An Information-Theoretic Asset Pricing Model *

Anisha Ghosh[†] Christian Julliard[‡] Alex P. Taylor[§]

March 13, 2019

Abstract

We show that a non-parametric estimate of the pricing kernel, extracted using an information-theoretic approach, delivers smaller out-of-sample pricing errors and a better cross-sectional fit than leading factor models. The information SDF (I-SDF) identifies sources of risk not captured by standard factors, generating very large annual alphas (10%-18%) and Sharpe ratios (0.90-1.3). I-SDFs extracted from a wide cross-section of equity portfolios are highly positively skewed and leptokurtic, and imply that about half of the observed risk premia represent a compensation for tail risk. The I-SDF offers a powerful benchmark relative to which competing theories and investment strategies can be evaluated.

Keywords: Pricing Kernel, Relative Entropy, Cross-sectional Asset Pricing, Factor Models, Factor Mimicking Portfolios, Alpha.
JEL Classification Codes: G11, G12, C13, C53

---

*We benefited from helpful comments from Andrew Ang, George Constantinides, Magnus Dahlquist, Francis Diebold, Ralph Koijen, Dong Lou, Ian Martin, Toby Moskowitz, Christopher Polk, Tarun Ramadorai, Robert Stambaugh, Romeo Tedongap, Raman Uppal, Jessica Wachter, Irina Zviadadze, and seminar participants at Wharton, ICEF, Imperial College Business School, LSE, Rotman School of Management, Stockholm School of Economics, ESSFM 2015 Conference in Gerzensee, and the EEA 2015 Annual Meeting. Any errors or omissions are the responsibility of the authors. Christian Julliard thanks the Economic and Social Research Council (UK) [grant number: ES/K002309/1] for financial support.

[†]Desautels Faculty of Management, McGill University; anisha.ghosh@mcgill.ca.

[‡]Department of Finance and FMG, London School of Economics, and CEPR; c.julliard@lse.ac.uk.

[§]Department of Finance, Manchester Business School; alex.taylor@mbs.ac.uk.

"Entropy, an alternative to variance as a measure of randomness, is playing an increasingly important role in asset pricing theory"

— John Y. Campbell (2014)

I Introduction

Asset prices contain information about the stochastic discounting of possible future states, i.e., about the pricing kernel, or stochastic discount factor (SDF). Based on this simple observation, and an information-theoretic approach, we propose a novel non-parametric method for the estimation of the pricing kernel, and we evaluate its out-of-sample empirical performance in pricing assets.

The proliferation of risk factors identified in the empirical asset pricing literature has brought forth concerns over data mining and spurious inference (see, e.g., Lewellen, Nagel, and Shanken (2010), Harvey and Liu (2015), McLean and Pontiff (2016), Bryzgalova (2016)), and highlights the risk of over-parameterization of the pricing kernel. Therefore, a non-parametric approach to the recovery of the pricing kernel is a potentially valuable alternative to the ad-hoc construction of risk factors. Moreover, given its strong empirical performance, it provides a benchmark model relative to which competing theories, as well as investment managers, can be evaluated.

The use of information-theoretic (or entropy based) techniques has become increasingly popular since their introduction in the finance literature by Stutzer (1995, 1996) and Kitamura and Stutzer (2002). Recent research has found the information-theoretic approach to be a useful tool for assessing the empirical plausibility of various paradigms in asset pricing. Examples include Julliard and Ghosh (2012), who rely on this entropy based inference approach to assess the empirical plausibility of the rare disasters hypothesis in explaining the equity premium puzzle; Backus, Chernov, and Zin (2014), who propose two entropy-based performance measures to assess candidate asset pricing models (see also Bansal and Lehmann (1997)); Borovicka, Hansen, and Scheinkman (2016), who propose using relative entropy minimization techniques to isolate a positive martingale component of the SDF process that contains information about long-term risk adjustments (see also Alvarez and Jermann (2005)); Ghosh, Julliard, and Taylor (2017), who use a relative entropy minimization approach to derive entropy bounds for the SDF that are tighter and more flexible than the seminal Hansen and Jagannathan (1991, 1997) bounds, as well as recover the unobservable components of the SDF (e.g. habits, return on total wealth, etc.) for a broad class of consumption-based asset pricing models.

In the above literature, the information-theoretic technique is used to construct an SDF *in sample*, which, by construction, correctly prices the cross-section of assets used in its estimation. Although this provides a very useful tool for the analysis of structural models, the fact that the test assets are priced perfectly means that standard tests of empirical asset pricing, where the success or failure of an asset pricing model is usually decided, are not meaningful. Indeed, it is an open question whether SDFs extracted using information-theoretic methods have, by the standards of conventional empirical asset pricing tests, any real ability to price assets out-of-sample. Given the increasing use of entropy techniques to assess the empirical plausibility of asset pricing models, we take this question to be of considerable importance. Therefore, in this paper we assess the ability of the entropy-based SDF to price broad cross sections of assets *out-of-sample*, and compare its performance with those of the leading factor models popular in the empirical asset pricing literature. Our results suggest that the entropy based SDF offers a powerful benchmark for asset pricing.

Building upon Ghosh, Julliard, and Taylor (2017), we show how the pricing kernel can be estimated in a non-parametric fashion using no arbitrage restrictions. Specifically, given the time series data of returns on a cross-section of assets, we utilize a model-free relative entropy minimization approach to estimate an SDF that prices the cross-section. The resulting SDF is a non-linear function of the asset returns and the Lagrange multipliers associated with the assets' cross-sectional pricing restrictions (i.e. the shadow value of relaxing the Euler equation restrictions).

We extend the SDF *out-of-sample* for the purposes of cross-sectional pricing and optimal asset allocation. In particular, using the Lagrange multipliers estimated in a training sample, we construct the out-of-sample SDF in a rolling fashion, and use it as the single factor to price the cross-section of test assets. Our approach does not require taking a stance on either the number or the identity of the underlying risk factors or on the functional form of the pricing kernel. Instead, the approach summarizes all the relevant pricing information (contained in, potentially multiple, priced risk factors) in the form of a single time series for the SDF. We refer to the out-of-sample SDF as the 'Information SDF' (I-SDF). The question then becomes whether this non-parametric approach to the recovery of the pricing kernel provides meaningful information for the pricing of assets out-of-sample, and if so, whether it can be considered a valuable alternative to the ad-hoc construction of risk factors commonly used as proxies for underlying sources of priced risk.

We estimate the I-SDF for diverse sets of equity portfolios that capture several well known asset pricing anomalies – including portfolios sorted on the basis of size, book-to-market-equity, momentum, industry, and long-term reversals – and analyze its ability to explain the cross-section of returns. Compared to leading multifactor models, such as the Fama–French 3-factor model (FF3), the Carhart 4-factor model (which adds the momentum factor to the FF3), or the more recent Fama–French 5-factor model (FF5), the I-SDF delivers smaller pricing errors on all the different sets of test assets despite being only a one-factor model. Moreover, it explains a larger fraction of the cross-sectional variation of the returns. These results hold for a variety of measures commonly used in the literature to assess the cross-sectional fit in addition to the standard OLS R^2. Most importantly, we show that the I-SDF (compared to the other factor models considered) more closely identifies – out-of-sample – the tangency portfolio, i.e. the maximum Sharpe ratio portfolio. Furthermore, we find that the I-SDF extracts novel pricing information not captured by the FF3, Carhart 4-factor, or FF5 models: it leads to an 'information anomaly,' generating large and statistically significant intercepts (14.0%–15.6% per annum at the monthly frequency) relative to these factor models, which only explain less than one-fourth of its time series variation. Finally, while the I-SDF is a nonlinear function of the asset returns and is, therefore, not tradeable, the tradeable I-SDF mimicking portfolio also generates statistically and economically large monthly Sharpe ratios (0.91-1.3 annualized) and αs (10.6%–18.6% annualized) with respect to the FF3, FF5, and momentum factors.

We also offer an economic interpretation of the I-SDF. Comparing the I-SDFs extracted from the different sets of equity portfolios reveals a strong commonality.[1] This commonality stems largely from the strong business cycle pattern of all the I-SDFs: the I-SDFs are smaller during expansions and larger during recessions. The first principal component, that explains between 45%-62% of the time series variation of the I-SDFs, also shows a clear business cycle pattern. However, the I-SDFs also exhibit another interesting commonality over and above the business cycle pattern: they are particularly high when the recessions are concomitant with big downward movements in the stock market, such as during the 1973-1975, 2001, and 2007-2009 recessions (unlike the 1981-1982 and 1990-1991 recessions that were not accompanied by big stock market crash episodes). This feature of the I-SDFs is also well captured by their first principal component. Our results suggest that while business cycle risk is an important source of systematic risk, it cannot fully explain the behavior of the I-SDFs, and that the true underlying SDF may not be a function of business cycle variables alone, but also of additional variables related to the performance of the overall stock market.

Interestingly, the I-SDFs, regardless of the cross section of assets they are extracted from, have a strongly non-Gaussian distribution: they are highly positively skewed and leptokurtic.[2] Furthermore, they imply that about one half of the observed risk premia of stocks represent a compensation for tail risk – a feature that is, once again, remarkably robust to the cross section of assets used to construct the I-SDF.

Finally, the impressive performance of the I-SDF at pricing cross sections of portfolio

[1] The average pairwise correlation between the I-SDFs is about 32% and 47%, respectively, at the monthly and quarterly frequencies.

[2] With the skewness coefficient varying from 3.47 to 18.95 (5.49-10.31), and the excess kurtosis varying from 20.45 to 420.5 (30.75 to 125.0), at the monthly (quarterly) frequency.

returns out of sample is not limited to the US equity market. We show that the I-SDF has comparable (or even better) performance at pricing other asset classes such as currencies and commodity futures, as well as international equity portfolios.

Our paper contributes to the extensive cross-sectional asset pricing literature that seeks to identify priced risk factors to explain the cross section of returns of different classes of financial assets. Harvey, Liu, and Zhu (2015) documents 316 risk factors discovered by academics. Lewellen, Nagel, and Shanken (2010) offer a critical assessment of asset pricing tests and conclude that although many of the proposed factors seem to perform well in terms of producing high cross-sectional R^2 and small pricing errors, this result is largely driven by the strong factor structure of the size and book-to-market-equity sorted portfolio returns (which are often used as the only test assets), which makes it quite likely for an arbitrarily chosen two or three factors, which have little correlation with the returns, to produce these results. Moreover, Bryzgalova (2016) shows that the apparent good performance of several factor models proposed in the literature might be the spurious outcome of a weak identification problem. We show that our I-SDF is robust to these concerns and that our approach provides a reliable benchmark against which competing models can be evaluated.

In spirit, our paper is close to, and builds upon, the long tradition of using asset prices to estimate the risk neutral probability measure (see, e.g. Jackwerth and Rubinstein (1996) and Ait-Sahalia and Lo (1998)) and use this information to extract an implied pricing kernel (see, e.g. Ait-Sahalia and Lo (2000), Hansen (2014), Rosenberg and Engle (2002), and Ross (2015)). The main advantages of our approach relative to this literature are that *a*) we do not need to rely exclusively on options data that are only available over a much shorter sample period, and *b*) we can construct an out-of-sample pricing kernel and maximum Sharpe ratio portfolio. Moreover, as we show, our method is very general and can be applied to other asset classes including currencies, commodities, as well as international equities.

Note that our paper focuses on establishing the superior out-of-sample performance of the relative entropy minimization approach in recovering the underlying SDF. Nevertheless, with the SDF in hand, one could address several open questions in macroeconomics and finance without taking a stand on preferences and data generating processes. For example, the I-SDF can be used to: measure the welfare cost of aggregate economic fluctuations;[3] estimate mutual and hedge funds alphas without the fear of results being driven by omitted factors and/or non-linear trading strategies;[4] assess the degree of international risk sharing across countries in a model free manner; shed light on how the pricing kernels constructed from different asset classes (e.g., stocks versus bonds) differ from one another, thereby offering guidance regarding the reasons (if any) for market incompleteness and segmentation.

The remainder of this paper is organized as follows. Section II describes our method of extracting the pricing kernel from a vector of asset returns, as well as the different inference methods used in the empirical analysis. The data used in the empirical analysis are described in Section III. The empirical performance of the I-SDF in explaining broad cross-sections of returns is presented in Section IV. Section V discusses the properties of the I-SDF. Section VI concludes with suggestions for future research.

II The Method

Our relative-entropy minimizing approach enables us to recover, for a given cross-section of assets, what we refer to as the *Information SDF* (I-SDF). Section II.1 describes the information-theoretic method used to construct the I-SDF. Section II.2 discusses the econometric tests used to assess the pricing performance of the I-SDF and compare its performance to some leading empirical asset pricing models commonly used in the literature.

[3] A topic that has so far largely been addressed by relying on parametric assumptions about investors preferences and the dynamics of the economy, and for which the results have proven to be quite sensitive to the precise assumptions.

[4] See e.g. Cochrane (2011).

II.1 Recovery of the Information SDF

The absence of arbitrage opportunities implies the existence of a strictly positive pricing kernel (also known as the stochastic discount factor), M, such that the expectation of the product of the kernel and a vector of excess returns, $\mathbf{R}_t^e \in \mathbb{R}^N$, is zero under the physical probability measure, $\mathbb{P}$:

$$\mathbf{0} = \mathbb{E}^{\mathbb{P}}\left[M_t \mathbf{R}_t^e\right] = \int M_t \mathbf{R}_t^e d\mathbb{P},$$

where $\mathbf{0}$ denotes a conformable vector of zeros. Under weak regularity conditions, the above restrictions on the SDF can be rewritten as

$$\mathbf{0} = \int \frac{M_t}{\bar{M}} \mathbf{R}_t^e \, d\mathbb{P} = \int \mathbf{R}_t^e \, d\mathbb{Q} \equiv \mathbb{E}^{\mathbb{Q}}\left[\mathbf{R}_t^e\right], \tag{1}$$

where $\bar{x} := \mathbb{E}\left[x_t\right]$, and $\frac{M_t}{\bar{M}} = \frac{d\mathbb{Q}}{d\mathbb{P}}$ is the Radon–Nikodym derivative of $\mathbb{Q}$ with respect to $\mathbb{P}$. This change of measure is legitimate if the measure $\mathbb{Q}$ is absolutely continuous with respect to $\mathbb{P}$.

Given the above, an estimate of the risk neutral probability measure $\mathbb{Q}$ can be obtained as the minimizer of its relative entropy with respect to the physical measure $\mathbb{P}$, i.e. as[5]

$$\arg\min_{\mathbb{Q}} D\left(\mathbb{Q}||\mathbb{P}\right) \equiv \arg\min_{\mathbb{Q}} \int \frac{d\mathbb{Q}}{d\mathbb{P}} \ln\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right) d\mathbb{P} \quad \text{s.t.} \quad \int \mathbf{R}_t^e \, d\mathbb{Q} = \mathbf{0}, \tag{2}$$

where $D\left(\mathbb{A}||\mathbb{B}\right) := \int \ln \frac{d\mathbb{A}}{d\mathbb{B}} d\mathbb{A} \equiv \int \frac{d\mathbb{A}}{d\mathbb{B}} \ln \frac{d\mathbb{A}}{d\mathbb{B}} d\mathbb{B}$ denotes the relative entropy of $\mathbb{A}$ with respect to $\mathbb{B}$, i.e. the Kullback–Leibler Information Criterion (KLIC) divergence between $\mathbb{A}$ and $\mathbb{B}$ (White (1982)). Note that $D\left(\mathbb{A}||\mathbb{B}\right)$ is always non-negative, and has a minimum at zero that is attained when $\mathbb{A}$ is identical to $\mathbb{B}$. This divergence measures the additional information content of $\mathbb{A}$ relative to $\mathbb{B}$ and, as pointed out by Robinson (1991), is very sensitive to any

[5]Minimizing the relative entropy to recover the risk neutral probability measure was first suggested by Stutzer (1995). Ghosh, Julliard, and Taylor (2017) extended the method to recover the unobserved component of the SDF for a broad class of consumption-based asset pricing models as well as to construct entropy bounds on the SDF and its components that are tighter and more flexible than the seminal Hansen–Jagannathan bounds.

deviation of one probability measure from another. Therefore, the optimization in equation (2) is a relative entropy minimization under the asset pricing restrictions coming from the Euler equations (1).

There are many different functions that could be used to measure the divergence between two probability measures ($\mathbb{Q}$ and $\mathbb{P}$ in this case). Equation (2) relies on one particular choice of this function. In fact, if the function is chosen to be quadratic, minimizing the divergence between $\mathbb{P}$ and $\mathbb{Q}$ is akin to the continuous updating GMM estimator. An important advantage of using the function in equation (2) is that the objective function can be re-written as a weighted sum of *all* the moments of the distribution of $\log(\mathbb{Q})$, whereas the quadratic choice of the function is akin to minimizing the variance. If the underlying pricing kernel is not lognormal, its variance would not be a sufficient statistic for its distribution and there is no obvious reason to estimate it to minimize the variance alone.

Ghosh, Julliard, and Taylor (2017) show that the above approach to the recovery of the pricing measure has several desirable properties. First, the estimation in equation (2) delivers a non-parametric maximum likelihood estimate of the risk neutral measure. A formal proof of this property of the estimator is as follows. Consider the following procedure for constructing the series $\{q_t\}_{t=1}^T$ (up to a positive scale). Given an integer $N >> 0$, distribute randomly to the various points in time $t = 1, ..., T$, the value $1/N$ in N independent draws. That is, draw a series of probability weights $\{\tilde{q}\}_{t=1}^T$, given by $\tilde{q}_t \equiv \frac{n_t}{N}$, where n_t measures the number of times that the value $1/N$ has been assigned to time t. Subsequently, check whether the drawn series $\{\tilde{q}\}_{t=1}^T$ satisfies the asset pricing restrictions $\sum_{t=1}^T R_t^e \tilde{q}_t = 0$. If it does, use this series as the estimator of $\{q_t\}_{t=1}^T$, and if it does not, draw another series. In other words, an estimate for q_t would correspond to the most likely outcome of the above procedure. Noticing that the distribution of the $\tilde{q}_t$ is, by construction, a multinomial distribution with support given by the historical sample, the likelihood of any particular sequence $\{\tilde{q}_t\}_{t=1}^T$ is given by:

$$L\left(\{\tilde{q}_t\}_{t=1}^T\right) = \frac{N!}{n_1!n_2!...n_T!} \times T^{-N} = \frac{N!}{N\tilde{q}_1!N\tilde{q}_2!...N\tilde{q}_T!} \times T^{-N}.$$

Hence, as $N \to \infty$ (and, therefore, the approach becomes more accurate), the log likelihood is given by[6]

$$\lim_{N\to\infty} \ln L\left(\{\tilde{q}_t\}_{t=1}^T\right) = -\sum_{t=1}^T \tilde{q}_t \ln \tilde{q}_t.$$

Therefore, taking into account the asset pricing constraints, the MLE of q_t solves

$$\{\hat{q}_t\}_{t=1}^T \equiv \arg\max - \sum_{t=1}^T \tilde{q}_t \ln \tilde{q}_t, \qquad \text{s.t. } \{\widetilde{q}_t\}_{t=1}^T \in \Delta^T, \ \sum_{t=1}^T \mathbf{R}_t^e \widetilde{q}_t = \mathbf{0}.$$

Note that the solution of the above likelihood maximization problem is also the solution of the relative entropy minimization problem in Equation (2) (see, e.g., Csiszar (1975)). Therefore, the KLIC minimization is equivalent to maximizing the likelihood for finding the risk neutral measure q_t.

The second desirable property of this approach to the recovery of the risk neutral measure is that, due to the presence of the logarithm in the objective function in equation (2), the use of relative entropy naturally enforces the non-negativity of the pricing kernel. Third, the approach satisfies Occam's razor, or the law of parsimony, since it adds the minimum amount of information needed for the pricing kernel to price assets. Fourth, it is straightforward to add conditioning information: given a vector of conditioning variables $\mathbf{Z}_{t-1}$, one simply has to multiply (element by element) the argument of the integral constraint in equation (2) by the conditioning variables in $\mathbf{Z}_{t-1}$. Fifth, there is no ex ante restriction on the number of assets that can be used in constructing M.[7] Sixth, as implied by Brown and Smith (1990), the use of entropy is desirable if one believes that tail events are an important component of the risk measure (minimum entropy estimators endogenously re-weight the observations to

[6]Recall that from Stirling's formula, we have

$$\lim_{N\tilde{q}_t\to\infty} \frac{N\tilde{q}_t!}{\sqrt{2\pi N\tilde{q}_t}\left(\frac{N\tilde{q}_t}{e}\right)^{N\tilde{q}_t}} = 1.$$

[7]The approach does not require a decomposition of M into short- and long-run components (cf. Alvarez and Jermann (2005)), and it does not rely on the existence of a continuum of options price data (cf. Ross (2015)).

appropriately account for tail events that happened to occur in the data with a frequency lower than their true probability).[8]

In this paper we focus on the out-of-sample asset pricing and investment performance of an SDF constructed using the above relative entropy minimization approach. In particular, note that since $\frac{M_t}{\overline{M}} = \frac{d\mathbb{Q}}{d\mathbb{P}}$, the optimization in equation (2) can be rewritten as

$$\arg\min_{M_t} \ \mathbb{E}^{\mathbb{P}}\left[M_t \ln M_t\right] \quad \text{s.t. } \mathbb{E}^{\mathbb{P}}\left[M_t \mathbf{R}_t^e\right] = \mathbf{0},$$

where, to simplify the exposition, we have used the normalization $\overline{M} = 1$ without loss of generality.[9] Given a sample of size T and a history of excess returns $\{\mathbf{R}_t^e\}_{t=1}^T$, the above expression can be made operational by replacing the expectation with a sample analogue, as is customary for moment based estimators,[10] obtaining

$$\arg\min_{\{M_t\}_{t=1}^T} \ \frac{1}{T}\sum_{t=1}^{T} M_t \ln M_t \quad \text{s.t. } \frac{1}{T}\sum_{t=1}^{T} M_t \mathbf{R}_t^e = \mathbf{0}. \tag{3}$$

The above formulation is convenient because a solution is easily obtainable via Fenchel's duality (see, e.g. Csiszar (1975)):

$$\widehat{M}_t \equiv M_t\left(\widehat{\theta}_T, \mathbf{R}_t^e\right) = \frac{T e^{\widehat{\theta}_T' \mathbf{R}_t^e}}{\sum_{t=1}^{T} e^{\widehat{\theta}_T' \mathbf{R}_t^e}}, \quad \forall t, \tag{4}$$

where $\widehat{\theta} \in \mathbb{R}^N$ is the vector of Lagrange multipliers that solve the unconstrained convex

[8] Based on this insight, Julliard and Ghosh (2012) used a relative entropy estimation approach to analyse the empirical plausibility of the rare events hypothesis to explain a host of asset pricing puzzles.

[9] This normalization is innocuous since the estimate of M_t is identified up to a strictly positive scale constant. This positive scale constant can be recovered from the Euler equation for the risk free rate as follows. The estimate $\widehat{M}_t = kM_t$, where M_t denotes the true underlying SDF and k is a positive constant. Therefore, $E\left[M_t\right] = E\left[\widehat{M}_t\right]/k = E\left[1/R_{f,t}\right]$. Hence, the constant k can be recovered as $k = E\left[\widehat{M}_t\right]/E\left[1/R_{f,t}\right]$.

[10] This amounts to assuming ergodicity for both the pricing kernel and asset returns.

problem

$$\widehat{\theta}_T := \arg\min_{\theta} \frac{1}{T}\sum_{t=1}^{T} e^{\theta' \mathbf{R}_t^e}, \tag{5}$$

and this last expression is the dual formulation of the entropy minimization problem in equation (3). The above duality result implies that the number of free parameters available in estimating $\{M_t\}_{t=1}^T$ is equal to the dimension of (the Lagrange multiplier) θ: that is, it is simply equal to the number of assets considered in the Euler equation.[11]

We use the above method to recover the time series of the SDF in a rolling out-of-sample fashion. In particular, for a given cross section of asset returns, we divide the time series of returns into rolling subsamples of length $\overline{T}$ and final date T_i, $i = 1, 2, 3, ...$, and constant $s := T_{i+1} - T_i$. In subsample i, we estimate the vector of Lagrange multipliers $\widehat{\theta}_{T_i}$ by solving the minimization in equation (5). Using the estimates of the Lagrange multipliers, $\widehat{\theta}_{T_i}$, the out-of-sample *Information SDF* (I-SDF) $M\left(\widehat{\theta}_{T_i}, \mathbf{R}_t^e\right)$ is obtained for the subsequent s periods (i.e. for t such that $T_i + 1 \leq t \leq T_{i+1}$) using equation (4). This process is repeated for each subsample to obtain the time series of the estimated kernel over the out-of-sample evaluation period.

This procedure is analogous in spirit to the canonical approach of forming portfolios (e.g. the SMB and HML portfolios) based on past asset return characteristics (e.g. by sorting on size and book-to-market-equity in the past calendar year). The key difference is that $M\left(\widehat{\theta}_{T_i}, \mathbf{R}_t^e\right)$ is a non-linear function of the portfolio $\widehat{\theta}'_{T_i}\mathbf{R}_t^e$ and the weights θ are chosen to deliver an MLE of the SDF in each (past) subsample.

In the empirical analysis, we set $s = 12$ months (4 quarters) for monthly (quarterly) data. This corresponds to an annual rebalancing of the portfolio. The size of the rolling window, $\overline{T}$, is set to 30 years.

---

[11] Note that since relative entropy is not symmetric, i.e., $D(\mathbb{Q}||\mathbb{P}) \neq D(\mathbb{P}||\mathbb{Q})$, we can reverse the roles of the probability measures $\mathbb{P}$ and $\mathbb{Q}$ in equation (2) to obtain an alternative definition of relative entropy and, therefore, a second approach to estimating the pricing kernel. This approach is described in Appendix A.1. The empirical results obtained using this approach are virtually indistinguishable from those obtained using equation (2) and are, hence, omitted for brevity.

II.2 Asset Pricing Tests

For a given cross-section of test assets, we construct the out-of-sample I-SDF using the procedure described in Section II.1. We evaluate the empirical performance of the I-SDF at monthly and quarterly frequencies. We compare the performance of the I-SDF to that of the one-factor CAPM, the three and five factor Fama–French models, and the Carhart four factor model (that adds the momentum factor to the three Fama-French factors).

We use the two-step method of Fama and MacBeth (1973) to assess the ability of each factor model to price the cross-section of test assets. In the first step, the factor loadings for the test assets are estimated from a time series regression of the excess returns on the factors:

$$\mathbf{R}_t^e = a + BF_t + \varepsilon_t.$$

In the second step, the factor risk premia are obtained from a cross-sectional regression of the average excess asset returns, $\mu \in \mathbb{R}^N$, on the factor loadings estimated from the first stage:

$$\mu = z\iota + B\gamma + \alpha = C\lambda + \alpha, \quad C := [\iota \ B]\,, \, \lambda' := [z \ \gamma']\,,$$

where ι denotes a conformable vector of ones, γ denotes a vector of regression slopes (that should be non-zero if the factors are priced), z is a scalar constant (that should be zero if the zero-beta rate matches the risk-free rate), and $\alpha \in \mathbb{R}^N$ is the vector of pricing errors (that should be zero if the factors price assets accurately).

Following the suggestions of Lewellen, Nagel, and Shanken (2010), we present several alternative measures of performance for the above cross-sectional regressions. First, we present the standard OLS cross-sectional adjusted R^2 (hereafter denoted by $\overline{R}^2_{OLS}$). This measure suffers from the shortcoming that if the returns have a strong factor structure (such as, e.g., the size and book-to-market-equity sorted portfolio returns), then an arbitrarily chosen set of two or three factors, that have little correlation with the returns, are quite likely to produce large values of this statistic. This is obviously less of an issue for our I-SDF since it is a one-factor model, but it is likely to affect the performance of the other factor models that we consider for comparison.

Second, we present the GLS adjusted R^2 (hereafter denoted by $\overline{R}^2_{GLS}$) that is obtained from the cross-sectional regression of $\widehat{V}^{-1/2}\mu$ on $\widehat{V}^{-1/2}[\iota\ B]$, where $V := Var(\mathbf{R}^e)$. The $\overline{R}^2_{GLS}$ for a model, unlike $\overline{R}^2_{OLS}$, is completely determined by the model-implied factors' proximity to the minimum variance frontier and, in general, presents a more stringent hurdle for models (Lewellen, Nagel, and Shanken (2010)).

Third, we present the cross-sectional T^2 statistic of Shanken (1985), given by $T^2 := \widehat{\alpha}' S_a^{+} \widehat{\alpha}$, where S_a^{+} is the pseudoinverse of the estimated $\Sigma_a := \left(1 + \gamma' \Sigma_F^{-1} \gamma\right) \frac{y \Sigma y}{T}$, $y := I - C\left(C'C\right)^{-1} C'$ and $\Sigma := Var(\varepsilon_t)$. The T^2 statistic has an asymptotic χ^2 distribution with $N - K - 1$ degrees of freedom, where K denotes the number of factors, and noncentrality parameter $\alpha' \Sigma_a^{+} \alpha = \alpha' \left(y \Sigma y\right)^{+} \alpha \frac{T}{\left(1 + \gamma' \Sigma_F^{-1} \gamma\right)}$, where Σ_F denotes the covariance matrix of the factors. We compute the p-value of this statistic under the null hypothesis that the model explains the vector of expected returns perfectly, i.e., the vector of pricing errors $\alpha = 0$.

Fourth, we present the quadratic $q := \alpha' \left(y \Sigma y\right)^{+} \alpha$ which measures how far a candidate model's factors are from the mean–variance frontier.[12] In particular, it is equal to the difference between the squared Sharpe ratio of the tangency portfolio of the test assets and the maximum squared Sharpe ratio attainable from the model-implied factors (or their mimicking portfolios in the case of non-traded factors).

Lastly, we present the simulated 90% confidence intervals for the statistics. The simulated confidence intervals are obtained using the approach suggested by Stock (1991) (see also Lewellen, Nagel, and Shanken (2010) for a detailed discussion). Consider first the construction of the confidence intervals for the $\overline{R}^2_{OLS}$. The simulations have two steps. First, we fix a true (population) cross-sectional R^2 that we want the model to have and alter the $(N \times 1)$ vector of expected returns, μ, to be $\mu = hC\lambda + e$, where $C \equiv [\iota,\ B]$, B denotes the vector of factor loadings in the historical sample, and $e \sim N\left(0, \sigma_e^2\right)$. The constants h and σ_e^2

[12]See Uppal and Zaffaroni (2015) for an alternative economic interpretation of this statistic.

are chosen to produce the right cross-sectional R^2 and maintain the historical cross-sectional dispersion of the average returns. Second, we jointly simulate an artificial time series of the factor and the returns of the same length as the historical data by sampling, with replacement, from the historical time series. We then use the two-pass regression method to estimate the sample cross-sectional R^2 of the simulated sample. We repeat the second step $1,000$ times to construct a sampling distribution of the R^2 statistic conditional on the given population R^2. This procedure is repeated for all values of the population R^2 between 0 and 1. The 90% confidence interval for the true R^2 represents all values of the population R^2 for which the estimated R^2 in the historical sample falls within the 5th and 95th percentiles of the sample distribution.

A confidence interval for q is found using a method similar to that used to obtain the confidence interval for the true (population) cross-sectional R^2. Specifically, a given population R^2 implies a specific value of q. We plot the sample distribution of the T^2 statistic as a function of q. The confidence interval for the true q represents all values of the q for which the estimated T^2 in the historical sample falls within the 5th and 95th percentiles of the sample distribution.

For the T^2 statistic, we present its finite-sample p-value, obtained from the above simulations, as the probability that the T^2 statistics in the simulated samples exceed the value of the statistic in the historical data for $q = 0$.

III Data Description

For most of the US equity portfolios, we assess the empirical performance of the extracted pricing kernel (the I-SDF) at monthly and quarterly frequencies. The out-of-sample evaluation covers the period 1963:07–2017:06. The start date 1963:07 is chosen to coincide with that in Fama and French (1993), Lewellen, Nagel, and Shanken (2010), as well as DeMiguel, Garlappi, and Uppal (2009). This facilitates a useful comparison of our results with the existing literature.

To illustrate the strength of our method, we analyse several cross-sections of equity portfolios.[13] In particular, we consider the 25 size and book-to-market-equity sorted portfolios, the 10 momentum-sorted portfolios, the 30 industry-sorted portfolios, and the 25 portfolios formed on long term reversal and size. We extract the I-SDF from, and use them to price, each of these cross-sections, as well as several combinations of these cross-sections.

Monthly returns data on the above portfolios are obtained from Kenneth French's data library. An estimate of the monthly risk free rate is subtracted from the portfolio returns to produce the excess returns. Our proxy for the risk-free rate is the one-month Treasury Bill rate, also obtained from Kenneth French's data library. The quarterly returns on the equity portfolios as well as the quarterly risk free rate are obtained by compounding the monthly returns within each quarter. The excess returns on the portfolios are then computed by subtracting the risk free rate.

We also evaluate the performance of the I-SDF for other asset classes, as well as additional US equity portfolios available over a shorter sample period. In particular, we consider (a) the 6 currency portfolios from Lustig, Roussanov, and Verdelhan (2011), formed by sorting the currencies of developed and emerging economies on the basis of their forward discounts and rebalancing every month. Monthly returns on these currency portfolios are available from 1983:11 to 2015:06, (b) portfolios of commodity futures from Asness, Moskowitz, and Pedersen (2013), the monthly returns on which are available over the period 1972:02–2010:06, (c) portfolios of global individual stocks from Asness, Moskowitz, and Pedersen (2013), also available over the 1972:02–2010:06 period, and (d) the smallest and largest deciles of US equity portfolios formed by sorting individual US stocks on the basis of accruals, market beta, net equity issuance, variance, and residual variance, the monthly returns on which are available from 1963:07 to 2017:06. Because the returns on the above cross sections (a)-(d) are available for much shorter time periods, we only present results at the monthly frequency.

[13] We focus on portfolios, rather than individual asset returns, since our estimation method requires a large time series dimension relative to the cross-sectional one.

IV Cross-Sectional Pricing

In this section, we evaluate the out-of-sample ability of the I-SDF to explain the cross-section of returns by presenting cross-sectional regression results for different sets of test assets.

Table 1 presents the cross-sectional pricing results when the test assets consist of the 25 size and book-to-market-equity sorted portfolios of Fama and French. Consider first Panel A, which presents the results at the monthly frequency. Row 1 shows that when the I-SDF is used as the sole factor, its estimated price of risk has the correct sign and is strongly statistically significant with an absolute value of the t-statistic in excess of 9. Harvey, Liu, and Zhu (2015) argue that a t-statistic of around 2.0 is too low a hurdle to establish the statistical significance of a given factor in the presence of extensive data mining. Using a new framework that allows for multiple tests, they show that a t-statistic greater then 3.0 would be required for a factor to be deemed as being statistically significant. Row 1 shows that the I-SDF has a t-statistic more than three times the value needed to establish statistical significance even after taking into account the possibility of data mining. Since the regression uses the monthly excess returns as the dependent variable, the intercept can be interpreted as the estimated monthly zero beta rate over and above the risk free rate. The estimated annualized zero beta rate is 3.6%. Although this is statistically significant, part of it may be attributable to the differences in lending and borrowing rates (1%–2%). Moreover, rows 2 and 3 show that the CAPM and the FF3 model produce substantially higher annualized intercepts of 14.4% and 15.6%, respectively. The I-SDF produces an $\overline{R}^2_{OLS}$ of 77.8% and, more importantly, $\overline{R}^2_{GLS}$ is very similar to $\overline{R}^2_{OLS}$, at 69.4%. Note that the GLS R^2 is high if and only if the factor is close to the mean–variance frontier and, in general, provides a more stringent hurdle for asset pricing models. The T^2 statistic shows that the model is not rejected at conventional significance levels. Lastly, the q statistic, which equals the difference between the squared Sharpe ratio of the tangency portfolio of the test assets and the squared Sharpe ratio of the factor-mimicking portfolio, is 0.050 and its 90% confidence interval includes 0, i.e., the I-SDF mimicking portfolio is statistically indistinguishable from

Table 1: 25 Fama–French Portfolios

Row	*const.*	λ_{sdf}	λ_{Rm}	λ_{SMB}	λ_{HML}	$\overline{R}^2_{OLS}$ (%)	$\overline{R}^2_{GLS}$ (%)	T^2	q
				Panel A: Monthly					
(1)	0.003	−0.327				77.8	69.4	27.8	0.050
	(6.98)	(−9.24)				[39.5,100]	[52.4,100]	(0.207)	[0.00,0.09]
(2)	0.012		−0.004			3.95	30.3	74.5	0.116
	(3.79)		(−1.41)			[−4.35,61.4]	[6.43,59.9]	(0.000)	[0.04,0.34]
(3)	0.013		−0.007	0.002	0.004	66.3	40.2	55.3	0.089
	(3.41)		(−1.99)	(3.23)	(6.02)	[21.1,90.9]	[20.5,90.8]	(0.002)	[0.03,0.16]
(4)	−0.001	−0.474	0.007	0.002	0.003	90.2	68.1	21.7	0.045
	(−0.310)	(−7.36)	(2.46)	(8.08)	(10.1)	[50.8,100]	[47.6,100]	(0.311)	[0.00,0.088]
				Panel B: Quarterly					
(1)	0.028	−10.2				35.2	31.3	44.4	0.308
	(17.4)	(−3.74)				[−1.22,100]	[12.3,70.5]	(0.451)	[0.00,0.60]
(2)	0.027		−0.003			−3.53	16.8	80.3	0.372
	(3.44)		(−0.427)			[−4.35,25.9]	[−0.57,43.5]	(0.000)	[0.08,0.97]
(3)	0.037		−0.019	0.006	0.011	70.3	25.0	58.4	0.304
	(3.22)		(−1.74)	(4.24)	(6.16)	[30.3,93.1]	[−7.50,66.3]	(0.003)	[0.08,0.71]
(4)	0.017	−9.60	0.001	0.007	0.010	77.7	30.6	39.1	0.267
	(1.43)	(−2.94)	(0.116)	(5.29)	(6.48)	[44.8,100]	[13.3,100]	(0.391)	[0.00,0.21]

Cross-sectional regressions of average excess returns of the 25 Fama–French portfolios on the estimated factor loadings for different asset pricing models. Panel A presents the monthly results and Panel B the quarterly results. In each panel, the first row presents the results when the factor is the information SDF. The information SDF is extracted from the 25 Fama–French portfolios using a relative entropy minimizing procedure, in a rolling out-of-sample fashion starting at 1963:07. Rows 2 and 3 present the results for the CAPM and the Fama–French 3-factor model, respectively. In row 4 the factors are the three Fama–French factors plus the information SDF. For each model, the table presents the intercept and slopes, along with t-statistics in parentheses. It also presents the OLS adjusted R^2 and the GLS adjusted R^2, along with the 90% confidence intervals for the true underlying population adjusted R^2 (in square brackets). The confidence intervals are constructed via simulations using the approach suggested by Stock (1991) and used by Lewellen, Nagel, and Shanken (2010). The last two columns present, respectively, Shanken's (1985) cross-sectional T^2 statistic along with its asymptotic p-value in parentheses, and the q statistic that measures how far the factor-mimicking portfolios are from the mean–variance frontier.

the maximum Sharpe ratio portfolio of the test assets.

In row 2, we present the results for the unconditional CAPM. The market risk premium has the wrong sign and is not statistically different from zero. The intercept, on the other hand, is strongly significant with an annualized value of 14.4%. The OLS and GLS $\overline{R}^2$ are much smaller at 3.95% and 30.3%, respectively, compared to those obtained with the I-SDF. The T^2 statistic is more than double that obtained with the I-SDF, and has a p-value of zero: i.e. the model is strongly rejected. The q statistic is closely related to the $\overline{R}^2_{GLS}$ and the T^2 statistics and, therefore, not surprisingly, provides similar conclusions: the 90% confidence interval for the q statistic implies a large unexplained Sharpe ratio between 0.20 and 0.58, i.e. the model fails to identify the maximum Sharpe ratio portfolio.

Row 3 presents the results for the FF 3-factor model. The results show that the market risk premium is not statistically significant but the risk premia associated with the factors proxying for risks related to size and book-to-market-equity are both significantly positive. However, the intercept is statistically and economically large, with an annualized value of 15.6%, similar to that obtained with the market risk factor alone in row 2. The $\overline{R}^2_{OLS}$ is high at 66.3% (although much smaller than the 77.8% value obtained with the I-SDF), consistent with existing empirical evidence that the 3 FF factors explain a large fraction of the time series and cross-sectional variation in the returns of the 25 FF portfolios. However, moving to a GLS cross-sectional regression, $\overline{R}^2$ drops sharply to 40.2%, consistent with the observation that a GLS regression offers a more stringent hurdle for models than does the OLS. This is in stark contrast to the I-SDF, which delivers very similar $\overline{R}^2$ using both the OLS and GLS procedures. The T^2 statistic is larger than that obtained with the I-SDF (55.3 vs 27.8), and has a p-value of zero, implying a statistical rejection of the model. The q statistic is also larger than that obtained with the I-SDF (0.089 vs 0.050). Moreover, the 90% confidence interval of the q statistic does not include 0, i.e. the maximum Sharpe ratio obtainable from the 3 FF factors is statistically smaller from the Sharpe ratio of the tangency portfolio of the test assets. Finally, the 95% confidence intervals for the true population $\overline{R}^2_{OLS}$ and $\overline{R}^2_{GLS}$ are tighter for the I-SDF than those obtained with the 3 FF factors. Specifically, the 95% confidence interval for the $\overline{R}^2_{OLS}$ includes the range [21.1, 90.9] for the 3 FF factors, whereas it includes a narrower range of [39.5, 100] for the I-SDF; the corresponding ranges for the $\overline{R}^2_{GLS}$ are [20.5, 90.8] and [52.4, 100] for the 3 FF factors and the I-SDF, respectively. Note that the confidence intervals for both the $\overline{R}^2_{OLS}$ and $\overline{R}^2_{GLS}$ include the highest possible value of 100 for the I-SDF but not for the 3 FF factors.

Row 4 presents the results when the I-SDF is used in conjunction with the 3 FF factors in the cross-sectional regression. Note that the risk premium for the I-SDF remains strongly statistically significant even in the presence of the 3 FF factors and its magnitude is very similar to that obtained when the I-SDF is used as the sole factor in row 1. Although $\overline{R}^2_{OLS}$

is higher, at 90.2% compared to 77.8% in row 1, the $\overline{R}^2_{GLS}$ for the two rows are very similar (68.1% vs 69.4%).

Similar results are obtained at the quarterly frequency in Panel B. The I-SDF delivers a strongly significant λ, the T^2 statistic implies that the pricing model is not rejected, and the low value of the q statistic implies that this factor is close to the the maximum Sharpe ratio portfolio. Nevertheless, the $\overline{R}^2_{OLS}$ and $\overline{R}^2_{GLS}$ are somewhat reduced compared to those at the monthly frequency. The CAPM, on the other hand, produces a negative $\overline{R}^2_{OLS}$, an $\overline{R}^2_{GLS}$ of 16.8%, and a T^2 statistic with a p-value of 0%. For the FF 3-factor, although $\overline{R}^2_{OLS}$ is high at 70.3%, the GLS $\overline{R}^2$ drops sharply, to only 25.0% (whereas that for the I-SDF is 31.3%). Moreover, the T^2 test rejects the FF 3-factor specification while the q statistic suggests that this factor model fails to identify the capital market line (while the I-SDF succeeds in this task). Lastly, combining the information factor with the FF3 factors leaves both the point estimate and the statistical significance of the information factor unaffected.

As shown in Lewellen, Nagel, and Shanken (2010), it is relatively easy to find two of three factors that produce large $\overline{R}^2_{OLS}$ for the 25 FF portfolios because of their strong factor structure. It is noteworthy that a single factor,[14] namely the I-SDF, does even better than the FF3 factors. Moreover, similar conclusions are obtained if, rather than relying on $\overline{R}^2_{OLS}$ alone, more stringent hurdles are imposed on the model via the $\overline{R}^2_{GLS}$, T^2, and q statistics, and their confidence bands.

We next show that the strong performance of our model holds not only for the size and book-to-market-equity sorted portfolios, but also for portfolios formed by sorting stocks on the basis of other characteristics, such as prior returns, industry, etc. Tables 2–4 present the cross-sectional regression results when the set of test assets consists of (a) the 10 momentum

[14]Of course, any multi-factor model can be rewritten as a single factor model (see e.g. Back (2010)), nevertheless this requires knowledge of the projection coefficients that are available only *ex post* to the econometrician. Hence, *ex ante*, the number of factors is a relevant metric for assessing the degrees of freedom that a model has for fitting the data. This point is further illustrated in Section A.2, where we show that a linear combination of the three FF factors (with the coefficients estimated in a rolling out of sample fashion from past data on the chosen cross section of assets) performs substantially worse at pricing the cross section out-of-sample relative to when the coefficients on the three factors are left unrestricted in the cross sectional regressions.

Table 2: 10 Momentum Portfolios

Row	*const.*	λ_{sdf}	λ_{Rm}	λ_{SMB}	λ_{HML}	λ_{MOM}	$\overline{R}^2_{OLS}$	$\overline{R}^2_{GLS}$	T^2	q
					Panel A: Monthly					
(1)	0.004 (15.5)	−0.27 (−10.3)					92.2 [66.3,100]	75.1 [16.0,100]	10.28 (0.325)	0.017 [0.00,0.05]
(2)	0.015 (2.51)		−0.010 (−1.68)				16.8 [−12.5,78.6]	6.4 [−7.85,42.4]	39.26 (0.000)	0.063 [0.03,0.43]
(3)	0.032 (1.97)		−0.022 (−1.37)	−0.003 (−0.24)	−0.031 (−1.27)	0.006 (5.93)	83.0 [−78.2,100]	19.2 [−25.5,92.2]	7.56 (0.386)	0.034 [0.00,0.34]
(4)	0.011 (1.44)	−0.25 (−2.95)	−0.003 (−0.45)	−0.002 (−0.36)	−0.018 (−1.78)	0.005 (11.8)	97.2 [−62.0,100]	79.5 [−38.3,100]	2.83 (0.640)	0.007 [0.00,0.318]
					Panel B: Quarterly					
(1)	0.011 (10.2)	−0.99 (−8.76)					89.4 [16.8,100]	80.7 [63.2,100]	7.14 (0.529)	0.043 [0.00,0.19]
(2)	0.041 (2.87)		−0.024 (−1.77)				19.1 [−12.5,80.9]	3.10 [−6.8,31.2]	40.08 (0.000)	0.200 [0.05,1.15]
(3)	0.105 (1.54)		−0.095 (−1.30)	0.062 (1.51)	0.036 (1.06)	0.020 (5.72)	78.3 [−72.8,94.6]	18.6 [−23.6,77.3]	3.91 (0.188)	0.105 [0.00,0.96]
(4)	0.038 (2.33)	−0.96 (−4.14)	−0.025 (−1.42)	0.036 (3.87)	0.015 (1.90)	0.021 (27.26)	99.0 [41.5,100]	93.5 [12.7,100]	0.58 (0.849)	0.007 [0.00,0.194]

Cross-sectional regressions of average excess returns of the 10 momentum-sorted portfolios on the estimated factor loadings for different asset pricing models. Panel A presents the monthly results and Panel B the quarterly results. The first row in each panel presents the results when the factor is the information SDF. The information SDF is extracted from the 10 momentum-sorted portfolios using a relative entropy minimizing procedure, in a rolling out-of-sample fashion starting at 1963:07. Rows 2 and 3 present the results for the CAPM and the Carhart 4-factor model, respectively. In row 4 the factors are the four Carhart factors plus the information SDF. For each model, the table presents the intercept and slopes, along with t-statistics in parentheses. It also presents the OLS adjusted R^2 and the GLS adjusted R^2, along with the 90% confidence intervals for the true underlying population adjusted R^2 (in square brackets). The confidence intervals are constructed via simulations using the approach suggested by Stock (1991) and used by Lewellen, Nagel, and Shanken (2010). The last two columns present, respectively, Shanken's (1985) cross-sectional T^2 statistic along with its asymptotic p-value in parentheses, and the q statistic that measures how far the factor-mimicking portfolios are from the mean–variance frontier.

Table 3: 25 Portfolios Formed on Long-Term Reversal and Size

Row	*const.*	λ_{sdf}	λ_{Rm}	λ_{SMB}	λ_{HML}	$\overline{R}^2_{OLS}$ (%)	$\overline{R}^2_{GLS}$ (%)	T^2	q
				Panel A: Monthly					
(1)	0.006 (8.06)	−0.23 (−2.38)				16.3 [−4.35,100]	66.9 [47.1,90.2]	21.12 (0.603)	0.034 [0.00,0.03]
(2)	0.005 (1.69)		0.003 (1.05)			0.43 [−4.35,37.4]	15.8 [0.15,33.1]	53.33 (0.001)	0.083 [0.03,0.18]
(3)	0.005 (2.15)		0.001 (0.38)	0.001 (2.15)	0.005 (4.82)	74.6 [34.9,100]	32.0 [1.66,100]	37.48 (0.019)	0.060 [0.01,0.10]
(4)	0.002 (1.29)	−0.368 (−6.70)	0.004 (2.52)	0.004 (6.77)	0.002 (2.45)	90.2 [71.2,100]	75.1 [92.3,100]	11.82 (0.862)	0.021 [0.00,0.005]
				Panel B: Quarterly					
(1)	0.025 (21.6)	−1.06 (−1.00)				−0.09 [−4.35,36.3]	44.0 [21.1,45.3]	38.52 (0.353)	0.184 [0.00,0.44]
(2)	0.011 (1.48)		0.013 (2.03)			11.5 [−4.35,70.8]	5.13 [−2.26,31.0]	59.65 (0.002)	0.283 [0.09,0.68]
(3)	0.012 (1.69)		0.006 (0.77)	0.004 (2.85)	0.016 (4.85)	79.3 [18.9,100]	18.7 [−7.61,84.3]	43.02 (0.018)	0.219 [0.04,0.49]
(4)	0.004 (0.738)	−2.80 (−4.28)	0.014 (2.36)	0.007 (5.20)	0.010 (3.64)	87.7 [58.0,100]	50.6 [−4.38,100]	21.50 (0.292)	0.127 [0.00,0.257]

Cross-sectional regressions of average excess returns of the 25 long term reversal and size sorted portfolios on the estimated factor loadings for different asset pricing models. Panel A presents the monthly results and Panel B the quarterly results. The first row in each panel presents the results when the factor is the information SDF. The information SDF is extracted from the 25 long term reversal and size sorted portfolios using a relative entropy minimizing procedure, in a rolling out-of-sample fashion starting at 1963:07. Row 2 presents the results for the CAPM, and row 3 for the Fama–French 3-factor model. In row 4 the factors are the three FamaFrench factors plus the information SDF. For each model, the table presents the intercept and slopes, along with t-statistics in parentheses. It also presents the OLS adjusted R^2 and the GLS adjusted R^2, along with the 90% confidence intervals for the true underlying population adjusted R^2 in square brackets below. The confidence intervals are constructed via simulations using the approach suggested by Stock (1991) and used by Lewellen, Nagel, and Shanken (2010). The last two columns present, respectively, Shanken's (1985) cross-sectional T^2 statistic along with its asymptotic p-value in parentheses, and the q statistic that measures how far the factor-mimicking portfolios are from the mean–variance frontier.

Table 4: Small, Large, Growth, Value, Winners, Losers, 10 Industry

Row	*const.*	λ_{sdf}	λ_{Rm}	λ_{SMB}	λ_{HML}	λ_{MOM}	$\overline{R}^2_{OLS}$	$\overline{R}^2_{GLS}$	T^2	q
Panel A: Monthly										
(1)	0.002	−1.07					92.9	87.9	6.03	0.013
	(5.65)	(−14.0)					[99.9,100]	[86.2,100]	(0.984)	[0.00,0.013]
(2)	0.008		−0.002				−4.09	3.51	65.23	0.101
	(2.39)		(−0.64)				[−7.14,57.1]	[−3.75,45.0]	(0.000)	[.038,.413]
(3)	0.004		0.002	0.002	0.001	0.009	83.8	39.3	30.43	0.050
	(1.94)		(1.24)	(2.17)	(1.35)	(8.66)	[11.4,100]	[−19.2,91.7]	(0.018)	[0.005,0.262]
(4)	−0.000	−1.27	0.006	0.003	0.0003	0.007	91.8	88.4	3.71	0.009
	(−0.118)	(−3.97)	(3.48)	(4.10)	(0.520)	(8.60)	[94.0,100]	[81.6,100]	(0.955)	[0.00,0.014]
Panel B: Quarterly										
(1)	0.016	−5.12					79.4	62.3	16.22	0.128
	(15.87)	(−7.66)					[56.1,100]	[28.6,100]	(0.555)	[0.00,0.554]
(2)	0.022		−0.004				−5.6	6.4	66.74	0.310
	(2.47)		(−0.45)				[−7.14,61.4]	[0.83,65.9]	(0.000)	[0.093,1.345]
(3)	0.013		0.006	0.005	0.004	0.025	81.8	39.7	28.64	0.157
	(1.87)		(0.86)	(2.29)	(1.52)	(7.88)	[25.0,100]	[−19.9,84.9]	(0.026)	[0.015,1.323]
(4)	0.010	−3.30	0.009	0.007	0.004	0.024	86.4	53.2	18.7	0.111
	(1.63)	(−3.21)	(1.46)	(3.19)	(1.97)	(8.54)	[41.5,100]	[−20.5,100]	(0.118)	[0.00,0.520]

Cross-sectional regressions of average excess returns of the 10 industry portfolios and the top and bottom deciles of portfolios sorted on the basis of size, book-to-market-equity, and momentum on the estimated factor loadings for different asset pricing models. Panel A presents the monthly results and Panel B the quarterly results. In each panel, row 1 presents the results when the factor is the information SDF. The information SDF is extracted from the 10 industry portfolios and the top and bottom deciles of portfolios sorted on the basis of size, book-to-market-equity, and momentum, using a relative entropy minimizing procedure, in a rolling out-of-sample fashion starting at 1963:07. Rows 2 and 3 present the results for the CAPM and the Carhart 4-factor model, respectively. In row 4 the factors are the three Fama-French factors plus the information SDF. For each model, the table presents the intercept and slopes, along with the t-statistics in parentheses. It also presents the OLS adjusted R^2 and the GLS adjusted R^2, along with the 90% confidence intervals for the true underlying population adjusted R^2 (in square brackets). The confidence intervals are constructed via simulations using the approach suggested by Stock (1991) and used by Lewellen, Nagel, and Shanken (2010). The last two columns present, respectively, Shanken's (1985) cross-sectional T^2 statistic along with its asymptotic p-value in parentheses, and the q statistic.

sorted portfolios, (b) the 25 portfolios formed on the basis of size and long-term reversal, and (c) the 10 industry portfolios and the smallest and largest deciles of portfolios formed on the basis of size, B/M, and momentum. The results, in each case, are very similar to those obtained with the 25 FF portfolios in Table 1.

Overall, Tables 2–4 show that: the I-SDF tends to produce smaller pricing errors and larger cross-sectional R^2s than the Fama–French 3-factor and the Carhart 4-factor models, despite being only a one-factor model; the risk premium associated with the I-SDF is statistically significant, even after controlling for the FF and Carhart factors; the T^2 statistic of the I-SDF implies that this factor is never rejected at standard confidence levels (while the other factor models considered are almost always rejected); the q statistic implies that the I-SDF successfully identifies the capital market line, i.e. the mimicking portfolio is statistically undistinguishable from the maximum Sharpe ratio portfolio (while the other factor models considered fail in this respect); in 9 cases out of 12 (or 12 out of 16 if Table 1 is included) the t-statistics of the information factor are larger than 3, hence clearing the higher hurdle for statistical significance recommended by Harvey and Liu (2015); and in 11 cases out of 12 (or 14 out of 16 if Table 1 is included) the 95% confidence intervals for the $\overline{R}^2_{OLS}$ and $\overline{R}^2_{GLS}$ are tighter for the I-SDF than those obtained with the 3 FF or 4 Carhart factors. Moreover, as an additional robustness check of the results in Tables 1 to 4, we have also obtained cross-sectional estimates using the Pen-FM (Penalized Fama–MacBeth) estimator of Bryzgalova (2016), that by design has the ability to detect spurious factors and shrink (in a 'lasso' fashion) their λ's to zero. Using this approach, we found virtually identical results for the information factor to those discussed above.[15]

Note that each of Tables 1–4 focused on US equities with long available histories of data. While US equities is undoubtedly the most widely studied asset class among both academics and practitioners, other asset classes such as currencies and commodities have gained prominence in recent times. Moreover, international financial markets are playing an

[15]We are thankful to Svetlana Bryzgalova for providing us with the necessary computer code to implement this test.

ever increasing role in the global landscape. In Table 5, we provide empirical evidence on the ability of the I-SDF to price currencies, commodities, and international equities out of sample. Because of the relatively short sample over which the returns data on these assets are available, we only present results at the monthly frequency and use a rolling window of 10 (rather than 30) years for the estimation of the I-SDF.

Row 1 presents the results for currency markets. The test assets are the six currency portfolios from Lustig, Roussanov, and Verdelhan (2011), formed by sorting the currencies of (at most) 35 developed and emerging economies on the basis of their forward discounts and rebalancing every month. Row 1 shows that the I-SDF produces an annualized intercept of −1.4%, which although marginally statistically significant with a t-statistic of 2.4, is economically small. The risk premium associated with the I-SDF, on the other hand, is strongly statistically significant with a t-statistic exceeding 6. Moreover, the risk premium has a similar magnitude to those obtained for the US equity market in Tables 1–4. The OLS $\overline{R}^2$ looks impressive at 88.6% and, more importantly, is quite close to the GLS $\overline{R}^2$ of 83.5%. Consistent with the above statistics, the T^2 statistic has a p-value of 60.4%, suggesting that the sum of squared pricing errors is not statistically different from zero.

Table 5: Performance of I-SDF on Other Asset Classes

Row	Assets	*const.*	λ_{sdf}	$\overline{R}^2_{OLS}$ (%)	$\overline{R}^2_{GLS}$ (%)	T^2	q
			Monthly				
(1)	Currencies	−.001 (−2.40)	−0.399 (−6.32)	88.6	83.5	2.73 (.604)	.013
(2)	Commodities	−.002 (−15.0)	−.114 (−15.8)	98.8	77.1	.175 (.916)	.001
(3)	Global Equities	.006 (19.33)	−.449 (−5.34)	90.2	63.5	2.17 (.337)	.008
(4)	Anomalies	−.002 (−2.0)	−.110 (−5.74)	88.9	92.2	1.20 (.753)	.003

Cross-sectional regressions of average excess returns listed in column 2 on the estimated factor loadings for the information SDF, at the monthly frequency. For each set of returns, the table presents the intercept and slopes, along with t-statistics in parentheses. It also presents the OLS adjusted R^2 and the GLS adjusted R^2. The last two columns present, respectively, Shanken's (1985) cross-sectional T^2 statistic along with its asymptotic p-value in parentheses, and the q statistic.

Row 2 presents the results for commodities. The test assets consist of a commodity index, constructed from 27 different commodity futures, and 3 additional portfolios formed by sorting the commodity futures on the basis of their book-to-market-equity or 'value' (details of the construction of these portfolios can be found in Asness, Moskowitz, and Pedersen (2013)). The results are quite similar to those obtained with currency portfolios in Row 1 – the estimated intercept is economically small with an annualized value of -1.8%; the risk premium associated with the I-SDF is strongly statistically significant with a t-statistic exceeding 15; the OLS and GLS $\overline{R}^2$ are both large at 98.8% and 77.1%, respectively; and the T^2 test fails to reject the model.

Row 3 presents the results for global equities. The test assets consist of a global equity index, constructed from individual stocks globally across 4 equity markets – the US, the UK, Europe, and Japan – and 3 additional portfolios formed by sorting the global individual stocks on the basis of their book-to-market-equity (details of the construction of these portfolios can be found in Asness, Moskowitz, and Pedersen (2013)). Row 3 shows that, in this case, the estimated intercept is statistically and economically large with an annualized value of 6.9%. However, the risk premium associated with the I-SDF is strongly statistically significant with a magnitude similar to those obtained in the US equity market in Tables 1–4. Moreover, the OLS and GLS $\overline{R}^2$ are both large at 90.2% and 63.5%, respectively, and the T^2 test fails to reject the model.

Row 4 presents results for certain additional classes of US equities, namely the highest and lowest decile portfolios formed by sorting stocks on the basis of accruals, market beta, net equity issuance, variance, and residual variance. The estimated intercept is economically small at 1.9% per annum; the risk premium associated with the I-SDF is strongly statistically significant with a t-statistic exceeding 5; the OLS and GLS $\overline{R}^2$ are both large at 88.9% and 92.2%, respectively; and the T^2 test fails to reject the model.

The above results suggest that the I-SDF accurately identifies the underlying sources of priced risk, for broad cross sections of assets and data frequencies. The latter finding is an important robustness check, since the method proposed in this paper relies on a large time series dimension (T) relative to the cross-sectional one (N). Hence, the stability of the results when the I-SDF is estimated at the quarterly frequency or from other asset classes with short histories is reassuring about the performance of the approach with shorter time series of returns data.

Finally, we end this section by evaluating whether the non-linear dependency of the I-SDF on the underlying assets carries valuable pricing information. We do so by comparing the I-SDF's pricing ability to the pricing performance of a linear projection of the I-SDF on the space of asset returns used to construct it. We perform two exercises to establish this point.

First, we compare the ability of the I-SDF and its linear projection to explain the deviations in the returns of different cross sections of assets from those implied by the CAPM, i.e. the portion of the asset returns left unexplained by the CAPM. Specifically, we recover the I-SDF using as assets returns the fitted residuals, $\alpha_p + \epsilon_{p,t}$, from time series regressions of the excess returns of a chosen cross section of assets on the market excess return:

$$R^e_{p,t+1} = \alpha_p + \beta_{p,m} R^e_{m,t+1} + \epsilon_{p,t+1}.$$

Under the assumption that the underlying pricing kernel can be decomposed multiplicatively into a component implied by the CAPM, M^{CAPM}_{t+1}, and a second residual component that is orthogonal to the market return, M^{resid}_{t+1}:

$$M_{t+1} = M^{CAPM}_{t+1} M^{resid}_{t+1},$$

the above approach of using the time series of fitted residuals from the CAPM regression to extract the I-SDF recovers the residual component of the SDF, M^{resid}_{t+1}, in the above equation.[16] Given the information-theoretic methodology used, the recovered M^{resid}_{t+1} represents the minimum amount of additional information that needs to be added to the CAPM-implied

[16]Note that, under the multiplicative decomposition of the SDF, we have

$$\mathbb{E}\left[M^{CAPM}_{t+1} M^{resid}_{t+1} R^e_{p,t+1}\right] = \mathbb{E}\left[\mathbb{E}(M^{CAPM}_{t+1} M^{resid}_{t+1} R^e_{p,t+1} | M^{resid}_{t+1}, \epsilon_{p,t+1})\right] = 0.$$

kernel, M_{t+1}^{CAPM}, in order to make it correctly price the cross section of asset returns out of sample.

The empirical performance of the I-SDF and its linear projection when the asset returns are the fitted CAPM residuals for each of the cross sections in Tables 1–4 is presented in Table 6, Panel A. Rows 1 and 2 show that the I-SDF performs substantially better than its projection at explaining the deviations of the average returns of the 25 FF portfolios from that implied by the CAPM – the OLS $\overline{R}^2$ is 81.2% for the I-SDF compared to only 19.1% for its linear projection; the GLS $\overline{R}^2$ is 62.6% for the I-SDF compared to being an order of magnitude smaller at 2.2% for its linear projection; the T^2 statistic fails to reject the hypothesis that the pricing errors are all zero for the I-SDF specification at the 5% level of significance, whereas the hypothesis is strongly rejected for the linear projection of the I-SDF; and the q statistic reveals that the unexplained Sharpe ratio of the tangency portfolio of the test assets using the linear projection of the I-SDF is more than one and a half times that of the I-SDF itself.

Similar results are obtained when the set of test assets consists of the deviations of the average returns of the 10 momentum portfolios from that implied by the CAPM (rows 3 and 4), or the deviations of the average returns of the 25 long-term reversal and size sorted portfolios from that implied by the CAPM (rows 5 and 6). The performance of the I-SDF and its linear projection are quite comparable when the set of test assets consists of the deviations of the average returns of the 10 industry portfolios and the smallest and largest deciles of size, book-to-market-equity, and momentum portfolios from that implied by the

Since $M_{t+1}^{CAPM} = 1 - \frac{\mu_m}{\sigma_m^2}(R_{m,t+1} - \mu_m)$ and M_{t+1}^{CAPM} is orthogonal to M_{t+1}^{resid}, we have

$$\mathbb{E}\left[\mathbb{E}(M_{t+1}^{CAPM} M_{t+1}^{resid} R_{p,t+1}^e | M_{t+1}^{resid}, \epsilon_{p,t+1})\right] = \mathbb{E}\left[M_{t+1}^{resid}\mathbb{E}(M_{t+1}^{CAPM} R_{p,t+1}^e | M_{t+1}^{resid}, \epsilon_{p,t+1})\right] = 0.$$

Now,

$$\mathbb{E}\left[M_{t+1}^{CAPM} R_{p,t+1}^e | M_{t+1}^{resid}, \epsilon_{p,t+1})\right] = \mathbb{E}\left[(1 - \frac{\mu_m}{\sigma_m^2}(R_{m,t+1} - \mu_m))R_{p,t+1}^e | M_{t+1}^{resid}, \epsilon_{p,t+1}\right] = \alpha_p + \epsilon_{p,t+1}$$

Therefore, we have $\mathbb{E}\left[M_{t+1}^{resid}\mathbb{E}(M_{t+1}^{CAPM} R_{p,t+1}^e | M_{t+1}^{resid}, \epsilon_{p,t+1})\right] = \mathbb{E}\left[M_{t+1}^{resid}(\alpha_p + \epsilon_{p,t+1})\right] = 0$. Thus, using the relative entropy minimization approach to extract the I-SDF using as asset returns the CAPM residuals for a chosen cross section of assets recovers the residual component, M_{t+1}^{resid}, of the pricing kernel for that cross section.

CAPM (rows 7 and 8).

As second exercise to evaluate the importance of the non-linearity of the I-SDF, we construct the pricing kernel, and it's linear projection, using a different set of asset from the ones employed in the cross-sectional asset pricing tests.

Table 6: Pricing of CAPM Residuals: I-SDF vs. its linear projection

Row	Assets	*const.*	λ_{sdf}	λ_{proj}	$\overline{R}^2_{OLS}$ (%)	$\overline{R}^2_{GLS}$ (%)	T^2	q
		Panel A: I-SDF Extracted From Different Cross-Sections, Monthly						
(1)	25 FF	.000 (1.58)	-0.441 (-10.24)		81.2	62.6	33.94 (.066)	.066
		.001 (1.69)		0.881 (2.58)	19.1	2.2	54.68 (.000)	.173
(2)	10 Momentum	.000 (1.07)	-0.418 (-14.6)		95.9	75.1	11.38 (.181)	.020
		.001 (1.56)		0.462 (8.76)	89.4	39.6	4.73 (.786)	.048
(3)	25 LTR & Size	.002 (6.12)	$-.292$ (-3.00)		25.0	61.2	23.80 (.415)	.039
		.001 (4.79)		.011 (2.30)	14.8	37.5	40.10 (.015)	.066
(4)	S, B, G, V, W, L, 10 Ind.	.001 (2.00)	$-.663$ (-10.33)		87.6	72.7	18.29 (0.194)	.033
		.000 (.813)		.004 (9.75)	84.0	74.0	17.13 (0.249)	.029
		Panel B: I-SDF Extracted From FF25, Monthly						
(1)	10 Momentum	.002 (6.57)	-0.973 (-15.7)		96.4	75.1	5.56 (.697)	.019
		.001 (1.78)		0.936 (6.75)	83.2	35.8	15.26 (.054)	.051
(2)	25 LTR & Size	.001 (4.72)	$-.273$ (-4.20)		41.0	17.6	47.12 (.002)	.080
		.002 (5.13)		.318 (1.32)	3.0	-7.2	55.09 (.000)	.106
(3)	S, B, G, V, W, L, 10 Ind.	.001 (1.13)	$-.674$ (-4.51)		56.4	29.6	31.70 (0.004)	.079
		.001 (1.96)		1.068 (6.09)	70.6	46.4	15.99 (0.314)	.063

Cross-sectional regressions of average CAPM residuals for the excess returns listed in column 2 on the estimated factor loadings for the information SDF and its mimicking portfolio, at the monthly frequency. Panel A presents results when the I-SDF is extracted using the fitted CAPM residuals for the same cross section of assets that it is then asked to price. Panel B, on the other hand, presents results when the I-SDF is extracted using the fitted CAPM residuals for the 25 FF portfolios and it is then asked to price the fitted CAPM residuals for different cross-sections of assets listed in column 2. For each set of returns, the table presents the intercept and slopes, along with t-statistics in parentheses. It also presents the OLS adjusted R^2 and the GLS adjusted R^2. The last two columns present, respectively, Shanken's (1985) cross-sectional T^2 statistic along with its asymptotic p-value in parentheses, and the q statistic.

In Panel B of Table 6 we present the performance of the I-SDF, and of its linear projection onto the set of assets used to construct it, in a scenario when the set of assets used to estimate the kernel differs from the set of test assets. In particular, the cross-section used to estimate the I-SDF (and its linear projection) consists of the fitted CAPM residuals for the 25 size and book-to-market-equity sorted portfolios of Fama and French, i.e. the portion of the returns on these portfolios that remain unexplained by the CAPM. The sets of test assets that the I-SDF and its linear projection are asked to price consist of the fitted CAPM residuals for the cross sections in Tables 2–4.

Row 1 shows that the I-SDF extracted using the CAPM residuals for the 25 FF portfolios performs very well in pricing the CAPM residuals for the 10 momentum portfolios – the estimated intercept is economically small at 2.4% per annum and can, in principle, be almost fully explained by the difference between borrowing and lending rates; the estimated price of risk for the I-SDF factor is strongly statistically significant with a t-statistic larger than 15 in absolute value; the OLS and GLS $\overline{R}^2$ are both large at 96.4% and 75.1%, respectively; the T^2 test fails to reject the model; and the point estimate of the q statistic implies an unexplained Sharpe ratio of only 0.13. Row 2 shows that the linear projection of the I-SDF performs substantially worse than the I-SDF itself at pricing the CAPM residuals for the 10 momentum portfolios – the GLS $\overline{R}^2$ is less than half of that obtained with the I-SDF (35.8% versus 75.1%) and, more importantly, the GLS $\overline{R}^2$ is almost two and a half times smaller than the OLS $\overline{R}^2$ (35.8% versus 83.2%); the T^2 test rejects the hypothesis that the pricing errors are all zero at the 10% level of significance; and the point estimate of the q statistic implies an unexplained Sharpe ratio of .23 compared to only .13 for the I-SDF.

When the set of test assets consists of the CAPM residuals for the 25 long-term reversal and size-sorted portfolios, the superior performance of the I-SDF (row 3) relative to its linear projection (row 4) is even more evident. The estimated price of risk for the I-SDF is strongly statistically significant with an absolute value of the t-statistic in excess of 4, whereas the price of risk for the linear projection is not statistically different from zero; the OLS $\overline{R}^2$ is 41.0% for the I-SDF compared to only 3% for its linear projection; and the GLS $\overline{R}^2$ is 17.6% for the I-SDF compared to being negative at −7.2% for its linear projection. The performance of the I-SDF and its linear projection are comparable when the set of test assets consists of the CAPM residuals for the 10 industry portfolios along with the smallest and largest deciles of size, book-to-market-equity, and momentum sorted portfolios (rows 5 and 6).

Overall, the results in Table 6 suggest that the non-linearity of the I-SDF contains valuable information about the underlying sources of priced risk – information that is partially lost by the linear projection of the I-SDF on to the set of asset returns used in its construction.

V Properties of the Information SDF

Having shown that the I-SDF is successful at pricing broad cross-sections of financial assets out-of-sample, we next turn to an investigation of the properties of the I-SDF.

V.1 Commonality of the Information SDFs

First, we analyse the common characteristics shared by the I-SDFs extracted using different cross-sections of asset returns. In particular, we focus on the cross-sections of equity portfolios used in Tables 1–4 that are meant to capture several well known asset pricing anomalies: 25 Size and Book-to-Market portfolios, 10 Momentum portfolios, 25 Long-term reversal and Size portfolios, as well as 10 Industry portfolios combined with the Small, Large, Growth, Value, Winners and Losers portfolios.

Figure 1 plots the time series of the quarterly I-SDFs (a similar pattern holds for the monthly ones) extracted from the four different cross-sections of asset returns, as well as the first principal component extracted from these four I-SDFs. To isolate visually the low frequency behavior of the I-SDFs, the figure reports three year moving averages of the I-SDFs. All the I-SDFs show a clear business cycle pattern: they are low during expansions and higher during NBER recession periods (depicted by the grey vertical bands). The first principal component extracted from the cross-section of the I-SDFs, that explains between

45%-62% of the time series variation of the I-SDFs (as reported in Table 7), also has a similar business cycle pattern.

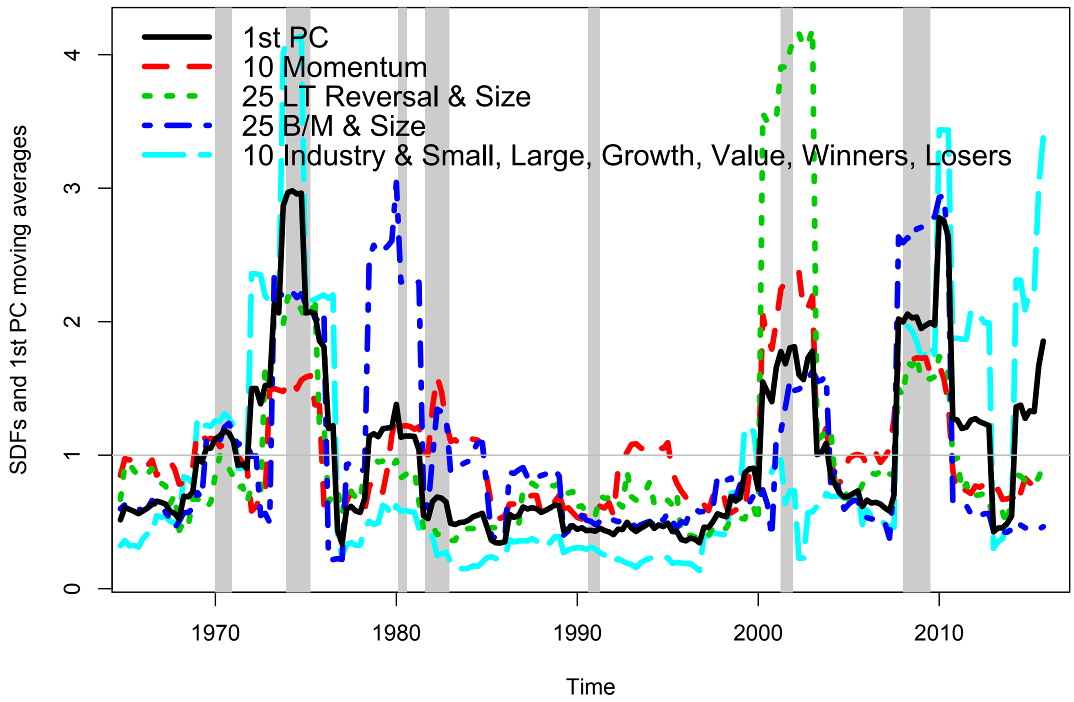


Figure 1: Three year moving average of the quarterly I-SDFs extracted from different cross-sections of assets and their first principal component. NBER recession periods are marked by the grey vertical bands.

However, Figure 1 also reveals another interesting commonalty between the I-SDFs over and above the co-movement related to the business cycle. The I-SDFs are especially high when the recession episodes are concomitant with periods of big stock market downturns, like the 1973-1975, 2001, and 2007-2009 recessions, unlike the 1981-1982 and 1990-1991 recessionary periods that were not accompanied by big stock market crash episodes. This feature of the I-SDFs is also well captured by their first principal component. Consider, for instance, the sharp contrast between the behavior of the I-SDFs during the 1990-1991 and 2007-2009 recessionary periods. Figure 1 shows that the I-SDFs were the highest during the

Table 7: Principal Components

	PC1	PC2	PC3	PC4
Panel A: Quarterly data				
Standard deviation	3.63	3.00	2.42	1.04
Proportion of Variance	0.45	0.31	0.20	0.04
Cumulative Proportion	0.45	0.76	0.96	1.00
Panel B: Monthly data				
Standard deviation	1.73	0.99	0.72	0.56
Proportion of Variance	0.62	0.21	0.11	0.06
Cumulative Proportion	0.62	0.83	0.94	1.00

Characteristics of principal components of I-SDFs extracted from four cross-sections of equity portfolios – 25 Size and Book-to-Market portfolios, 10 Momentum portfolios, 25 Long-term reversal and Size portfolios, as well as 10 Industry portfolios combined with the Small, Large, Growth, Value, Winners and Losers portfolios – starting at 1963:07. Panel A presents the quarterly results and Panel B the monthly results.

latter period while they were close to their lowest values during the former recession period. Both of these periods were characterized by very low average quarterly real consumption growth of -0.27% and -0.40%, respectively. However, the behavior of the stock market was markedly different between these periods: the average quarterly real stock market return was 1.1% over the 1990-1991 period while it was substantially negative at -7.0% over the recent financial crisis of 2007-2009. As yet another example, compare the 1973-1975 and 1990-1991 recessions. Average quarterly real consumption growth was 0.0% during the former period while it was substantially lower at -0.3% over the latter period. However, the stock market performed much worse in the earlier recession, with an average return of -6.3% over the 1973-1975 period. Figure 1 shows that the I-SDFs were much higher in the former period than in the latter, despite consumption growth being much higher in the former period.

Our results suggest that while business cycle risk is an important source of systematic risk, it cannot fully explain the behavior of the I-SDFs. We find evidence that the true underlying SDF may not solely be a function of the business cycle, but also of additional variables related to the performance of the stock market (that is only imperfectly correlated with the business cycle).

Table 8 reports the sample correlations between the I-SDFs plotted in Figure 1, their first principal component, and a dummy variable that takes the value one during the NBER

Table 8: Correlation matrix

	10 Mom.	25 LT Rev. & Size	10 Ind. & 6 Anomalies	Recession	PC1
		Panel A: Quarterly data			
25 B/M & Size	0.27	0.30	0.42	0.26	0.87
10 Momentum		0.83	0.57	0.45	0.65
25 LTR & Size			0.40	0.26	0.65
10 Ind. & 6 Anomalies				0.42	0.72
Recession					0.40
		Panel B: Monthly data			
25 B/M & Size	0.14	0.19	0.21	0.24	0.26
10 Momentum		0.65	0.40	0.33	0.64
25 LTR & Size			0.32	0.17	0.55
10 Ind. & 6 Anomalies				0.05	0.95
Recession					0.15

Cross-correlations between the I-SDFs extracted from different sets of equity portfolios, correlations between the I-SDFs and the first principal component and a recession dummy. Panels A and B report results at the quarterly and monthly frequencies, respectively.

recession periods and zero otherwise. The average correlation between quarterly (monthly) I-SDFs is about 47% (32%), all I-SDFs are highly correlated with their first principal component (the average quarterly and monthly correlations are, respectively, 72% and 60%) and show substantial correlations with the recession dummy (the average quarterly and monthly correlations are, respectively, 35% and 20%), and the latter is positively correlated with the first principal component.

V.2 I-SDF and Tail Risk

The I-SDFs have a strong non-Gaussian distribution, regardless of the set of equity portfolios used for the estimation. Table 9 presents the summary statistics from the distributions of the I-SDFs.[17] Consider first Panel A that reports results at the monthly frequency. The I-SDFs are strongly positively skewed with the coefficient of skewness varying from 3.47, when the FF25 portfolios are used as test assets, to 18.95 when the set of assets consists of the 10 industry portfolios and the smallest and largest deciles of size, book-to-market equity, and momentum sorted portfolios. The coefficient of excess kurtosis varies from 20.45 to 420.5

[17] Recall that the mean of the I-SDFs is normalized to unity.

from the former to the latter set of test assets. Similar results are obtained at the quarterly frequency in Panel B. These results suggest that tail risk is an important source of priced risk.

With the estimated I-SDFs at hand, it is also possible to quantify the share of the risk premium of an asset attributable to tail risk compensation. Recall that, given an SDF M_t, lack of arbitrage opportunities implies that the risk premium on an asset is i

$$\mathbb{E}\left[R^e_{i,t}\right] = -\frac{\int \left(M_t - \overline{M}\right) R^e_{i,t}\mathbf{1}_{\left\{\left(M_t,\, R^e_{i,t}\right)\in\mathbb{A}\right\}} d\mathbb{P} + \int \left(M_t - \overline{M}\right) R^e_{i,t}\mathbf{1}_{\left\{\left(M_t,\, R^e_{i,t}\right)\notin\mathbb{A}\right\}} d\mathbb{P}}{\mathbb{E}\left[1/R^f_{t-1}\right]}$$

for any arbitrary set $\mathbb{A}$, where $\mathbf{1}_{\{\}}$ denotes the indicator function that takes value 1 when the condition in brackets is satisfied, and the numerator of the above expression is simply the covariance between the SDF and the excess return. As a consequence, using our estimates of the I-SDFs, choosing a set in the tail of the empirical distribution of $\widehat{M_t}$ and $R^e_{i,t}$, and replacing the integral with a sample analogue, we can compute the share of the asset's observed risk premium generated by tail risk.

Table 9: Summary Statistics of I-SDFs

Row	Assets	Median	Volatility	Skewness	Kurtosis	Max	Min
			Panel A: Monthly				
(1)	FF25	0.85	0.72	3.47	20.45	7.80	0.049
(2)	10 Momentum	0.87	0.92	11.10	159.8	15.71	0.097
(3)	25 LTR and Size	0.88	0.91	14.84	300.5	20.11	0.084
(4)	S, L , G, V, W , L, 10 Ind	0.81	1.61	18.95	420.5	37.83	0.083
			Panel B: Quarterly				
(1)	FF25	0.16	3.37	5.49	30.75	23.95	0.000
(2)	10 Momentum	0.64	1.43	5.66	42.35	14.39	0.025
(3)	25 LTR and Size	0.43	2.83	10.31	125.0	37.32	0.004
(4)	S, L , G, V, W , L, 10 Ind	0.35	2.78	6.20	42.26	23.45	0.001

Summary statistics from the distribution of the I-SDFs constructed from various cross-sections of assets (listed in Column 2). Panels A and B present the results at monthly and quarterly frequencies, respectively.

Table 10 reports the share of the risk premium of the market return (Panel A), and of

Table 10: Quantifying Tail Risk

Row	Assets	$M_{0.95}$	$M_{0.90}$	$R^e_{0.95}$	$R^e_{0.90}$	$M_{0.95}$ & $R^e_{0.95}$	$M_{0.90}$ & $R^e_{0.90}$
		Panel A: Tail Risk of the Market Portfolio					
(1)	FF 25	.412	.436	.389	.440	.402	.525
		[.051]	[.100]	[.049]	[.099]	[.012]	[.022]
(2)	10 Mom	.224	.341	.360	.456	.281	.421
		[.051]	[.100]	[.049]	[.099]	[.017]	[.039]
(3)	25 LTR & Size	.477	.420	.426	.503	.450	.571
		[.059]	[.100]	[.049]	[.099]	[.012]	[.026]
(4)	S,L,G,V,W,L & 10 Ind	.348	.447	.515	.585	.455	.610
		[.051]	[.100]	[.049]	[.099]	[.012]	[.032]
		Panel B: Tail Risk of the Cross-Section					
(1)	FF 25	.406	.406	.424	.510	.398	.551
		[.067]	[.033]	[.097]	[.126]	[.117]	[.219]
(2)	10 Mom	.373	.462	.307	.389	.211	.327
		[.508]	[.426]	[.293]	[.418]	[.274]	[.502]
(3)	25 LTR & Size	.414	.462	.454	.557	.422	.540
		[.051]	[.047]	[.049]	[.054]	[.069]	[.075]
(4)	S,L,G,V,W,L & 10 Ind	.404	.484	.438	.507	.368	.466
		[.276]	[.220]	[.214]	[.276]	[.291]	[.473]

Contribution of tail risk to the overall risk premium of the aggregate stock market portfolio (Panel A) and to the risk premium of the cross-section listed in column 2 (Panel B). The contribution of tail risk is computed as the ratio of (a) the covariance between the I-SDF and the excess returns on the market (Panel A) or each asset in the cross section (Panel B), computed using only those observations that lie in the tail of the I-SDF (Columns 3 and 4), the excess returns (Columns 5 and 6) or both (Columns 7 and 8), and (b) the overall covariance between the I-SDF and the excess returns on the market (Panel A) or each asset in the cross section (Panel B), computed using all the available observations. Each row presents the results when the I-SDF is extracted using the cross section listed in column 2 of that row. In Panel A, the two numbers in each cell denote the ratio for the overall market excess return, along with the fraction of the total number of observations belonging to the corresponding tail in square brackets below. In Panel B, the two numbers in each cell denote the mean of the ratio across all the assets in the cross section, along with the standard deviation of these ratios in square brackets below.

the cross-section of asset returns (Panel B) attributable to tail risk in monthly data.[18] The first two columns show that the 5% (10%) most extreme realizations of the SDF drive about 36% (41%) of the total market risk premium and about 40% (45%) of the cross-sectional risk premia. Columns three and four show that the 5% (10%) most extreme realizations of market returns (top panel) and asset returns (bottom panel) generate about 41-42% (49-50%) of the observed risk premia. Finally, the last two columns report the share of observed risk premia generate by the joint tail of asset returns and SDF. The joint tail of the SDF and the market portfolio (Panel A) accounts on average for about 40-53% of the market risk premium – i.e. around 2% of the possible states generate roughly half of the market equity

[18]We focus on monthly data only due to the larger number of observations available in this case, but quarterly results are very similar to the ones reported.

premium – and the joint tail of the SDF and asset returns (Panel B) accounts on average for about 35-47% of the asset risk premia.

Overall, Table 10 indicates that (independently from the cross-section used to estimate the I-SDF) a very large part of observed risk premia are due to a compensation for tail risk. A large literature in asset pricing highlights the crucial role of tail events in explaining several observed features of stock market data. Examples include the rare disasters paradigm in a representative agent setting that models negative skewness in aggregate consumption growth (see, e.g., Barro (2006)) as well as models with incomplete markets that feature negative skewness in the cross-sectional distribution of idiosyncratic labor income growth (see, e.g., Mankiw (1982), Constantinides and Ghosh (2017)). By their very definition, such disaster events are rare making it difficult to rigorously calibrate such models. Our methodology offers a data-driven approach to calibrating models featuring tail risks as an important source of systematic risk.

V.3 I-SDF: an Information Anomaly

We next proceed to show that the I-SDF contains novel pricing information not captured by standard multifactor asset pricing models, such as the FF 3- and 5-factor and the Carhart 4-factor models. Table 11 presents time series regressions of the I-SDF, constructed from each set of test assets in Tables 1–4 (and indicated in the second column), on the FF3 factors. Whenever the assets used to construct the I-SDF include momentum-sorted portfolios, we also include the momentum factor as a regressor in addition to the FF3 factors. If the factors fully explain the variation in the I-SDF, the intercepts from the time series regressions should be indistinguishable from zero and the R^2 of the regressions should be high.

Note that the I-SDF, while being a function of asset returns, is not directly a traded asset or portfolio of assets. Therefore, in Table 11, we also present the regression results for the mimicking portfolio, maximally correlated with the I-SDF, which we refer to as the *Information Portfolio* (I-P).

The I-P is constructed from a time series regression of the I-SDF on to the set of asset returns used in its construction. In particular, the I-SDF $\widehat{M}_t$ is projected onto the space of excess returns to obtain the vector of portfolio weights $\omega_T \in \mathbb{R}^N$ (normalized to sum to unity).[19] That is, the mimicking portfolio weights ω_T are given by

$$\omega_T := -\frac{\widehat{\mathbf{b}}_T}{\left|\widehat{\mathbf{b}}_T' \iota\right|}, \quad \left[\widehat{a}_T, \widehat{\mathbf{b}}_T'\right] := \underset{\left\{a_T, \mathbf{b}_T'\right\}}{\arg\min} \frac{1}{T} \sum_{t=1}^{T} \left(\widehat{M}_t - a_T - \mathbf{b}_T \mathbf{R}_t^e\right)^2, \tag{6}$$

where ι denotes a conformable column vector of ones. Using the portfolio weights vector, the I-P is obtained as $R_t^{IP} = \omega_T' \mathbf{R}_t^e$ over the out-of-sample evaluation period.

Note that, in the case of time series regressions of the I-P on a set of candidate risk factors, the intercepts have the interpretation of a standard α.

Panel A presents the results at the monthly frequency. In rows 1–2, the 25 size and book-to-market-equity sorted portfolios are used to extract the kernel and its mimicking portfolio. Row 1 shows that the 3 FF factors explain only 17.6% of the variation in the I-SDF. Moreover, the estimated intercept is strongly statistically significant, with an annualized value of 14.0%. Note that since the I-SDF is not a tradeable factor, the intercept is not interpretable as a tradeable alpha. Row 2 shows that the FF factors can explain a larger fraction of the variation in the I-P than in the I-SDF (36.3% versus 17.6%). However, even in this case, more than sixty percent of the variation is left unexplained by the FF factors. Moreover, the estimated intercept, which in this case has the interpretation of a standard α, is statistically and economically large, at 18.6% per annum. These results, together with the observation that the I-SDF performs substantially better at pricing the cross-section of the 25 size and B/M sorted portfolios (Table 1), suggest that the FF factors do not fully capture the sources of priced risk even for the size and book-to-market portfolios.

Similar results are obtained for the 10 momentum sorted portfolios (rows 3–4), the 25

[19] Here, T denotes the length of the entire out-of-sample evaluation period. Note that there is a look-ahead bias in the construction of the I-SDF mimicking portfolio, the I-P. A tradeable version of the I-P can be easily constructed and has performance very similar to that reported here. These results are available from the authors upon request.

Table 11: Explaining I-SDF and I-P with FF3 and Momentum Factors

Row	Assets	α_{sdf}	α_{IP} (%)	β_{Rm}	β_{SMB}	β_{HML}	β_{MOM}	$\overline{R}^2_{OLS}$ (%)
			Panel A: Monthly					
(1)	FF25	1.17 (40.4)		−4.52 (−6.51)	−1.86 (−1.89)	−11.60 (−10.95)		17.6
(2)	FF25		1.55 (8.12)	0.45 (9.96)	0.21 (3.22)	1.25 (18.03)		36.3
(3)	10 Momentum	1.17 (33.5)		−4.86 (−5.83)	1.73 (1.50)	−3.23 (−2.53)	−11.7 (−14.2)	24.4
(4)	10 Momentum		1.31 (4.09)	0.74 (9.69)	−0.42 (−3.96)	0.27 (2.35)	1.73 (22.8)	46.7
(5)	25 LTR & Size	1.16 (29.9)		−4.06 (−4.40)	−1.36 (−1.05)	−9.14 (−6.51)		7.20
(6)	25 LTR & Size		1.28 (6.11)	0.43 (8.62)	−0.13 (−1.79)	0.94 (12.37)		22.7
(7)	S, B, G, V, W, L, 10 Ind.	1.30 (19.4)		−8.49 (−5.31)	−6.22 (−2.80)	−12.1 (−4.91)	−14.9 (−9.38)	15.1
(8)	S, B, G, V, W, L, 10 Ind.		0.88 (5.02)	0.68 (16.26)	0.50 (8.61)	0.96 (15.0)	1.17 (28.23)	62.7
			Panel B: Quarterly					
(1)	FF25	4.49 (4.38)		3.59 (0.27)	2.93 (0.15)	−29.8 (−1.69)		0.31
(2)	FF25		8.20 (4.57)	−0.13 (−0.57)	−0.07 (−0.19)	1.07 (3.44)		5.32
(3)	10 Momentum	1.63 (13.85)		−8.04 (−5.39)	5.07 (2.35)	−3.01 (−1.53)	−12.16 (−8.21)	28.4
(4)	10 Momentum		3.00 (4.32)	0.77 (8.83)	−0.30 (−2.37)	0.21 (1.81)	1.21 (13.85)	53.0
(5)	25 LTR & Size	2.37 (5.45)		−5.87 (−1.02)	9.99 (1.19)	−17.1 (−2.28)		1.56
(6)	25 LTR & Size		5.06 (5.23)	0.27 (2.13)	−0.51 (−2.74)	0.81 (4.86)		11.2
(7)	S, B, G, V, W, L, 10 Ind.	3.35 (7.84)		−8.53 (−1.58)	−19.48 (−2.49)	−22.34 (−3.12)	−33.06 (−6.16)	15.6
(8)	S, B, G, V, W, L, 10 Ind.		3.50 (3.69)	0.43 (3.55)	0.75 (4.29)	0.96 (6.01)	1.52 (12.78)	45.0

The table presents the intercept and slope coefficients, along with the t-statistics in parentheses, as well as the OLS adjusted R^2, from time series regressions of the information SDF (odd rows) and I-P (even rows) on the Fama–French and Carhart factors. Each row presents the results when the information SDF and I-P are constructed using the cross-section of assets listed in column 2. Since the α_{IP} is presented in percentage terms, α_{sdf} is presented as the intercept multiplied by 100 for the sake of comparability. Note that, since the I-P weights in equation (6) are proportional to minus the projection coefficients, and are normalized to sum to 1, one would expect the betas of I-P and I-SDF to have opposite signs and their magnitudes are not directly comparable. Panels A and B present the results at monthly and quarterly frequencies.

portfolios formed on the basis of size and long term reversal (rows 5–6), and the smallest and largest deciles of portfolios formed on the basis of size, B/M, and momentum and the 10 industry-sorted portfolios (rows 7–8). The $\overline{R}^2_{OLS}$ from the I-SDF regressions vary from 7.2% (for the size and long-term reversal sorted portfolios) to 24.4% (for the 10 momentum-sorted portfolios), showing that a substantial proportion of the variability in the I-SDFs cannot be explained by the movements in the FF3 and momentum factors. The $\overline{R}^2_{OLS}$ from the

I-P regressions are higher, varying from 22.7%–62.7%, but still a substantial fraction of the variability is left unexplained by the standard multifactor models. The estimated annualized intercepts are all statistically significant and economically large, varying from 14.0%–15.6% for the I-SDF and from 10.6%–18.6% for the I-P.

As a robustness check (not reported), we also added as regressors the profitability and investment factors of Fama and French (2015), obtaining very similar results, in terms of intercepts and measures of fit, to the ones reported in the table.[20]

The results obtained at the quarterly frequency in Panel B are largely similar. In fact, the FF3 or Carhart 4 factors explain an even smaller fraction of the variability of the I-SDF at a quarterly frequency compared to that at a monthly frequency. For two out of the four sets of test assets, $\overline{R}^2_{OLS}$ is less than 2%, and the estimated intercepts are statistically significant in all four cases. For the I-P regressions, $\overline{R}^2_{OLS}$ is lower at the quarterly frequency in three out of the four sets of test assets. The estimated α's are all statistically significant and economically large, varying from 12.0% to 32.8% (annualized).

V.4 I-SDF and the Maximum Sharpe Ratio Portfolio

Finally, we further investigate the ability of the I-SDF to identify correctly the capital market line, i.e. the maximum Sharpe ratio portfolio of the test assets. Note that if an SDF prices a set of test assets perfectly in-sample, then its mimicking portfolio equals the mean-variance tangency portfolio of the assets – i.e. the maximum Sharpe ratio portfolio. However, the same is not necessarily true out-of-sample. Therefore, we assess the empirical performance of the I-P constructed as the mimicking portfolio of the *out-of-sample* I-SDF (and with portfolio weights computed according to equation (6)).

We evaluate the performance of the I-P using the same performance measures as in DeMiguel, Garlappi, and Uppal (2009), namely (i) the Sharpe ratio and (ii) the certainty-equivalent (CEQ) return for the expected utility of a mean–variance investor. The Sharpe

[20] Intercepts and $\overline{R}^2_{OLS}$ are virtually identical when the dependent variable is the I-SDF, while for the I-P the αs are somewhat reduced and the $\overline{R}^2_{OLS}$ is minimally increased.

ratio is defined as $\widehat{SR}_{I\text{-}P} = \widehat{\mu}_{I\text{-}P}/\widehat{\sigma}_{I\text{-}P}$, where $\widehat{\mu}_{I\text{-}P}$ and $\widehat{\sigma}_{I\text{-}P}$ are the sample mean and sample standard deviation, respectively, of the excess returns on the I-P. The CEQ return is defined as the risk free rate that would make an investor with mean–variance preferences and coefficient of risk aversion $\gamma = 1$ indifferent between the risky I-P and the risk free rate: $\widehat{CEQ}_{I\text{-}P} = \widehat{\mu}_{I\text{-}P} - \frac{\gamma}{2}\widehat{\sigma}^2_{I\text{-}P}$.

For each cross-section of assets used to construct the I-SDF, we compute the Sharpe Ratio, the CEQ return, and the first four moments of the corresponding I-P. The results are presented in Table 12. Panels A and B report results at the monthly and quarterly frequencies, respectively. As a benchmark to facilitate comparison, we also compute the corresponding statistics for the $1/N$ portfolio of the test assets. In addition to the equally-weighted portfolio, we also compare the performance of the I-P to other standard benchmarks, including: the market portfolio (row 2), the value and size risk factor mimicking portfolios (HML and SMB in rows 3 and 4, respectively) of Fama and French (1993) that are meant to exploit the value and size premia; the momentum portfolio (row 6) of Carhart (1997); and the combined value and momentum portfolio (row 9) that is meant to exploit the negative correlation between value and momentum strategies (see Asness, Moskowitz, and Pedersen (2013)).

Consider first row 1 of Panel A, where the I-P is constructed from the 25 size and book-to-market sorted portfolios. Its 2.3% monthly (27.6% annual) average return is more than three times that of the $1/N$ portfolio (presented in parentheses below), almost 5 times that of the market portfolio, almost 8 times that of the HML (row 3) portfolio, almost 12 times that of the SMB portfolio (row 4), more than three times that of the momentum portfolio (row 6), and more than four times that of the value and momentum strategies combined (row 9). These very high returns are obtained with a volatility that is only about one and a half times larger than that of the market and momentum portfolios.

Moreover, the I-P's monthly Sharpe ratio is 0.380 (≈ 1.3 annualized), while the Sharpe ratio of the corresponding $1/N$ benchmark (presented in parentheses below) is only 0.144 monthly (or 0.49 annualized), i.e. less than one-half that of the I-P. The I-P's Sharpe ratio

Table 12: Summary Statistics of Information Portfolio & Returns

Row	Assets	Mean	Volatility	Sharpe Ratio	Skewness	Kurtosis	CEQ
	Panel A: Monthly						
(1)	R^{IP} (FF25)	0.023 (0.007)	0.060 (0.051)	0.380 (0.144)	0.182 (−0.549)	4.37 (5.50)	0.021 (0.006)
(2)	Market - Risk Free	0.005	0.044	0.118	−0.531	4.968	0.004
(3)	HML	0.003	0.028	0.123	0.073	5.066	0.003
(4)	SMB	0.002	0.031	0.072	0.494	8.456	0.002
(5)	R^{IP} (10 Momentum)	0.028 (0.005)	0.107 (0.047)	0.264 (0.1095)	−1.64 (−0.321)	13.35 (4.782)	0.023 (0.003)
(6)	Momentum Portfolio	0.007	0.042	0.156	−1.333	13.56	0.006
(7)	R^{IP} (25 Long-Term Reversal & Size)	0.018 (0.008)	0.059 (0.051)	0.304 (0.154)	−1.00 (−0.412)	11.662 (5.764)	0.016 (0.007)
(8)	R^{IP} (S, B, G, V, W, L, 10 Industry)	0.024 (0.006)	0.070 (0.045)	0.348 (0.128)	−0.724 (−0.478)	6.956 (4.945)	0.022 (0.005)
(9)	HML & Momentum	0.005	0.023	0.217	−0.958	10.78	0.005
	Panel B: Quarterly						
(1)	R^{IP} (FF25)	0.090 (0.023)	0.258 (0.101)	0.349 (0.233)	−0.929 (−0.269)	5.19 (3.70)	0.057 (0.018)
(2)	Market - Risk Free	0.016	0.085	0.193	−0.496	3.735	0.012
(3)	HML	0.010	0.058	0.179	0.352	4.725	0.008
(4)	SMB	0.008	0.056	0.140	0.303	2.851	0.006
(5)	R^{IP} (10 Momentum)	0.065 (0.016)	0.133 (0.090)	0.489 (0.181)	−0.420 (−0.310)	3.599 (3.887)	0.056 (0.012)
(6)	Momentum Factor	0.019	0.077	0.245	−1.344	10.40	0.016
(7)	R^{IP} (25 Long-Term Reversal & Size)	0.059 (0.025)	0.144 (0.101)	0.413 (0.247)	−1.94 (−0.14)	18.38 (3.981)	0.049 (0.020)
(8)	R^{IP} (S, B, G, V, W, L, 10 Industry)	0.086 (0.018)	0.169 (0.088)	0.513 (0.210)	−0.223 (−0.392)	7.209 (3.906)	0.072 (0.014)
(9)	HML & Momentum	0.015	0.044	0.334	−0.682	7.695	0.014

Mean, volatility, Sharpe ratio, skewness, kurtosis, and CEQ statistic for the portfolios listed in column 2: the information portfolio, R^{IP}, constructed from various cross-sections of assets (listed in parentheses) with the corresponding statistics for an equally-weighted portfolio of the underlying assets presented in parentheses below; the market minus the risk free rate portfolio; the value portfolio (HML); the size portfolio (SMB); the momentum portfolio; and the value and momentum portfolio. Panels A and B present the results at monthly and quarterly frequencies.

not only outperforms the $1/N$ benchmark, but also the market portfolio, by a factor of more than three, the HML portfolio by a factor of more than three, the SMB portfolio by a factor of more than five, the momentum factor by a factor of more than two, and the combined Value and Momentum portfolio (that has an annualized SR of about 0.75) by a factor of 1.75.

Note that the very high returns and Sharpe ratio of the I-P in row 1 do not seem to be a compensation for negative skewness and tail risk: the I-P's skewness is positive (about 0.182), while that of the market and the combined HML and momentum portfolios are negative (and very large for momentum based strategies), and its kurtosis is similar to that of the market and HML portfolios, and much smaller than those of the momentum, HML plus momentum, and SMB strategies.

Similar conclusions are obtained using the CEQ return as the measure of performance. A mean–variance investor with $\gamma = 1$ would need an annualized risk free rate of 25.2% (or about 2.1% monthly) in order to not invest in the I-P, whereas a risk free rate of only 7.2% (or about 0.6% monthly) is required for such an agent to not invest in the $1/N$ portfolio. Similarly, annual (monthly) risk free rates of only 4.8% (0.4%), 3.6% (0.3%), 7.2% (0.6%), and 6.0% (0.5%), respectively, are required in order to be indifferent between the risk free rate and the market, the HML, the momentum, and the HML plus momentum portfolios.

To show that the performance of the I-P in row 1 is not driven by just a subset of the data, Panel A of Figure 2 plots the path of $1 invested in the I-P over the entire out-of-sample evaluation period. Note that because we use excess returns in the construction of the I-P, this corresponds to a long–short strategy that is short $1 in the risk free rate and uses the proceeds to invest in the optimal portfolio of risky assets. For comparison, and since the plotted I-P is constructed using the FF25 portfolios (hence it might exploit the size and value premia), we also plot the path of $1 invested in the HML and SMB portfolios, as well as the excess return on the market and the equally weighted portfolios. Note also that the graph is in log scale, so that the slopes of the various lines are directly comparable across the various strategies at each point in time. As is evident from the figure, the I-P outperforms, by a wide margin, each of the benchmarks. Moreover, the I-P outperformance is robust across sub-periods: the average slope of the I-P line is higher in virtually all the 10-year sub-periods. For robustness, Panel B of Figure 1 presents the same cumulated returns as Panel A but with the benchmark portfolios leveraged in order to have the same volatility as the Information Portfolio. The figure shows that the I-P clearly outperforms the various benchmarks. Moreover, the I-P tends to have less severe contractions in returns than the other portfolios during, and following, market-wide crashes (vertical dot-dashed lines in the

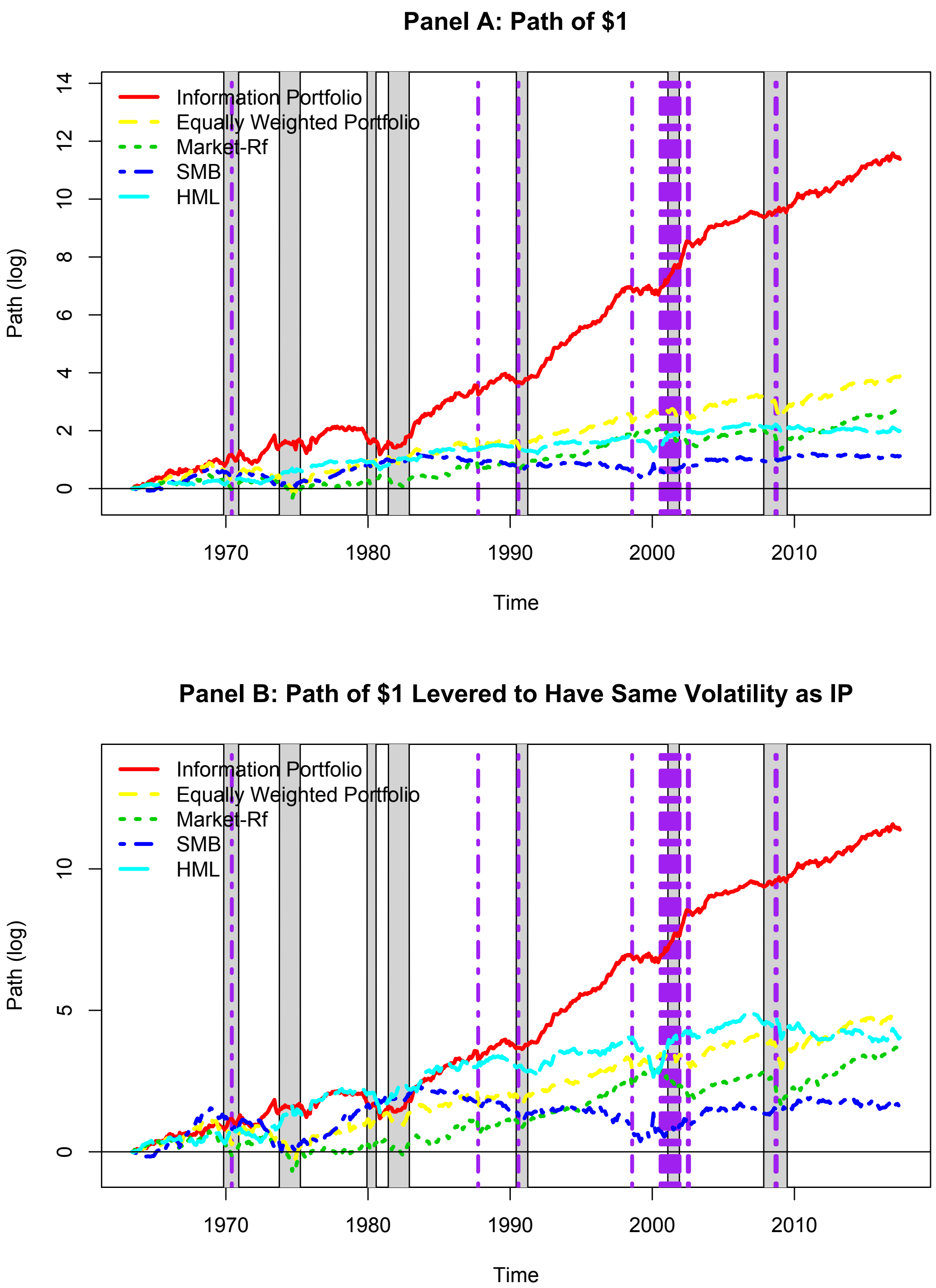


Figure 2: Panel A: cumulated log returns of a zero wealth $1 invested in: information portfolio (red solid line); market portfolio in excess of the risk free rate (green dotted line); SMB portfolio (dark blue dash-dot line); HML portfolio (pale blue long-dash line); $1/N$ portfolio (yellow dashed line). Panel B: same series as Panel A but with portfolios leveraged to the same volatility as the information portfolio. The information portfolio is non-parametrically extracted at a monthly frequency from the 25 Fama–French portfolios using a relative entropy minimization procedure in a rolling out-of-sample fashion starting at 1963:07. Shaded areas indicate NBER recession dates while the vertical dot-dashed lines indicate market crashes identified using the Mishkin and White (2002) approach.

figure).[21]

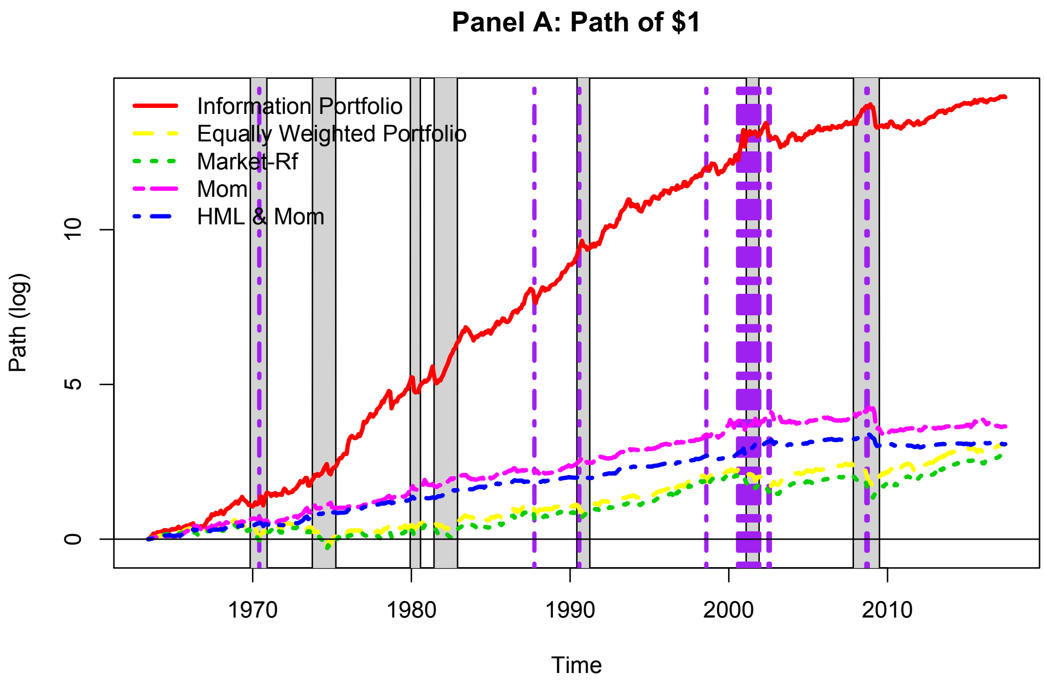
These results suggest that the relative entropy minimization approach to the recovery of the underlying SDF advocated in this paper leads to a very good performance, both in terms of out-of-sample pricing as well as constructing an optimally diversified portfolio. Moreover, the I-P outperforms the various benchmark portfolios not only when it is constructed using the FF25 portfolios, but also when different cross-sections are used. In particular, row 5 of Table 12 presents the results when the I-P is constructed from the 10 momentum sorted portfolios. In this case, the returns on the I-P are even higher: about 2.8% per month. Moreover, once again, the I-P has a Sharpe ratio more than double that of the $1/N$ portfolio and a CEQ return more than 7 times higher, and similarly outperforms the market, HML, momentum, and HML plus momentum strategies. Similar results are obtained when the I-P is constructed from the 25 long-term reversal and size sorted portfolios (row 7).

The I-P portfolio shows a similarly strong performance (in terms of SR and CEQ) when it is constructed using the the Small, Big, Growth, Value, Winners, and Losers portfolios as well as the 10 industry sorted portfolios (row 8). To show once again that this result is not driven by a particular sub-sample, and in order to offer a time series comparison of this portfolio with the momentum, and the joint value and momentum strategies, Figure 3 plots the path of $1 invested in the I-P (from row 8 of Table 12) over the entire out-of-sample evaluation period. Comparing both unleveraged as well as leveraged strategies (in Panels A and B, respectively), it is clear that the I-P outperforms the momentum, and value plus momentum, strategies in each 10-year sub-period.

Furthermore, Panel B of Table 12 shows that results similar to those discussed above are obtained when the information portfolio is estimated using quarterly data.

Overall, our results show that the I-P typically outperforms the naïve $1/N$ portfolio as well as other standard benchmarks, in terms of the Sharpe ratio and CEQ return. Moreover, these

[21]We follow Mishkin and White (2002) and identify a stock market crash as a period in which either the Dow Jones Industrial, the S&P500, or the NASDAQ index drops by at least 20 percent in a time window of either one day, five days, one month, three months, or one year.



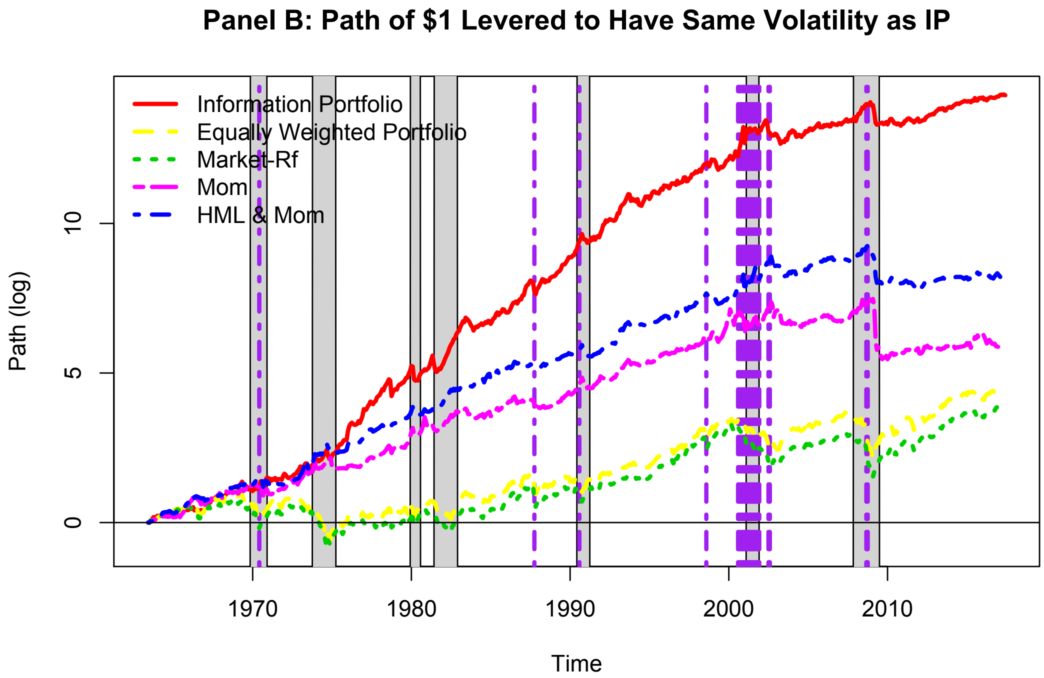


Figure 3: Panel A: cumulated log returns of a zero wealth $1 invested in: information portfolio (red solid line); market portfolio in excess of the risk free rate (green dotted line); $1/N$ portfolio (yellow dashed line); momentum portfolio (purple long-dash-dot line); value and momentum portfolio (dark blue dash-dot line). Panel B: same series as Panel A but with portfolios leveraged to the same volatility as the information portfolio. The information portfolio is non-parametrically extracted at a monthly frequency from the Small, Big, Value, Growth, Winners, Losers and 10 Industry portfolios, using a relative entropy minimization procedure in a rolling out-of-sample fashion starting at 1963:07. Shaded areas indicate NBER recession dates. Vertical dot-dashed lines indicate market crashes identified using the Mishkin and White (2002) approach.

results seem quite robust with respect to the set of risky assets used for its construction, the data frequency, and the subsample considered. This is consistent with the findings in Section IV that, compared to leading multifactor models, the I-SDF comes closest to identifying the tangency portfolio out-of-sample. Moreover, note that the above results have been obtained without searching for either an optimal rolling window or an optimal rebalancing frequency.

VI Conclusion

Given a cross section of asset returns, we show how an information-theoretic approach can be used to estimate, non-parametrically, an out-of-sample pricing kernel. We show that this 'information SDF' prices asset returns as well as, or better than, commonly employed multi-factor models (FF3, Carhart 4-factor, and FF5 models) and that, unlike these factor models, it seems to more closely pin down the tangency portfolio out-of-sample, as a correct SDF should. Moreover, the I-SDF extracts novel pricing information not captured by the Fama–French and momentum factors (which explain only a small share of the I-SDF's time variation). While the I-SDF is a nonlinear function of the asset returns and is, therefore, non-tradeable, a tradeable mimicking portfolio produces statistically and economically large monthly Sharpe ratios, varying from 0.91-1.3 per annum, and αs, varying from 10.6%–18.6% per annum with respect to the FF3, momentum, and FF5 factors. These results hold independently of the set of test assets used. The I-SDF offers a useful benchmark against which competing theories and investment strategies can be evaluated.

References

Ait-Sahalia, Y., and A. W. Lo (1998): "Nonparametric Estimation of State-Price Densities Implicit in Financial Asset Prices," *Journal of Finance*, 53(2), 499–547.

——— (2000): "Nonparametric risk management and implied risk aversion," *Journal of Econometrics*, 94(1-2), 9–51.

Alvarez, F., and U. J. Jermann (2005): "Using Asset Prices to Measure the Persistence of the Marginal Utility of Wealth," *Econometrica*, 73(6), 1977–2016.

Asness, C. S., T. J. Moskowitz, and L. H. Pedersen (2013): "Value and Momentum Everywhere," *The Journal of Finance*, 68(3), 929–985.

Back, K. (2010): *Asset Pricing and Portfolio Choice Theory*, Financial Management Association Survey and Synthesis. Oxford University Press.

Backus, D., M. Chernov, and S. E. Zin (2014): "Sources of Entropy in Representative Agent Models," *The Journal of Finance*, 69(1), 51–99.

Bansal, R., and B. N. Lehmann (1997): "Growth-Optimal Portfolio Restrictions On Asset Pricing Models," *Macroeconomic Dynamics*, 1(02), 333–354.

Barro, R. J. (2006): "Rare Disasters and Asset Markets in the Twentieth Century," *Quarterly Journal of Economics*, 121(3), 823–866.

Borovicka, J., L. P. Hansen, and J. A. Scheinkman (2016): "Misspecified Recovery," *Journal of Finance*, 71, 2493–2544.

Brown, D. E., and R. L. Smith (1990): "A Correspondence Principle for Relative Entropy Minimization," *Naval Research Logistics*, 37(2).

Bryzgalova, S. (2016): "Spurious Factors in Linear Asset Pricing Models," manuscript.

Campbell, J. Y. (2014): "Empirical Asset Pricing: Eugene Fama, Lars Peter Hansen, and Robert Shiller," *Scandinavian Journal of Economics*.

Carhart, M. M. (1997): "On Persistence in Mutual Fund Performance," *Journal of Finance*, 52(1), 57–82.

Cochrane, J. H. (2011): "Presidential Address: Discount Rates," *The Journal of Finance*, 66(4), 1047–1108.

Constantinides, G., and A. Ghosh (2017): "Asset Pricing With Countercyclical Household Consumption Risk," *Journal of Finance*, 72, 415–460.

Csiszar, I. (1975): "I-Divergence Geometry of Probability Distributions and Minimization Problems," *Annals of Probability*, 3, 146–158.

DeMiguel, V., L. Garlappi, and R. Uppal (2009): "Optimal Versus Naive Diversification: How Inefficient is the 1/N Portfolio Strategy?," *Review of Financial Studies*, 22(5), 1915–1953.

Fama, E. F., and K. R. French (1993): "Common Risk Factors in the Returns on Stocks and Bonds," *The Journal of Financial Economics*, 33, 3–56.

——— (2015): "A five-factor asset pricing model," *Journal of Financial Economics*, 116(1), 1–22.

Fama, E. F., and J. MacBeth (1973): "Risk, Return and Equilibrium: Empirical Tests," *Journal of Political Economy*, 81, 607–636.

Ghosh, A., C. Julliard, and A. P. Taylor (2017): "What is the Consumption-CAPM Missing? An Information-Theoretic Framework for the Analysis of Asset Pricing Models," *The Review of Financial Studies*, 30, 442–504.

Hansen, L. P. (2014): "Nobel Lecture: Uncertainty Outside and Inside Economic Models," *Journal of Political Economy*, 122(5), 945 – 987.

Hansen, L. P., and R. Jagannathan (1991): "Implications of Security Market Data for Models of Dynamic Economies," *Journal of Political Economy*, 99(2), 225–62.

Hansen, L. P., and R. Jagannathan (1997): "Asssessing Specification Errors in Stochastic Discount Factor Models," *The Journal of Finance*, 52, 557–590.

Harvey, C. R., and Y. Liu (2015): "Lucky Factors," Duke University manuscript.

Harvey, C. R., Y. Liu, and H. Zhu (2015): ". . . and the Cross-Section of Expected Returns," *Review of Financial Studies*.

Jackwerth, J. C., and M. Rubinstein (1996): "Recovering Probability Distributions from Option Prices," *Journal of Finance*, 51(5), 1611–32.

Julliard, C., and A. Ghosh (2012): "Can Rare Events Explain the Equity Premium Puzzle?," *Review of Financial Studies*, 25(10), 3037–3076.

Kitamura, Y., and M. Stutzer (2002): "Connections between entropic and linear projections in asset pricing estimation," *Journal of Econometrics*, 107(1-2), 159–174.

Lewellen, J., S. Nagel, and J. Shanken (2010): "A skeptical appraisal of asset pricing tests," *Journal of Financial Economics*, 96(2), 175–194.

Lustig, H., N. Roussanov, and A. Verdelhan (2011): "Common Risk Factors in Currency Markets," *Review of Financial Studies*, 24(11), 3731–3777.

Mankiw, N. G. (1982): "Hall's Consumption Hypothesis and Durable Goods," *Journal of Monetary Economics*, pp. 417–26.

McLean, R. D., and J. Pontiff (2016): "Does Academic Research Destroy Stock Return Predictability?," *The Journal of Finance*, 71(1), 5–32.

Mishkin, F. S., and E. N. White (2002): "U.S. Stock Market Crashes and Their Aftermath: Implications for Monetary Policy," NBER Working Papers 8992, National Bureau of Economic Research, Inc.

Robinson, P. M. (1991): "Consistent Nonparametric Entropy-based Testing," *Review of Economic Studies*, 58(3), 437–453.

Rosenberg, J. V., and R. F. Engle (2002): "Empirical pricing kernels," *Journal of Financial Economics*, 64(3), 341–372.

Ross, S. A. (2015): "The Recovery Theorem," *The Journal of Finance*, 70(2), 615–648.

Shanken, J. (1985): "Multivariate Tests of the Zero-Beta CAPM," *Journal of Financial Economics*, 14, 327–348.

Stock, J. (1991): "Confidence Intervals for the Largest Autoregressive Root in US Macroeconomic Time Series," *Journal of Monetary Economics*, 28, 435–459.

Stutzer, M. (1995): "A Bayesian approach to diagnosis of asset pricing models," *Journal of Econometrics*, 68(2), 367 – 397.

——— (1996): "A simple Nonparametric Approach to Derivative Security Valuation," *Journal of Finance*, LI(5), 1633–1652.

Uppal, R., and P. Zaffaroni (2015): "Portfolio Choice with Model Misspecification: A Foundation for Alpha and Beta Portfolios," EDHEC Manuscript.

White, H. (1982): "Maximum Likelihood of Misspecified Models," *Econometrica*, 50, 1–25.

A Appendix

A.1 An Alternative Minimum Entropy Pricing Kernel

The definition of relative entropy, or KLIC, implies that this discrepancy metric is not symmetric, that is, generally $D(\mathbb{A}||\mathbb{B}) \neq D(\mathbb{B}||\mathbb{A})$ unless $\mathbb{A}$ and $\mathbb{B}$ are identical (in which case their divergence would be zero). This implies that for measuring the information divergence between $\mathbb{Q}$ and $\mathbb{P}$, we can also interchange the roles of $\mathbb{Q}$ and $\mathbb{P}$ in equation (2) to recover $\mathbb{Q}$ as

$$\arg\min_{\mathbb{Q}} D(\mathbb{P}||\mathbb{Q}) \equiv \arg\min_{\mathbb{Q}} \int \ln \frac{d\mathbb{P}}{d\mathbb{Q}} d\mathbb{P} \quad \text{s.t.} \quad \int \mathbf{R}_t^e \, d\mathbb{Q} = \mathbf{0}. \tag{7}$$

Since $\frac{M_t}{\overline{M}} = \frac{d\mathbb{Q}}{d\mathbb{P}}$, the optimization in equation (7) can be rewritten as

$$\arg\min_{M_t} \ \mathbb{E}^{\mathbb{P}}[\ln M_t] \quad \text{s.t.} \ \mathbb{E}^{\mathbb{P}}[M_t \mathbf{R}_t^e] = \mathbf{0}.$$

where, to simplify the exposition, we have used the innocuous normalization $\overline{M} = 1$. Replacing the expectation with a sample analogue yields

$$\arg\min_{\{M_t\}_{t=1}^T} \ \frac{1}{T}\sum_{t=1}^{T} \ln M_t \quad \text{s.t.} \ \frac{1}{T}\sum_{t=1}^{T} M_t \mathbf{R}_t^e = \mathbf{0}. \tag{8}$$

Thanks to Fenchel's duality theorem (see, e.g. Csiszar (1975)) this entropy minimization is solved by

$$\widehat{M}_t \equiv M_t\left(\widehat{\theta}_T, \mathbf{R}_t^e\right) = \frac{1}{T(1+\widehat{\theta}_T' \mathbf{R}_t^e)}, \quad \forall t \tag{9}$$

where $\widehat{\theta}_T \in \mathbb{R}^N$ is the solution to

$$\arg\min_{\theta} \ -\frac{1}{T}\sum_{t=1}^{T} \log(1+\theta' \mathbf{R}_t^e),$$

and this last expression is the dual formulation of the entropy minimization problem in equation (8). Note also that this dual problem is analogous to estimating the so-called growth-optimal portfolio (i.e. the portfolio with the maximum log return).

We next proceed to show that the entropy minimization problem in Equation (7) also delivers a maximum likelihood estimate of the risk neutral measure. Let the vector $\mathbf{z}_t$ be a sufficient statistic for the state of the economy at time t. Given $\mathbf{z}_t$, the equilibrium quantities, such as asset returns $\mathbf{R}^e$ and the sdf M, are just a mapping from $\mathbf{z}$ on to the real line, that is,

$$M\left(\mathbf{z}\right):\mathbf{z}\rightarrow\mathbb{R}_{+},\quad \mathbf{R}^{e}\left(\mathbf{z}\right):\mathbf{z}\rightarrow\mathbb{R}^{N},\quad M_{t}\equiv M\left(\mathbf{z}_{t}\right),\quad \mathbf{R}_{t}^{e}\equiv\mathbf{R}^{e}\left(\mathbf{z}_{t}\right),$$

where $\mathbf{z}_t$ is the time t realization of $\mathbf{z}$.

Equipped with the above definition, we can rewrite the Euler equation (1) as

$$0=\mathbb{E}\left[\mathbf{R}_{t}^{e}M_{t}\right]\equiv\int\mathbf{R}_{t}^{e}M_{t}dP=\int\mathbf{R}^{e}\left(\mathbf{z}\right)M\left(\mathbf{z}\right)p\left(\mathbf{z}\right)d\mathbf{z},\tag{10}$$

where $p\left(\mathbf{z}\right)$ is the pdf associated with the physical measure P. Moving to the risk-neutral measure we have

$$0=\mathbb{E}\left[\mathbf{R}_{t}^{e}M_{t}\right]=\mathbb{E}^{Q}\left[\mathbf{R}_{t}^{e}\right]=\int\mathbf{R}^{e}\left(\mathbf{z}\right)q\left(\mathbf{z}\right)d\mathbf{z},\tag{11}$$

where $q\left(\mathbf{z}\right)$ is the pdf associated with the risk-neutral measure Q and $M=dQ/dP$. Note that

$$D\left(P||Q\right)=\int\ln\frac{dP}{dQ}dP=\int p\left(\mathbf{z}\right)\ln p\left(\mathbf{z}\right)d\mathbf{z}-\int p\left(\mathbf{z}\right)\ln q\left(\mathbf{z}\right)d\mathbf{z}.$$

Since the first term on the right-hand side of the above expression does not involve q, $D\left(P||Q\right)$ is minimized, with respect to Q, by choosing the distribution that maximizes the second term, that is,

$$Q^{*}\equiv\arg\min_{Q}D\left(P||Q\right)\equiv\arg\max_{q}\ \mathbb{E}\left[\ln q\left(\mathbf{z}\right)\right]\ \text{ s.t. }\ \mathbb{E}^{Q}\left[\mathbf{R}_{t}^{e}\right]=0.$$

That is, the minimum entropy estimator in Equation (7) maximizes the expected – risk-neutral – log likelihood. Approximating the continuous distribution $q\left(\mathbf{z}\right)$ with a multinomial distribution $\{q_t\}_{t=1}^{T}$ that assigns probability weight q_t to the time t realizations of $\mathbf{z}$, a non-parametric maximum likelihood estimator of Q can be obtained as

$$\begin{aligned}\{q_t^{*}\}_{t=1}^{T}&=\arg\max\quad\frac{1}{T}\sum_{t=1}^{T}\ln q_t\\ &\text{s.t. }\ q_t\in\Delta^{T}\equiv\left\{(q_1,\,q_2,\ldots,\,q_T):q_t\geqslant0,\,\sum_{t=1}^{T}q_t=1\right\}\ \text{and (11) holds,}\end{aligned}\tag{12}$$

provided that

$$\frac{1}{T}\sum_{t=1}^{T}\ln q_t\underset{T\rightarrow\infty}{\overset{p.}{\longrightarrow}}\mathbb{E}\left[\ln q\left(\mathbf{z}\right)\right].$$

Since the correlation of the SDF estimates obtained with either equations (4) or (9) is extremely high (more than 95%), and the pricing performances of the two are almost indistinguishable, to simplify the exposition we present only the results based on the former.

A.2 Performance of the Fama-French Model-Implied SDF

In Section IV, the cross-sectional pricing performance of the I-SDF is compared to those of popular multifactor models such as the Fama-french three factor model. Note that, as explained in Section II, the I-SDF is constructed in a rolling out-of-sample fashion using past data on a chosen cross section of portfolios. One may, therefore, potentially argue that the empirical approach in Section IV does not place the I-SDF and the Fama-French three factor model on an equal footing, because the former is constructed using past data on that particular cross section of portfolios that it is then asked to price out-of-sample, while the latter is not allowed to use such information.

We, therefore, also present the empirical performance of the three factor model when a *Fama-French model-implied SDF* (hereafter referred to as the FF-SDF) is constructed as a linear function of the three risk factors, with the coefficients estimated in a rolling out-of-sample fashion using only past returns data on the particular cross section of portfolios. Specifically, we define the model-implied SDF as:

$$M_t^{FF} = \gamma_0 + \sum_{j=1}^{3} \gamma_j f_{j,t}, \tag{13}$$

where the coefficients γ_j, $j = 0, 1, 2, 3$, are estimated in a rolling out-of-sample fashion using only past returns data on the cross-section of portfolios, so as to satisfy the Euler equation restrictions for these portfolios:

$$0 = E[(R_{i,t} - R_{f,t})(\gamma_0 + \sum_{j=1}^{3} \gamma_j f_{j,t})]. \tag{14}$$

The resulting M_t^{FF} is then used as the single risk factor in the cross sectional regressions to assess its empirical performance.

The results are presented in Table A.1. The table shows that the performance of the one-factor FF-SDF is substantially worse than when the FF factors are treated as three separate risk factors with unrestricted coefficients in the cross-sectional regressions (Tables 1–4). These results hold both at the monthly and quarterly frequencies, regardless of the cross section of test assets used. In particular, the estimated price of risk for the FF-SDF varies wildly, both in magnitude and sign, based on the particular cross section of assets used: at the monthly frequency, it varies from −131.9 for the 10 momentum-sorted portfolios to 2.08 for the 25 long term reversal and size sorted portfolios; at the quarterly frequency, it varies from −97.22 for the 10 momentum-sorted portfolios to 2.54 for the cross section comprised of the smallest and largest deciles of the size, book-to-market-equity, and momentum sorted portfolios and the 10 industry portfolios. In 3 out of the 8 cases in Table A1, the estimated price of risk is not statistically significant at conventional significance levels; in 5 out of the 8 cases, the t-statistic for the estimate is smaller than 3, the threshold advocated in the literature to take into account the effects of data mining. The estimated intercepts, on the other hand, are all statistically significant and economically large, varying from 7.2% (4.8%) to 13.2% (8.0%) at the monthly (quarterly) frequency. The magnitudes of the $\overline{R}^2_{OLS}$ and $\overline{R}^2_{GLS}$ are substantially smaller than those obtained in Tables 1–4. Moreover, in 4 out of 8 cases in Table A1, the 95% confidence intervals for the $\overline{R}^2_{OLS}$ and $\overline{R}^2_{GLS}$ cover almost the entire permissible region, i.e. [0.00, 100], pointing to the low power of inference methods for this particular factor.

Table A.1: Performance of FF Model-Implied SDF

Row	Assets	*const.*	$\lambda_{sdf_{FF}}$	$\overline{R}^2_{OLS}$ (%)	$\overline{R}^2_{GLS}$ (%)	T^2	q
			Panel A: Monthly				
(1)	FF 25	.006	−0.112	16.11	21.55	82.11	0.130
		(7.91)	(−2.37)	[−4.35,85.4]	[1.23,54.82]	(0.000)	[0.04,0.39]
(2)	10 Mom	.011	−131.9	43.49	7.92	1.44	0.063
		(4.85)	(−2.82)	[−12.5,100]	[2.70,100]	(0.986)	[0.00,0.32]
(3)	25 ME & LTR	.009	2.08	33.02	14.34	39.98	0.083
		(21.86)	(3.58)	[−1.22,100]	[1.70,100]	(0.400)	[0.00,0.12]
(4)	S,L,G,V,W,L & 10 Ind	.007	0.194	−4.89	4.61	63.54	0.100
		(3.16)	(0.549)	[−7.14,58.2]	[0.65,49.6]	(0.024)	[0.004,0.36]
			Panel B: Quarterly				
(1)	FF 25	.018	−0.745	50.33	14.41	65.69	0.380
		(12.20)	(−5.03)	[40.0,100]	[0.53,82.3]	(0.023)	[0.019,0.80]
(2)	10 Mom	.012	−97.22	4.73	14.37	7.02	0.179
		(2.53)	(−1.20)	[−12.5,100]	[−0.09,100]	(0.475)	[0.00,1.13]
(3)	25 ME & LTR	.016	−3.56	52.89	5.36	36.16	0.281
		(8.08)	(−5.29)	[6.09,100]	[0.40,100]	(0.180)	[0.00,0.63]
(4)	S,L,G,V,W,L & 10 Ind	.020	2.54	0.50	9.89	53.40	0.298
		(8.54)	(1.04)	[−7.14,87.1]	[3.06,74.09]	(0.021)	[0.012,1.13]

Cross-sectional regressions of average excess returns listed in column 2 on the estimated factor loadings for the Fama-French model-implied SDF. The SDF is extracted from the portfolios listed in Column 2 in a rolling out-of-sample fashion starting at 1963:07. Panel A presents the monthly results and Panel B the quarterly results. For each cross section, the table presents the intercept and slopes, along with t-statistics in parentheses. It also presents the OLS adjusted R^2 and the GLS adjusted R^2, along with the 90% confidence intervals for the true underlying population adjusted R^2 (in square brackets). The confidence intervals are constructed via simulations using the approach suggested by Stock (1991) and used by Lewellen, Nagel, and Shanken (2010). The last two columns present, respectively, Shanken's (1985) cross-sectional T^2 statistic along with its simulated p-value in parentheses; and the q statistic, along with its 90% confidence interval, that measures how far the factor-mimicking portfolios are from the mean–variance frontier.