# New Entropy Restrictions and the Quest for Better Specified Asset Pricing Models*

Gurdip Bakshi[a†] Fousseni Chabi-Yo[b‡]

[a]*Smith School of Business, University of Maryland, College Park, MD 20742, USA*

[b]*Isenberg School of Management, University of Massachusetts, Amherst, MA 01002, USA*

This draft, December 3, 2017 (First draft, January 2013)

---

**Abstract**

This paper proposes the entropy of $m^2$ ($m$ is the stochastic discount factor) as a metric to evaluate asset pricing models. We develop a bound on the entropy of $m^2$ when $m$ correctly prices a finite number of returns and consider models that pass the lower bound on $m$, yet fail the lower bound on $m^2$. Interpreting our results, we elaborate on the distinction between the entropy of $m^2$ versus the entropy of $m$. We further show that the entropy of $m^2$ represents an upper bound on the expected excess (log) return of the security with the payoff of $m$.

KEY WORDS: Stochastic discount factors, lower entropy bounds, individual asset pricing models

JEL CLASSIFICATION CODES: C51, C52, G12.

*We are deeply indebted to the referee and Jennifer Conrad (the editor) for their efforts in completely revamping this paper. The authors acknowledge helpful discussions with Ravi Bansal, Mike Chernov, Riccardo Colacito, John Crosby, Xiaohui Gao, Steve Heston, Ravi Jagannathan, Mark Loewenstein, Dilip Madan, George Panayotov, Leonid Kogan, Alberto Rossi, Ken Singleton, Georgios Skoulakis, Rene Stulz, Michael Stutzer, Adrien Verdelhan (AFA 2015 discussant), Jessica Wachter, Jinming Xue, and Lu Zhang. Seminar participants at the University of Maryland and Ohio State University provided useful suggestions. We have also benefited from participants at the 2015 American Finance Association meeting (Boston) and 2014 Econometric Society meetings (Minneapolis). Any remaining errors are our responsibility alone. All computer codes are available from the authors.

†Tel.: +1-301-405-2261. *E-mail address:* gbakshi@rhsmith.umd.edu

‡Tel.: +1-614-292-8477. *E-mail address:* fchabiyo@isenberg.umass.edu

# 1. Introduction

The lower bound statistics constructed from observed asset returns have found their applicability in evaluating individual asset pricing models. Taking a cue from the theory underlying the volatility bound of Hansen and Jagannathan (1991), a wave of research has emphasized entropy bounds on the stochastic discount factor and its correlated permanent and transitory components. This line of inquiry includes, among others, Alvarez and Jermann (2005), Hansen and Scheinkman (2009), Bakshi and Chabi-Yo (2012), Backus, Chernov, and Zin (2014), Borovička, Hansen, and Scheinkman (2016), and Almeida and Garcia (2017).

Our idea is to focus on a new moment – the entropy of the *square* of the stochastic discount factor (hereby SDF, or simply $m$) – and we develop the corresponding lower entropy bound.

The novelty of our paper is to show that the entropy of $m^2$ can be employed as a statistic to reject some models that pass the lower bound on the entropy of $m$. The defining attribute of the entropy of $m^2$, from an asset pricing perspective, is that it represents an upper bound on the expected excess (log) gross return of the security with a payoff of $m$. Additionally, we show that the lower bound on the entropy of $m^2$ can be derived from a vector of traded asset returns, and the proposed bound is distinct from others with no analytical analogs.

The developments here are relevant, as the quest for well-specified SDFs has dominated the agenda in asset pricing. Despite substantial progress, identifying the desirable properties of the SDFs, in addition to their link to economic fundamentals, remains a tall order.

In our framework, we consider models that meet the minimum entropy criterion on the entropy of $m$ and then ask the question: Is the same model poised for acceptance based on the bound on the entropy of $m^2$? Our empirical analysis results in the finding that it is possible for a model to respect the lower bound on the entropy of $m$ but fail to satisfy the lower bound on the entropy of $m^2$. In this sense, our niche is to show that the lower bound on the entropy of $m^2$ (constructed from observed asset returns) can offer a way to evaluate individual asset pricing models.[1]

Our approach also clarifies how certain variables can help to improve the workings of a model. For example, a baseline model with market, size, book-to-market, operating profitability, and investment portfolios is rejected (respectively, is not rejected) according to the lower bound on the entropy of $m^2$ (respectively, $m$), with eight out of nine bootstrap $p$-values below 0.1 (our Table 2). To the contrary, adding momentum variables to the baseline model is associated with a threefold increase in the bootstrap $p$-values, with five out of nine $p$-values above 0.1 (our Table 3). Our applications indicate that the bound on the entropy of $m^2$ is tighter than the bound on the entropy of $m$. This feature further motivates the use of the bound statistic based on the entropy of $m^2$.

## 2. Entropy of $m^2$

Let $m_{t,t+1}$ be the SDF between time $t$ and $t+1$. Moreover, let $E_t[.]$ ($E[.]$) represent conditional (unconditional) expectation. We consider SDFs that conditionally price the risk-free bond:

$$E_t[m_{t,t+1} \times 1] \equiv q_t = \frac{1}{R_{t+1,f}}, \tag{1}$$

where $R_{t+1,f}$ is the gross return of the risk-free bond.

Our results are not affected by the length of the period for $m_{t,t+1}$, which can be arbitrary.

[1] Our characterizations (see equations (5) and (9)) show that, compared with the entropy of $m$, the entropy of $m^2$ assigns heavier weights to the higher moments (beyond the mean) of $m$. Intuitively, a candidate model is amiss if it does not generate higher-moment effects consistent with the entropy bound on $m^2$.

Hence, we use the notations $m$ and $m_{t,t+1}$ interchangeably. The unconditional variance of $m_{t,t+1}$ is denoted $\text{var}[m]$.

## *2.1. The role of entropy L[m] in testing asset pricing models*

In this subsection, we consider the entropy of $m$, which serves as a bridge for highlighting our results on the entropy of $m^2$. The unconditional entropy of $m$, denoted by $L[m]$, is defined as[2]

$$L[m] \;\equiv\; \log(E[m]) \;-\; E[\log(m)]. \tag{2}$$

The entropy measure $L[m]$ is related to *Jensen's gap*, $J\{m\}$, defined as

$$J\{m\} \;\equiv\; E[f\{m\}] \;-\; f\{E[m]\} \;\geq 0, \text{ applied to the } \textit{convex function } f\{m\} = -\log(m). \tag{3}$$

In contrast, the variance measure used in Hansen and Jagannathan (1991) is related to Jensen's gap applied to the convex function $f\{m\} = m^2$.

An expansion-based interpretation of entropy obtains by expressing $L[m] = \log(E[e^{\log(m)}]) - E[\log(m)]$. To outline this depiction, we take a Taylor expansion of $\exp(n\log(m))$ around $E[n\log(m)]$, for a positive integer $n$, which implies that

$$\begin{aligned}\exp(\log(m^n)) \;&=\; e^{nE[\log(m)]}\left(1 + n\left(\log(m) - E\left[\log(m)\right]\right) + \frac{n^2}{2!}\left(\log(m) - E\left[\log(m)\right]\right)^2\right.\\ &\qquad \left. + \frac{n^3}{3!}\left(\log(m) - E\left[\log(m)\right]\right)^3 + \frac{n^4}{4!}\left(\log(m) - E\left[\log(m)\right]\right)^4 + \ldots\right). \end{aligned} \tag{4}$$

[2]The asset pricing implications of entropy are explored in the studies of Stutzer (1995), Bansal and Lehmann (1997), Alvarez and Jermann (2005), Bakshi and Chabi-Yo (2012), Backus, Chernov, and Zin (2014), Borovička, Hansen, and Scheinkman (2016), Qin, Linetsky, and Nie (2016), Orlowski, Sali, and Trojani (2016), Almeida and Garcia (2017), Christensen (2017), Ghosh, Julliard, and Taylor (2017), and Bakshi, Chabi-Yo, and Gao (2017).

Applying expectation and taking logs on both sides in equation (4), we get $L[m^n] = \log(E[m^n]) - E[\log(m^n)]$ (upon rearranging). Hence, with $n = 1$, it holds that

$$L[m] \quad = \quad \log\Big(1 + \sum_{j=2}^{\infty} \frac{\mu_{\log(m)}^{[j]}}{j!}\Big), \quad \text{where } \mu_{\log(m)}^{[j]} \equiv E[(\log(m) - E[\log(m)])^j]. \tag{5}$$

$L[m]$ can be viewed as encapsulating the central moments of $\log(m)$, and distributions of $\log(m)$ that incorporate fatter tails tend to support a higher $L[m]$. When $\log(m)$ follows a normal distribution (i.e., $m$ is distributed lognormal), the correspondence between entropy and the central moments of $\log(m)$ is exact[3] and is given by $L[m] = \frac{1}{2}\text{var}[\log(m)]$.

How is the entropy measure $L[m]$ used in the tests of asset pricing models? Alvarez and Jermann (2005, page 2008, equation (A.1)) and Backus, Chernov, and Zin (2014, page 57, equation (5)) propose the following lower entropy bound:

$$\underbrace{L[m]}_{\text{Entropy from model}} \quad \geq \quad \underbrace{E[\log(R_{t+1}^m) - \log(R_{t+1,f})]}_{\text{Based on observed returns}}. \tag{6}$$

The bound on $L[m]$ in equation (6) is denominated in units of expected log gross return of a generic portfolio (*any* arbitrary portfolio with gross return, $R_{t+1}^m$, that satisfies correct pricing) in excess of log gross return of a risk-free bond. In contrast, Almeida and Garcia (2017) offer a result in which the entropy bound can be determined as $\inf_m \; (-E[\log(m)])$, where $m$ is extracted from an optimization problem while ensuring that $m$ correctly prices a vector of asset returns.

[3]This follows from employing the expression of the moment generating function of the normal distribution applied to $\log(m)$, i.e., $E[e^{\log(m)}] = \exp(E[\log(m)] + \frac{1}{2}\text{var}[\log(m)])$, using the operation of log and rearranging to use the definition of entropy in equation (2).

## 2.2. *Rationale for studying the entropy $L[m^2]$ and the bound on $L[m^2]$*

This subsection has two objectives. First, we provide a rationale for studying $L[m^2]$. Second, we propose a lower bound on $L[m^2]$ as a metric for evaluating asset pricing models.

The unconditional entropy of $m^2$ is defined as

$$L[m^2] \;\equiv\; \log(E[m^2]) \,-\, E[\log(m^2)], \tag{7}$$

which can be expressed in terms of *Jensen's gap* as

$$J\{m\} \;=\; E[f\{m^2\}] \,-\, f\{E[m^2]\}, \text{ applied to the } \textit{convex function } f\{m\} = -\log(m). \tag{8}$$

Additionally, based on equation (4), setting $n = 2$, and comparing with equation (7), we note that

$$L\left[m^2\right] \;=\; \log\left(1 + \frac{2^2}{2!}\mu^{[2]}_{\log(m)} + \frac{2^3}{3!}\mu^{[3]}_{\log(m)} + \frac{2^4}{4!}\mu^{[4]}_{\log(m)} + \ldots\right). \tag{9}$$

Viewed through the prism of the central moments of $\log(m)$, equation (9) shows that $L[m^2]$ assigns a bigger weight to each central moment of $\log(m)$ than $L[m]$. The relative merits of $L[m]$ and $\text{var}[m]$ are addressed in Alvarez and Jermann (2005, page 1985) on the grounds that additional higher moments affect $L[m]$ besides $\text{var}[\log(m)]$.

The assumption of lognormality of $m$ implies that $\mu^{[3]}_{\log(m)} = 0$, $\mu^{[4]}_{\log(m)} = 3(\mu^{[2]}_{\log(m)})^2$, $\mu^{[5]}_{\log(m)} = 0$, $\mu^{[6]}_{\log(m)} = 15(\mu^{[2]}_{\log(m)})^3$, and so on. Then $L[m^2] = 4(\frac{1}{2}\text{var}[\log(m)])$ and, hence, $L[m^2] = 4L[m]$.

To establish the link among the entropy of $m^2$, the entropy of $m$, and variance of $m$, we subtract twice of $L[m]$ in equation (3) from $L[m^2]$ in equation (7), and note that

$$L[m^2] \;=\; 2\,L[m] \;+\; \log\left(1+\frac{\mathrm{var}[m]}{(E[m])^2}\right). \tag{10}$$

Equation (10) elicits three observations. First, Hansen and Jagannathan (1991) derive the lower bound on $\mathrm{var}[m]$, when the SDF correctly prices finitely many assets. Second, the lower bound on $L[m]$ is not known when the SDF correctly prices finitely many assets (and is presented here in equation (20)). Third, when the bounds are not unique, lower bounding the parts in a sum might not be a proper way to lower bound the sum in the right-hand side of equation (10).[4]

We now present a theoretical lower bound on $L[m^2]$ that can be inferred from a vector of traded asset returns, when the SDF is required to correctly price finitely many returns. To do so, consider the following set $\mathbb{S}$ of the SDFs:

$$\mathbb{S} = \{m_{t,t+1} > 0 : E\left[m_{t,t+1}\right] = E[q_t] \text{ and } E\left[m_{t,t+1}\mathbf{R}_{t,t+1}\right] = \mathbf{1}\}, \tag{11}$$

where $\mathbf{1}$ is a vector column of ones. The restriction $E[m_{t,t+1}] = E[q_t]$ follows from the feature that the SDFs conditionally price the risk-free bond. Equation (11) further requires the SDFs to unconditionally price an additional $N \times 1$ vector of gross returns.

**Result 1** *The entropy of $m^2$ satisfies*

$$\begin{aligned} L[m^2] \;\geq\; \mathbb{LB}_{m^2} \equiv 2\left(E\left[\log\left(\frac{(\mathbf{1}-E\left[q_t\right]E\left[\mathbf{R}_{t,t+1}\right])'\Sigma^{-1}}{\mathbf{1}'\Sigma^{-1}\left(\mathbf{1}-E\left[q_t\right]E\left[\mathbf{R}_{t,t+1}\right]\right)}\mathbf{R}_{t,t+1}\right)\right] \;-\; \log\left((E[q_t])^{-1}\right)\right) \\ +\; \log\left(1+(\mathbf{1}-E\left[q_t\right]E\left[\mathbf{R}_{t,t+1}\right])'\Sigma^{-1}\left(\mathbf{1}-E\left[q_t\right]E\left[\mathbf{R}_{t,t+1}\right]\right)/\left(E[q_t]\right)^2\right), \end{aligned} \tag{12}$$

[4]Suppose you have a function $K[m] = H[m] + G[m]$. If one could bound the function as $H[m] > h^*$ and $G[m] > g^*$, then $h^* + g^*$ can only be a unique lower bound for $K[m]$, if $h^*$ and $g^*$ are *unique* lower bounds.

*where $\Sigma$ is the variance-covariance matrix of $\mathbf{R}_{t,t+1}$.*

**Proof**: See Appendix A. ■

The lower entropy bound in equation (12) summarizes the properties of the distribution of $m$. Define, for brevity, the $N \times 1$ vector of constants

$$\mathbf{a} \equiv \frac{\Sigma^{-1}\left(\mathbf{1} - E\left[q_t\right] E\left[\mathbf{R}_{t,t+1}\right]\right)}{\mathbf{1}' \underbrace{\Sigma^{-1}\left(\mathbf{1} - E\left[q_t\right] E\left[\mathbf{R}_{t,t+1}\right]\right)}_{\equiv \mathbf{y}}}, \tag{13}$$

which allows us to write the first term of the lower bound in equation (12) as

$$\log(\mathbf{a}'\mathbf{R}_{t,t+1}) = \log(E[\mathbf{a}'\mathbf{R}_{t,t+1}]) + \log\left(1+\tilde{u}\right) \quad \text{with} \quad \tilde{u} \equiv \frac{\mathbf{a}'\mathbf{R}_{t,t+1} - E[\mathbf{a}'\mathbf{R}_{t,t+1}]}{E[\mathbf{a}'\mathbf{R}_{t,t+1}]}. \tag{14}$$

Taking a Taylor expansion $\log(1+\tilde{u}) = \sum_{i=1}^{\infty} \frac{(-1)^{1+i}}{i} \tilde{u}^i$, we observe that

$$E[\log(\mathbf{a}'\mathbf{R}_{t,t+1})] = \log(E[\mathbf{a}'\mathbf{R}_{t,t+1}]) + \sum_{i=1}^{\infty} \frac{(-1)^{1+i}}{i} \frac{E[(\mathbf{a}'\mathbf{R}_{t,t+1} - E[\mathbf{a}'\mathbf{R}_{t,t+1}])^i]}{(E[\mathbf{a}'\mathbf{R}_{t,t+1}])^i}. \tag{15}$$

Hence, the lower bound in equation (12) can be alternatively expressed as

$$\begin{aligned} \mathbb{LB}_{m^2} \;\; &= \;\; 2\{\log(E[\mathbf{a}'\mathbf{R}_{t,t+1}]) - \log((E[q_t])^{-1})\} \; + \; \log(1 + \mathbf{y}'\Sigma\mathbf{y}/\left(E\left[q_t\right]\right)^2) \\ &\quad + \; 2\sum_{i=1}^{\infty} \frac{(-1)^{1+i}}{i} \frac{E[(\mathbf{a}'\mathbf{R}_{t,t+1} - E[\mathbf{a}'\mathbf{R}_{t,t+1}])^i]}{(E[\mathbf{a}'\mathbf{R}_{t,t+1}])^i} \end{aligned} \tag{16}$$

and is computable from the time-series of asset returns. Observe that the bound on $L[m^2]$ amplifies the effect of higher-order return moments (relative to the lower bound on the entropy of $m$).

Ghosh, Julliard, and Taylor (2017) construct entropy bounds when the SDF can be factorized into observable and model-specific unobservable components. Analogously, the bound on $\text{var}[m]$, that is, the Hansen and Jagannathan (1991) bound, and our bound on $L[m^2]$ in equation (12), constitute relevant metrics for evaluating individual asset pricing models.

Liu (2014, Proposition 1 and Collorary 1) derives an upper bound on $E[m^\delta]$, when $\delta \in [0,1]$, and a lower bound on $E[m^\delta]$, when $\delta < 0$, where $\delta$ is expressed in terms of the risk aversion parameter $\gamma \equiv \frac{1}{1-\delta}$. Further, our bound on $L[m^2]$ offers a distinction to the noncentral moment bounds considered in Snow (1991, equations (7) and (12)).

## 2.3. *Economic interpretations of $L[m^2]$*

Focusing on theoretical and economic rationale, we next show that $L[m^2]$ encodes information about the expected excess (log) return of a security that entitles the investor to a payoff of $m_{t,t+1}$. The security with SDF payoff has return

$$r_{t,t+1}^{\text{SDF}} \equiv \frac{m_{t,t+1}}{E_t[m_{t,t+1}^2]} - 1 \tag{17}$$

and is a hedging asset with $\log(R_{t+1,f}) - E_t[\log(1+r_{t,t+1}^{\text{SDF}})] \geq 0$.

We establish the following result.

**Result 2** *The expected excess (log) return of a security with SDF payoff is related to $L[m^2]$ as follows:*

$$L[m^2] \geq \underbrace{E[\log(R_{t+1,f}) - \log(1+r_{t,t+1}^{\text{SDF}})]}_{\text{Expected excess (log) return of SDF security}} \geq 0. \tag{18}$$

**Proof**: See Appendix B. ■

The expected excess (log) return of a security that pays $m$ is limited by the entropy of $m^2$, similar to how the Sharpe ratio is upper bounded by the volatility of $m$ (e.g., Cochrane (2005, page 20)). Our result can be traced to equation (B3) of Appendix B, which shows that the conditional expected excess (log) return of the SDF security reflects the difference between the conditional entropy of $m^2$ and the conditional entropy of $m$.

Contemplating our Result 2, there is one additional issue to be addressed: When does the bound in equation (18) bind? The next result formalizes the *special* restrictions on the SDF, denoted as $m^\bullet$, under which the bound on $L[m^2]$ becomes tight.

**Problem 1** *Identify an SDF, denoted by $m^\bullet$, satisfying $L[(m^\bullet)^2] = \mathbb{LB}_{m^2}$, where $\mathbb{LB}_{m^2}$ is derived by using the pricing restriction that $m \in \mathbb{S}$.* ■

Appendix C shows that a solution to Problem 1 is of the form

$$m^\bullet_{t,t+1} = \frac{\psi}{\left(\mathbf{a}'\mathbf{R}_{t,t+1}\right)^\nu}. \tag{19}$$

We determine the parameters $\nu$ and $\psi$ sequentially via equations (C13) and (C14), where vector $\mathbf{a}$ is defined in equation (13). Additionally, the solution for $m^\bullet_{t,t+1}$ requires that $m^\bullet_{t,t+1}$ and $\mathbb{LB}_{m^2}$ be from a vector of gross returns $\mathbf{R}_{t,t+1}$. While our result is in terms of entropy, it is analogous, in terms of variance, to the approach of Hansen and Jagannathan (1991).

Finally, the lower bound on $L[m^3]$ is presented in (D14) of Appendix D. Whereas the bound on $L[m^2]$ in (12) is based on the mean and variance-covariance matrix of $\mathbf{R}_{t,t+1}$, the bound on $L[m^3]$ is associated with the expectation of demeaned $\mathbf{y}'\mathbf{R}_{t,t+1}$ raised to the power $3/2$. Our preference for choosing $L[m^2]$ as a dispersion measure to study asset pricing models is guided by empirical considerations. First, the bound on $L[m^2]$ is tractable to implement relative to the one on $L[m^3]$.

Second, we show that the bound on $L[m^2]$ is tighter than the bound on $L[m]$ in our applications.

# 3. Revealing the value added of $L[m^2]$ over $L[m]$

We study the value added of $L[m^2]$ over $L[m]$ from different angles.

## *3.1. Building an empirical case for considering $L[m^2]$*

Appendix E shows that the lower bound on $L[m]$ is given by (where vector **a** is determined from the properties of $\mathbf{R}_{t,t+1}$ via the calculation in equation (13))

$$L[m] \geq \mathbb{LB}_m \equiv E[\log(\mathbf{a}^{'}\mathbf{R}_{t,t+1})] \; - \; \log((E[q_t])^{-1}). \tag{20}$$

Equation (20) is consistent with a theory that every $m$ that prices the vector of assets in set $\mathbb{S}$ (in equation (11)) must have $L[m] \geq \mathbb{LB}_m$ and simultaneously must have $L[m^2] \geq \mathbb{LB}_{m^2}$.

In this light, we next explore a framework in which an individual asset pricing model can pass the lower entropy bound on $m$, but the $L[m^2]$ statistic potentially serves as a way of rejecting it. These exercises emphasize the value of $L[m^2]$.

**Example 1** [*Linear SDF with returns of the market and returns of options*]: The model is

$$m_{t,t+1} \; = \; \boldsymbol{\eta}'[R_{t+1,f} \;\; \mathbf{Q}_{t,t+1}], \qquad \text{with} \qquad \mathbf{Q}_{t,t+1} = \begin{pmatrix} \frac{S_{t+1}}{S_t} \\ \frac{\max(S_t e^{-0.03} - S_{t+1}, 0)}{P_t[S_t e^{-0.03}]} \\ \frac{\max(S_t e^{-0.01} - S_{t+1}, 0)}{P_t[S_t e^{-0.01}]} \\ \frac{\max(S_{t+1} - S_t e^{0.01}, 0)}{C_t[S_t e^{0.01}]} \\ \frac{\max(S_{t+1} - S_t e^{0.03}, 0)}{C_t[S_t e^{0.03}]} \end{pmatrix}, \tag{21}$$

where $S_t$ is the price of the equity market portfolio (proxied here by the S&P 500 index) and $P_t[S_t e^{-d}]$ ($C_t[S_t e^{d}]$) is the price of a put (call) on the market portfolio with strike that is $d\%$ out-of-the-money (OTM), for $d = 0.03$ or $d = 0.01$.

Following Cochrane (2005, pages 65 and 66), we estimate the projection coefficients in equation (21) as $\boldsymbol{\eta} = [-5.72 \;\; 6.63 \;\; 0.21 \;\; 0.07 \;\; -0.14 \;\; 0.04]$. ♣

Table 1 shows that the model in equation (21) generates positive entries of $m_{t,t+1}$ with an annualized volatility of 108% ($0.31 \times \sqrt{12}$). The distribution of $m_{t,t+1}$ is positively skewed and heavy tailed. This SDF manifests a decreasing region as well as an increasing region in the return of the market portfolio. The increasing region is a consequence of the positive projection coefficient on the 3% OTM call.[5]

For our exercises, we take a stand on the set of returns $\mathbf{R}_{t,t+1}$ that enter the construction of the lower bounds $\mathbb{LB}_m$ and $\mathbb{LB}_{m^2}$. Here, we take a collection of return time-series that capture various dimensions of the equity market universe, for example, those highlighted in, among others, Fama and French (2017). These observable characteristics, in addition to the market portfolio, include 25 portfolios formed on the basis of size, in conjunction with (1) book-to-market, (2) momentum, (3) accrual, (4) investment, (5) long-term reversal, (6) net issues, (7) operating profitability, (8) variance, and (9) residual variance.

Thus, we compute $\mathbb{LB}_m$ and $\mathbb{LB}_{m^2}$ using each of the nine sets of return portfolios comprising 26 assets and matched to our option sample of 1990:01 to 2015:12 (with 311 observations). The

[5]How do we reconcile the use of projected SDFs that include option returns? For one, there is some precedence in considering such models in the derivatives literature in which long and short equity positions exhibit exposures to the downside and the upside (e.g., Bakshi, Madan, and Panayotov (2010), Chabi-Yo (2012), and Christoffersen, Heston, and Jacobs (2013)). In addition, Vanden (2004) builds a model in which the agent faces nonnegative wealth constraints and shows that the SDF is a linear function of payoffs on puts and calls.

question is: If the model in equation (21) is deemed acceptable by the $\mathbb{LB}_m$ metric, is it rejected by $\mathbb{LB}_{m^2}$?

To gauge statistical significance, we formulate the following two null hypotheses:

$$L[m] \;-\; \mathbb{LB}_m \;\geq\; 0 \quad \text{and} \tag{22}$$

$$L[m^2] \;-\; \mathbb{LB}_{m^2} \;\geq\; 0, \tag{23}$$

which we test using a bootstrap procedure. First, we generate, with replacement, the return time-series of $(R_{t+1,f}\ \mathbf{Q}_{t,t+1}\ \mathbf{R}_{t,t+1})$. Next, in each bootstrap draw, we compute the projection coefficients $\boldsymbol{\eta}$ and compute $L[m]$ and $L[m^2]$ from the time-series of $m_{t,t+1}$. Finally, we compare $L[m]$ ($L[m^2]$) to the regenerated value of $\mathbb{LB}_m$ ($\mathbb{LB}_{m^2}$).

The bootstrap procedure is repeated 100,000 times, and we count the proportion of bootstrap draws that satisfy $L[m] - \mathbb{LB}_m \geq 0$ and $L[m^2] - \mathbb{LB}_{m^2} \geq 0$. We record this proportion as the empirical $p$-value in Table 1, and low $p$-values imply *rejection* of the null hypothesis. The bootstrap $p$-values reported in Table 1 indicate that one *cannot reject* that this model passes $\mathbb{LB}_m$ (the minimum bootstrap $p$-value is 0.374) but is rejected in seven out of nine instances according to $\mathbb{LB}_{m^2}$ (using the $p$-value of 0.1 as the cutoff value).

**Example 2** [*SDF is exponentially linear in returns based on Fama and French (2015)*]: Let $\mathbf{er}_{t,t+1} \equiv \mathbf{Q}_{t,t+1} - \mathbf{1}R_{t+1,f}$ be a $9 \times 1$ vector of excess returns, where $\mathbf{Q}_{t,t+1}$ is based on market, plus the low and high extreme portfolios featured in Fama and French (2015). Our model is

$$m_{t,t+1} \;=\; \exp(\lambda_0 - 1 + \boldsymbol{\lambda}'\mathbf{er}_{t,t+1}), \tag{24}$$

where the constants $(\lambda_0, \boldsymbol{\lambda}) \in \mathbb{R}^{10}$ solve the minimization problem $\inf_{(\lambda_0,\boldsymbol{\lambda})} -E[q_t]\lambda_0 + E[\exp(\lambda_0 - 1 + \boldsymbol{\lambda}'\mathbf{er}_{t,t+1})]$. This form of SDF can be traced to a minimum discrepancy problem described in Borovička, Hansen, and Scheinkman (2016), Almeida and Garcia (2017), and Bakshi, Chabi-Yo, and Gao (2017), where the objective is to minimize $E[m\log(m)]$ subject to $E[m\,\mathbf{er}] = \mathbf{0}$ and $E[m - E[q_t]] = 0$. The Lagrange multipliers on the constraints are $\boldsymbol{\lambda}$ and $\lambda_0$, respectively. ♣

Table 2 strengthens our arguments by showing that the null hypothesis of $L[m] \geq \mathbb{LB}_m$ is not rejected with a minimum bootstrap $p$-value of 0.289. In contrast, based on the reported bootstrap $p$-values, we can reject that $L[m^2] = 0.1207$ is greater than $\mathbb{LB}_{m^2}$ in eight out of nine instances. This model generates a positive skewness of 1.1 and a kurtosis of 9.8.

**Example 3** [*SDF is exponentially linear in the baseline returns in Fama and French (2015), augmented with momentum portfolios*]: Let $\mathbf{er}^{\text{aug}}_{t,t+1} \equiv \mathbf{Q}_{t,t+1} - \mathbf{1}R_{t+1,f}$ be a $11 \times 1$ vector of excess returns. Here $\mathbf{Q}_{t,t+1}$ is based on market, the low and high extreme portfolios featured in Fama and French (2015), and the portfolio of extreme past losers and winners. More concretely,

$$m_{t,t+1} \;=\; \exp(\lambda_0 - 1 + \boldsymbol{\lambda}'\mathbf{er}^{\text{aug}}_{t,t+1}), \tag{25}$$

where $(\lambda_0, \boldsymbol{\lambda}) \in \mathbb{R}^{12}$ solves the minimization problem outlined in Example 2. ♣

Table 3 is telling about the pricing implications of incorporating the momentum portfolios. First, this model produces sufficient $L[m]$: the point estimate of 0.0569 always exceeds the lower bounds anchored to the nine portfolios. Observe further that the bootstrap $p$-values rise across the board versus the Table 2 counterparts and support a minimum value of 0.704. Moreover, the addition of momentum portfolios almost doubles the model-based entropy of $m^2$ to 0.2428, consistent with a time-series of $m_{t,t+1}$, which manifests pronounced deviations from lognormality.

The takeaway is that the model in equation (25) is associated with reduced likelihood of rejecting the null hypothesis $L[m^2] - \mathbb{LB}_{m^2} \geq 0$, with four out of nine $p$-values now *less* than 0.1 (with size and momentum, size and net issues, size and variance, and size and residual variance). From another perspective, the $p$-values are about three times larger than those in Table 2.

### *3.2. Models that pass the bound on $L[m^2]$ also pass the bound on $L[m]$*

In this section, we consider an additional experiment in which a candidate SDF is chosen to automatically satisfy the lower bound on $L[m^2]$. We then investigate if the same model of the SDF fails to pass the lower bound on $L[m]$.

To pursue our line of inquiry, we return to Problem 1 and consider the following null hypothesis:

$$L[m^\bullet] \; - \; \mathbb{LB}_m \geq 0 \quad \text{(where } m^\bullet \text{ is consistent with } L[(m^\bullet)^2] = \mathbb{LB}_{m^2}\text{)}. \tag{26}$$

We then employ the following steps and a bootstrap procedure to assess statistical significance:

- First, using a $12 \times 1$ vector of gross returns (the same as those adopted in Table 3), we estimate $\nu$ and $\psi$ using equations (C13) and (C14) and construct the SDF $m^\bullet$ according to (19). The resulting $m^\bullet$, which is consistent with $L[(m^\bullet)^2] = \mathbb{LB}_{m^2}$, supports a volatility (monthly) of 0.35.

- Second, to test whether $L[m^\bullet] - \mathbb{LB}_m \geq 0$, we construct $\mathbb{LB}_m$ based on a different set of gross returns (e.g., risk-free return, the market, and 10 portfolios sorted on momentum).

- Third, we jointly generate, with replacement, all the returns used in the construction of $m^\bullet$ (we again estimate $\nu$ and $\psi$) and $\mathbb{LB}_m$. We then compare $\mathbb{LB}_m$ to the model-based entropy $L[m^\bullet]$ in each bootstrap draw. If the criterion in equation (26) is satisfied, we assign it a value of one.

The bootstrap procedure is repeated 100,000 times, and we count the proportion of bootstrap draws that satisfy $L[m^\bullet] - \mathbb{LB}_m \geq 0$. We record this proportion as the empirical $p$-value, and low $p$-values imply *rejection* of the null hypothesis. The key conclusion from Table 4 is that the constructed model that passes the lower bound on $L[m^2]$ also passes the lower bound on $L[m]$. Across the nine different sets of returns used to construct $\mathbb{LB}_{m^2}$, the lowest $p$-value is 0.764.

### *3.3. Models that pass the bound on $L[m]$ need not pass the bound on $L[m^2]$*

Next, we deviate from the preceding analysis by *restricting* the model-based $L[m]$ to be equal to $\mathbb{LB}_m$. We reinforce the idea that models can pass the lower entropy bound on $m$ yet fail the lower entropy bound on $m^2$. To do so, we formalize the following problem.

**Problem 2** *Identify an SDF, denoted by $m^G$, satisfying $L[m^G] = \mathbb{LB}_m$, where $\mathbb{LB}_m$ is derived by using the pricing restriction that $m \in \mathbb{S}$.* ■

The relevance of Problem 2 stems from the feature that $m^{\mathrm{G}}$, by construction, respects the lower bound on $L[m]$ but leaves open the possibility that $m^{\mathrm{G}}$ can be rejected using the lower bound on $L[m^2]$. A solution to Problem 2 is (see Appendix F)

$$m^{\mathrm{G}}_{t,t+1} = \frac{\beta}{\left(\mathbf{a}'\mathbf{R}_{t,t+1}\right)^{\alpha}} \quad \text{for some parameters } \alpha \text{ and } \beta, \tag{27}$$

where $\alpha$ and $\beta$ sequentially solve the equations

$$0 = E[q_t] - \exp((\alpha - 1)E[\log(\mathbf{a}^{'}\mathbf{R}_{t,t+1})]) \times E[(\mathbf{a}^{'}\mathbf{R}_{t,t+1})^{-\alpha}] \quad \text{and} \tag{28}$$

$$0 = \beta - \exp((\alpha - 1)\, E[\log(\mathbf{a}^{'}\mathbf{R}_{t,t+1})]). \tag{29}$$

The class of SDFs proposed in equation (27) is a power function of the portfolio return $\mathbf{a}^{'}\mathbf{R}_{t,t+1}$ with exponent $\alpha$ and absent the cross-product terms between asset returns in $\mathbf{R}_{t,t+1}$. The solution relies on the feature that $L[m^{\mathrm{G}}]$ equates to $\mathbb{LB}_m$ and $E[q_t] = E[m^{\mathrm{G}}_{t,t+1}]$. Since $\mathbf{R}_{t,t+1}$ is arbitrary, the set of $m^{\mathrm{G}}$ satisfying Problem 2 is potentially large.

Turning to our empirical exercises and implementations, Table 5 considers the same $12 \times 1$ vector of gross returns to construct $m^{\mathrm{G}}$ as that used in Tables 3 and 4. Next, we solve equations (28) and (29), which guarantees that $L[m^{\mathrm{G}}]$ matches $\mathbb{LB}_m$. In so doing, we circumvent moving parts, whereby the considered model satisfies the lower bound on $L[m]$ by construction. With this protocol for a comparison, we investigate whether the model passes the lower bound on $L[m^2]$.

Table 5 shows that exponent $\alpha$ is 2.920, whereas the centering coefficient, $\beta$, is 1.051. To evaluate statistical significance, we formulate the null hypothesis as

$$L[(m^{\mathrm{G}})^2] \;-\; \mathbb{LB}_{m^2} \geq 0, \quad (\text{where } m^{\mathrm{G}} \text{ is consistent with } L[m^{\mathrm{G}}] = \mathbb{LB}_m). \tag{30}$$

We test $L[(m^{\mathrm{G}})^2] - \mathbb{LB}_{m^2} \geq 0$ using a procedure that compares model-based entropy $L[(m^G)^2]$ to $\mathbb{LB}_{m^2}$ in each bootstrap draw. If the criterion in equation (30) is satisfied, we assign it a value of one. The procedure is repeated 100,000 times, and we count the proportion of bootstrap draws that satisfy $L[(m^G)^2] - \mathbb{LB}_{m^2} \geq 0$. We record this proportion as the empirical $p$-value.

The message from Table 5 is that it is feasible for a model to pass the lower bound on $L[m]$ and simultaneously fail to pass the lower bound on $L[m^2]$. Overall, our exercises reveal the relevance of $L[m^2]$ as a tool to empirically evaluate an asset pricing model.

What could be a possible rationale for our findings? First, $L[m^2]$ puts more weight on the higher-order moments of the SDF distribution. Second, we show that $L[m^2] - L[m]$ captures the tail behavior of the SDF distribution. Loosely speaking, the bound on $L[m^2]$ is relatively tighter from the vantage point of testing asset pricing models.

### *3.4. Additional clarifications*

We offer additional clarifications regarding our findings in Tables 1 through 5.

**A. Problem 1 and Problem 2 and connections to Hansen and Jagannathan (1991):** Our Table 4 (Table 5) considers $m^{\bullet}$ ($m^{\mathrm{G}}$) that meet the restriction on $\mathbb{LB}_{m^2}$ ($\mathbb{LB}_m$) but poses a testable hypothesis regarding the restriction on $\mathbb{LB}_m$ ($\mathbb{LB}_{m^2}$). We underscore that the returns used to build $m^{\bullet}$ ($m^{\mathrm{G}}$) are not the same as those used to construct $\mathbb{LB}_m$ ($\mathbb{LB}_{m^2}$). Our approach aligns with Hansen and Jagannathan (1991) (see our Internet Appendix (Section A)), who propose a minimum variance SDF consistent with a lower bound on the variance of SDFs when SDFs correctly price a set of returns.

**B. Robustness of our findings in Tables 4 and 5:** We consider an alternative setting in which the number of assets used to construct $\mathbb{LB}_m$ ($\mathbb{LB}_{m^2}$) is increased to 27 as opposed to 12 in Table 4 (Table 5). At the same time, we use the returns from the same well-understood 12 portfolios to construct the SDFs $m^{\bullet}_{t,t+1}$ and $m^{\mathrm{G}}_{t,t+1}$. The bootstrap $p$-values in Table 6 indicate that the hypothesis $L[m^{\bullet}] - \mathbb{LB}_m \geq 0$ is not rejected whereas $L[(m^{\mathrm{G}})^2] - \mathbb{LB}_{m^2} \geq 0$ is rejected.

**C. Other asset pricing models:** The criterion for model assessment proposed in our paper is adept at flagging models when the $m$ underlying a model is not lognormally distributed. Theoretically, when $\log(m)$ is normally distributed, $L[m^2] = 4L[m]$. In such an economic environment, models that fail the bound on $L[m]$ can be expected to fail the bound on $L[m^2]$. The Internet Appendix (Section C, Table Internet Appendix-II) highlights this feature in the context of some models, showing that the setting of normally distributed shocks does not offer a proper playing field to understand the distinctions between $L[m^2]$ and $L[m]$.[6] This prompted our approaches outlined in Tables 1 through 5 and emphasize the merits of using $L[m^2]$ on empirical grounds.

# 4. Conclusions

A central problem in finance is the specification of the SDF (denoted by $m_{t,t+1}$). We study this problem by providing new asset pricing restrictions that are based on the entropy of $m_{t,t+1}^2$.

In our analysis, we address the conceptual differences between the unconditional entropy of $m_{t,t+1}^2$ versus the unconditional entropy of $m_{t,t+1}$. In particular, we establish that the unconditional entropy of $m_{t,t+1}^2$ is the upper bound on the unconditional expected excess (log) gross return of a security that pays $m_{t,t+1}$. The entropy restrictions we develop are based on the ability of the SDF to *jointly* price the risk-free bond and a set of risky assets.

Our focus is on showing that a model can meet the lower bound on the entropy of $m_{t,t+1}$ and can be rejected based on the lower bound on the entropy of $m_{t,t+1}^2$. Thus, the unconditional

[6] Colacito, Ghysels, Meng, and Siwasarit (2016) propose a model of log SDF that depends on innovation in consumption growth by modeling the growth rate of consumption as a skew-normal variable with time-varying parameters. They show that the introduction of time-varying skewness substantially increases equity risk premiums and produces sizable variation in conditional risk premiums. The SDF in their model is conditionally lognormal, but the unconditional distribution of the SDF is not lognormal.

entropy restrictions on $m_{t,t+1}^2$ proposed in this paper allow one to flag asset pricing models that are not flagged according to the entropy of $m_{t,t+1}$. Our approach can be extended to incorporate conditioning information.

# References

Almeida, C., Garcia, R., 2017. Robust economic implications of nonlinear pricing kernels. Management Science 63, 3361–3380.

Alvarez, F., Jermann, U., 2005. Using asset prices to measure the persistence of the marginal utility of wealth. Econometrica 73, 1977–2016.

Backus, D., Chernov, M., Zin, S., 2014. Sources of entropy in representative agent models. Journal of Finance 69, 51–99.

Bakshi, G., Chabi-Yo, F., 2012. Variance bounds on the permanent and transitory components of stochastic discount factors. Journal of Financial Economics 105, 191–208.

Bakshi, G., Chabi-Yo, F., Gao, X., 2017. A recovery that we can trust? Deducing and testing the restrictions of the recovery theorem. Review of Financial Studies (forthcoming).

Bakshi, G., Madan, D., Panayotov, G., 2010. Returns of claims on the upside and the viability of U-shaped pricing kernels. Journal of Financial Economics 97, 130–154.

Bansal, R., Lehmann, B., 1997. Growth-optimal portfolio restrictions on asset pricing models. Macroeconomic Dynamics 1, 333–354.

Bansal, R., Yaron, A., 2004. Risks for the long run: A potential resolution of asset pricing puzzles. Journal of Finance 59, 1481–1509.

Bekaert, G., Engstrom, E., 2017. Asset return dynamics under bad environment-good environment fundamentals. Journal of Political Economy 125, 713–760.

Borovička, J., Hansen, L., Scheinkman, J., 2016. Misspecified recovery. Journal of Finance 71, 2493–2544.

Campbell, J., Cochrane, J., 1999. Force of habit: A consumption-based explanation of aggregate stock market behavior. Journal of Political Economy 107, 205–251.

Campbell, J., Vuolteenaho, T., 2004. Bad beta, good beta. American Economic Review 94, 1249–1275.

Chabi-Yo, F., 2012. Pricing kernels with stochastic skewness and volatility risk. Management Science 58, 624–640.

Chan, Y., Kogan, L., 2002. Catching up with the Joneses: Heterogeneous preferences and the dynamics of asset prices. Journal of Political Economy 110, 1255–1285.

Chapman, D., 1998. Habit formation and aggregate consumption. Econometrica 66, 1223–1230.

Christensen, T., 2017. Nonparametric stochastic discount factor decomposition. Econometrica 86, 1501–1536.

Christoffersen, P., Heston, S., Jacobs, K., 2013. Capturing option anomalies with a variance-dependent pricing kernel. Review of Financial Studies 26, 1963–2006.

Cochrane, J., 2005. Asset Pricing. Princeton University Press, Princeton, NJ.

Colacito, R., Ghysels, E., Meng, J., Siwasarit, W., 2016. Skewness in expected macro fundamentals and the predictability of equity returns: Evidence and theory. Review of Financial Studies 20, 2069–2109.

Epstein, L., Zin, S., 1991. Substitution, risk aversion, and the temporal behavior of consumption and asset returns: An empirical analysis. Journal of Political Economy 99, 263–286.

Fama, E., French, K., 2015. A five-factor asset pricing model. Journal of Financial Economics 116, 1–22.

Fama, E., French, K., 2017. Choosing factors. Unpublished working paper. University of Chicago and Dartmouth College.

Ghosh, A., Julliard, C., Taylor, A., 2017. What is the consumption-CAPM missing? An information-theoretic framework for the analysis of asset pricing models. Review of Financial Studies 30, 442–504.

Hansen, L., Heaton, J., Li, N., 2008. Consumption strikes back? Measuring long-run risk. Journal of Political Economy 116, 260–302.

Hansen, L., Jagannathan, R., 1991. Implications of security market data for dynamic economies. Journal of Political Economy 99, 225–261.

Hansen, L., Scheinkman, J., 2009. Long-term risk: An operator approach. Econometrica 77, 177–234.

Liu, Y., 2014. Index option returns and generalized entropy bounds. Unpublished working paper. SSRN # 2149265.

Orlowski, P., Sali, A., Trojani, F., 2016. Arbitrage free dispersion. Unpublished working paper. University of Lugano and University of Geneva.

Pohl, W., Schmedders, K., Wilms, O., 2015. Higher-order effects in asset-pricing models with long-run risks. Unpublished working paper. University of Zurich.

Qin, L., Linetsky, V., Nie, Y., 2016. Long forward probabilities, recovery and the term structure of bond risk premiums. Unpublished working paper (http://ssrn.com/abstract=271366). Northwestern University.

Santos, T., Veronesi, P., 2010. Habit formation, the cross-section of stock returns, and the cash flow risk puzzle. Journal of Financial Economics 98, 385–413.

Snow, K., 1991. Diagnosing asset pricing models using the distribution of asset returns. Journal of Finance 46, 955–983.

Stutzer, M., 1995. A Bayesian approach to diagnosis of asset pricing models. Journal of Finance 68, 367–397.

van Zwet, W., 1966. Convex Transformations of Random Variables. Mathematical Centre Tracts, Amsterdam.

Vanden, J., 2004. Options trading and the CAPM. Review of Financial Studies 17, 207–238.

Wachter, J., 2013. Can time-varying risk of rare disasters explain aggregate stock market volatility?. Journal of Finance 68, 987–1035.

Zhou, G., Zhu, Y., 2009. A long-run risks model with long-and short-run volatilities: Explaining predictability and volatility risk premium. Unpublished working paper. Washington University.

## Appendix A: Proof of Result 1

We adopt the following notation to streamline equation presentation and the steps of the proof:

$$\mathbf{y} \equiv \Sigma^{-1} \left(\mathbf{1} - E\left[q_t\right] E\left[\mathbf{R}_{t,t+1}\right]\right) \qquad \text{and} \qquad \mathbf{a} \equiv \frac{\mathbf{y}}{\mathbf{1}'\mathbf{y}}, \tag{A1}$$

where $q_t = E_t\left[m_{t,t+1}\right]$ and $\Sigma$ is the variance-covariance matrix of $\mathbf{R}_{t,t+1}$. We assume that $\mathbf{a}'\mathbf{R}_{t,t+1}$ is strictly positive. We define

$$\mathrm{er}_R \ \equiv \ E[\log(\mathbf{a}'\mathbf{R}_{t,t+1})] \ - \ \log((E[q_t])^{-1}). \tag{A2}$$

**Proof of the entropy bound on $m_{t,t+1}^2$ in equation (12)**. By the definition of entropy: $L[m^2] = \log(E[m^2]) - E[\log(m^2)]$. Then

$$\begin{aligned}
L[m_{t,t+1}^2] &= \log\left(E[m_{t,t+1}^2]\right) - 2\log\left(E\left[q_t\right]\right) + 2L\left[m_{t,t+1}\right], \\
&= \log\left(1 + \frac{E[m_{t,t+1}^2] - \left(E\left[q_t\right]\right)^2}{\left(E\left[q_t\right]\right)^2}\right) + 2L\left[m_{t,t+1}\right], \\
&= \log\left(1 + \frac{\mathrm{var}\left[m_{t,t+1}\right]}{\left(E[q_t]\right)^2}\right) + 2L\left[m_{t,t+1}\right], \\
&\geq \log\left(1 + \frac{\mathrm{var}[m_{t,t+1}]}{\left(E\left[q_t\right]\right)^2}\right) + 2\,\mathrm{er}_R. \quad (\text{as } L[m] \geq \mathrm{er}_R;\ \text{see } (A2) \text{ and } (E5))
\end{aligned} \tag{A3}$$

Because $E\left[m_{t,t+1}\mathbf{R}_{t,t+1}\right] = \mathbf{1}$ and setting $q_t = E_t\left[m_{t,t+1}\right]$, it follows that

$$E\left[m_{t,t+1}\left(\mathbf{R}_{t,t+1} - E\left(\mathbf{R}_{t,t+1}\right)\right)\right] = \left(\mathbf{1} - E\left[q_t\right] E\left[\mathbf{R}_{t,t+1}\right]\right). \tag{A4}$$

Multiplying equation (A4) by $(\mathbf{1}-(E[q_t])E[\mathbf{R}_{t,t+1}])'\Sigma^{-1}$ yields

$$(\mathbf{1}-(E[q_t])E[\mathbf{R}_{t,t+1}])'\Sigma^{-1}(\mathbf{1}-E[q_t]E[\mathbf{R}_{t,t+1}]) = E\left[m_{t,t+1}(\mathbf{1}-(E[q_t])E[\mathbf{R}_{t,t+1}])'\Sigma^{-1}(\mathbf{R}_{t,t+1}-E[\mathbf{R}_{t,t+1}])\right]. \tag{A5}$$

Applying the Cauchy-Schwarz to the right-hand side of equation (A5), it follows that

$$\begin{aligned}
\mathrm{var}[m_{t,t+1}] &\geq (\mathbf{1}-(E[q_t])E[\mathbf{R}_{t,t+1}])'\Sigma^{-1}(\mathbf{1}-(E[q_t])E[\mathbf{R}_{t,t+1}]), \\
&\geq (\mathbf{1}-(E[q_t])E[\mathbf{R}_{t,t+1}])'\Sigma^{-1}\Sigma\Sigma^{-1}(\mathbf{1}-(E[q_t])E[\mathbf{R}_{t,t+1}]), \\
&\geq \mathbf{y}'\Sigma\mathbf{y}.
\end{aligned} \tag{A6}$$

Combining the expressions in equations (A3) and (A6), we obtain the bound on $L[m^2_{t,t+1}]$ presented in equation (12) of Result 1. ■

## Appendix B: Proof of equation (18) of Result 2

The gross return of the security with SDF payoff is $1+r^{\mathrm{SDF}}_{t,t+1} = \frac{m_{t,t+1}}{E_t[m^2_{t,t+1}]}$, and it satisfies the Euler equation with $E_t[m_{t,t+1}(1+r^{\mathrm{SDF}}_{t,t+1})] = 1$.

There are two parts of the proof of equation (18). First, we establish that $L[m^2] \geq E[\log(R_{t+1,f})] - E[\log(1+r^{\mathrm{SDF}}_{t,t+1})]$. Next, we show that the security that pays the SDF is a hedging asset with expected return lower than the risk-free return. Let $L_t[m]$ denote the conditional entropy.

Taking logs of the expression for $1+r^{\mathrm{SDF}}_{t,t+1}$ and, adding and subtracting $\log(m^2_{t,t+1})$, we obtain

$$\log(1+r^{\mathrm{SDF}}_{t,t+1}) = \log(m_{t,t+1}) - \log(E_t[m^2_{t,t+1}]) + \log(m^2_{t,t+1}) - 2\log(m_{t,t+1}). \tag{B1}$$

Netting out $\log(m_{t,t+1})$ and taking expectations on both sides of equation (B1), we have

$$E_t[\log(1+r^{\text{SDF}}_{t,t+1})] + E_t[\log(m_{t,t+1})] \;=\; \overbrace{E_t[\log(m^2_{t,t+1})] - \log(E_t[m^2_{t,t+1}])}^{-L_t[m^2_{t,t+1}] \text{ from eq. (7)}}. \tag{B2}$$

Subtracting $\log(E_t[m_{t,t+1}])$ from both sides of equation (B2) and rearranging, it follows that

$$\begin{aligned} L_t[m^2_{t,t+1}] + \log(E_t[m_{t,t+1}]) &= -E_t[\log(1+r^{\text{SDF}}_{t,t+1})] + \underbrace{\log(E_t[m_{t,t+1}]) - E_t[\log(m_{t,t+1})]}_{L_t[m_{t,t+1}]\geq 0} && \text{(B3)} \\ &\geq -E_t[\log(1+r^{\text{SDF}}_{t,t+1})]. && \text{(B4)} \end{aligned}$$

Rearranging, we obtain the following expression for the conditional entropy of $m^2_{t,t+1}$:

$$L_t[m^2_{t,t+1}] \;\geq\; -E_t[\log(1+r^{\text{SDF}}_{t,t+1})] \;-\; \log(E_t[m_{t,t+1}]). \tag{B5}$$

Since the gross return of the risk-free bond satisfies $E_t\left[m_{t,t+1}\right] = 1/R_{t+1,f}$, we obtain

$$L_t[m^2_{t,t+1}] \;\geq\; \log(R_{t+1,f}) \;-\; E_t[\log(1+r^{\text{SDF}}_{t,t+1})]. \tag{B6}$$

Taking unconditional expectations on both sides of equation (B6),

$$E[L_t[m^2_{t,t+1}]] \;\geq\; E[\log(R_{t+1,f})] \;-\; E[\log(1+r^{\text{SDF}}_{t,t+1})]. \tag{B7}$$

Now exploit the following relation

$$E[L_t[m^2_{t,t+1}]] \;\leq\; L[m^2_{t,t+1}] \quad \text{since} \quad L[u^2] = E[L_t[u^2]] \;+\; L[E_t[u^2]] \;\text{ for any random variable } u. \tag{B8}$$

Therefore, $L[m_{t,t+1}^2] \geq E[\log(R_{t+1,f})] - E[\log(1+r_{t,t+1}^{\text{SDF}})]$. Our measure is tied to the maximum expected (log) return on a security that pays the SDF. ♣

Completing the picture, we need to show that the security with SDF payoff is a hedging asset. Observe that $E_t[m_{t,t+1}^2] \geq (E_t[m_{t,t+1}])^2$, because $\text{var}(m_{t,t+1}) > 0$. Hence,

$$E_t[1+r_{t,t+1}^{\text{SDF}}] = \frac{E_t[m_{t,t+1}]}{E_t[m_{t,t+1}^2]} \ \leq \ \frac{E_t[m_{t,t+1}]}{(E_t[m_{t,t+1}])^2} = \frac{1}{E_t[m_{t,t+1}]} = R_{t+1,f}. \tag{B9}$$

Equation (B9), thus, shows that $E_t[1+r_{t,t+1}^{\text{SDF}}] \leq R_{t+1,f}$, which implies that

$$\log\left(E_t\left[1+r_{t,t+1}^{\text{SDF}}\right]\right) \ \leq \ \overbrace{\log\left(\frac{1}{E_t[m_{t,t+1}]}\right)}^{\log(R_{t+1,f})}. \tag{B10}$$

By an application of Jensen's inequality,

$$E_t[\log(1+r_{t,t+1}^{\text{SDF}})] \ \leq \ \log\left(E_t\left[1+r_{t,t+1}^{\text{SDF}}\right]\right) \ \leq \ \log(R_{t+1,f}). \tag{B11}$$

It then follows that

$$\log(R_{t+1,f}) - E_t[\log(1+r_{t,t+1}^{\text{SDF}})] \geq 0. \quad \text{Therefore,} \quad E[\log(R_{t+1,f})] - E[\log(1+r_{t,t+1}^{\text{SDF}})] \geq 0.$$

The proof is complete. ■

## Appendix C: Proof of equation (19)

Our Result 1 implies that $L[m^2] \geq \mathbb{LB}_{m^2}$, where

$$\mathbb{LB}_{m^2} \equiv 2\left(E[\log(\mathbf{a}^{'}\mathbf{R}_{t,t+1})] - \log((E[q_t])^{-1})\right) + \log(1+\mathbf{y}^{'}\Sigma\mathbf{y}/(E[q_t])^2), \tag{C1}$$

with $\mathbf{a}$ and $\mathbf{y}$ displayed in equation (A1).

To investigate the restrictions under which the bound becomes binding, we need to show that there exists $m^{\bullet}_{t,t+1}$ with

$$1+r^{\bullet\mathrm{SDF}}_{t,t+1} \;=\; \frac{m^{\bullet}_{t,t+1}}{E[(m^{\bullet}_{t,t+1})^2]} \;\text{ such that }\; \mathbb{LB}_{m^2} = E[\log(\frac{R_{t+1,f}}{1+r^{\bullet\mathrm{SDF}}_{t,t+1}})]. \tag{C2}$$

Guided by Result 1, the identity (C2) becomes

$$\mathbb{LB}_{m^2} \equiv E\left[\log\left(((E[q_t])\,\mathbf{a}^{'}\mathbf{R}_{t,t+1})^2(1+\mathbf{y}^{'}\Sigma\mathbf{y}/(E[q_t])^2)\right)\right] = E[\log(\frac{R_{t+1,f}}{1+r^{\bullet\mathrm{SDF}}_{t,t+1}})]. \tag{C3}$$

Upon simplification,

$$\frac{R_{t+1,f}}{1+r^{\bullet\mathrm{SDF}}_{t,t+1}} = (\mathbf{a}^{'}\mathbf{R}_{t,t+1})^2\left((E[q_t])^2+\mathbf{y}^{'}\Sigma\mathbf{y}\right). \tag{C4}$$

Thus,

$$1+r^{\bullet\mathrm{SDF}}_{t,t+1} = \frac{R_{t+1,f}\,\mho^{\bullet}}{\left(\mathbf{a}^{'}\mathbf{R}_{t,t+1}\right)^2} \qquad \text{with} \qquad \mho^{\bullet} \equiv \frac{1}{(E[q])^2+\mathbf{y}^{'}\Sigma\mathbf{y}}. \tag{C5}$$

With the return of the SDF security of the form $\frac{m^{\bullet}_{t,t+1}}{E[(m^{\bullet}_{t,t+1})^2]}$, one needs to find an SDF $m^{\bullet}_{t,t+1}$ that satisfies

$$\frac{\mho^{\bullet}\,R_{t+1,f}}{\left(\mathbf{a}^{'}\mathbf{R}_{t,t+1}\right)^2} = \frac{m^{\bullet}_{t,t+1}}{E[(m^{\bullet}_{t,t+1})^2]}. \tag{C6}$$

We conjecture and then verify that a solution is as given in equation (19). Direct substitution implies that

$$\frac{m^{\bullet}_{t,t+1}}{E[(m^{\bullet}_{t+1})^2]} = \frac{\frac{1}{(\mathbf{a}'\mathbf{R}_{t,t+1})^{\nu}}}{E\left[\frac{\psi}{(\mathbf{a}'\mathbf{R}_{t,t+1})^{2\nu}}\right]}. \tag{C7}$$

With this step, we replace the SDF in identity (C6) and obtain

$$\frac{\frac{1}{(\mathbf{a}'\mathbf{R}_{t,t+1})^{\nu}}}{E\left[\frac{\beta}{(\mathbf{a}'\mathbf{R}_{t,t+1})^{2\nu}}\right]} = \frac{\mho^{\bullet} R_{t+1,f}}{(\mathbf{a}'\mathbf{R}_{t,t+1})^2}. \tag{C8}$$

Proceeding we take logs and then apply the expectations operator. Thus, it follows that

$$E\left[\log\left(\frac{\frac{1}{(\mathbf{a}'\mathbf{R}_{t,t+1})^{\nu}}}{E\left[\frac{1}{(\mathbf{a}'\mathbf{R}_{t,t+1})^{2\nu}}\right]}\right)\right] - \log(\psi) = E\left[\log\left(\frac{\mho^{\bullet} R_{t+1,f}}{(\mathbf{a}'\mathbf{R}_{t,t+1})^2}\right)\right]. \tag{C9}$$

To enforce the correct pricing of the risk-free return, it must hold that

$$E[q_t] = \psi E\left[\frac{1}{(\mathbf{a}'\mathbf{R}_{t,t+1})^{\nu}}\right]. \text{ Hence, } \log(\psi) = \log\left(E\left[q_t\right]\right) - \log\left(E\left[\frac{1}{(\mathbf{a}'\mathbf{R}_{t,t+1})^{\nu}}\right]\right). \tag{C10}$$

With the aid of equations (C10) and (C9), we determine that

$$E\left[\log\left(\frac{\frac{1}{(\mathbf{a}'\mathbf{R}_{t,t+1})^{\nu}}}{E\left[\frac{1}{(\mathbf{a}'\mathbf{R}_{t,t+1})^{2\nu}}\right]}\right)\right] - \log\left(E\left[q_t\right]\right) + \log\left(E\left[\frac{1}{(\mathbf{a}'\mathbf{R}_{t,t+1})^{\nu}}\right]\right) = E\left[\log\left(\frac{\mho^{\bullet} R_{t+1,f}}{(\mathbf{a}'\mathbf{R}_{t,t+1})^2}\right)\right]. \tag{C11}$$

This expression simplifies to

$$E\left[\log\left(\frac{1}{\left(\mathbf{a}'\mathbf{R}_{t,t+1}\right)^{\nu}}\right)\right]-\log\left(E\left[\frac{1}{\left(\mathbf{a}'\mathbf{R}_{t,t+1}\right)^{2\nu}}\right]\right)-\log\left(E\left[q_t\right]\right)$$
$$+\log\left(E\left[\frac{1}{\left(\mathbf{a}'\mathbf{R}_{t,t+1}\right)^{\nu}}\right]\right)-E\left[\log\left(\frac{\mho^{\bullet}R_{t+1,f}}{\left(\mathbf{a}'\mathbf{R}_{t,t+1}\right)^{2}}\right)\right]=0. \tag{C12}$$

Therefore, we determine $\nu$ and $\psi$ as solutions to the equations

$$E\left[\log\left(\frac{1}{\left(\mathbf{a}'\mathbf{R}_{t,t+1}\right)^{\nu}}\right)\right]-\log\left(E\left[\frac{1}{\left(\mathbf{a}'\mathbf{R}_{t,t+1}\right)^{2\nu}}\right]\right)-\log\left(E\left[q_t\right]\right)$$
$$+\log\left(E\left[\frac{1}{\left(\mathbf{a}'\mathbf{R}_{t,t+1}\right)^{\nu}}\right]\right)-E\left[\log\left(\frac{\mho^{\bullet}R_{t+1,f}}{\left(\mathbf{a}'\mathbf{R}_{t,t+1}\right)^{2}}\right)\right]=0, \tag{C13}$$

and

$$E\left[\log\left(\frac{\frac{1}{\left(\mathbf{a}'\mathbf{R}_{t,t+1}\right)^{\nu}}}{E\left[\frac{1}{\left(\mathbf{a}'\mathbf{R}_{t,t+1}\right)^{2\nu}}\right]}\right)\right]-\log\left(E\left[q_t\right]\right)+\log\left(E\left[\frac{1}{\left(\mathbf{a}'\mathbf{R}_{t,t+1}\right)^{\nu}}\right]\right)-E\left[\log\left(\frac{\mho^{\bullet}R_{t+1,f}}{\left(\mathbf{a}'\mathbf{R}_{t,t+1}\right)^{2}}\right)\right]=0. \tag{C14}$$

In this case, the lower bound on $L[m^2]$ becomes binding. ■

## Appendix D: Expression for the lower bound on $L[m^3]$

Our end objective is to derive a lower bound on the entropy of $m^3$ presented in equation (D14).

By definition, it follows that

$$\begin{aligned}
L[m_{t,t+1}^3] &= \log(E\left[m_{t,t+1}^3\right]) - E[\log\left(m_{t,t+1}^3\right)], && \text{(D1)}\\
&= \log(E[m_{t,t+1}^3]) - 3\log(E[q_t]) + 3\log\left(E\left[q_t\right]\right) - 3E[\log(m_{t,t+1})], && \text{(D2)}\\
&= \log(E[m_{t,t+1}^3]) - 3\log(E[q_t]) + 3(\log\left(E\left[m_{t,t+1}\right]\right) - E[\log(m_{t,t+1})]), && \text{(D3)}\\
&= \log\left(E\left[m_{t,t+1}^3\right]\right) - 3\log\left(E\left[q_t\right]\right) \; + \; 3L[m_{t,t+1}], && \text{(D4)}\\
&= \log(E[(\frac{m_{t,t+1}}{E[m_{t,t+1}]})^3]) + 3L[m_{t,t+1}]. && \text{(D5)}
\end{aligned}$$

Exploiting equation (A4), we note that

$$E\left[m_{t,t+1}\left(\mathbf{R}_{t,t+1} - \mathbf{R}_{t,t+1}\right)\right] = \left(\mathbf{1} - E\left[q_t\right] E\left[\mathbf{R}_{t,t+1}\right]\right). \tag{D6}$$

By multiplying equation (D6) by $\frac{1}{E[m_{t,t+1}]}$, one obtains

$$E\left[\frac{m_{t,t+1}}{E\left[m_{t,t+1}\right]}\left(\mathbf{R}_{t,t+1} - E\left[\mathbf{R}_{t,t+1}\right]\right)\right] = \left(\frac{\mathbf{1}}{E\left[q_t\right]} - E\left[\mathbf{R}_{t,t+1}\right]\right). \tag{D7}$$

Now multiply equation (D7) by $\left(\frac{\mathbf{1}}{E[q_t]} - E\left[\mathbf{R}_{t,t+1}\right]\right)' \Sigma^{-1}$. As a consequence

$$\begin{aligned}
& E\left[\frac{m_{t,t+1}}{E\left[m_{t,t+1}\right]}\left(\frac{\mathbf{1}}{E\left[q_t\right]} - E\left[\mathbf{R}_{t,t+1}\right]\right)' \Sigma^{-1}\left(\mathbf{R}_{t,t+1} - E\left[\mathbf{R}_{t,t+1}\right]\right)\right] \\
= & \left(\frac{\mathbf{1}}{E\left[q_t\right]} - E\left[\mathbf{R}_{t+1}\right]\right)' \Sigma^{-1}\left(\frac{\mathbf{1}}{E\left[q_t\right]} - E\left[\mathbf{R}_{t,t+1}\right]\right). \qquad \text{(D8)}
\end{aligned}$$

Next apply the Hölder inequality to the left-hand side of equation (D8). When $p,q \in [1,\infty]$ and $\frac{1}{p}+\frac{1}{q}=1$, it is true that

$$
\begin{aligned}
& \left| E\left[\frac{m_{t,t+1}}{E\left[m_{t,t+1}\right]}\left(\frac{\mathbf{1}}{E\left[q_t\right]}-E\left[\mathbf{R}_{t,t+1}\right]\right)^{\prime}\Sigma^{-1}\left(\mathbf{R}_{t,t+1}-E\left[\mathbf{R}_{t,t+1}\right]\right)\right]\right| \\
\leq & \left(E\left[\left|\frac{m_{t,t+1}}{E\left[m_{t,t+1}\right]}\right|^{p}\right]\right)^{\frac{1}{p}}\left(E\left[\left|\left(\frac{\mathbf{1}}{E\left[q_t\right]}-E\left[\mathbf{R}_{t,t+1}\right]\right)^{\prime}\Sigma^{-1}\left(\mathbf{R}_{t,t+1}-E\left[\mathbf{R}_{t,t+1}\right]\right)\right|^{q}\right]\right)^{\frac{1}{q}}.
\end{aligned}
\tag{D9}
$$

Replace equation (D8) in the left-hand side of equation (D9). To be precise,

$$
\begin{aligned}
& \left|\left(\frac{\mathbf{1}}{E\left[q_t\right]}-E\left[\mathbf{R}_{t+1}\right]\right)^{\prime}\Sigma^{-1}\left(\frac{\mathbf{1}}{E\left[q_t\right]}-E\left[\mathbf{R}_{t+1}\right]\right)\right| \\
\leq & \left(E\left[\left|\frac{m_{t,t+1}}{E\left[m_{t,t+1}\right]}\right|^{p}\right]\right)^{\frac{1}{p}}\left(E\left[\left|\left(\frac{\mathbf{1}}{E\left[q_t\right]}-E\left[\mathbf{R}_{t+1}\right]\right)^{\prime}\Sigma^{-1}\left(\mathbf{R}_{t,t+1}-E\left[\mathbf{R}_{t,t+1}\right]\right)\right|^{q}\right]\right)^{\frac{1}{q}}.
\end{aligned}
\tag{D10}
$$

The implication is that

$$
\frac{\left|\left(\frac{\mathbf{1}}{E[q_t]}-E\left[\mathbf{R}_{t+1}\right]\right)^{\prime}\Sigma^{-1}\left(\frac{\mathbf{1}}{E[q_t]}-E\left[\mathbf{R}_{t+1}\right]\right)\right|}{\left(E\left[\left|\left(\frac{\mathbf{1}}{E[q_t]}-E\left[\mathbf{R}_{t+1}\right]\right)^{\prime}\Sigma^{-1}\left(\mathbf{R}_{t,t+1}-E\left[\mathbf{R}_{t,t+1}\right]\right)\right|^{q}\right]\right)^{\frac{1}{q}}} \leq \left(E\left[\left(\frac{m_{t,t+1}}{E\left[m_{t,t+1}\right]}\right)^{p}\right]\right)^{\frac{1}{p}}. \tag{D11}
$$

The following result then holds by setting $p=3$:

$$
\left\{\frac{\left|\left(\frac{\mathbf{1}}{E[q_t]}-E\left[\mathbf{R}_{t+1}\right]\right)^{\prime}\Sigma^{-1}\left(\frac{\mathbf{1}}{E[q_t]}-E\left[\mathbf{R}_{t+1}\right]\right)\right|}{\left(E\left[\left|\left(\frac{\mathbf{1}}{E[q_t]}-E\left[\mathbf{R}_{t+1}\right]\right)^{\prime}\Sigma^{-1}\left(\mathbf{R}_{t,t+1}-E\left[\mathbf{R}_{t,t+1}\right]\right)\right|^{\frac{3}{2}}\right]\right)^{\frac{2}{3}}}\right\}^{3} \leq E\left[\left(\frac{m_{t,t+1}}{E\left[m_{t,t+1}\right]}\right)^{3}\right]. \tag{D12}
$$

We can, therefore, bound the expression

$$L[m^3_{t,t+1}] \geq 3\log\left(\frac{\left|\left(\frac{\mathbf{1}}{E[q_t]}-E\left[\mathbf{R}_{t+1}\right]\right)^{'}\Sigma^{-1}\left(\frac{\mathbf{1}}{E[q_t]}-E\left[\mathbf{R}_{t+1}\right]\right)\right|}{\left(E\left[\left|\left(\frac{\mathbf{1}}{E[q_t]}-E\left[\mathbf{R}_{t+1}\right]\right)^{'}\Sigma^{-1}\left(\mathbf{R}_{t,t+1}-E\left[\mathbf{R}_{t,t+1}\right]\right)\right|^{\frac{3}{2}}\right]\right)^{\frac{2}{3}}}\right)+3L\left[m_{t,t+1}\right], \tag{D13}$$

which is equivalent to the expression for the bound on $L[m^3]$ shown below in equation (D14):

$$L[m^3_{t,t+1}] \geq 3\ \log\left(\frac{1}{E[q_t]}\frac{\mathbf{y}^{'}\Sigma\mathbf{y}}{\left(E\left[\left|\left(\mathbf{y}^{'}\mathbf{R}_{t,t+1}-E\left[\mathbf{y}^{'}\mathbf{R}_{t,t+1}\right]\right)\right|^{\frac{3}{2}}\right]\right)^{\frac{2}{3}}}\right)+6\,\mathrm{er}_R. \tag{D14}$$

This result holds since $L[m_{t,t+1}] \geq 2\mathrm{er}_R$ (see equation (A3)). ■

## Appendix E: Lower bound on $L[m]$

The central feature of the ensuing lower bound on the entropy of $m_{t,t+1}$ is that it is based on $m_{t,t+1}$ correctly price $N+1$ assets. Internet Appendix (Section B, Table Internet Appendix-I) illustrate the sharpness of these bounds.

We maintain the following notation (as in equation (A1)):

$$\mathbf{y} \equiv \Sigma^{-1}\left(\mathbf{1}-E\left[q_t\right]E\left[\mathbf{R}_{t,t+1}\right]\right), \qquad \text{and} \qquad \mathbf{a} \equiv \frac{\mathbf{y}}{\mathbf{1}'\mathbf{y}}, \tag{E1}$$

where $\Sigma$ is the variance-covariance matrix of $\mathbf{R}_{t,t+1}$.

**Proof of the bound on the entropy of $m_{t,t+1}$:** Consider the following return:

$$\begin{aligned} E[\log(m_{t,t+1}\mathbf{a}^{'}\mathbf{R}_{t,t+1})] &\leq \log(E[m_{t,t+1}\mathbf{a}^{'}\mathbf{R}_{t,t+1}]), \\ &\leq \log(\mathbf{a}^{'}E\left[m_{t,t+1}\mathbf{R}_{t,t+1}\right]), \\ &\leq \log(\mathbf{a}^{'}\mathbf{1}) = \log(1), \\ &\leq 0. \end{aligned} \tag{E2}$$

From equation (E2) and noting that $\log(m_{t,t+1}\mathbf{a}^{'}\mathbf{R}_{t,t+1}) = \log(m_{t,t+1}) + \log(\mathbf{a}^{'}\mathbf{R}_{t,t+1})$, we deduce that

$$E\left[\log\left(\mathbf{a}^{'}\mathbf{R}_{t,t+1}\right)\right] \leq -E\left[\log\left(m_{t,t+1}\right)\right]. \tag{E3}$$

Adding $\log\left(E\left[m_{t,t+1}\right]\right)$ to both sides of equation (E3) yields

$$\log\left(E\left[m_{t,t+1}\right]\right) + E[\log(\mathbf{a}^{'}\mathbf{R}_{t,t+1})] \;\leq\; \log(E[m_{t,t+1}]) - E\left[\log\left(m_{t,t+1}\right)\right] = L\left[m_{t,t+1}\right]. \tag{E4}$$

Since $q_t = E_t\left[m_{t,t+1}\right]$, equation (E4) simplifies to

$$L\left[m_{t,t+1}\right] \;\geq\; E[\log(\mathbf{a}^{'}\mathbf{R}_{t,t+1})] \;-\; \log\left(\frac{1}{E[q_t]}\right), \text{ where } \mathbf{a}^{'} \text{ is as defined in equation (E1).} \tag{E5}$$

The result depends on the variance-covariance matrix of returns and the mean return vector. ■

## Appendix F: Proof of solution to Problem 2

Observe that

$$L[m^{\mathrm{G}}_{t,t+1}] = E[\log(\frac{E[q_t]}{m^{\mathrm{G}}_{t,t+1}})], \tag{F1}$$

and it holds that $L[m^{\mathrm{G}}_{t,t+1}] = E[\log(\mathbf{a}^{'}\mathbf{R}_{t,t+1})] - \log((E[q_t])^{-1})$. This implies that

$$E\left[\log\left(\frac{\frac{\mathbf{a}^{'}\mathbf{R}_{t,t+1}}{(E[q_t])^{-1}}}{\frac{E[q_t]}{m^{\mathrm{G}}_{t,t+1}}}\right)\right] = 0. \tag{F2}$$

Hence,

$$E\left[\log(m^{\mathrm{G}}_{t,t+1}\mathbf{a}^{'}\mathbf{R}_{t,t+1})\right] = 0. \tag{F3}$$

Now conjecture that $m^{\mathrm{G}}_{t,t+1} = \frac{\beta}{\left(\mathbf{a}^{'}\mathbf{R}_{t,t+1}\right)^{\alpha}}$. To ensure the SDF prices correctly the risk-free return

$$E[q_t] = E[m^{\mathrm{G}}_{t,t+1}] = \beta E[(\mathbf{a}^{'}\mathbf{R}_{t,t+1})^{-\alpha}]. \tag{F4}$$

Next, to ensure that equation (F3) holds,

$$E[\log(\frac{\beta}{\left(\mathbf{a}^{'}\mathbf{R}_{t,t+1}\right)^{\alpha}}\mathbf{a}^{'}\mathbf{R}_{t,t+1})] = 0. \tag{F5}$$

This implies that

$$\log(\beta) - (\alpha-1)E[\log(\mathbf{a}^{'}\mathbf{R}_{t,t+1})] = 0. \text{ Rearranging, } \beta = \exp((\alpha-1)E[\log(\mathbf{a}^{'}\mathbf{R}_{t,t+1})]). \tag{F6}$$

Replace $\beta$ in equation (F4):

$$E\left[q_t\right] = \exp\Big((\alpha-1)E[\log(\mathbf{a}^{'}\mathbf{R}_{t,t+1})]\Big)\,E\left(\frac{1}{\left(\mathbf{a}^{'}\mathbf{R}_{t,t+1}\right)^{\alpha}}\right). \tag{F7}$$

We have the corroborated the form of the solution to Problem 2. ■

Table 1
**Model of SDF with returns of risk-free bond, equity market portfolio, and out-of-the-money puts and calls**

The sample period for this exercise is 1990:01 to 2015:12 (311 observations). The model is

$$m_{t,t+1} = \boldsymbol{\eta}'[R_{t+1,f} \;\; \mathbf{Q}_{t,t+1}].$$

The $5 \times 1$ vector of gross returns $\mathbf{Q}_{t,t+1}$ contains

- gross return of the market (S&P 500 index);
- gross return of a 3% out-of-the-money put on the S&P 500 index;
- gross return of a 1% out-of-the-money put on the S&P 500 index;
- gross return of a 1% out-of-the-money call on the S&P 500 index;
- gross return of a 3% out-of-the-money call on the S&P 500 index.

The projection coefficients are estimated to be (i.e., Cochrane (2005, pages 65 and 66))

$$\boldsymbol{\eta} = [-5.72 \;\; 6.63 \;\; 0.21 \;\; 0.07 \;\; -0.14 \;\; 0.04].$$

We present the monthly properties of $m$ and then $L[m]$ and $L[m^2]$. The lower bounds on $\mathbb{LB}_m$ and $\mathbb{LB}_{m^2}$ are extracted over the sample period 1990:01 to 2015:12. The reported $p$-values are based on a bootstrap and tests the null hypothesis that the model-based entropy is greater than the lower bound. In each bootstrap draw, we randomly select, with replacement, raw asset returns $\mathbf{R}_{t,t+1}$ together with $\mathbf{Q}_{t,t+1}$ and recompute the projection coefficients $\boldsymbol{\eta}$. Then we generate the time-series of $m_{t,t+1}$ according to equation (21) and compute the corresponding lower bounds according to equations (12) and (20). Here we perform 100,000 bootstrap trials. The set $\mathbb{S}$ in equation (11) is based on risk-free return plus a $26\times1$ vector of returns. The $p$-values in bold indicate rejection of the null hypothesis.

| Model | $L[m]$ | $L[m^2]$ | Properties of the SDF (monthly) | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | Mean | Std. | Skewness | Kurtosis | Min. | Max. |
| Returns of risk-free bond, equity, puts, and calls | 0.0314 | 0.1567 | 0.9976 | 0.31 | 3.7 | 21.4 | 0.65 | 2.85 |

| Set of assets in $\mathbb{S}$ | 1990:01–2015:12 | | | |
|---|---|---|---|---|
| $R_{t+1,f}$ plus market plus | $\mathbf{R}_{t,t+1}$ is of dimension $26 \times 1$ | | | |
| | $\mathbb{LB}_m$ | Bootstrap $p$-value | $\mathbb{LB}_{m^2}$ | Bootstrap $p$-value |
| 25 Size×B/M | 0.0298 | 0.572 | 0.3975 | **0.018** |
| 25 Size×Momentum | 0.0266 | 0.570 | 0.2668 | **0.057** |
| 25 Size×Accrual | 0.0196 | 0.652 | 0.1779 | 0.203 |
| 25 Size×Investment | 0.0368 | 0.395 | 0.3509 | **0.029** |
| 25 Size×Long-Term Reversal | 0.0352 | 0.374 | 0.2692 | **0.095** |
| 25 Size×Operating Profitability | 0.0226 | 0.633 | 0.2012 | 0.195 |
| 25 Size×Net Issues | 0.0348 | 0.433 | 0.4030 | **0.011** |
| 25 Size×Variance | 0.0204 | 0.767 | 0.3138 | **0.038** |
| 25 Size×Residual Variance | 0.0226 | 0.666 | 0.2762 | **0.050** |

Table 2

**Model where log SDF is linear in a set of baseline excess returns**

The sample period for this exercise is 1963:07 to 2017:06 (648 observations). The specific model we consider takes the form

$$m_{t,t+1} = \exp(\lambda_0 - 1 + \boldsymbol{\lambda}' \mathbf{er}_{t,t+1}),$$

where $\mathbf{er}_{t,t+1} \equiv \mathbf{Q}_{t,t+1} - \mathbf{1}R_{t+1,f}$ is a $9 \times 1$ vector of excess returns. The excess return $\mathbf{er}_{t,t+1}$ are based on $\mathbf{Q}_{t,t+1}$ from the five-factor model of Fama and French (2015):

- gross return of the market;
- gross returns of the two extreme low, and high, size portfolios;
- gross returns of the two extreme low, and high, book-to-market portfolios;
- gross returns of the two extreme low, and high, operating profitability portfolios;
- gross returns of the two extreme low, and high, investment portfolios.

The constants $(\lambda_0, \boldsymbol{\lambda})$ are a solution to the minimization problem $\inf_{(\lambda_0,\boldsymbol{\lambda})} -E[q_t]\lambda_0 + E[\exp(\lambda_0 - 1 + \boldsymbol{\lambda}'\mathbf{er}_{t,t+1})]$. This form of SDF obtains by minimizing the objective $E[m\log(m)]$ subject to $E[m\,\mathbf{er}] = \mathbf{0}$ and $E[m - E[q_t]] = 0$. The Lagrange multipliers on the equality constraints are $\boldsymbol{\lambda}$ and $\lambda_0$, respectively. Our procedure implies that

$$\lambda_0 = 1.01, \text{ and } \quad \boldsymbol{\lambda} = [36.3\ 0.0\ -0.2\ -13.8\ -10.0\ -0.1\ -0.2\ -10.1\ -4.2].$$

We present the monthly properties of $m$, and the model-based $L[m]$, and $L[m^2]$. The lower bounds on $\mathbb{LB}_m$ and $\mathbb{LB}_{m^2}$ are over the sample period 1963:07 to 2017:06. In each bootstrap draw, we randomly select, with replacement, raw asset returns $\mathbf{R}_{t,t+1}$ together with $\mathbf{Q}_{t,t+1}$ and recompute $(\lambda_0, \boldsymbol{\lambda})$. Then we generate the time-series of $m_{t,t+1}$, according to equation (25), and also compute the corresponding lower bounds according to equations (12) and (20). Here we perform 100,000 bootstrap trials. The $p$-values in bold indicate rejection of the null hypothesis.

| | $L[m]$ | $L[m^2]$ | Properties of the SDF (monthly): Mean | Std. | Skewness | Kurtosis | Min. | Max. |
|---|---|---|---|---|---|---|---|---|
| $m_{t,t+1} = \exp(\lambda_0 - 1 + \boldsymbol{\lambda}' \mathbf{er}_{t,t+1})$ | 0.0304 | 0.1207 | 0.988 | 0.25 | 1.1 | 9.8 | 0.33 | 2.31 |

| Set of assets in $\mathbb{S}$ <br> $R_{t+1,f}$ plus market plus | 1963:07–2017:06 <br> $\mathbf{R}_{t,t+1}$ is of dimension $26 \times 1$ | | | |
|---|---|---|---|---|
| | $\mathbb{LB}_m$ | Bootstrap $p$-value | $\mathbb{LB}_{m^2}$ | Bootstrap $p$-value |
| 25 Size×B/M | 0.0243 | 0.692 | 0.2311 | **0.039** |
| 25 Size×Momentum | 0.0352 | 0.337 | 0.2795 | **0.021** |
| 25 Size×Accrual | 0.0238 | 0.644 | 0.2011 | **0.052** |
| 25 Size×Investment | 0.0279 | 0.569 | 0.2589 | **0.028** |
| 25 Size×Long-Term Reversal | 0.0176 | 0.745 | 0.0944 | **0.086** |
| 25 Size×Operating Profitability | 0.0402 | 0.753 | 0.3413 | 0.201 |
| 25 Size×Net Issues | 0.0191 | 0.289 | 0.1281 | **0.013** |
| 25 Size×Variance | 0.0332 | 0.435 | 0.3115 | **0.017** |
| 25 Size×Residual Variance | 0.0374 | 0.357 | 0.3399 | **0.015** |

Table 3

**Model where log SDF is linear in a set of baseline excess returns, augmented with momentum portfolios**

The sample period for this exercise is 1963:07 to 2017:06 (648 observations). The specific model we consider takes the form

$$m_{t,t+1} = \exp(\lambda_0 - 1 + \boldsymbol{\lambda}' \mathbf{er}^{\text{aug}}_{t,t+1}),$$

where $\mathbf{er}^{\text{aug}}_{t,t+1} \equiv \mathbf{Q}_{t,t+1} - \mathbf{1}R_{t+1,f}$ is a $11 \times 1$ vector of excess returns. The excess return $\mathbf{er}^{\text{aug}}_{t,t+1}$ are based on $\mathbf{Q}_{t,t+1}$ from the baseline five-factor model of Fama and French (2015) augmented with two momentum portfolios (reflecting the gross returns of past losers and winners):

- gross return of the market;
- gross returns of the two extreme low, and high, size portfolios;
- gross returns of the two extreme low, and high, book-to-market portfolios;
- gross returns of the two extreme low, and high, operating profitability portfolios;
- gross returns of the two extreme low, and high, investment portfolios;
- gross returns of the two extreme low, and high, momentum portfolios.

The constants $(\lambda_0, \boldsymbol{\lambda})$ are a solution to the minimization problem $\inf_{(\lambda_0,\boldsymbol{\lambda})} -E[q_t]\lambda_0 + E[\exp(\lambda_0 - 1 + \boldsymbol{\lambda}'\mathbf{er}^{\text{aug}}_{t,t+1})]$. This form of SDF obtains by minimizing the objective $E[m\log(m)]$ subject to $E[m\,\mathbf{er}^{\text{aug}}] = \mathbf{0}$ and $E[m - E[q_t]] = 0$. The Lagrange multipliers on the equality constraints are $\boldsymbol{\lambda}$ and $\lambda_0$, respectively. Our procedure implies that

$$\lambda_0 = 1.01, \text{ and } \quad \boldsymbol{\lambda} = [53.8 \;\; 4.2 \;\; 0.1 \;\; -20.4 \;\; -15.0 \;\; -3.3 \;\; 0.3 \;\; -12.0 \;\; -2.9 \;\; 0.1 \;\; -5.9].$$

We present the monthly properties of $m$, and the model-based $L[m]$, and $L[m^2]$. The lower bounds on $\mathbb{LB}_m$ and $\mathbb{LB}_{m^2}$ are over the sample period 1963:07 to 2017:06. In each bootstrap draw, we randomly select, with replacement, raw asset returns $\mathbf{R}_{t,t+1}$ together with $\mathbf{Q}_{t,t+1}$ and recompute $(\lambda_0, \boldsymbol{\lambda})$. Then we generate the time-series of $m_{t,t+1}$, according to equation (25), and also compute the corresponding lower bounds according to equations (12) and (20). Here we perform 100,000 bootstrap trials. The $p$-values in bold indicate rejection of the null hypothesis.

| | $L[m]$ | $L[m^2]$ | Properties of the SDF (monthly) Mean | Std. | Skewness | Kurtosis | Min. | Max. |
|---|---|---|---|---|---|---|---|---|
| $m_{t,t+1} = \exp(\lambda_0 - 1 + \boldsymbol{\lambda}' \mathbf{er}^{\text{aug}}_{t,t+1})$ | 0.0569 | 0.2428 | 0.981 | 0.36 | 3.0 | 27.2 | 0.16 | 4.66 |

| Set of assets in $\mathbb{S}$ / $R_{t+1,f}$ plus market plus | 1963:07–2017:06, $\mathbf{R}_{t,t+1}$ is of dimension $26 \times 1$: $\mathbb{LB}_m$ | Bootstrap $p$-value | $\mathbb{LB}_{m^2}$ | Bootstrap $p$-value |
|---|---|---|---|---|
| 25 Size×B/M | 0.0243 | 0.978 | 0.2311 | 0.275 |
| 25 Size×Momentum | 0.0352 | 0.802 | 0.2795 | **0.064** |
| 25 Size×Accrual | 0.0238 | 0.946 | 0.2011 | 0.396 |
| 25 Size×Investment | 0.0279 | 0.927 | 0.2589 | 0.147 |
| 25 Size×Long-Term Reversal | 0.0176 | 0.970 | 0.0944 | 0.638 |
| 25 Size×Operating Profitability | 0.0402 | 0.975 | 0.3413 | 0.810 |
| 25 Size×Net Issues | 0.0191 | 0.704 | 0.1281 | **0.019** |
| 25 Size×Variance | 0.0332 | 0.795 | 0.3115 | **0.052** |
| 25 Size×Residual Variance | 0.0374 | 0.735 | 0.3399 | **0.040** |

Table 4
**Model that passes the bound on $L[m^2]$ also passes the bound on $L[m]$**

The lower bound on $L[m^2]$, denoted by $\mathbb{LB}_{m^2}$, is based on equation (12), while the lower bound on $L[m]$, denoted by $\mathbb{LB}_m$, is based on equation (20). The criterion is

$$\text{Model passes} = \begin{cases} \text{If } L[m] \geq \mathbb{LB}_m \\ \text{fails otherwise} \end{cases} \quad \text{and/or,} \;\; \text{Model passes} = \begin{cases} \text{If } L[m^2] \geq \mathbb{LB}_{m^2} \\ \text{fails otherwise.} \end{cases}$$

The model of the SDF that passes the bound $\mathbb{LB}_{m^2}$ is of the type (see Problem 1)

$$m^{\bullet}_{t,t+1} = \frac{\psi}{\left(\mathbf{a}'\mathbf{R}_{t,t+1}\right)^{\nu}} \qquad \text{(where } m^{\bullet} \text{ is consistent with } L[(m^{\bullet})^2] = \mathbb{LB}_{m^2}\text{).}$$

The parameters $\nu$ and $\psi$ solve equations (C13) and (C14), respectively. We use the following returns (a total of 12) to construct the SDF $m^{\bullet}_{t,t+1}$: (i) gross return of the risk-free bond, (ii) gross return of the market, (iii) gross returns of the two extreme low, and high, size portfolios, (iv) gross returns of the two extreme low, and high, book-to-market portfolios, (v) gross returns of the two extreme low, and high, operating profitability portfolios; (vi) gross returns of the two extreme low, and high, investment portfolios; (vii) gross returns of the two extreme low, and high, momentum portfolios. The sample period is 1963:07 to 2017:06. To construct $\mathbb{LB}_m$, we appeal to different sets of gross returns (a total of 12):

A. Set A: risk-free return, market, and 10 portfolios sorted by size;
B. Set B: risk-free return, market, and 10 portfolios sorted by B/M;
C. Set C: risk-free return, market, and 10 portfolios sorted by momentum;
D. Set D: risk-free return, market, and 10 portfolios sorted by accrual;
E. Set E: risk-free return, market, and 10 portfolios sorted by variance;
F. Set F: risk-free return, market, and 10 portfolios sorted by residual variance;
G. Set G: risk-free return, market, and 10 portfolios sorted by net issues;
H. Set H: risk-free return, market, and 10 portfolios sorted by long-term reversal;
I. Set I: risk-free return, market, and 10 portfolios sorted by investment;
J. Set J: risk-free return, market, and 10 portfolios sorted by operating profitability.

The reported bootstrap $p$-value is the proportion of bootstrap draws (out of 100,000) and tests the null hypothesis that $L[m^{\bullet}] \geq \mathbb{LB}_m$ (low $p$-values imply rejection, in bold).

| | | Properties of $m^{\bullet}$ | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| $\nu$ | $\psi$ | $L[m^{\bullet}]$ | $L[(m^{\bullet})^2]$ | Mean | Std. | Skewness | Kurtosis | Min. | Max. |
| 4.035 | 1.059 | 0.0433 | 0.1587 | 0.996 | 0.35 | 4.2 | 39.4 | 0.22 | 4.85 |

$p$-values for the null hypothesis $L[m^{\bullet}] \geq \mathbb{LB}_m$

| | A | B | C | D | E | F | G | H | I | J |
|---|---|---|---|---|---|---|---|---|---|---|
| $p$-value | 0.988 | 0.995 | 0.949 | 0.997 | 0.964 | 0.764 | 0.995 | 0.995 | 0.991 | 0.996 |
| $\mathbb{LB}_m$ | 0.0134 | 0.0124 | 0.0239 | 0.0129 | 0.0176 | 0.0308 | 0.0153 | 0.0105 | 0.0166 | 0.0097 |

Table 5

**Model that passes the bound on $L[m]$, but may or may not pass the bound on $L[m^2]$**

The lower bound on $L[m^2]$, denoted by $\mathbb{LB}_{m^2}$, is based on equation (12), while the lower bound on $L[m]$, denoted by $\mathbb{LB}_m$, is based on equation (20). The criterion is

$$\text{Model passes} = \begin{cases} \text{If } L[m] \geq \mathbb{LB}_m \\ \text{fails otherwise} \end{cases} \quad \text{and/or, Model passes} = \begin{cases} \text{If } L[m^2] \geq \mathbb{LB}_{m^2} \\ \text{fails otherwise.} \end{cases}$$

The model of the SDF that passes the bound $\mathbb{LB}_m$ is of the type (see Problem 2)

$$m^{\mathrm{G}}_{t,t+1} = \frac{\beta}{\left(\mathbf{a}'\mathbf{R}_{t,t+1}\right)^{\alpha}} \quad (\text{where } m^{\mathrm{G}} \text{ is consistent with } L[m^{\mathrm{G}}] = \mathbb{LB}_m).$$

The parameters $\alpha$ and $\beta$ solve $0 = E[q_t] - \exp((\alpha-1)E[\log(\mathbf{a}'\mathbf{R}_{t,t+1})])E[(\mathbf{a}'\mathbf{R}_{t,t+1})^{-\alpha}]$, and $0 = \beta - \exp((\alpha-1)E[\log(\mathbf{a}'\mathbf{R}_{t,t+1})])$. We use the following returns (a total of 12) to construct the SDF $m^{\mathrm{G}}_{t,t+1}$: (i) gross return of the risk-free bond, (ii) gross return of the market, (iii) gross returns of the two extreme low, and high, size portfolios, (iv) gross returns of the two extreme low, and high, book-to-market portfolios, (v) gross returns of the two extreme low, and high, operating profitability portfolios; (vi) gross returns of the two extreme low, and high, investment portfolios; (vii) gross returns of the two extreme low, and high, momentum portfolios. The sample period is 1963:07 to 2017:06. To construct $\mathbb{LB}_{m^2}$, we appeal to different sets of gross returns (a total of 12):

A. Set A: risk-free return, market, and 10 portfolios sorted by size;

B. Set B: risk-free return, market, and 10 portfolios sorted by B/M;

C. Set C: risk-free return, market, and 10 portfolios sorted by momentum;

D. Set D: risk-free return, market, and 10 portfolios sorted by accrual;

E. Set E: risk-free return, market, and 10 portfolios sorted by variance;

F. Set F: risk-free return, market, and 10 portfolios sorted by residual variance;

G. Set G: risk-free return, market, and 10 portfolios sorted by net issues;

H. Set H: risk-free return, market, and 10 portfolios sorted by long-term reversal;

I. Set I: risk-free return, market, and 10 portfolios sorted by investment;

J. Set J: risk-free return, market, and 10 portfolios sorted by operating profitability.

The reported bootstrap $p$-value is the proportion of bootstrap draws (out of 100,000) and tests the null hypothesis that $L[(m^{\mathrm{G}})^2] \geq \mathbb{LB}_{m^2}$ (low $p$-values imply rejection, in bold).

| | | Properties of $m^{\mathrm{G}}$ | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| $\alpha$ | $\beta$ | $L[m^{\mathrm{G}}]$ | $L[(m^{\mathrm{G}})^2]$ | Mean | Std. | Skewness | Kurtosis | Min. | Max. |
| 2.920 | 1.051 | 0.0220 | 0.0963 | 0.996 | 0.23 | 2.9 | 25.0 | 0.33 | 3.16 |

$p$-values for the null hypothesis $L[(m^{\mathrm{G}})^2] \geq \mathbb{LB}_{m^2}$

| | A | B | C | D | E | F | G | H | I | J |
|---|---|---|---|---|---|---|---|---|---|---|
| $p$-value | 0.510 | 0.604 | **0.066** | 0.425 | 0.245 | **0.034** | 0.317 | 0.703 | 0.276 | 0.689 |
| $\mathbb{LB}_{m^2}$ | 0.0767 | 0.0708 | 0.1452 | 0.0947 | 0.1094 | 0.181 | 0.1098 | 0.0639 | 0.111 | 0.0633 |

Table 6

**Robustness of conclusions in Tables 4 and 5**

Here we depart by increasing the number of assets used to construct $\mathbb{LB}_m$ ($\mathbb{LB}_{m^2}$) in Table 4 (Table 5). Specifically, we consider risk-free return, market, and 25×1 vector of returns. The model of the SDF that passes the bound $\mathbb{LB}_{m^2}$ is of the type (see Problem 1)

$$m^{\bullet}_{t,t+1} = \frac{\psi}{\left(\mathbf{a}'\mathbf{R}_{t,t+1}\right)^{\nu}} \qquad \text{(where } m^{\bullet} \text{ is consistent with } L[(m^{\bullet})^2] = \mathbb{LB}_{m^2}\text{)}.$$

The model of the SDF that passes the bound $\mathbb{LB}_m$ is of the type (see Problem 2)

$$m^{\mathrm{G}}_{t,t+1} = \frac{\beta}{\left(\mathbf{a}'\mathbf{R}_{t,t+1}\right)^{\alpha}} \qquad \text{(where } m^{\mathrm{G}} \text{ is consistent with } L[m^{\mathrm{G}}] = \mathbb{LB}_m\text{)}.$$

We use the following returns (a total of 12) to construct the SDF $m^{\bullet}_{t,t+1}$ and $m^{\mathrm{G}}_{t,t+1}$: (i) gross return of the risk-free bond, (ii) gross return of the market, (iii) gross returns of the two extreme low, and high, size portfolios, (iv) gross returns of the two extreme low, and high, book-to-market portfolios, (v) gross returns of the two extreme low, and high, operating profitability portfolios; (vi) gross returns of the two extreme low, and high, investment portfolios; (vii) gross returns of the two extreme low, and high, momentum portfolios. The sample period is 1963:07 to 2017:06. The reported bootstrap $p$-value is the proportion of bootstrap draws (out of 100,000) and tests the null hypothesis that $L[m^{\bullet}] \geq \mathbb{LB}_m$ or $L[(m^{\mathrm{G}})^2] \geq \mathbb{LB}_{m^2}$. Low $p$-values imply rejection and are shown in bold.

| | $\nu$ | $\psi$ | $L[m^{\bullet}]$ | $L[(m^{\bullet})^2]$ | Mean | Std. | Skewness | Kurtosis | Min. | Max. |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | Properties of $m^{\bullet}$ | | | | | | | |
| Setting of Table 4 | 4.035 | 1.059 | 0.0433 | 0.1587 | 0.9961 | 0.35 | 4.2 | 36.4 | 0.22 | 4.85 |
| | | | Properties of $m^{\mathrm{G}}$ | | | | | | | |
| | $\alpha$ | $\beta$ | $L[m^{\mathrm{G}}]$ | $L[(m^{\mathrm{G}})^2]$ | Mean | Std. | Skewness | Kurtosis | Min. | Max. |
| Setting of Table 5 | 2.920 | 1.051 | 0.022 | 0.0963 | 0.9961 | 0.23 | 2.9 | 22.0 | 0.33 | 3.16 |

| $R_{t+1,f}$ plus market plus | $\mathbb{LB}_m$ | Bootstrap $p$-value $L[m^{\bullet}] \geq \mathbb{LB}_m$ | $\mathbb{LB}_{m^2}$ | Bootstrap $p$-value $L[(m^{\mathrm{G}})^2] \geq \mathbb{LB}_{m^2}$ |
|---|---|---|---|---|
| 25 Size × B/M | 0.0243 | 0.977 | 0.2311 | **0.018** |
| 25 Size × Momentum | 0.0352 | 0.737 | 0.2795 | **0.006** |
| 25 Size × Accrual | 0.0238 | 0.932 | 0.2011 | **0.014** |
| 25 Size × Investment | 0.0279 | 0.917 | 0.2589 | **0.010** |
| 25 Size × Long-Term Reversal | 0.0203 | 0.963 | 0.1633 | **0.030** |
| 25 Size × Operating Profitability | 0.0191 | 0.971 | 0.1281 | **0.070** |
| 25 Size × Net Issues | 0.0402 | 0.617 | 0.3413 | **0.003** |
| 25 Size × Variance | 0.0332 | 0.727 | 0.3115 | **0.010** |
| 25 Size × Residual Variance | 0.0374 | 0.650 | 0.3399 | **0.006** |

# New Entropy Restrictions and the Quest for Better Specified Asset Pricing Models

**Internet Appendix: Not for Publication**

**Abstract**

Section A of the Internet Appendix studies the implications of enforcing the Hansen-Jagannathan lower bound, whereas Section B sheds light on the sharpness of the lower bounds on $L[m]$. Section C presents three asset pricing models (i) difference habit, (ii) recursive utility with stochastic variance, and (iii) recursive utility with constant jump intensity. Our empirical assessment shows that each model is rejected based on the lower bounds on $L[m]$ and $L[m^2]$.

# I. Internet Appendix

## *A. Implications of enforcing the Hansen-Jagannathan lower bound*

The problem in Hansen and Jagannathan (1991, page 235) is to find the SDF with minimum variance

$$\min_{m \in \mathbb{S}} \left\{ E\left[m^2\right] - \left(E\left[m\right]\right)^2 \right\}. \tag{IA-1}$$

Consider a portfolio $p$ depicted by the return

$$R^p_{t,t+1} = \mathbf{a}^{'}\mathbf{R}_{t,t+1} \qquad \text{with } \mathbf{a} = \frac{\mathbf{y}}{\mathbf{1}'\mathbf{y}} \qquad \text{and } \mathbf{y} = \Sigma^{-1}\left(\mathbf{1} - E\left[q_t\right]E\left[\mathbf{R}_{t,t+1}\right]\right). \tag{IA-2}$$

Analogous to how we construct $m^{\bullet}$ and $m^{G}$ in Problems 1 and 2, we use portfolio (IA-2) to construct $m^{\text{HJ}}_{t,t+1}$ consistent with the Hansen and Jagannathan (1991) lower bound. We conjecture and then verify the solution.

$$\begin{aligned} m^{\text{HJ}}_{t,t+1} &= \beta_0 + \beta_{\text{HJ}} R^p_{t,t+1}, & \text{(IA-3)} \\ &= \beta_0 + \beta_{\text{HJ}} (\mathbf{a}^{'}\mathbf{R}_{t,t+1}), & \text{(IA-4)} \end{aligned}$$

where $\beta_0$ and $\beta_{\text{HJ}}$ are constant parameters.

The first restriction is that the variance of $m^{\text{HJ}}_{t,t+1}$ equate to the Hansen and Jagannathan (1991, equation (12)) minimum variance, given by $\sigma_{\text{HJ}}^2 = \mathbf{y}^{'}\Sigma\mathbf{y}$,

$$\beta_{\text{HJ}}^2 \overbrace{\text{var}\left(\mathbf{a}^{'}\mathbf{R}_{t,t+1}\right)}^{\mathbf{a}^{'}\Sigma\mathbf{a}} = \overbrace{\sigma_{\text{HJ}}^2}^{\mathbf{y}^{'}\Sigma\mathbf{y}}, \quad \text{implying that} \tag{IA-5}$$

$$\beta_{\text{HJ}}^2 \mathbf{y}^{'}\Sigma\mathbf{y} = \sigma_{\text{HJ}}^2 (\mathbf{1}'\mathbf{y})^2. \tag{IA-6}$$

Therefore, we obtain

$$\beta_{\text{HJ}}^2 = \left(\mathbf{1}'\mathbf{y}\right)^2 \text{ and hence } \beta_{\text{HJ}} = \mathbf{1}'\mathbf{y}. \tag{IA-7}$$

Next, we enforce the restriction on the mean of the SDF:

$$\underbrace{\beta_0 + \left(\mathbf{1}'\mathbf{y}\right) E\left[\mathbf{a}^{'}\mathbf{R}_{t,t+1}\right]}_{E\left[m_{t,t+1}^{\text{HJ}}\right]} = E\left[q_t\right], \tag{IA-8}$$

which yields that

$$\beta_0 = E\left[q_t\right] - \left(\mathbf{1}'\mathbf{y}\right) E[\mathbf{a}^{'}\mathbf{R}_{t,t+1}]. \tag{IA-9}$$

The end result is the expression for the minimum variance SDF of the type

$$m_{t,t+1}^{\text{HJ}} = E\left[q_t\right] + \left(\mathbf{1}'\mathbf{y}\right)\left(\mathbf{a}^{'}\mathbf{R}_{t,t+1} - E[\mathbf{a}^{'}\mathbf{R}_{t,t+1}]\right), \tag{IA-10}$$

$$= E\left[q_t\right] + \mathbf{y}^{'}\left(\mathbf{R}_{t,t+1} - E[\mathbf{R}_{t,t+1}]\right). \tag{IA-11}$$

The restrictions (IA-5) and (IA-8) are used to construct $m^{\text{HJ}}$, and are in the flavor of how we used restrictions (C3) and (C10) to construct $m^{\bullet}$ (analogously, we use equations (F2) and (F4) to construct $m^{\text{G}}$).

By construction $m^{\text{HJ}}_{t,t+1}$ prices correctly the generic portfolio $\mathbf{a}^{'}\mathbf{R}_{t,t+1}$, as verified below:

$$\begin{aligned}
E[m^{\text{HJ}}_{t,t+1}(\mathbf{a}^{'}\mathbf{R}_{t,t+1})] &= \frac{\mathbf{1}}{(\mathbf{1}'\mathbf{y})}E[m^{\text{HJ}}_{t,t+1}(\mathbf{y}^{'}\mathbf{R}_{t,t+1})], \\
&= \frac{\mathbf{1}}{(\mathbf{1}'\mathbf{y})}\left(E\left[\left(E\left[q_t\right]+\left(\mathbf{y}^{'}\mathbf{R}_{t,t+1}-E\left[\mathbf{y}^{'}\mathbf{R}_{t,t+1}\right]\right)\right)\left(\mathbf{y}^{'}\mathbf{R}_{t,t+1}\right)\right]\right), \\
&= \frac{\mathbf{1}}{(\mathbf{1}'\mathbf{y})}\left\{E\left[q_t\right]\left(\mathbf{y}^{'}E\left[\mathbf{R}_{t,t+1}\right]\right)+\text{var}[\mathbf{y}^{'}\mathbf{R}_{t,t+1}]\right\}. \qquad \text{(IA-12)}
\end{aligned}$$

Next, we note that

$$\text{var}[\mathbf{y}^{'}\mathbf{R}_{t,t+1}] = \mathbf{y}^{'}\Sigma\mathbf{y} = \left(\mathbf{1}-E\left[q_t\right]E\left[\mathbf{R}_{t,t+1}\right]\right)^{'}\Sigma^{-1}\left(\mathbf{1}-E\left[q_t\right]E\left[\mathbf{R}_{t,t+1}\right]\right) \qquad \text{(IA-13)}$$

and further that

$$\begin{aligned}
\mathbf{y}^{'}\left(E\left[q_t\right]\right)E\left[\mathbf{R}_{t,t+1}\right] &= \left(\mathbf{1}-E\left[q_t\right]E\left[\mathbf{R}_{t,t+1}\right]\right)^{'}\Sigma^{-1}E\left[\mathbf{R}_{t,t+1}\right], \\
&= \left(\mathbf{1}-E\left[q_t\right]E\left[\mathbf{R}_{t,t+1}\right]\right)^{'}\Sigma^{-1}\left(E\left[q_t\right]E\left[\mathbf{R}_{t,t+1}\right]-\mathbf{1}+\mathbf{1}\right), \\
&= -\mathbf{y}^{'}\Sigma\mathbf{y}+\left(\mathbf{1}-E\left[q_t\right]E\left[\mathbf{R}_{t,t+1}\right]\right)^{'}\Sigma^{-1}\mathbf{1}, \\
&= -\mathbf{y}^{'}\Sigma\mathbf{y}+\mathbf{y}^{'}\mathbf{1}. \qquad \text{(IA-14)}
\end{aligned}$$

Equations (IA-13) and (IA-14) together imply that

$$E[m^{\text{HJ}}_{t,t+1}(\mathbf{a}^{'}\mathbf{R}_{t,t+1})] = \frac{\mathbf{1}}{(\mathbf{1}'\mathbf{y})}\left\{-\mathbf{y}^{'}\Sigma\mathbf{y}+\mathbf{y}^{'}\mathbf{1}+\mathbf{y}^{'}\Sigma\mathbf{y}\right\} = \frac{(\mathbf{1}^{'}\mathbf{y})}{(\mathbf{1}'\mathbf{y})} = 1.$$

While the minimum variance SDF prices correctly the return of the portfolio $\mathbf{a}^{'}\mathbf{R}_{t,t+1}$ , does $m^{\text{HJ}}_{t,t+1}$

price correctly the set of return $\mathbf{R}_{t,t+1}$? To price the set of return $\mathbf{R}_{t,t+1}$, we must have

$$E\left[m^{\text{HJ}}_{t,t+1}\mathbf{R}_{t,t+1}\right] = 1 \text{ or } E\left[q_t\right]E\left[\mathbf{R}_{t,t+1}\right] + \text{cov}\left(\left(\mathbf{1} - E\left[q_t\right]E\left[\mathbf{R}_{t,t+1}\right]\right)'\Sigma^{-1}\mathbf{R}_{t,t+1}, \mathbf{R}_{t,t+1}\right) = 1. \tag{IA-15}$$

Theoretically, this equality holds only when the dimension of $\mathbf{R}_{t,t+1}$ is one.

In summary, our theoretical results indicate a one-way implication: when the pricing restrictions in set (11) are used to construct the Hansen and Jagannathan bound, the variance of the minimum variance SDF is identical to the bound. The converse need not hold; that is, the obtained minimum variance SDF does not necessary price the $N+1$ assets employed to construct the minimum variance SDF. ■

## B. *Sharpness of our entropy bound on $L[m]$*

How sharp is our bound on $L[m]$ compared to the bound constructed from a generic portfolio return in Backus, Chernov, and Zin (2014, Column 2 of Table I).

Table Internet Appendix-I reports our lower bounds on $L[m]$ and the associated bootstrap *p*-values. We consider several $N$ (the dimensionality of $\mathbf{R}_{t,t+1}$) and draw two conclusions. First, our bounds on $L[m]$ are quantitatively sharper, implying greater hurdles on pricing models (e.g., compare bounds in Panel V versus those in Panels I through IV). Second, the bounds obtained with a portfolio are far less stringent than the corresponding bounds that rely on the SDFs correctly pricing each of the assets composing the portfolio. This can be seen by comparing the bound displayed in row (c) versus (i) and between row (d) versus (j). ■

## C. Example asset pricing models

Our goal is to learn about the properties of $m_{t,t+1}$, and their consistency with bound restrictions. Additionally, we compare $L[m^2]$ to $4L[m]$. We focus on three models:

(i) Difference habit,

(ii) Recursive utility with stochastic variance, and

(iii) Recursive utility with constant jump intensity.

Some of the model solutions require loglinearization, whose effects are explored and elaborated in the study of Pohl, Schmedders, and Wilms (2015).

### C.1. Difference habit model

The shocks in the difference habit model model are normally distributed, and the SDF is (Campbell and Cochrane (1999))

$$m_{t,t+1} \;=\; \beta\, g_{t+1}^{\rho-1} \left(\frac{s_{t+1}}{s_t}\right)^{\rho-1}, \tag{IA-16}$$

where $g_{t+1}$ is consumption growth, $\beta$ is the time discount parameter, and $1-\rho$ is the coefficient of relative risk aversion. Define $s_t \equiv 1-\exp(z_t)$ and $z_t \equiv \log(h_t) - \log(c_t)$, where $s_t$ is the surplus ratio corresponding to $z_t$, and the habit $h_{t+1}$ is known at $t$. The laws of motion for $h_t$ and $g_t$ are

$$\log(h_{t+1}) = \log(h) + \eta\,[B]\,\log(c_t) \quad \text{and} \quad \log(g_{t+1}) = \log(g) + \gamma[B]\,\upsilon^{\frac{1}{2}}\omega_{gt+1}, \tag{IA-17}$$

where $B$ is the lag operator, such that $B\{s_{t+1}\} = s_t$, with backshift operators $\gamma[B] = \sum_{j=0}^{\infty} \gamma_j B^j$ and $\eta\,[B] = \sum_{j=0}^{\infty} \eta_j B^j$. Moreover, $\upsilon$ denotes the constant variance of $\log(g_t)$, and $\omega_{gt+1}$ is i.i.d. standard normal variable.

Loglinear approximation of $\log(s_t)$, in conjunction with equation (IA-17), leads to the following dynamics:

$$\log(s_{t+1}) - \log(s_t) = \left(\frac{s-1}{s}\right)(\eta\,[B]\,B - 1)\log(g_{t+1}). \tag{IA-18}$$

Completing the model description, we define the state variable $x_t = (\gamma[B] - \gamma_0)\,\upsilon^{\frac{1}{2}}\,\omega_{gt+1}$, which governs the following dynamics of the log consumption growth:

$$x_t = \gamma_1\,\upsilon^{\frac{1}{2}}\,\omega_{gt} + \varphi_g\,x_{t-1} \qquad \text{with} \qquad \varphi_g = \frac{\gamma_2}{\gamma_1}. \tag{IA-19}$$

Models that accommodate habit have shown promise in matching salient attributes of the asset market data, including the equity premium, procyclicality of stock prices, counter-cyclicality of stock volatility, and return predictability at long horizons (e.g., see, among others, Bekaert and Engstrom (2017), Chapman (1998), Chan and Kogan (2002), and Santos and Veronesi (2010)).

*C.2. Recursive utility models*

The recursive utility models are adopted from Backus, Chernov, and Zin (2014):

$$U_t = \left[(1-\beta)\,c_t^{\rho} \;+\; \beta\,(\mu_t\,[U_{t+1}])^{\rho}\right]^{\frac{1}{\rho}}, \tag{IA-20}$$

with certainty equivalent function $\mu_t\,[U_{t+1}] = \left(E_t\left[U_{t+1}^{\alpha}\right]\right)^{\frac{1}{\alpha}}$. Moreover, $\beta$ is the time preference parameter, $\frac{1}{1-\rho}$ is the intertemporal elasticity of substitution, and $1-\alpha$ is the coefficient of relative risk aversion.

The shocks $\omega_{gt}$, $z_{gt}$, and $\omega_{ht}$ are standard normal random variables, independent of each other and across time. Additionally, the jump component $z_{gt}$ is a Poisson mixture of normals: conditional on the number of jumps $j$, $z_{gt}$ is normal, with mean $j\theta$ and variance $j\delta^2$. The probability of $j \geq 0$ jumps at date $t$ is $e^{h_{t-1}} h_{t-1}^{j}/j!$, and the jump intensity, $h_{t-1}$, is the mean of $j$.

With backshift operators characterized by $\nu[B] = \sum_{j=0}^{\infty} \nu_j B^j$ and $\psi[B] = \sum_{j=0}^{\infty} \psi_j B^j$, the state-variables in this model obey the following dynamics:

$$\log(g_t) = \log(g) + \gamma[B]\,\upsilon_{t-1}^{1/2}\omega_{gt} + \psi[B]\,z_{gt} - \psi[1]\,h\theta, \qquad h_t = h + \eta[B]\,\omega_{ht}, \tag{IA-21}$$

$$\upsilon_t = \upsilon + \nu[B]\,\omega_{\upsilon t}, \quad z_{gt}|j \sim \mathcal{N}\left(j\theta, j\delta^2\right), \quad P[j] = \exp(-h_{t-1})\,\frac{(h_{t-1})^j}{j!}. \tag{IA-22}$$

*A. Recursive utility model with stochastic variance.* Set $h = 0$, $\eta[B] = 0$, $\psi[B] = 0$ in equations (IA-21) and (IA-22). For tractability, we consider the evolution of the transformed variable:

$$x_t = \varphi_g x_{t-1} + \gamma_1 \upsilon_{t-1}^{1/2}\omega_{gt}. \tag{IA-23}$$

*B. Recursive utility model with constant jump intensity*: In equations (IA-21) and (IA-22), set $\nu[B] = 0$.

Models that incorporate recursive preferences in conjunction with stochastic variance or jumps in the consumption growth dynamics have proved successful in explaining asset pricing quantities. We refer the reader to, among others, Epstein and Zin (1991), Bansal and Yaron (2004), Campbell and Vuolteenaho (2004), Hansen, Heaton, and Li (2008), Wachter (2013), and Zhou and Zhu

(2009).

*C.3. Empirical evidence and connection to our findings*

How do the models under consideration fare when viewed from the perspective of data-based lower bounds on the entropy of $m$, entropy of $m^2$, and the volatility of $m$?

Our implementation of the models with difference habit (hereby DH), recursive utility with stochastic variance (hereby RU-SV), and recursive utility with constant jump intensity (hereby RU-CJI) follows the calibration procedure in Backus, Chernov, and Zin (2014, respectively, Model (4) in Table 2, Model (1) in Table 3, and Model (4) in Table 4). The corresponding model parameterizations are displayed in our Table Internet Appendix-III, which indicates that each model reasonably calibrates to consumption growth data.

Aided by our analytical representations, we generate the paths for $m_{t,t+1}$. The paths are based on the model parameters in Table Internet Appendix-III and shocks driving the fundamentals (e.g., $\omega_{\upsilon t}$ and $\omega_{gt}$ for the RU-SV). Then we obtain the sample averages of the series $\{m^2_{t,t+1}, m_{t,t+1} : t = 1, \ldots T\}$, and accordingly compute the entropies $L[m^2_{t,t+1}]$, $L[m_{t,t+1}]$, and the volatilities of $m_{t,t+1}$.

Next, we draw 50,000 paths for the shocks driving a model and, hence, obtain 50,000 paths for $m_{t,t+1}$. Panels A, B, and C of Table Internet Appendix-II report the entropies and volatilities across the models, obtained by averaging the entropies over the replications. The $p$-values, shown in square brackets, represent the proportion of replications for which the model-based entropy and volatility measures exceeds the corresponding lower bound obtained from the returns data in 50,000 replications of a simulation over 966 months.

How successful are the three models in generating $L[m]$ that is consistent with the data? Panel

A of Table Internet Appendix-II reveals an $L[m]$ of 0.0196, 0.0217, and 0.0190, respectively, for the DH, RU-SV, and RU-CJI models. Based on our data-based performance measure, computed based on SET B, all the models are rejected (as seen by the bootstrap $p$-values).

Such an implication from our bound, calculated using the return properties of the risk-free bond, the equity market, and the 25 portfolios sorted by size and momentum, differ from a finding in Backus, Chernov, and Zin (2014). Specifically, the data-based lower bound in Backus, Chernov, and Zin (2014, Table 1) are generally of an order lower than the average conditional entropy $E[L_t[m]]$ obtained from asset pricing models. In particular, all of the 11 $E[L_t[m]]$ in Backus, Chernov, and Zin (2014, Tables II through IV) exceed the lower bound inferred from the returns on a generic portfolio taken to be the S&P 500 index.

How does one explain this discrepancy? We note that the magnitude of the lower bound on $L[m]$ in the calculations of Backus, Chernov, and Zin (2014, Table 1, row S&P 500) is 0.0040, whereas it is 0.0367, based on our lower bound and SET B. It bears emphasizing that the lower bound on $L[m]$ constructed from the returns of a (single) generic portfolio may provide an insufficient hurdle in evaluating the merits of an asset pricing model. The bounds on $L[m]$ agree in suggesting that the models are misspecified.

Elaborating further, we now argue that considering the entropy $L[m^2]$ in the model assessment can provide an important contrast to our findings based on the entropy $L[m]$. One noteworthy result is that the entropy $L[m^2]$ of the RU-CJI model is about 15-fold higher than the other two models that do not incorporate the random jump feature in the dynamics of the consumption growth. For example, the DH, RU-SV, and RU-CJI models generate $L[m^2]$ of 0.0785, 0.0869, and 1.4331, respectively (see the entries in Panel B of Table Internet Appendix-II). We further note that since the lower bound restriction implied from asset prices is 0.1956, the DH and RU-SV models are rejected at the 5% level. However, the RU-CJI model with constant jump intensity cannot be rejected at the 5% level, which is a point of departure based on the entropy $L[m]$.

Accordingly, one question emerges: Why does the RU-CJI fail to explain features of $m$, as reflected in asset prices when $L[m]$-based performance measure is used, while the model is successful in explaining features of $m$, as reflected in asset prices when the $L[m^2]$-based performance measure is used? To investigate a source of model performance, we note that the entropy measure $L[m^2]$ is substantially more sensitive to tail asymmetries and tail size of the $m$ distribution as opposed to the entropy measure $L[m]$.

Taking such a trait of entropies into consideration, we report the moments of $m_{t,t+1}$ for each of the models in Panel D of Table Internet Appendix-II. The unexpected finding is that the RU-CJI model embeds excessive levels of skewness and kurtosis of $m_{t,t+1}$, while generating variance that is almost 90 times its DH and RU-SV model counterparts. Our contention is that the inordinate levels of the higher-order moments of $m_{t,t+1}$ give rise to the reported $L[m^2]$ of 1.4331 for the RU-CJI model.

How should one interpret a model, such as the RU-CJI, that calibrates well to the first moment, the second moment, and the autocorrelation of consumption growth but does not produce finite central moments for the distribution of $m_{t,t+1}$? This result arises because a convex transform of a random variable, which is here Poisson-distributed, increases the skewness to the right (see van Zwet (1966, page 10, Theorem 2.2.1)).

To see this analytically, we can use the density of the Poisson random variable to show that $E_t[(m_{t,t+1})^k] = E_t\left[e^{k\log(m_{t,t+1})}\right] = E_t\left[E_t\left[e^{k\log(m_{t,t+1})}|j\right]\right] = e^{G[k]}E_t\left[e^{H[k]j}\right]$, for constants $G[k]$

and $H[k]$. Note that $e^{H[k]j}$ is a convex transformation of the Poisson variable $J$, and, for certain parameterizations, does not admit finite higher-moments of $m_{t,t+1}$. The inordinate amounts of skewness and kurtosis do not appear to be a reasonable depiction of valuation operators, which are likely to be characterized by exponential, rather than power, tails.

Finally, consider the volatility bound on $m$ using the Hansen and Jagannathan (1991, equation (12)). As seen from Panel C of Table Internet Appendix-II, the DH and RU-SV models are rejected, but the RU-CJI model is not rejected for reasons discussed, namely, that the RU-CJI model embeds an unreasonable volatility, skewness, and kurtosis of $m$.

*C.4. Details: SDF of the difference habit model*

Using a loglinear approximation of $\log(s_t)$,

$$\log(m_{t,t+1}) = D_0 + (\rho - 1)\,\frac{1}{s}\,(1 - (1-s)\,\eta\,[B]\,B)\,\gamma\,[B]\,\upsilon^{\frac{1}{2}}\omega_{gt+1}, \qquad \text{(IA-24)}$$

$$\text{where } D_0 = \log(\beta) + (\rho - 1)\log(g) + (\rho - 1)\,\frac{(s-1)}{s}\left(\frac{\eta_0}{1-\varphi_h} - 1\right)\log(g).$$

Given the approximation $\log(s_t) \approx 1 + \frac{(s-1)}{s} z_t$, the dynamics of the surplus consumption ratio are

$$\log(s_{t+1}) - \log(s_t) = \frac{(s-1)}{s}\,(\eta\,[B]\,B - 1)\log(g_{t+1}). \qquad \text{(IA-25)}$$

Therefore, we may write the log SDF as

$$
\begin{aligned}
\log(m_{t,t+1}) &= \log(\beta) + (\rho-1)\log(g) + (\rho-1)\frac{(s-1)}{s}\left(\eta\left[B\right]B-1\right)\log(g) \\
&\quad + (\rho-1)\frac{1}{s}\left(1-(1-s)\,\eta\left[B\right]B\right)\gamma[B]\,\upsilon^{\frac{1}{2}}\,\omega_{gt+1}. \qquad \text{(IA-26)}
\end{aligned}
$$

We have the expression. ■

*C.5. Details: SDF of the recursive utility models*

Based on equations (IA-20) and (IA-22), we note that $\omega_{gt}$, $z_{gt}$, and $\omega_{ht}$ are standard normal random variables, independent of each other and across time. The jump component $z_{gt}$ is a Poisson mixture of normals: conditional on the number of jumps $j$, $z_{gt}$ is normal with mean $j\theta$ and variance $j\delta^2$. The probability of $j \geq 0$ jumps at date $t+1$ is $e^{h_t} h_t^j / j!$ expands to

$$
\begin{aligned}
m_{t,t+1} &= \exp\left(\chi_0 + a_g\left[B\right]\upsilon_t^{\frac{1}{2}}\omega_{gt+1} + a_z\left[B\right]z_{gt+1} + a_\upsilon\left[B\right]\omega_{\upsilon t+1} + a_h\left[B\right]\omega_{ht+1}\right), \qquad \text{(IA-27)} \\
\chi_0 &= \log(\beta) + (\rho-1)\log(g) \\
&\quad - (\alpha-\rho)\left(D\upsilon - Jh\right) - (\alpha-\rho)\left(\alpha/2\right)\left(\left(Db_1\nu\left[b_1\right]\right)^2 + \left(Jb_1\eta\left[b_1\right]\right)^2\right), \qquad \text{(IA-28)}
\end{aligned}
$$

where $a_g\left[B\right]$, $a_z\left[B\right]$, $a_\upsilon\left[B\right]$, and $a_h\left[B\right]$ are backshift operators defined as follows:

$$
a_g\left[B\right] = (\rho-1)\,\gamma[B] + (\alpha-\rho)\,\gamma[b_1], \qquad a_z\left[B\right] = (\rho-1)\,\psi\left[B\right] + (\alpha-\rho)\,\psi\left[b_1\right], \qquad \text{(IA-29)}
$$

$$
a_\upsilon\left[B\right] = (\alpha-\rho)\,D\left(b_1\nu\left[b_1\right] - \nu\left[B\right]B\right), \qquad a_h\left[B\right] = (\alpha-\rho)\,J\left(b_1\eta\left[b_1\right] - \eta\left[B\right]B\right), \qquad \text{(IA-30)}
$$

$$
D = (\alpha/2)\left(\gamma[b_1]\right)^2, \quad \text{and} \qquad J = \left(\frac{e^{\alpha\psi[b_1]\theta + (\alpha\psi[b_1]\delta)^2} - 1}{\alpha}\right). \qquad \text{(IA-31)}
$$

The functions $\eta[b_1]$, $\nu[b_1]$, and $\gamma[b_1]$ are polynomial functions of $b_1$:

$$\eta[b_1] = \sum_{j=0}^{\infty} b_1^j \eta_j, \quad \gamma[b_1] = \sum_{j=0}^{\infty} b_1^j \gamma_j, \quad \nu[b_1] = \sum_{j=0}^{\infty} b_1^j \nu_j, \quad \psi[b_1] = \sum_{j=0}^{\infty} b_1^j \psi_j, \tag{IA-32}$$

with $\gamma_0 = 1$, where

$$\sum_{j=1}^{\infty} \gamma_j < \infty, \quad \sum_{j=1}^{\infty} \eta_j < \infty, \quad \sum_{j=1}^{\infty} \nu_j < \infty, \quad \sum_{j=1}^{\infty} \psi_j < \infty, \tag{IA-33}$$

and

$$\nu[B] = \sum_{j=0}^{\infty} \nu_j B^j \qquad \text{and} \qquad \psi[B] = \sum_{j=0}^{\infty} \psi_j B^j. \tag{IA-34}$$

**A. Recursive utility with stochastic variance:** The SDF is a special case of (IA-27) with $h = 0$, $\eta[B] = 0$, $J = 0$. The SDF takes the form

$$m_{t,t+1} = \exp\begin{pmatrix} H_0 + (\rho-1)\gamma[B]\upsilon_t^{\frac{1}{2}}\omega_{gt+1} + (\alpha-\rho)\gamma[b_1]\upsilon_t^{\frac{1}{2}}\omega_{gt+1} \\ +(\alpha-\rho)Db_1\nu[b_1]\omega_{\upsilon t+1} - (\alpha-\rho)D\nu[B]B\omega_{\upsilon t+1} \end{pmatrix},$$

with

$$H_0 = \log(\beta) + (\rho-1)\log g - (\alpha-\rho)(D\upsilon) - (\alpha-\rho)(\alpha/2)\left((Db_1\nu[b_1])^2\right). \tag{IA-35}$$

Now we define

$$x_t = (\gamma[B]-\gamma_0)\upsilon_t^{\frac{1}{2}}\omega_{gt+1}. \tag{IA-36}$$

The state variable $x_t$ dynamics is

$$x_t = \varphi_g x_{t-1} + \gamma_1 \upsilon_{t-1}^{\frac{1}{2}} \omega_{gt}, \quad \text{with} \quad \gamma_j = \varphi_g \gamma_{j-1} \text{ for } j \geq 2 \quad \text{and} \quad \varphi_g = \frac{\gamma_2}{\gamma_1}. \tag{IA-37}$$

It can be shown that the dynamics of the state variable $\upsilon_t$ is

$$\upsilon_t - \upsilon = \varphi_\upsilon (\upsilon_{t-1} - \upsilon) + \nu_0 \omega_{\upsilon t}, \quad \text{for } j \geq 2 \quad \text{and} \quad \varphi_\upsilon = \frac{\nu_1}{\nu_0}. \tag{IA-38}$$

The SDF can be expressed as

$$m_{t,t+1} = \exp(H_1 + H_2 x_t + H_3 x_{t+1} + H_4 \upsilon_t + H_5 \upsilon_{t+1}), \tag{IA-39}$$

where

$$H_1 = H_0 + (\alpha - \rho) D\upsilon + (\alpha - \rho) D b_1 \nu[b_1] \frac{(\varphi_\upsilon - 1)}{\nu_0} \upsilon, \tag{IA-40}$$

$$H_2 = (\rho - 1) - ((\alpha - \rho)\gamma[b_1] + (\rho - 1)) \frac{\varphi_g}{\gamma_1}, \tag{IA-41}$$

$$H_3 = \frac{(\rho - 1)}{\gamma_1} + \frac{(\alpha - \rho)\gamma[b_1]}{\gamma_1}, \tag{IA-42}$$

$$H_4 = (\alpha - \rho) D \left( -b_1 \nu[b_1] \frac{\varphi_\upsilon}{\nu_0} - 1 \right), \text{ and} \tag{IA-43}$$

$$H_5 = (\alpha - \rho) D b_1 \frac{\nu[b_1]}{\nu_0}. \tag{IA-44}$$

■

**B. Recursive utility model with constant jump intensity:** Consider the consumption growth dynamics with $\nu[B]=0$ (in this case $\upsilon_t=\upsilon$). It can be shown that the SDF reduces to

$$m_{t,t+1}=\exp\left(\begin{array}{c}\chi_0\\ +(\rho-1)x_t+((\rho-1)\gamma_0+(\alpha-\rho)\gamma[b_1])\upsilon^{\frac{1}{2}}\omega_{gt+1}\\ +(\rho-1)(\psi[B]-\psi_0)z_{gt+1}+((\rho-1)\psi_0+(\alpha-\rho)\psi[b_1])z_{gt+1}\\ +(\alpha-\rho)Jb_1\eta[b_1]\omega_{ht+1}-(\alpha-\rho)(h_t-h)J\end{array}\right). \quad \text{(IA-45)}$$

Now denote

$$\widetilde{x}_t=(\psi[B]-\psi_0)z_{gt+1}. \quad \text{(IA-46)}$$

The law of motion of $\widetilde{x}_t$ becomes

$$\widetilde{x}_t=\varphi_z\widetilde{x}_{t-1}+\psi_1 z_{gt}, \quad \text{with} \quad \varphi_z=\frac{\psi_2}{\psi_1} \quad \text{and} \quad \psi_{j+2}=\varphi_z\psi_{j+1} \quad \text{for } j\geq 1. \quad \text{(IA-47)}$$

The SDF in equation (IA-45) reduces to

$$m_{t,t+1}=\exp\left(G_0+G_1x_t+G_2\widetilde{x}_{t-1}+G_3z_{gt}+G_4h_t+G_5z_{gt+1}+G_6\upsilon^{\frac{1}{2}}\omega_{gt+1}+G_7\omega_{ht+1}\right), \quad \text{(IA-48)}$$

with

$$\begin{array}{rclrcl} G_0 &=& \chi_0+(\alpha-\rho)hJ, & G_1 &=& (\rho-1),\\ G_2 &=& (\rho-1)\varphi_z, & G_3 &=& (\rho-1)\psi_1,\\ G_4 &=& -(\alpha-\rho)J, & G_5 &=& (\rho-1)\psi_0+(\alpha-\rho)\psi[b_1],\\ G_6 &=& (\rho-1)\gamma_0+(\alpha-\rho)\gamma[b_1], & G_7 &=& (\alpha-\rho)Jb_1\eta[b_1]. \end{array}$$

■

Table Internet Appendix-I

**Sharpness of our entropy bounds on $m_{t,t+1}$, when SDFs correctly price each of the $N+1$ assets**

Reported are the lower entropy bounds with the one-sided $p$-values in $\langle . \rangle$. Our lower entropy bound on $m_{t,t+1}$ is based on equation (20) and relies on the ability of the SDF to correctly price *each of the* $N+1$ assets (the risk-free bond and $N$ risky assets). The Backus, Chernov, and Zin (2014, equation (5)) lower bound on the entropy of $m_{t,t+1}$ (denoted by BCZ) is based on the expression $E[\log(R^m_{t,t+1})] - \log(R_{t+1,f})$, where $R^m_{t,t+1}$ is the return on a single risky asset or a benchmark portfolio (i.e., which we proxy by the value-weighted equity market return or equally weighted portfolio of 25 Fama-French size and book-to-market portfolios). $R_{t+1,f}$ is the gross return of the three-month Treasury bond. We employ different assets and $N$ in the construction of the bounds. For example, in Panel I, the $N$ risky assets are based on two data sets: SET A contains the value-weighted market returns, together with the 25 Fama-French size and book-to-market portfolios, while SET B contains the value-weighted market returns together with the 25 Fama-French size and momentum portfolios. The sample period is from July 1931 to December 2011 (966 observations). To compute these $p$-values, we first use the block bootstrap with a block size of 20 to generate 50,000 samples from the original data. Then we compute the lower bounds in each sample and tabulate the proportion of bootstrap samples for which the lower bound is less than zero.

| | Lower bound on $m_{t,t+1}$ | |
|---|---|---|
| | Bound | $p$-value |
| *Panel I. SDF correctly prices each of the $N+1$ assets, and we set $N=26$* | | |
| (a) Set A: Market, 25 size & B/M | **0.023** | $\langle 0.000 \rangle$ |
| (b) Set B: Market, 25 size & momentum | **0.037** | $\langle 0.003 \rangle$ |
| *Panel II. SDF correctly prices each of the $N+1$ assets, and we set $N=25$* | | |
| (c) Set C: 25 size & B/M | **0.022** | $\langle 0.000 \rangle$ |
| (d) Set D: 25 size & momentum | **0.029** | $\langle 0.000 \rangle$ |
| *Panel III. SDF correctly prices each of the $N+1$ assets, and we set $N=11$* | | |
| (e) Set E: Market, 10 momentum | **0.020** | $\langle 0.000 \rangle$ |
| *Panel IV. SDF correctly prices each of the $N+1$ assets, and we set $N=2$* | | |
| (f) Set F: Market, Low Momentum | **0.010** | $\langle 0.000 \rangle$ |
| (g) Set G: Market, high Momentum | **0.014** | $\langle 0.010 \rangle$ |
| *Panel V. SDF correctly prices each of the $N+1$ assets, and we set $N=1$* | | |
| | (BCZ, Eq. 5) | |
| (h) Set H: Market portfolio only | **0.005** | $\langle 0.005 \rangle$ |
| (i) Set I: EWI portfolio of 25 size & B/M | **0.007** | $\langle 0.001 \rangle$ |
| (j) Set J: EWI portfolio of 25 size & momentum | **0.007** | $\langle 0.001 \rangle$ |

Table Internet Appendix-II

**Model comparisons using bounds**

Reported are the results for bounds on the entropy of $m$, the entropy of $m^2$, and the volatility of $m$, for three models:

- the difference habit (denoted by DH),
- the recursive utility with stochastic variance (denoted by RU-SV),
- and the recursive utility with constant jump intensity (denoted by RU-CJI).

The one-sided $p$-values shown in square brackets represent the proportion of replications for which the model-based quantity (entropy or volatility) exceeds, in 50,000 replications, the lower bound computed from observed asset prices. Our lower bound on the entropy of $m^2$ is based on equation (12) and relies on the ability of the SDF to correctly price $N+1$ assets (the risk-free bond and $N$ risky assets). The $N$ risky assets are based on SET B, which contains the value-weighted market returns, together with the 25 Fama-French size and momentum portfolios. The sample period is from July 1931 to December 2011. The lower bound on the entropy of $m$ is based on equation (20) and also relies on the ability of the SDF to correctly price $N+1$ assets. The lower bound on the volatility of $m$ is based on Hansen and Jagannathan (1991, equation (12)). We focus on SET B, as it corresponds to the maximum lower bound on entropy measures (as in our Table Internet Appendix-I). Panel D presents the variance, skewness, and kurtosis of $m$, which are consistent with model parameterizations in Table Internet Appendix-III. The one-sided $p$-values $\langle . \rangle$, reported below the lower bounds, represent the proportion of bootstrap samples for which the lower bound is less than zero.

| | Habit model | Recursive utility models | | |
|---|---|---|---|---|
| | DH | RU-SV | RU-CJI | Lower bounds (Set B) |
| *Panel A: Entropy of m* | | | | |
| $L[m]$ | 0.0196 | 0.0217 | 0.0190 | 0.0367 |
| | [0.000] | [0.000] | [0.000] | $\langle 0.003 \rangle$ |
| *Panel B: Entropy of $m^2$* | | | | |
| $L[m^2]$ | 0.0785 | 0.0869 | 1.4331 | 0.1956 |
| | [0.000] | [0.000] | [1.000] | $\langle 0.003 \rangle$ |
| *Panel C: Volatility bound* | | | | |
| Hansen and Jagannathan (1991) | 0.0415 | 0.0444 | 3.344 | 0.1292 |
| | [0.000] | [0.000] | [1.000] | $\langle 0.000 \rangle$ |
| *Panel D: Moments of the $m_{t,t+1}$ distribution* | | | | |
| Variance | 0.0403 | 0.0444 | 3.3438 | |
| Skewness | 0.6041 | 0.6476 | $+\infty$ | |
| Kurtosis | 3.6447 | 3.8061 | $+\infty$ | |

Table Internet Appendix-III

**Parameters employed in model implementation**

Displayed in this table are the parameters that govern preferences and the dynamics of consumption growth. These parameters are adopted from Tables 2, 3, and 4 of Backus, Chernov, and Zin (2014), and likewise $\log(g)$ and $\eta_0$ are taken from their page 16. Our implementation of the models with difference habit (hereby DH), recursive utility with stochastic variance (hereby RU-SV), and recursive utility with constant jump intensity (hereby RU-CJI) follows Backus, Chernov, and Zin (2014, respectively, Model (4) in Table 2, Model (1) in Table 3, and Model (4) in Table 4). We use US annual real personal consumption expenditures as a proxy for aggregate consumption over the sample period of 1931:07 to 2011:12 (966 observations). To compare model implications with the data, we simulate a finite sample of consumption growth, $c_{t+1}/c_t$, over 966 months. Following convention, we then compute the annualized consumption growth as $\exp(\sum_{j=1}^{12}\log(c_{t+j}/c_{t+j-1}))$. The reported model mean, standard deviation, and auto-correlation are based on the annualized consumption growth.

| Parameter | DH | RU-SV | RU-CJI | Data implied |
|---|---|---|---|---|
| *Panel A: Preferences* | | | | |
| $\rho$ | -9.0000 | 0.3333 | 0.3333 | |
| $\alpha$ | | -9.0000 | -9.0000 | |
| $\beta$ | 0.9980 | 0.9980 | 0.9980 | |
| $\varphi_h$ | 0.9000 | | | |
| $s$ | 0.5000 | | | |
| *Panel B: Consumption growth dynamics* | | | | |
| $\gamma_0$ | 1.0000 | 1.0000 | 1.0000 | |
| $\log(g)$ | 0.0015 | 0.0015 | 0.0015 | |
| $\eta_0$ | 0.1000 | | | |
| $\gamma_1$ | 0.0271 | 0.0271 | 0.0281 | |
| $\varphi_g$ | 0.9790 | 0.9790 | 0.9690 | |
| $\upsilon^{1/2}$ | 0.0099 | 0.0099 | 0.0079 | |
| $\nu_0$ | | $0.23\times10^{-5}$ | | |
| $\varphi_\upsilon$ | | 0.9870 | | |
| $h$ | | | 0.0008 | |
| $\theta$ | | | -0.1500 | |
| $\delta$ | | | 0.1500 | |
| $\psi_0$ | | | 1.0000 | |
| $b_1$ | | 0.9977 | 0.9979 | |
| *Panel C: Consumption growth* | | | | |
| Mean (annualized) | 1.0192 | 1.0190 | 1.0189 | 1.0339 |
| Std. Dev. (annualized) | 0.0416 | 0.0415 | 0.0369 | 0.0287 |
| Autocorrelation | 0.2424 | 0.2433 | 0.1771 | 0.2386 |