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

Gurdip Bakshi\textsuperscript{a}† Fousseni Chabi-Yo\textsuperscript{b}‡

\textsuperscript{a}Smith School of Business, University of Maryland, College Park, MD 20742, USA
\textsuperscript{b}Fisher College of Business, Ohio State University, Columbus, OH 43210, USA

First draft, October 2012
This draft, December 14, 2015

Abstract

This paper proposes the entropy of \( m^2 \) (\( m \) is the stochastic discount factor) as a metric to evaluate asset pricing models. One key result is that the entropy of \( m^2 \) represents the maximum expected excess (log) return of the security with the payoff of \( m \). We formalize the sense in which the entropy of \( m^2 \) is distinct from the volatility of \( m \) and the entropy of \( m \). Furthermore, we develop a lower bound on the entropy of \( m^2 \) under the setting that \( m \) correctly prices a finite number of asset returns.

KEY WORDS: Stochastic discount factors, lower bounds, eigenfunction problem.

JEL CLASSIFICATION CODES: C51, C52, G12.

\textsuperscript{*}The authors acknowledge helpful discussions with Ravi Bansal, Mike Chernov, Riccardo Colacito, Steve Heston, Mark Loewenstein, Dilip Madan, George Panayotov, Alberto Rossi, Rene Stulz, Michael Stutzer, Adrien Verdelhan (AFA 2015 discussant), Jessica Wachter, 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. We thank David Backus, Mike Chernov, and Stan Zin for allowing us to cross-check our computer codes.

\textsuperscript{†}Tel.: +1-301-405-2261. \textit{E-mail address:} gbakshi@rhsmith.umd.edu

\textsuperscript{‡}Tel.: +1-614-292-8477. \textit{E-mail address:} chabi-yo_1@fisher.osu.edu
1. Introduction

Consider a security with payoff \( m_{t,t+1} \) and, hence, a time-\( t \) price of \( E_t[m^2_{t,t+1}] \) (e.g., Cochrane (2005, page 18)), where \( m_{t,t+1} \) is the stochastic discount factor (SDF) between time \( t \) and \( t+1 \), and \( E_t[\cdot] \) represents conditional expectation. The security with SDF payoff has return \( r^{\text{SDF}}_{t,t+1} = \frac{m_{t,t+1}}{E_t[m^2_{t,t+1}]} - 1 \), and is a hedging asset with gross return lower than the gross return of the risk-free bond, \( R^f_t \). Thus, we define

\[
er_{t,t+1} \equiv \log(R^f_t) - E_t[\log(1 + r^{\text{SDF}}_{t,t+1})]
\]

as the expected excess (log) return of the SDF security. (1)

Our core ideas are twofold. First, we establish that the entropy of \( m^2_{t,t+1} \) represents the maximum expected excess (log) return of the security with SDF payoff. Second, we develop a theoretical lower bound on the entropy of \( m^2_{t,t+1} \) to be used in assessing asset pricing models. We further show that the theoretical lower bound on the entropy of \( m^2_{t+1} \) can be extracted from a vector of traded asset returns.

Our investigation poses two central questions. First, what is the information content of the entropy of \( m^2_{t,t+1} \) relative to the volatility of \( m_{t,t+1} \) in Hansen and Jagannathan (1991)? Second, is there room for our entropy measure, given that the entropy of \( m_{t,t+1} \) has been analyzed by Stutzer (1995), Bansal and Lehmann (1997), Alvarez and Jermann (2005), and Backus, Chernov, and Zin (2014)?

Whereas the volatility of \( m_{t,t+1} \), the entropy of \( m_{t,t+1} \), and the entropy of \( m^2_{t,t+1} \) each aims to characterize the expected excess return of some reference asset, our focus on the asset with SDF payoff provides a new way to differentiate among pricing models. We develop arguments to show that our rationale for considering the entropy of \( m^2_{t,t+1} \) rests on strong economic foundations. Additionally, we are able to extend our methods to analyzing the entropy of the square of the permanent (martingale) component of the SDF.

A noteworthy element of our approach is that the bound on the entropy of \( m^2_{t,t+1} \) is constructed based on the notion that \( m_{t,t+1} \) correctly prices the risk-free bond, the long-term discount bond, and, finitely, many risky assets. Enabling a crucial dimension of model assessment, we show that a model that is rejected using the bound on the entropy of \( m_{t,t+1} \) may not be rejected based on the bound on the entropy of \( m^2_{t,t+1} \). Complementing our work, we show that an important diagnostic check is whether the higher moments of
\(m_{t,t+1}\) from a model are reasonable.

Our exercises are important, as the quest for well-performing SDFs has dominated the agenda in asset pricing. Despite substantial progress, identifying the desirable properties of the SDFs, and the embedded permanent component, in addition to their link to economic fundamentals, remains an unresolved issue. Overall, our bounds convey economic interpretations, can encapsulate data considerations that transcend model calibrations, and our framework can incorporate statistical concerns in model assessment.

**Related literature:** Our work belongs to a branch of asset pricing that explores the relevance of volatility bounds and entropy bounds to distinguish among models. The new angle in our paper is that the entropy of \(m_{t,t+1}^2\) is related to the maximum expected excess (log) return of the security that pays the SDF. Our approach also lies within the tradition of examining the SDFs, together with their correlated permanent and transitory components (e.g., Alvarez and Jermann (2005), Hansen and Scheinkman (2009), and Bakshi and Chabi-Yo (2012)), and we propose new entropy restrictions to evaluate asset pricing models.

In the vein of Hansen and Jagannathan (1991), our entropy representations are cast in a framework in which the SDF correctly prices finitely many asset returns (when an SDF correctly prices a portfolio, it is not tantamount to correctly pricing each of the assets constituting the portfolio). The new entropy bounds are parameterized in terms of both a vector of expected returns and a variance covariance matrix of returns, and they have no analytical analogs.

### 2. A new entropy bound when the SDF correctly prices finitely many returns

We employ a result in Alvarez and Jermann (2005, Proposition 1) and Hansen and Scheinkman (2009, page 200), who establish that \(m_{t,t+1}\) admits a multiplicative decomposition:

\[
m_{t,t+1} = m_{t,t+1}^P m_{t,t+1}^T \quad \text{with} \quad E[m_{t,t+1}^P] = 1 \quad \text{and} \quad m_{t,t+1}^T = (R_{t,t+1,\infty})^{-1},
\]
where \( m_{t,t+1}^P \) (or \( m_{t,t+1}^T \)) is the permanent (transitory) component of the stochastic discount factor \( m_{t,t+1} > 0 \), \( R_{t,t+1,\infty} \) is the gross return of an infinite-maturity discount bond, and \( E[\cdot] \) is unconditional expectation.

Alvarez and Jermann (2005, Proposition 1) show that the \( m_{t,t+1}^P \) component of the SDF is unique when \( m_{t,t+1}^T = (R_{t,t+1,\infty})^{-1} \). It is the case of uniquely identified \( m_{t,t+1}^P \) that is of economic interest.

The components \( m_{t,t+1}^P \) and \( m_{t,t+1}^T \) can be correlated, and, if they exist, can be obtained by solving the eigenfunction problem of Hansen and Scheinkman (2009, Corollary 6.1). In the context of parameterized models, both Hansen (2012) and Christensen (2014) show that an appropriately solved eigenfunction problem will ensure a unique \( m_{t,t+1}^P \).

Our objective is to propose a bound on the entropy of \( m_{t,t+1}^2 \), and its permanent component, to evaluate models when \( m_{t,t+1} \) is required to price many asset returns. Hansen and Jagannathan (1991, equation (3)) show that excluding the full pricing information in the analysis can weaken the implications for \( m_{t,t+1} \).

2.1. Rationale for a new entropy measure in asset pricing tests and economic interpretation

We assume \( E[m_{t,t+1}^{1+\theta}] < +\infty \) for some \( \theta \geq 1 \). The entropy of the random variable \( m_{t,t+1} \) is defined as:

\[
L[m_{t,t+1}] = \log(E[m_{t,t+1}]) - E[\log(m_{t,t+1})].
\]

The purpose of this subsection is to motivate a new entropy measure, specifically \( L[m_{t,t+1}^2] \) (or \( L[(m_{t,t+1}^P)^2] \)), as a metric for evaluating asset pricing models, defined as:

\[
L[m_{t,t+1}^2] = \log(E[m_{t,t+1}^2]) - E[\log(m_{t,t+1}^2)].
\]

Our results are not affected by the length of the period for \( m_{t,t+1} \), which can be arbitrary. Guided by this feature, we use the notations \( m \) and \( m_{t,t+1} \) interchangeably.
To fix ideas, we note that the measure $L[m^2]$ is related to Jensen’s gap, $J\{m\}$, defined as:

$$J\{m\} \equiv E[f\{m\}] - f\{E[m]\} \geq 0,$$

applied to the convex function $f\{m\} = h\{g\{m\}\}$, \(5\)

where $h\{m\} = -\log(m)$, $g\{m\} = m^2$, and $f\{m\}$ is a composition of two convex functions. \(6\)

A second-order Taylor expansion of $h\{g\{m\}\}$ around $E[g\{m\}]$ is

$$h\{g\{m\}\} \approx h\{E[g\{m\}]\} + (g\{m\} - E[g\{m\}]) h'\{E[g\{m\}]\} + \frac{1}{2} (g\{m\} - E[g\{m\}])^2 h''\{E[g\{m\}]\}.$$ \(7\)

Therefore, for $h\{m\} = -\log(m)$ and $g\{m\} = m^2$, one obtains

$$L[m^2] = \frac{E[h\{g\{m\}\}]}{\text{Jensen's gap}} - h\{E[g\{m\}]\} \approx \frac{1}{2} \text{Var}[m^2] \left( \frac{1}{E[m^2]} \right)^2.$$ \(8\)

In contrast, the variance measure used in Hansen and Jagannathan (1991) is related to Jensen’s gap, applied to the convex function $f\{m\} = m^2$. The entropy measure featured in Stutzer (1995), Bansal and Lehmann (1997), Alvarez and Jermann (2005), Backus, Chernov, and Zin (2014), and Ghosh, Julliard, and Taylor (2012) corresponds to $f\{m\} = -\log(m)$, with $L[m] \approx \frac{1}{2} \text{Var}[m] \left( \frac{1}{E[m]} \right)^2$. The upshot is that the composition $f\{m\} = h\{g\{m\}\}$ magnifies convexity and widens Jensen’s gap in comparison to $f\{m\} = -\log(m)$.

Equation (8) reveals that $L[m^2]$ can be seen as reflecting the dispersion of $m^2$ around its expectation. In particular, a probability distribution of $m$ that incorporates fatter tails tend to support a higher $L[m^2]$.

Our choice of the convex function $f\{m\} = -\log(m^2)$ in equation (5), which generates the entropy of $m^2$, belongs to the family of Cressie and Read (1984) functions, defined, for some $\theta$, as $\text{CR}\{m\} = \frac{1}{\theta(1+\theta)} (m^{\theta+1} - 1)$, and considered also in the work of Borovicka, Hansen, and Scheinkman (2015, Section 8.1). One obtains $-\log(m^2) = \lim_{\theta \to -1} \text{CR}\{m^2\}$, in contrast to $-\log(m) = \lim_{\theta \to -1} \text{CR}\{m\}$.

There is an important economic interpretation associated with $L[m^2]$. In particular, $L[m^2]$ encodes information about the expected excess (log) return of a fundamental asset, namely, the security that entitles
the investor a payoff of \( m_{t,t+1} \) and has return \( r_{t,t+1}^{\text{SDF}} \) (as defined in equation (1)). We now prove.

**Theorem 1** The expected excess (log) return of a security with SDF payoff is related to \( L[m_{t,t+1}^2] \) as follows:

\[
L[m_{t,t+1}^2] \geq E[\log(R_t^f)] - E[\log(1 + r_{t,t+1}^{\text{SDF}})] \geq 0.
\]  

(9)

**Proof:** See Appendix A.

The entropy of \( m^2 \) is the maximum expected excess (log) return of a security that pays \( m \). Attesting to their importance, claims on the upside of \( m \) are considered by Bakshi, Madan, and Panayotov (2010, equation (12)), and some related claims on \( m \) are at the center of Schneider (2015). Moreover, the security with a payoff of \( m \) is not the same as the security that pays \( m \) with minimum variance, as formalized by Hansen and Richards (1987, Lemma 3.1, page 596).

Our interpretation of the dispersion measure in equation (9) is distinct from others:

- Hansen and Jagannathan (1991, equations (15) and (16), page 235) show that \( \sqrt{\text{Var}[m]/(E[m])^2} \) is bounded by minus the Sharpe ratio of a security that pays the SDF with minimum variance.

- Alvarez and Jermann (2005, equation (A.1), page 2008) and Backus, Chernov, and Zin (2014, equation (5), page 57) show that the entropy of \( m \) is denominated in units of expected (log) gross return of a generic portfolio in excess of (log) gross return of a risk-free bond.

Our insight is that \( L[m^2] \) maps to the expected excess (log) return of the security with SDF payoff.\(^1\)

To highlight the nature of the information contained in \( L[m^2] \), we additionally note that the three dispersion measures are interrelated (subtract twice of \( L[m] \) in equation (3) from \( L[m^2] \) in equation (4)) as:

\[
L[m^2] = 2L[m] + \log \left( 1 + \frac{\text{Var}[m]}{(E[m])^2} \right).
\]  

(10)

\(^1\)We can interpret \( L[(m'_{t,t+1})^2] \). The security with the payoff \( m_{t,t+1}'/m_{t,t+1}^2 \) has a time \( t \) price of \( E_t[(m'_{t,t+1})^2] \), and a well-defined return: \( r_{t,t+1}^{PSDF} \equiv \frac{m_{t,t+1}'/m_{t,t+1}^2}{E[(m'_{t,t+1})^2]} - 1 \). Adapting the proof of Theorem 1, we can show that \( L[(m'_{t,t+1})^2] \geq E[\log(R_{t,t+1,\infty})] - E[\log(1 + r_{t,t+1}^{PSDF})] \). This contrasts Alvarez and Jermann (2005, Proposition 3, equation (4)), who establish that \( L[m_{t,t+1}' \log \) return of a generic portfolio in excess of expected (log) return of a long-term discount bond.
The relation in equation (10) prompts three observations. First, Hansen and Jagannathan (1991) derive the lower bound on \( \text{Var}[m] \), when the SDF correctly prices finitely many assets. Second, the lower bound on \( L[m] \) is not known in the general case, when the SDF correctly prices finitely many assets and is presented here in equation (28). Third, when the bounds are not unique, then lower bounding the parts in a sum may not be a proper way to lower bound the sum in equation (10).\(^2\) We will present a theoretical lower bound on \( L[m^2] \) and \( L[(m_P)^2] \) that can be inferred from a vector of traded asset returns.

While developing the implications of the new entropy measure, we ask another question: what do we gain when \( L[m^2] \) (or \( L[(m_P)^2] \)) is applied to asset pricing problems? We note that \( L[m^2] \) (or \( L[(m_P)^2] \)) offers flexibility in detecting non-normalities in \( \log(m) \) (\( \log(m_P) \)). Using the fact that when \( m \) is lognormal \( 2L[m] = \text{Var}[\log(m)] \approx \text{Var}[m] \), together with the approximation \( \log(1+x) \approx x \), it holds that

\[
L[m^2] = 2L[m] + \log \left(1 + \frac{2L[m]}{(E[m])^2}\right) \approx 4L[m], \quad \text{(when } E[m] \approx 1) \tag{11}
\]

\[
L[(m_P)^2] = 2L[m_P] + \log \left(1 + \frac{2L[m_P]}{(E[m_P])^2}\right) \approx 4L[m_P]. \quad \text{\(m_P\) is a martingale with } E[m_P] = 1 \tag{12}
\]

In Colacito, Ghysels, and Meng (2013, equation (11)), \( \log(m) \) is not normal and \( L[m^2] \neq 4L[m] \), illustrating that \( L[m^2] \) could be a suitable candidate for evaluating models under deviations from lognormality.

One may ask why not choose the entropy of any \( m^j \), for \( j > 2 \), as test functions to study asset pricing models. For example, it can be shown that the entropy of \( m^3 \) corresponds to the maximum expected excess (log) return of the squared contract on \( m \), i.e., a security with payoff equal \( m^2 \) and price \( E_t(m^3) \). While the entropy of \( m^3 \) may prove useful in some applications, we maintain focus on \( L[m_{t,t+1}^2] \) for reasons discussed.

Overall, our results indicate that \( L[m^2] \) is a viable measure of dispersion, capturing the expected squared deviation of \( m^2 \) around its mean. Provided \( L[m] \neq 0 \), \( L[m^2] \) is also different from the volatility measure of Hansen and Jagannathan (1991). Additionally, \( L[m^2] \) is different from \( 4L[m] \), when \( m \) is not lognormally distributed. An essential distinguishing trait of \( L[m^2] \) is its ability to more effectively cope with the effect of distributional non-normalities in \( \log(m) \).

\(^2\)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.
2.2. Characterizing the bounds on the entropies \( L[m_{t,t+1}^2] \) and \( L[(m_{t,t+1}^P)^2] \)

This subsection features a theoretical bound on \( L[m_{t,t+1}^2] \) and \( L[(m_{t,t+1}^P)^2] \), when the SDF is required to correctly price finitely many returns. To develop these bounds, consider the following set \( S \) of SDFs:

\[
S = \{ m_{t,t+1} > 0 : E_t[m_{t,t+1} = q_t], \ E[m_{t,t+1}R_{t,t+1,\infty}] = 1, \ \text{and} \ E[m_{t,t+1}R_{t,t+1}] = 1 \},
\]

where 1 is a vector column of ones. Moreover, \( R_{t,t+1} \) is an \( N \times 1 \) vector of gross returns that excludes the risk-free bond and the infinite-maturity discount bond.

We further postulate that some SDFs that belong to the set \( S \) can be uniquely decomposed into permanent and transitory components:

\[
S_P = \{ m_{t,t+1} \in S : m_{t,t+1} = m_{t,t+1}^P m_{t,t+1}^T, \ \text{and} \ m_{t,t+1}^T = (R_{t,t+1,\infty})^{-1} \}.
\]

Equally important, equation (13) requires the SDF to correctly price each of the \( N + 2 \) distinct assets. Therefore, due to the pricing of additional risky assets, set \( S \) is smaller than its counterparts based on pricing three assets, for example, a risk-free bond, a long-term bond, and a generic portfolio of risky assets.

**Theorem 2** The following theoretical bounds are germane to asset pricing models:

(a) The entropy of \( m_{t,t+1}^2 \) satisfies:

\[
L[m_{t,t+1}^2] \geq \mathrm{LB}_{m^2} \equiv 2 \left( E \left[ \log \left( \frac{(1 - E[q_t]E[R_{t,t+1}])\Sigma^{-1}(1 - E[q_t]E[R_{t,t+1}])}{1\Sigma^{-1}(1 - E[q_t]E[R_{t,t+1}])} \right) \right] - \log \left( E[q_t]^{-1} \right) \right) + \log \left( 1 + (1 - E[q_t]E[R_{t,t+1}])\Sigma^{-1}(1 - E[q_t]E[R_{t,t+1}]) / (E[q_t]^2) \right),
\]

where \( \Sigma \) is the variance-covariance matrix of \( R_{t,t+1} \).
(b) The entropy of \( (m_{t,t+1}^P)^2 \) satisfies

\[
L[(m_{t,t+1}^P)^2] \geq \mathbb{E}[(m^P)^2] \equiv 2 \left( E \left[ \log \left( \frac{(1 - E[q_t]E[R_{t,t+1}])^T \Sigma^{-1} (1 - E[q_t]E[R_{t,t+1}])}{(1 - E[R_{t,t+1}])} \right) - E \left[ \log (R_{t,t+1,\infty}) \right] \right) + \log \left( 1 + (1 - E[R_{t,t+1}/R_{t,t+1,\infty}])^T \Sigma_P^{-1} (1 - E[R_{t,t+1}/R_{t,t+1,\infty}]) \right),
\]

(16)

where \( \Sigma_P \) is the variance-covariance matrix of \( R_{t,t+1}/R_{t,t+1,\infty} \).

**Proof:** See Appendix B.

The entropy bounds stipulated in equations (15) and (16) summarize properties of the distribution of \( m_{t,t+1} \) and \( m_{t,t+1}^P \) and, hence, contain information that could help to gauge asset pricing models. Moreover, the lower bounds presented in the right-hand side of equations (15) and (16) in Theorem 2 are computable from the time-series of asset returns and discount bonds, and are model-free.

One may interpret the lower bound on \( L[m^2] \) in equation (15) of Theorem 2 as having two economically meaningful components. The first term surrogates an excess rate of return, whereas the second term is proportional to a Sharpe ratio-related component. Equations (15) and (16) indicate that the bound on \( L[m^2] \) and \( L[(m^P)^2] \) is determined by the vector of mean returns and a quadratic form of mean and variance of the vector of returns.

The lower bound on \( L[(m^P)^2] \) in equation (16) is distinct from the lower bound on \( \text{Var}[m^P] \) in Bakshi and Chabi-Yo (2012, equation (6)). Analogously, the lower bound on \( \text{Var}[m] \), that is, the Hansen and Jagannathan (1991, equation (12)) bound, and our lower bound on \( L[m^2] \) in equation (15), constitute distinctly relevant metrics for evaluating asset pricing models. Moreover, Ghosh, Julliard, and Taylor (2012, Section II.1) construct entropy bounds when the SDF can be factorized into observable and model-specific unobservable components. Our entropy bounds on the SDF are distinct from their bounds, allow correlated multiplicative components, and can be inferred from the returns data.

Liu (2012, 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 \( \gamma \equiv \frac{1}{1-\delta} \).
Moreover, our results in Theorem 2 can be contrasted to the single-return-based bound on the generalized entropy function in Liu (2012, equations (11) and (12)). In addition, our entropy bound on \( L[m^2] \) offers a distinction to the bounds considered in Snow (1991, equations (7) and (12)). Specifically, our bounds are easier to implement and do not involve solving an optimization problem.

3. Analyzing asset pricing models

Our benchmark for assessing whether a model produces sufficient entropy are the bounds in Theorem 2. We compare our results with those from the bound on the entropy of \( m \) and the volatility of \( m \). Moreover, we consider a bootstrap procedure to judge whether a model statistically meets the data-based lower bounds.

3.1. Asset pricing models

Our goal is to learn about the properties of \( m_{t,t+1} \) and \( m^{\rho}_{t,t+1} \), and their consistency with bound restrictions. We focus on three models: (i) difference habit, (ii) recursive utility with stochastic variance, and (iii) recursive utility with constant jump intensity. Our analysis can be expanded to consider other models.

3.1.1. Difference habit model

In the difference habit model (e.g., Campbell and Cochrane (1999)), the SDF is

\[
m_{t,t+1} = \beta g_{t+1}^{\rho-1} \left( s_{t+1}/s_t \right)^{\rho-1},
\]

(17)

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 time \( t \). The laws of motion for \( h_t \) and \( g_t \) are

\[
\log(h_{t+1}) = \log(h) + \eta [B] \log(c_t) \quad \text{and} \quad \log(g_{t+1}) = \log(g) + \gamma [B] u^z \omega_{g,t+1},
\]

(18)
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, $\nu$ 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 (18), leads to the following dynamics:

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

(19)

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

$$
x_t = \gamma_1 \nu^{1/2} \omega_{gt} + \varphi_g x_{t-1} \quad \text{with} \quad \varphi_g = \frac{\gamma_2}{\gamma_1}.
$$

(20)

The key step is to solve the eigenfunction problem to recover $m^p_{t,t+1}$: $E_t[m_{t,t+1} e_{t+1}] = \zeta e_t$, where $\zeta$ is the eigenvalue and $e_{t+1}$ is the eigenfunction (Hansen and Scheinkman (2009, Definition 6.1)). For the SDF of the habit model specified in equation (17), the permanent component is:

$$
m^p_{t,t+1} = \exp(-D_1 + D_2 x_{t-1} + D_3 x_t + D_4 x_{t+1}),
$$

(21)

where $D_1$ through $D_4$ are defined in equations (A17) through (A20) of the Internet Appendix (Section I).

We employ equation (21) to compute the left-hand side of equation (16) of Theorem 2. 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 (2012), Chapman (1998), Chan and Kogan (2002), and Santos and Veronesi (2010)).
3.1.2. Recursive utility models

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

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

(22)

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

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 dynamics:

\[
\begin{align*}
\log(g_t) &= \log(g) + \gamma [B] \omega_{gt} + \psi [B] z_{gt} - \psi [1] h \theta, \quad h_t = h + \eta [B] \omega_{ht}, \\
v_t &= v + v [B] \omega_{vt}, \quad z_{gt} \mid j \sim \mathcal{N}(j \theta, j \delta^2), \quad P[j] = \exp(-h_{t-1}) \binom{h_{t-1}}{j} / j!, \\
\end{align*}
\]

(23, 24)

where \( \omega_{gt}, z_{gt}, \) 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 \( t \) is \( e^{h_{t-1}} h_{t-1}^j / j! \), and the jump intensity, \( h_{t-1} \), is the mean of \( j \).

A. Recursive utility model with stochastic variance. Set \( h = 0, \eta [B] = 0, \psi [B] = 0 \) in equations (23) and (24). For tractability, we consider the evolution of the transformed variable:

\[ x_t = \Phi_g x_{t-1} + \gamma_1 \nu_{t-1} \omega_{gt}. \]

(25)

Then the permanent component of the SDF of the recursive utility model with stochastic variance is:

\[ m_{t,t+1}^P = \exp(H_6 + (H_2 - \tau_0)x_t + (H_3 + \tau_0)x_{t+1} + (H_4 - \tau_1)x_t + (H_5 + \tau_1)x_{t+1}), \]

(26)

where \( H_2 \) through \( H_6, \tau_0, \) and \( \tau_1 \) are presented in the Internet Appendix (Section II).
B. Recursive utility model with constant jump intensity: In equations (23) and (24), set \( \nu[B] = 0 \). Then the permanent component of the SDF of the recursive utility model with constant jump intensity is:

\[
m^P_{t,t+1} = \exp \left( G_9 - G_8 h_t + (G_5 + \xi_1) z_{gt+1} + (G_6 + \xi_2 \gamma_1) \nu^2 \omega_{gt+1} + (G_7 + \xi_0 \eta_0) \omega_{ht+1} \right),
\]

where \( G_5 \) through \( G_9 \), \( \xi_0 \) through \( \xi_3 \), and \( \eta_0 \) are presented in the Internet Appendix (Section II).

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. Notable applications include, among others, Epstein and Zin (1991), Bansal and Yaron (2004), Campbell and Vuolteenaho (2004), Hansen, Heaton, and Li (2008), Martin (2013), Wachter (2013), and Zhou and Zhu (2009). Wachter (2013) emphasizes that her model can reconcile the size of the equity premium, the behavior of equity volatility, and the return predictability of Treasury bonds, pointing to a possible link between seemingly disparate phenomena from equity and bond markets.

3.2. There is empirical rationale for considering the entropy of \( L[m^2] \) and \( L[(m^P)^2] \)

Pertinent to our empirical inquiry is first the question: How meaningful are our entropy bounds on \( m^2_{t,t+1} \) and \( (m^P_{t,t+1})^2 \)? To answer this question, we need to show that entropy \( L[m^2_{t,t+1}] \) (or \( L[(m^P_{t,t+1})^2] \)) contains information beyond that which is contained in entropy \( L[m_{t,t+1}] \) (or \( L[m^P_{t,t+1}] \)).

When \( m \) is lognormally distributed in the setting of equation (11): \( L[m^2]/4L[m] - 1 = 0 \). One implication of this restriction is that the bound on \( L[m^2] \) is proportional to the bound on \( L[m] \), which is amenable to validation from the data. However, a problem that arises is that our bound on \( L[m^2] \) is based on \( m \) correctly pricing \( N + 2 \) returns, whereas the available bound on \( L[m] \) from Backus, Chernov, and Zin (2014, equation (5), page 57) is based on \( m \) correctly pricing a generic portfolio return. Such a difference in the formulation of the bounds can confound inference in our exercises. The following results are aimed at remedying the situation by making the bounds on \( L[m] \) and \( L[m^P] \) sharper. Specifically, when \( m_{t,t+1} \) correctly prices \( N + 2 \)
returns, the entropy of $m_{t,t+1}$ and $m_{t,t+1}^P$ satisfy (proof is in Section III of the Internet Appendix):

$$L[m_{t,t+1}] \geq \mathbb{L}[m] \equiv E \left[ \log \left( \frac{\Sigma^{-1} (1 - E[q_t]E[R_{t,t+1}])' R_{t,t+1}}{\Sigma^{-1} (1 - E[q_t]E[R_{t,t+1}])} \right) \right] - \log \left( (E[q_t])^{-1} \right),$$  \hspace{1cm} (28)

$$L[m_{t,t+1}^P] \geq \mathbb{L}[m^P] \equiv E \left[ \log \left( \frac{\Sigma^{-1} (1 - E[q_t]E[R_{t,t+1}])' R_{t,t+1}}{\Sigma^{-1} (1 - E[q_t]E[R_{t,t+1}])} \right) \right] - E \left[ \log (R_{t,t+1,\infty}) \right],$$  \hspace{1cm} (29)

where, as before, $\Sigma$ is the variance-covariance matrix of $R_{t,t+1}$.

Combining the right-hand sides of equation (15) and equation (28) (and isomorphically equation (16) and equation (29) for $m_{t,t+1}^P$), we construct the quantities:

$$\Pi_m \equiv \frac{\mathbb{L}[m^2]}{4 \mathbb{L}[m]} - 1$$  \hspace{1cm} and  \hspace{1cm} $$\Pi_{m^P} \equiv \frac{\mathbb{L}[m^P]^2}{4 \mathbb{L}[m^P]} - 1.$$  \hspace{1cm} (30)

The hypothesis $\Pi_m = 0$ amounts to testing whether $L[m^2]$ and $L[m]$ impound the same information, when $m_{t,t+1}$ correctly prices $N+2$ returns (a proof is in Section IV of the Internet Appendix).

Table 1 provides a point estimate of $\Pi_m$, and also $\Pi_{m^P}$, for three sets of $R_{t,t+1}$, and a bootstrap $p$-value that tests whether $\Pi_m = 0$ versus $\Pi_m \neq 0$.

Our empirical analysis illustrates that the hypothesis of $\Pi_m = 0$ is rejected, whereby the data-based lower bound on $L[m_{t,t+1}^2]$ can depart from its $4L[m_{t,t+1}]$ counterparts by as much as 56.17%. The reported $p$-values are based on a block bootstrap, with a block size of 20, with 50,000 replications from the data. The hypothesis of $\Pi_{m^P} = 0$ is also rejected. Our evidence provides some rationale for considering $L[m^2]$ and $L[(m^P)^2]$ in assessing asset pricing models.

3.3. Implementation and calculation of model-based entropies

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 Appendix A, which indicates that each model reasonably calibrates to consumption growth data.

Aided by the analytical representations of \( m_{t,t+1} \) (for example, as derived in equation (21)), we generate the paths for \( m_{t,t+1} \), along with those of \( m_{t,t+1} \), over 966 months corresponding to our returns data over 1931:07 to 2011:12. The paths are based on the model parameters in Table Appendix A and shocks driving the fundamentals (e.g., \( \omega_{ut} \) and \( \omega_{gt} \) for the RU-SV). Then we obtain the sample averages of the series \( \{m_{t,t+1}^2, m_{t,t+1}, (m_{t,t+1})^2, m_{t,t+1}^2 : t = 1, \ldots 966\} \), and accordingly compute the entropies \( L[m_{t,t+1}^2], L[m_{t,t+1}], L[(m_{t,t+1})^2], L[m_{t,t+1}^2] \), and also the volatilities of \( m_{t,t+1} \) and \( m_{t,t+1}^2 \).

Next, we draw 50,000 paths for the shocks driving a model and, hence, obtain 50,000 paths for \( m_{t,t+1} \) and \( m_{t,t+1}^2 \). Panels A, B, and C of Table 2 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.

3.4. Model assessment based on the lower bound on \( L[m] \) and \( L[m^P] \)

The next question to ask is: how successful are the three models in generating \( L[m] \) that is consistent with the data? Panel A of Table 2 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, that is, the lower bound on \( L[m] \), displayed on the right-hand side of equation (28), computed based on SET B, all the models are rejected at the 5% level (as seen by the bootstrap \( p \)-values).

Such an implication from our bound, calculated using the return properties of the risk-free bond, the long-term discount 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 $L[m_{t,t+1}]$ constructed from the returns of a (single) generic portfolio may provide an insufficient hurdle in evaluating the merits of an asset pricing model. When the entropy calculations exploit the information in the distribution of the return vector $R_{t,t+1}$, it imposes stronger implications for $m_{t,t+1}$.

We now ask: Are the properties of $m^P$ implicit in the models consistent with the entropy bound $L[m^P]$? We find that the $L[m^P]$ produced by the DH, RU-SV, and RU-CJI models are 0.0203, 0.0237, and 0.0197, respectively, while the data-based average lower bound on $L[m^P]$ is 0.0348. The reported $p$-values indicate that all the three models are rejected at the 5% level; namely, the models generate insufficient entropy $L[m^P]$. In essence, the bounds on both $L[m]$ and $L[m^P]$ agree in suggesting that the models are misspecified.

### 3.5. Model assessment based on the bound on $L[m^2]$ and $L[(m^P)^2]$ yields insights

Elaborating further, we now argue that considering the entropy $L[m^2]$ (or $L[(m^P)^2]$) in the model assessment can provide an important contrast to our findings based on the entropy $L[m]$ (or $L[m^P]$).

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 $L[m_{t,t+1}^2]$ of 0.0785, 0.0869, and 1.4331, respectively (see the entries in Panel B of Table 2).

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 $L[m]$. 

15
Accordingly, one question emerges: Why does the RU-CJI fail to explain features of \( m \) (and likewise \( m^p \)), 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 \( L[m^2] \)-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 \( m_{t,t+1} \) and \( m^p_{t,t+1} \) for each of the models in Panel D and Panel E of Table 2. 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 \( L[m_{t,t+1}^2] \) (\( L[(m^p_{t,t+1})^2] \)) of 1.4331 (1.4858) 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 both \( m_{t,t+1} \) and \( m^p_{t,t+1} \)? 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 invoke the density of the poisson random variable to show that

\[
E_t[(m_{t,t+1})^k] = E_t[e^{k\log(m_{t,t+1})}] = E_t[E_t[e^{k\log(m_{t,t+1})|J}]] = e^{G[k]}E_t[e^{H[k]}J],
\]

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.

How general are our conclusions with respect to the RU-CJI model? Specifically, are there model combinations that produce reasonable higher-order moments of \( m^p_{t,t+1} \) (or \( m^p_{t,t+1} \)), calibrate well to consumption growth data, and yet deliver high entropies? To probe this issue, we vary the jump distribution parameters

\[3\] As noted in equation (10), the entropy of \( m \), the entropy of \( m^2 \), and the variance of \( m \) are related by the expression: \( \exp[L[m^2]] - 2L[m] - 1 = \frac{\text{Var}[m]}{E[m]} \). Such a relation suggests that it may be possible for a model to satisfy the bound on \( \text{Var}[m] \), but not the bound on \( L[m^2] \), and vice versa, and similarly for \( L[m] \).
(θ, δ, h) of the consumption growth dynamics (see equation (24)), and report some illustrative results in Table Appendix B. The takeaway message is that jump parameterizations (among the 27 parameter combinations) that yield reasonable levels of skewness and kurtosis of \( m_{t,t+1}^P \) do not appear to produce enough entropies to satisfy the lower bound on \( L[m_{t,t+1}^P] \) and \( L[(m_{t,t+1}^P)^2] \).

Finally, consider the volatility bound on \( m \) using the Hansen and Jagannathan (1991, equation (12)) and the volatility bound on \( m^P \) using Bakshi and Chabi-Yo (2012, equation (6)). As seen from Panel C of Table 2, the DH, RU-SV 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 \( m \).

4. Conclusions

A central problem in finance is the specification of the stochastic discount factor \( (m_{t,t+1}) \). 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 entropy of \( m_{t,t+1}^2 \) versus the variance measure and the entropy of \( m_{t,t+1} \). In particular, we establish that the entropy of \( m_{t,t+1}^2 \) is the maximum expected excess (log) return of a security that pays \( m_{t,t+1} \). The entropy restrictions we develop are based on the ability of the stochastic discount factor to jointly price the risk-free bond, the long-term discount bond, and a set of risky assets. Our approach is flexible and one can refine the bounds framework to incorporate conditioning information.
References


Appendix A: Proof of equation (9) of Theorem 1

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

There are two parts of the proof of equation (9). First, we establish that \(L[m^2_{t,t+1}] \geq E[\log(R_t^f)] - E[\log(1 + r_{t,t+1}^{SDF})]\). Next, we show that the security that pays the SDF is a hedging asset with expected return lower than the risk-free return.

Taking logs of the expression for \(1 + r_{t,t+1}^{SDF}\), and adding and subtracting \(\log(m_{t,t+1}^2)\), we obtain:

\[
\log(1 + r_{t,t+1}^{SDF}) = \log(m_{t,t+1}) - \log(E_t[m_{t,t+1}^2]) + \log(m_{t,t+1}^2) - 2\log(m_{t,t+1}).
\]  

(31)

Netting out \(\log(m_{t,t+1})\) and taking expectations on both sides of equation (31), we have:

\[
E_t[\log(1 + r_{t,t+1}^{SDF})] + E_t[\log(m_{t,t+1})] = E_t[\log(m_{t,t+1}^2)] - \log(E_t[m_{t,t+1}^2]).
\]  

(32)

Subtracting \(\log(E_t[m_{t,t+1}])\) from both sides of equation (32) and rearranging, it follows that

\[
L_t[m_{t,t+1}^2] + \log(E_t[m_{t,t+1}]) = -E_t[\log(1 + r_{t,t+1}^{SDF})] + \log(E_t[m_{t,t+1}]) - E_t[\log(m_{t,t+1})].
\]  

(33)

\[
\geq -E_t[\log(1 + r_{t,t+1}^{SDF})].
\]  

(34)

Rearranging, we obtain the expression for the conditional entropy of \(m_{t,t+1}^2\):

\[
L_t[m_{t,t+1}^2] \geq -E_t[\log(1 + r_{t,t+1}^{SDF})] - \log(E_t[m_{t,t+1}]).
\]  

(35)

Since the gross return of the risk-free bond satisfies \(E_t[m_{t,t+1}] = 1/R_t^f\), we obtain

\[
L_t[m_{t,t+1}^2] \geq \log(R_t^f) - E_t[\log(1 + r_{t,t+1}^{SDF})].
\]  

(36)
Taking unconditional expectations on both sides of equation (36):

\[ E[L_t[m^2_{t,t+1}]] \geq E[\log(R^f_t)] - E[\log(1 + r^{SDF}_{t,t+1})]. \]  

(37)

Now exploit the following relation:

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

(38)

Therefore, \( L[m^2_{t,t+1}] \geq E[\log(R^f_t)] - E[\log(1 + r^{SDF}_{t,t+1})]. \) 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 \( E_t[m^2_{t,t+1}] \geq (E_t[m_{t,t+1}])^2, \) because \( \text{Var}(m_{t,t+1}) > 0. \) Hence,

\[ E_t[1 + r^{SDF}_{t,t+1}] = \frac{E_t[m_{t,t+1}]}{E_t[m^2_{t,t+1}]} \leq \frac{E_t[m_{t,t+1}]}{(E_t[m_{t,t+1}])^2} = \frac{1}{E_t[m_{t,t+1}]} = R^f_t. \]  

(39)

Equation (39), thus, shows that \( E_t[1 + r^{SDF}_{t,t+1}] \leq R^f_t, \) which implies

\[ \log \left( E_t[1 + r^{SDF}_{t,t+1}] \right) \leq \log \left( \frac{1}{E_t[m_{t,t+1}]} \right). \]  

(40)

By an application of Jensen’s inequality,

\[ E_t[\log(1 + r^{SDF}_{t,t+1})] \leq \log \left( E_t[1 + r^{SDF}_{t,t+1}] \right) \leq \log(R^f_t). \]  

(41)

It then follows that

\[ \log(R^f_t) - E_t[\log(1 + r^{SDF}_{t,t+1})] \geq 0. \]  

Therefore,

\[ E[\log(R^f_t)] - E[\log(1 + r^{SDF}_{t,t+1})] \geq 0. \]  

(42)

The proof is complete.
Appendix B: Proof of Theorem 2

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

\[ y \equiv \Sigma^{-1} (1 - E[q_t] E[R_{t+1}]), \quad y_p \equiv \Sigma_p^{-1} (1 - E[R_{t+1} / R_{t+1,\infty}]), \quad \text{and} \quad a \equiv \frac{y}{\gamma y}, \]  

(43)

where \( q_t = E_t[m_{t+1}] \), \( \Sigma \) is the variance-covariance matrix of \( R_{t+1} \), and \( \Sigma_p \) is the variance-covariance matrix of \( R_{t+1} / R_{t+1,\infty} \). We assume that \( a' R_{t+1} \) is strictly positive. Further define,

\[ e_{rR} \equiv E \left[ \log \left( a' R_{t+1} \right) \right] - \log \left( (E[q_t])^{-1} \right), \]

(44)

\[ e_{r\infty} \equiv E \left[ \log \left( R_{t+1,\infty} \right) \right] - \log \left( (E[q_t])^{-1} \right). \]

(45)

Proof of the entropy bound on \( m_{t+1}^2 \) in equation (15). By the definition of entropy: \( L[m^2] = \log(E[m^2]) - E[\log(m^2)] \). Then

\[
L[m_{t+1}^2] = \log \left( E[m_{t+1}^2] \right) - 2 \log \left( E[q_t] \right) + 2L[m_{t+1}^2],
\]

\[
= \log \left( 1 + \frac{E[m_{t+1}^2] - (E[q_t])^2}{(E[q_t])^2} \right) + 2L[m_{t+1}^2],
\]

\[
= \log \left( 1 + \frac{\text{Var} \left[ m_{t+1}^2 \right]}{(E[q_t])^2} \right) + 2L[m_{t+1}^2],
\]

\[
\geq \log \left( 1 + \frac{\text{Var} \left[ m_{t+1}^2 \right]}{(E[q_t])^2} \right) + 2e_{rR}. \quad \text{(as } L[m_{t+1}] \geq e_{rR}; \text{ see (44) and (C5)})
\]

(46)

Because \( E[m_{t+1} R_{t+1}] = 1 \) and setting \( q_t = E[m_{t+1}] \), it follows that

\[
E \left[ m_{t+1} \left( R_{t+1} - E \left[ R_{t+1} \right] \right) \right] = \left( 1 - E[q_t] E \left[ R_{t+1} \right] \right).
\]

(47)

Multiplying equation (47) by \( (1 - (E[q_t]) E \left[ R_{t+1} \right])' \Sigma^{-1} \) yields

\[
\left( 1 - (E[q_t]) E \left[ R_{t+1} \right] \right)' \Sigma^{-1} \left( 1 - E[q_t] E \left[ R_{t+1} \right] \right) = E \left[ m_{t+1} \left( 1 - (E[q_t]) E \left[ R_{t+1} \right] \right)' \Sigma^{-1} \left( R_{t+1} - E \left[ R_{t+1} \right] \right) \right].
\]

(48)
Applying the Cauchy-Schwarz inequality to the right-hand side of equation (48), it can be shown that

\[
\text{Var}[m_{t,t+1}] \geq (1 - (E[q_t]) E[R_{t,t+1}])^{-1} (1 - (E[q_t]) E[R_{t,t+1}]),
\]

\[
\geq (1 - (E[q_t]) E[R_{t,t+1}])^{-1} \Sigma^{-1} (1 - (E[q_t]) E[R_{t,t+1}]),
\]

\[
\geq y' \Sigma y.
\]

(49)

Combining the expressions in equations (46) and (49), we obtain the bound on \(L[m_{t,t+1}^2]\) presented in equation (15) of Theorem 2.

Proof of the entropy bound on \((m_{t,t+1}^p)^2\) in equation (16) of Theorem 2. We write

\[
L[(m_{t,t+1}^p)^2] = \log(E[(m_{t,t+1}^p)^2]) - E[\log((m_{t,t+1}^p)^2)],
\]

\[
= \log(E[(m_{t,t+1}^p)^2]) - 2E[\log(m_{t,t+1}^p)],
\]

\[
= \log(E[(m_{t,t+1}^p)^2]) + 2L[m_{t,t+1}^p],
\]

\[
= \log(1 + \text{Var}[m_{t,t+1}^p]) + 2L[m_{t,t+1}^p].
\]

(50)

We show in equation (C9) that \(L[m_{t,t+1}^p] \geq \text{er}_R - \text{er}_\infty\) (the complete expressions for \(\text{er}_R\) and \(\text{er}_\infty\) are in equations (44) and (45), respectively). Therefore, we deduce that

\[
L[(m_{t,t+1}^p)^2] \geq \log(1 + \text{Var}[m_{t,t+1}^p]) + 2(\text{er}_R - \text{er}_\infty).
\]

(51)

Since \(E[m_{t,t+1}R_{t,t+1}] = E[m_{t,t+1}^p R_{t,t+1}^{R_{t,t+1}}] = 1\), we then obtain:

\[
E \left[ m_{t,t+1}^p \left( \frac{R_{t,t+1}}{R_{t,t+1,\infty}} - E \left[ \frac{R_{t,t+1}}{R_{t,t+1,\infty}} \right] \right) \right] = 1 - E \left[ \frac{R_{t,t+1}}{R_{t,t+1,\infty}} \right].
\]

(52)

Multiplying each side of equation (52) by \(\left(1 - E \left[ \frac{R_{t,t+1}}{R_{t,t+1,\infty}} \right] \right)^{-1} \Sigma_p^{-1}\), we get

\[
\left(1 - E \left[ \frac{R_{t,t+1}}{R_{t,t+1,\infty}} \right] \right)^{-1} \Sigma_p^{-1} \left(1 - E \left[ \frac{R_{t,t+1}}{R_{t,t+1,\infty}} \right] \right) = E \left[ m_{t,t+1}^p \left(1 - E \left[ \frac{R_{t,t+1}}{R_{t,t+1,\infty}} \right] \right)^{-1} \Sigma_p^{-1} \left( \frac{R_{t,t+1}}{R_{t,t+1,\infty}} - E \left[ \frac{R_{t,t+1}}{R_{t,t+1,\infty}} \right] \right) \right].
\]

(53)
Applying the Cauchy-Schwarz inequality to the right-hand side of equation (53), we note that

\[
\text{Var}[m_{t,t+1}^p] \geq \left(1 - E\left[\frac{R_{t,t+1}}{R_{t,t+1=\infty}}\right]\right) \Sigma_p^{-1} \left(1 - E\left[\frac{R_{t,t+1}}{R_{t,t+1=\infty}}\right]\right),
\]

\[
\geq \left(1 - E\left[\frac{R_{t,t+1}}{R_{t,t+1=\infty}}\right]\right) \left(\Sigma_p^{-1}\right)^t \Sigma_p^{-1} \left(1 - E\left[\frac{R_{t,t+1}}{R_{t,t+1=\infty}}\right]\right),
\]

\[
\geq y_p^t \Sigma_p y_p. \quad \text{(where noting } y_p \equiv \Sigma_p^{-1} \left(1 - E\left[\frac{R_{t,t+1}}{R_{t,t+1=\infty}}\right]\right)) \quad (54)
\]

Inserting the bound derived in equation (54) into equation (51) leads to the bound in equation (16) of Theorem 2.
Table 1  
**Relevance of our entropy bounds on** $m_{t,t+1}^2$ **and** $(m_{t,t+1}^P)^2$  
The logic of this test is that when the SDF (its permanent component) is lognormally distributed, then $L[m_{t,t+1}^2] = 4L[m_{t,t+1}]$ (or $L[(m_{t,t+1}^P)^2] = 4L[m_{t,t+1}^P]$). Accordingly, we consider the quantities

$$
\Pi_m \equiv \frac{\mathbb{LB}_{m^2}}{4\mathbb{LB}_m} - 1 \\
\Pi_{m^P} \equiv \frac{\mathbb{LB}((m^P)^2)}{4\mathbb{LB}_{m^P}} - 1,
$$

where $\mathbb{LB}_{m^2}$ and $\mathbb{LB}_m$ are defined in equations (15) and equation (28), respectively, and $\mathbb{LB}((m^P)^2)$ and $\mathbb{LB}_{m^P}$ are defined in equation (16) and equation (29), respectively.

In our implementation, we proxy $R_{t,t+1,\infty}$ by the monthly return of a 30-year Treasury bond. $\Sigma_p$ is the variance-covariance matrix of $R_{t,t+1}/R_{t,t+1,\infty}$, whereas $\Sigma$ is the variance-covariance matrix of $R_{t,t+1}$.

The entropy calculations are based on the SDF correctly pricing each of the $N + 2$ assets, and our computation of $\Pi_m$ and $\Pi_{m^P}$ relies on three data sets for $R_{t,t+1}$: SET A contains the value-weighted market returns together with the 25 Fama-French size and book-to-market portfolios (i.e, $N = 26$), SET B contains the value-weighted market returns together with the 25 Fama-French size and momentum portfolios (i.e, $N = 26$), and SET C contains only the value-weighted equity market returns (i.e, $N = 1$). The sample period is from July 1931 to December 2011 (966 observations). To compute the $p$-values reported in parentheses, we employ a block bootstrap with a block size of 20 to generate $\hat{b}=50,000$ samples from the original data. We then compute $\Pi_b = \Pi$ for $b = 1, \ldots, \hat{b}$, the cross-sectional average $\bar{\Pi}$, and the standard error $se(\bar{\Pi}) = \text{std}(\Pi) / \sqrt{\hat{b}}$ of $\bar{\Pi}$. Accordingly, we compute the $t$ statistic as $(\bar{\Pi} - 0) / se(\bar{\Pi})$. The absolute value of the $t$-statistic is then used to compute the two-sided $p$-value.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Pi_m$</td>
<td>56.17% (0.000)</td>
<td>33.26% (0.000)</td>
<td>21.89% (0.000)</td>
</tr>
<tr>
<td>$\Pi_{m^P}$</td>
<td>51.85% (0.000)</td>
<td>33.05% (0.000)</td>
<td>13.79% (0.021)</td>
</tr>
</tbody>
</table>
Table 2

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 companion results are presented for the permanent component of the SDF. 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 bounds on the entropy of \( m^2 \) and \((m^p)^2\) are based on equations (15) and (16) of Theorem 2, and rely on the ability of the SDF to correctly price \( N + 2 \) assets (the risk-free bond, the long-term discount 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 bounds on the entropy of \( m \) and \( m^p \) are based on equations (28) and (29), respectively, and also rely on the ability of the SDF to correctly price \( N + 2 \) assets (the risk-free bond, the long-term discount bond, and \( N \) risky assets). The lower bound on the volatility of \( m \) are based on Hansen and Jagannathan (1991, equation (12)) and the lower bound on the volatility of \( m^p \) are based on Bakshi and Chabi-Yo (2012, equation (6)). We focus on SET B, as it corresponds to the maximum lower bound on entropy measures (as in our Table Internet Appendix-I). Panels D and E present the variance, skewness, and kurtosis of \( m \) and \( m^p \), which are consistent with model parameterizations in Table Appendix A. The one-sided \( p \)-values (\( \langle \cdot \rangle \)) reported below the lower bounds, represent the proportion of bootstrap samples for which the lower bound is less than zero.

<table>
<thead>
<tr>
<th></th>
<th>Habit model</th>
<th>Recursive utility models</th>
<th>Lower bounds (Set B)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DH</td>
<td>RU-SV</td>
<td>RU-CJI</td>
</tr>
<tr>
<td><strong>Panel A: Entropies of ( m ) and ( m^p ), when ( m ) correctly prices ( N + 2 ) returns</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( L[m] )</td>
<td>0.0196</td>
<td>0.0217</td>
<td>0.0190</td>
</tr>
<tr>
<td></td>
<td>[0.000]</td>
<td>[0.000]</td>
<td>[0.000]</td>
</tr>
<tr>
<td>( L[m^p] )</td>
<td>0.0203</td>
<td>0.0237</td>
<td>0.0197</td>
</tr>
<tr>
<td></td>
<td>[0.000]</td>
<td>[0.000]</td>
<td>[0.000]</td>
</tr>
<tr>
<td><strong>Panel B: Entropies of ( m^2 ) and ((m^p)^2)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( L[m^2] )</td>
<td>0.0785</td>
<td>0.0869</td>
<td>1.4331</td>
</tr>
<tr>
<td></td>
<td>[0.000]</td>
<td>[0.000]</td>
<td>[1.000]</td>
</tr>
<tr>
<td>( L[(m^p)^2] )</td>
<td>0.0811</td>
<td>0.095</td>
<td>1.4858</td>
</tr>
<tr>
<td></td>
<td>[0.000]</td>
<td>[0.000]</td>
<td>[1.000]</td>
</tr>
<tr>
<td><strong>Panel C: Volatility bounds</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hansen and Jagannathan (1991)</td>
<td>0.0415</td>
<td>0.0444</td>
<td>3.344</td>
</tr>
<tr>
<td></td>
<td>[0.000]</td>
<td>[0.000]</td>
<td>[1.000]</td>
</tr>
<tr>
<td>Bakshi and Chabi-Yo (2012)</td>
<td>0.0403</td>
<td>0.0487</td>
<td>3.248</td>
</tr>
<tr>
<td></td>
<td>[0.000]</td>
<td>[0.000]</td>
<td>[1.000]</td>
</tr>
<tr>
<td><strong>Panel D: Moments of the ( m_{t,t+1} ) distribution</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variance</td>
<td>0.0403</td>
<td>0.0444</td>
<td>3.3438</td>
</tr>
<tr>
<td>Skewness</td>
<td>0.6041</td>
<td>0.6476</td>
<td>+( \infty )</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>3.6447</td>
<td>3.8061</td>
<td>+( \infty )</td>
</tr>
<tr>
<td><strong>Panel E: Moments of the ( m^p_{t,t+1} ) distribution</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variance</td>
<td>0.0415</td>
<td>0.0487</td>
<td>3.2480</td>
</tr>
<tr>
<td>Skewness</td>
<td>0.6142</td>
<td>0.6778</td>
<td>+( \infty )</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>3.6654</td>
<td>3.8786</td>
<td>+( \infty )</td>
</tr>
</tbody>
</table>
Table Appendix A

**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 autocorrelation are based on the annualized consumption growth.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>DH</th>
<th>RU-SV</th>
<th>RU-CJI</th>
<th>Data implied</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1931:07 to 2011:12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parameter</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Panel A: Preferences</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\rho)</td>
<td>-9.0000</td>
<td>0.3333</td>
<td>0.3333</td>
<td></td>
</tr>
<tr>
<td>(\alpha)</td>
<td>-9.0000</td>
<td>-9.0000</td>
<td>-9.0000</td>
<td></td>
</tr>
<tr>
<td>(\beta)</td>
<td>0.9980</td>
<td>0.9980</td>
<td>0.9980</td>
<td></td>
</tr>
<tr>
<td>(\phi_h)</td>
<td>0.9000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(s)</td>
<td>0.5000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Panel B: Consumption growth dynamics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\gamma_0)</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td></td>
</tr>
<tr>
<td>(\log(g))</td>
<td>0.0015</td>
<td>0.0015</td>
<td>0.0015</td>
<td></td>
</tr>
<tr>
<td>(\eta_0)</td>
<td>0.1000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\gamma_1)</td>
<td>0.0271</td>
<td>0.0271</td>
<td>0.0281</td>
<td></td>
</tr>
<tr>
<td>(\phi_{\gamma})</td>
<td>0.9790</td>
<td>0.9790</td>
<td>0.9690</td>
<td></td>
</tr>
<tr>
<td>(\nu^{1/2})</td>
<td>0.0099</td>
<td>0.0099</td>
<td>0.0079</td>
<td></td>
</tr>
<tr>
<td>(\nu_0)</td>
<td>0.23 \times 10^{-5}</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\phi_{\nu})</td>
<td>0.9870</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(h)</td>
<td></td>
<td></td>
<td></td>
<td>0.0008</td>
</tr>
<tr>
<td>(\theta)</td>
<td></td>
<td>-0.1500</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\delta)</td>
<td></td>
<td>0.1500</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\psi_0)</td>
<td></td>
<td></td>
<td>1.0000</td>
<td></td>
</tr>
<tr>
<td>(b_1)</td>
<td></td>
<td>0.9977</td>
<td>0.9979</td>
<td></td>
</tr>
<tr>
<td>Panel C: Consumption growth</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (annualized)</td>
<td>1.0192</td>
<td>1.0190</td>
<td>1.0189</td>
<td>1.0339</td>
</tr>
<tr>
<td>Std. Dev. (annualized)</td>
<td>0.0416</td>
<td>0.0415</td>
<td>0.0369</td>
<td>0.0287</td>
</tr>
<tr>
<td>Autocorrelation</td>
<td>0.2424</td>
<td>0.2433</td>
<td>0.1771</td>
<td>0.2386</td>
</tr>
</tbody>
</table>
Table Appendix B

**Impact of alternative jump parameterizations in the RU-CJI model**

Here we vary $\theta$, $\delta$, and $h$, which govern the distribution of jumps (see equation (24)) in the consumption growth dynamics for the RU-CJI model. We keep other parameters of the RU-CJI model to those specified in Table Appendix A. For each set of parameters, the reported values are averages across 50,000 replications. For each replication, we simulate the path of consumption growth $c_{t+1}/c_t$ over 966 months.

Following convention, we then compute the annualized consumption growth as $\exp\left(\frac{1}{12} \sum_{j=1}^{12} \log(c_{t+j}/c_{t+j-1})\right)$.

The reported model mean and standard deviation are based on the annualized consumption growth. For each parameter set, we also report the average values of entropy $L[m_{t+1}^p]$ and $L[(m_{t+1}^p)^2]$, as well as the central moments of the permanent component of the SDF. The bolded parameter set corresponds to Backus, Chernov, and Zin (2014, Model (4), Table 4).

<table>
<thead>
<tr>
<th>$\theta$</th>
<th>$\delta$</th>
<th>$h$</th>
<th>Entropies</th>
<th>Moments of $m_{t+1}^p$</th>
<th>$\frac{c_{t+1}}{c_t}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>$L[m_t^p]$</td>
<td>$L[(m_t^p)^2]$</td>
<td>Variance</td>
</tr>
<tr>
<td>-0.15</td>
<td>0.02</td>
<td>0.0002</td>
<td>0.011</td>
<td>0.046</td>
<td>0.024</td>
</tr>
<tr>
<td>-0.15</td>
<td>0.02</td>
<td>0.0004</td>
<td>0.011</td>
<td>0.050</td>
<td>0.027</td>
</tr>
<tr>
<td>-0.15</td>
<td>0.02</td>
<td>0.0008</td>
<td>0.012</td>
<td>0.057</td>
<td>0.033</td>
</tr>
<tr>
<td>-0.15</td>
<td>0.07</td>
<td>0.0002</td>
<td>0.011</td>
<td>0.053</td>
<td>0.030</td>
</tr>
<tr>
<td>-0.15</td>
<td>0.07</td>
<td>0.0004</td>
<td>0.012</td>
<td>0.062</td>
<td>0.039</td>
</tr>
<tr>
<td>-0.15</td>
<td>0.07</td>
<td>0.0008</td>
<td>0.013</td>
<td>0.082</td>
<td>0.057</td>
</tr>
<tr>
<td>-0.15</td>
<td>0.15</td>
<td>0.0002</td>
<td>0.013</td>
<td>0.403</td>
<td>0.459</td>
</tr>
<tr>
<td>-0.15</td>
<td>0.15</td>
<td>0.0004</td>
<td>0.015</td>
<td>0.764</td>
<td>1.083</td>
</tr>
<tr>
<td>-0.15</td>
<td>0.15</td>
<td>0.0008</td>
<td>0.020</td>
<td>1.486</td>
<td>3.248</td>
</tr>
<tr>
<td>-0.07</td>
<td>0.02</td>
<td>0.0002</td>
<td>0.011</td>
<td>0.043</td>
<td>0.022</td>
</tr>
<tr>
<td>-0.07</td>
<td>0.02</td>
<td>0.0004</td>
<td>0.011</td>
<td>0.043</td>
<td>0.022</td>
</tr>
<tr>
<td>-0.07</td>
<td>0.02</td>
<td>0.0008</td>
<td>0.011</td>
<td>0.044</td>
<td>0.023</td>
</tr>
<tr>
<td>-0.07</td>
<td>0.07</td>
<td>0.0002</td>
<td>0.011</td>
<td>0.044</td>
<td>0.023</td>
</tr>
<tr>
<td>-0.07</td>
<td>0.07</td>
<td>0.0004</td>
<td>0.011</td>
<td>0.046</td>
<td>0.024</td>
</tr>
<tr>
<td>-0.07</td>
<td>0.07</td>
<td>0.0008</td>
<td>0.011</td>
<td>0.049</td>
<td>0.027</td>
</tr>
<tr>
<td>-0.07</td>
<td>0.15</td>
<td>0.0002</td>
<td>0.012</td>
<td>0.115</td>
<td>0.096</td>
</tr>
<tr>
<td>-0.07</td>
<td>0.15</td>
<td>0.0004</td>
<td>0.012</td>
<td>0.188</td>
<td>0.177</td>
</tr>
<tr>
<td>-0.07</td>
<td>0.15</td>
<td>0.0008</td>
<td>0.014</td>
<td>0.333</td>
<td>0.356</td>
</tr>
<tr>
<td>-0.02</td>
<td>0.02</td>
<td>0.0002</td>
<td>0.011</td>
<td>0.043</td>
<td>0.022</td>
</tr>
<tr>
<td>-0.02</td>
<td>0.02</td>
<td>0.0004</td>
<td>0.011</td>
<td>0.043</td>
<td>0.022</td>
</tr>
<tr>
<td>-0.02</td>
<td>0.02</td>
<td>0.0008</td>
<td>0.011</td>
<td>0.043</td>
<td>0.022</td>
</tr>
<tr>
<td>-0.02</td>
<td>0.07</td>
<td>0.0002</td>
<td>0.011</td>
<td>0.043</td>
<td>0.022</td>
</tr>
<tr>
<td>-0.02</td>
<td>0.07</td>
<td>0.0004</td>
<td>0.011</td>
<td>0.044</td>
<td>0.022</td>
</tr>
<tr>
<td>-0.02</td>
<td>0.07</td>
<td>0.0008</td>
<td>0.011</td>
<td>0.045</td>
<td>0.023</td>
</tr>
<tr>
<td>-0.02</td>
<td>0.15</td>
<td>0.0002</td>
<td>0.011</td>
<td>0.069</td>
<td>0.048</td>
</tr>
<tr>
<td>-0.02</td>
<td>0.15</td>
<td>0.0004</td>
<td>0.012</td>
<td>0.096</td>
<td>0.075</td>
</tr>
<tr>
<td>-0.02</td>
<td>0.15</td>
<td>0.0008</td>
<td>0.013</td>
<td>0.149</td>
<td>0.131</td>
</tr>
</tbody>
</table>
New Entropy Restrictions and the Quest for Better Specified Asset Pricing Models

Gurdip Bakshi and Fousseni Chabi-Yo

Internet Appendix: Not for Publication

Abstract

The Internet Appendix provides detailed steps and expressions for some results presented in the paper. Section I (Section II) provides the solution of the eigenfunction problem for the difference habit formation (recursive utility) model. Section III provides a proof of the lower bound on the entropy of \( m \) and the lower bound on the entropy of \( m^p \) (equations (28) and (29), respectively), when \( m \) is required to correctly price finitely many assets. We also shed light on the sharpness of the lower bounds. Section IV is devoted to developing the testable restriction \( \frac{LB_m}{LB_m} - 1 = 0 \) (as presented in equation (30)).
I. Permanent component of the SDF of the difference habit model

For the law of motions of $h_t$ and $g_t$ in equation (18), we define the backshift operators $\eta[B]$ and $\gamma[B]$:

$$
\eta[B] = \sum_{j=0}^{\infty} \eta_j B^j \quad \text{and} \quad \gamma[B] = \sum_{j=0}^{\infty} \gamma_j B^j, \quad (A1)
$$

with $\eta_0 = 1 - \varphi_h$ and $\eta_{j+1} = \varphi_h \eta_j$, $j \geq 0$, and $\gamma_0 = 1$. Invoking a log linear approximation of $\log(s_t)$,

$$
\log(m_{t+1}) = D_0 + (\rho - 1) \frac{1}{s} (1 - (1 - s) \eta[B] B) \gamma[B] \nu^2 \omega_{gt+1}, \quad (A2)
$$

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

Given the log linear approximation $\log(s_t) \approx 1 + \frac{\log(g)}{s - 2}$, the dynamics of the surplus consumption ratio is

$$
\log(s_{t+1}) - \log(s_t) = \frac{(s-1)}{s} (\eta[B] B - 1) \log(g_{t+1}). \quad (A3)
$$

Therefore, we may write the log SDF as

$$
\log(m_{t+1}) = \log(\beta) + (\rho - 1) \log(g) + (\rho - 1) \left( \frac{s-1}{s} (\eta[B] B - 1) \log(g) \right)
+ (\rho - 1) \frac{1}{s} (1 - (1 - s) \eta[B] B) \gamma[B] \nu^2 \omega_{gt+1}. \quad (A4)
$$

To solve for the permanent and transitory components of the SDF, we write the log SDF as

$$
\log(m_{t+1}) = \log(\beta) + (\rho - 1) \log(g) + (\rho - 1) \left( \frac{s-1}{s} (\eta[B] B - 1) \log(g) \right)
- (\rho - 1) \frac{1}{s} (1 - s) \eta[B] B \nu - (\rho - 1) \frac{1}{s} (1 - s) \eta[B] B \nu^2 \omega_{gt+1}
+ (\rho - 1) \frac{1}{s} \nu^2 \omega_{gt+1}. \quad (A5)
$$
where

\[ x_t = (\gamma[B] - \gamma_0) \omega_{gt+1}, \quad \text{implying} \quad x_{t+1} - \varphi_g x_t = \gamma_1 \omega_{gt+1}. \quad (A6) \]

We simplify the log SDF as

\[
\log(m_{t,t+1}) = D_0 + (\rho - 1) \frac{1}{s} \left( 1 - \frac{1}{\gamma_1} \varphi_g \right) x_t + \frac{1}{s} \gamma_1 x_{t+1} \\
+ (\rho - 1) \frac{1}{s} (1 - s) \frac{1}{\gamma_1} \varphi_g - 1 \eta[B] x_{t-1} - (\rho - 1) \frac{1}{s} (1 - s) \frac{1}{\gamma_1} \eta[B] x_t. \quad (A7) \]

Consider the eigenfunction problem:

\[ E_t [m_{t,t+1} e_{t+1}] = \zeta e_t, \quad \text{where} \quad \zeta \text{ is the eigenvalue and } e_{t+1} \text{ is the eigenfunction}. \quad (A8) \]

Accordingly, the permanent and transitory components of the SDF are

\[ m_{t,t+1}^P = m_{t,t+1} \left( \frac{e_{t+1}}{\zeta e_t} \right) \quad \text{and} \quad m_{t,t+1}^T = \frac{\zeta e_t}{e_{t+1}}. \quad (A9) \]

We conjecture that \( e_{t+1} \) in equation (A8) is of the form:

\[ \log(e_{t+1}) = \delta[B] x_{t+1}, \quad \text{where} \quad \delta[B] = \sum_{j=0}^\infty \delta_j B^j \quad \text{with} \quad \delta_0 = 1. \quad (A10) \]

To verify the solution, we expand to the following:

\[
\log(m_{t,t+1}) + \log \left( \frac{e_{t+1}}{e_t} \right) = D_0 + (\rho - 1) \frac{1}{s} \left( 1 - \frac{1}{\gamma_1} \varphi_g \right) x_t + \frac{1}{s} \gamma_1 x_{t+1} \\
- (\rho - 1) \frac{1}{s} (1 - s) \frac{1}{\gamma_1} \varphi_g - 1 \eta[B] x_{t-1} - (\rho - 1) \frac{1}{s} (1 - s) \frac{1}{\gamma_1} \eta[B] x_t \\
+ (\rho - 1) \frac{1}{s} (1 - s) \frac{1}{\gamma_1} \eta[B] + \delta_0) x_{t+1}. \quad (A11) \]
Upon simplifying the expectation involving the eigenfunction problem, we derive $\zeta$ as

$$
\log(\zeta) = \log(\beta) + (\rho - 1) \log(g) + (\rho - 1) \frac{(s-1)}{s} (\eta B B^{-1}) \log(g) + \frac{1}{2} \left( (\rho - 1) \frac{1}{s \gamma_1} + \delta_0 \right) \gamma_1^2 \nu \\
+ \left( (\rho - 1) \frac{1}{s} \left( \frac{1 - \frac{1}{\gamma_1}}{\gamma_0} \right) + \left( (\rho - 1) \frac{1}{s \gamma_1} + \delta_0 \right) \varphi_s \right) x_t + (\rho - 1) \frac{1}{s} (1-s) \left( \frac{1}{\gamma_1} \varphi_s - 1 \right) \eta B x_{t-1} \\
+ \left( -(\rho - 1) \frac{1}{s} (1-s) \frac{1}{\gamma_1} \eta [B] - \delta [B] \right) x_t + (\delta [B] - \delta_0) x_{t+1}.
$$

(A12)

Using the identification approach, we deduce

$$
\log(\zeta) = D_0 + \frac{1}{2} ((\rho - 1) (s \gamma_1)^{-1} + \delta_0) \gamma_1^2 \nu, \tag{A13}
$$

and

$$
\delta_1 = - \left( (\rho - 1) \frac{1}{s} + \delta_0 \varphi_s \right) - \left( -(\rho - 1) \frac{1}{s} (1-s) \frac{1}{\gamma_1} \eta_0 - \delta_0 \right), \\
\delta_{j+1} = -(\rho - 1) \frac{1}{s} (1-s) \left( \frac{1}{\gamma_1} \varphi_s - 1 \right) \eta_{j-1} - \left( -(\rho - 1) \frac{1}{s} (1-s) \frac{1}{\gamma_1} \eta_j - \delta \right) \text{ for } j \geq 1. \tag{A14}
$$

Exploiting the solution to the eigenfunction function, we derive the transitory component of the SDF as

$$
m_{t,j+1}^T = \exp(D_0 + D_1 + D_5 (x_t - x_{t+1})). \tag{A15}
$$

Equation (A15) implies the permanent component in equation (21), where

$$
D_0 = \log(\beta) + (\rho - 1) \log(g) + (\rho - 1) \frac{(s-1)}{s} \left( \frac{\eta_0}{1 - \varphi_s} - 1 \right) \log(g), \tag{A16}
$$

$$
D_1 = \frac{1}{2} (\rho - 1) (s \gamma_1)^{-1} + \delta_0 \gamma_1^2 \nu, \tag{A17}
$$

$$
D_2 = (\rho - 1) \frac{1}{s} (1-s) \left( \frac{1}{\gamma_1} \varphi_s - 1 \right) \eta [B], \tag{A18}
$$

$$
D_3 = - \delta [B] - (\rho - 1) \frac{1}{s} (1-s) \frac{1}{\gamma_1} \eta [B] + (\rho - 1) \frac{1}{s} \left( 1 - \frac{1}{\gamma_1} \varphi_s \right), \tag{A19}
$$

$$
D_4 = (\rho - 1) (s \gamma_1)^{-1} + \delta [B], \tag{A20}
$$

$$
D_5 = \delta [B]. \tag{A21}
$$
II. Permanent component of the SDF of the recursive utility models

We derive the expressions of the permanent component of the SDFs, as presented in equations (26) and (27).

Based on equations (22) and (24), 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 \( e^{\chi_0} h_t^j / j! \) expands to

\[
m_{t,t+1} = \exp \left( \chi_0 + a_g \omega_{gt}^2 + a_z z_{gt} + a_\nu \omega_{ht} \right),
\]

where \( a_g, a_z, a_\nu \) are backshift operators defined as follows:

\[
a_g = (\rho - 1) \gamma + (\alpha - \rho) \gamma_1, \quad a_z = (\rho - 1) \psi + (\alpha - \rho) \psi_1, \quad a_\nu = (\alpha - \rho) D (b_1 v_1 - v_B),
\]

\[
D = (\alpha/2) (\gamma_1^2), \quad \text{and} \quad J = \left( \frac{e^{\alpha \psi_1 \theta + (\alpha \psi_1 \delta)^2} - 1}{\alpha} \right).
\]

The functions \( \eta, \nu, \gamma \) are polynomial functions of \( b_1 \):

\[
\eta = \sum_{j=0}^{\infty} b_j \eta_j, \quad \gamma = \sum_{j=0}^{\infty} b_j \gamma_j, \quad \nu = \sum_{j=0}^{\infty} b_j \nu_j, \quad \psi = \sum_{j=0}^{\infty} b_j \psi_j.
\]

This ends the proof. ■
with $\gamma_0 = 1$, where

$$
\sum_{j=1}^{\infty} \gamma_j < \infty, \quad \sum_{j=1}^{\infty} \eta_j < \infty, \quad \sum_{j=1}^{\infty} \nu_j < \infty, \quad \sum_{j=1}^{\infty} \psi_j < \infty, \quad (B7)
$$

and

$$
\nu[B] = \sum_{j=0}^{\infty} \nu_j B^j \quad \text{and} \quad \psi[B] = \sum_{j=0}^{\infty} \psi_j B^j. \quad (B8)
$$

A. Recursive utility with stochastic variance: The SDF is a special case of (B1) with $h = 0, \eta[B] = 0, J = 0$. The SDF takes the form:

$$
m_{t,t+1} = \exp \left( H_0 + (\rho - 1) \sum_{j=1}^{\infty} \gamma_j \nu_j^2 \omega_{g,t+1} + (\alpha - \rho) \sum_{j=1}^{\infty} \gamma_j \nu_j^2 \omega_{g,t+1}^2 \right),
$$

with

$$
H_0 = \log(\beta) + (\rho - 1) \log g - (\alpha - \rho) (D\nu) - (\alpha - \rho) (\alpha/2) \left( (Db_1 \nu[B])^2 \right). \quad (B9)
$$

Now, define

$$
x_t = (\gamma[B] - \gamma_0) \nu_t^2 \omega_{g,t+1}. \quad (B10)
$$

The state variable $x_t$ dynamics is:

$$
x_t = \varphi_g x_{t-1} + \gamma_1 \nu_{t-1}^2 \omega_{g,t}, \quad \text{with} \quad \gamma_j = \varphi_g \gamma_{j-1} \quad \text{for} \quad j \geq 2 \quad \text{and} \quad \varphi_g = \frac{\gamma_2}{\gamma_1}. \quad (B11)
$$

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

$$
\nu_t - \nu = \varphi_0 (\nu_{t-1} - \nu) + \nu_0 \omega_{u,t}, \quad \text{for} \quad j \geq 2 \quad \text{and} \quad \varphi_0 = \frac{\nu_1}{\nu_0}. \quad (B12)
$$
The SDF can be expressed as

\[ m_{t,t+1} = \exp(H_1 + H_2 x_t + H_3 x_{t+1} + H_4 \nu_t + H_5 \nu_{t+1}), \]  

(B13)

where

\[ H_1 = H_0 + (\alpha - \rho) D \nu + (\alpha - \rho) D b_1 \nu [b_1] \frac{(\phi_0 - 1)}{\nu_0} \nu, \]  

(B14)

\[ H_2 = (\rho - 1) - ((\alpha - \rho) \gamma [b_1] + (\rho - 1)) \frac{\phi_0}{\gamma_1}, \]  

(B15)

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

(B16)

\[ H_4 = (\alpha - \rho) D \left(-b_1 \nu [b_1] \frac{\phi_0}{\nu_0} - 1\right), \]  

(B17)

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

(B18)

Proceeding, we now solve the eigenfunction problem specified in equations (A8) and (A9). We conjecture that \( \log(e_{t+1}) = \tau_0 x_{t+1} + \tau_1 \nu_{t+1} \). Hence,

\[ \log(m_{t,t+1} e_{t+1}/e_t) = H_1 + (H_2 - \tau_0) x_t + (H_3 + \tau_0) x_{t+1} + (H_4 - \tau_1) \nu_t + (H_5 + \tau_1) \nu_{t+1} \]  

(B19)

and

\[ \log(\zeta) = H_1 + (H_5 + \tau_1) \nu (1 - \phi_0) + \frac{1}{2} (H_5 + \tau_1)^2 \nu_0^2 + (H_2 - \tau_0 + (H_3 + \tau_0) \phi_0) x_t \]

\[ + \left( (H_4 - \tau_1) + \frac{1}{2} (H_5 + \tau_0)^2 \gamma_1^2 + (H_5 + \tau_1) \phi_0 \right) \nu_t. \]  

(B20)

Using the identification approach, we arrive at the expressions:

\[ \log(\zeta) = H_1 + (H_5 + \tau_1) \nu (1 - \phi_0) + \frac{1}{2} (H_5 + \tau_1)^2 \nu_0^2 \]  

(B21)
\( \tau_0 = \frac{H_2 + H_3 \varphi_g}{1 - \varphi_g} \) and \( \tau_1 = \frac{H_4 + \frac{1}{2} (H_3 + \varpi_0)^2 \chi_1 + H_5 \varphi_v}{1 - \varphi_v}. \) \hfill (B22)

With these results, we are in a position to state the transitory and permanent components as:

\[
\begin{align*}
m^T_{t, t+1} &= \exp \left( H_1 + (H_5 + \tau_1) \varpi (1 - \varphi_v) + \frac{1}{2} (H_3 + \tau_1)^2 \varpi_0^2 + \tau_0 (x_t - x_{t+1}) + \tau_1 (\varpi_t - \varpi_{t+1}) \right), \\
m^P_{t, t+1} &= \exp \left( (H_2 - \tau_0) x_t + (H_3 + \tau_0) x_{t+1} + (H_4 - \tau_1) \varpi_t + (H_5 + \tau_1) \varpi_{t+1} \right). \quad \hfill (B23)
\end{align*}
\]

Setting \( H_6 \equiv -(H_5 + \tau_1) \varpi (1 - \varphi_v) - (H_5 + \tau_1)^2 \varpi_0^2 / 2, \) we have equation (26).

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

\[
m_{t, t+1} = \exp \left( \begin{array}{c}
\chi_0 \\
+ (\rho - 1) x_t + ((\rho - 1) \gamma_0 + (\alpha - \rho) \gamma [b_1]) \omega_{g(t+1)} \\
+ (\rho - 1) (\psi [B] - \psi_0) \omega_{g(t+1)} + ((\rho - 1) \psi_0 + (\alpha - \rho) \psi [b_1]) \omega_{g(t+1)} \\
+ (\alpha - \rho) J b_1 \eta [b_1] \omega_{h(t+1)} - (\alpha - \rho) (h_t - h) J
\end{array} \right). \quad \hfill (B24)
\]

Now denote

\( \tilde{x}_t = (\psi [B] - \psi_0) \omega_{g(t+1)}. \) \hfill (B25)

The law of motion of \( \tilde{x}_t \) becomes

\[
\tilde{x}_t = \phi_z \tilde{x}_{t-1} + \psi_1 \omega_{g(t+1)} \quad \text{with} \quad \phi_z = \frac{\psi_2}{\psi_1} \quad \text{and} \quad \psi_{j+2} = \phi_z \psi_{j+1} \quad \text{for} \ j \geq 1. \quad \hfill (B26)
\]

The SDF in equation (B24) reduces to

\[
m_{t, t+1} = \exp \left( G_0 + G_1 x_t + G_2 \tilde{x}_{t-1} + G_3 \omega_{g(t+1)} + G_4 h_t + G_5 \omega_{g(t+1)} + G_6 \omega_{g(t+1)}^2 + G_7 \omega_{h(t+1)} \right), \quad \hfill (B27)
\]
with

\[ G_0 = \chi_0 + (\alpha - \rho) hJ, \quad G_1 = (\rho - 1), \]
\[ G_2 = (\rho - 1) \varphi_2, \quad G_3 = (\rho - 1) \psi_1, \]
\[ G_4 = -(\alpha - \rho) J, \quad G_5 = (\rho - 1) \psi_0 + (\alpha - \rho) \psi[b_1], \]
\[ G_6 = (\rho - 1) \gamma_0 + (\alpha - \rho) \gamma[b_1], \quad G_7 = (\alpha - \rho) Jb_1 \eta[b_1]. \]

For the eigenfunction problem in equations (A8)–(A9), i.e., \( E_t[m_{t+1}e_{t+1}] = \zeta e_t \), we conjecture that the eigenfunction is of the form:

\[ e_{t+1} = \exp \left( \zeta_0 h_{t+1} + \zeta_1 \zeta_{t+1} + \zeta_2 x_{t+1} + \zeta_3 \eta_t \right). \]  \( \text{(B28)} \)

Algebraic manipulation yields the expression:

\[
m_{t+1} \frac{e_{t+1}}{e_t} = \exp \left( G_0 + \zeta_0 h - \zeta_0 \varphi h + (G_1 - \zeta_2 + \zeta_2 \varphi_8) x_t + G_2 \bar{x}_{t-1} + (G_3 - \zeta_1 + \zeta_3 \psi_1) \zeta_{t+1} \right) \]
\[ + (G_4 - \zeta_0 + \zeta_0 \varphi h) h_t + (\zeta_3 \psi_7 - \zeta_3) \bar{x}_{t-1} \]
\[ \times \exp \left( (G_5 + \zeta_1) \zeta_{t+1} + (G_6 + \zeta_2 \gamma_1) \nu^2 \omega_{t+1} + (G_7 + \zeta_0 \eta_0) \omega_{t+1} \right). \]  \( \text{(B29)} \)

Upon further manipulation of equation (B29), we get

\[
\zeta = \xi \times \exp \left( G_0 + \zeta_0 h - \zeta_0 \varphi h + (G_1 - \zeta_2 + \zeta_2 \varphi_8) x_t + G_2 \bar{x}_{t-1} + (G_3 - \zeta_1 + \zeta_3 \psi_1) \zeta_{t+1} \right) \]
\[ + (G_4 - \zeta_0 + \zeta_0 \varphi h) h_t + (\zeta_3 \psi_7 - \zeta_3) \bar{x}_{t-1} \]  \( \text{(B30)} \)

with

\[
\xi = E_t \left( \exp \left( (G_5 + \zeta_1) \zeta_{t+1} + (G_6 + \zeta_2 \gamma_1) \nu^2 \omega_{t+1} + (G_7 + \zeta_0 \eta_0) \omega_{t+1} \right) \right). \]  \( \text{(B31)} \)

One may observe that

\[
\zeta = (E_t \exp ((G_5 + \zeta_1) \zeta_{t+1})) \left( E_t \exp ((G_6 + \zeta_2 \gamma_1) \nu^2 \omega_{t+1}) \right) \left( E_t ((G_7 + \zeta_0 \eta_0) \omega_{t+1}) \right) \]  \( \text{(B32)} \)
\[ = E_t \left( \exp \left( (G_5 + \zeta_1) \theta + \frac{1}{2} (G_5 + \zeta_1)^2 \delta^2 \right) \right) \exp \left( \frac{1}{2} (G_6 + \zeta_2 \gamma_1)^2 \nu + \frac{1}{2} (G_7 + \zeta_0 \eta_0)^2 \right) \]
and

\[
E_t \left( \exp \left( (G_5 + \varsigma_1) \theta + \frac{1}{2} (G_5 + \varsigma_1)^2 \delta^2 \right)^j \right) = \exp (G_8 h_t), \quad (B33)
\]

with

\[
G_8 = e^{((G_5 + \varsigma_1) \theta + \frac{1}{2} (G_5 + \varsigma_1)^2 \delta^2)} - 1. \quad (B34)
\]

As a consequence, equation (B32) simplifies to

\[
\xi = \exp \left( G_8 h_t + \frac{1}{2} (G_6 + \varsigma_2 \gamma_1)^2 v + \frac{1}{2} (G_7 + \varsigma_0 \eta_0)^2 \right). \quad (B35)
\]

We substitute equation (B35) in equation (B30) and rearrange to obtain:

\[
\log(\varsigma) = G_0 + \varsigma_0 h - \varsigma_0 \Phi h + (G_1 - \varsigma_2 + \varsigma_2 \Phi_g) x_t
\]
\[
+ (G_3 - \varsigma_1 + \varsigma_3 \Psi_1) z_{t\ell} + (G_4 - \varsigma_0 + \varsigma_0 \Phi_h + G_8) h_t
\]
\[
+ (\varsigma_3 \Phi_z - \varsigma_3 + G_2) \bar{x}_{t-1} + \frac{1}{2} (G_6 + \varsigma_2 \gamma_1)^2 v + \frac{1}{2} (G_7 + \varsigma_0 \eta_0)^2. \quad (B36)
\]

Using the identification approach, we then have

\[
\log(\varsigma) = G_0 + \varsigma_0 h (1 - \Phi_h) + \frac{1}{2} (G_6 + \varsigma_2 \gamma_1)^2 v + \frac{1}{2} (G_7 + \varsigma_0 \eta_0)^2 \quad (B37)
\]

and

\[
G_1 - \varsigma_2 + \varsigma_2 \Phi_g = 0, \quad G_4 - \varsigma_0 + \varsigma_0 \Phi_h + G_8 = 0, \quad (B38)
\]
\[
G_3 - \varsigma_1 + \varsigma_3 \Psi_1 = 0, \quad \varsigma_3 \Phi_z - \varsigma_3 + G_2 = 0.
\]

Finally, we get

\[
\varsigma_0 = \frac{G_8 + G_9}{1 - \Phi_h}, \quad \varsigma_1 = G_3 + \varsigma_3 \Psi_1, \quad \varsigma_2 = \frac{G_1}{1 - \Phi_g}, \quad \varsigma_3 = \frac{G_2}{1 - \Phi_z}. \quad (B39)
\]
The transitory component is, therefore, 
\[ m_{t,t+1}^T = \zeta \exp (e_t - e_{t+1}), \]
and we obtain:
\[ m_{t,t+1}^T = \zeta \exp (\xi_0 (h_t - h_{t+1}) + \xi_1 (z_{gt} - z_{gt+1}) + \xi_2 (x_t - x_{t+1}) + \xi_3 (x_t - 1)), \]  
(B40)

We can establish the relation in equation (27) by setting $G_9 \equiv -\frac{1}{2} (G_6 + \xi_2 \gamma_1)^2 v - \frac{1}{2} (G_7 + \xi_0 \eta_0)^2$.

III. Lower bound on $L[m_{t,t+1}]$ and $L[m_{t,t+1}^P]$ in equations (28) and (29)

We present a lower bound on the entropy of $m_{t,t+1}$ and a lower bound on the entropy of $m_{t,t+1}^P$ when $m_{t,t+1}$ correctly price $N + 2$ assets. Then we illustrate the sharpness of these bounds.

We maintain the following notation:

\[ y \equiv \Sigma^{-1} (1 - E [q_i] E [R_{t,t+1}]), \quad \text{and} \quad a \equiv \frac{y}{1'y}, \]  
(C1)

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

A. Proof of the bound on the entropy of $m_{t,t+1}$

Consider the following return
\[
E \left[ \log \left( m_{t,t+1} a' R_{t,t+1} \right) \right] \leq \log \left( E \left[ m_{t,t+1} a' R_{t,t+1} \right] \right),
\]  
(C2)
\[
\leq \log \left( a' E [m_{t,t+1} R_{t,t+1}] \right),
\]
\[
\leq \log \left( a' 1 \right) = \log (1),
\]
\[
\leq 0.
\]

From equation (C2) and noting that $\log (m_{t,t+1} a' R_{t,t+1}) = \log (m_{t,t+1}) + \log (a' R_{t,t+1})$, we deduce that
\[
E \left[ \log \left( a' R_{t,t+1} \right) \right] \leq -E \left[ \log (m_{t,t+1}) \right],
\]  
(C3)
Adding \( \log(E[m_{t,t+1}]) \) to both sides of equation (C3) yields

\[
\log(E[m_{t,t+1}]) + E[\log(a'R_{t,t+1})] \leq \log(E[m_{t,t+1}]) - E[\log(m_{t,t+1})] = L[m_{t,t+1}]. \tag{C4}
\]

Since \( q_t = E_t[m_{t,t+1}] \), equation (C4) simplifies to

\[
L[m_{t,t+1}] \geq E[\log(a'R_{t,t+1})] - \log(1/E[q_t]), \text{ where } a' \text{ is as defined in equation (C1)}. \tag{C5}
\]

We have a result that depends on the variance-covariance matrix of returns and the mean return vector. ■

**B. Proof of the bound on the entropy of \( m_{t,t+1}^P \)**

We note that

\[
E\left[ \log\left( m_{t,t+1}^P a'R_{t,t+1} \right) \right] = E\left[ \log\left( \frac{m_{t,t+1}^P a'R_{t,t+1}}{R_{t,t+1,\infty}} \right) \right]. \tag{C6}
\]

Invoking Jensen’s inequality, we have

\[
E\left[ \log\left( \frac{m_{t,t+1}^P a'R_{t,t+1}}{R_{t,t+1,\infty}} \right) \right] = E\left[ \log\left( m_{t,t+1} a'R_{t,t+1} \right) \right], \tag{C7}
\]

\[
\leq \log\left( a'E[m_{t,t+1} R_{t,t+1}] \right),
\]

\[
\leq \log\left( a' \mathbf{1} \right) = 0.
\]

From equation (C7), we deduce

\[
E\left[ \log\left( \frac{a'R_{t,t+1}}{R_{t,t+1,\infty}} \right) \right] \leq -E[\log(m_{t,t+1}^P)] = L[m_{t,t+1}^P]. \tag{C8}
\]

Hence,

\[
L[m_{t,t+1}^P] \geq E[\log(a'R_{t,t+1})] - E[\log(R_{t,t+1,\infty})]. \tag{C9}
\]

This bound is the generalization to \( N + 2 \) assets. ■
C. Sharpness of our entropy bounds

How sharp is our bound on $L[m_{t,t+1}]$ compared to the bound constructed from a generic portfolio return in Backus, Chernov, and Zin (2014, Column 2 of Table I) and the corresponding bound on $L[m_{t,t+1}^P]$ in Alvarez and Jermann (2005) (based on pricing the risk-free bond return, the long-term bond return, and a generic portfolio return)? To address this question, Table Internet Appendix-I reports our lower bounds on $L[m_{t,t+1}]$ and $L[m_{t,t+1}^P]$ and the associated bootstrap $p$-values.

We consider several $N$ (the dimensionality of $R_{t,t+1}$) and draw two conclusions from our computations in Table Internet Appendix-I:

First, our bounds on $L[m_{t,t+1}]$ and $L[m_{t,t+1}^P]$ are quantitatively sharper with $N > 1$, 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 comprising the portfolio. This can be seen by comparing the bound displayed in row (c) versus (i) and between row (d) versus (j).

IV. Proof of equation (30)

Let $\mathbb{L}_B^{m_2}$ represent the lower bound on $L[m^2]$ and $\mathbb{L}_B^{m}$ the lower bound on $L[m]$, as depicted in the right-hand sides of equation (15) and equation (28), respectively. We wish to show that

$$\frac{L[m^2]}{4L[m]} - 1 = 0 \quad \text{implies the data-based restriction that} \quad \frac{\mathbb{L}_B^{m_2}}{4\mathbb{L}_B^{m}} - 1 = 0. \quad (C10)$$

Since $\mathbb{L}_B^{m_2}$ is the lower bound on $L[m^2]$, we have

$$\mathbb{L}_B^{m_2} \leq L[m^2] = 4L[m]. \quad \text{Hence,} \quad \frac{1}{4} \mathbb{L}_B^{m_2} \leq L[m]. \quad (C11)$$
Since $\text{LB}_m$ is the highest bound among all lower bounds on $L[m]$: 

$$\frac{1}{4} \text{LB}_m \leq \text{LB}_m.$$  \hspace{1cm} (C12)

Correspondingly, since $\text{LB}_m$ is the lower bound on $L[m]$, we must have

$$\text{LB}_m \leq L[m] = \frac{1}{4} L[m^2].$$  \hspace{1cm} (C13)

That is, $4\text{LB}_m \leq L[m^2]$. 

Since $\text{LB}_{m^2}$ is the highest among all lower bounds on $L[m^2]$, we have

$$4\text{LB}_m \leq \text{LB}_{m^2}.$$  \hspace{1cm} (C14)

Combining (C12) and (C14) leads to $\frac{\text{LB}_{m^2}}{4\text{LB}_m} - 1 = 0$. 

$\blacksquare$
Table Internet Appendix-I
Sharpness of our entropy bounds on $m_{t,t+1}$ and $m^P_{t,t+1}$, when SDFs correctly price each of the $N+2$ assets

Reported are the lower entropy bounds with the one-sided $p$-values in $\langle \rangle$. Our lower entropy bounds on $m_{t,t+1}$ and $m^P_{t,t+1}$ are based on equations (28) and (29), respectively, and rely on the ability of the SDF to correctly price each of the $N+2$ assets (the risk-free bond, the long-term discount 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: $E[\log(R^m_{t,t+1})]$, while the Alvarez and Jermann (2005, equation (4)) lower bound on the entropy of $m^P_{t,t+1}$ (denoted by AJ) is based on the expression: $E[\log(R^m_{t,t+1})] - E[\log(R_{t,t+1,\infty})]$, where $R^m_{t,t+1}$ is the return on a single risky asset or a benchmark portfolio (i.e., which we proxy, for instance, by the value-weighted equity market return or equally weighted portfolio of 25 Fama-French size and book-to-market portfolios). Moreover, $R_{t,t+1,\infty}$ is the return on an infinite-maturity bond, which we proxy by the return of a 30-year Treasury bond. $R^f_t$ 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.

<table>
<thead>
<tr>
<th>Panel</th>
<th>SDF correctly prices each of the $N+2$ assets, and we set $N$</th>
<th>Lower bound on $m_{t,t+1}$</th>
<th>$p$-value</th>
<th>Lower bound on $m^P_{t,t+1}$</th>
<th>$p$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel I</td>
<td>Market, 25 size &amp; B/M</td>
<td>0.023</td>
<td>(0.000)</td>
<td>0.021</td>
<td>(0.000)</td>
</tr>
<tr>
<td></td>
<td>(b) Market, 25 size &amp; momentum</td>
<td>0.037</td>
<td>(0.003)</td>
<td>0.035</td>
<td>(0.003)</td>
</tr>
<tr>
<td>Panel II</td>
<td>25 size &amp; B/M</td>
<td>0.022</td>
<td>(0.000)</td>
<td>0.020</td>
<td>(0.000)</td>
</tr>
<tr>
<td></td>
<td>(d) 25 size &amp; momentum</td>
<td>0.029</td>
<td>(0.000)</td>
<td>0.027</td>
<td>(0.000)</td>
</tr>
<tr>
<td>Panel III</td>
<td>Market, 10 momentum</td>
<td>0.020</td>
<td>(0.000)</td>
<td>0.018</td>
<td>(0.001)</td>
</tr>
<tr>
<td>Panel IV</td>
<td>Market, Low Momentum</td>
<td>0.010</td>
<td>(0.000)</td>
<td>0.008</td>
<td>(0.000)</td>
</tr>
<tr>
<td></td>
<td>(g) Market, high Momentum</td>
<td>0.014</td>
<td>(0.010)</td>
<td>0.012</td>
<td>0.011</td>
</tr>
<tr>
<td>Panel V</td>
<td>Market portfolio only (BCZ, Eq. 5)</td>
<td>0.005</td>
<td>(0.005)</td>
<td>0.003</td>
<td>(0.066)</td>
</tr>
<tr>
<td></td>
<td>(i) EWI portfolio of 25 size &amp; B/M</td>
<td>0.007</td>
<td>(0.001)</td>
<td>0.005</td>
<td>(0.018)</td>
</tr>
<tr>
<td></td>
<td>(j) EWI portfolio of 25 size &amp; momentum</td>
<td>0.007</td>
<td>(0.001)</td>
<td>0.005</td>
<td>(0.021)</td>
</tr>
</tbody>
</table>

14