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

By

Anisha Ghosh
Christian Julliard
Alex P. Taylor

October 2011

Anisha Ghosh is an Assistant Professor of Finance at Carnegie Mellon University's Tepper School of Business. Her research interests span dynamic asset pricing, financial econometrics, and macroeconomics. She has given invited talks at leading academic institutions and her research has been published in top journals. She was awarded a Ph.D. by the Department of Economics at the London School of Economics. Christian Julliard is an Assistant Professor/Lecturer in the Department of Finance and a senior research associate of the Financial Market Group (FMG), at the London School of Economics. He is also a research affiliate of the International Macroeconomics and Financial Economics programmes of the Centre for Economic Policy Research (CEPR), and an associated editor of Economica. His research has been published in top economics and finance journal. He was awarded a Ph.D. by the Department of Economics at Princeton University where he was also affiliated with the Bendheim Center for Finance and the Woodrow Wilson School of Public and International Affairs. His research interests span macroeconomics, finance, and the frontier of applied econometrics. Alex Taylor is a lecturer in finance at the Manchester Accounting and Finance Group, University of Manchester. He was an undergraduate at St Peters College, University of Oxford, and obtained his doctorate at Imperial College, University of London. His research interests include asset pricing, predictability of market returns, and credit risk. Any opinions expressed here are those of the authors and not necessarily those of the FMG. The research findings reported in this paper are the result of the independent research of the authors and do not necessarily reflect the views of the LSE.

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

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

August 25, 2011

Abstract

We study a broad class of asset pricing models in which the stochastic discount factor (SDF) can be factorized into an observable component and a potentially unobservable, model-specific, one. Exploiting this decomposition we derive new entropy bounds that restrict the admissible regions for the SDF and its components. Without using this decomposition, to a second order approximation, entropy bounds are equivalent to the canonical Hansen-Jagannathan bounds. However, bounds based on our decomposition have higher information content, are tighter, and exploit the restriction that the SDF is a positive random variable. Our information-theoretic framework also enables us to extract a non-parametric estimate of the unobservable component of the SDF. Empirically, we find it to have a business cycle pattern, and significant correlations with both financial market crashes unrelated to economy-wide contractions, and the Fama-French factors. We apply our methodology to some leading consumption-based models, gaining new insights about their empirical performance.

Keywords: Pricing Kernel, Stochastic Discount Factor, Consumption Based Asset Pricing, Entropy Bounds.
JEL Classification Codes: G11, G12, G13, C52.

[†]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.

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 risk 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) and 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)).

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)). 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 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 models with housing risk where the additional component consists of 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)) and ambiguity 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 way to analyze dynamic asset pricing models for which the SDF can be factorized into an observable component and a potentially unobservable one. Our analysis utilizes an information-theoretic entropy approach to assess the empirical plausibility of candidate SDFs of this form, and provides the most likely estimate of the time series of the unobserved pricing kernel.

First, we construct entropy bounds that restrict the admissible regions for the SDF and its unobservable component. Dynamic equilibrium asset pricing models generally impose strong assumptions on the preferences of consumers and the dynamics of the state variables driving asset prices in order to identify the SDF. In contrast, we rely on a model-free no-arbitrage approach to construct bounds on the SDF and its components. 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) have 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 still possible), and we can accommodate an asset space of arbitrary dimension. Our methodology can also be used to construct bounds (Ψ-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 Ψ-bounds are the tightest ones, and can be satisfied if and only if the model is actually

3

able to price assets correctly.

Second, we show how the relative entropy minimization approach used for the construction of the bounds can be used to extract non-parametrically the time series of both the SDF and its unobservable component. This methodology identifies the *most likely*, in a 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), and Rosenberg and Engle (2002)). Compared to this literature, our nonparametric approach offers two 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, therefore connecting our work to macro-finance modeling. Empirically, 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, our estimated time series for the unobservable pricing kernel is highly correlated with the Fama and French (1993) factors, independently from the sample frequency and the set of 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 French factors.

Third, we apply our methodology to some of the leading 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 Ψ-bound, which is a bound on the unobservable component of the pricing kernel, 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

4

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).

Finally, the methodology developed in this paper has considerable 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 heterogenous agents, limited stock market participation, and fragile beliefs.

The remainder of the paper is organized as follows. Section II presents the information-theoretic methodology and Section II.1 introduces 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. The models considered, and their mapping into our framework, are presented in Section IV.1 while the empirical results are presented in Section IV.2. 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 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

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

5

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

$$M_t = m(\theta, t) \times \psi_t \tag{2}$$

where $m(\theta, t)$ is a known non-negative function of data observable at time t and the parameters vector $\theta \in \mathbb{R}^k$, and ψ_t is a potentially unobservable component. In the most common case $m(\theta, t)$ is simply a function of consumption growth, i.e. $m(\theta, t) = m(\Delta c_t, \theta)$ 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

$$\mathbf{0} = \mathbb{E}\left[m(\theta, t)\psi_t \mathbf{R}_t^e\right] \equiv \int m(\theta, t)\psi_t \mathbf{R}_t^e dP \tag{3}$$

where $\mathbb{E}$ is the unconditional rational expectation operator, $\mathbf{R}_t^e \in \mathbb{R}^N$ 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

$$\mathbf{0} = \int m(\theta, t)\frac{\psi_t}{\bar{\psi}}\mathbf{R}_t^e \, dP = \int m(\theta, t)\mathbf{R}_t^e \, d\Psi \equiv \mathbb{E}^\Psi\left[m(\theta, t)\mathbf{R}_t^e\right]$$

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

The transformation above implies that, given a set of consumption and asset returns data, for any θ we can estimate the Ψ probability measure as

$$\hat{\Psi} \equiv \arg\min_{\Psi} D(\Psi||P) \equiv \arg\min_{\Psi} \int \frac{d\Psi}{dP}\ln\frac{d\Psi}{dP}dP \quad \text{s.t.} \quad \mathbf{0} = \mathbb{E}^\Psi\left[m(\theta, t)\mathbf{R}_t^e\right]. \tag{4}$$

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 Ψ 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 Ψ is identical to P, that is when all the information needed to price assets is contained in $m(\theta, t)$ and ψ_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

$$\hat{Q} \equiv \underset{Q}{\arg\min}\, D(Q||P) \equiv \underset{\Psi}{\arg\min} \int \frac{dQ}{dP} \ln \frac{dQ}{dP} dP \;\; \text{s.t.} \;\; \mathbf{0} = \int \mathbf{R}_t^e dQ \equiv \mathbb{E}^Q[\mathbf{R}_t^e] \tag{5}$$

under the restriction that Q and P are absolutely continuous.

Moreover, since relative entropy is not symmetric, we can also recover Ψ and Q as

$$\hat{\Psi} \;\; \equiv \;\; \underset{\Psi}{\arg\min}\, D(P||\Psi) \equiv \underset{\Psi}{\arg\min} \int \ln \frac{dP}{d\Psi} dP \;\; \text{s.t.} \;\; \mathbf{0} = \mathbb{E}^\Psi[m(\theta, t)\mathbf{R}_t^e] \tag{6}$$

$$\hat{Q} \;\; \equiv \;\; \underset{Q}{\arg\min}\, D(P||Q) \equiv \underset{Q}{\arg\min} \int \ln \frac{dP}{dQ} dP \;\; \text{s.t.} \;\; \mathbf{0} = \mathbb{E}^Q[\mathbf{R}_t^e] \tag{7}$$

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 Ψ 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 ψ_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 ψ_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 Jagan-nathan (1991).[1]

Third, our approach to uncover the ψ_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 $\mathbf{Z}_{t-1}$, 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 $\mathbf{Z}_{t-1}$.

Fifth, there is no ex-ante restriction on the number of assets that can be used in constructing ψ_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.

Finally, this approach is numerically simple when implemented via duality (see e.g. Csiszar (1975)). That is, when implementing the entropy minimiza-tion in Equation (4) each element of the series $\{\psi_t\}_{t=1}^T$ can be estimated, up to a positive constant scale factor, as

$$\hat{\psi}_t = \frac{e^{\lambda(\theta)' m(\theta,t) \mathbf{R}_t^e}}{\sum_{t=1}^T e^{\lambda(\theta)' m(\theta,t) \mathbf{R}_t^e}}, \quad \forall t \tag{8}$$

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

$$\lambda(\theta) \equiv \arg\min_\lambda \frac{1}{T} \sum_{t=1}^T e^{\lambda' m(\theta,t) \mathbf{R}_t^e}, \tag{9}$$

---

[1] Hansen and Jagannathan (1991) offer an alternative bound that imposes this restric-tion, but it is computationally cumbersome (the minimum variance portfolio is basically an option) and generally not applied.

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 ψ_t, up to a constant positive scale factor, as being equal to

$$\hat{\psi}_t = \frac{1}{T(1 + \lambda(\theta)'m\,(\theta, t)\,\mathbf{R}_t^e)}, \quad \forall t \tag{10}$$

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

$$\lambda(\theta) \equiv \arg\min_{\lambda} -\sum_{t=1}^{T} \log(1 + \lambda'm\,(\theta, t)\,\mathbf{R}_t^e), \tag{11}$$

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) λ – 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 Extremum 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 ψ 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, $M_t^*\left(\bar{M}\right)$, 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 $M_t^*\left(\bar{M}\right)$ is a linear function of returns, it does not generally satisfy the restriction.[2]

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,[3] 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 Gaussian distributions is proportional to their variances; 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

---

[2] We call the bound in Definition 3 the "canonical" HJ-bound since Hansen and Jagannathan (1991) also provide an alternative bound, that imposes the non-negativity of the pricing kernel, but that is not generally used due to its computational complexity.

[3] The (sufficient, but not necessary) regularity conditions required for the approximation result stated above are typically satisfied in consumption-based asset pricing models.

structure of the pricing kernel. Below 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 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 θ, and ψ_t is a potentially unobservable component. A large class of equilibrium asset pricing models including ones with standard 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 θ, we define the following bounds:*

1. *$M1$-bound:*

$$
D\left(P||\frac{M_t}{\bar{M}}\right) \equiv \int -\ln\frac{M_t}{\bar{M}}dP \;\geqslant\; D\left(P||\frac{m(\theta,t)\psi_t^*}{\overline{m(\theta,t)\psi_t^*}}\right)
$$

$$
\equiv \int -\ln\frac{m(\theta,t)\psi_t^*}{\overline{m(\theta,t)\psi_t^*}}dP
$$

   where ψ_t^* solves Equation (6) and $\overline{m(\theta,t)\psi_t^*} \equiv \mathbb{E}\left[m(\theta,t)\psi_t^*\right].$

2. *$M2$-bound:*

$$
D\left(\frac{M_t}{\bar{M}}||P\right) \equiv \int \frac{M_t}{\bar{M}}\ln\frac{M_t}{\bar{M}}dP \;\geqslant\; D\left(\frac{m(\theta,t)\psi_t^*}{\overline{m(\theta,t)\psi_t^*}}||P\right)
$$

$$
\equiv \int \frac{m(\theta,t)\psi_t^*}{\overline{m(\theta,t)\psi_t^*}}\ln\frac{m(\theta,t)\psi_t^*}{\overline{m(\theta,t)\psi_t^*}}dP
$$

   where ψ_t^* solves Equation (4).

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

$$D\left(P||\frac{m\left(\theta,t\right)\psi_t^*}{\overline{m\left(\theta,t\right)\psi_t^*}}\right) \geq D\left(P||Q^*\right) \text{ and } D\left(\frac{m\left(\theta,t\right)\psi_t^*}{\overline{m\left(\theta,t\right)\psi_t^*}}||P\right) \geq D\left(Q^*||P\right)$$

by construction,[4] 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 $m\left(\theta,t\right)$.

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

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

1. *Ψ1-bound:*

$$D\left(P||\frac{\psi_t}{\overline{\psi}}\right) \equiv -\int \ln\frac{\psi_t}{\overline{\psi}}dP \geqslant D\left(P||\frac{\psi_t^*}{\overline{\psi}}\right)$$

   where ψ_t^* solves Equation (6);

2. *Ψ2-bound*

$$D\left(\frac{\psi_t}{\overline{\psi}}||P\right) \equiv \int \frac{\psi_t}{\overline{\psi}}\ln\frac{\psi_t}{\overline{\psi}}dP \geqslant D\left(\frac{\psi_t^*}{\overline{\psi}}||P\right)$$

   where ψ_t^* solves Equation (4).

Besides providing an additional check for any candidate SDF, the Ψ-bounds are useful in that a simple comparison of $D\left(\frac{\psi_t^*}{\overline{\psi}}||P\right)$, $D\left(\frac{m(\theta,t)}{\overline{m(\theta,t)}}||P\right)$ and $D\left(Q^*||P\right)$ can provide a very informative decomposition in terms of

---

[4]Cf. Definition 4 in Appendix A.2.

the entropy contribution to the pricing kernel, that is logically similar to the widely used variance decomposition analysis. For example, if $D\left(\frac{\psi_t^*}{\psi}||P\right)$ happens to be close to $D\left(Q^*||P\right)$, while $D\left(\frac{m(\theta,t)}{m(\theta,t)}||P\right)$ is substantially smaller, the decomposition would imply that most of the ability of the candidate SDF to price assets comes from the ψ_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 ψ_t component – the Ψ-bounds will offer a tight selection criterion since, under the null of the model being true, we should have $D\left(\frac{\psi_t^*}{\psi}||P\right) = D\left(P||\frac{\psi_t^*}{\psi}\right) = 0$ 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 θ both the Q-bounds and the HJ-bounds will be satisfied, but nevertheless the SDF grossly violates the pricing restrictions in the Euler Equations (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 ψ_t component. Such a bound, presented in Definition 5 of Appendix A.2, identifies a minimum variance $\psi_t^*\left(\bar\psi\right)$ component with standard deviation given by

$$\sigma_{\psi^*} = \bar\psi\sqrt{\mathbb{E}\left[\mathbf{R}_t^e m\left(\theta,t\right)\right]' Var\left(\mathbf{R}_t^e m\left(\theta,t\right)\right)^{-1}\mathbb{E}\left[\mathbf{R}_t^e m\left(\theta,t\right)\right]}. \qquad (12)$$

This bound, as the entropy based Ψ-bound in Definition 2, uses information about the structure of the SDF but, differently from the latter, does not constrains ψ_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 Ψ-bound.

Equation (12), viewed as a second order approximation to the entropy Ψ-bound, 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

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

where $Var\left(M_t^*\left(\bar{M}\right)\right)$ 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 ψ_t in Definition 5 is also satisfied since, from Equation (12), we have that under the assumption of constant ψ_t the bound can be satisfied only if $\mathbb{E}\left[\mathbf{R}_t^e m(\theta_0, t)\right] \equiv \mathbb{E}\left[\mathbf{R}_t^e M_t(\theta_0)\right] = \mathbf{0}$, that is only if the candidate SDF is able to price asset returns.

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

We first illustrate our methodology for the Consumption-CAPM (C-CAPM) of Breeden (1979) 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,

$$M_{t+1} = \delta \left(C_{t+1}/C_t\right)^{-\gamma}, \tag{13}$$

where δ denotes the subjective discount factor, γ 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) and 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 consump-

tion risk of the stock market if consumption is slow to respond to returns. 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:

$$M_{t+1}^S = \delta^{1+S} \left(C_{t+1+S}/C_t\right)^{-\gamma} R_{t+1,t+1+S}^f, \tag{14}$$

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 (14), unlike the one in Equation (13), 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, $\theta = \gamma$, $M_t = m\left(\theta, t\right) = \left(C_{t+1}/C_t\right)^{-\gamma}$, and $\psi_t = \delta$, a constant, for all t. For the ultimate consumption risk model, $\theta = \gamma$, $m\left(\theta, t\right) = \left(C_{t+1+S}/C_t\right)^{-\gamma}$, 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[5] 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 1, 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 γ and the HJ, $Q1$, $M1$, and $\Psi1$ bounds for the contemporaneous consumption risk model in Equation (13). The black curve

---

[5]See Appendix A.3 for a thorough data description.

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, ψ_t, is a constant hence it has zero relative entropy for all values of γ, as shown by the orange straight line with triangles. The blue curve with "+" signs and the yellow curve with inverted triangles show the relative entropy as a function of the risk aversion coefficient of the filtered SDF and its missing component, respectively. The model satisfies the HJ bound for very high values of $\gamma \geqslant 64$, as shown by the green dotted-dashed vertical line. It satisfies the $Q1$ bound for even higher values of $\gamma \geqslant 72$, as shown by the red dashed vertical line. The minimum value of γ at which the $M1$ bound is satisfied is given by the value corresponding to the intersection of the black and blue curves, i.e. it is the minimum value of γ 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 $\Psi1$ bound identifies the minimum value of γ 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 γ, the model fails to satisfy the $\Psi1$ bound for any value of γ.

Panel B shows that very similar results are obtained for the $Q2$, $M2$, and $\Psi2$ bounds. The $Q2$ and $M2$ bounds are satisfied for values of γ at least as large as 73 and 99, respectively, while the $\Psi2$ bound is not satisfied for any value of γ. 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 Ψ-bounds are tighter than the M-bounds.

Figure 2 presents analogous results to Figure 1 for the ultimate consumption risk model in Equation (14). Panel A shows that the HJ, $Q1$, and $M1$ bounds are satisfied for $\gamma \geqslant 22$, 23, 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 1, Panel A, for the contemporaneous consumption risk model. As for the latter model, the $\Psi1$ bound is not satisfied for any value of γ. Panel B shows that the $Q2$ and $M2$ bounds are satisfied for $\gamma \geqslant 24$ and 47, respectively, while the $\Psi2$ bound is not satisfied for any value of γ.
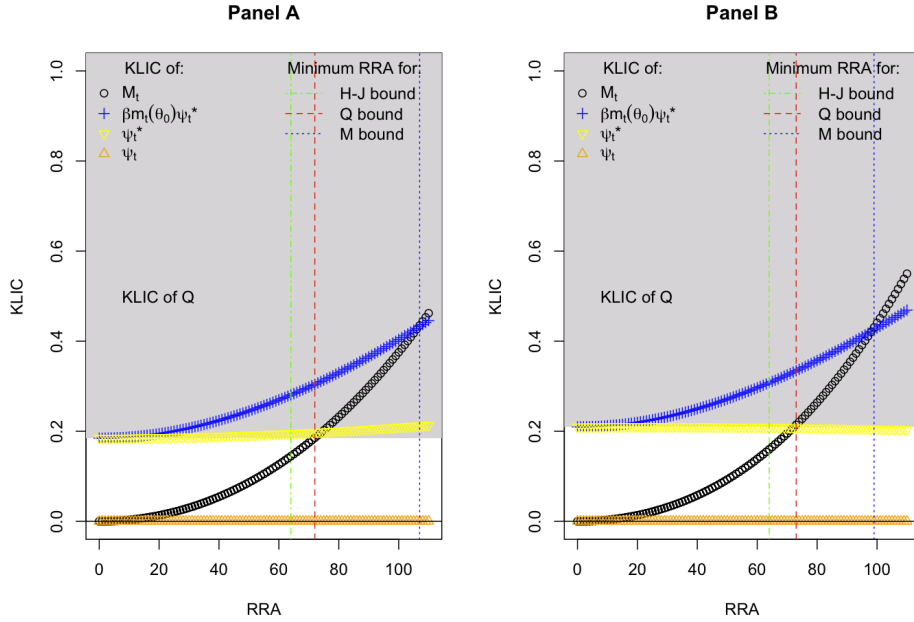
16

Figure 1: 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.

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 ($\hat{\gamma} = 1.5$), and at this value of the coefficient the model is able to explain about 60% of the cross-sectional variation across the 25 Fama-French portfolios, all the bounds reject the model for low RRA, and the Ψ bounds are not satisfied for any level of RRA. This stresses the power of the model evaluation approach proposed.

Figure 3, 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, ψ_t^*, estimated using Equation (6). The black solid line plots the filtered SDF, $M_t = (C_t/C_{t-1})^{-\gamma} \psi_t^*$. The grey shaded areas represent NBER-dated recessions while the green dashed vertical lines correspond to the major stock
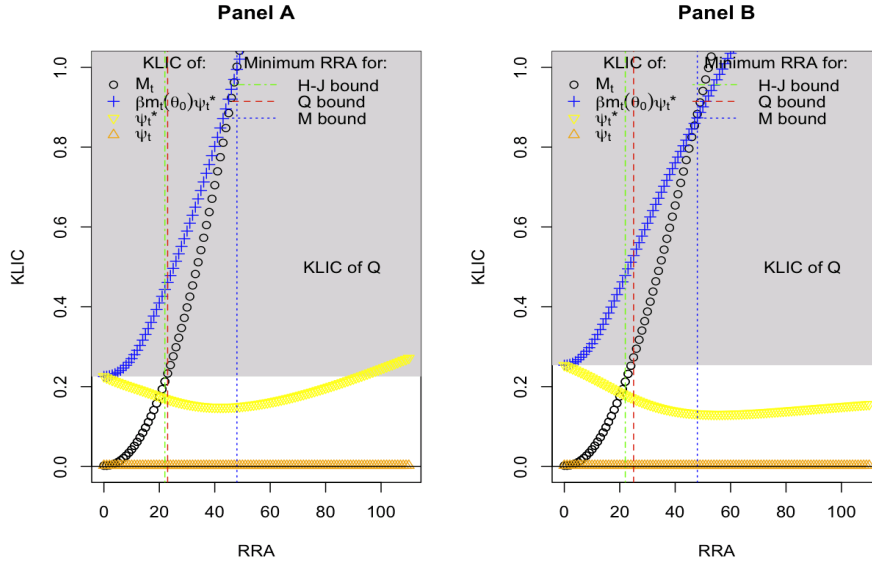
Figure 2: The figure plots the KLIC of the filtered and model-implied SDFs and their unobservable components as a function of the risk aversion coefficient and the entropy bounds for the ultimate consumption risk CCAPM.

market crashes identified in Mishkin and White (2002). 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, ψ, 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 $\gamma = 10$.

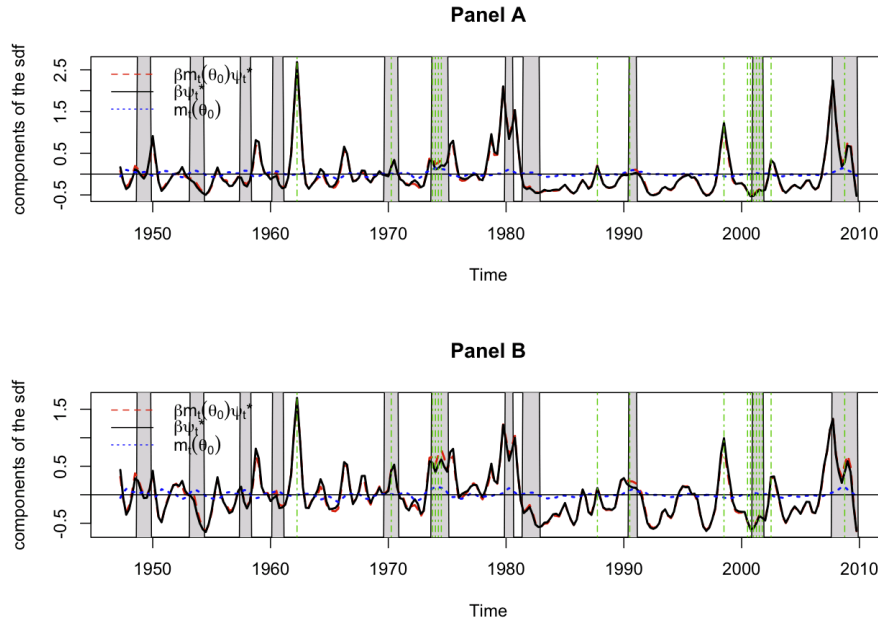Finally, Figure 4, Panel A plots the time series of the filtered SDF and

Figure 3: The figure plots the (demeaned) time series of the filtered SDF and its components for the standard CCAPM for γ=10. Shaded areas are NBER recession periods. Vertical dashed lines are the stock market crashes identified by Mishkin and White (2002).

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, differently 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

19
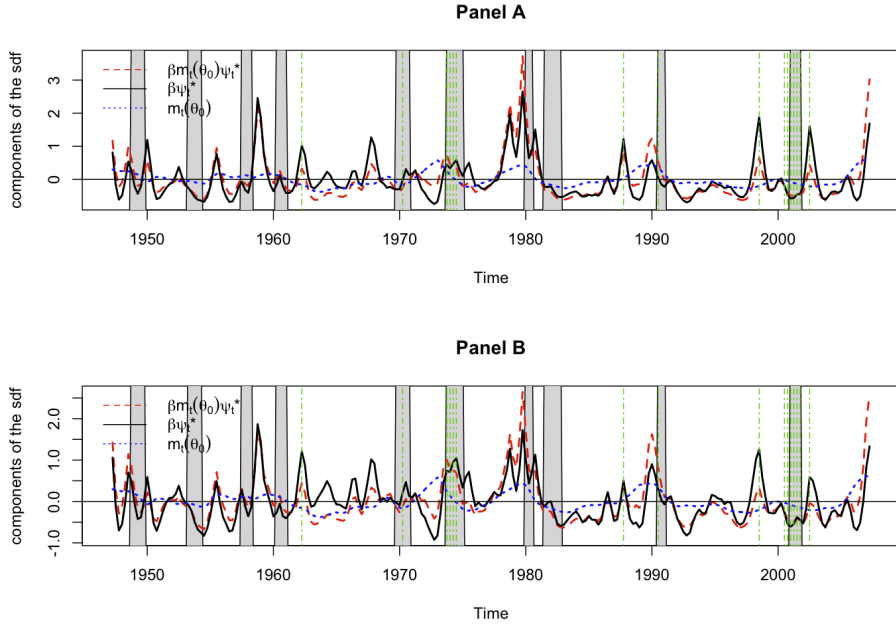
Figure 4: The figure plots the (demeaned) time series of the filtered SDF and its components for the ultimate consumption risk CCAPM for $\gamma=10$. Shaded areas are NBER recession periods. Vertical dashed lines are the stock market crashes identified by Mishkin and White (2002).

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.

20

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, ψ_t, that is model-specific:

$$M_t = (C_t/C_{t-1})^{-\gamma}\,\psi_t.$$

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, for each model we compute the values of the power coefficient, γ, at which the model-implied SDF satisfies the HJ, Q, M, and Ψ bounds. To simplify the exposition, we focus on one-dimensional bounds as a function of the risk aversion parameter, γ, 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. Second, 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.

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

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

where δ is the subjective time discount factor, γ is the curvature parameter, $S_t = \frac{C_t - X_t}{C_t}$ denotes the surplus consumption ratio and X_t is the habit component. Taking logs we have

$$\ln M_t = \ln \delta - \gamma \Delta c_t - \gamma \Delta s_t, \tag{15}$$

where lower case letters denote the natural logarithms of the upper case letters. Therefore, in this model, the expression for $\ln(\psi_t)$ is given by:

$$\ln \psi_t = \ln \delta - \gamma \Delta s_t. \tag{16}$$

Note that the missing component, ψ, depends on the surplus consumption ratio, S, that is not observed. To obtain the time series of ψ, we extract the surplus consumption ratio from observed consumption data as follows. In this model, the aggregate consumption growth is assumed to follow an $i.i.d.$ process:

$$\Delta c_t = g + v_t, \quad v_t \sim i.i.d.N\left(0, \sigma^2\right).$$

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

$$s_t = (1 - \phi)\,\overline{s} + \phi s_{t-1} + \lambda\left(s_{t-1}\right) v_t, \tag{17}$$

where $\overline{s}$ is the steady state log surplus consumption ratio and

$$\lambda\left(s_t\right) = \begin{cases} \frac{1}{\overline{S}}\sqrt{1 - 2\left(s_t - \overline{s}\right)}, & \text{if } s_t \leq s_{max} \\ 0, & \text{if } s_t > s_{max} \end{cases},$$

$$s_{max} = \overline{s} + \frac{1}{2}\left(1 - \overline{S}^2\right), \qquad \overline{S} = \sigma\sqrt{\frac{\gamma}{1-\phi}}.$$

For each value of γ, we use the calibrated values of the model preference parameters (δ, ϕ) in Campbell and Cochrane (1999), the sample mean (g) and volatility (σ) of the consumption growth process, and the innovations in real consumption growth, $\widehat{v}_t = \Delta c_t - g$, to extract the time series of the surplus consumption ratio using Equation (17) and obtain the time series of the model-implied SDF and its missing component from Equations (15) and (16).

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 μ_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:

$$dY_t = k\left(\overline{Y} - Y_t\right)dt - \alpha\left(Y_t - \lambda\right)\left[dc_t - E\left(dc_t\right)\right], \qquad (18)$$

where $\overline{Y}$ is the long run mean of the inverse surplus consumption ratio and k controls the speed of mean reversion. For each value of γ, we use the calibrated values of the model parameters $\left(\delta, k, \overline{Y}, \alpha, \lambda\right)$ in Menzly, Santos, and Veronesi (2004), the sample values of μ_c and σ_c, and the innovations in real consumption growth, $\widehat{dB_t} = \frac{[dc_t - E(dc_t)]}{\sigma_c}$, 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 from, respectively, Equations (15) and (16).

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

$$\ln M_{t+1} = \theta \log \delta - \frac{\theta}{\rho} \Delta c_{t+1} + (\theta - 1) r_{c,t+1}, \qquad (19)$$

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

The aggregate consumption and dividend growth rates, Δc_{t+1} and Δd_{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, σ_t^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 risk free rate, $r_{f,t}$, are affine functions of the state variables, x_t and σ_t^2. Therefore, Constantinides and Ghosh (2010) argue that these affine functions may be inverted to express the unobservable state variables, x_t and σ_t^2, in terms of the observables, $z_{m,t}$ and $r_{f,t}$. Following this approach, the pricing kernel in Equation (19) can be expressed, in log-linearized form, as a function of observable variables

$$\ln M_{t+1} = c_1 - \gamma \Delta c_{t+1} + c_3 \left(r_{f,t+1} - \kappa_1 r_{f,t} \right) + c_4 \left(z_{m,t+1} - \kappa_1 z_{m,t} \right), \qquad (20)$$

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, is obtained as

$$\ln M_t^q = -\gamma \Delta^q c_t + \ln \psi_t \qquad (21)$$

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

$$\ln \psi_t = 3c_1 + \sum_{i=0}^{2} \left[c_3 \left(r_{f,t-i} - \kappa_1 r_{f,t-i-1} \right) + c_4 \left(z_{m,t-i} - \kappa_1 z_{m,t-i-1} \right) \right]. \quad (22)$$

For each value of γ, 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 ψ component in Equations (21) and (22).

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

In this model, the pricing kernel is given by:

$$M_t = \delta \left(C_t/C_{t-1} \right)^{-\gamma} \left(A_t/A_{t-1} \right)^{\frac{\gamma\rho-1}{\rho-1}},$$

where A_t is the expenditure share on non-housing consumption, γ^{-1} is the intertemporal elasticity of substitution and ρ is the intratemporal elasticity of substitution between housing services and non-housing consumption.

Taking logs we have:

$$\ln M_t = \ln \delta - \gamma \Delta c_t + \frac{\gamma\rho - 1}{\rho - 1} \Delta a_t. \quad (23)$$

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

$$\ln \psi_t = \ln \delta + \frac{\gamma\rho - 1}{\rho - 1} \Delta a_t, \quad (24)$$

For each value of γ, we use the calibrated values of the model parameters (δ, ρ) in Piazzesi, Schneider, and Tuzel (2007) to obtain the time series of the model-implied SDF and its missing component from Equations (23) and (24), 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 the information bounds on the SDF and its unobservable component and the extracted time series of the SDF depend on the set of test assets used in their construction. Since the Euler equation holds for any traded asset as well as any adapted portfolio of the 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.

To assess the empirical plausibility of the asset pricing models described in the previous section using our methodology, we proceed in two ways. First, for each model we compute the minimum values of the power coefficient, γ, at which the model-implied SDF satisfies the HJ, Q, M, and Ψ bounds. Table I 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, $Q1$, $M1$, and $\Psi1$ bounds. The first row in each panel presents the bounds for the Campbell and Cochrane (1999) external habit model (henceforth referred to as CC). Panel A shows that when the excess return on the market portfolio is used in the construction of the bounds, the minimum value of γ at which the pricing kernel satisfies the HJ, $Q1$, $M1$, and $\Psi1$ 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, $Q1$, $M1$, and $\Psi1$ 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

26

satisfied for $\gamma \geqslant 9.8$, implying that the risk aversion varies over $[43.9, \infty)$, the M-bound is satisfied for $\gamma \geqslant 9.9$, implying that the risk aversion varies over $[44.2, \infty)$, and the Ψ-bound is satisfied for $\gamma \geqslant 13.9$, implying that the risk aversion varies over $[52.5, \infty)$. A similar ordering of the bounds is obtained when the set of assets consist 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 pointing to the robustness of our methodology.

The second row in each panel presents the bounds for the Menzly, Santos, and Veronesi (2004) external habit model (henceforth referred to as MSV). 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 Bansal and Yaron (2004) long run risks model (henceforth referred to as BY). Panel A shows that when the excess return on the market portfolio is used in the construction of the bounds, the minimum value of γ 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, γ 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 Ψ 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 Pi-

Table I: Bounds for RRA, Quarterly Data 1947:Q2-2009:Q4

	HJ-Bound	$Q1/Q2$-Bounds	$M1/M2$-Bounds	$\Psi1/\Psi2$-Bounds
Panel A: Market Portfolio				
CC	1.4	1.4/1.4	1.4/1.4	1.4/1.4
MSV	11.4	11.2/11.5	12.4/12.6	15.7/16.2
BY	3.0	3.0/3.0	3.0/3.0	3.0/3.0
PST	19.2	19.2/19.4	24.4/24.3	16.2/16.5
Panel B: FF 25 Portfolios				
CC	7.3	9.8/9.8	9.9/9.9	13.9/13.9
MSV	27.8	31.7/32.1	33.9/33.9	53.3/53.9
BY	4.0	5.0/5.0	5.0/5.0	5.0/5.0
PST	64.3	75.0/74.5	87.2/83.8	70.9/72.5
Panel C: 10 Size Portfolios				
CC	1.8	1.8/1.9	1.8/1.9	2.0/2.1
MSV	13.6	13.3/13.7	14.6/15.1	18.5/19.2
BY	3.0	4.0/4.0	4.0/4.0	4.0/4.0
PST	27.8	28.0/28.4	34.7/34.7	23.9/24.5
Panel D: 10 BM Portfolios				
CC	2.8	3.3/3.2	3.4/3.2	3.8/3.7
MSV	17.0	18.2/18.1	19.9/19.7	26.1/26.5
BY	4.0	4.0/4.0	4.0/4.0	4.0/4.0
PST	33.3	36.0/35.4	43.5/41.5	31.2/31.3
Panel E: 10 Momentum Portfolios				
CC	5.4	6.8/6.8	6.9/6.9	8.6/8.7
MSV	24.0	26.1/26.7	29.1/29.1	39.1/40.9
BY	4.0	5.0/4.0	5.0/4.0	5.0/4.0
PST	51.7	57.3/57.2	71.7/72.0	48.1/50.3
Panel F: 10 Industry Portfolios				
CC	3.0	3.0/3.3	3.1/3.3	3.5/3.8
MSV	17.6	17.3/18.2	19.1/20.0	26.4/27.6
BY	4.0	4.0/4.0	4.0/4.0	4.0/4.0
PST	36.2	37.0/37.9	46.7/47.1	35.5/36.6

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)

azzesi, Schneider, and Tuzel (2007) housing model (henceforth referred to as PST). When the set of test assets consists of the excess return on the market portfolio, the HJ, $Q1\,(Q2)$, $M1\,(M2)$, and $\Psi1\,(\Psi2)$ 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 I 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 II reports analogous bounds as in Table I at the annual frequency. The table shows that, at the annual frequency, the HJ, Q, M, and Ψ bounds are satisfied for much smaller values of the utility curvature parameter, γ, 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 I. However, in line with the theoretical predictions, the Q-bound is generally tighter than the HJ-bound, and the M-bound is tighter than the Q-bound.

Our second 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, γ. Therefore, for each model, we fix γ at its calibrated value and extract the time series of the SDF and its components.

Table III reports the results at the quarterly frequency. In order to ex-

Table II: Bounds for RRA, Annual Data 1930-2009

	HJ-Bound	$Q1/Q2$-Bounds	$M1/M2$-Bounds	$\Psi1/\Psi2$-Bounds
Panel A: Market Portfolio				
CC	0.1	0.1/0.1	0.1/0.1	0.1/0.1
MSV	0.1	0.1/0.1	0.1/0.1	0.1/0.1
BY	4.0	4.0/4.0	4.0/4.0	4.0/4.0
PST	8.6	9.9/9.5	14.9/15.0	6.8/6.8
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	4.0	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	16.3/14.2

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 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)

amine the models' ability to explain the cross-section of asset returns, we do not consider the market return on its own but focus instead on multiple test assets. Panels A, B, C, D, and E report results for the following sets of test assets: 25 Fama-French, 10 size-sorted, 10 book-to-market-equity-sorted, 10 momentum-sorted, and 10 industry-sorted portfolios, respectively. 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, $\{M_t^m = m_t\psi_t^m\}_{t=1}^T$.

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 ψ^*, the correlation between the filtered and model-implied ψ is only 0.02 when ψ^* 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 ψ^* 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, γ 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 ψ^* 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 ψ^* 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 (22), in the long run risks model the ψ component of the SDF is an exponentially affine function of the market-wide log

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

	Correlation of filtered and model SDF		Cross-sectional R^2	
	$\rho\left(\ln\psi_t^*,\ln\psi_t^m\right)$	$\rho\left(\ln M_t^*,\ln M_t^m\right)$	no intercept	free intercept
Panel A: Fama-French 25 Portfolios				
CC	0.02/0.06	0.05/0.08	-0.17	0.17
MSV	0.00/0.04	0.02/0.06	-0.64	0.0003
$BY^{rest.}$ *(unrest.)*	0.10/0.11 (0.27) (0.29)	0.11/0.12 (0.30) (0.31)	-1.01 (-0.48)	0.03 (0.02)
PST	$-0.04/-0.07$	0.06/0.01	-1.79	0.05
Panel B: 10 Size-Sorted Portfolios				
CC	0.19/0.17	0.25/0.23	-0.13	0.83
MSV	0.20/0.22	0.24/0.25	0.73	0.88
$BY^{rest.}$ *(unrest.)*	0.38/0.38 (0.80) (0.82)	0.37/0.38 (0.85) (0.86)	0.71 (0.89)	0.84 (0.94)
PST	$-0.22/-0.20$	$-0.01/-0.02$	-0.56	0.53
Panel C: 10 BM-Sorted Portfolios				
CC	0.16/0.15	0.20/0.20	0.27	0.60
MSV	0.14/0.16	0.18/0.19	0.06	0.24
$BY^{rest.}$ *(unrest.)*	0.34/0.33 (0.58) (0.65)	0.34/0.34 (0.65) (0.60)	-0.50 (-0.03)	0.07 (0.02)
PST	$-0.08/-0.09$	0.09/0.07	-2.31	0.27
Panel D: 10 Momentum-Sorted Portfolios				
CC	0.08/0.06	0.12/0.10	-0.64	0.07
MSV	0.06/0.08	0.10/0.12	-0.56	0.15
$BY^{rest.}$ *(unrest.)*	0.12/0.11 (0.44) (0.44)	0.13/0.12 (0.50) (0.48)	-1.06 (-0.40)	0.60 (0.29)
PST	$-0.12/-0.14$	0.07/0.02	-0.75	0.36
Panel E: 10 Industry-Sorted Portfolios				
CC	0.05/0.04	0.11/0.09	-11.1	0.003
MSV	0.06/0.11	0.10/0.12	-5.39	0.08
$BY^{rest.}$ *(unrest.)*	0.23/0.27 (0.55) (0.58)	0.25/0.28 (0.61) (0.62)	-1.58 (-1.17)	0.01 (0.11)
PST	0.13/0.12	0.19/0.17	-0.69	0.01

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)

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 (Row 3). Second, we provide results where the parameter vector is treated as free (in parentheses in Row 3). The parameter γ 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 ψ^* 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, γ 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 ψ^* 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 IV reports results analogous to those in Table III at the annual frequency. The results are largely similar to those in Table III. 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 III and IV report the cross-sectional R^2's 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 often varies wildly for the same model, and often take on large negative values when an intercept is not allowed in the cross-sectional regression, when evaluated using different sets of assets. This is in stark contrast with the results based on entropy bounds in Tables I and II, that tend instead to give consistent results for each model across different sets of assets (even though all models seem to perform better, along this dimension, at the annual frequency).

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 BM-sorted portfolios both in the presence and absence of an intercept.

Overall, Tables III and IV 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.

Table IV: Correlations of Filtered and Model SDFs, 1930-2009

	Correlation of filtered and model SDF		Cross-sectional R^2	
	$\rho\left(\ln \psi_t^*, \ln \psi_t^m\right)$	$\rho\left(\ln M_t^*, \ln M_t^m\right)$	no intercept	free intercept
Panel A: Fama-French 6 Portfolios				
CC	$0.15/0.16$	$0.23/0.22$	-0.75	0.03
MSV	$-0.04/-0.07$	-0.04/-0.07	-0.98	0.21
$BY^{rest.}$	$0.33/0.36$	$0.24/0.29$	0.35	0.35
(unrest.)	$(0.37)\ (0.42)$	$(0.62)\ (0.56)$	(0.50)	(0.66)
PST	$-0.04/-0.01$	-0.01/-0.05	-0.79	0.21
Panel B: 10 Size-Sorted Portfolios				
CC	$-0.03/-0.03$	$0.07/0.06$	-3.88	0.17
MSV	$0.06/-0.01$	$0.06/-0.01$	0.08	0.85
$BY^{rest.}$	$0.47/0.50$	$0.36/0.40$	0.86	0.96
(unrest.)	$(0.59)\ (0.61)$	$(0.68)\ (0.68)$	(0.80)	(0.95)
PST	$0.17/0.13$	-0.01/-0.08	0.11	0.91
Panel C: 10 BM-Sorted Portfolios				
CC	$0.07/0.03$	$0.16/0.10$	-3.12	0.01
MSV	$-0.08/-0.06$	-0.07/-0.06	-2.44	0.05
$BY^{rest.}$	$0.52/0.53$	$0.41/0.47$	0.40	0.47
(unrest.)	$(0.52)\ (0.54)$	$(0.61)\ (0.53)$	(0.73)	(0.82)
PST	$0.22/0.34$	$0.08/0.08$	0.18	0.57
Panel D: 10 Momentum-Sorted Portfolios				
CC	$0.26/0.27$	$0.34/0.33$	0.50	0.78
MSV	$0.09/0.07$	$0.09/0.08$	-1.51	0.40
$BY^{rest.}$	$0.41/0.50$	$0.31/0.41$	-0.33	0.31
(unrest.)	$(0.48)\ (0.53)$	$(0.68)\ (0.66)$	(-0.14)	(0.08)
PST	$-0.07/-0.06$	-0.03/-0.06	-0.45	0.01
Panel E: 10 Industry-Sorted Portfolios				
CC	$-0.04/-0.04$	$0.03/0.003$	-4.61	0.22
MSV	$-0.04/-0.10$	-0.04/-0.10	-4.72	0.24
$BY^{rest.}$	$0.26/0.39$	$0.20/0.34$	-1.24	0.17
(unrest.)	$(0.31)\ (0.41)$	$(0.37)\ (0.37)$	(-0.30)	(0.25)
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)

Tables V and VI 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^m and the fitted value from the regression. Similarly, Columns 2 and 3 present the correlation of the filtered SDF and its missing component with the three FF factors, respectively. These columns provide interesting results because the FF factors have been very successful at explaining the cross-sectional variation in returns between different classes of financial assets.

Consider first Table V. 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 0.54 and 0.59 when ψ^* 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 is asset returns. Moreover, Column 3 reveals that this high correlation is due almost entirely to the missing component, ψ^*, 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 ψ^*.

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 be-

Table V: Correlations with FF3, 1947:Q2-2009:Q4

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

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 VI: Correlations with FF3, 1930-2009

	Correlation With FF3		
	$\ln M_t^m$	$\ln M_t^*$	$\ln \psi_t^*$
Panel A: Fama-French 6 Portfolios			
CC	0.19	0.73/0.78	0.72/0.77
MSV	0.12	0.73/0.78	0.72/0.77
$BY^{rest.}$	0.73	0.77/0.77	0.68/0.72
(unrest.)	(0.81)	(0.78) (0.78)	(0.68) (0.72)
PST	0.35	0.81/0.76	0.65/0.67
Panel B: 10 Size-Sorted Portfolios			
CC	0.19	0.82/0.85	0.82/0.85
MSV	0.12	0.83/0.86	0.83/0.86
$BY^{rest.}$	0.73	0.77/0.77	0.71/0.73
(unrest.)	(0.84)	(0.78) (0.79)	(0.71) (0.74)
PST	0.35	0.75/0.72	0.64/0.66
Panel C: 10 BM-Sorted Portfolios			
CC	0.19	0.71/0.75	0.72/0.74
MSV	0.12	0.72/0.76	0.72/0.75
$BY^{rest.}$	0.73	0.67/0.60	0.59/0.58
(unrest.)	(0.83)	(0.67) (0.60)	(0.59) (0.58)
PST	0.35	0.64/0.23	0.50/0.33
Panel D: 10 Momentum-Sorted Portfolios			
CC	0.19	0.55/0.63	0.58/0.61
MSV	0.12	0.55/0.62	0.57/0.61
$BY^{rest.}$	0.73	0.69/0.69	0.51/0.57
(unrest.)	(0.85)	(0.73) (0.73)	(0.60) (0.64)
PST	0.35	0.73/0.70	0.50/0.55
Panel E: 10 Industry-Sorted Portfolios			
CC	0.19	0.49/0.53	0.49/0.53
MSV	0.12	0.50/0.54	0.50/0.55
$BY^{rest.}$	0.73	0.42/0.39	0.38/0.42
(unrest.)	(0.86)	(0.42) (0.38)	(0.36) (0.40)
PST	0.35	0.41/0.27	0.34/0.37

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)

tween 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 VI reveals that very similar results are obtained at the annual frequency. Tables V and VI 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 γ and the choice of the set of assets.

One thing to notice in Tables V and VI is that our filtered SDF and ψ^* are consistently highly correlated with the FF factors independently from the sample frequency and the cross-section of assets used for the estimation. 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.

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, 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 the information contained in a large cross-section of asset returns.

Second, using a relative entropy minimization approach, we also 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.

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) and 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 reasonable 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 substantial 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

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

AMEMIYA, T. (1985): *Advanced econometrics*. Harvard University Press, Cambridge: Mass.

AÏT-SAHALIA, Y., AND A. W. LO (1998): "Nonparametric Estimation of State-Price Densities Implicit in Financial Asset Prices," *Journal of Finance*, 53(2), 499–547.

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

BACKUS, D., M. CHERNOV, AND S. E. ZIN (2011): "Sources of Entropy in Representative Agent Models," Working Paper 17219, National Bureau of Economic Research.

BAKSHI, G. S., AND F. CHABI-YO (2011): "Variance Bounds on the Permanent and Transitory Components of Stochastic Discount Factors," *Journal of Financial Economics*, Forthcoming.

BANSAL, R., AND A. YARON (2004): "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles," *Journal of Finance*, 59(4), 1481–1509.

BREEDEN, D. T. (1979): "An Intertemporal Asset Pricing Model with Stochastic Consumption and Investment Opportunities," *Journal of Financial Economics*, 7, 265–96.

BREEDEN, D. T., M. R. GIBBONS, AND R. H. LITZENBERGER (1989): "Empirical Test of the Consumption-Oriented CAPM," *The Journal of Finance*, 44(2), 231–262.

BROWN, D. E., AND R. L. SMITH (1990): "A Correspondance Principle for Relative Entropy Minimization," *Naval Research Logistics*, 37(2).

BRUNNERMEIER, M. K., AND C. JULLIARD (2007): "Money Illusion and Housing Frenzies," *Review of Financial Studies*, 20(5).

CAMPBELL, J. Y. (1996): "Understanding Risk and Return," *Journal of Political Economy*, 104(2), 298–345.

CAMPBELL, J. Y., AND J. H. COCHRANE (1999): "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior," *Journal of Political Economy*, 107(2), 205–51.

COCHRANE, J. H. (1996): "A Cross-Sectional Test of an Investment-Based Asset Pricing Model," *Journal of Political Economy*, 104(3), 572–621.

CONSTANTINIDES, G. M., AND A. GHOSH (2010): "Asset Pricing Tests with Long Run Risks in Consumption Growth," Manuscript, University of Chicago.

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

DUFFIE, D. (2005): *Dynamic Asset Pricing Theory.* New Age International (P) Ltd.

EPSTEIN, L. G., AND S. E. ZIN (1989): "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica*, 57, 937–968.

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

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

HANSEN, L. P., AND T. J. SARGENT (2010): "Fragile Beliefs and the Price of Uncertainty," *Quantitative Economics*, 1(1), 129–162.

HAYASHI, F. (2000): *Econometrics.* Princeton University Press: Princeton.

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

JAYNES, E. T. (1957a): "Information Theory and Statistical Mechanics," *The Physics Review*, 106(4), 620–630.

——— (1957b): "Information theory and statistical mechanics II," *The Physical Review*, 108(2), 171–190.

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

KREPS, D. M., AND E. L. PORTEUS (1978): "Temporal Resolution of Uncertainty and Dynamic Choice Theory," *Econometrica*, 46(1), 185–200.

LETTAU, M., AND S. LUDVIGSON (2001a): "Consumption, Aggregate Wealth, and Expected Stock Returns," *Journal of Finance*, 56(3), 815–49.

——— (2001b): "Resurrecting the (C)CAPM: A Cross-Sectional Test When Risk Premia Are Time-Varying," *Journal of Political Economy*, 109, 1238–1286.

LUSTIG, H. N., AND S. G. V. NIEUWERBURGH (2005): "Housing Collateral, Consumption Insurance, and Risk Premia: An Empirical Perspective," *The Journal of Finance*, 60(3), pp. 1167–1219.

MANKIW, N. G., AND M. D. SHAPIRO (1986): "Risk and Return: Consumption Beta Versus Market Beta," *Review of Economics and Statistics*, 68, 452–59.

MEHRA, R., AND E. C. PRESCOTT (1985): "The Equity Premium: A Puzzle," *Journal of Monetary Economics*, 15(2), 145–61.

MENZLY, L., T. SANTOS, AND P. VERONESI (2004): "Understanding Predictability," *Journal of Political Economy*, 112(1), 1–47.

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

OWEN, A. B. (1988): "Empirical Likelihood Ratio Confidence Intervals for a Single Functional," *Biometrika*, 75(2), 237–249.

——— (1991): "Empirical Likelihood for Linear Models," *The Annals of Statistics*, 19(4), 1725–1747.

——— (2001): *Empirical Likelihood*. Chapman and Hall.

PARKER, J. A., AND C. JULLIARD (2005): "Consumption Risk and the Cross-Section of Expected Returns," *Journal of Political Economy*, 113(1).

PIAZZESI, M., M. SCHNEIDER, AND S. TUZEL (2007): "Housing, consumption and asset pricing," *Journal of Financial Economics*, 83(3), 531–569.

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

RUBINSTEIN, M. (1976): "The Valuation of Uncertain Income Streams and the Pricing of Options," *Bell Journal of Economics*, 7, 407–25.

SCHERVISH, M. J. (1995): *Theory of statistics*, Springer series in statistics. Springer-Verlag.

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

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

ULRICH, M. (2010): "Observable Long-Run Ambiguity and Long-Run Risk," Columbia University Manuscript.

WEIL, P. (1989): "The Equity Premium Puzzle and the Risk-Free Rate Puzzle," *Journal of Monetary Economics*, 24, 401–421.

A   Appendix

A.1   Maximum Likelihood Analogy

The approaches in Equations (4) and (6) deliver maximum likelihood estimates of the ψ_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 Δ^{T-1}

$$\Delta^{T-1} \equiv \left\{ (\psi_1, \psi_2,..., \psi_T) : \psi_t \geqslant 0, \ \sum_{t=1}^{T} \psi_t = 1 \right\},$$

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

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

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 ψ_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) $\left\{ \tilde{\psi} \right\}_{t=1}^{T}$ given by

$$\tilde{\psi}_t \equiv \frac{n_t}{N}$$

where n_t measures the number of times that the value $1/N$ has been assigned to time t. Second, check whether the drawn series $\left\{ \tilde{\psi} \right\}_{t=1}^{T}$ satisfies the pricing restriction $\sum_{t=1}^{T} m\left(\theta, t\right) R_t^e \tilde{\psi}_t = 0$. 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 ψ_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 $\left\{\tilde{\psi}_t\right\}_{t=1}^T$ 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 log 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[6] 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 ψ_t would solve

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

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 KLIC minimization is equivalent to maximizing the likelihood in an unbiased procedure for finding the ψ_t component of the pricing kernel.

A.2 Additional Bounds and Derivations

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

$$M_t^*\left(\bar{M}\right) \equiv \underset{\left\{M_t\left(\bar{M}\right)\right\}_{t=1}^T}{\arg\min} \ \sqrt{Var\left(M_t\left(\bar{M}\right)\right)} \ s.t. \ \mathbf{0} = \mathbb{E}\left[\mathbf{R}_t^e M_t\left(\bar{M}\right)\right] \qquad (25)$$

---

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

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

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

The solution of the problem in Equation (25) is

$$M_t^*\left(\bar{M}\right) = \bar{M} + \left(\mathbf{R}_t^e - \mathbb{E}\left[\mathbf{R}_t^e\right]\right)' \beta_{\bar{M}},$$

where $\beta_{\bar{M}} = Cov\left(\mathbf{R}_t^e\right)^{-1}\left(-\bar{M}\mathbb{E}\left[\mathbf{R}_t^e\right]\right)$.

Definition 4 (Q-bounds) *We define the following risk neutral probability bounds for any candidate stochastic discount factor M_t.*

1. *Q1-bound:*

$$D\left(P||\frac{M_t}{\bar{M}}\right) \equiv \int -\ln\frac{M_t}{\bar{M}}dP \geqslant D\left(P||Q^*\right)$$

   where Q^* solves Equation (7).

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

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

   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.[7] Assuming that there exists a $\mu_p < \infty$ and a $\mu_q < \infty$ such that:*

1. *(Existence of maxima)*

$$\left.\frac{\partial \ln p}{\partial x}\right|_{x=\mu_p} = 0, \quad \left.\frac{\partial \ln q}{\partial x}\right|_{x=\mu_q} = 0;$$

2. *(Finite second moments)*

$$-\left[\left.\frac{\partial^2 \ln p}{\partial x^2}\right|_{x=\mu_p}\right]^{-1} \equiv \sigma_p^2 < \infty, \quad -\left[\left.\frac{\partial^2 \ln q}{\partial x^2}\right|_{x=\mu_q}\right]^{-1} \equiv \sigma_q^2 < \infty.$$

---

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

We have that, in the limit of the small time interval, a second order approximation of the Q-bounds yields[8]

$$D\left(P||\frac{M_t}{\bar{M}}\right) \quad \propto \quad Var\left(M_t\right), \tag{26}$$

$$D\left(\frac{M_t}{\bar{M}}||P\right) \quad \propto \quad Var\left(M_t\right). \tag{27}$$

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 $Q1$ and $Q2$ bounds, respectively, as

$$D\left(P||\frac{M_t}{\bar{M}}\right) \equiv \int \ln\frac{dP}{dQ}dP = \int p\ln\frac{p}{q}dx \tag{28}$$

and

$$D\left(\frac{M_t}{\bar{M}}||P\right) \equiv \int \frac{dQ}{dP}\ln\frac{dQ}{dP}dP = \int \ln\frac{dQ}{dP}dQ = \int q\ln\frac{q}{p}dx. \tag{29}$$

If q and p are twice continuously differentiable and there exists a μ_p and a μ_q such that

$$\left.\frac{\partial\ln p}{\partial x}\right|_{x=\mu_p} = 0, \quad \left.\frac{\partial\ln q}{\partial x}\right|_{x=\mu_q} = 0,$$

and

$$-\left[\left.\frac{\partial^2\ln p}{\partial x^2}\right|_{x=\mu_p}\right]^{-1} \equiv \sigma_p^2 < \infty, \quad -\left[\left.\frac{\partial^2\ln q}{\partial x^2}\right|_{x=\mu_q}\right]^{-1} \equiv \sigma_q^2 < \infty$$

we have from a second order Taylor approximation that

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

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

---

[8]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 (27) 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 (27) is exact.

47

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 (28) we have that in the limit of the small time interval

$$
\begin{aligned}
\int p \ln \frac{p}{q} dx &\approx \int \left[-\frac{1}{2} \frac{(x - \mu_p)^2}{\sigma^2} + \frac{1}{2} \frac{(x - \mu_q)^2}{\sigma^2} \right] p dx \\
&= \frac{1}{2\sigma^2} \left[-\sigma^2 + \int (x - \mu_q)^2 p dx \right] \\
&= \frac{1}{2\sigma^2} \Big\{ -\sigma^2 + \int \Big[(x - \mu_p)^2 + (\mu_p - \mu_q)^2 \\
&\quad + 2 (\mu_p - \mu_q)(x - \mu_p) \Big] p dx \Big\} \\
&= \frac{1}{2\sigma^2} (\mu_p - \mu_q)^2 = \frac{1}{2\sigma^2} \sigma^2 \sigma_\xi^2 = \frac{1}{2} \sigma_\xi^2
\end{aligned}
$$

where the density ξ is a (strictly positive) martingale defined by $\xi \equiv \frac{dQ}{dP}$, 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 (29) we have

$$\int q \ln \frac{q}{p} dx = \frac{1}{2} \sigma_\xi^2.$$

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

$$D\left(P \| \frac{M_t}{\overline{M}}\right) \propto Var(M_t), \quad D\left(\frac{M_t}{\overline{M}} \| P\right) \propto Var(M_t).$$

∎

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

$$\psi_t^*\left(\bar{\psi}\right) \equiv \underset{\{\psi_t(\bar{\psi})\}_{t=1}^{T}}{\arg\min} \sqrt{Var\left(\psi_t\left(\bar{\psi}\right)\right)} \text{ s.t. } \mathbf{0} = \mathbb{E}\left[\mathbf{R}_t^e m\left(\theta, t\right) \psi_t\left(\bar{\psi}\right)\right]$$

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

48

The solution of the above minimization for a given θ is

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

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

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

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 portfolios 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 nondurables and services excluding housing services.