What is the Consumption-CAPM missing?
An Information-Theoretic Framework for the Analysis of Asset Pricing Models

Anisha Ghosh*  Christian Julliard†  Alex P. Taylor‡

*Tepper School of Business, Carnegie Mellon University; anishagh@andrew.cmu.edu.
†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.
What is the Consumption-CAPM missing?
An Information-Theoretic Framework for the Analysis of Asset Pricing Models

First draft: November 5, 2010. This draft: August 20, 2013

Abstract
We consider asset pricing models in which the SDF can be factorized into an observable component and a potentially unobservable one. Using a relative entropy minimization approach, we extract non-parametrically the most likely time series of the SDF and its potentially unobservable component. Empirically, we find it to have a business cycle pattern, and significant correlations with financial market crashes and the Fama-French factors. Moreover, we derive new (entropy) bounds for the SDF and its components that are tighter, and have higher information content, than the commonly used ones. We apply our methodology to the study of several consumption-based models.

Keywords: Pricing Kernel, Stochastic Discount Factor, Consumption Based Asset Pricing, Entropy Bounds.

JEL Classification Codes: G11, G12, G13, C52.
I Introduction

The absence of arbitrage opportunities implies the existence of a pricing kernel, also known as the stochastic discount factor (SDF), such that the equilibrium price of a traded security can be represented as the conditional expectation of the future payoff discounted by the pricing kernel. The standard consumption-based asset pricing model, within the representative agent and time-separable power utility framework, identifies the pricing kernel as a simple parametric function of consumption growth. However, pricing kernels based on consumption growth alone cannot explain (i) the historically observed levels of returns, giving rise to the Equity Premium and Risk Free Rate Puzzles (e.g. Mehra and Prescott (1985), Weil (1989)), and (ii) the cross-sectional dispersion of returns between different classes of financial assets (e.g. Hansen and Singleton (1983), Mankiw and Shapiro (1986), Breeden, Gibbons, and Litzenberger (1989), Campbell (1996), Cochrane (1996)).

Nevertheless, there is considerable empirical evidence that consumption risk does matter for explaining asset returns (e.g. Lettau and Ludvigson (2001a, 2001b), Parker and Julliard (2005), Hansen, Heaton, and Li (2008), Savov (2011)). Therefore, a burgeoning literature has developed based on modifying the preferences of investors and/or the structure of the economy. In such models the resulting pricing kernel can be factorized into an observable component consisting of a parametric function of consumption growth, and a potentially unobservable, model-specific, component. Prominent examples in this class include: the external habit model where the additional component consists of a function of the habit level (Campbell

1Recently, Julliard and Ghosh (2012) show that pricing kernels based on consumption growth alone cannot explain either the equity premium puzzle, or the cross-section of asset returns, even after taking into account the possibility of rare disasters.
and Cochrane (1999); Menzly, Santos, and Veronesi (2004)); the long run
risks model based on recursive preferences where the additional component
consists of the return on total wealth (Bansal and Yaron (2004)); and mod-
els with housing risk where the additional component consists of the growth
in the expenditure share on non-housing consumption (Piazzesi, Schneider,
and Tuzel (2007)). The additional, and potentially unobserved, component
may also capture deviations from rational expectations (e.g. Brunnermeier
and Julliard (2007)), models with robust control (e.g. Hansen and Sargent
(2010)), heterogeneous agents (e.g. Constantinides and Duffie (1996)), am-
biguity aversion (e.g. Ulrich (2010)), as well as a liquidity factor arising
from solvency constraints (e.g. Lustig and Nieuwerburgh (2005)).

In this paper, we propose a new methodology to analyze dynamic as-
set pricing models for which the SDF can be factorized into an observable
component and a potentially unobservable one. Our no-arbitrage approach
allows us to a) filter from the data the most likely estimate\(^2\) of the time
series of the unobserved pricing kernel and, b) construct entropy bounds to
assess the empirical plausibility of candidate SDFs.

First, we show that, given a set of asset returns and consumption data,
a relative entropy minimization approach can be used to extract, non-
parametrically, the time series of both the SDF and its unobservable com-
ponent. This methodology identifies the most likely, in an information-
theoretic sense, time series of the SDF and its unobservable component.
Along this dimension our paper is close in spirit to, and innovates upon, the
long tradition of using asset prices to estimate the risk neutral probability
measure (see e.g. Jackwerth and Rubinstein (1996), and Aït-Sahalia and
Lo (1998)) and use this information to extract an implied pricing kernel
(see e.g. Aït-Sahalia and Lo (2000), Rosenberg and Engle (2002), and Ross
\(^2\)In the sense of being analogous to a non parametric maximum likelihood procedure.
Empirically, our estimated time series for the unobservable pricing kernel is highly correlated with the Fama and French (1993) factors, for a variety of sample frequencies and assets used in the estimation. This suggests that our approach does a good job in identifying the pricing kernel, and provides a rationalization of the empirical success of the Fama and French factors. Moreover, the estimated most likely SDF ($M^*$) has a clear business cycle pattern but also shows significant and sharp reactions to stock market crashes (even if these crashes do not necessarily result in economy wide contractions). This point is illustrated in Figure 1 that reports the business cycle and residual components of the most likely (log) SDF filtered using the different asset pricing models we consider in our paper.

Second, we construct entropy bounds that restrict the admissible regions for the SDF and its unobservable component. Our results complement and improve upon the seminal work by Hansen and Jagannathan (1991), that provide minimum variance bounds for the SDF. The use of an entropy metric is also closely related to the works of Stutzer (1995, 1996), that first suggested to construct entropy bounds based on asset pricing restrictions, and Alvarez and Jermann (2005), who derive a lower bound for the volatility of the permanent component of investors’ marginal utility of wealth (see also Backus, Chernov, and Zin (2011), Bakshi and Chabi-Yo (2011) and Kitamura and Stutzer (2002)). We show that, in the mean-standard deviation space, a second order approximation of the risk neutral entropy bounds ($Q$-bounds) has the canonical Hansen-Jagannathan bounds as a special case, but are generally tighter since they naturally impose the non negativity restriction on the pricing kernel. Using the multiplicative structure of the pricing kernel, we are able to provide bounds ($M$-bounds) that have higher information content, and are tighter, than both the Hansen and Jagannathan (1991)
and the risk neutral entropy bounds. Moreover, our approach improves upon Alvarez and Jermann (2005) in that a decomposition of the pricing kernel into permanent and transitory components is not required (but is still possible), and we can accommodate an asset space of arbitrary dimension. Our methodology can also be used to construct bounds ($\Psi$-bounds) for the potentially unobserved component of the pricing kernel. We show that for models in which the pricing kernel is a function of observable variables only, the $\Psi$-bounds are the tightest ones, and can be satisfied if and only if the model is actually able to correctly price assets.
Third, we apply our methodology to some of the most well known consumption-based asset pricing models, gaining new insights about their empirical performance. For the standard time separable power utility model, we show that the pricing kernel satisfies the Hansen and Jagannathan (1991) bound for large values of the risk aversion coefficient, and the $Q$ and $M$ bounds for even higher levels of risk aversion. However, the $\Psi$-bound is tighter and is not satisfied for any level of risk aversion. We show that these findings are robust to the use of the long run consumption risk measure of Parker and Julliard (2005), despite the fact that this measure of consumption risk is able to explain a substantial share of the cross-sectional variation in asset returns with a small risk aversion coefficient. Considering more general models of dynamic economies, such as models with habit formation, long run risks in consumption growth, and complementarities in consumption, we find empirical support for the long run risks framework of Bansal and Yaron (2004) and Hansen, Heaton, and Li (2008). Moreover, the empirical application illustrates that inference based on the entropy bounds delivers results that are much more stable, in evaluating the plausibility of a given model across different sets of assets, than the cross-sectional $R^2$ (that instead tends to vary wildly for the same model).

Compared to the previous literature, our nonparametric approach offers four main advantages: i) it can be used to extract information not only from options, but also from any type of financial assets; ii) instead of relying exclusively on the information contained in financial data, it allow us to also exploit the information about the pricing kernel contained in the time series of aggregate consumption, thereby connecting our results to macrofinance modeling; iii) the relative entropy extraction of the SDF is akin to a nonparametric maximum likelihood procedure and thereby provides the most likely estimate of its time series; iv) the methodology has consider-
able generality, and may be applied to any model that delivers well-defined Euler equations and for which the SDF can be factorized into an observable component and an unobservable one (these include investment-based asset pricing models, and models with heterogeneous agents, limited stock market participation, and fragile beliefs).

The remainder of the paper is organized as follows. Section II presents the information-theoretic methodology, the entropy bounds developed, and their properties. Section III uses the Consumption-CAPM with power utility as an illustrative example of the application of our methodology. Section IV applies the methodology developed in this paper to the analysis of more general models of dynamic economies. Section V concludes and discusses extensions. The Appendix contains proofs, additional details on the methodology, and a thorough data description.

II Entropy and the Pricing Kernel

In the absence of arbitrage opportunities, there exists a strictly positive pricing kernel, $M_{t+1}$, or stochastic discount factor (SDF), such that the equilibrium price, $P_{it}$, of any asset $i$ delivering a future payoff, $X_{it+1}$, is given by

$$P_{it} = \mathbb{E}_t [M_{t+1} X_{it+1}] .$$

(1)

where $\mathbb{E}_t$ is the rational expectation operator conditional on the information available at time $t$. For a broad class of models, the SDF can be factorized as follows

$$M_t = m(\theta, t) \times \psi_t$$

(2)

where $m(\theta, t)$ is a known, strictly positive, function of data observable at time $t$ and the parameter vector $\theta \in \mathbb{R}^k$, and $\psi_t$ is a potentially unobserv-
able component. In the most common case, $m(\theta, t)$ is simply a function of consumption growth, i.e. $m(\theta, t) = m(\theta, \Delta c_t)$ where $\Delta c_t \equiv \log \frac{C_t}{C_{t-1}}$ and $C_t$ denotes the time $t$ consumption flow.

Equations (1) and (2) imply that, for any set of tradable assets, the following vector of Euler equations must hold in equilibrium

$$0 = \mathbb{E}[m(\theta, t) \psi_t R_t^\psi] \equiv \int R_t^\psi dP$$

where $\mathbb{E}$ is the unconditional rational expectation operator, $R_t^\psi \in \mathbb{R}^N$ is a vector of excess returns on different tradable assets, and $P$ is the unconditional physical probability measure. Under weak regularity conditions the above pricing restrictions for the SDF can be rewritten as

$$0 = \int m(\theta, t) \frac{\psi_t}{\psi} R_t^\psi dP = \int m(\theta, t) R_t^\psi d\Psi \equiv \mathbb{E}^\Psi [m(\theta, t) R_t^\psi]$$

where $\bar{x} \equiv \mathbb{E}[x_t]$, and $\frac{\psi_t}{\psi} = \frac{\partial \Psi}{\partial P}$ is the Radon-Nikodym derivative of $\Psi$ with respect to $P$. For the above change of measure to be legitimate, we need absolute continuity of the measures $\Psi$ and $P$.

Therefore, given a set of consumption and asset returns data, for any $\theta$, one can obtain a – maximum likelihood – estimate of the $\Psi$ probability measure as follows:

$$\hat{\Psi} \equiv \arg \min_{\Psi} D(\Psi||P) \equiv \arg \min_{\Psi} \int \frac{d\Psi}{dP} \ln \frac{d\Psi}{dP} dP \ \text{s.t.} \ 0 = \mathbb{E}^\Psi [m(\theta, t) R_t^\psi].$$

The above is a relative entropy (or Kullback-Leibler Information Criterion (KLIC)) minimization under the asset pricing restrictions coming from the Euler equations. That is, we can estimate the unknown measure $\Psi$ as the one that adds the minimum amount of additional information needed for the pricing kernel to price assets. Note also that $D(\Psi||P)$ is always non
negative, and has a minimum at zero that is reached when $\Psi$ is identical to $P$, that is when all the information needed to price assets is contained in $m(\theta, t)$ and $\psi_t$ is simply a constant term.

The above approach can also be used, as first suggested by Stutzer (1995), to recover the risk neutral probability measure $(Q)$ from the data as

$$
\hat{Q} \equiv \arg \min_Q D(Q\|P) \equiv \arg \min_Q \int \frac{dQ}{dP} \ln \frac{dQ}{dP} dP \quad \text{s.t.} \quad 0 = \int \mathbf{R}_t^Q dQ \equiv \mathbb{E}^Q [\mathbf{R}_t^Q]
$$

(5)

under the restriction that $Q$ and $P$ are absolutely continuous.

Moreover, since relative entropy is not symmetric, we can also recover $\Psi$ and $Q$ as

$$
\hat{\Psi} \equiv \arg \min_{\Psi} D(P\|\Psi) \equiv \arg \min_{\Psi} \int \ln \frac{dP}{d\Psi} dP \quad \text{s.t.} \quad 0 = \mathbb{E}^\Psi [m(\theta, t) \mathbf{R}_t^\Psi]
$$

$$
\hat{Q} \equiv \arg \min_{Q} D(P\|Q) \equiv \arg \min_{Q} \int \ln \frac{dP}{dQ} dP \quad \text{s.t.} \quad 0 = \mathbb{E}^Q [\mathbf{R}_t^Q]
$$

(6)

Note that the approaches in Equations (4) and (6) can identify $\{\psi_t\}_{t=1}^T$ only up to a positive scale constant.

But why should relative entropy minimization be an appropriate criterion for recovering the unknown measures $\Psi$ and $Q$? There are several reasons for this choice.

First, as formally shown in Appendix A.1, the approaches in Equations (4) and (6) deliver the maximum likelihood estimate of the $\psi_t$ component of the pricing kernel – that is, the most likely estimate given the data at hand. That is, the above KLIC minimization is equivalent to maximizing the likelihood in an unbiased procedure for finding the $\psi_t$ component of the pricing kernel. Note that this is also the rationale behind the principle of maximum entropy (see e.g. Jaynes (1957a, 1957b)) in physical sciences and Bayesian probability that states that, subject to known testable constraints –
the asset pricing Euler restrictions in our case—the probability distribution that best represent our knowledge is the one with maximum entropy, or minimum relative entropy in our notation.

Second, the use of relative entropy, due to the presence of the logarithm in the objective functions in Equations (4)-(7), naturally imposes the non negativity of the pricing kernel. This, for example, is not imposed in the identification of the minimum variance pricing kernel of Hansen and Jagannathan (1991).³

Third, our approach to uncover the $\psi_t$ component of the pricing kernel satisfies the Occam’s razor, or law of parsimony, since it adds the minimum amount of information needed for the pricing kernel to price assets. This is due to the fact that the relative entropy is measured in units of information.

Fourth, it is straightforward to add conditioning information to construct a conditional version of the entropy bounds presented in the next section: given a vector of conditioning variables $Z_{t-1}$, one simply has to multiply (element by element) the argument of the integral constraints in Equations (4), (5), (6) and (7) by the conditioning variables in $Z_{t-1}$.

Fifth, there is no ex-ante restriction on the number of assets that can be used in constructing $\psi_t$, and the approach can naturally handle assets with negative expected rates of return (cf. Alvarez and Jermann (2005)).

Sixth, as implied by the work of Brown and Smith (1990), the use of entropy is desirable if we think that tail events are an important component of the risk measure.⁴

Finally, this approach is numerically simple when implemented via du-

³Hansen and Jagannathan (1991) offer an alternative bound that imposes this restriction, but it is computationally cumbersome (the minimum variance portfolio is basically an option). See also Hansen, Heaton, and Luttmer (1995).

⁴Brown and Smith (1990) develop what they call “a Weak Law of Large Numbers for rare events,” that is they show that the empirical distribution that would be observed in a very large sample converges to the distribution that minimizes the relative entropy.
ality (see e.g. Csiszar (1975)). That is, when implementing the entropy minimization in Equation (4) each element of the series \( \{ \psi_t \}_{t=1}^T \) can be estimated, up to a positive constant scale factor, as

\[
\hat{\psi}_t = \frac{e^{\lambda(\theta')m(\theta, t)}R_t^\psi}{\sum_{t=1}^T e^{\lambda(\theta')m(\theta, t)}R_t^\psi}, \quad \forall t
\]

where \( \lambda(\theta) \in \mathbb{R}^N \) is the solution to

\[
\lambda(\theta) \equiv \arg \min_{\lambda} \frac{1}{T} \sum_{t=1}^T e^{\lambda m(\theta, t)}R_t^\psi,
\]

and this last expression is the dual formulation of the entropy minimization problem in Equation (4).

Similarly, the entropy minimization in Equation (6) is solved by setting each \( \psi_t \), up to a constant positive scale factor, as being equal to

\[
\hat{\psi}_t = \frac{1}{T(1 + \lambda(\theta')m(\theta, t)R_t^\psi)}, \quad \forall t
\]

where \( \lambda(\theta) \in \mathbb{R}^N \) is the solution to the following unconstrained convex problem

\[
\lambda(\theta) \equiv \arg \min_{\lambda} - \sum_{t=1}^T \log(1 + \lambda' m(\theta, t)R_t^\psi),
\]

and this last expression is the dual formulation of the entropy minimization problem in Equation (6).

Note also that the above duality results imply that the number of free parameters available in estimating \( \{ \psi \}_{t=1}^T \) is equal to the dimension of (the Lagrange multiplier) \( \lambda \) – that is, it is simply equal to the number of assets considered in the Euler equation.

Moreover, since the \( \lambda(\theta)'s \) in Equations (9) and (11) are akin to Ex-
tremum Estimators (see e.g. Hayashi (2000, Ch. 7)), under standard regularity conditions (see e.g. Amemiya (1985, Theorem 4.1.3)), one can construct asymptotic confidence intervals for both $\{\psi_t\}_{t=1}^T$ and the entropy bounds presented in the next Section.

II.1 Entropy Bounds

Based on the relative entropy estimation of the pricing kernel and its component $\psi$ outlined in the previous section, we now turn our attention to the derivation of a set of entropy bounds for the SDF, $M$, and its components.

Dynamic equilibrium asset pricing models identify the SDFs as parametric functions of variables determined by the consumers’ preferences and the dynamics of state variables driving the economy. A substantial research effort has been devoted to developing diagnostic methods to assess the empirical plausibility of candidate SDFs in pricing assets as well as provide guidance for the construction and testing of other – more realistic – asset pricing theories.

The seminal work by Hansen and Jagannathan (1991) identifies, in a model-free no-arbitrage setting, a variance minimizing benchmark stochastic discount factor, $M_t^* (\bar{M})$, whose variance places a lower bound on the variances of other SDFs (see Definition 3 in Appendix A.2). The $HJ$-bounds offer a natural benchmark for evaluating the potential of an equilibrium asset pricing model since, by construction, any SDF that is consistent with observed data should have a variance that is not smaller than the one identified by the bound. However, the identified minimum variance SDF does not impose the non negativity constraint on the pricing kernel and, since $M_t^* (\bar{M})$ is a linear function of returns, it does not generally satisfy the restriction.5

5We call the bound in Definition 3 the “canonical” $HJ$-bound since Hansen and Jagannathan (1991, 1997) also provide an alternative bound, that imposes the non-negativity
As noticed in Stutzer (1995), using the Kullback-Leibler Information Criterion minimization in Equation (5), one can construct an entropy bound for the risk neutral probability measure that naturally imposes the non negativity constraint on the pricing kernel. In Definition 4 in Appendix A.2 we generalize the idea of using an entropy minimization approach to construct risk neutral bounds – $Q$-bounds – for the pricing kernel. These bounds, like the $HJ$-bound, use only the information contained in asset returns but, differently from the latter, they impose the restriction that the pricing kernel must be positive. Moreover, under mild regularity conditions, we show that (see Remark 1 in Appendix A.2), to a second order approximation, the problem of constructing canonical $HJ$-bounds and $Q$-bounds are equivalent, in the sense that approximated $Q$-bounds identify the minimum variance bound for the SDF. The intuition behind this result is simple: a) a second order approximation of (the log of) a smooth pdf delivers an approximately Gaussian distribution (see e.g. Schervish (1995)); b) the relative entropy of a Gaussian distribution is proportional to its variance; c) the diffusion invariance principle (see e.g. Duffie (2005, Appendix D)) implies that in the continuous time limit the change of measure does not change the volatility.

Both the $HJ$ and $Q$ bounds described above use only information about asset returns and neither information about consumption growth, nor the structure of the pricing kernel. Instead, we propose a novel approach that, while also imposing the non negativity of the pricing kernel, a) takes into account more information about the form of the pricing kernel, therefore delivering sharper bounds, and b) allows us to construct information bounds for the individual components of the SDFs.

Consider an SDF that, as in Equation (2), can be factorized into two of the pricing kernel, but that is computationally complex.

\[^{6}\text{The (sufficient, but not necessary) regularity conditions required for the approximation result stated above are typically satisfied in consumption-based asset pricing models.}\]
components, i.e. \( M_t = m(\theta, t) \times \psi_t \) where \( m(\theta, t) \) is a known non negative function of observable variables (generally consumption growth) and the parameter vector \( \theta \), and \( \psi_t \) is a potentially unobservable component. A large class of equilibrium asset pricing models, including ones with time separable power utility with a constant coefficient of relative risk aversion, external habit formation, recursive preferences, durable consumption goods, housing, and disappointment aversion, fall into this framework. Based on the above factorization of the SDF we can define the following bounds.

**Definition 1 (M-bounds)** For any candidate stochastic discount factor of the form in Equation (2), and given any choice of the parameter vector \( \theta \), we define the following bounds:

1. **M1-bound:**
   \[
   D \left( P \| \frac{M_t}{M} \right) = \int - \ln \frac{M_t}{M} dP \geq D \left( P \| \frac{m(\theta, t) \psi_t^*}{m(\theta, t) \psi_t^*} \right) = \int - \ln \frac{m(\theta, t) \psi_t^*}{m(\theta, t) \psi_t^*} dP
   \]
   where \( \psi_t^* \) solves Equation (6) and \( m(\theta, t) \psi_t^* \equiv \mathbb{E}[m(\theta, t) \psi_t^*] \).

2. **M2-bound:**
   \[
   D \left( \frac{M_t}{M} \| P \right) = \int \frac{M_t}{M} \ln \frac{M_t}{M} dP \geq D \left( \frac{m(\theta, t) \psi_t^*}{m(\theta, t) \psi_t^*} \| P \right) = \int \frac{m(\theta, t) \psi_t^*}{m(\theta, t) \psi_t^*} \ln \frac{m(\theta, t) \psi_t^*}{m(\theta, t) \psi_t^*} dP
   \]
   where \( \psi_t^* \) solves Equation (4).

The above bounds for the SDF are tighter than the \( Q \)-bounds since, denoting with \( Q^* \) the minimum entropy risk neutral probability measure,
we have that

\[ D \left( P \left| \frac{m(\theta, t) \psi_t^*}{m(\theta, t) \psi_t} \right. \right) \geq D \left( P \left| Q^* \right. \right) \quad \text{and} \quad D \left( \frac{m(\theta, t) \psi_t^*}{m(\theta, t) \psi_t} \left| P \right. \right) \geq D \left( Q^* \left| P \right. \right) \]

by construction,\(^7\) and are also more informative since not only is the information contained in asset returns used in their construction, but also the structure of the pricing kernel in Equation (2) and the information contained in \( m(\theta, t) \).

Information about the SDF can also be elicited by constructing bounds for the \( \psi_t \) component itself. Given the \( m(\theta, t) \) component, these bounds identify the minimum amount of information that \( \psi_t \) should add for the pricing kernel \( M_t \) to be able to price asset returns.

**Definition 2 (\( \Psi \)-bounds)** For any candidate stochastic discount factor of the form in Equation (2), and given any choice of the parameter vector \( \theta \), two lower bounds for the relative entropy of \( \psi_t \) are defined as:

1. \( \Psi1 \)-bound:

\[ D \left( P \left| \frac{\psi_t}{\psi} \right. \right) = - \int \frac{\psi_t}{\psi} \ln \frac{\psi_t}{\psi} dP \geq D \left( P \left| \frac{\psi_t^*}{\psi} \right. \right) \]

where \( \psi_t^* \) solves Equation (6);

2. \( \Psi2 \)-bound

\[ D \left( \frac{\psi_t}{\psi} \left| P \right. \right) = \int \frac{\psi_t}{\psi} \ln \frac{\psi_t}{\psi} dP \geq D \left( \frac{\psi_t^*}{\psi} \left| P \right. \right) \]

where \( \psi_t^* \) solves Equation (4).

Besides providing an additional check for any candidate SDF, the \( \Psi \)-bounds are useful in that a simple comparison of \( D \left( \frac{\psi_t}{\psi} \left| P \right. \right) \), \( D \left( \frac{m(\theta, t)}{m(\theta, t)} \left| P \right. \right) \)

\(^7\)Cf. Definition 4 in Appendix A.2.
and $D(Q^*||P)$ can provide a very informative decomposition in terms of the entropy contribution to the pricing kernel, that is logically similar to the widely used variance decomposition analysis. For example, if $D\left(\frac{\psi_t}{\overline{\psi}}||P\right)$ happens to be close to $D(Q^*||P)$, while $D\left(\frac{m(\theta,t)}{m(\overline{\theta},t)}||P\right)$ is substantially smaller, the decomposition would imply that most of the ability of the candidate SDF to price assets comes from the $\psi_t$ component.

Moreover, note that if we want to evaluate a model of the form $M_t = m(\theta,t)$ – i.e. a model without the unobservable $\psi_t$ component – the $\Psi$-bounds will offer a tight selection criterion since, under the null of the model being true, we should have $D\left(\frac{\psi_t}{\overline{\psi}}||P\right) = D\left(P||\frac{\psi_t}{\overline{\psi}}\right) = 0$ and this is a tighter bound than the $HJ$, $Q$ and $M$ bounds defined above. The intuition for this is simple: $Q$-bounds (and $HJ$-bounds) require the model under test to deliver at least as much relative entropy (variance) as the minimum relative entropy (variance) SDF, but they do not require that the $m(\theta,t)$ under scrutiny should also be able to price the assets. That is, it might be the case – as in practice we will show is the case – that for some values of $\theta$ both the $Q$-bounds and the $HJ$-bounds will be satisfied, but nevertheless the SDF grossly violates the pricing restrictions in the Euler Equation (3).

Note that in principle a volatility bound, similar to the Hansen and Jagannathan (1991) bound for the pricing kernel, can be constructed for the $\psi_t$ component. Such a bound, presented in Definition 5 of Appendix A.2, identifies a minimum variance $\psi_t^*\left(\overline{\psi}\right)$ component with standard deviation given by

$$
\sigma_{\psi^*} = \overline{\psi} \sqrt{\mathbb{E}[R_t^e m(\theta,t)]^{\prime} \text{Var}(R_t^e m(\theta,t))^{-1} \mathbb{E}[R_t^e m(\theta,t)]}.
$$

(12)

This bound, as the entropy based $\Psi$-bounds in Definition 2, uses information about the structure of the SDF but, differently from the latter, does not
constrains $\psi_t$ and $M_t$ to be non-negative as implied by economic theory. Moreover, using the same approach employed in Remark 1, this last bound can be obtained as a second order approximation of the entropy based $\Psi$-bounds.

Equation (12), viewed as a second order approximation to the entropy $\Psi$-bounds, makes clear why bounds based on the decomposition of the pricing kernel as $M_t = m(\theta, t) \psi_t$ offer sharper inference than bounds based on only $M_t$. Consider for example the case in which the candidate SDF is of the form $M_t = m(\theta, t)$, that is $\psi_t = 1$ for any $t$. In this case, it can easily happen that there exists a $\tilde{\theta}$ such that

$$Var \left( M_t \left( \tilde{\theta} \right) \right) \equiv Var \left( m \left( \tilde{\theta}, t \right) \right) \geq Var \left( M^*_t \left( \tilde{M} \right) \right)$$

where $Var \left( M^*_t \left( \tilde{M} \right) \right)$ is the Hansen and Jagannathan (1991) bound in Definition 3 of Appendix A.2, that is there exists a $\tilde{\theta}$ such that the $HJ$-bound is satisfied. Nevertheless, the existence of such a $\tilde{\theta}$ does not imply that the candidate SDF is able to price asset returns. This would be the case if and only if the volatility bound for $\psi_t$ in Definition 5 is also satisfied since, from Equation (12), we have that under the assumption of constant $\psi_t$ the bound can be satisfied only if $E \left[ R_t^\xi m (\theta_0, t) \right] \equiv E \left[ R_t^\xi M_t (\theta_0) \right] = 0$, that is only if the candidate SDF is able to price asset returns.

**II.1.1 Entropy Bounds and the Second Hansen-Jagannathan distance**

Our entropy bounds are also connected to the *second* Hansen and Jagannathan distance. Given a model that defines a SDF $M$, Hansen and Jagannathan (1997) assumes that portfolio payoffs are elements of an Hilbert space and consider the minimum squared deviation between $M$ and a pric-
ing kernel \( q \in \mathcal{M} \) (or \( \mathcal{M}^+ \) if nonnegativity is imposed), where \( \mathcal{M} \) denotes the set of all admissible stochastic discount factors. That is, the second HJ distance is defined as

\[
d^2_{HJ} := \min_{q \in \mathcal{M}} \| M - q \|^2 = \min_{q \in \mathcal{M}} \mathbb{E} \left[ (M_t - q_t)^2 \right].
\]

Note that \( q \in \mathcal{M} \) can be rewritten as \( q \in L^2 \) satisfying the pricing restriction (1), that is

\[
d^2_{HJ} := \min_{q \in L^2} \mathbb{E} \left[ (M_t - q_t)^2 \right] \quad \text{s.t.} \quad 0 = \mathbb{E} [q_t \mathbb{R}_t] = \mathbb{E}^Q [\mathbb{R}_t].
\]

Note that the constraint above is the same one used for constructing our risk neutral entropy bounds. Nevertheless, the nature of the minimization is different since in constructing the entropy bounds we search for the most likely pricing kernel, while in the above definition \( q \) is chosen to minimize the second moment of the deviation from the candidate SDF \( M \).

The entropy bounds focus on the space of distribution functions, and since there is a one to one mapping between distributions and moments generating functions, one would expect that the entropy bounds we propose would carry information not only about second moments, but also about all the other moments of the stochastic discount factors. As we are about to show, this is indeed the case.

Consider the risk neutral entropy \( Q_1 \)-bound. Given a candidate SDF \( M \), this defines the entropy distance

\[
d_{Q1} = D \left( P || \hat{Q} \right) - D \left( P || M_t \right)
\]

where \( \hat{Q} \) solves Equation (7) and we have normalized \( M_t \) to have unit mean to simplify the exposition. This distance must be non-positive for the can-
didate SDF $M_t$ to satisfy the $Q_1$-bound (since to satisfy the bound one needs $D(P||M_t) \geq D\left(P||\hat{Q}\right)$). Denoting by $\hat{q}$ the minimum entropy SDF identified by the bound, the above distance can be rewritten as

$$d_{Q_1} = \ln \mathbb{E} \left[ e^{\ln \hat{q}} \right] - \ln \mathbb{E} \left[ e^{\ln M_t} \right] + \int \ln M_t dP - \int \ln \hat{q} dP,$$

where we used the fact that by construction $\mathbb{E} [M_t] = \mathbb{E} [\hat{q}] = 1$. Note also that, by construction, $\hat{q} \in \mathcal{M}^+$, that is the relative entropy minimization identifies an admissible SDF in the Hansen and Jagannathan (1997) sense.

To illustrate the links between our bounds and the second HJ distance, we follow the cumulant expansion approach of Backus, Chernov, and Zin (2011). Recall that the cumulant generating function (i.e. the log of the moment generating function) of a random variable $\ln x_t$ is

$$k^x(s) = \ln \mathbb{E} \left[ e^{s \ln x_t} \right]$$

and, with appropriate regularity conditions, it admits the power series expansion

$$k^x(s) = \sum_{j=1}^{\infty} \frac{\kappa_j s^j}{j!}, \quad (13)$$

where the $j$-th cumulant, $\kappa_j$, is the $j$-th derivative of $k^x(s)$ evaluated at $s = 0$. That is, $\kappa_j^x$ capture the $j$-th moment of the variable $\ln x_t$, i.e. $\kappa_1^x$ reflects the mean of the variable, $\kappa_2^x$ the variance, $\kappa_3^x$ the skewness, $\kappa_4^x$ the kurtosis and so on. For instance, if $\ln x_t \sim N(\mu_x; \sigma_x^2)$, we have $\kappa_1^x = \mu_x$, $\kappa_2^x = \sigma_x^2$, $\kappa_{j>2}^x = 0$.

We can therefore rewrite $d_{Q_1}$ and $D\left(P||\hat{Q}\right)$ as

$$d_{Q_1} = k^x(1) - \kappa_1^x - (k^M(1) - \kappa_1^M) = \frac{\kappa_2^\hat{q} - \kappa_2^M}{2!} + \frac{\kappa_3^\hat{q} - \kappa_3^M}{3!} + \frac{\kappa_4^\hat{q} - \kappa_4^M}{4!} + \ldots$$

19
\[
D \left( P \| Q \right) = \frac{\kappa_2^q}{2!} + \frac{\kappa_3^q}{3!} + \frac{\kappa_4^q}{4!} + ..
\]

That is, like in the second HJ distance, the first term in the infinite sums above captures the Gaussian terms of the distributions, while the other terms captures the nonnormal elements. That is, the second HJ distance looks at the second moments deviations, the relative entropy approach of the Q1 bounds looks at all the moments of the candidate SDF, \( M \), and the minimum entropy one, \( \hat{q} \).

But note that there is an important conceptual difference between the second HJ distance, and the risk neutral \( D \left( P \| Q \right) \) bound. The former, looks form the minimum adjustment – in a least square sense – that makes \( M - \lambda^t \mathbf{R}^t \) an admissible stochastic discount factor (where \( \lambda \) arises form the linear projection of \( M \) on the space of returns). The latter instead identifies the most likely SDF given the data, \( \hat{q} \), and offers it as a benchmark to which a candidate SDF \( M \) can be compared.

The idea of minimum adjustment of the second HJ distance is also strongly connected to our \( M \) bounds. Consider the decomposition \( M_t = m(\theta, t) \psi_t \) in its extreme form: \( M_t \equiv m(\theta, t) \) i.e. the case in which the candidate SDF is fully observable and, under the null of the model under scrutiny, \( \psi \) should simply be a constant. In this case the \( M_1 \)-bound defines the distance

\[
d_{M_1} = D \left( P \| M_t \hat{\psi}_t \right) - D \left( P \| M_t \right) \equiv D \left( P \| \psi_t \right)
\]

where \( \hat{\psi} = \frac{d \Phi}{d P} \), \( \hat{\Psi} \) solves Equation (6), and we have normalized \( \hat{\psi}_t \) to have unit mean (to simplify exposition). Note also that, by construction, \( M_t \hat{\psi}_t \in \mathcal{M} \) (or \( \mathcal{M}^+ \) if \( M \) is nonnegative), that is once again the relative entropy minimization identifies an admissible SDF in the Hansen and Jagannathan
sense. Using the cumulant expansion, the above distance can be rewritten as

\[ d_{M1} = \frac{\kappa_2^{\hat{\psi}}}{2!} + \frac{\kappa_3^{\hat{\psi}}}{3!} + \frac{\kappa_4^{\hat{\psi}}}{4!} + \ldots \]  

(14)

where \( \kappa_j^{\hat{\psi}} \) denotes the \( j \)-th cumulant of \( \log \hat{\psi} \), and \( \hat{\psi} \) and solves

\[ \arg \min_{\{\psi_t\}_{t=1}} \frac{\kappa_2^{\hat{\psi}}}{2!} + \frac{\kappa_3^{\hat{\psi}}}{3!} + \frac{\kappa_4^{\hat{\psi}}}{4!} + \ldots \quad \text{s.t.} \quad 0 = E \left[ m(\theta, t) R_t^\Psi \right]. \]  

(15)

The above implies that the \( \hat{\psi} \) component identified by our \( M1 \) bound has a very similar interpretation to the second Hansen and Jagannathan distance: it provides the minimum adjustment – in the entropy sense – that would make \( m(\theta, t) \hat{\psi}_t \) an admissible stochastic discount factor. The key difference between the second HJ bound and our \( M1 \) is that the former focuses only on the minimum second moment deviation, while our bound takes into consideration not only the second moment (captured by the \( \kappa_2^{\hat{\psi}} \) cumulant in equation (14)), but also all other possible moments (captured by the \( \kappa_{j>2}^{\hat{\psi}} \) cumulants). This implies that if skewness, kurtosis, tail probabilities etc. are relevant for asset pricing, our approach is more desirable than the least square one. Moreover, note that the cumulant generating function cannot be a finite-order polynomial of degree greater than two (see Theorem 7.3.5 of Lukacs (1970)). That is, if mean and variance are not sufficient statistics for the distribution of the true SDF, then all the other higher moments become relevant for characterizing the SDF, and their relevance for asset pricing is captured by our entropy approach given the one to one mapping between relative entropy and cumulants.

The above cumulant decomposition also allows us to assess the relevance of higher moments for pricing asset returns. In particular, with the estimated \( \left\{ \ln \hat{\psi}_t \right\}_{t=1}^T \) at hand, we can estimate its moments using sample
analogues, use these moments to compute the cumulants, and finally compute the contribution of the $j$-th cumulant to the total entropy of $\hat{\psi}$ as

$$ \frac{\kappa_j^\psi / j!}{\sum_{s=2}^\infty \kappa_s^\psi / s!} \equiv \frac{\kappa_j^\psi / j!}{D\left(P || \hat{\psi}\right)} $$

as well as the total contribution of cumulants of order bigger than $j$ as

$$ \frac{\sum_{s=j+1}^\infty \kappa_s^\psi / s!}{\sum_{s=2}^\infty \kappa_s^\psi / s!} \equiv \frac{D\left(P || \hat{\psi}\right) - \sum_{s=2}^j \kappa_s^\psi / s!}{D\left(P || \hat{\psi}\right)}. $$

These statistics are important for comparing the informativeness of our bounds relative to the second HJ distance since, if the minimum variance deviation had all the relevant information for pricing asset returns, we would expect

$$ \frac{D\left(P || \hat{\psi}\right) - \kappa_2^\psi / 2!}{D\left(P || \hat{\psi}\right)} \cong 0 \quad \text{and} \quad \frac{\kappa_j^\psi / j!}{D\left(P || \hat{\psi}\right)} \cong 0 \quad \forall j > 2. $$

As we will show in the empirical section below, this is not the case.

### III An Illustrative Example: the C-CAPM with Power Utility

We first illustrate our methodology for the Consumption-CAPM (C-CAPM) of Breeden (1979), Lucas (1978) and Rubinstein (1976), when the utility function is time and state separable with a constant coefficient of relative risk aversion. For this specification of preferences, the SDF takes the form,

$$ M_{t+1} = \delta (C_{t+1}/C_t)^{-\gamma}, $$

(18)
where \( \delta \) denotes the subjective time discount factor, \( \gamma \) is the coefficient of relative risk aversion, and \( C_{t+1}/C_t \) denotes the real per capita aggregate consumption growth. Empirically, the above pricing kernel fails to explain i) the historically observed levels of returns, giving rise to the Equity Premium and Risk Free Rate Puzzles (e.g. Mehra and Prescott (1985), Weil (1989)), and ii) the cross-sectional dispersion of returns between different classes of financial assets (e.g. Mankiw and Shapiro (1986), Breeden, Gibbons, and Litzenberger (1989), Campbell (1996), Cochrane (1996)).

Parker and Julliard (2005) argue that the covariance between contemporaneous consumption growth and asset returns understates the true consumption risk of the stock market if consumption is slow to respond to return innovations. They propose measuring the risk of an asset by its ultimate risk to consumption, defined as the covariance of its return and consumption growth over the period of the return and many following periods. They show that, while the ultimate consumption risk would correctly measure the risk of an asset if the C-CAPM were true, it may be a better measure of the true risk if consumption responds with a lag to changes in wealth. The ultimate consumption risk model implies the following SDF:

\[
M_{t+1}^S = \delta^{1+S} (C_{t+1+S}/C_t)^{-\gamma} R_{t+1,t+1+S}^{f},
\]

where \( S \) denotes the number of periods over which the consumption risk is measured and \( R_{t+1,t+1+S}^{f} \) is the risk free rate between periods \( t + 1 \) and \( t + 1 + S \). Note that the standard C-CAPM obtains when \( S = 0 \). Parker and Julliard (2005) show that the specification of the SDF in Equation (19), unlike the one in Equation (18), explains a large fraction of the variation in expected returns across assets for low levels of the risk aversion coefficient.

The functional forms of the above two SDFs fit into our framework in
Equation (2). For the contemporaneous consumption risk model, \( \theta = \gamma \), \( M_t = m(\theta, t) = (C_{t+1}/C_t)^{-\gamma} \), and \( \psi_t = \delta \), a constant, for all \( t \). For the ultimate consumption risk model, \( \theta = \gamma \), \( m(\theta, t) = (C_{t+1+S}/C_t)^{-\gamma} \), and \( \psi_t = \delta^{1+S}R_{t+1,t+1+S}^f \). Therefore, for each model, we construct entropy bounds for the SDF and its components using quarterly data\(^8\) on per capita real personal consumption expenditures on nondurable goods and returns on the 25 Fama-French portfolios over the post war period 1947:1-2009:4 and compare them with the \( HJ \) bound. We also obtain the non-parametrically extracted (called "filtered" hereafter) SDF and its components for \( \gamma = 10 \). For the ultimate consumption risk model, we set \( S = 11 \) quarters because the fit of the model is the greatest at this value as shown in Parker and Julliard (2005).

Figure 2, Panel A plots the relative entropy (or KLIC) of the filtered and model-implied SDFs and their unobservable components as a function of the risk aversion coefficient \( \gamma \) and the \( HJ \), \( Q1 \), \( M1 \), and \( \Psi1 \) bounds for the contemporaneous consumption risk model in Equation (18). The black curve with circles shows the relative entropy of the model-implied SDF as a function of the risk aversion coefficient. For this model, the missing component of the SDF, \( \psi_t \), is a constant hence it has zero relative entropy for all values of \( \gamma \), as shown by the orange straight line with triangles. The blue curve with "+" signs and the yellow curve with inverted triangles show, respectively, the relative entropy as a function of the risk aversion coefficient of the filtered SDF and its missing component. The model satisfies the \( HJ \) bound for very high values of \( \gamma \geq 64 \), as shown by the green dotted-dashed vertical line. It satisfies the \( Q1 \) bound for even higher values of \( \gamma \geq 72 \), as shown by the red dashed vertical line. The minimum value of \( \gamma \) at which the \( M1 \) bound is satisfied is given by the value corresponding to the intersection

\(^8\)See Appendix A.3 for a thorough data description.
of the black and blue curves, i.e. it is the minimum value of $\gamma$ for which the relative entropy of the model-implied SDF exceeds that of the filtered SDF. The figure shows that this corresponds to $\gamma = 107$. Finally, the $\Psi_1$ bound identifies the minimum value of $\gamma$ for which the missing component of the model-implied SDF has a higher relative entropy than the missing component of the filtered SDF. Since the former has zero relative entropy while the latter has a strictly positive value for all values of $\gamma$, the model fails to satisfy the $\Psi_1$ bound for any value of $\gamma$.

Figure 2: The figure plots the KLIC of the filtered and model-implied SDFs and their unobservable components as a function of the risk aversion coefficient and the entropy bounds for the standard CCAPM.

Panel B shows that very similar results are obtained for the $Q_2$, $M_2$, and $\Psi_2$ bounds. The $Q_2$ and $M_2$ bounds are satisfied for values of $\gamma$ at least as large as 73 and 99, respectively, while the $\Psi_2$ bound is not satisfied for any value of $\gamma$. Overall, as suggested by the theoretical predictions, the $Q$-bounds are tighter than the $HJ$-bound, the $M$-bounds are tighter than
the $Q$-bounds, and the $\Psi$-bounds are tighter than the $M$-bounds.

We also construct confidence bands for above the relative entropy bounds using 1,000 simulations of the same length as the historical time series. The 95% confidence bands for the $Q_1$ and $Q_2$ bounds extend over the intervals $[70.0, 109.0]$ and $[69.5, 109.0]$, respectively and those for the $M_1$ and $M_2$ bounds cover the intervals $[94.5, 157.5]$ and $[86.0, 150.0]$, respectively. Finally, the $\Psi_1$ and $\Psi_2$ bounds are not satisfied for any finite value of the risk aversion coefficient in any of the simulated samples. The simulation results reveal two points. First, it demonstrates the robustness of our approach - the two different definitions of relative entropy produce very similar results. Second, the confidence bands are quite tight in contrast with the large values of the standard error typically obtained when using GMM type approaches to estimate the risk aversion parameter.

Figure 3 presents analogous results to Figure 2 for the ultimate consumption risk model in Equation (19). Panel A shows that the $HJ$, $Q_1$, and $M_1$ bounds are satisfied for $\gamma \geq 22, 23, \text{ and } 46$, respectively. These are almost three times, more than three times, and more than two times smaller, respectively, than the corresponding values in Figure 2, Panel A, for the contemporaneous consumption risk model. As for the latter model, the $\Psi_1$ bound is not satisfied for any value of $\gamma$. Panel B shows that the $Q_2$ and $M_2$ bounds are satisfied for $\gamma \geq 24 \text{ and } 47$, respectively, while the $\Psi_2$ bound is not satisfied for any value of $\gamma$. The simulated 95% confidence bands for the $Q_1$ and $Q_2$ bounds extend over the intervals $[23.0, 35.0]$ and $[24.0, 37.0]$, respectively, and those for the $M_1$ and $M_2$ bounds cover the intervals $[36.0, 60.0]$ and $[40.0, 74.0]$, respectively. Also, similar to the contemporaneous consumption risk model, the $\Psi_1$ and $\Psi_2$ bounds are not satisfied for any finite value of the risk aversion coefficient in any of the simulated samples.
It is important to notice that, even though the best fitting level for the RRA coefficient for the ultimate consumption risk model is smaller than 10 ($\gamma = 1.5$), and at this value of the coefficient the model is able to explain about 60% of the cross-sectional variation in returns across the 25 Fama-French portfolios, all the bounds reject the model for low RRA, and the $\Psi$ bounds are not satisfied for any level of RRA. This stresses the power of the proposed approach.

The above results indicate that our entropy bounds are not only theoretically, but also empirically, tighter than the HJ variance bounds. Using the cumulants decomposition introduced in the previous section, we can identify the information content added by taking into account higher moments of the SDF and its components. In particular, the statistics in equations (16) (blue dot-dash line) and (17) (red dashed line) are plotted in the left panels of Figure 4 (for $S = 0$) and Figure 5 (for $S = 11$).
Figure 4: $S = 0$. Left panel: cumulants contribution to $D \left( P \mid \hat{P} \right)$. Right panel: densities of $m_t := \left( \frac{C_{t+1}}{C_t} \right)^{-10}$ and $M_t := \left( \frac{C_{t+1}}{C_t} \right)^{-10} \hat{\psi}_t$.

Figure 5: $S = 11$. Left panel: cumulants contribution to $D \left( P \mid \hat{P} \right)$. Right panel: densities of $m_t := \left( \frac{C_{t+12}}{C_t} \right)^{-10} R_{t+1,t+12}^f$ and $M_t := \left( \frac{C_{t+12}}{C_t} \right)^{-10} R_{t+1,t+12}^f \hat{\psi}_t$. 
The Figures show that the contribution of the second moment to $D \left( P || \hat{\Psi} \right)$ is large – being in the 74-78% range – but that higher moments also play a very important role, with their cumulated contribution being in the 22–26% range. Among these higher moments, the lion’s share goes to the Skewness, with it’s individual contribution being about 18% for both $S = 0$ and $S = 11$.

The relevance of Skewness is also outlined in the right panels of Figure 4 (for $S = 0$) and Figure 5 (for $S = 11$) where the (Epanechnikov kernel estimates of the) densities of $m_t := \left( \frac{C_{t+1+s}}{C_t} \right)^{-10} R_{t+1,t+1+S}^f$ and $M_t := \left( \frac{C_{t+1+s}}{C_t} \right)^{-10} R_{t+1,t+1+S}^I \hat{\psi}_t$ are reported. The figures illustrate that, beside the increase in variance generated by $\hat{\psi}$, there is also a substantial increase of the skewness for our estimated most likely pricing kernel. This point is also outlined in figures 6 (for $S = 0$) and 7 (for $S = 11$) where the left panels report the cumulant decomposition of the entropy of $m_t$ while the right panel reports the cumulant decomposition for $M_t := m_t \hat{\psi}_t$. The figure shows that the sources of entropy of our most likely pricing kernel ($m_t \hat{\psi}_t$) are very different than the ones of the consumption growth component alone ($m_t$): almost all (99%) the entropy of $m_t$ is generated by it’s second moment, while higher cumulants have basically no role; instead, about a quarter (24 – 25%) of the entropy of $m_t \hat{\psi}_t$ is generated by the third and higher cumulants.

We now turn to the analysis of the time series properties of the candidate SDFs considered. Figure 8, Panel A plots the time series of the filtered SDF and its components estimated using Equation (6) for $\gamma = 10$ for the contemporaneous consumption risk model. The blue dotted line plots the component of the SDF that is a parametric function of consumption growth, $m(\theta, t) = (C_t/C_{t-1})^{-\gamma}$. The red dashed line plots the filtered unobservable component of the SDF, $\hat{\psi}^*_t$, estimated using Equation (6). The black solid
Figure 6: $S = 0$. Left panel: cumulants contribution to $D \left( P \| \left( \frac{C_{t+1}}{C_t} \right)^{-10} \right)$. Right panel: cumulants contribution to $D \left( P \| \left( \frac{C_{t+1}}{C_t} \right)^{-10} \hat{\psi}_t \right)$.

Figure 7: $S = 11$. Left panel: cumulants contribution to $D \left( P \| \left( \frac{C_{t+12}}{C_t} \right)^{-10} R_{t+1,t+12}^f \right)$. Right panel: cumulants contribution to $D \left( P \| \left( \frac{C_{t+12}}{C_t} \right)^{-10} R_{t+1,t+12}^f \hat{\psi}_t \right)$. 

30
Figure 8: The figure plots the (demeaned) time series of the filtered SDF and its components for the standard CCAPM for $\gamma=10$. Shaded areas are NBER recession periods. Vertical dashed lines are the stock market crashes identified by Mishkin and White (2002).

Panel A

Panel B

The figure reveals two main points. First, the estimated SDF has a clear business cycle pattern, but also shows significant and sharp reactions to financial market crashes that do not result in economy-wide contractions. Second, the time series of the SDF almost coincides with that of the unobservable component. In fact, the correlation between the two time series is 0.996. The observable consumption growth component of the SDF, on the other hand, has a correlation of only 0.06 with the SDF. Therefore, most of the variation in the SDF comes from variation in the unobservable component, $\psi$, and not from the consumption growth component. In fact,
the volatility of the SDF and its unobservable component are very similar with the latter explaining about 99% of the volatility of the former, while the volatility of the consumption growth component accounts for only about 1% of the volatility of the filtered SDF. Similar results are obtained in Panel B that plots the time series of the filtered SDF and its components estimated using Equation (4) for $\gamma = 10$.

![Graph](image)

Figure 9: The figure plots the (demeaned) time series of the filtered SDF and its components for the ultimate consumption risk CCAPM for $\gamma=10$. Shaded areas are NBER recession periods. Vertical dashed lines are the stock market crashes identified by Mishkin and White (2002).

Finally, Figure 9, Panel A plots the time series of the filtered SDF and its components estimated using Equation (6) for $\gamma = 10$ for the ultimate consumption risk model. The figure shows that, as in the contemporaneous consumption risk model, the estimated SDF has a clear business cycle pattern, but also shows significant and sharp reactions to financial market crashes that do not result in economy wide contractions. However, dif-
ferently from the latter model, the time series of the consumption growth component is much more volatile and more highly correlated with the SDF. The volatility of the consumption growth component is 21.7\%, more than 2.5 times higher than that for the standard model. The correlation between the filtered SDF and its consumption growth component is 0.37, an order of magnitude bigger than the correlation of 0.06 in the contemporaneous consumption risk model. This explains the ability of the model to account for a much larger fraction of the variation in expected returns across the 25 Fama-French portfolios for low levels of the risk aversion coefficient. In fact, the cross-sectional $R^2$ of the model is 54.1\% (for $\gamma = 10$), an order of magnitude higher than the value of 5.2\% for the standard model. However, the correlation between the ultimate consumption risk SDF and its unobservable component is still very high at 0.92, showing that the model is missing important elements that would further improve its ability to explain the cross-section of returns. Similar results are obtained in Panel B that plots the time series of the filtered SDF and its components estimated using Equation (4) for $\gamma = 10$.

Overall, the results show that our methodology provides useful diagnostics for dynamic asset pricing models. Moreover, the very similar results obtained using the two different types of relative entropy minimization in Equations (4) and (6), suggest robustness of our approach.

IV Application to More General Models of Dynamic Economies

Our methodology provides useful diagnostics to assess the empirical plausibility of a large class of consumption-based asset pricing models where the SDF, $M_t$, can be factorized into an observable component consisting of a parametric function of consumption, $C_t$, as in the standard time-separable
power utility model, and a potentially unobservable one, $\psi_t$, that is model-specific:

$$M_t = \left(\frac{C_t}{C_{t-1}}\right)^{-\gamma} \psi_t.$$

In this section, we apply it to a set of "winners" asset pricing models, i.e. frameworks that can successfully explain the Equity Premium and the Risk Free Rate Puzzles with "reasonable" calibrations. In particular, we consider the external habit formation models of Campbell and Cochrane (1999) and Menzly, Santos, and Veronesi (2004), the long-run risks model of Bansal and Yaron (2004), and the housing model of Piazzesi, Schneider, and Tuzel (2007). We apply our methodology to assess the empirical plausibility of these models in two ways. First, since our methodology identifies the most likely time-series of the SDF, we compare this time-series with the model-implied time-series of the SDF for each model. Moreover, we show that, independently from the set of assets used for its estimation, the most likely SDF is highly correlated with the Fama-French factors, it has a clear business cycle pattern, and shows sharp and significant reaction to financial market crashes (even if these crashes do not map into a strong contraction in aggregate consumption). Second, for each model we compute the values of the power coefficient, $\gamma$, at which the model-implied SDF satisfies the $HJ$, $Q$, $M$, and $\Psi$ bounds. To simplify the exposition, we focus on one-dimensional bounds as a function of the risk aversion parameter, $\gamma$, while fixing the other parameters at the authors’ preferred values. We show that, as suggested by the theoretical predictions, the $Q$-bounds are generally tighter than the $HJ$-bound, and the $M$-bounds are always tighter than both $HJ$ and $Q$ bounds.

In the next Sub-Section we present the models considered. The reader familiar with these models can go directly to Section IV.2, that reports the
empirical results, without loss of continuity. A detailed data description is presented in Appendix A.3.

IV.1 The Models Considered

IV.1.1 External Habit Formation Model: Campbell and Cochrane (1999)

In this model, identical agents maximize power utility defined over the difference between consumption and a slow-moving habit or time-varying subsistence level. The SDF is given by

\[ M_t = \delta \left( \frac{C_t}{C_{t-1}} \right)^{-\gamma} \left( \frac{S_t}{S_{t-1}} \right)^{-\gamma}, \]

where \( \delta \) is the subjective time discount factor, \( \gamma \) is the curvature parameter that provides a lower bound on the time varying coefficient of relative risk aversion, \( S_t = \frac{C_t-X_t}{C_t} \) denotes the surplus consumption ratio and \( X_t \) is the habit component. In this model, the expression for \( \ln(\psi_t) \) is given by:

\[ \ln \psi_t = \ln \delta - \gamma \Delta s_t. \] (20)

where \( \Delta s_t := \ln \left( \frac{S_t}{S_{t-1}} \right) \). Note that, in this model, the missing component, \( \psi \), depends on the surplus consumption ratio, \( S_t \), that is not observed.

To obtain the time series of \( \psi \), we extract the surplus consumption ratio from observed consumption data as follows. In this model, the aggregate consumption growth is assumed to follow an \textit{i.i.d.} process:

\[ \Delta c_t = g + v_t, \quad v_t \sim \text{i.i.d.} N \left( 0, \sigma^2 \right). \]

The log surplus consumption ratio evolves as a heteroskedastic AR(1) process:

\[ s_t = (1 - \phi) \bar{s} + \phi s_{t-1} + \lambda (s_{t-1}) v_t, \] (21)
where \( \bar{s} \) is the steady state log surplus consumption ratio and

\[
\lambda(s_t) = \begin{cases} 
\frac{1}{2} \sqrt{1 - 2(s_t - \bar{s})} - 1, & \text{if } s_t \leq s_{max} \\
0, & \text{if } s_t > s_{max}
\end{cases},
\]

\[
s_{max} = \bar{s} + \frac{1}{2} \left( 1 - \bar{S}^2 \right), \quad \bar{S} = \sigma \sqrt{\frac{\gamma}{1 - \phi}}.
\]

For each value of \( \gamma \), we use the calibrated values of the model preference parameters \((\delta, \phi)\) in Campbell and Cochrane (1999), the sample mean \((g)\) and volatility \((\sigma)\) of the consumption growth process, and the innovations in real consumption growth, \(\hat{\nu}_t = \Delta c_t - g\), to extract the time series of the surplus consumption ratio using Equation (21) and obtain the time series of the model-implied SDF and its missing component.

### IV.1.2 External Habit Formation Model: Menzly, Santos, and Veronesi (2004)

In this model, the SDF and its missing component are analogous to those in the Campbell and Cochrane (1999) model. The aggregate consumption growth is also assumed to follow an \( i.i.d. \) process:

\[
dc_t = \mu_c dt + \sigma_c dB_t,
\]

where \( \mu_c \) is the mean consumption growth, \( \sigma_c > 0 \) is a scalar, and \( B_t \) is a Brownian motion. The point of departure from the Campbell and Cochrane (1999) framework is that the Menzly, Santos, and Veronesi (2004) model assumes that the inverse surplus consumption ratio, \( Y_t \equiv \frac{1}{S_t} \), follows a mean reverting process that is perfectly negatively correlated with innovations in consumption growth:

\[
dY_t = k \left( \bar{Y} - Y_t \right) dt - \alpha (Y_t - \lambda) \left[ dc_t - E \left( dc_t \right) \right], \quad (22)
\]
where $\bar{Y}$ is the long run mean of the inverse surplus consumption ratio and $k$ controls the speed of mean reversion. For each value of $\gamma$, we use the calibrated values of the model parameters ($\delta$, $k$, $\bar{Y}$, $\alpha$, $\lambda$) in Menzly, Santos, and Veronesi (2004), the sample values of $\mu_c$ and $\sigma_c$, and the innovations in real consumption growth, $dB_t = \frac{dc_t - E(dc_t)}{\sigma_c}$, to extract the time series of the surplus consumption ratio, and that allows us to compute the time series of the model-implied SDF and its missing component.

**IV.1.3 Long-Run Risks Model: Bansal and Yaron (2004)**

The Bansal and Yaron (2004) long-run risks model assumes that the representative consumer has the version of Kreps and Porteus (1978) preferences adopted by Epstein and Zin (1989) and Weil (1989) for which the SDF is given by

$$\ln M_{t+1} = \theta \log \delta - \frac{\theta}{\rho} \Delta c_{t+1} + (\theta - 1)r_{c,t+1},$$

(23)

where $r_{c,t+1}$ is the unobservable log gross return on an asset that delivers aggregate consumption as its dividend each period, $\delta$ is the subjective time discount factor, $\rho$ is the elasticity of intertemporal substitution, $\theta = \frac{1-\gamma}{1-1/\rho}$, and $\gamma$ is the relative risk aversion coefficient.

The aggregate consumption and dividend growth rates, $\Delta c_{t+1}$ and $\Delta d_{t+1}$ respectively, are modeled as containing a small persistent expected growth rate component, $x_t$, that follows an AR(1) process with stochastic volatility, and fluctuating variance, $\sigma_x^2$, that evolves according to a homoscedastic linear mean reverting process.

For the log-linearized version of the model, the log price-consumption ratio, $z_t$, the log price-dividend ratio, $z_{m,t}$, and the log gross risk free rate, $r_{f,t}$, are affine functions of the state variables, $x_t$ and $\sigma_x^2$. Therefore, Constantinides and Ghosh (2011) argue that these affine functions may be inverted to express the unobservable state variables, $x_t$ and $\sigma_x^2$, in terms of
the observables, $z_{m,t}$ and $r_{f,t}$. Following this approach, the pricing kernel in Equation (23) can be expressed, in log-linearized form, as a function of observable variables

$$\ln M_{t+1} = c_1 - \gamma \Delta c_{t+1} + c_3 (r_{f,t+1} - \kappa_1 r_{f,t}) + c_4 (z_{m,t+1} - \kappa_1 z_{m,t}), \quad (24)$$

where the parameters $c = (c_1, c_3, c_4)'$ are functions of the time-series and preference parameters.

The model is calibrated at the monthly frequency. Since, due to data availability, we assess the empirical plausibility of models at the quarterly and annual frequencies, we obtain the pricing kernels at these frequencies by aggregating the monthly kernels. For instance, the quarterly pricing kernel, $M_q^q$, is obtained as

$$\ln M_q^q = -\gamma \Delta_q c_t + \ln \psi_t \quad (25)$$

where $\Delta_q c_t$ denotes quarterly log-consumption difference and $\ln \psi_t$ is given by

$$\ln \psi_t = 3c_1 + \sum_{i=0}^{2} [c_3 (r_{f,t-i} - \kappa_1 r_{f,t-i-1}) + c_4 (z_{m,t-i} - \kappa_1 z_{m,t-i-1})]. \quad (26)$$

For each value of $\gamma$, we use the calibrated parameter values from Bansal and Yaron (2004) and the time series of the price-dividend ratio and risk free rate to obtain the time series of the SDF and its $\psi$ component in Equations (25) and (26).

**IV.1.4 Housing: Piazzesi, Schneider, and Tuzel (2007)**

In this model, the pricing kernel is given by:

$$M_t = \delta \left( C_t / C_{t-1} \right)^{-\gamma} \left( A_t / A_{t-1} \right)^{\frac{\gamma - 1}{\gamma - 1}},$$
where $A_t$ is the expenditure share on non-housing consumption, $\gamma^{-1}$ is the intertemporal elasticity of substitution and $\rho$ is the intratemporal elasticity of substitution between housing services and non-housing consumption.

Taking logs we have:

$$\ln M_t = \ln \delta - \gamma \Delta c_t + \frac{\gamma \rho - 1}{\rho - 1} \Delta a_t. \quad (27)$$

Therefore, in this model, the expression for $\ln \psi_t$ is given by:

$$\ln \psi_t = \ln \delta + \frac{\gamma \rho - 1}{\rho - 1} \Delta a_t, \quad (28)$$

For each value of $\gamma$, we use the calibrated values of the model parameters ($\delta$, $\rho$) in Piazzesi, Schneider, and Tuzel (2007) to obtain the time series of the model-implied SDF and its missing component from Equations (27) and (28), respectively.

**IV.2 Empirical Results**

For our empirical analysis, we focus on two data samples: an annual data sample starting at the onset of the Great Depression (1929 – 2009), and a quarterly data sample starting in the post World War II period (1947 : Q1 – 2009 : Q4). A detailed data description is presented in Appendix A.3. Note that, in any finite sample, the extracted time series of the SDF, as well as the information bounds on the SDF and its unobservable component, depend on the set of test assets used for their construction. Since the Euler equation holds for any traded asset as well as any adapted portfolio of assets, this gives an infinitely large number of moment restrictions. Nevertheless, econometric considerations necessitate the choice of only a subset of assets to be used. As a consequence, in our empirical analysis, we compute bounds, and extract the time series of the SDF and its components, using a large variety of
cross-sections of test assets, and we show that the empirical findings are quite robust to the set of test assets used.

IV.2.1 The Time Series of the Most Likely SDF

Our first approach to assessing the empirical plausibility of these models is based on the observation that our methodology identifies the most likely time-series of the SDF, which we call the filtered SDF. We compare the filtered SDF with the model-implied SDF for each model. Note that the filtered SDF and its missing component depend on the local curvature of the utility function, $\gamma$. Therefore, for each model, we fix $\gamma$ at the authors’ calibrated value, and extract the time series of the SDF and its components.

Table I reports the results at the quarterly frequency. In order to examine the models’ ability to explain the cross-section of asset returns, we focus on multiple sets of test assets. Panels A, B, C, D, and E report, respectively, results for the following sets of assets: 25 Fama-French, 10 size-sorted, 10 book-to-market-equity-sorted, 10 momentum-sorted, and 10 industry-sorted portfolios. The first column reports the correlation between the filtered time series of the missing component, $\{\psi_t^*\}_{t=1}^T$, of the SDF and the corresponding model-implied time series, $\{\psi_t^m\}_{t=1}^T$. The second column shows the correlation between the filtered SDF, $\{M_t^* = m_t\psi_t^*\}_{t=1}^T$, where $m_t = (C_t/C_{t-1})^{-\gamma}$, and the model-implied SDF, $\{M_t^m = m_t\psi_t^m\}_{t=1}^T$. The 95% confidence intervals for these correlations, obtained from 1000 simulations of the same length as the historical time series, are reported in square brackets below.

Consider first the results for the CC external habit model that are presented in the first row of each panel. For this model, the utility curvature parameter is set to the calibrated value of $\gamma = 2$. Panel A, Column 1 shows that when the 25 FF portfolios are used in the extraction of $\psi^*$, the cor-
The table reports the correlation between the extracted and the model-implied stochastic discount factors and their missing components using quarterly data over 1947:2-2009:4 and a different set of portfolios in each Panel. The acronyms CC, MSV, BY and PST, denote respectively the models of Campbell and Cochrane (1999), Menzly, Santos, and Veronesi (2004), Bansal and Yaron (2004) and Piazzesi, Schneider, and Tuzel (2007).

### Table I: Correlation of Filtered and Model SDFs, 1947:Q2-2009:Q4

<table>
<thead>
<tr>
<th>Panel: Fama-French 25 Portfolios</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \rho (\ln \psi_t^n, \ln \psi_t^m) )</td>
</tr>
<tr>
<td>CC</td>
</tr>
<tr>
<td>MSV</td>
</tr>
<tr>
<td>BY\textsuperscript{rest} (unrest.)</td>
</tr>
<tr>
<td>MSV</td>
</tr>
<tr>
<td>BY\textsuperscript{rest} (unrest.)</td>
</tr>
<tr>
<td>PST</td>
</tr>
</tbody>
</table>

### Panel B: 10 Size-Sorted Portfolios

<table>
<thead>
<tr>
<th>Panel: 10 BM-Sorted Portfolios</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \rho (\ln \psi_t^n, \ln \psi_t^m) )</td>
</tr>
<tr>
<td>CC</td>
</tr>
<tr>
<td>MSV</td>
</tr>
<tr>
<td>BY\textsuperscript{rest} (unrest.)</td>
</tr>
<tr>
<td>PST</td>
</tr>
</tbody>
</table>

### Panel C: 10 Momentum-Sorted Portfolios

<table>
<thead>
<tr>
<th>Panel: 10 Industry-Sorted Portfolios</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \rho (\ln \psi_t^n, \ln \psi_t^m) )</td>
</tr>
<tr>
<td>CC</td>
</tr>
<tr>
<td>MSV</td>
</tr>
<tr>
<td>BY\textsuperscript{rest} (unrest.)</td>
</tr>
<tr>
<td>PST</td>
</tr>
</tbody>
</table>

### Panel D: 10 Industry-Sorted Portfolios

<table>
<thead>
<tr>
<th>Panel: 10 Industry-Sorted Portfolios</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \rho (\ln \psi_t^n, \ln \psi_t^m) )</td>
</tr>
<tr>
<td>CC</td>
</tr>
<tr>
<td>MSV</td>
</tr>
<tr>
<td>BY\textsuperscript{rest} (unrest.)</td>
</tr>
<tr>
<td>PST</td>
</tr>
</tbody>
</table>

The table reports the correlation between the extracted and the model-implied stochastic discount factors and their missing components using quarterly data over 1947:2-2009:4 and a different set of portfolios in each Panel. The acronyms CC, MSV, BY and PST, denote respectively the models of Campbell and Cochrane (1999), Menzly, Santos, and Veronesi (2004), Bansal and Yaron (2004) and Piazzesi, Schneider, and Tuzel (2007).
relation between the filtered and model-implied $\psi$ is only 0.02 when $\psi^*$ is estimated using Equation (6). Column 2 shows that the correlation between the filtered and model-implied SDFs is marginally higher at 0.05. When $\psi^*$ is estimated using Equation (4), the correlations are very similar at 0.06 and 0.08, respectively. Panels $B - E$ show that the correlations between the filtered and model-implied SDFs and their missing components remain small for all the other sets of portfolios.

The second row in each panel presents the results for the MSV external habit model. In this case, $\gamma$ is set equal to 1 which is the calibrated value in the model. Row 2 in each panel shows that the results for the MSV model are very similar to those for the CC model. When $\psi^*$ is estimated using Equation (6), the correlations between the filtered and model-implied missing components of the SDFs are small, varying from 0.00 for the 25 FF portfolios to 0.20 for the size-sorted portfolios. The correlations between the filtered and model-implied SDFs are marginally higher, varying from 0.02 for the 25 FF portfolios to 0.24 for the size-sorted portfolios. Similar results are obtained when $\psi^*$ is estimated using Equation (4).

The third row in each panel presents the results for the BY long run risks model. As shown in Equation (26), in the long run risks model the $\psi$ component of the SDF is an exponentially affine function of the market-wide log price-dividend ratio and its lag, and the log risk free rate and its lag. But the parameters of the affine relation are functions of the underlying model parameters, some of which are not “deep” preference parameters but instead characterizations of the data generating processes. Since the parameters of the data generating processes could be in principle different in different samples, we present two types of results for the SDF of the BY model. First, we present results where the restrictions on the vector of parameters of the affine relation implied by the BY calibration are imposed
Second, we provide results where the parameter vector is treated as free (in parentheses in Row 3). The parameter $\gamma$ is set equal to the BY calibrated value of 10. Row 3, Panel A, Column 1 shows that when the 25 FF portfolios are used in filtering the SDF, the correlation between the filtered and model-implied missing components of the SDFs is 0.10 (0.11) when the restrictions are imposed on the coefficients vector $c$, and $\psi^*$ is estimated using Equation (6) (Equation (4)). This is an order of magnitude higher than the values obtained for the CC and MSV models in Rows 1 and 2, respectively. When the coefficients $c$ are treated as free parameters, the correlation more than doubles from 0.10 (0.11) to 0.27 (0.29). Column 2 shows that the correlation between the filtered and model-implied SDFs is 0.11 (0.12) in the presence of the restrictions and is more than two times higher at 0.30 (0.31) when the restrictions are not imposed.

Similar results are obtained in Panels B-E for the other sets of test assets. The correlation between the filtered and model-implied missing components of the SDF varies from 0.12 (0.11) for the 10 momentum-sorted portfolios to 0.38 (0.38) for the size-sorted portfolios for the restricted specification. These are often an order of magnitude higher than the correlations obtained for the CC and MSV models. For the unrestricted specification, the correlations more than double, varying from 0.44 (0.44) for the 10 momentum-sorted portfolios to 0.80 (0.82) for the size-sorted portfolios. These results show that the SDF implied by the long run risks model correlates much more strongly with the non-parametrically extracted most likely time series of the SDF than the external habit models of CC and MSV.

The fourth row in each panel presents the results for the PST housing model. In this case, $\gamma$ is set equal to 16 which is the calibrated value in the original paper. Column 1 shows that the correlations between the filtered and model-implied missing components of the SDFs are very small and often
have the wrong sign, varying from $-0.22 (-0.20)$ for the size-sorted portfolios to $0.13 (0.12)$ for the industry-sorted portfolios when $\psi^*$ is estimated using Equation (6) (Equation (4)). The correlations between the filtered and model-implied SDFs are marginally higher varying from $-0.01 (-0.02)$ for the size-sorted portfolios to $0.19 (0.17)$ for the industry-sorted portfolios.

Table II reports results analogous to those in Table I at the annual frequency. The results are largely similar to those in Table I. The table shows that, at the annual frequency, the SDF implied by the long run risks model correlates even more strongly with the filtered SDF relative to the external habit and housing models.

The last two columns of Tables I and II report the cross-sectional $R^2$’s, along with the simulated 95% confidence bands in square brackets below, implied by the model-specific SDFs for the different sets of test assets. The cross-sectional $R^2$ are obtained by performing a cross-sectional regression of the historical average returns on the model-implied expected returns. Column 3 reports the cross-sectional $R^2$ when there is no intercept in the regression while Column 4 presents results when an intercept is included. The results reveal that the cross-sectional $R^2$’s vary wildly for the same model, and often take on large negative values when an intercept is not allowed in the cross-sectional regression, or when evaluated using different sets of assets. Moreover, they have very wide confidence intervals. As we show in the next section, this is in stark contrast with the results based on entropy bounds in Tables VI and VII, that tend instead to give consistent results and tighter confidence bands for each model across different sets of assets and samples.

A notable exception to the poor cross-sectional performance of the models considered is that, at the annual frequency, the BY model, unlike the CC, MSV, and PST models, has stable cross-sectional $R^2$ for the size and
Table II: Correlations of Filtered and Model SDFs, 1930-2009

<table>
<thead>
<tr>
<th></th>
<th>Correlation of filtered and model SDF</th>
<th>Cross-sectional $R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\rho (\ln \psi_i^<em>, \ln \psi_{m,i}^</em>)$</td>
<td>$\rho (\ln M_i^<em>, \ln M_{i,m}^</em>)$</td>
</tr>
<tr>
<td><strong>Panel A: Fama-French 6 Portfolios</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$CC$</td>
<td>0.15 / 0.16</td>
<td>0.23 / 0.22</td>
</tr>
<tr>
<td>$MSV$</td>
<td>[-14.35] / [-15.36]</td>
<td>[-0.82] / [-11.41]</td>
</tr>
<tr>
<td>$BY_{rest.}$</td>
<td>0.33 / 0.36</td>
<td>0.24 / 0.29</td>
</tr>
<tr>
<td>$PST$</td>
<td>-0.04 / -0.01</td>
<td>-0.01 / -0.05</td>
</tr>
</tbody>
</table>

| **Panel B: 10 Size-Sorted Portfolios** | | | |
| $CC$ | -0.026 / -0.03 | 0.075 / 0.06 | -0.88 | 0.172 |
| $BY_{rest.}$ | 0.47 / 0.50 | 0.36 / 0.40 | 0.36 | 0.96 |
| $PST$ | 0.17 / 0.13 | -0.01 / -0.08 | 0.115 | 0.914 |

| **Panel C: 10 BM-Sorted Portfolios** | | | |
| $CC$ | 0.07 / 0.03 | 0.16 / 0.10 | -0.12 | 0.01 |
| $MSV$ | [-24.32] / [-29.31] | [-17.39] / [-24.35] | [-7.09] / [-7.75] | [0.88] / [0.89] |
| $BY_{rest.}$ | 0.52 / 0.53 | 0.41 / 0.47 | 0.40 | 0.47 |
| $PST$ | 0.22 / 0.34 | 0.08 / 0.08 | 0.18 | 0.57 |

| **Panel D: 10 Momentum-Sorted Portfolios** | | | |
| $CC$ | 0.26 / 0.27 | 0.34 / 0.33 | 0.90 | 0.00 |
| $MSV$ | [0.09] / [0.07] | [0.09] / [0.08] | [-1.5] | 0.40 |
| $BY_{rest.}$ | 0.41 / 0.50 | 0.31 / 0.41 | -0.33 | 0.31 |
| $PST$ | -0.07 / -0.06 | -0.03 / -0.06 | -0.45 | 0.01 |

| **Panel E: 10 Industry-Sorted Portfolios** | | | |
| $CC$ | -0.03 / -0.04 | 0.03 / 0.06 | -4.62 | 0.23 |
| $BY_{rest.}$ | 0.26 / 0.39 | 0.20 / 0.34 | -1.24 | 0.17 |
| $PST$ | 0.12 / 0.12 | -0.07 / -0.20 | -7.15 | 0.56 |

The table reports the correlation between the extracted and the model-implied stochastic discount factors and their missing components using annual data over 1930-2009 and a different set of portfolios in each Panel. The acronyms $CC$, $MSV$, $BY$ and $PST$, denote respectively the models of Campbell and Cochrane (1999), Menzly, Santos, and Veronesi (2004), Bansal and Yaron (2004) and Piazzesi, Schneider, and Tuzel (2007).
BM-sorted portfolios both in the presence and absence of an intercept.

Overall, Tables I and II make two main points. First, they demonstrate the robustness of our estimation methodology – very similar results are obtained using Equations (6) and (4). Second, they show that the long run risks model implies an SDF that is the most highly correlated with the filtered SDF – the most likely SDF given the data.

The correlations between model specific SDFs and filtered SDFs discussed above would have little significance if the filtered discount factors had no clear economic interpretation. In order to address this concern, we show below that our filtered pricing kernel has clear economic content since a) it is always highly correlated with the Fama-French factors, that can be interpreted as proxies for the true unknown sources of systematic risk, b) it implies that the most likely SDF should have a strong business cycle pattern, and c) react significantly to financial market crashes.

Tables III and IV report the correlations between the filtered and model-implied SDFs and the three Fama-French (FF) factors at the quarterly and annual frequencies, respectively. Column 1 presents the correlation between the model-implied SDF and the three FF factors. This is computed by performing a linear regression of the model-implied time series of the SDF, \( \{M_t^m\}_{t=1}^T \), on the three FF factors and computing the correlation between \( M_t^m \) and the fitted value from the regression. Similarly, Columns 2 and 3 present the correlations of the filtered SDF and its missing component with the three FF factors, respectively.

Consider first Table III. Row 1 of each panel shows that for the CC model, the correlation between the model-implied SDF and the three FF factors is small at 0.18. Panel A, Row 1, Column 2 shows that, while the model-implied SDF correlates poorly with the FF factors, the filtered SDF correlates very highly with the factors having a correlation coefficient of
Table III: Correlations with FF3, 1947:Q2-2009:Q4

<table>
<thead>
<tr>
<th></th>
<th>Correlation With FF3</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \ln M_t^{\alpha} )</td>
<td>( \ln M_t^{\gamma} )</td>
<td>( \ln \psi_t^{\ast} )</td>
<td></td>
</tr>
<tr>
<td><strong>Panel A: Fama-French 25 Portfolios</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( CC )</td>
<td>0.18</td>
<td>0.54/0.59</td>
<td>0.54/0.59</td>
<td></td>
</tr>
<tr>
<td>( MSV )</td>
<td>0.21</td>
<td>0.54/0.59</td>
<td>0.54/0.59</td>
<td></td>
</tr>
<tr>
<td>( BY^{rest} )</td>
<td>0.45</td>
<td>0.54/0.58</td>
<td>0.52/0.57</td>
<td></td>
</tr>
<tr>
<td>( (unrest.) )</td>
<td>(0.87)</td>
<td>(0.54)</td>
<td>(0.53)</td>
<td>(0.57)</td>
</tr>
<tr>
<td>( PST )</td>
<td>0.07</td>
<td>0.49/0.52</td>
<td>0.45/0.50</td>
<td></td>
</tr>
<tr>
<td><strong>Panel B: 10 Size-Sorted Portfolios</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( CC )</td>
<td>0.18</td>
<td>0.88/0.89</td>
<td>0.87/0.89</td>
<td></td>
</tr>
<tr>
<td>( MSV )</td>
<td>0.21</td>
<td>0.87/0.89</td>
<td>0.87/0.89</td>
<td></td>
</tr>
<tr>
<td>( BY^{rest} )</td>
<td>0.45</td>
<td>0.89/0.90</td>
<td>0.86/0.88</td>
<td></td>
</tr>
<tr>
<td>( (unrest.) )</td>
<td>(0.90)</td>
<td>(0.89)</td>
<td>(0.86)</td>
<td>(0.88)</td>
</tr>
<tr>
<td>( PST )</td>
<td>0.07</td>
<td>0.81/0.82</td>
<td>0.75/0.76</td>
<td></td>
</tr>
<tr>
<td><strong>Panel C: 10 BM-Sorted Portfolios</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( CC )</td>
<td>0.18</td>
<td>0.83/0.86</td>
<td>0.83/0.86</td>
<td></td>
</tr>
<tr>
<td>( MSV )</td>
<td>0.21</td>
<td>0.83/0.86</td>
<td>0.83/0.86</td>
<td></td>
</tr>
<tr>
<td>( BY^{rest} )</td>
<td>0.45</td>
<td>0.84/0.86</td>
<td>0.81/0.85</td>
<td></td>
</tr>
<tr>
<td>( (unrest.) )</td>
<td>(0.91)</td>
<td>(0.84)</td>
<td>(0.81)</td>
<td>(0.85)</td>
</tr>
<tr>
<td>( PST )</td>
<td>0.07</td>
<td>0.87/0.89</td>
<td>0.84/0.87</td>
<td></td>
</tr>
<tr>
<td><strong>Panel D: 10 Momentum-Sorted Portfolios</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( CC )</td>
<td>0.18</td>
<td>0.52/0.52</td>
<td>0.51/0.51</td>
<td></td>
</tr>
<tr>
<td>( MSV )</td>
<td>0.21</td>
<td>0.52/0.52</td>
<td>0.51/0.51</td>
<td></td>
</tr>
<tr>
<td>( BY^{rest} )</td>
<td>0.45</td>
<td>0.55/0.53</td>
<td>0.50/0.50</td>
<td></td>
</tr>
<tr>
<td>( (unrest.) )</td>
<td>(0.92)</td>
<td>(0.55)</td>
<td>(0.50)</td>
<td>(0.50)</td>
</tr>
<tr>
<td>( PST )</td>
<td>0.07</td>
<td>0.53/0.51</td>
<td>0.43/0.43</td>
<td></td>
</tr>
<tr>
<td><strong>Panel E: 10 Industry-Sorted Portfolios</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( CC )</td>
<td>0.18</td>
<td>0.65/0.69</td>
<td>0.64/0.68</td>
<td></td>
</tr>
<tr>
<td>( MSV )</td>
<td>0.21</td>
<td>0.65/0.69</td>
<td>0.65/0.68</td>
<td></td>
</tr>
<tr>
<td>( BY^{rest} )</td>
<td>0.45</td>
<td>0.66/0.69</td>
<td>0.62/0.65</td>
<td></td>
</tr>
<tr>
<td>( (unrest.) )</td>
<td>(0.88)</td>
<td>(0.66)</td>
<td>(0.62)</td>
<td>(0.65)</td>
</tr>
<tr>
<td>( PST )</td>
<td>0.07</td>
<td>0.53/0.55</td>
<td>0.47/0.51</td>
<td></td>
</tr>
</tbody>
</table>

The table reports the correlations between the 3 Fama-French factors and the model-implied SDF, the filtered SDF, and the missing component of the filtered SDF using quarterly data over 1947:2-2009:4 and a different set of portfolios in each Panel. The acronyms \( CC \), \( MSV \), \( BY \) and \( PST \), denote respectively the models of Campbell and Cochrane (1999), Menzly, Santos, and Veronesi (2004), Bansal and Yaron (2004) and Piazzesi, Schneider, and Tuzel (2007).
Table IV: Correlations with FF3, 1930-2009

<table>
<thead>
<tr>
<th></th>
<th>$\ln M_1^0$</th>
<th>$\ln M_1^*$</th>
<th>$\ln \psi_1^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Fama-French 6 Portfolios</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$CC$</td>
<td>0.19</td>
<td>0.73/0.78</td>
<td>0.72/0.77</td>
</tr>
<tr>
<td>$MSV$</td>
<td>0.12</td>
<td>0.73/0.78</td>
<td>0.72/0.77</td>
</tr>
<tr>
<td>$BY_{rest.}$</td>
<td>0.73</td>
<td>0.77/0.77</td>
<td>0.68/0.72</td>
</tr>
<tr>
<td>(unrest.)</td>
<td>(0.81)</td>
<td>(0.78)/(0.78)</td>
<td>(0.68)/(0.72)</td>
</tr>
<tr>
<td>$PST$</td>
<td>0.35</td>
<td>0.81/0.76</td>
<td>0.65/0.67</td>
</tr>
<tr>
<td><strong>Panel B: 10 Size-Sorted Portfolios</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$CC$</td>
<td>0.19</td>
<td>0.82/0.85</td>
<td>0.82/0.85</td>
</tr>
<tr>
<td>$MSV$</td>
<td>0.12</td>
<td>0.83/0.86</td>
<td>0.83/0.86</td>
</tr>
<tr>
<td>$BY_{rest.}$</td>
<td>0.73</td>
<td>0.77/0.77</td>
<td>0.71/0.73</td>
</tr>
<tr>
<td>(unrest.)</td>
<td>(0.84)</td>
<td>(0.78)/(0.79)</td>
<td>(0.71)/(0.74)</td>
</tr>
<tr>
<td>$PST$</td>
<td>0.35</td>
<td>0.75/0.72</td>
<td>0.64/0.66</td>
</tr>
<tr>
<td><strong>Panel C: 10 BM-Sorted Portfolios</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$CC$</td>
<td>0.19</td>
<td>0.71/0.75</td>
<td>0.72/0.74</td>
</tr>
<tr>
<td>$MSV$</td>
<td>0.12</td>
<td>0.72/0.76</td>
<td>0.72/0.75</td>
</tr>
<tr>
<td>$BY_{rest.}$</td>
<td>0.73</td>
<td>0.67/0.60</td>
<td>0.59/0.58</td>
</tr>
<tr>
<td>(unrest.)</td>
<td>(0.83)</td>
<td>(0.67)/(0.60)</td>
<td>(0.59)/(0.58)</td>
</tr>
<tr>
<td>$PST$</td>
<td>0.35</td>
<td>0.64/0.23</td>
<td>0.50/0.33</td>
</tr>
<tr>
<td><strong>Panel D: 10 Momentum-Sorted Portfolios</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$CC$</td>
<td>0.19</td>
<td>0.55/0.63</td>
<td>0.58/0.61</td>
</tr>
<tr>
<td>$MSV$</td>
<td>0.12</td>
<td>0.55/0.62</td>
<td>0.57/0.61</td>
</tr>
<tr>
<td>$BY_{rest.}$</td>
<td>0.73</td>
<td>0.69/0.69</td>
<td>0.51/0.57</td>
</tr>
<tr>
<td>(unrest.)</td>
<td>(0.85)</td>
<td>(0.73)/(0.73)</td>
<td>(0.60)/(0.64)</td>
</tr>
<tr>
<td>$PST$</td>
<td>0.35</td>
<td>0.73/0.70</td>
<td>0.50/0.55</td>
</tr>
<tr>
<td><strong>Panel E: 10 Industry-Sorted Portfolios</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$CC$</td>
<td>0.19</td>
<td>0.49/0.53</td>
<td>0.49/0.53</td>
</tr>
<tr>
<td>$MSV$</td>
<td>0.12</td>
<td>0.50/0.54</td>
<td>0.50/0.55</td>
</tr>
<tr>
<td>$BY_{rest.}$</td>
<td>0.73</td>
<td>0.42/0.39</td>
<td>0.38/0.42</td>
</tr>
<tr>
<td>(unrest.)</td>
<td>(0.86)</td>
<td>(0.42)/(0.38)</td>
<td>(0.36)/(0.40)</td>
</tr>
<tr>
<td>$PST$</td>
<td>0.35</td>
<td>0.41/0.27</td>
<td>0.34/0.37</td>
</tr>
</tbody>
</table>

The table reports the correlations between the 3 Fama-French factors and the model-implied SDF, the filtered SDF, and the missing component of the filtered SDF using annual data over 1930-2009 and a different set of portfolios in each Panel. The acronyms $CC$, $MSV$, $BY$, and $PST$, denote respectively the models of Campbell and Cochrane (1999), Menzly, Santos, and Veronesi (2004), Bansal and Yaron (2004) and Piazzesi, Schneider, and Tuzel (2007).
0.54 and 0.59 when $\psi^*$ is estimated using Equations (6) and (4), respectively. This is reassuring for our methodology because, as is well known, the FF factors are successful in explaining a large fraction of the cross-sectional dispersion in asset returns. Moreover, Column 3 reveals that this high correlation is due almost entirely to the missing component, $\psi^*$, and not $m$ – the correlation between the filtered SDF and the FF factors is the same as that between the filtered missing component of the SDF and the FF factors. The results in Panels $B - E$ are largely similar – the filtered SDF and its missing component have high correlation with the FF factors for all the different sets of test assets, varying from 0.52 (0.52) for the momentum-sorted portfolios to 0.87 (0.89) for the size-sorted portfolios, and the high correlation is almost entirely due to the missing component $\psi^*$.

Row 2 in each panel shows that for the MSV model, the correlation between the model-implied SDF and the FF factors is small at 0.21. Also, the filtered SDF correlates strongly with the FF factors and this is almost entirely driven by the missing component of the SDF and not the consumption growth component.

Row 3 in each panel shows that for the BY model, the correlation between the model-implied SDF and the FF factors is 0.45 in the presence of the restrictions. This is more than double the correlations obtained for the CC and MSV models. Moreover, the correlation further doubles when the restrictions are not imposed varying from 0.87 – 0.92.

Finally, Row 4 in each panel shows that for the PST model, the correlation between the model-implied SDF and the FF factors is very small at 0.07. The filtered SDF, on the other hand, correlates strongly with the FF factors which is almost entirely driven by the missing component of the SDF and not the consumption growth component.

Table IV reveals that very similar results are obtained at the annual
frequency. Tables III and IV demonstrate the robustness of our estimation methodology – the filtered time series of the SDF and its missing component is quite robust to the choice of the utility curvature parameter $\gamma$ and the choice of the set of assets.

One thing to notice in Tables III and IV is that our filtered SDF and $\psi^*$ are consistently highly correlated with the FF factors independently from the sample frequency and the cross-section of assets used for the estimation (even assets, like the Industry portfolios, that are not well priced by the FF factors). This finding has two important implications. First, it suggests that our estimation approach successfully identifies the unobserved pricing kernel, since there is substantial empirical evidence that the FF factors do proxy for asset risk sources. Second, our finding provides a rationalization of the empirical success of the FF factors in pricing asset returns.

The reason behind the stable correlation results between our filtered SDFs and the three Fama French factors seems to be the fact that, independently from the set of assets used for the filtering, the most likely SDF tends to have a very similar time series behaviour. In particular, it shows a clear business cycle pattern, and significant and sharp reactions to stock market crashes (even if these crashes do not necessarily result in economy wide contractions). This feature of the filtered SDFs is illustrated in Figure 1 (annual frequency) and Figure 10 (quarterly frequency). In each figure we report the business cycle component (Panel A) and the residual component of the filtered $M^*$ for the different models.\textsuperscript{9} At the annual frequency (Figure 1), independently from the model considered, both the business cycle and residual components are extremely similar across the models. At the quarterly frequency, the components are once again very similar with the

\textsuperscript{9}The decomposition into a business cycle and a residual component is obtained by applying the Hodrick and Prescott (1997) filter to the estimated $M^*$.\textsuperscript{9}
slight exception of MSV that delivers an overall more volatile $M^*$. 

Figure 10: Business cycle (Panel A) and residual (Panel B) components of the most likely SDF ($M^*$) filtered using quarterly data over the period 1947:Q1-2009:Q4 for the different models considered: Bansal and Yaron (2004) (BY), Campbell and Cochrane (1999) (CC), Menzly, Santos, and Veronesi (2004) (MSV), and Piazzesi, Schneider, and Tuzel (2007) (PST). Shaded areas denote NBER recession years, and vertical dashed lines indicate the major stock market crashes identified by Mishkin and White (2002).

In Table V we compare the business cycle and market crash features of the filtered SDFs with the model implied ones. For each model considered, and for both the filtered ($M^*$) and model implied ($M$) pricing kernels, the table reports the risk neutral probabilities of recessions (Column 3), and stock market crashes non-concomitant with recessions (Column 4).

Focusing on quarterly data (Panel A), Column 3 shows that the filtered SDFs ($M^*$) imply a risk neutral probability of a recession in the 21.8%-24.6% range. Comparing this with the model implied probabilities reveals
Table V: Recession and Market Crash Probabilities of $M$ and $M^*$

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>CC $M$</td>
<td>25.2%</td>
<td>2.4%</td>
</tr>
<tr>
<td>$M^*$</td>
<td>23.5%</td>
<td>3.8%</td>
</tr>
<tr>
<td>BY $M$</td>
<td>55.3%</td>
<td>3.4%</td>
</tr>
<tr>
<td>$M^*$</td>
<td>24.6%</td>
<td>3.6%</td>
</tr>
<tr>
<td>MSV $M$</td>
<td>21.6%</td>
<td>2.4%</td>
</tr>
<tr>
<td>$M^*$</td>
<td>21.8%</td>
<td>6.9%</td>
</tr>
<tr>
<td>PST $M$</td>
<td>19.6%</td>
<td>2.8%</td>
</tr>
<tr>
<td>$M^*$</td>
<td>23.0%</td>
<td>3.5%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: Annual Data, 1929-2009</th>
</tr>
</thead>
<tbody>
<tr>
<td>CC $M$</td>
</tr>
<tr>
<td>$M^*$</td>
</tr>
<tr>
<td>BY $M$</td>
</tr>
<tr>
<td>$M^*$</td>
</tr>
<tr>
<td>MSV $M$</td>
</tr>
<tr>
<td>$M^*$</td>
</tr>
<tr>
<td>PST $M$</td>
</tr>
<tr>
<td>$M^*$</td>
</tr>
</tbody>
</table>

The table reports probability of recessions and stock market crashes non-concomitant with recessions implied by the model ($M$) and filtered ($M^*$) SDFs at quarterly (Panel A) and annual (Panel B) frequencies

that, with the exception of BY, all the other model-implied pricing kernels imply a similar probability of recessions. However, Column 4 shows that these kernels, with the exception of BY, fail to show the significant and sharp reaction to stock market crashes exhibited by the filtered SDFs: the probabilities of stock market crashes non-concomitant with recessions implied by the filtered SDFs are between 25% and 187% higher than those implied by the model specific kernels. Panel B reports similar findings at the annual frequency (but, in this case, MSV and PST imply significantly smaller recession probabilities than both $M^*$ and the BY and CC models).

The above results suggest that the explanatory power of these models for asset pricing would be improved by augmenting the pricing kernels with

\[^{10}\]Note that at the annual frequency, a year is designated as a recession year if at least one of its quarters is in an NBER recession period.
a component that exhibits sharp reactions to market crashes that are not perfectly correlated with the business cycle.

IV.2.2 Entropy Bounds Analysis

Our second approach to assess the empirical plausibility of the asset pricing models considered relies on the entropy bounds derived in Section II.1. For each model we compute the minimum values of the power coefficient, $\gamma$, at which the model-implied SDF satisfies the $HJ$, $Q$, $M$, and $\Psi$ bounds. We also compute the 95% confidence bands for these bounds from 1000 simulations of the same length as the historical time series. Table VI reports the results at the quarterly frequency. Panels $A$, $B$, $C$, $D$, $E$, and $F$ report results when the set of assets used in the construction of the bounds include the market, 25 Fama-French, 10 size-sorted, 10 book-to-market-equity-sorted, 10 momentum-sorted, and 10 industry-sorted portfolios, respectively. Consider first the results for the $HJ$, $Q_1$, $M_1$, and $\Psi_1$ bounds. The first row in each panel presents the bounds for the CC model. Panel $A$ shows that when the excess return on the market portfolio is used in the construction of the bounds, the minimum value of $\gamma$ at which the pricing kernel satisfies the $HJ$, $Q_1$, $M_1$, and $\Psi_1$ bounds is 1.4 in all four cases. However, when the set of test assets consists of the excess returns on the 25 Fama-French portfolios, Panel $B$ shows that the $HJ$, $Q_1$, $M_1$, and $\Psi_1$ bounds are satisfied for a minimum value of $\gamma = 7.3, 9.8, 9.9$, and $13.9$, respectively. Therefore, as suggested by the theoretical predictions, the $Q$-bound is tighter than the $HJ$-bound, and the $M$-bound is tighter than the $Q$-bound. Note that in this model, the coefficient of risk aversion is $\frac{\gamma}{S_t}$, where $S_t$ is the surplus consumption ratio. For $\gamma = 2$, the calibrated value in CC, the risk aversion varies over $[20, \infty)$. Panel $B$ reveals that the $Q$-bound is satisfied for $\gamma \geq 9.8$, implying that the risk aversion varies over $[43.9, \infty)$, the $M$-bound
is satisfied for $\gamma \geq 9.9$, implying that the risk aversion varies over $[44.2, \infty)$, and the $\Psi$-bound is satisfied for $\gamma \geq 13.9$, implying that the risk aversion varies over $[52.5, \infty)$. A similar ordering of the bounds is obtained when the set of assets consists of the 10 size-sorted, 10 book-to-market-equity-sorted, 10 momentum-sorted, and 10 industry-sorted portfolios in Panels C, D, E, and F, respectively. Also, very similar results are obtained for the $Q2$, $M2$, and $\Psi2$ bounds, stressing the robustness of our methodology.

The second row in each panel presents the bounds for the MSV model. When the set of test assets consists of the excess return on the market portfolio, the $HJ$, $Q1$, $M1$, and $\Psi1$ bounds are satisfied for a minimum value of $\gamma = 11.4$, 11.2, 12.4, and 15.7, respectively. For the 25 Fama-French portfolios, the bounds are much higher at 27.8, 31.7, 33.9, and 53.3, respectively. Therefore, this model requires very high values of the local curvature of the utility function to explain the equity premium and the cross-section of asset returns. In fact, this model requires much higher levels of risk aversion compared to the CC model for each of the set of test assets. As in the case of the CC model, very similar results are obtained for the $Q2$, $M2$, and $\Psi2$ bounds.

The third row in each panel presents the bounds for the BY model. Panel A shows that when the excess return on the market portfolio is used in the construction of the bounds, the minimum value of $\gamma$ at which the pricing kernel satisfies the $HJ$, $Q1$, $M1$, and $\Psi1$ bounds is 3.0 in all four cases. When the set of test assets consists of the excess returns on the 25 Fama-French portfolios, Panel B shows that the $HJ$ bound is satisfied for a minimum value of $\gamma = 4.0$ while the $Q1$, $M1$, and $\Psi1$ bounds are satisfied for a minimum value of $\gamma = 5.0$. Similar results are obtained for the other sets of portfolios and for the $Q2$, $M2$, and $\Psi2$ bounds. In this model, $\gamma$ represents the coefficient of relative risk aversion. Therefore, the results in
Panels $A - F$ reveal that the model-implied pricing kernel satisfies the $HJ$, $Q$, $M$, and $\Psi$ bounds for reasonable values of the risk aversion coefficient for all sets of test assets.

Finally, the fourth row in each panel presents the bounds for the PST model. When the set of test assets consists of the excess return on the market portfolio, the $HJ$, $Q_1$ ($Q_2$), $M_1$ ($M_2$), and $\Psi_1$ ($\Psi_2$) bounds are satisfied for a minimum value of $\gamma = 19.2$, 19.2 (19.4), 24.4 (24.3), and 16.2 (16.5), respectively. For the 25 Fama-French portfolios, the bounds are much higher at 64.3, 75.0 (74.5), 87.2 (83.8), and 70.9 (72.5), respectively. Therefore, this model requires very high levels of risk aversion to explain the equity premium and the cross-section of asset returns.

Overall, Table VI demonstrates that, in line with the theoretical underpinnings of the various bounds, the $Q$-bound is generally tighter than the $HJ$-bound because it naturally exploits the restriction that the SDF is a strictly positive random variable. The $M$-bound is tighter than the $Q$-bound because it formally takes into account the ability of the SDF to price assets. This relative ordering holds for a variety of different dynamic asset pricing models. Furthermore, the results suggest that while the external habit models of CC and MSV, as well as the housing model of PST require high levels of risk aversion to satisfy the bounds, the long run risks model of BY satisfies the bounds for reasonable levels of risk aversion for all the sets of test assets.

Table VII reports analogous bounds as in Table VI at the annual frequency. The table shows that, at the annual frequency, the $HJ$, $Q$, $M$, and $\Psi$ bounds are satisfied for much smaller values of the utility curvature parameter, $\gamma$, for each of the models considered and for each set of test assets. There is also less dispersion between the bounds compared to the quarterly data in Table VI. However, in line with the theoretical predictions, the $Q$-
bound is tighter than the $HJ$-bound, and the $M$-bound is tighter than the $Q$-bound.

Note that the above bound results have tight confidence bands and are much more consistent, in evaluating the plausibility of a given model across different sets of assets, than the cross-sectional $R^2$ measures reported in Tables I and II that vary wildly for the same model and have very wide confidence intervals.

V Conclusion

In this paper, we propose an information-theoretic approach to assess the empirical plausibility of candidate SDFs for a large class of dynamic asset pricing models. The models we consider are characterized by having a pricing kernel that can be factorized into an observable component, consisting in general of a parametric function of consumption growth, and a potentially unobservable one that is model-specific.

Based on this decomposition of the pricing kernel, we provide three major contributions.

First, using a relative entropy minimization approach, we show how to extract non-parametrically the time series of both the SDF and its unobservable component. Given the data, this methodology identifies the most likely – in the information theoretic sense – time series of the SDF and its unobservable component. Applying this methodology to the data we find that the estimated SDF has a clear business cycle pattern, but also shows significant and sharp reactions to financial market crashes that do not result in economy wide contractions. Moreover, we find that the non-parametrically extracted SDF, independently from the set of assets used for its construction, is highly correlated with the risk factors proposed in Fama and French (1993). This provides a rationalization of the empirical success
of the Fama French factors in pricing asset returns, and suggests that our filtering procedure does successfully identify the unobserved component of the SDF.

Second, we construct a new set of entropy bounds that build upon and improve the ones suggested in the previous literature in that a) they naturally impose the non negativity of the pricing kernel, b) they are generally tighter and have higher information content, and c) allow to utilize jointly the information contained in consumption data and a large cross-section of asset returns.

Third, applying the methodology developed in this paper to a large class of dynamic asset pricing models, we find that the external habit models of Campbell and Cochrane (1999) and Menzly, Santos, and Veronesi (2004), as well as the housing model of Piazzesi, Schneider, and Tuzel (2007), require very high levels of risk aversion to satisfy the bounds, while the long run risks model of Bansal and Yaron (2004) satisfies the bounds for low levels of risk aversion. These results are robust to the choice of test assets used in the construction of the bounds as well as the frequency of the data. Moreover, comparing the non-parametrically extracted SDF with those implied by the above asset pricing models, we again find empirical support for the long run risks framework.

The methodology developed in this paper is considerably general, and may be applied to any model that delivers well-defined Euler equations like models with heterogenous agents, limited stock market participation, and fragile beliefs, as long as the SDF can be factorized into an observable component and a potentially unobservable one.

References


A Appendix

A.1 Maximum Likelihood Analogy

The approaches in Equations (4) and (6) deliver maximum likelihood estimates of the \( \psi_t \) component of the pricing kernel. To formally see the analogy between our approach and an MLE procedure, let’s consider the two entropy minimization problems separately.

First, note that normalizing \( \{ \psi_t \}_{t=1}^T \) to lie in the unit simplex \( \Delta^{T-1} \)

\[
\Delta^{T-1} = \left\{ (\psi_1, \psi_2, ..., \psi_T) : \psi_t \geq 0, \sum_{t=1}^T \psi_t = 1 \right\},
\]

the solution of the estimation problem in Equation (6) also solves the following optimization

\[
\left\{ \tilde{\psi}_t \right\}_{t=1}^T \equiv \arg \max \frac{1}{T} \sum_{t=1}^T \ln \psi_t, \text{ s.t. } \{ \psi_t \}_{t=1}^T \in \Delta^{T-1}, \sum_{t=1}^T m(\theta, t) R^c_t \psi_t = 0.
\]

But the objective function above is simply the non parametric log likelihood (aka empirical likelihood) of Owen (1988, 1991, 2001) maximized under the asset pricing restrictions for a vector of asset returns.

Second, to see why the estimation problem in Equation (4) also delivers a maximum likelihood estimate of the \( \psi_t \) component, consider the following procedure for constructing (up to a scale) the series \( \{ \psi_t \}_{t=1}^T \). First, given an integer \( N >> 0 \), distribute to the various points in time \( t = 1, ..., T \), at random and with equal probabilities, the value \( 1/N \) in \( N \) independent draws. That is, draw a series of values (probability weights) \( \{ \tilde{\psi}_t \}_{t=1}^T \) given by

\[
\tilde{\psi}_t = \frac{n_t}{N},
\]

where \( n_t \) measures the number of times that the value \( 1/N \) has been assigned to time \( t \). Second, check whether the drawn series \( \{ \tilde{\psi}_t \}_{t=1}^T \) satisfies the pricing restriction \( \sum_{t=1}^T m(\theta, t) R^c_t \tilde{\psi}_t = 0 \). If it does, use this series as the estimator of \( \{ \psi_t \}_{t=1}^T \), and if it doesn’t draw another series. Obviously, a more efficient way of finding an estimate for \( \psi_t \) would be to choose the most likely outcome of the above procedure. Noticing that the distribution of the \( \tilde{\psi}_t \) is, by construction, a multinomial distribution with support given by the data sample, we have that the likelihood of any particular sequence \( \{ \tilde{\psi}_t \}_{t=1}^T \) is

\[
L \left( \left\{ \tilde{\psi}_t \right\}_{t=1}^T \right) = \frac{N!}{n_1! n_2! ... n_T!} \times T^{-N} = \frac{N!}{N \tilde{\psi}_1! N \tilde{\psi}_2! ... N \tilde{\psi}_T!} \times T^{-N}.
\]
Therefore, the most likely value of $\left\{ \tilde{\psi}_t \right\}_{t=1}^T$ maximizes the logarithmic likelihood

$$\ln L \left( \left\{ \tilde{\psi}_t \right\}_{t=1}^T \right) \propto \frac{1}{N} \left( \ln N! - \sum_{t=1}^T \ln \left( N \tilde{\psi}_t! \right) \right).$$

Since the above procedure of assigning probability weights will become more and more accurate as $N$ grows bigger, we would ideally like to have $N \to \infty$. But in this case one can show\(^{11}\) that

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

Therefore, taking into account the constraint for the pricing kernel, the maximum likelihood estimate (MLE) of the time series of $\tilde{\psi}_t$ would solve

$$\left\{ \tilde{\psi}_t \right\}_{t=1}^T \equiv \arg \max_{\left\{ \tilde{\psi}_t \right\}_{t=1}^T} -\sum_{t=1}^T \tilde{\psi}_t \ln \tilde{\psi}_t, \quad \text{s.t.} \quad \left\{ \tilde{\psi}_t \right\}_{t=1}^T \in \Delta^T, \sum_{t=1}^T m(\theta, t) R_{\theta}^T \tilde{\psi}_t = 0.$$

But the solution of the above MLE problem is also the solution of the relative entropy minimization problem in Equation (4) (see e.g. Csiszar (1975)). That is, the KIC minimization is equivalent to maximizing the likelihood in an unbiased procedure for finding the $\tilde{\psi}_t$ component of the pricing kernel.

### A.2 Additional Bounds and Derivations

**Definition 3 (Canonical HJ-bound)** for each $E[M_t] = \bar{M}$, the Hansen and Jagannathan (1991) minimum variance SDF is

$$M^*_t(\bar{M}) = \arg \min_{\{M_t(\bar{M})\}_{t=1}^T} \sqrt{Var \left( M_t(\bar{M}) \right)} \quad \text{s.t.} \quad 0 = E \left[ R_t^c M_t(\bar{M}) \right] \quad (29)$$

and any candidate stochastic discount factor $M_t$ must satisfy $Var(M_t) \geq Var \left( M^*_t(\bar{M}) \right)$.

The solution of the problem in Equation (29) is

$$M^*_t(\bar{M}) = \bar{M} + (R_t^c - E [R_t])' \beta_{\bar{M}},$$

where $\beta_{\bar{M}} = Cov (R_t^c)^{-1} \left( -\bar{M} E [R_t^c] \right)$.

\(^{11}\)Recall that from Stirling’s formula we have:

$$\lim_{N \tilde{\psi}_1 \to \infty} \frac{N \tilde{\psi}_1!}{\sqrt{2\pi N \tilde{\psi}_1} \left( \frac{N \tilde{\psi}_1}{e} \right)^{N \tilde{\psi}_1}} = 1.$$
**Definition 4 (Q-bounds)** We define the following risk neutral probability bounds for any candidate stochastic discount factor $M_t$.

1. **Q1-bound:**
   
   $$D\left(P \mid \frac{M_t}{M}\right) \equiv \int -\ln \frac{M_t}{M} dP \geq D\left(P \mid Q^*\right)$$

   where $Q^*$ solves Equation (7).

2. **Q2-bound (Stutzer (1995))**:

   $$D\left(\frac{M_t}{M} \mid P\right) \equiv \int \frac{M_t}{M} \ln \frac{M_t}{M} dP \geq D\left(Q^* \mid P\right)$$

   where $Q^*$ solves Equation (5).

**Remark 1 (HJ-bounds as approximated Q-bounds).** Let $p$ and $q$ denote the densities of the state $x$ associated, respectively, with the physical, $P$, and the risk neutral, $Q$, probability measures. Assuming that:

(A.1) $q$ and $p$ are twice continuously differentiable;

and that there exists a $\mu_p < \infty$ and a $\mu_q < \infty$ such that:

(A.2) **(Existence of maxima)**

$$\frac{\partial \ln p}{\partial x} \bigg|_{x=\mu_p} = 0, \quad \frac{\partial \ln q}{\partial x} \bigg|_{x=\mu_q} = 0; \quad (30)$$

(A.3) **(Finite second moments)**

$$- \left[ \frac{\partial^2 \ln p}{\partial x^2} \bigg|_{x=\mu_p} \right]^{-1} \equiv \sigma^2_p < \infty, \quad - \left[ \frac{\partial^2 \ln q}{\partial x^2} \bigg|_{x=\mu_q} \right]^{-1} \equiv \sigma^2_q < \infty. \quad (31)$$

We have that, in the limit of the small time interval, a second order approximation of the Q-bounds yields:

$$D\left(P \mid \frac{M_t}{M}\right) \propto \text{Var} \left(M_t\right), \quad (32)$$

$$D\left(\frac{M_t}{M} \mid P\right) \propto \text{Var} \left(M_t\right). \quad (33)$$

---

12 For expositional simplicity, we focus on a scalar state variable, but the result is straightforward to extend to a vector state.

13 For the Q2 bound only, using the dual objective function of the entropy minimization problem, Stutzer (1995) provides a similar approximation result to the one in Equation (33) that is valid when the variance bound is sufficiently small. Moreover, for the case of Gaussian iid returns, Kitamura and Stutzer (2002) show that the approximation of the Q2 bound in Equation (33) is exact.
Proof of Remark 1. Denote by $p$ and $q$ the densities associated, respectively, with the physical probability measure $P$ and the risk neutral measure $Q$. We can then rewrite the $Q_1$ and $Q_2$ bounds, respectively, as

$$D \left( P \mid M_t \right) = \int \ln \frac{dP}{dQ} dP = \int p \ln \frac{p}{q} dx$$

(34)

and

$$D \left( \frac{M_t}{M} \mid P \right) = \int \frac{dQ}{dP} \ln \frac{dQ}{dP} dP = \int \ln \frac{dQ}{dP} dQ = \int q \ln \frac{q}{p} dx.$$  

(35)

Given conditions A.1-A.3, we have from a second order Taylor approximation that

$$\ln q \propto \frac{1}{2} \frac{\partial^2 \ln q}{\partial x^2} \bigg|_{x=\mu_q} (x - \mu_q)^2 \equiv -\frac{1}{2} \frac{(x - \mu_q)^2}{\sigma_q^2}$$

$$\ln p \propto \frac{1}{2} \frac{\partial^2 \ln p}{\partial x^2} \bigg|_{x=\mu_p} (x - \mu_p)^2 \equiv -\frac{1}{2} \frac{(x - \mu_p)^2}{\sigma_p^2}.$$  

That is, $q$ and $p$ are approximately (to a second order) Gaussian

$$q \approx N \left( \mu_q, \sigma_q^2 \right), \quad p \approx N \left( \mu_p, \sigma_p^2 \right).$$

Note also that in the limit of the small time interval, by the diffusion invariance principle, we have $\sigma_q^2 = \sigma_p^2 = \sigma^2$. Therefore, plugging the above approximation into Equation (34), we have that in the limit of the small time interval

$$\int p \ln \frac{p}{q} dx \approx \int \left[ -\frac{1}{2} \frac{(x - \mu_p)^2}{\sigma^2} + \frac{1}{2} \frac{(x - \mu_q)^2}{\sigma^2} \right] p dx$$

$$= \frac{1}{2\sigma^2} \left[ -\sigma^2 + \int (x - \mu_q)^2 p dx \right]$$

$$= \frac{1}{2\sigma^2} \left\{ -\sigma^2 + \int \left[ (x - \mu_p)^2 + (\mu_p - \mu_q)^2 \right. \right.$$  

$$+ 2(\mu_p - \mu_q)(x - \mu_p) \left. \right] p dx \right\}$$

$$= \frac{1}{2\sigma^2} \left( \mu_p - \mu_q \right)^2 = \frac{1}{2\sigma^2} \sigma^2 \sigma^2 \xi = \frac{1}{2} \sigma^2 \xi$$

where the density $\xi$ is a (strictly positive) martingale defined by $\xi = \frac{dQ}{dP}$, and the one to the last equality comes from the change of drift implied by the Girsanov’s Theorem (see e.g. Duffie (2005, Appendix D)).
Similarly, from Equation (35) we have

\[ \int q \ln \frac{q}{p} dq = \frac{1}{2} \sigma_{\xi}^2. \]

Since \( Q \) and \( P \) are equivalent measures, \( M_t \propto \xi_t \). Therefore, in the limit of the small time interval \( \text{Var}(M_t) \propto \sigma_{\xi}^2 \), implying

\[ D \left( \frac{P}{M} \right) \propto \text{Var} (M_t), \quad D \left( \frac{M_t}{P} \right) \propto \text{Var} (M_t). \]

\[ \]

**Definition 5 (Volatility bound for \( \psi_t \))** For each \( E[\psi_t] = \bar{\psi} \), the minimum variance \( \psi_t \) is

\[ \psi_t^* \left( \bar{\psi} \right) \equiv \arg \min_{\{\psi_t(\bar{\psi})\}^T_{t=1}} \sqrt{\text{Var} \left( \psi_t \left( \bar{\psi} \right) \right)} \quad \text{s.t.} \quad 0 = \mathbb{E} \left[ R_t^\theta m (\theta, t) \psi_t \left( \bar{\psi} \right) \right] \]

and any candidate SDF must satisfy the condition \( \text{Var} (\psi_t) \geq \text{Var} \left( \psi_t^* \left( \bar{\psi} \right) \right) \).

The solution of the above minimization for a given \( \theta \) is

\[ \psi_t^* \left( \bar{\psi} \right) = \bar{\psi} + \left( R_t^\theta m (\theta, t) - \mathbb{E} \left[ R_t^\theta m (\theta, t) \right] \right)' \beta_{\bar{\psi}} \]

where \( \beta_{\bar{\psi}} = \text{Var} \left( R_t^\theta m (\theta, t) \right)^{-1} \left( -\bar{\psi} \mathbb{E} \left[ R_t^\theta m (\theta, t) \right] \right) \) and the lower volatility bound is given by

\[ \sigma_{\psi^*} \equiv \sqrt{\text{Var} \left( \psi_t^* \left( \bar{\psi} \right) \right)} = \bar{\psi} \sqrt{\mathbb{E} \left[ R_t^\theta m (\theta, t) \right]' \text{Var} \left( R_t^\theta m (\theta, t) \right)^{-1} \mathbb{E} \left[ R_t^\theta m (\theta, t) \right]} \].

**A.3 Data Description**

At the quarterly frequency, we use 6 different sets of assets: i) the market portfolio, ii) the 25 Fama-French portfolios, iii) the 10 size-sorted portfolios, iv) the 10 book-to-market-equity-sorted portfolios, v) the 10 momentum-sorted portfolios, and vi) the 10 industry-sorted portfolios. At the annual frequency, we use the same sets of assets except the 25 Fama-French portfolio that are replaced by the 6 portfolios formed by sorting stocks on the basis of size and book-to-market-equity because of the small time series dimension available at the annual frequency.

Our proxy for the market return is the Center for Research in Security Prices (CRSP) value-weighted index of all stocks on the NYSE, AMEX, and NASDAQ. The proxy for the risk-free rate is the one-month Treasury Bill rate obtained from the CRSP files. The returns on all the portfolios are obtained from Kenneth French’s data library. Quarterly (annual) returns for the above assets are computed by compounding monthly returns within each quarter (year), and converted to real using the personal consumption
deflator. Excess returns on the assets are then computed by subtracting the risk free rate.

Finally, for each dynamic asset pricing model, the information bounds and the non-parametrically extracted and model-implied time series of the SDF depend on consumption data. For the standard Consumption-CAPM of Breeden (1979) and Rubinstein (1976), the external habit models of Campbell and Cochrane (1999) and Menzly, Santos, and Veronesi (2004), and the long-run risks model of Bansal and Yaron (2004), we use per capita real personal consumption expenditures on nondurable goods from the National Income and Product Accounts (NIPA). We make the standard “end-of-period” timing assumption that consumption during quarter $t$ takes place at the end of the quarter. For the housing model of Piazzesi, Schneider, and Tuzel (2007) aggregate consumption is measured as expenditures on non-durables and services excluding housing services.
### Table VI: Bounds for RRA, Quarterly Data 1947:Q2-2009:Q4

<table>
<thead>
<tr>
<th></th>
<th>HJ-Bound</th>
<th>Q1/Q2-Bounds</th>
<th>M1/M2-Bounds</th>
<th>Ψ1/Ψ2-Bounds</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Market Portfolio</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>CC</strong></td>
<td>1.4</td>
<td>1.4</td>
<td>1.4</td>
<td>1.4</td>
</tr>
<tr>
<td></td>
<td>[0.5, 3.5]</td>
<td>[0.5, 3.5]</td>
<td>[0.5, 3.5]</td>
<td>[0.5, 3.5]</td>
</tr>
<tr>
<td><strong>MSV</strong></td>
<td>11.4</td>
<td>11.2</td>
<td>12.4</td>
<td>15.7</td>
</tr>
<tr>
<td></td>
<td>[5.0, 17.5]</td>
<td>[5.0, 18.0]</td>
<td>[5.5, 19.0]</td>
<td>[7.5, 24.5]</td>
</tr>
<tr>
<td><strong>BY</strong></td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
</tr>
<tr>
<td></td>
<td>[3.0, 3.0]</td>
<td>[3.0, 3.0]</td>
<td>[3.0, 3.0]</td>
<td>[3.0, 3.0]</td>
</tr>
<tr>
<td><strong>PST</strong></td>
<td>19.2</td>
<td>19.2</td>
<td>24.4</td>
<td>16.2</td>
</tr>
<tr>
<td></td>
<td>[3.0, 38.0]</td>
<td>[3.0, 39.0]</td>
<td>[4.0, 50.0]</td>
<td>[3.0, 31.0]</td>
</tr>
</tbody>
</table>

| **Panel B: FF 25 Portfolios** |
| **CC**   | 7.3      | 9.8          | 9.9          | 13.9         |
|          | [9.5, 21.5] | [9.5, 21.5] | [9.5, 22.0] | [13.2, 43.1] |
| **MSV**  | 27.8     | 37.7         | 33.9         | 53.3         |
|          | [30.6, 47.0] | [30.5, 51.0] | [31.5, 56.0] | [49.6, 81.0] |
| **BY**   | 4.0      | 5.0          | 5.0          | 5.0          |
|          | [5.0, 5.0] | [5.0, 5.0]  | [5.0, 5.0]  | [5.0, 5.0]  |
| **PST**  | 64.3     | 75.0         | 87.2         | 70.9         |
|          | [72.5, 122.0] | [69.5, 122.0] | [78.5, 136.5] | [70.0, 125.5] |

| **Panel C: 10 Size Portfolios** |
| **CC**   | 1.8      | 2.0          | 2.0          | 2.1          |
|          | [1.0, 5.6] | [1.0, 6.5]  | [1.0, 5.6]  | [1.0, 5.7]  |
| **MSV**  | 13.6     | 14.6         | 18.5         | 19.2         |
|          | [11.0, 25.5] | [12.0, 25.5] | [12.6, 36.0] | [16.0, 36.5] |
| **BY**   | 3.0      | 4.0          | 4.0          | 4.0          |
|          | [4.0, 4.0] | [4.0, 4.0]  | [4.0, 4.0]  | [4.0, 4.0]  |
| **PST**  | 27.8     | 34.7         | 34.7         | 23.9         |
|          | [24.5, 58.5] | [24.5, 58.5] | [28.5, 73.5] | [22.0, 53.0] |

| **Panel D: 10 BM Portfolios** |
| **CC**   | 2.8      | 3.4          | 3.8          | 3.7          |
|          | [2.5, 8.0] | [2.5, 8.0]  | [3.0, 10.0] | [2.5, 10.0] |
| **MSV**  | 17.0     | 19.9         | 26.1         | 26.5         |
|          | [15.0, 28.0] | [15.0, 29.0] | [21.5, 40.0] | [21.5, 44.0] |
| **BY**   | 4.0      | 4.0          | 4.0          | 4.0          |
|          | [4.0, 4.0] | [4.0, 4.0]  | [4.0, 4.0]  | [4.0, 4.0]  |
| **PST**  | 33.3     | 43.5         | 41.5         | 31.2         |
|          | [30.6, 64.0] | [35.0, 79.0] | [34.0, 79.0] | [27.0, 58.0] |

| **Panel E: 10 Momentum Portfolios** |
| **CC**   | 5.4      | 6.9          | 6.9          | 8.6          |
|          | [5.0, 13.0] | [5.0, 13.5] | [5.0, 13.5] | [6.5, 18.0] |
| **MSV**  | 24.0     | 29.1         | 39.1         | 40.9         |
|          | [22.0, 36.0] | [23.5, 39.5] | [32.5, 57.5] | [34.0, 63.5] |
| **BY**   | 4.0      | 5.0          | 5.0          | 5.0          |
|          | [5.0, 4.0] | [5.0, 4.0]  | [5.0, 4.0]  | [5.0, 4.0]  |
| **PST**  | 51.7     | 71.7         | 72.0         | 48.1         |
|          | [48.5, 87.0] | [57.0, 113.0] | [42.0, 78.5] | [43.5, 82.5] |

| **Panel F: 10 Industry Portfolios** |
| **CC**   | 3.0      | 3.3          | 3.3          | 3.5          |
|          | [2.5, 5.7] | [2.5, 8.5]  | [2.5, 5.7]  | [2.5, 10.5] |
| **MSV**  | 17.6     | 19.1         | 26.4         | 27.6         |
|          | [14.0, 27.0] | [15.5, 29.0] | [21.5, 41.5] | [21.5, 44.0] |
| **BY**   | 4.0      | 4.0          | 4.0          | 4.0          |
|          | [4.0, 4.0] | [4.0, 4.0]  | [4.0, 4.0]  | [4.0, 4.0]  |
| **PST**  | 36.2     | 46.7         | 47.1         | 35.5         |
|          | [31.0, 65.5] | [37.5, 86.5] | [20.0, 63.5] | [30.5, 66.5] |

The table reports the values of the utility curvature parameter at which the model-implied SDF satisfies the HJ, Q, M, and Ψ bounds using quarterly data over 1947:2-2009:4 and a different set of portfolios in each Panel. The acronyms CC, MSV, BY and PST, denote respectively the models of Campbell and Cochrane (1999), Menzly, Santos, and Veronesi (2004), Bansal and Yaron (2004) and Piazzesi, Schneider, and Tuzel (2007).
Table VII: Bounds for RRA, Annual Data 1930-2009

<table>
<thead>
<tr>
<th></th>
<th>( HJ )-Bound</th>
<th>( Q1/Q2 )-Bounds</th>
<th>( M1/M2 )-Bounds</th>
<th>( \Psi1/\Psi2 )-Bounds</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Market Portfolio</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( CC )</td>
<td>0.1 / 0.1</td>
<td>[0.2, 0.7] / [0.3, 0.5]</td>
<td>0.1 / 0.1</td>
<td>0.1 / 0.1</td>
</tr>
<tr>
<td>( MSV )</td>
<td>0.1 / 0.1</td>
<td>[0.1, 0.1] / [0.1, 0.1]</td>
<td>0.1 / 0.1</td>
<td>0.1 / 0.1</td>
</tr>
<tr>
<td>( BY )</td>
<td>4.0</td>
<td>4.0/4.0</td>
<td>4.0/4.0</td>
<td>4.0/4.0</td>
</tr>
<tr>
<td>( PST )</td>
<td>8.6</td>
<td>9.9 / 9.5</td>
<td>14.9 / 15.0</td>
<td>6.8 / 6.8</td>
</tr>
</tbody>
</table>

| **Panel B: FF 6 Portfolios** | | | | |
| \( CC \) | 0.3 | 1.0 / 0.6 | 1.0 / 0.6 | 1.2 / 0.7 |
| \( MSV \) | 0.2 | 0.2 / 0.2 | 0.2 / 0.2 | 0.2 / 0.2 |
| \( BY \) | 5.0 | 5.0/5.0 | 5.0/5.0 | 5.0/5.0 |
| \( PST \) | 12.4 | 16.8 / 15.2 | 20.5 / 17.7 | 13.2 / 12.0 |

| **Panel C: 10 Size Portfolios** | | | | |
| \( CC \) | 0.1 | 0.5 / 0.3 | 0.5 / 0.3 | 0.5 / 0.3 |
| \( MSV \) | 0.2 | 0.2 / 0.2 | 0.2 / 0.2 | 0.2 / 0.2 |
| \( BY \) | 4.0 | 4.0/4.0 | 4.0/4.0 | 4.0/4.0 |
| \( PST \) | 10.4 | 13.6 / 12.2 | 15.7 / 13.7 | 11.5 / 10.3 |

| **Panel D: 10 BM Portfolios** | | | | |
| \( CC \) | 0.2 | 0.8 / 0.5 | 0.8 / 0.5 | 0.9 / 0.5 |
| \( MSV \) | 0.2 | 0.2 / 0.2 | 0.2 / 0.2 | 0.2 / 0.2 |
| \( BY \) | 5.0 | 5.0/5.0 | 5.0/5.0 | 5.0/5.0 |
| \( PST \) | 11.2 | 15.8 / 13.8 | 17.7 / 15.8 | 12.7 / 11.2 |

| **Panel E: 10 Momentum Portfolios** | | | | |
| \( CC \) | 0.4 | 1.4 / 0.9 | 1.4 / 0.9 | 1.5 / 1.0 |
| \( MSV \) | 0.2 | 0.2 / 0.2 | 0.2 / 0.2 | 0.2 / 0.2 |
| \( BY \) | 5.0 | 5.0/5.0 | 5.0/5.0 | 5.0/5.0 |
| \( PST \) | 14.3 | 18.3 / 16.9 | 21.1 / 18.5 | 13.9 / 12.9 |

| **Panel F: 10 Industry Portfolios** | | | | |
| \( CC \) | 0.4 | 1.7 / 1.0 | 1.7 / 1.0 | 2.2 / 1.2 |
| \( MSV \) | 0.2 | 0.2 / 0.2 | 0.2 / 0.2 | 0.2 / 0.2 |
| \( BY \) | 5.0 | 5.0/5.0 | 5.0/5.0 | 5.0/5.0 |
| \( PST \) | 14.1 | 19.7 / 17.4 | 22.0 / 18.9 | 16.5 / 14.2 |

The table reports the values of the utility curvature parameter at which the model-implied SDF satisfies the HJ, Q, M, and \( \Psi \) bounds using annual data over 1930-2009 and a different set of portfolios in each Panel. The acronyms \( CC \), \( MSV \), \( BY \) and \( PST \) denote respectively the models of Campbell and Cochrane (1999), Menzly, Santos, and Veronesi (2004), Bansal and Yaron (2004) and Piazzesi, Schneider, and Tuzel (2007).