geq 480$ and even for lower N in Panel A. The CH test has somewhat weaker power, particularly for the higher return perturbation volatility of Panel B. However, it too has good power for larger $N \geq 780$ and its power is perfect for $N \geq 480$ in Panel A.

The next simulations allow both the mean and volatility of the true SDFs to differ and also introduce stochastic dominance by allowing the SDF with the larger mean to have a smaller volatility. Thus, the riskless rate is set to .1% (5%) for the first (second) SDF and the volatility is set to 10% (25%). Results are in Table II, columns 9-12. Again, BF has excellent power except for N=240 in Panel B (higher perturbation volatility.) KW has decent power in Panel A for large N (720 and 960) and for large T (120) but its power deteriorates in Panel B. WE has weak power throughout. CH has excellent power for $N \geq 480$ in Panel A, for $N \geq 720$ in Panel B and even for a few cases with T = 30.

Finally, we document power with a larger number of asset groups. We choose five groups to match some of our later empirical tests. To make the tests face a tough challenge, we set up the experiment so that just one of the groups has a stochastically dominant SDF, the other four having SDFs with the same mean and variance. Asset group #1 has a stochastically dominant SDF with a riskless return of 0.1% and a standard deviation of 10%. Groups #2 through #5 each have SDFs with a riskless return of 5% and a standard deviation of 25%.

Table II, columns 13-16, report the results. The power is somewhat lower in most cases than in the two group tests reported in columns 9-12. KW's power seems to have fallen the most but WE is not very powerful in either case, particularly with the higher perturbation volatility in Panel B. However, BF still has good power except for lower N in Panel B while CH has excellent power in Panel A when $N \geq 480$ and for $T = 30$. Its power is also quite good for $N = 960$ even in Panel B.

III. B. 2. Test Power when the SDF equation is false for at least one asset group

This subsection considers the test power consequences of one asset group being aberrant in the sense that the basic SDF expectation (1) is not equal to unity. Since the previous subsection considers cross-group differences in the SDF itself, this section assumes that the SDF has the same distribution across all assets but the SDF/gross return product is not the same. Such a situation implies both incomplete markets and an arbitrage opportunity, which can be seen most intuitively by noting that the riskless rates must differ among asset groups.

In the interest of space, we will only consider a single simulation of this type with two asset groups. Power is tabulated as before with various combinations of the number of assets per group and the number of time series observations and two choices of return perturbation volatility.

To parameterize the error in the basic SDF equation, we set $E_{t-1}(\tilde{m}_t \tilde{R}_{i,t}) = 1 + \delta$ for each asset i in the aberrant group, with $\delta \neq 0$; δ takes on the values 0.05 and 0.1. In the other (normal) group, $\delta = 0$. The riskless rate is 0.1% per period and SDF volatility is 15% per period. The riskless rate and volatility are the same in the SDFs for each group.

Table III presents the results. For the KW and WE tests, power now improves with T, even when it is close to N while the opposite is true for CH (left side of table, perturbation volatility of 1%.) The BF test has virtually no power for all values of N and T, (not reported), essentially reflecting the fact that the variances of the underlying SDFs are the same in both asset groups. In contrast, the WE tests exhibit power in excess of 90% when N is close to 1,000, T = 120. When δ = 0.1, power is very good for all three tests except for CH and N = 240. Power is also excellent for CH when N $\geq$ 480. However, for larger return perturbation volatility, (right panels of Table III), power is quite poor for the lower value of δ.

III.B.3. Conclusions about Test Power

In summary, from all the above simulations, we learn that very large cross-sectional sample, N close to 1,000, provides robust power under a variety of conditions including the time series length, T, and the return perturbation volatility. When SDFs have disparate means and variances across asset groups, the tests provide decent power when return perturbation volatility is low, except when T approaches N/2 and the degrees-of-freedom start to become problematic. The power is generally very poor when the return perturbation variance is large and T is a large fraction of N.

When the SDF equation is false by 10% (relative to the predicted value of 1.0) in one asset group, while the SDFs have the same distribution across groups, the KW, WE and CH tests have good power for large N.

IV. Data

We collect monthly return observations on U.S. bonds, stocks, currencies (per US$), commodities and real estate (REITs, or real estate investment trusts), for July 2002 through December 2013, 138 months in all. The data begin in July 2002 because the Trace database starts reporting bond returns in that month. Stocks are sampled randomly from those on the CRSP database. We purposely select equities with low leverage to make them as different as possible from bonds, although we also select an equal-size random sample of other equities for later comparison.[23] Currencies and commodities are drawn from the Datastream and Real Estate Investment Trusts (REITs) from the CRSP database. In the cross-sectional sample, there are 956

[23] The average leverage (book debt/total assets) ratio is 10.21% for the 956 low-leverage equities.

low-leverage stocks, 123 bonds, 37 spot exchange rates per US$, 47 commodities, and 89 REITs that have simultaneous observations for every month.

We also have data from DataStream for equity markets in China and India, the Shanghai and Mumbai Exchanges. These countries are ideal for our tests because their equity markets feature a large number of individual assets. Also, given the political, financial and structural diversity of the two countries, we thought *a priori* that they could conceivably be segmented, particularly in light of restrictions on trading by foreigners (in both countries) and the obstacles faced by nationals who wish to trade abroad (such as foreign exchange regulations.) Concurrent return observations are available for 927 Indian stocks and 269 Chinese stocks from June 1994 through June 2016, a total of 265 months. The number of Chinese stocks is on the low side to assure good test power according to the simulations in Section III.B so we should be cautious if the test turns out to support integration of the two markets.

V. Empirical Tests of Market Integration Using the SDF Estimator

V.A. Equity Markets of India and China

There are adequate time series observations, so to examine whether integration has changed over time, we separate the data into two non-contiguous time periods, June 1994 to July 2005 (133 months) and August 2005 to June 2016 (132 months). There are 927 Indian stocks and 269 Chinese stocks in each sub-period.

Table IV reports the results for four different test statistics. The Kruskal/Wallis (1952) (hereafter KW) test indicates whether one set of SDF estimates stochastically dominates any other and it also provides a test of the difference in medians. The Welch (1951) (WE) test is for the equality of means and the Brown/Forsythe (1974) (BF) test is for the equality of variances. The Chi-Square test is for whether the estimated SDFs vectors are equal element by element. In all cases, a low p-value would reject the null hypothesis that the estimated SDFs are equal in, respectively, medians, mean, variances, or elements.

As the table reveals, there is strong evidence against the hypothesis that China and India share the same SDF. The means, medians, and elements are all significantly different during both sub-periods. The only hint that these two markets are becoming closer over time is the Brown/Forsyth test, which is statistically significant in the first sub-period but not the second. This suggests that the SDFs in the two markets more recently have similar volatilities. There is also a tiny increase in the Chi-Square test but it remains highly significant.

V.B. Tests across US asset classes

The SDF paradigm should apply to any partition of the available assets, but we begin the US tests with what should be a tough challenge. We estimate SDFs from each asset class independently and then test whether they are the same across asset classes. Our SDF estimator requires more assets than time periods, so we are limited to time series samples shorter than the number of individual assets in the smallest class, which is currencies with 37. Hence, the 138 available months are separated into roughly equal subsamples, 34 observations in the first two subsamples and 35 observations in the next two. We realize these tests probably lack power because N-T, the degrees-of-freedom, is quite small for some asset classes. Nonetheless, we believe they are worth reporting while recognizing their likely limitations. The results are in Table V.

The Kruskal/Wallis (1952) (hereafter KW) test indicates whether one set of SDF estimates stochastically dominates any other and it also provides a test of the difference in medians. There are five sets of sample SDFs, one for each asset class, which implies that the KW Chi-Square variate under the null hypothesis (H_0: no SDF dominates another) has four degrees-of-freedom. According to the KW test results reported in Table V, there is no stochastic dominance in any of the four sub-periods. The sample medians are not significantly different from one another. Hence this test does not reject SDF uniqueness for these assets and time periods.

Table V also reports tests for the equality of means and variances across the five sets of SDF estimates, the Welch (1951) (WE) test for means and the Brown/Forsythe (1974) (BF) test for variances. In agreement with the non-parametric KW test, the WE test finds no evidence of a difference in means for the SDFs estimated independently from the five asset classes. None of the p-values indicates significance.

The WE test allows for unequal volatilities across asset classes. This is fortunate because the BF test for differences in variances rejects the null in every sub-period. The Chi-Square element by element test essentially agrees with the BF test except that its p-value is marginally significant in the first sub-period. Evidently, although the sample SDFs appear to be located with their means and medians close to one another,[24] at least one asset class has sample SDFs with significantly larger or smaller variance than the others. This is apparently sufficient to induce significant differences in the elements of some estimated SDF vectors. In order to ascertain which asset class (or classes) is responsible, Table VI reports the time series standard deviations of the sample SDFs.[25]

It appears that currencies and commodities display larger estimated SDF volatilities than equities, bonds, and real estate. One possible explanation is that the sampling error in estimated SDFs is larger for asset classes with fewer constituent members. There seems to be a strong negative connection between the number of available assets and the volatility. Currencies have the smallest number of individual assets (only 37) and commodities are next (with 47.) This explanation is buttressed by the simulation results in Section III.A, which reveal a material improvement in the quality of our estimator with the number of assets, holding constant the time series sample size.

V.C. Tests with larger samples of assets and time periods

To investigate the possible confounding impact of sampling error, we conduct two further experiments. First, we compare stocks and bonds alone, without reference to the other three asset classes. Stocks and bonds dominate the sample with 956 and 123 individual assets, respectively. Recall that the number of assets N in a group must exceed the number of time series observations T. Since there are only 123 bonds available, we cannot use all 138 time series observations at

---

[24] The Welch test for equal means is valid even when variances are unequal.

[25] The simulations in section III.B suggest that test power might not be very good for small collections of assets such as 37 for currencies and 47 for commodities. However, since the BF test rejects strongly, power per se does not appear to be a problem.

once, so we simply divide them in half, 69 months in the first sub-sample (July 2002 – March 2008) and 69 in the second (April 2008 – December 2013.) Table VII provides the results.

As the Table VII Brown/Forsyth tests indicate, the increased sample sizes, both in number of asset and in time periods, does not overturn the previous result that the SDFs have divergent volatilities. However, unlike the results in Table V, the Chi-Square test no longer detects significant differences in the SDF vector elements.

Table VIII reports the volatilities, which are considerably larger for bonds than for stocks in both sub-periods. Evidently, the number of bonds remains too small compared to the number of time periods, which probably implies more sampling error and hence higher estimated SDF volatility for bonds.

In the second experiment, we abandon a strict asset class categorization in order to estimate sample SDFs using all available monthly observations at once and roughly equal-sized groups of assets. This increases the time series sample size from T=34 or T=35 (as in Table V) or T=69 (in Table VII) to T=138. Since the number of assets N in a group must exceed the number of time series observations T, it becomes necessary to mix stocks, which are the most numerous, in with the four other asset types in separate groups. There are 1252 individual assets of all types available for the 138 sample months, so five roughly equal-sized groups would contain, respectively, 250, 250, 250, 251, and 251 individual assets.

We compose the groups in the following manner: To group #1, we assign 250 equities, selected randomly; in group #2 we mingle 127 randomly-selected equities with all available (123) bonds; similarly, group #3 has 213 equities and 37 currencies; group #4 has 204 equities and 47 commodities; and group #5 has 162 equities and 89 REITs. The results are reported in Table IX. The Welch (1947) test for means and the Brown/Forsythe (1974) test for variances are in agreement with the non-parametric KW test. The Chi-Square test agrees with a p-value close to 0.5. Consequently, in these tests there is no evidence of a significant difference in the SDFs estimated from the different asset groupings. After taking account of sampling error disparities across test groups, there is no evidence of SDF differences even though the groups are heterogeneous in the sense of including five distinct asset classes.

However, there is one caveat. Our simulations in section III.B reveal that test power might not be very large when there are only 240 assets in each group unless sampling error is rather small. Thus far, we have not attempted to disentangle sampling return perturbation volatility from volatility in the true SDF. The variance of the estimated SDF is the sum of the two variances; we investigate this issue in Section V.E just below.

V.D. Tests with more equities and greater power

In the hope of achieving more test power, we conduct two final empirical experiments. In the first, we divide the sample of low-leverage equities into two equal-sized groups of 478 stocks each and work with all available 138 time series observations. Section III.B suggests that this choice of N and T should have good power. In the second test, we expand N even further by collecting a second group of 956 equities, randomly sampled from remaining CRSP stocks that **do** have significant levels of leverage.[26] Conceptually, this second test should be an exacting hurdle for SDF theory because the two groups of equities differ markedly in their leverage ratios.[27]

The results for both tests are reported in Table X. Panel A reports that none of the four tests, KW, WE, BF, and CH, rejects the null hypothesis that the SDFs are the same in the two groups of 478 low-leverage equities at a high level of significance, though the BF test is on the margin with a p-value of 0.084. Panel B reports a stronger inference; even with very different leverage (and, as consequence, likely different levels of riskiness), there is no evidence of a difference in SDFs. In all cases, the p-values are far from indicating significant rejection of the null hypothesis.

In conclusion, for a battery of tests with differing asset classes, differing group sizes, and diverse time series sample sizes, the unique SDF paradigm holds up well for U.S. assets. SDF uniqueness cannot be proved absolutely true, of course, but it is not rejected by our tests after

[26] The average leverage ratio (book debt/total assets) for this second group of stocks is 32.51%; the low leverage group has an average ratio of 10.21%.
[27] There are, of course, many other ways to construct heterogeneous groups of equities (size, beta, etc.) for similar cross-group tests, which we leave for future research.

properly accounting for sampling variation. Our tests here are conducted with U.S. data spanning a single recent decade, so more comprehensive tests with longer samples are clearly on the agenda for future research.

V.E. Properties of the estimated SDF; disentangling sampling error and true SDF volatility and the Hansen/Jagannathan variance bound

In the previous sub-section, we find with demonstrably powerful tests that SDF estimates from low- and higher-leveraged stocks are not significantly different. This does not prove that the SDF is unique, but that proposition cannot be rejected by those tests. In this section, we temporarily assume that the SDF **is** unique, which enables us to shed light on the properties of SDF estimates. It also permits the disentanglement of volatility in the true SDF from sampling error volatility in the estimated SDF and it allows us to check whether our estimates satisfy the Hansen/Jagannathan bounds

In conformance with previous notation, we now let $\hat{\mathbf{m}}(\mathrm{L})$ denote the vector of estimated SDFs from the low-leverage stocks and $\hat{\mathbf{m}}(\mathrm{H})$ denote the estimated SDFs from the higher-leverage stocks. When the H and L markets are integrated, an element of these vectors at time t can be expressed as

$$\hat{\mathbf{m}}_{\mathbf{j,t}} = \mathbf{E}_{\mathbf{t-1}}(\mathbf{m}_{\mathbf{t}}) + \boldsymbol{\upsilon}_{\mathbf{m,t}} + \boldsymbol{\upsilon}_{\mathbf{j,t}} \text{ , j=L,H} \tag{16}$$

where $\boldsymbol{\upsilon}_{\mathbf{m,t}}$ is the unexpected component of the true SDF at time t and $\boldsymbol{\upsilon}_{\mathbf{j,t}}$ is the estimation error in the sample estimated SDF for group j (j=L,H). No element on the right side of (16) is correlated with any other, so the time series variance of the estimated SDF is

$$\mathrm{Var}(\hat{\mathrm{m}}_{\mathrm{j}}) = \mathrm{Var}[\mathrm{E}(\mathrm{m})] + \mathrm{Var}(\upsilon_{\mathrm{m}}) + \mathrm{Var}(\upsilon_{\mathrm{j}}) \text{ , j=L,H.} \tag{17}$$

Assuming that the estimation errors for L and H are independent of each other,

$$\mathbf{Cov}(\hat{\mathbf{m}}_{\mathrm{L}}, \hat{\mathbf{m}}_{\mathbf{H}}) = \mathbf{Var}[\mathrm{E}(\mathbf{m})] + \mathbf{Var}(\boldsymbol{\upsilon}_{\mathbf{m}}) \text{ .} \tag{18}$$

The right side of (18) is the total volatility induced by the true SDF, including the intertemporal evolution of its expectation and its period-by-period unexpected component. Subtracting this result from (17) provides estimation error variances for j=L and j=H.

The SDF paradigm implies that $E_{t-1}(m_t)=1/(1+R_{F,t})$ for the riskless rate R_F at time t-1. During the time period of our sample, 2002-2013, the riskless rate had historically low variation over time, so Var[E(m)] should be relatively small compared to $Var(\upsilon_m)$, which should dominate (18).

Estimated over July 2002 through December 2013, the standard deviations of $\hat{m}_L$ and $\hat{m}_H$ are, respectively, 0.580 and 0.696 per month and correlation of $\hat{m}_L$ and $\hat{m}_H$ is 0.3002. This implies a standard deviation of SDF components, the square root of (18), equal to 0.3481. The standard deviations for the estimation errors for L and H are then, respectively, 0.4636 and 0.6031. Not surprisingly, higher leverage equities are associated with more volatile estimation errors. Both $\hat{m}_L$ and $\hat{m}_H$ are slightly autocorrelated at the first lag, autocorrelations of .181 and .157, respectively, but neither is statistically significant. In other respects, they seem to possess no particularly bizarre properties; for example their excess kurtoses are -0.266 and 0.436 and their skewnesses are 0.371 and 0.457, respectively.

The Hansen/Jagannathan (1991) variance-related bound requires that $\sigma(m)/E(m)$ be larger than the largest possible Sharpe ratio.[28] Recent opinions, Welch (2000), seem to be that the excess return on the best possible portfolio is no more than about 7% per annum (or even lower lately) and the portfolio's standard deviation may be around 16% per annum, so the largest Sharpe ratio is no more than 0.44. The sample means of $\hat{m}_L$ and $\hat{m}_H$ are, respectively, 0.9945 and 0.9961, both approximately unity. Our annualized SDF standard deviation is $0.348\sqrt{12}$, which comfortably satisfies the HS bounds. This inference contrasts strongly with previous research that has specified SDF proxies that depend on macroeconomic data. Evidently, SDFs that depends on

[28] Hansen and Jagannathan also derived bounds involving moment other than the first and second. See Snow (1991) for empirical estimation with a variety of bounds.

returns, such as ours and Long's Numeraire portfolio, are sufficiently volatile. This is a puzzle that clearly deserves further investigation.

The means and standard deviations of the estimated SDFs provide a simple test that the true SDFs are strictly positive (and, consequently, that there are no arbitrage opportunities.) The monthly observed standard deviation, 0.3481, implies that a negative SDF realization would be about 2.37 standard deviations below the mean. This implies that the probability of a negative SDF is about 0.0088, less than one percent.

To obtain a visual image of the evolution of the SDF, it is appropriate to first expunge estimation error. This is not possible for each individual time series observation, but one can adjust the overall series to have the true SDF volatility as estimated by (18). We simply need to find an attenuation coefficient, γ such that $\mathrm{Var}(\gamma\hat{m}) \cong \mathrm{Var}(\upsilon_M)$, which assumes that the riskless rate's variance is sufficiently small that it can be ignored; hence, $\gamma = \left[\mathrm{Var}(\upsilon_m) / \mathrm{Var}(\hat{m})\right]^{1/2}$. The adjustment entails the transformation

$$\hat{\hat{m}} = \bar{m} + \gamma(\hat{m} - \bar{m}), \tag{19}$$

where the double "chapeau" denotes the transformed SDF and $\bar{m}$ is the sample mean. For the low and high leverage equity groups, the attenuation coefficients are .6005 and .4999, respectively. This attenuation provides an adjusted standard deviations of exactly 0.3481 for both the Low and Higher leverage equity groups.

Figure IV plots the two adjusted SDF series using a 12-month moving average to smooth out short-term fluctuations. There is clearly a connection between the two series, which is not a surprise because our test above could not reject the hypothesis that they are the same. There is, however, something of a puzzle here in that the SDF is larger than 1.0 toward the end of the 2000 decade for both series. Of course, this is the *ex post* SDF, including the unexpected component. The expected SDF would presumably be much smoother.[29] Higher SDF values from 2006 through

[29] Neither series has a unit root according to the usual tests.

2010 are, perhaps, not all that surprising as they preceded and accompanied the recent economic contraction.

V.F. Estimated SDFs and Returns on a Market Index

In one further validation experiment, we estimate the relation over time between SDF estimates from low- and higher-leveraged stocks and observed returns on the S&P 500 index. This is motivated because the SDF paradigm stipulates that in each period there should be a relation between the aggregate market portfolio's return, $R_{W,t}$, and the SDF m_t, of the following form:

$$\mathbf{m_t = a_{t-1} - b_{t-1} R_{W,t}} \;,$$

where the coefficients are time varying and strictly positive; Cf. Cochrane (2001a, pp. 139-140.)

Unfortunately, we have only estimates of the two variables in the relation above, our estimate $\hat{\mathbf{m}}$ for m and the S&P 500 return for R_W. Moreover, we know nothing about the time variation in the coefficients, a and b, and are obliged to adopt the perhaps forlorn hope they are relatively constant.

Operationally, we run two proximate regressions,

$$\mathbf{\hat{m}_{j,t} = a_j - b_j R_{S\&P,t}}$$

with j=L (H) for Lower- (Higher-) leveraged equities, using 138 monthly observations, July 2002 through December 2013. Perhaps surprising, given the possible problems with this specification, we find $-b_L$=-3.48 (t-statistic=-3.17) and $-b_H$=-2.68 (t-statistic=-1.98.) Both slope coefficients have the right sign and are significant, though b_H's significance could be regarded as marginal. Clearly, there is more estimation error in $\hat{\mathbf{m}}$ for the higher-leveraged equities. As one would expect, the intercept terms are both very close to 1.0 and are highly significant, t-statistics of 21.0 and 17.1, respectively. However, the explanatory power is rather low, adjusted R-squares of 6.20% and 2.10%, respectively.

VI. Conclusions

The stochastic discount factor (SDF) paradigm predicts that the same SDF should price all assets in a given period when markets are complete. We develop tests of SDF uniqueness by first deriving an SDF estimator that can be calculated from observed returns and is agnostic with respect to macroeconomic state variables, preferences, and the form of the multivariate distribution of returns, including its factor structure. We emphasize that an SDF estimator needs not depend on the above-mentioned traditional elements because the SDF is a mathematical function within an integral equation.

Our SDF estimator is theoretically biased in finite samples and has a standard error that depends on both the number of asset, N, and the number of time periods, T, used in its construction. Hence, to examine the estimator's qualities, we resort to simulations. We find that the estimator is accurate when N-T is relatively large with N>2T and N near 1,000.

Equipped with an agnostic SDF estimator, we suggest four different tests that can potentially detect when SDFs differ significantly across groups of assets. Simulations are presented to assess the power of these tests. For large N relative to T, the suggested tests have excellent power that approaches 100% depending on various parameters such as the volatility of the true SDFs and the sampling variation in returns. We also present evidence that our SDF estimator works well when returns have thick tails and differ significantly in their means, volatilities and correlations with each other and also, as mentioned above, when there is a multi-factor structure of returns.

We first apply our estimator to test the equality of SDFs in the Chinese and Indian stock markets during two sub-samples of approximately eleven years each, from 1994 to 2016 in total. We find strong evidence that the SDFs are different in these countries in both sub-periods, with only a hint that their SDF volatilities (but not their means or other features) are becoming somewhat closer in the second sub-period. This seems to agree with common intuition that China and India remain segregated because of structural and regulatory differences.

We then apply our estimator and tests to data on U.S. equities, bonds, commodities, currencies, and real estate (REITs) over a common time period, 138 months from July 2002 through December 2013. As the simulations predict, asset classes with few individual assets (a low N), such as commodities and currencies, produce sample SDFs with larger volatilities. However, even in this case, there is no evidence that SDF means are different across asset classes. This result suggests that excessive SDF volatility in smaller asset classes might be attributable to sampling variation.

This explanation is corroborated by reorganizing the individual assets into larger grouping; sampling error is thereby reduced and there are no longer any significant violations of SDF uniqueness. The same result is obtained in a further test with larger numbers of individual assets (close to 1,000). Owing to data availability, such a test can be done only with equities. We find that two large groups of equities, one group with minimal leverage and the other with average leverage, are priced with SDFs that are not statistically distinguishable. We also find that these U.S.-based SDFs comfortably satisfy the Hansen/Jagannathan variance bound and are significantly non-negative.

Overall, the SDF paradigms's main prediction, that the same SDF prices all assets during the same time period, cannot be rejected with our tests using U.S. data in various asset classes during the 2002-2013 time period. The results are consistent with U.S. financial markets being integrated sufficiently across diverse asset classes to prevent the detection of incompleteness. Also, they suggest that cross-asset class arbitrage opportunities are difficult to uncover over all segments of the U.S. financial market.

Future research will determine whether the same inferences are obtained with other international data and with samples from other time periods.

Table I

Simulated Performance Information for the SDF Estimator

To assess our SDF estimator, we simulate true SDFs with mean=1/(1 + riskless interest rate) and various time series volatilities. Gross asset returns are simulated so that their mean values multiplied by the SDFs are equal to 1.0 but errors perturb their sample values. See Appendix A for details. The performance of the SDF estimator is measured by the correlation between true and sample SDFs and by Theil's (1966) U_2 statistic, which is closely related to the mean square prediction error. Linear regressions are reported in Panel A where the dependent variable is the correlation and in Panel B where the dependent variable is U_2. In Panel C, the dependent variable is the sample time series standard deviation of the estimated SDFs. Panel D reports the implied riskless rate from the reciprocals of the estimated SDFs. There are 2,880 parameter combinations, each with an independently-simulated set of true SDFs and returns.

Variable	Coefficient	T-Statistic
A: Correlation between true and estimated SDFs		
T, Time Periods	-1.104E-03	-11.034
N, Assets	1.505E-04	12.036
True SDF Volatility	1.295	18.581
Perturbation Volatility	-1.744	-49.302
Riskless Rate	5.799E-01	0.244
Adjusted R^2	0.488	
B: U_2 from comparing true and estimated SDFs		
T, Time Periods	2.068E-03	54.927
N, Assets	-2.964E-04	-62.989
True SDF Volatility	1.344E-01	5.126
Perturbation Volatility	6.237E-01	46.860
Riskless Rate	2.423E-01	0.271
Adjusted R^2	0.782	
C: Standard Deviation of Estimated SDFs		
T, Time Periods	2.917E-03	53.945
N, Assets	-4.270E-04	-63.166
True SDF Volatility	5.180E-02	1.376
Perturbation Volatility	4.899E-01	25.622
Riskless Rate	2.360E-01	0.184
Adjusted R^2	0.731	
D: Riskless Rate Inferred from Estimated SDFs		
T, Time Periods	-3.106E-06	-0.229
N, Assets	2.401E-06	1.415
True SDF Volatility	1.950E-02	2.063
Perturbation Volatility	1.164E-02	2.427
Riskless Rate	7.797E-01	2.422
Adjusted R^2	0.008	

Table II

Test Power with a Non-Unique Stochastic Discount Factors; Differing Means and/or Volatilities

Asset groups, each with N individual assets, have T simultaneous time series observations. The true stochastic discount factors (SDFs) differ across groups. SDF estimators are computed from the sample return observations in each group and then compared with the Kruskal/Wallis (KW), Brown/Forsythe (BF), Welch (WE), and Chi-Square (CH) tests. Power is the percentage of correct rejections of the null hypothesis (H_0: no difference in the SDFs) in 1,000 replications with a type I error of five percent. Perturbation volatility is 1% (2%) in Panels A (B). In the two-group tests (columns 1-12), SDFs means are determined by the reciprocals of unity plus riskless rates of .1% and 5% per period and SDF volatilities are 10% and 25% (standard deviation per period.) Any non-reported test not for a comparison has minimal power. The first comparisons are for equal volatilities and differing means, column 3 (4-6) for 25% (5%) volatilities. The BF test should not and does not have power in this case. Column 7-8 report tests for equal SDF means but different volatilities. In columns 9-12, one SDF stochastically dominates the other, with higher mean and lower volatility. In the five-group test, columns 13-16, one group's SDF stochastically dominates the other four.

		Two-Group Tests										Five-Group Test			
Riskless Rate		0.10% 5%	0.10% 5%			0.10% 0.10%		0.10% 5%				one group @ .1% four groups @ 5%			
SDF Volatility		25% 25%	5% 5%			10% 25%		10% 25%				one group @ 10% four groups @ 25%			
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
T	N	CH	KW	WE	CH	BF	CH	KW	BF	WE	CH	KW	BF	WE	CH
Panel A: Perturbation Volatility: 1%															
30	240	79.7	98.2	99.8	77.4	99.9	99.9	21.9	99.4	9.0	100.0	6.7	75.0	7.7	100.0
60	240	16.8	100.0	100.0	9.5	99.9	98.9	43.2	99.9	15.8	99.8	17.1	99.7	14.4	96.2
90	240	1.3	100.0	100.0	0.2	100.0	82.1	59.4	100.0	21.4	92.2	26.6	99.8	19.2	58.4
120	240	0.0	99.7	100.0	0.0	100.0	29.4	67.8	100.0	28.8	43.1	31.8	99.9	20.9	9.4
30	480	100.0	100.0	100.0	100.0	100.0	100.0	22.1	100.0	7.5	100.0	7.0	84.3	9.4	100.0
60	480	99.5	100.0	100.0	99.7	100.0	100.0	50.9	100.0	19.0	100.0	24.7	99.9	18.9	100.0
90	480	94.6	100.0	100.0	95.0	100.0	100.0	73.1	100.0	28.6	100.0	46.9	100.0	34.2	100.0
120	480	63.4	100.0	100.0	61.7	100.0	100.0	85.9	100.0	44.0	100.0	62.5	100.0	42.1	100.0
30	720	100.0	100.0	100.0	100.0	100.0	100.0	21.4	100.0	8.7	100.0	9.0	84.4	9.5	100.0
60	720	100.0	100.0	100.0	100.0	100.0	100.0	53.7	100.0	19.5	100.0	29.4	100.0	22.3	100.0
90	720	100.0	100.0	100.0	100.0	100.0	100.0	75.5	100.0	30.4	100.0	51.5	100.0	34.8	100.0
120	720	100.0	100.0	100.0	100.0	100.0	100.0	88.7	100.0	45.8	100.0	69.9	100.0	50.8	100.0
30	960	100.0	100.0	100.0	100.0	99.9	100.0	22.7	99.9	8.6	100.0	8.7	86.6	9.7	100.0
60	960	100.0	100.0	100.0	100.0	100.0	100.0	54.0	100.0	17.2	100.0	27.2	100.0	20.8	100.0
90	960	100.0	100.0	100.0	100.0	100.0	100.0	76.9	100.0	33.3	100.0	48.0	100.0	35.5	100.0
120	960	100.0	100.0	100.0	100.0	100.0	100.0	89.2	100.0	45.7	100.0	73.1	100.0	51.5	100.0

Table II (Continued)

		Two-Group Tests										Five-Group Test			
Riskless Rate		0.10%	0.10%			0.10%		0.10%				one group @ .1%			
		5%	5%			0.10%		5%				four groups @ 5%			
SDF Volatility		25%	5%			10%		10%				one group @ 10%			
		25%	5%			25%		25%				four groups @ 25%			
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
T	N	CH	KW	WE	CH	BF	CH	KW	BF	WE	CH	KW	BF	WE	CH
Panel B: Perturbation Volatility: 2%															
30	240	1.0	19.0	18.8	0.0	88.0	40.2	12.8	83.4	4.8	48.1	2.4	45.1	4.5	25.0
60	240	0.0	17.8	12.6	0.0	91.4	2.1	20.2	85.8	7.0	3.5	1.7	72.3	2.7	0.1
90	240	0.0	6.5	0.6	0.0	85.1	0.1	17.0	75.4	5.2	0.1	1.8	63.9	1.1	0.0
120	240	0.0	2.2	0.0	0.0	76.0	0.0	9.6	58.4	3.3	0.0	0.5	55.1	0.3	0.0
30	480	15.6	83.8	95.8	9.8	98.5	92.2	19.5	95.1	7.2	96.1	5.0	65.5	8.0	88.5
60	480	0.6	97.3	100.0	0.0	99.9	62.1	37.8	99.5	13.5	73.2	13.4	96.8	11.9	41.7
90	480	0.0	98.4	100.0	0.0	100.0	19.2	49.3	99.8	18.8	30.7	18.9	97.8	12.9	4.7
120	480	0.0	98.4	100.0	0.0	99.9	2.6	59.6	99.5	24.2	5.8	27.7	99.0	18.4	0.1
30	720	54.6	97.3	99.9	47.0	99.0	99.4	22.4	98.5	10.9	99.9	6.5	71.0	8.8	98.0
60	720	8.8	99.9	100.0	1.5	99.9	95.7	47.7	99.7	18.2	98.6	17.4	99.0	16.4	93.1
90	720	0.8	100.0	100.0	0.0	100.0	79.9	60.8	99.9	25.5	92.6	34.7	100.0	24.5	63.8
120	720	0.3	100.0	100.0	0.0	100.0	62.8	77.2	100.0	37.8	81.6	46.5	100.0	31.8	41.9
30	960	89.4	99.5	100.0	87.8	99.3	99.9	23.7	99.2	9.3	99.9	5.6	76.7	8.7	99.9
60	960	38.9	100.0	100.0	32.1	100.0	99.9	47.4	100.0	18.5	100.0	24.2	99.8	19.0	99.7
90	960	7.0	100.0	100.0	2.3	100.0	99.1	69.9	100.0	30.1	99.8	41.3	100.0	31.0	97.2
120	960	2.0	100.0	100.0	0.1	100.0	97.9	81.0	100.0	42.6	99.8	59.7	100.0	43.9	92.7

Table III

Test Power With Non-Unique Stochastic Discount Factors; $E(\tilde{m}\tilde{R}) \neq 1$ for One Group

Two asset groups, each with N individual assets, have T simultaneous time series observations. The groups share a common SDF whose mean is determined by the reciprocal of unity plus a riskless rate of .1% per period and SDF volatility is 15% (standard deviation per period.) But the basic SDF equation is false for one of the two groups; i.e., for that group, $E(\tilde{m}\tilde{R}) = 1 + \delta$ with $\delta \neq 0$. SDF estimators are computed from the sample return observations in each group and then compared with the Kruskal/Wallis (KW), Welch (WE), and Chi-Square (CH) tests. The BF test is not reported since it has no power in this case. Power is the percentage of correct rejections of the null hypothesis (H_0: no difference in the SDFs) in 1,000 replications with a type I error of five percent. Perturbation volatility is 1% (2%) in the left (right) section; δ = .05 (.10) in Panel A (B).

Perturbation Volatility		1%			2%		
		Panel A: δ = .05					
T	N	KW	WE	CH	KW	WE	CH
30	240	0.3	0.1	84.3	0.3	0.0	0.5
60	240	5.1	2.0	19.5	0.2	0.0	0.0
90	240	23.8	18.4	0.4	0.2	0.0	0.0
120	240	36.8	46.8	0.0	0.4	0.0	0.0
30	480	0.8	0.3	100.0	0.4	0.0	16.3
60	480	13.1	9.2	99.9	3.4	0.3	0.1
90	480	56.5	72.8	98.3	12.4	5.3	0.0
120	480	91.7	99.7	78.7	29.4	29.5	0.0
30	720	0.8	0.2	100.0	0.5	0.2	58.6
60	720	13.9	12.3	100.0	8.4	3.7	5.6
90	720	72.8	84.4	100.0	33.4	40.0	0.2
120	720	98.3	100.0	100.0	66.4	93.0	0.1
30	960	0.7	0.2	100.0	0.6	0.2	91.8
60	960	18.0	15.4	100.0	10.9	9.1	48.2
90	960	77.6	86.5	100.0	50.7	65.0	5.7
120	960	99.3	100.0	100.0	84.9	98.9	1.3
		Panel B: δ = .10					
30	240	91.2	94.4	100.0	69.5	83.7	82.9
60	240	100.0	100.0	100.0	97.8	100.0	13.6
90	240	100.0	100.0	100.0	99.3	100.0	0.1
120	240	100.0	100.0	97.8	98.6	100.0	0.0
30	480	95.5	95.7	100.0	90.0	95.4	100.0
60	480	100.0	100.0	100.0	100.0	100.0	99.9
90	480	100.0	100.0	100.0	100.0	100.0	97.7
120	480	100.0	100.0	100.0	100.0	100.0	74.6
30	720	94.0	95.9	100.0	92.6	98.1	100.0
60	720	100.0	100.0	100.0	100.0	100.0	100.0
90	720	100.0	100.0	100.0	100.0	100.0	100.0
120	720	100.0	100.0	100.0	100.0	100.0	100.0
30	960	95.7	97.3	100.0	94.7	96.9	100.0
60	960	100.0	100.0	100.0	100.0	100.0	100.0
90	960	100.0	100.0	100.0	100.0	100.0	100.0
120	960	100.0	100.0	100.0	100.0	100.0	100.0

Table IV
Tests With Chinese and Indian Equities

Stochastic discount factors (SDFs) are estimated with individual equities traded in concurrent months on the Chinese and Indian stock exchanges over June 1994 through July 2016, (265 months.) There are 927 Indian stocks and 269 Chinese stocks. The total sample is divided into roughly equal subsamples with 133 monthly observations in the first and 132 in the second. Differences in the SDFs estimated for China and India are tested for stochastic dominance with the non-parametric Kruskal/Wallis (1952) statistic. Means and variances are compared with, respectively, the Welch (1951) and Brown/Forsythe (1974) tests. A Hausman (1978) type Chi-Square tests whether the estimated SDF vectors are equal element by element. P-values are for the null hypothesis that the asset classes are all priced with the same SDFs. A low p-value rejects the null.

Sub-Period	Stochastic Dominance (Kruskal/Wallis)	Equal Means (Welch)	Equal Variances (Brown/Forsythe)	Equal Elements (Chi-Square)
	P-value for identical SDFs in Chinese and Indian Equities			
Jun‘94-July‘05	0.0000	0.0000	0.0363	0.0000
Aug‘05-Jul‘26	0.0000	0.0000	0.2889	0.0004

Table V
Tests With Five U.S. Asset Classes

Stochastic discount factors (SDFs) are estimated for five different asset classes in the U.S., equities, bonds, currencies, commodities, and real estate (REITs), using simultaneous monthly observations for individual assets, July 2002 through December 2013, (138 months.) The total sample is divided into four similarly-sized subsamples with 34 monthly observations in the first two subsamples and 35 observations in last two. Differences across asset classes in the estimated SDFs are tested for stochastic dominance with the non-parametric Kruskal/Wallis (1952) statistic. Means and variances are compared with, respectively, the Welch (1951) and Brown/Forsythe (1974) tests. A Hausman (1978) type Chi-Square tests whether estimated SDF vectors are equal element by element. This Chi-Square test compares each asset pair and is considered significant if the minimum p-value is below the type I error divided by a Bonferroni correction; i.e., p-value less than .05/10 = .005 (with ten pairs being compared.) The minimum across ten pairs is reported. P-values are for the null hypothesis that the asset classes are all priced with the same SDFs. A low p-value rejects the null.

Sub-Period	Stochastic Dominance (Kruskal/Wallis)	Equal Means (Welch)	Equal Variances (Brown/Forsythe)	Equal Elements (Chi-Square)
	P-value for identical SDFs in all five asset classes			
Jul‘02-Apr‘05	0.976	1.000	0.000	0.008
May‘05-Feb‘08	0.956	1.000	0.000	<0.001
Mar‘08-Jan‘11	0.756	1.000	0.000	<0.001
Feb‘11-Dec‘13	0.817	1.000	0.000	<0.001

Table VI
Volatility of Sample SDFs by U.S. Asset Class and Sub-Period

The time series standard deviation is reported for sample SDFs estimated simultaneously with five different U.S. asset classes in four sequential sub-periods. The number of available assets, N, is reported in the second line.

	Equities	Bonds	Currencies	Commodities	Real Estate
	956	123	37	47	89
Sub-Period	Time Series Standard Deviation of Estimated SDF				
Jul ‘02-Apr ‘05	0.429	0.785	2.504	3.391	1.117
May ‘05-Feb ‘08	0.406	0.692	5.464	4.286	1.039
Mar ‘08-Jan ‘11	0.335	0.402	7.630	2.432	1.068
Feb ‘11-Dec ‘13	0.430	0.620	11.767	2.123	0.931

Table VII
Tests With U.S. Stocks and Bonds

Stochastic discount factors (SDFs) are estimated for U.S. equities and bonds using simultaneous monthly observations for individual assets, July 2002 through December 2013, (138 months.) The total sample is divided into roughly two equal sub-samples with 69 monthly observations each. Differences between stocks and bonds in the estimated SDFs are tested for stochastic dominance with the non-parametric Kruskal/Wallis (1952) statistic. Means and variances are compared with, respectively, the Welch (1951) and Brown/Forsythe (1974) tests. A Hausman (1978) type Chi-Square tests whether estimated SDF vectors are equal element by element. P-values are for the null hypothesis that bonds and stocks are priced with the same SDFs. A low p-value rejects the null.

Sub-Period	Stochastic Dominance (Kruskal/Wallis)	Equal Means (Welch)	Equal Variances (Brown/Forsythe)	Equal Elements (Chi-Square)
	P-value for identical SDFs in bonds and stocks			
Jul‘02-Mar‘08	0.578	0.944	0.0000	0.111
Apr‘08-Dec‘13	0.927	0.968	0.0002	0.328

Table VIII
Volatilities of Estimated SDFs for U.S. Stocks and Bonds

The time series standard deviation is reported for SDFs estimated simultaneously with U.S. stocks and bonds in two sequential sub-periods. The number of available assets, N, is reported in the second line.

	Equities	Bonds
	956	123
Sub-Period	SDF Volatility	
Jul ‘02-Mar ‘08	0.480	1.019
April ‘08-Dec ‘13	0.420	0.720

Table IX

Tests With Mingled Groups of U.S. Assets in Different Classes

Stochastic discount factors (SDFs) are estimated for five groupings of U.S. assets from different classes using simultaneous monthly gross return observations, July 2002 through December 2013, (138 months.) There are 1252 assets of all types available; they are assigned to five roughly equal sized groups of 250, 250, 250, 251, and 251 so that the number of assets in each group exceeds the time series sample size, which permits the calculation of estimated SDFs for each group separately. The composition of each group is reported in the second part of the table. The 956 available equities are assigned randomly to groups and mingled with all available assets of another type in groups 2-5. Differences in estimated SDFs across asset groups are tested for stochastic dominance with the non-parametric Kruskal/Wallis (1952) statistic. A Hausman (1978) type Chi-Square tests whether estimated SDF vectors are equal element by element. This Chi-Square test compares each asset pair and is considered significant if the minimum p-value is below the type I error divided by a Bonferroni correction; i.e., p-value less than .05/10 = .005 (with ten pairs being compared.) The minimum across ten pairs is reported. Means and variances are compared with, respectively, the Welch (1951) and Brown/Forsythe (1974) tests. P-values are for the null hypothesis that the asset classes are all priced with the same SDFs. A low p-value rejects the null.

Stochastic Dominance (Kruskal/Wallis)	Equal Means (Welch)	Equal Variances (Brown/Forsythe)	Equal Elements (Chi-Square)
P-value for identical SDFs in all five asset groups			
0.996	1.000	0.370	0.489

Group	Composition
1	250 Equities
2	127 Equities and 123 Bonds
3	213 Equities and 37 Currencies
4	204 Equities and 47 Commodities
5	162 Equities and 89 REITs

Table X

Tests With Larger Samples of U.S. Equities

Stochastic discount factors (SDFs) are estimated for two groupings of U.S. equities using simultaneous monthly gross return observations, July 2002 through December 2013, (138 months.) In Panel A, 956 low-leverage equities are randomly assigned to two groups of 478 each and SDFs are estimated separately from each group. In Panel B, SDFs estimated with 956 low-leverage equities are compared to SDFs estimated with 956 randomly-selected (and different) equities that have typical leverage levels.[30] Differences in estimated SDFs across asset groups are tested for stochastic dominance with the non-parametric Kruskal/Wallis (1952) statistic. Means and variances are compared with, respectively, the Welch (1951) and Brown/Forsythe (1974) tests. A Hausman (1978) type Chi-Square tests whether estimated SDF vectors are equal element by element. Low p-values in the table would reject the null hypothesis that all groups are priced with the same SDFs.

Sample Period	Stochastic Dominance (Kruskal/Wallis)	Equal Means (Welch)	Equal Variances (Brown/Forsythe)	Equal Elements (Chi-Square)
	A: Two groups of 478 low-leverage equities			
Jul ‘02-Dec ‘13	0.547	0.999	0.084	0.481
	B: Low- vs. higher-leverage groups of 956 equities each			
	0.679	0.995	0.808	0.457

[30] The Lower (Higher) leverage group has an average book debt to assets ratio of 10.21% (32.51%).

References

Adrian, Tobias, Richard K. Crump, and Emanuel Moench. "Pricing the Term Structure with Linear Regressions," *Journal of Financial Economics* 110(1) (2013), 110-138.

Aït-Sahalia, Yacine and Andrew W. Lo. "Nonparametric Estimation of State-Price Densities Implicit in Financial Asset Prices," *Journal of Finance* 53 (1998), 499-547.

Aït-Sahalia, Yacine and Andrew W. Lo. "Nonparametric Risk Management and Implied Risk Aversion," *Journal of Econometrics* 94 (2000), 9-51.

Araujo, Fabio, Joao Victor Issler, and Marcelo Fernandes. "Estimating the Stochastic Discount Factor without a Utility Function," Working Paper, Graduate School of Economics EPGE Getulio Vargas Foundation, March 2005.

Araujo, Fabio and Joao Victor Issler. "A Stochastic Discount Factor Approach to Asset Pricing Using Panel Data Asymptotics," Graduate School of Economics EPGE Getulio Vargas Foundation, 2011.

Brown, Morton B. and Alan B. Forsythe. "Robust Tests for Equality of Variances," *Journal of the American Statistical Association* 69 (1974), 364–367.

Brown, Stephen J. and Mark I. Weinstein. "A New Approach to Testing Asset Pricing Models: The Bilinear Paradigm," *Journal of Finance* 38(3) (1983), 711-43.

Campbell, John Y. "Intertemporal Asset Pricing without Consumption Data," *American Economic Review* 83(1993), 487-512.

Campbell, John Y. "Empirical Asset Pricing: Eugene Fama, Lars Peter Hansen, and Robert Shiller," Working Paper, Harvard University, May 2014.

Chapman, David A. "Approximating the Asset Pricing Kernel," *Journal of Finance* 52 (1997), 1383-1410.

Chen, Xiaohong and Sydney Ludvigson,. "Land of Addicts? An Empirical Investigation of Habit-Based Asset Pricing Models," *Journal of Applied Econometrics* 24 (2009), 1057-1093.

Chen, Zhiwu, and Peter J. Knez, "A Pricing Operator-Based Testing Foundation for a Class of Factor Pricing Models," *Mathematical Finance* 4 (1994), 121-141.

Chen, Zhiwu, and Peter J. Knez. "Measurement of Market Integration and Arbitrage," *Review of Financial Studies* 8 (1995), 287-325.

Chen, Zhiwu, and Peter J. Knez. "Portfolio Performance Measurement: Theory and Applications," *Review of Financial Studies* 9 (1996), 511-555.

Cochrane, John H. "A Cross-Sectional Test of an Investment-Based Asset Pricing Model," *Journal of Political Economy* 104 (1996), 241-255.

Cochrane, John H. “Asset Pricing,” Revised Edition, Princeton: Princeton University Press (2005).

Cochrane, John H. “A Rehabilitation of Stochastic Discount Factor Methodology,” Working Paper No. 8533, National Bureau of Economic Research (2001b).

Cochrane, John H., and Christopher L. Culp. “Equilibrium Asset Pricing and Discount Factors: Overview and Implications for Derivatives Valuation and Risk Management,” ch. 5, in *Modern Risk Management: A History*. Peter Field, ed. London: Risk Books (2003).

Cochrane, John H., and Lars Peter Hansen. “Asset Pricing Explorations for Macroeconomics,” Working Paper No. 4408, National Bureau of Economic Research (June 1992.)

Da, Zhi and Hayong Yun. “Electricity Consumption and Asset Prices,” Working Paper, Notre Dame University (2010).

DeSantis, Gerard. “Volatility Bounds for Stochastic Discount Factors: Tests and Implications from International Financial Markets,” Ph.D. dissertation, Department of Economics, University of Chicago (1993).

Engle, Robert F., and Sharon Kozicki. “Testing for Common Features,” *Journal of Business and Economic Statistics* 11 (1993), 369-80.

Ferson, Wayne. “Theory and Empirical Testing of Asset Pricing Models,” in Jarrow, R. A., V. Maksimovic, and W. T. Ziemba (eds.), *Handbooks in OR and MS,* Vol. 9, Finance, Amsterdam: Elsevier Science, 1995.

Hansen, Lars Peter, and Ravi Jagannathan. “Implications of Security Market Data for Models of Dynamic Economies,” *Journal of Political Economy* 99 (1991), 225-262.

Hausman, J.A. “Specification Tests in Econometrics,” Econometrica 46 (6) (1978), 1251-1271.

He, Jia, Lilian Ng and Chu Zhang. “Asset Pricing Specification Errors and Performance Evaluation,” *European Finance Review* 3 (1999), 205-232.

Hotelling, Harold. “The Generalization of Student’s Ratio,” *Annals of Mathematical Statistics* 2 (3) (1931), 360–378.

Kan, Raymond, and Guofu Zhou. “A Critique of the Stochastic Discount Factor Methodology,” *Journal of Finance* 54 (1999), 1221-1248

Korteweg, Arthur, and Stefan Nagel. “Risk-Adjusting the Returns to Venture Capital,” *Journal of Finance* 71(3) (2016), 1437–1470

Kozak, Serhiy, Stefan Nagel, and Shrihari Santosh. “Interpreting Factor Models,” University of Michigan, Working Paper, November 2015

Kruskal, William and Allan Wallis, “Use of Ranks in One-Criterion Variance Analysis,” *Journal of the American Statistical Association* 47 (260) (1952), 583-621.

Lettau, Martin, and Sydney Ludvigson. “Resurrecting the C(CAPM); A Cross-Sectional Test When Risk Premia are Time-Varying,” *Journal of Political Economy* 109(6) (2001), 1238-1287.

Lettau, Martin, and Sydney Ludvigson. “Consumption, Aggregate Wealth and Expected Stock Returns,” *Journal of Finance* 56 (2001), 815-849.

Long, John. “The Numeraire Portfolio,” *Journal of Financial Economics* 26(1) (1990), 29-69.

McCulloch, J. Huston, “The Risk-Neutral Measure and Option Pricing under Log-Stable Uncertainty,” Ohio State University Economics Department working paper, (2003).

Mehra, Raj, and Edward Prescott. “The Equity Premium: A Puzzle,” *Journal of Monetary Economics* 15 (1985), 145-161.

Nagel, Stefan and Kenneth J. Singleton. “Estimation and Evaluation of Conditional Asset Pricing Models,” *Journal of Finance* 66(3) (2011), 873-909.

Polyanin, Andrei D., and Alexander V. Manzhirov, *Handbook of Integral Equations*, Boca Raton, CRC Press, 1998.

Roll, Richard. “Evidence on the Growth Optimum Model,” *Journal of Finance* 28(3) (1973), 551-566.

Rosenberg, Joshua, and Robert F. Engle. “Empirical Pricing Kernels,” *Journal of Financial Economics* 64 (2002), 341-372.

Ross, Stephen A. “The Arbitrage Theory Of Capital Asset Pricing,” *Journal of Economic Theory* 13 (1976), 341–360.

Singleton, Kenneth J. “Specification and Estimation of Intertemporal Asset Pricing Models,” in Briedman, B., and F. Hahn (eds.), *Handbook of Monetary Economics*, Vol, 1, Amsterdam, Elsevier Science, 1990.

Singleton, Kenneth J. “*Empirical Dynamic Asset Pricing*,” Princeton: Princeton University Press, 2006.

Snow, Karl N. “Diagnosing Asset Pricing Models Using the Distribution of Asset Returns.” *Journal of Finance*, Papers and Proceedings 46 (1991), 955-983.

Theil, Henri. “*Applied Economic Forecasting*,” Chicago: Rand McNally, 1966.

Welch, B. L. “On the Comparison of Several Mean Values: An Alternative Approach,” *Biometrika* 38 (1951), 330-336.

Welch, Ivo, “View of Financial Economists on the Equity Premium and Other Issues,” *Journal of Business* 73 (2000), 501-537.

Figure I
The Estimated and True SDFs with Small Return Perturbations

To demonstrate the SDF estimator, the perturbation in equation (A-3) of Appendix A is set to a very small value, .01% per period. The true SDF has a mean dictated by a riskless rate of .4% per period and its standard deviation is 4% per period. Returns have a mean and standard deviation per period of .8% and 8%, respectively. The number of assets, N, is 120 and the number of time periods, T, is 60, so there are sixty estimated and true SDFs plotted.

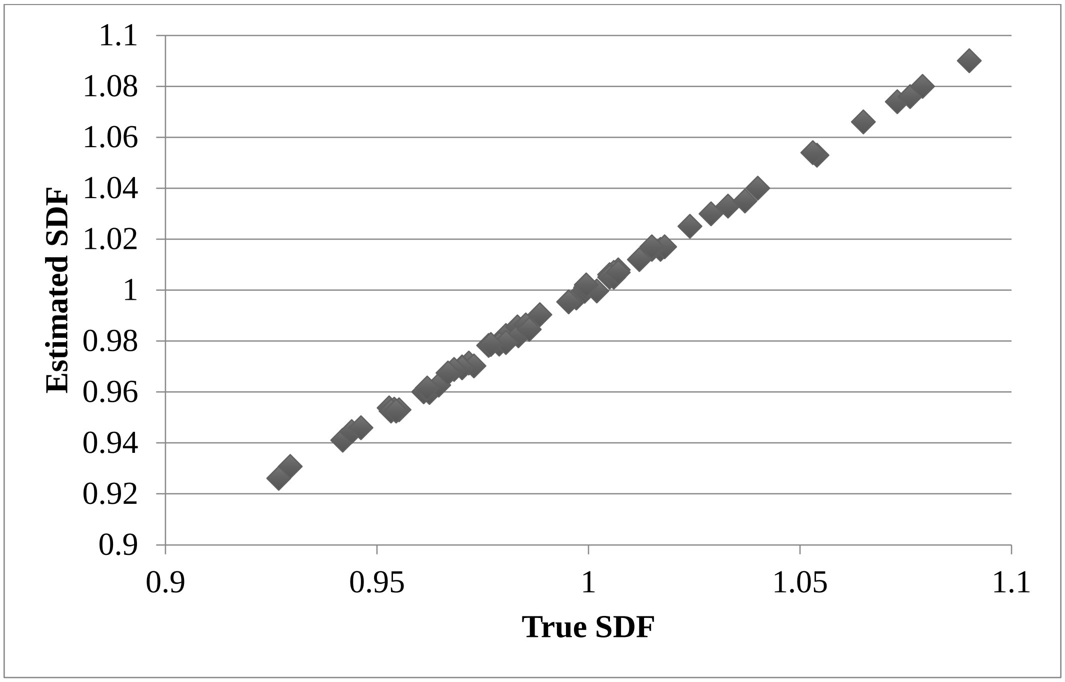
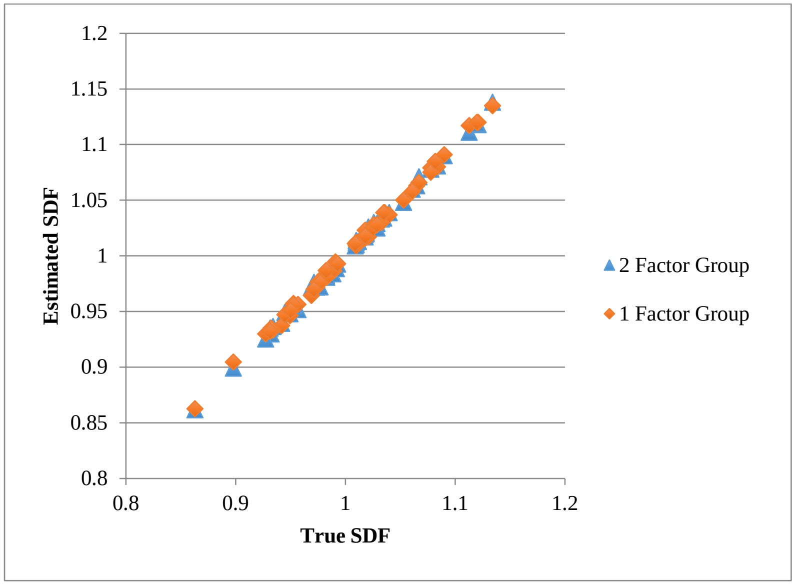


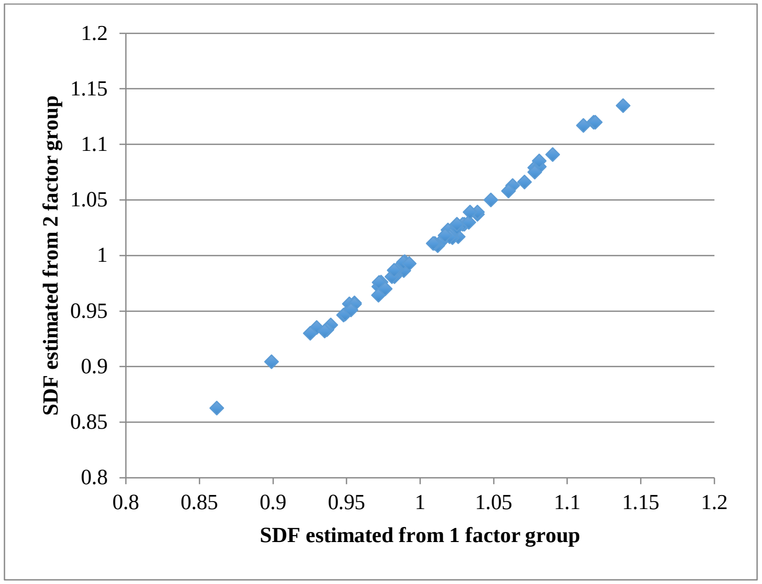
Figure II

The Estimated and True SDFs with for Asset Groups
with Diverse Factor Structures and Levels of Return Perturbation Volatility

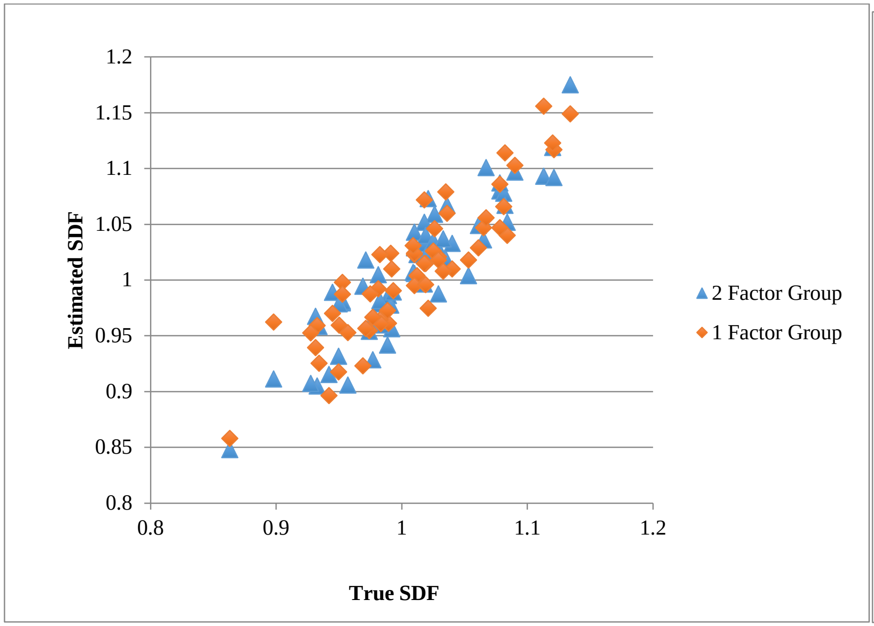
There is a unique SDF that prices all assets. It has a mean dictated by a riskless rate of .4% per period and a standard deviation is 4% per period. One group of assets has returns driven by a two-factor structure while the other group of assets has a single-factor structure. See Appendix A, Section A.1.a, for details. The number of assets, N, is 120 and the number of time periods, T, is 60, so there are sixty estimated and true SDFs plotted. In the first panel below, the return perturbations are very small, a standard deviation of 0.01% per period. The second panel has return perturbations with ten times as much volatility, a standard deviation of 0.1% per period. All other parameter values for the simulations are specified in Appendix A. The first plot below shows each group's estimated SDF plotted against the true SDF. The second plot shows the estimated SDFs for the two asset groups plotted against each other.

Panel A, Small Perturbations





Panel B, Larger Perturbations



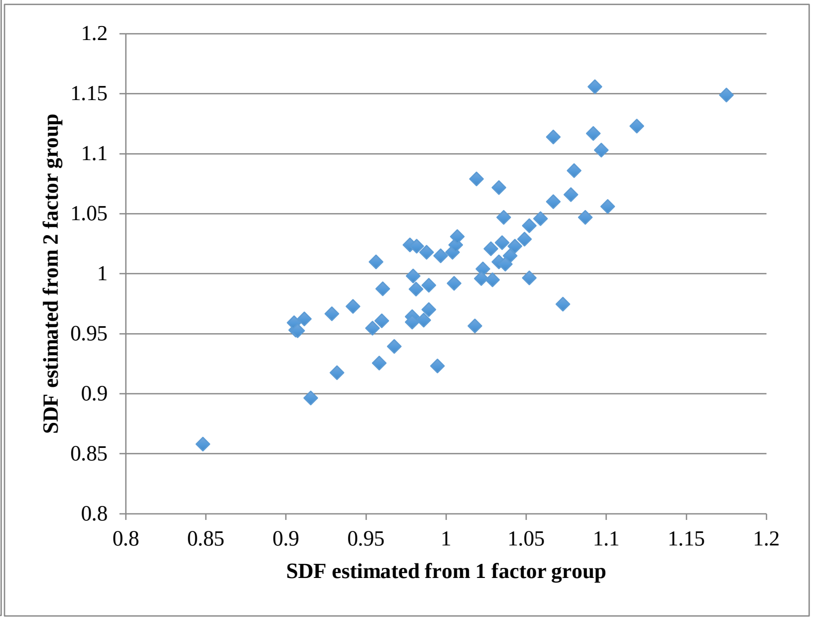
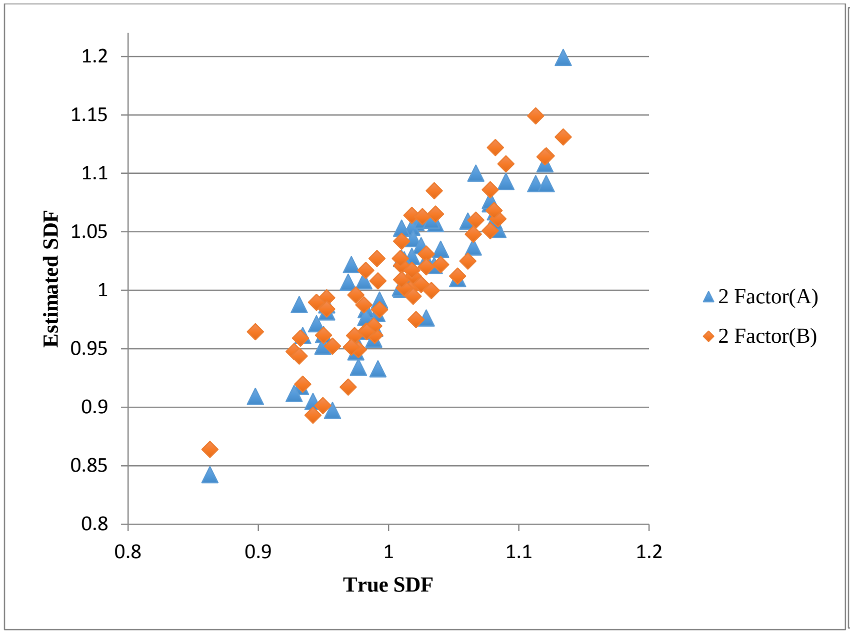


Figure III

The Estimated and True SDFs with for Two Asset Groups
Both with Two-Factor Structures but Whose Factors are Unrelated

There is a unique SDF that prices all assets. It has a mean dictated by a riskless rate of .4% per period and a standard deviation is 4% per period. Both groups of assets have returns driven by a two-factor structure but the factors are unrelated across groups. See Appendix A, Section A.1.b, for details. The number of assets, N, is 120 and the number of time periods, T, is 60, so there are sixty estimated and true SDFs plotted. The return perturbations are relatively large, a standard deviation of 0.1% per period, the same as in Panel B of Figure II above. All other parameter values for the simulations are specified in Appendix A. The first plot below shows each group's estimated SDF plotted against the true SDF. The second plot shows the estimated SDFs for the two asset groups plotted against each other.



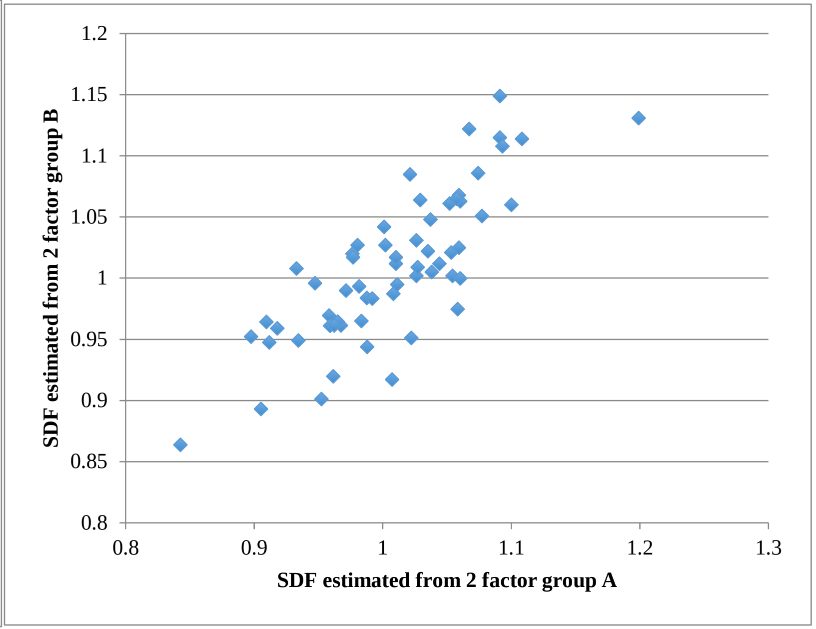
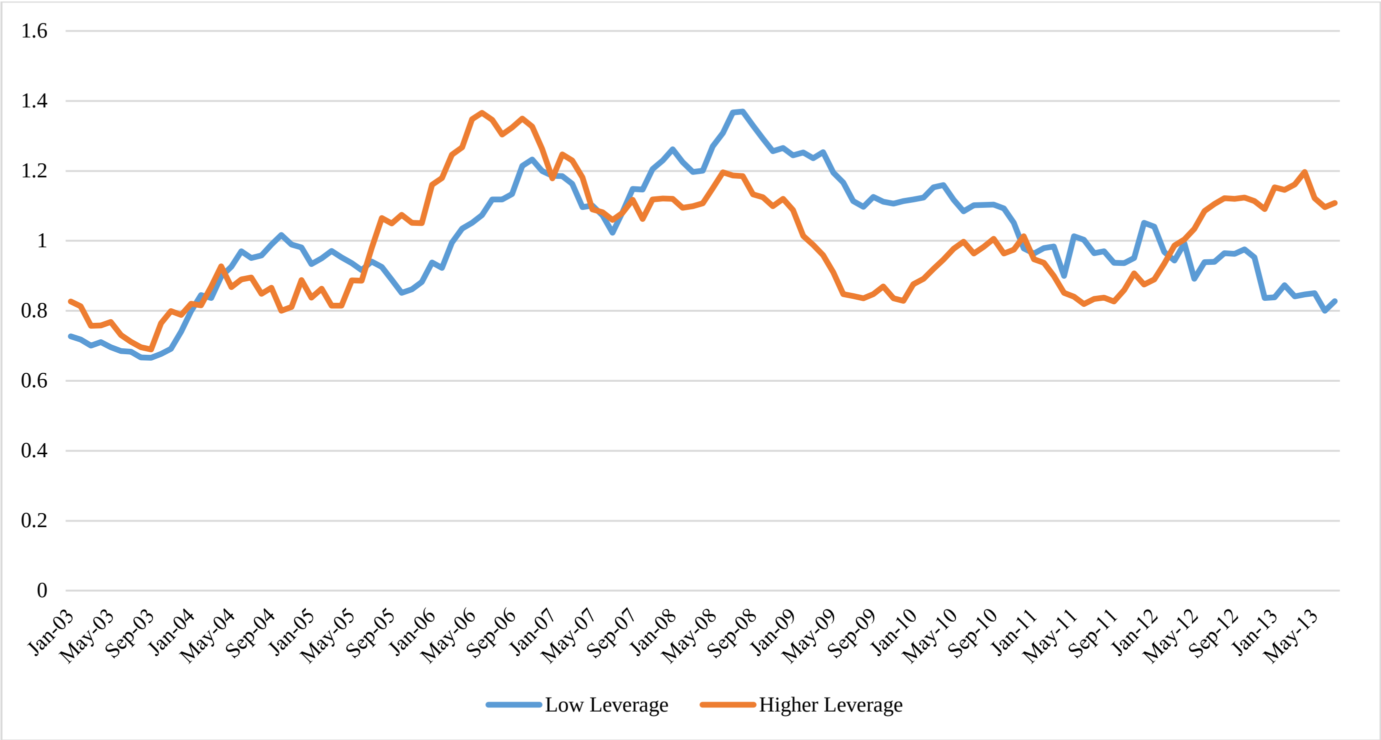


Figure IV

Time Series Plots of Estimated SDFs from Low- and Higher-Leveraged Equities

Two groups of equities, each with 956 individual firms, are used to estimate Stochastic Discount Factors (SDFs) with data from July 2002 through December 2013. One group is selected from firms with the lowest average leverage ratios over the 138 sample months. The other group is randomly selected from other firms and hence has higher leverage. The average leverage ratio for the first (second) group is 10.2% (32.5%) book debt divided by total assets. The estimated SDFs from each group are adjusted so that their time series standard deviations are equal to the implied standard deviation of the true SDF, which according to SDF theory and consistent with the tests in section IV.C, is the same for the two groups. The plot depicts 12-month moving averages centered on the first day of the labeled month.



Appendix A
Details of Simulations

Simulations discussed at various points in the paper are described in detail in this appendix.

A.1. Simulations when the Stochastic Discount Factor (SDF) is unique.

Step 1 is to generate a time series sample of "true" SDF realizations of length T. Specifically, we select a gross riskless rate, R_F, (1+the riskless return), and generate the SDF at time t as

$$m_t = \frac{1}{R_F}\exp(\xi_t - \sigma_\xi^2/2) \text{ , (t=1,…,T)} \tag{A-1}$$

where ξ is a IID random variable with mean zero and standard deviation σ_ξ. The exponential in (A-1) has a mean of 1.0 if ξ is normally distributed, which we assume to be the case initially[31] and, in accordance with SDF theory and the absence of arbitrage, (A-1) provides a strictly positive m_t.

In Step 2, initial gross unscaled returns are generated to be strictly positive (thus assuming limited liability) with a pre-specified mean and volatility (which are assumed to be the same for all individual assets); i.e., for asset i,

$$\hat{R}_{i,t} = \mu\exp(\zeta_{i,t} - \sigma_\zeta^2/2) \text{ , (t=1,…T; i=1,…N)} \tag{A-2}$$

where μ is the expected gross return (1 + the net return) and σ_ζ is the standard deviation of the unscaled gross return $\hat{R}$. We find in simulations (in Appendix B, robustness checks section) that imposition of equal means and variances at this stage has an immaterial effect because the final scaled returns used in all subsequent calculations are computed as

$$\mathbf{R_{i,t}} = \frac{\hat{\mathbf{R}}_{\mathbf{i,t}}}{\sum_{\mathbf{t=1}}^{\mathbf{T}} \mathbf{m_t}\hat{\mathbf{R}}_{\mathbf{i,t}}/\mathbf{T}}\,\mathbf{exp}(\vartheta_{\mathbf{i,t}} - \sigma_\vartheta^2/2) \tag{A-3}$$

where ϑ is an IID return perturbation with mean zero and standard deviation σ_ϑ. As required by SDF theory, (A-3) implies that

$$E\left[\frac{1}{T}\sum_{t=1}^{T} m_t R_{i,t}\right] = 1 \text{ .}$$

[31] In Appendix B, robustness checks section, we consider non-normally distributed variation whose simulations are detailed later in this Appendix.

However, because of the perturbations added as shown in (A-3), the sample average return/SDF product, (the expression within brackets above) is not exactly unity and differs from unity by an amount that varies across individual assets.

Final gross returns on N assets are generated independently for T time periods according to (A-3). Consequently, except for their common dependence on the average SDF, the returns in this simulation are uncorrelated with each other. We consider the consequences of this assumption below where we present analogous simulations with correlated returns that are generated by assets that conform to a factor structure.

The final simulation step uses the estimator (equation 9 in the text) with the final returns from (A-3) to obtain $\hat{m}_t$ (t=1,…,T), for comparison with the true values from (A-1), m_t (t=1,…,T).

A.1.a. Divergent Factor Structures

The second set of simulations reported in section I.D of the text first presumes that there are two asset classes that share a common factor but that the second asset class is also driven by a second factor that has no influence on the first asset class. In other words, instead of the uncorrelated returns as in (A-2), we have

$$\hat{R}_{i,t} = \exp[R_F + \beta_{i,1} f_{1,t} + \beta_{i,2} f_{2,t} + \zeta_{i,t} - \phi_i] \tag{A-4}$$

where the exponentiation correction factor is

$$\phi_i = \mathrm{Var}(\beta_{i,1} f_{1,t} + \beta_{i,2} f_{2,t} + \zeta_{i,t}) / 2.$$

The mean return for each individual asset is dictated by the riskless rate, R_F, plus the mean of the first factor, which we assume is equal to a constant risk premium of .6% per period plus the riskless rate of .4% per period. The mean of the second factor is zero along with the mean of the idiosyncratic return, $\zeta_{i,t}$. The time series standard deviation is four percent per period for the factors and for the idiosyncratic return. For assets in the first group, $\beta_{i,2} = 0, \ldots \forall i$, but only their mean is zero for assets in the second group. Otherwise, the cross-sectional standard deviation of both β is 0.1. The mean of the first factor $\beta_{i,1}$ is 1.0 for both asset groups.

A.1.b. Independent Factor Structures

The third set of simulations in I.D assumes a two-factor structure for both groups of assets, but the factors are independent of each other across groups. In this case, both β's in (A-4) are non-zero for most assets. The cross-sectional means of $\beta_{i,1}$ and $\beta_{i,2}$ are 1.0 and zero, respectively. Their cross-sectional standard deviations are both 0.1.

A.2. Simulations when Returns have Thick Tails.

We generate "true" SDFs as in section A.1 with a lognormal distribution as in equation (A-1) and the same panoply of parameters. Initial gross returns are also generated in the same way, as in equation (A-2).

But equation (A-3) is replaced by

$$\mathbf{R}_{i,t} = \frac{\hat{\mathbf{R}}_{i,t}}{\sum_{t=1}^{T} \mathbf{m}_t \hat{\mathbf{R}}_{i,t} / \mathbf{T}} + \vartheta_{i,t} \qquad \text{(A-5)}$$

in which the zero mean IID return perturbation ϑ is now additive and is distributed according to a truncated Cauchy distribution with a scale parameter that varies from .005 to .045 in .005 increments (i.e., nine different values.) The scale parameter is a measure of the Cauchy distribution's spread; it replaces the standard deviation used for the same purpose with the Gaussian. However, it is not associated with a second moment because the Cauchy has an infinite mean and all higher moments are also infinite.

A truncated Cauchy possesses finite moments but still exhibits extreme outcomes compared to a Gaussian. In the simulations here, we truncate the Cauchy tails, retaining only the middle 95% of simulated values.[32] With a 95% truncation and the scale parameters listed above, gross returns are guaranteed to remain strictly positive.

The return perturbation in (A-5) is additive, in contrast to the previously multiplicative lognormal return perturbation as in (A-3). This choice is necessitated by the extremely large positive values, even with truncation, that would result from taking the exponential of a Cauchy variate. We are not aware of a satisfactory method of correcting for the induced bias. In the Gaussian case, one simply subtracts half of the variance (see equations (A-1) through (A-3)), but there is no corresponding correction using the Cauchy scale for the same purpose. An additive return perturbation finesses this difficulty because it is symmetric and not exposed to the amplification of exponentiation.[33]

[32] The simulations first select a cumulative distribution function p-value, a number between zero and 1.0, and then calculate the inverse Cauchy corresponding to that p. If the p is less than .025 or greater than .975, it is discarded and another p is randomly chosen.

[33] Since the Cauchy mean does not exist, one often uses the median, but a Cauchy with median of zero always has an exponentiated median of 1.0. However, the exponentiated truncated Cauchy can have an extremely large mean.

Appendix B
Robustness Checks

In this Appendix, we investigate the qualities of our SDF estimator with alternative assumptions about returns. Sub-section B.1. examines the consequences of thick tails, a phenomenon that is seemingly ubiquitous for financial asset returns. Sub-section B.2. looks in more detail at the impact of returns that are cross-sectionally correlated and have different means and variances. Sub-section B.3. provides further insights about the behavior of our estimator when returns in different groups have diverse factor structures. (These latter two subsections, B.2 and B.3 simply add conformation to the "Demonstrations of Concepts" reported in Section II.C.) Finally, subsection B.4 suggests a refined estimator that may be useful when estimation error is highly heterogeneous.

B.1. Thick-tailed returns

In the simulations of section III, returns are log-normally distributed, so a natural question is whether our SDF estimator behaves as well when returns are characterized by very large or very small returns, well beyond those typically observed under a Gaussian regime. Our estimator does involve a cross-product matrix that contains squared returns, so it might be sensitive to extreme observations.

To examine this issue, we repeat the simulations of III.A holding everything the same except for the return perturbations, which are now assumed to follow a truncated Cauchy distribution. The details are in section A.2 of Appendix A. Table B.1, which corresponds to Table I, presents the results with truncated Cauchy return perturbations. Comparing Panels A and B of the two tables, one observes that the results are virtually unchanged qualitatively and are even more significant with thick-tailed return perturbations. All the variables have the same signs and all the significant variables (which is everything except the riskless rate) are still significant.

There is one change in Panel C, which shows the influence of various parameters on the volatility of the estimated SDFs. In Table B.1, the true SDF's volatility has become significant. In Panel D, which explains the inferred riskless rate, the SDF volatility and the Cauchy return perturbation scale parameter are not significant while the true riskless rate is more significant. Earlier, we speculated that the volatilities might be showing up in Panel D of Table I because of

Jensen's inequality in the riskless rate's reciprocal estimation, but instead, that result appears to be related to multiplicative return perturbations.

We again find no effect from the difference in means fractional component of the MSE. The averages of the standard deviation difference fractional component and the correlation fractional component are similar, 0.191 and 0.808, respectively.

As for the 2,880 means of true and estimated SDFs, they are still very close, with even a slightly higher correlation, 0.9998, and almost identical averages and standard deviations. The average correlation has risen to 0.439 and the maximum and minimum correlations over the 2,880 parameter combinations are now, respectively, 0.995 and -0.409.

In summary, thick-tailed returns do not seem to compromise the qualities of our estimator. Its seeming improvement with thick tails, however, may be partly attributable to the return perturbation being additive rather than multiplicative and to a set of Cauchy scale parameters that rendered the return perturbations less severe. Regardless of such caveats, however, there seems to be little cause for concern when returns exhibit thick tails.

B.2. Correlated Returns with Unequal Means and Volatilities

The simulated returns in section III are independent of one another and have the same expected values and volatilities. Section II-C presented some illustrations with correlated returns that have disparate means and divergent standard deviations. We now extend these illustrations for a wider set of parameters.

Perhaps the simplest way, (and the way we choose), to simulate returns with such characteristics is to employ the venerable market model. Accordingly, we assume that each initial gross return is obtained from the following model:

$$1+\hat{R}_{i,t} = \exp[R_F + \beta_i(R_{M,t} - R_F) + \varsigma_{i,t} - (\beta_i^2\sigma_M^2 + \sigma_\varsigma^2)/2] \tag{B-1}$$

where R_F is the net risk-free rate (not 1+R), $R_{M,t}$ is a normally distributed "market" common return in period t, β_i is the slope coefficient or "beta" for asset i and $\varsigma_{i,t}$ is a normally distributed IID "idiosyncratic" return for asset i in period t. The last term on the right of (B-1), in parentheses, is a volatility correction for exponentiation.

For each set of parameters, we generate a new set of market returns, idiosyncratic returns, and betas. Then, the simulation proceeds as before, making sure that the average initial gross return from (B-1), multiplied by the SDF, is equal to 1.0 and then adding sampling return perturbation as in equation (A-3) of Appendix A, Section A-1.

The betas are assumed to be cross-sectionally normally distributed with a mean of unity and a standard deviation of 0.1, which implies that most betas fall between 0.8 and 1.2. Since the beta is different for each asset, the expected returns vary cross-sectionally as well.

The market returns are assumed to have a mean equal to the risk free rate plus a premium equal to 0.6% per month and a standard deviation of 4% per month, approximately 13.9% per annum. The idiosyncratic returns are assumed to have a standard deviation of 8% per month, so the market model R-square is 20%, which is in the usual range for equities.

Results are reported in Table B.2. They are virtually identical in Panels A and B with the earlier results in Table I of Section III.A. Thus, inducing correlation and different mean returns and volatilities has no impact whatever on the correlations between true and estimated SDFs and on Theil's U_2 statistic. There are some minor differences in Panels C and D, however. The standard deviation of estimated SDFs (Panel C) now shows significance for the true SDF volatility. The inferred riskless rate (Panel D) shows more significance for the true riskless rate and the number of assets and less significance for the return perturbation volatility. However, these differences are relatively small in magnitude.

The other indicators are also very similar, as one would expect given the similar results in Tables I and B.2. For example, the correlations between true and estimated SDFs range from a maximum of 0.961 to a minimum of -.567. The mean difference fractional component of the MSE is very close to zero in all cases (it's maximum is only 0.0015), which implies that there is no material bias in the estimated SDFs.

In summary, returns that are correlated and differ in their means and volatilities present no difficulties for our SDF estimator.

B.3. Returns with Factor Structures

There seems to be a widely held intuitive notion about a connection between the SDF and the factor structure of different asset classes. For instance, if bonds are driven by fewer underlying risk factors than equities, SDFs must, allegedly, be different for bonds and stocks. But this intuition is not valid. Of course, if the bond and stock market are not integrated and there are cross-market arbitrage opportunities, there would not be a unique SDF common to both markets. But this is not directly attributable to their possibly disparate factor structures. We must try to remember that the SDF appears in an integral equation and there is no mathematical reason for it to be influenced by the factor structure of returns. We shall show now that it is not.

To examine the issue for a wider set of parameters than the illustrations in section I.C, this subsection provides additional simulations wherein the returns are generated by a two-factor structure. The results could be compared with the simulations in section B.2 where returns are generated by a single-factor structure. Here, the return generating function is

$$1+R_{i,t}=\exp[R_F+\beta_{i,1}f_{1,t}+\beta_{i,2}f_{2,t}+\varsigma_{i,t}-(\beta_{i,1}^2\sigma_{f_1}^2+\beta_{i,2}^2\sigma_{f_2}^2+\sigma_{\varsigma}^2)/2] \qquad \text{(B-2)}$$

where $f_{1,t}=R_{M,t}-R_F$, the first factor, has the same distribution as the market excess return in equation (B-1), (the single-factor model), and $\beta_{i,1}$ has the same cross-sectional distribution. The second factor in (B-2) is assumed to have the same volatility and mean return as the first factor but $\beta_{i,2}$ has a cross-sectional mean of zero and a standard deviation of 0.1, (which is the same cross-sectional volatility as $\beta_{i,1}$.) Clearly, the returns generated by a two-factor model will have higher volatilities. All other parameters are the same as in section B.2, including the distribution of the regression disturbances. The results are reported in Table B.3.

Comparing Table B.3 for returns with a two-factor structure against Table B.2 where returns have a single-factor structure, we observe that the coefficients and t-statistics of the various parameters are virtually the same in Panels A, B, and C. The correlation between true SDFs and our sample estimates depend in an almost identical fashion to T, N, the true SDF volatility, and the perturbations volatility. The same is true of Theil's U_2 (Panel B) and for the standard deviation of the estimated SDF (Panel C.)

The only material difference seems to be for the inferred riskless rate. As shown in Panel D, N and T have switched places in terms of significance as have the two volatilities. The true riskless rate is still significant though the significance level has fallen along with the adjusted R-square, which, however, is very modest. It seems likely that the differences displayed in Panel D of Tables B.2 and B.3 is partly attributable to sampling error.[34]

The bottom line is that the factor structure of returns has no relevance for our SDF estimator. This is not all that surprising since the SDF, when unique, should be impervious to any conceivable return generating process. It is reassuring, however, to see that it actually is.

B.4. An Improved SDF Estimator for Cross-Sectionally Heteroscedastic Errors[35]

As explained in Section II.B, our SDF estimator minimizes the sum of squared average surprises in the SDF*Gross Return product. The surprise for asset j in period t is $\varepsilon_{j,t} = m_t R_{j,t} - 1$ (equation 3), its average over T periods is $\bar{\varepsilon}_j = (1/T)\sum_{t=1}^{T}\varepsilon_{j,t}$ and the variance of its average, assuming no serial correlation but allowing for non-stationarity, is $\mathrm{Var}(\bar{\varepsilon}_j) = (1/T^2)\sum_{t=1}^{T}\mathrm{Var}_t(\varepsilon_{j,t})$. This suggests that a weighted estimator, with weights proportional to the precisions of each asset's average surprise, might very well have smaller sampling error, particularly when there is suspected to be a wide discrepancy across assets in $\mathrm{Var}(\bar{\varepsilon}_j)$.

The only problem is that the surprise for each asset cannot be observed without knowing the value of the SDF in each period; hence, an iterative approach is required. We implement the iteration as follows: (1) calculate an SDF with the unweighted estimator (9); (2) calculate an estimated $\mathrm{Var}(\bar{\varepsilon}_j)$ for each asset using this initial SDF; (3) weight the returns for asset j and the jth position in the vector of 1.0's by the precision, $1/\sqrt{\mathrm{Var}(\bar{\varepsilon}_j)}$; (4) obtain the new SDF and new measures of surprise with the weighted observations; (5) repeat until there are miniscule changes in the resulting SDF.

We apply this approach with the simulated two-factor structure described in section A.1.b of Appendix A, while adding cross-sectional heteroscedasticity to the perturbation volatility in (A-3.) Specifically, the perturbation volatility for j is $\sigma_{j,\vartheta} = \bar{\sigma}_\vartheta \exp(\tilde{\eta}_j)$, where $\bar{\sigma}_\vartheta$ = .01% is the

[34] Remember that the additional factor adds as much noise as the first factor in the Table B.2 results.
[35] We are indebted to Robert Engle for suggesting this refinement.

average perturbation volatility across assets and $\tilde{\eta}_j$ is a single draw from a normal distribution adjusted so that $E[\exp(\tilde{\eta}_j)] = 1$. We varied the volatility of $\tilde{\eta}_j$ over a range of multiples of $\bar{\sigma}_\vartheta$ and found very little improvement for multiples from 10 to 1000. For example, with a multiple of 500, the correlations between the true SDF and an estimated SDFs improved from .5984 for original unweighted estimator to .5998 with the weighted one.

Of course, things could be different with larger average perturbation volatility and with assets whose SDF estimation errors are grossly heteroscedastic.

Table B.1
Simulated Performance Information for the SDF Estimator with Thick-Tailed Returns

We simulate true SDFs with mean=1/(1+riskless interest rate) and various time series volatilities. Gross asset returns are simulated so that their mean values multiplied by the SDFs are equal to 1.0, but errors perturb their sample values. The errors are generated from a Cauchy distribution with various scale parameters and truncation that retains only the middle 95%. See Appendix A Section A.2. The performance of the SDF estimator is measured by the correlation between true and sample SDFs and by Theil's (1966) U_2 statistic, which is closely related to the mean square prediction perturbation. Linear regressions are reported in Panel A where the dependent variable is the correlation and in Panel B where the dependent variable is U_2. In Panel C, the dependent variable is the sample time series standard deviation of the estimated SDFs. Panel D reports the implied riskless rate from the reciprocals of the estimated SDFs. There are 2,880 parameter combinations, each with an independently-simulated set of true SDFs and returns.

Variable	Coefficient	T-Statistic
A: Correlation between true and estimated SDFs		
T, Time Periods	-2.129E-03	-30.089
N, Assets	2.767E-04	31.286
True SDF Volatility	1.474	34.709
Perturbation Scale	-0.1788	-97.287
Riskless Rate	-0.5534	-0.330
Adjusted R^2	0.813	
B: U_2 from comparing true and estimated SDFs		
T, Time Periods	1.772E-03	59.132
N, Assets	-2.483E-04	-66.289
True SDF Volatility	0.1337	7.434
Perturbation Scale	6.895	88.548
Riskless Rate	-0.4715	-0.663
Adjusted R^2	0.846	
C: Standard Deviation of Estimated SDFs		
T, Time Periods	2.049E-03	47.071
N, Assets	-2.946E-04	-54.145
True SDF Volatility	0.3602	13.793
Perturbation Scale	4.646	41.088
Riskless Rate	-1.518	-1.471
Adjusted R^2	0.709	
D: Riskless Rate Inferred from Estimated SDFs		
T, Time Periods	-1.783E-05	-1.277
N, Assets	1.906E-06	1.092
True SDF Volatility	6.402E-03	0.764
Perturbation Scale	-2.588E-02	-0.713
Riskless Rate	1.395	4.212
Adjusted R^2	0.00600	

Table B.2
Simulated Performance Information for the SDF Estimator
When Returns Are Correlated and Have Unequal Means and Variances

We simulate true SDFs with mean=1/(1 + riskless interest rate) and various time series volatilities. Gross asset returns are simulated so that their mean values multiplied by the SDFs are equal to 1.0, but errors perturb their sample values. The initial returns are lognormal and generated by an underlying one-factor market model with a dispersion in betas, a market index whose mean exceeds the risk-free rate by 0.6% per month and has a volatility of 4% per month. The market model R-square is 0.2. The performance of the SDF estimator is measured by the correlation between true and sample SDFs and by Theil's (1966) U_2 statistic, which is closely related to the mean square prediction error. Linear regressions are reported in Panel A where the dependent variable is the correlation and in Panel B where the dependent variable is U_2. In Panel C, the dependent variable is the sample time series standard deviation of the estimated SDFs. Panel D reports the implied riskless rate from the reciprocals of the estimated SDFs. There are 2,880 parameter combinations, each with an independently-simulated set of true SDFs and returns, including betas, market returns and idiosyncratic returns.

Variable	Coefficient	T-Statistic
A: Correlation between true and estimated SDFs		
T, Time Periods	-1.160E-03	-11.621
N, Assets	1.459E-04	11.692
True SDF Volatility	1.263	18.161
Perturbation Volatility	-1.702	-48.219
Riskless Rate	-1.532	-0.647
Adjusted R^2	0.479	
B: U_2 from comparing true and estimated SDFs		
T, Time Periods	2.093E-03	55.377
N, Assets	-2.912E-04	-61.646
True SDF Volatility	0.162	6.171
Perturbation Volatility	0.620	46.421
Riskless Rate	0.308	0.344
Adjusted R^2	0.780	
C: Standard Deviation of Estimated SDFs		
T, Time Periods	2.948E-03	54.193
N, Assets	-4.210E-04	-61.925
True SDF Volatility	8.702E-02	2.297
Perturbation Volatility	0.496	25.778
Riskless Rate	-0.605	-0.469
Adjusted R^2	0.729	
D: Riskless Rate Inferred from Estimated SDFs		
T, Time Periods	-1.796E-05	-1.330
N, Assets	3.682E-06	2.181
True SDF Volatility	1.946E-02	2.069
Perturbation Volatility	5.584E-03	1.169
Riskless Rate	1.419	4.430
Adjusted R^2	0.011	

Table B.3

Simulated Performance Information for the SDF Estimator When Returns Are Correlated, with Unequal Means and Variances, and have a Two-Factor Structure

We simulate true SDFs with mean=1/(1 + riskless interest rate) and various time series volatilities. Gross asset returns are simulated so that their mean values multiplied by the SDFs are equal to 1.0, but errors perturb their sample values. The initial returns are lognormal and generated by an underlying two-factor model with the same cross-sectional dispersion in both factor betas. The first factor is a market index whose mean exceeds the risk-free rate by 0.6% per month and has a volatility of 4% per month. The second factor has zero mean but also a volatility of 4% per month. The idiosyncratic volatility is the same as in Table B.1. The performance of the SDF estimator is measured by the correlation between true and sample SDFs and by Theil's (1966) U_2 statistic, which is closely related to the mean square prediction error. Linear regressions are reported in Panel A where the dependent variable is the correlation and in Panel B where the dependent variable is U_2. In Panel C, the dependent variable is the sample time series standard deviation of the estimated SDFs. Panel D reports the implied riskless rate from the reciprocals of the estimated SDFs. There are 2,880 parameter combinations, each with an independently-simulated set of true SDFs and returns, including betas, market returns and idiosyncratic returns.

Variable	Coefficient	T-Statistic
A: Correlation between true and estimated SDFs		
T, Time Periods	-1.182E-03	-12.010
N, Assets	1.535E-04	12.482
True SDF Volatility	1.237	18.048
Perturbation Volatility	-1.667	-47.928
Riskless Rate	2.262	0.097
Adjusted R^2	0.479	
B: U_2 from comparing true and estimated SDFs		
T, Time Periods	2.084E-03	54.679
N, Assets	-2.943E-04	-61.778
True SDF Volatility	0.143	5.380
Perturbation Volatility	0.611	45.307
Riskless Rate	-0.224	-0.248
Adjusted R^2	0.775	
C: Standard Deviation of Estimated SDFs		
T, Time Periods	2.938E-03	53.354
N, Assets	-4.230E-04	-61.447
True SDF Volatility	5.665E-02	1.477
Perturbation Volatility	0.487	25.017
Riskless Rate	-0.933	-0.715
Adjusted R^2	0.723	
D: Riskless Rate Inferred from Estimated SDFs		
T, Time Periods	-4.632E-05	-3.418
N, Assets	-6.477E-07	-0.382
True SDF Volatility	5.391E-03	0.571
Perturbation Volatility	1.048E-02	2.187
Riskless Rate	0.780	2.426
Adjusted R^2	0.00735	