1. Preview of Hansen Jagannathan Bound

Cochrane J H, Hansen L P. Asset pricing explorations for macroeconomics[J]. **NBER macroeconomics annual**, 1992, 7: 115-165.

1.1 Why Hansen Jagannathan Bound?

Asset market data are often ignored in evaluating macroeconomic models, and aggregate quantity data are often avoided in empirical investigations of asset market returns. While there may be short-term benefits to proceeding along separate lines, we argue that security market data are among the most sensitive and, hence, attractive proving grounds for models of the aggregate economy.

Not only can security data be informative for macroeconomic modeling, but macroeconomic modeling should be also valuable in interpreting the cross-sectional and time-series behavior of asset returns.

A large body of empirical work on asset pricing aims **simply at reducing** asset valuation to the pricing of a relatively small number of "factors," **without explicit reference to the fundamental sources of risk**. While these dimensionality-reduction exercises can be quite useful in some contexts, it is difficult, if not impossible, to evaluate the significance of apparent asset-pricing anomalies without specifying an underlying valuation model that ties asset prices to fundamental features of the underlying economic environment, that is, without using some dynamic economic model.

1.2 How to get?

Simplest version.

We study a regression of a discount factor m onto a constant and the vector x of asset payoffs observed by an econometrician:

$$m = a + x'b + e$$

where a is a constant term, b is a vector of slope coefficients, and e is the regression error. The standard least-squares formula for the regression coefficients gives:

$$b \equiv [cov(x, x)]^{-1} cov(x, m)$$
$$a \equiv Em - Ex'b$$

Without **direct data on the stochastic discount factor** m, these regression coefficients cannot be estimated in the usual fashion. Instead, we can exploit the fact that y must be a valid discount factor to infer them. The pricing relation q = E(mx) implies:

$$cov(x, m) = q - E(m)E(x)$$

After substitution, we obtain:

$$b = [cov(x, x)]^{-1}[q - E(m)E(x)]$$

Hence, asset information alone can be used to construct the regression coefficients b, given $E(m)$.

Because the right-hand side variables of a regression are **uncorrelated with residuals** by construction:

$$var(m) = var(x'b) + var(e)$$

It follows that $var(x'b)^{1/2}$ gives a lower bound on the standard deviation of m. Thus, **we have a lower bound on the standard deviation of all admissible stochastic discount factors** y in J with the pre-specified mean, Em.

If there is no real risk-free interest rate. As a consequence, the price of a unit payoff is not known, and Em cannot be inferred from the asset market data. Instead, **we must calculate the lower bound on the standard deviation of m for each possible value of the mean**. This computation leads to the **lower envelope** of the set of means and standard deviations of admissible discount factors.

1.3 Flaws

- The restriction on SDF derived in Hansen and Jagannathan(1991) all pertain to *first and second moments*. (**Author**)
- Suppose some clever economist does discover nonlinear transformation of the aggregate consumption series that plots in the Hansen-Jagannathan feasible region fro discount factors. Does this mean we should adopt the corresponding model of utility?(**Campbell 1**)
- This paper takes the second moment of realized returns as given, rather than as something to be explained. But to a macroeconomist, it is more natural to think of realized returns as endogenous, determined along with other variables.(**Campbell 2**)
- Although there are no *false positives* in the Hansen-Jagannathan diagnostic tests, there are *false negatives* which is very prominent in the habit formation model.(**Philippe Weil 1**)
- Because its very **theoretical agnosticism**, the Hansen-Jagannathan framework in itself *offers very few insights* as to why specific models pass or fail the Hansen-Jagannathan tests. (**Philippe Weil 2**)

1.3.1 Conditional or unconditional?

> *Incomplete market models and subsistence level models may perform adequately in terms of the **unconditional bounds** on the stochastic discount factor. However, as emphasized in the paper, **unconditional bounds represent only one of many possible moment restrictions**. With*

2. Introduction of entropy

2.1 From relative entropy and probability

The distinction between true and risk-adjusted(risk neutral) probabilities is central to asset pricing. The relative entropy of the risk-adjusted distribution is then

$$
\begin{aligned}
L_t(p^\star_{t,t+n}/p_{t,t+n}) &= \sum_{n=1}^{n} p_{t,t+n} log(p_{t,t+n}/p^\star_{t,t+n}) \\
&= -E_t(log(p^\star_{t,t+n}/p_{t,t+n}))
\end{aligned}
$$

$$
with \quad E_t(p^\star_{t,t+n}/p_{t,t+n}) = 1
$$
$$
\implies L_t(p^\star_{t,t+n}/p_{t,t+n}) = log(E_t(p^\star_{t,t+n}/p_{t,t+n})) - E_t(log(p^\star_{t,t+n}/p_{t,t+n}))
$$

$$
with \quad \frac{p^\star_{t,t+n}}{p_{t,t+n}} = \frac{m_{t+n}}{E_t^p(m_{t+n})}
$$
$$
\implies L_t(m_{t,t+n}) = log E_t m_{t,t+n} - E_t log m_{t,t+n}
$$

Take unconditional expectation

$$
\implies E(L_t(m_{t,t+n})) = E(log E_t m_{t,t+n}) - E log m_{t,t+n}
$$

Scale this by horizon n, and we get the so-called ***entropy*** in the article:

$$
I(n) = n^{-1} E L_t(m_{t,t+n})
$$

2.2 From a unified form of volatility

2.2.1 Generalization of variance

Jensen's gap, applied to ***convex function***:

$$L[m] = E[f(m)] - f(E[m])$$
$$f(m) = -log(m) \qquad entropy \ measure$$
$$f(m) = m^2 \qquad variance \ measure$$

> ✏️ $L(m) \geq 0$ **can be used as a *loss function***

2.2.1.1 Advantages

1. Entropy contains ***higher moments information*** which is the whole point.
2. Entropy extends more easily to ***multiple periods*** than standard deviation
3. ***Easy computation***. Many popular asset pricing models are loglinear, or nearly so. Logarithmic measures are conveniently computed for them.

2.3 Exploration for entropy $L(m)$

2.3.1 A unified form

Unified form of Jensen's gap, also, ***convex function***:

$$L[m^2] = E[f(m^2)] - f(E[m^2])$$
$$f(m) = -log(m)$$

Taylor Expansion of $e^{log(m^n)}$ around $E[nlog(m)]$:

$$e^{log(m^n)} = e^{nE[log(m)]}\{1 + n(log(m) - E[log(m)]) + \frac{n^2}{2!}(log(m) - E[log(m)])^2$$
$$+ \frac{n^3}{3!}(log(m) - E[log(m)])^3 + ...\}$$

Take expectation and taking logs on both side:

$$L[m^n] = log(E[m^n]) - E[log(m^n)]$$
$$= log(1 + \sum_{j=2}^{n} \frac{n^j}{j!}\mu_{log(m)}^{[j]})$$
$$where \qquad \mu_{log(m)}^{[j]} = E[(log(m) - E[log(m)])^j]$$

The equation shows that $L[m^n]$ assigns a bigger weight to each central moment of log(m) than $L[m]$

2.3.2 Entropy Spectrum

From time series perspective, we have two ends of the entropy spectrum, the short end $I(1)$ which is the essential ingredient of entropy and the long end $I(\infty)$. The difference between them is **horizon dependence**.

3. Empirical results for source of entropy

3.1 Derivation and Details

3.1.1 Entropy

From above, the definition of entropy

$$I(n) = n^{-1}EL_t(m_{t,t+1})$$

First result, **entropy bound**, connects **one-period** entropy to **one-period** excess return:

$$I(1) = EL_t(m_{t,t+1}) \geq E(\underbrace{logr_{t,t+1} - logr^1_{t,t+1}}_{excess\ log\ return})$$

where $r_{t,t+1}$ is the return on a one-period bound. In words, mean excess log return are bounded above by the entropy of pricing kernel.

3.1.2 Horizon Dependence

Second result, **Horizon Dependence**, use the behavior of entropy over different time horizons to characterize the **dynamics** of the pricing kernel.

$$\begin{aligned} H(n) &= I(n) - I(1) \\ &= n^{-1}EL_t(m_{t,t+n}) - EL_t(m_{t,t+1}) \\ &= -E(y^n_t - y^1_t) \end{aligned}$$

In iid case, $EL_t(m_{t,t+n}) = nEL_t(m_{t,t+1})$, so $H(n) = 0$, This is a generalization of a well-known property of random walks: *the variance is proportional to the time interval*.

In other cases, horizon dependence reflects **departure from the iid case**, and in this sense is a measure of the pricing kernel's dynamics.

Perhaps the most useful feature of horizon dependence is that it is observable and can be connected to bond yields, *which give us some insight of its property*.

y_t^n 反映了在n horizons 内定价核的平均定价能力，因此$H(n)$反映了定价核在n horizons 内的**定价能力变化**。

【接下来请欣赏抽象粗糙的推导过程，主要为了方便理解。】

$$from \qquad H(n) = -E(y_t^n - y_t^1)$$

$$E_{t+1}(m_{t+1,t+2}\underbrace{R_{t+1,t+2}}_{f t+1}) = 1$$

$$E_{t+1}(m_{t+1,t+2}) = \frac{1}{f_{t+1}}$$

$$with \qquad ny_t^n = \sum_{j=1}^{n} f_t^{j-1}$$

$$y_t^n = n^{-1}\left(\frac{1}{E_t(m_1)} + \frac{1}{E_{t+1}(m_2)} + \dots \frac{1}{E_{t+n-2}(m_{n-1})}\right)$$

$$y_t^1 = \frac{1}{E_t(m_1)}$$

> **Alvarez and Jermann(2005):** '*Prices of long-term bonds are particularly informative about the persistence of pricing kernels because they are the markets forecast of the long-term changes in the pricing kernel*.'

> ✎ **Horizon Dependence, as a measure of mean of** $m_{t,t+1}$

> ✎ **Entropy 与 Horizon Dependence用于衡量模型内核是否贴合数据** . *In modern asset pricing theory, a pricing kernel accounts for asset returns. The reverse is also true: asset returns contain information about the pricing kernel that gave rise to them.*

3.1.3 Choose the benchmark

Facts about excess returns (% per month)

Asset	Mean	Standard Deviation	Skewness	Excess Kurtosis
S&P 500	0.40	5.56	−0.40	7.90
Fama-French (small, low)	−0.30	11.40	0.28	9.40
Fama-French (small, high)	0.90	8.94	1.00	12.80
S&P 500 6% OTM δ-puts	−1.84	5.38	2.77	16.64
Pound Sterling	0.35	3.16	−0.50	1.50
5 year bond	0.15	1.90	0.10	4.87

- **Entropy**: log excess return is less than 1%
- **Horizon Dependence**:   term spread is more than 0.1%

These properties of returns are estimates. The exact number is not critical, but it is helpful to have *a clear numerical benchmark*.

3.2 Representative Agent Models

3.2.1 SDF Case

To explain equity premium or other financial facts, the **macro -finance** literature explored a wide range of alternative preferences and market structure. A sampling with a prominent example of each case:

1. **Habits** (Campbell and Cochrane 1999a, 1999b)
2. **Recursive Utility** (Epstein and Zin 1989)
3. Long run risks (Bansal and Yaron 2004; Bansal, kiku, and Yaron 2012)
4. idiosyncratic risk (Constantinides and Duffie 1996)
5. Heterogenous preference (Garleanu and Panageas 2015)
6. **Rare Disasters** (Reitz 1988; Barro 2006)
7. Utility non separable across goods (Piazzesi, Schneider, and Tuzel 2007)
8. Leverage; balance-sheet; "institutional finance"(Brunnermeier 2009, Krishnamurthy and He2013, many others)
9. Ambiguity aversion, min-max preferences (Hansen and Sargent 2001)
10. Behavioral finance; probability mistakes (Shiller 1981,2014)

These approaches look different, but in the end the ideas are quite similar. ***Each of them boils down to a generalization of marginal utility or discount factor***, most of the same form,

$$M_{t+1} = \beta(\frac{C_{t+1}}{C_t})^{-\gamma}Y_{t+1}$$

Take logs and we get:

$$logM_{t+1} = constant - \gamma logg_t + logY_{t+1}$$

So when we model SDF, we can handle g_t ***or*** Y_{t+1} ***, this subsection mostly handles latter, and next subsection handles former.***

For example, habit model:

$$logh_{t+1} = logh + \eta(B)logc_t$$

3.2.2 LogLinear Case

> ✏️ **LogLinear or numerical solution?**

- **Models with Constant Variance**

$$logg_t = logg + \gamma(B)v^{1/2}w_t$$

where $\gamma_0 = 1, \sum_j \gamma_j^2 < \infty$ and innovation w_t are iid with mean zero, variance one.
- **Models with Stochastic Variance**

$$logg_t = logg + \gamma(B)v_{t-1}^{1/2}w_t$$
$$v_t = v + v(B)w_{vt}$$

- **Models with Jumps**
  Disaster model

$$logg_t = logg + \gamma(B)v^{1/2}w_{gt} + \psi(B)z_{gt} - \psi(1)h_{t-1}\theta$$
$$h_t = h + \eta(B)w_{ht}$$

> There is a whole resurgent interest in "***rare disasters***" that I have not covered. Plus, there is a lot of macro/asset pricing literature that is trying to put in various financial frictions at the core of both. Some of this is an attempt to understand the 2008 financial crisis, a worthy subject. ***The big question is whether such frictions are important for every day asset pricing.***

3.2.3 A unified form

Contains everything.

$$logm_{t+1} = logm + \alpha_g(B)(v_t/v)^{1/2}w_{gt+1} + \alpha_z(B)z_{gt+1} + \alpha_v(B)w_{vt+1} + \alpha_h(B)w_{ht+1}$$
$$v_t = v + v(B)w_{vt}$$
$$h_t = h + \eta(B)w_{vt}$$

3.2.4 Philosophical insights

> **Cochrane J H. Macro-finance[J].** *Review of Finance*, **2017, 21(3): 945-985.**

Different microeconomic stories for the *same aggregate* outcomes have different policy implications.How to choose the best among so many models?

One might distinguish models by which data for Y turn out to work best. But **most of the candidates are highly correlated with each other** – most models end up adding a recession state variable, and it is practically a defining feature of recessions that many variables move together – so telling models apart will be hard this way. That fact also means that telling them apart is less important than may seem.

There is some hope in formally testing models – do their moment conditions and cross equation restrictions hold? – and in checking models' additional assumptions – do conditional moments vary as much and in the way that long-run risk or rare disaster models specify? **But though most models are easily rejected, those rejections correspond to economically uninteresting moments**. And by **publication selection bias** if nothing else, models are cleverly constructed that there auxiliary assumptions are not easily falsified; the variation in moments they require is small, hard to measure, or depends on rare events.

The models also differ in their **tractability, elegance, and the number and fragility of extra assumptions** needed to get from theory to central facts. I think it is a **mistake** to embrace too quickly a formalistic scientism that ignores these features. In explaining which models become popular throughout economics, tractability, elegance, and parsimony **matter more than probability values of test statistics**. Economics needs simple tractable models that help to capture the bewildering number of mechanisms people like to talk about.

There is some wisdom in the joke about **the drunk who looks for his keys under the light, not in the dark where he lost them**. Black boxes are not convincing. Elegance matters. **Economic models are more quantitative parables than scientifically precise models, and elegant parables are more convincing**. Models that need an extra assumption for every fact are less convincing than are models that tie several facts together with a small number of assumptions.

[**not quite understand**]Financial economics is always in danger of being simply an interpretive or poetic discipline: Markets went down, sentiment must have fallen. Markets went down, risk aversion must have risen. Markets went down, there must have been selling pressure. Markets went down, the

Gods must be displeased. Models that rejectably tie their central explanations to other data, and cannot "explain" any event are more convincing.

Partially answer my question about empirical results.

3.3 Empirical Results

3.3.1 Vasicek model

> ✏️ **Benchmark model**

The defined pricing kernel:

$$logm_{t,t+1} = logm + \sum_{j=0}^{\infty} a_j w_{t+1-j}$$

$$-f_t^n = logm + k(A_n) + [a(B)/B^{n+1}]_+ w_t$$

for $n \geq 0$ and $A_n = \sum_{j=0}^{n} a_j$; the subscript "+" means ignore negative powers of B.

$k(s)$ is **cumulant generating function**:

$$k(s) = logE(e^{sw_t})$$

$$= \sum_{i=1}^{\infty} \kappa_{j,t} s^j / j!$$

The cumulant $\kappa_{j,t}$ is the j th derivative of $k(s)$ at $s = 0$; $\kappa_{1,t}$ is the mean; $\kappa_{2,t}$ is the variance; the third and fourth cumulants capture skewness and excess kurtosis, respectively.

So then ω_t is normal:

$$k(s) = s^2/2$$

So the mean forward rate and mean yields:

$$with \quad y_t^n = n^{-1} \sum_{j=1}^{n} f_t^{j-1}$$

$$-E(f_t^n) = logm + k(A_n)$$

$$-E(y_t^n) = logm + n^{-1} \sum_{j=1}^{n} k(A_{j-1})$$

With simple manipulation, we can get $I(n)$:

$$I(n) = n^{-1} E log q_t^n - E log m_{t+1}$$

From vasicek model, we can get:

$$E log(m_{t+1}) = log m$$

so:

$$I(n) = n^{-1} \sum_{j=1}^{n} k(A_{j-1})$$
$$I(1) = k(A_0)$$
$$= a_0^2/2$$
$$H(n) = I(n) - I(1)$$
$$= n^{-1} \sum_{j=1}^{n} [k(A_{j-1}) - k(A_0)]$$

In this setting, **_the initial coefficient a_0 governs one-period entropy and the others (a_j for $j \geq 1$) combine with it to govern horizon dependence._**

3.3.1.1 The values for vasicek

- We make $log m_{t+1}$ an ARMA(1,1) process, and we need to estimate three parameters (a_0, a_1, φ), because $a_{j+1} = \varphi a_j$.
- From Chernov and Mueller report, $|a_1| = 0.878 \times 10^{-3}$, and we set $\varphi = 0.85$. And all the parameters we need to estimate is a_0 .
- Finally, we use the benchmark horizon dependence value to get a_0 .
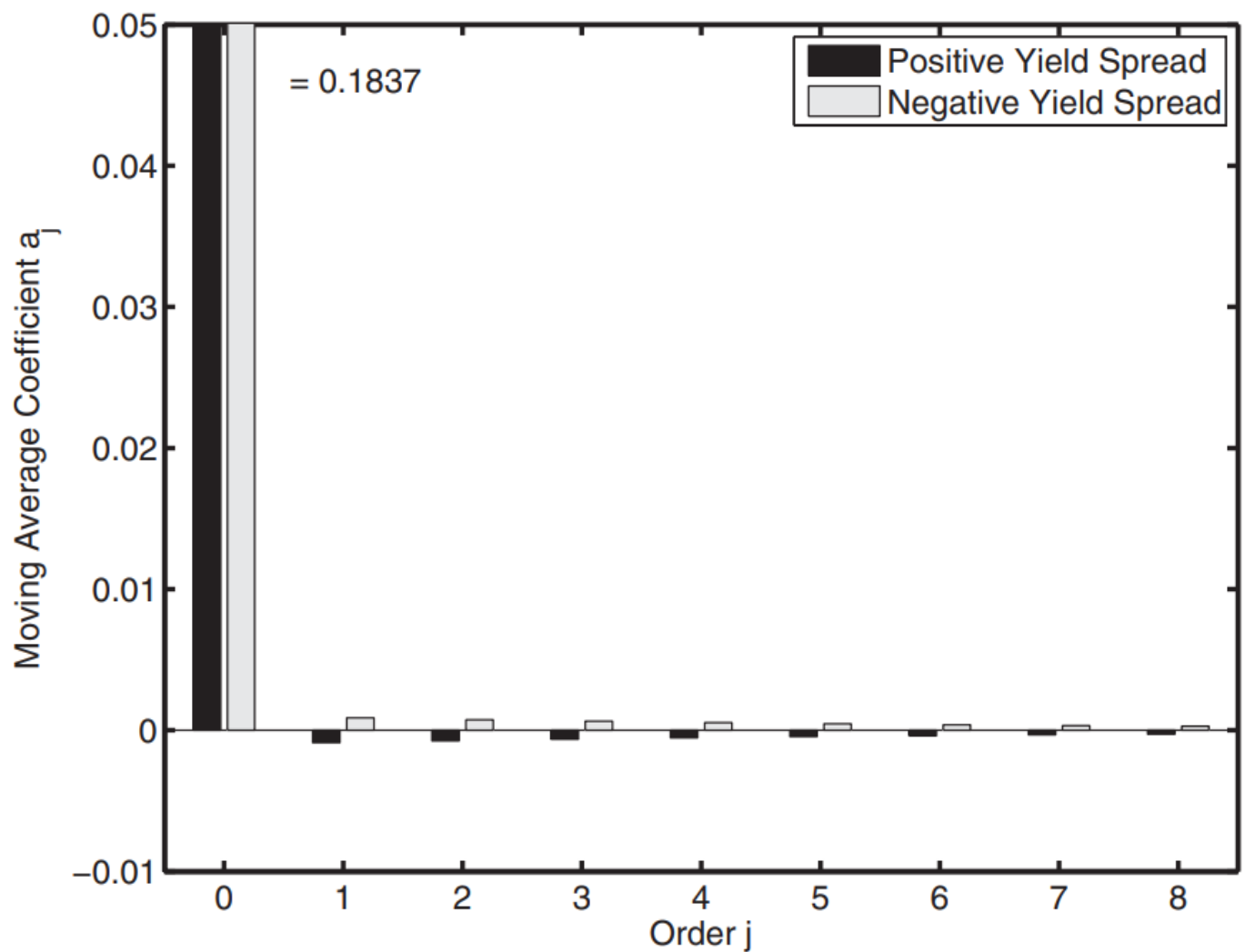
3.3.2 Flaws for models

我们通过给定bound值，反推出vasicek model参数，也就是说vasicek model是能够刚好满足bound的模型，以及对应的bench 参数值。

对于其余模型，根据现实消费等数据计算模型参数，测试其是否满足bound。
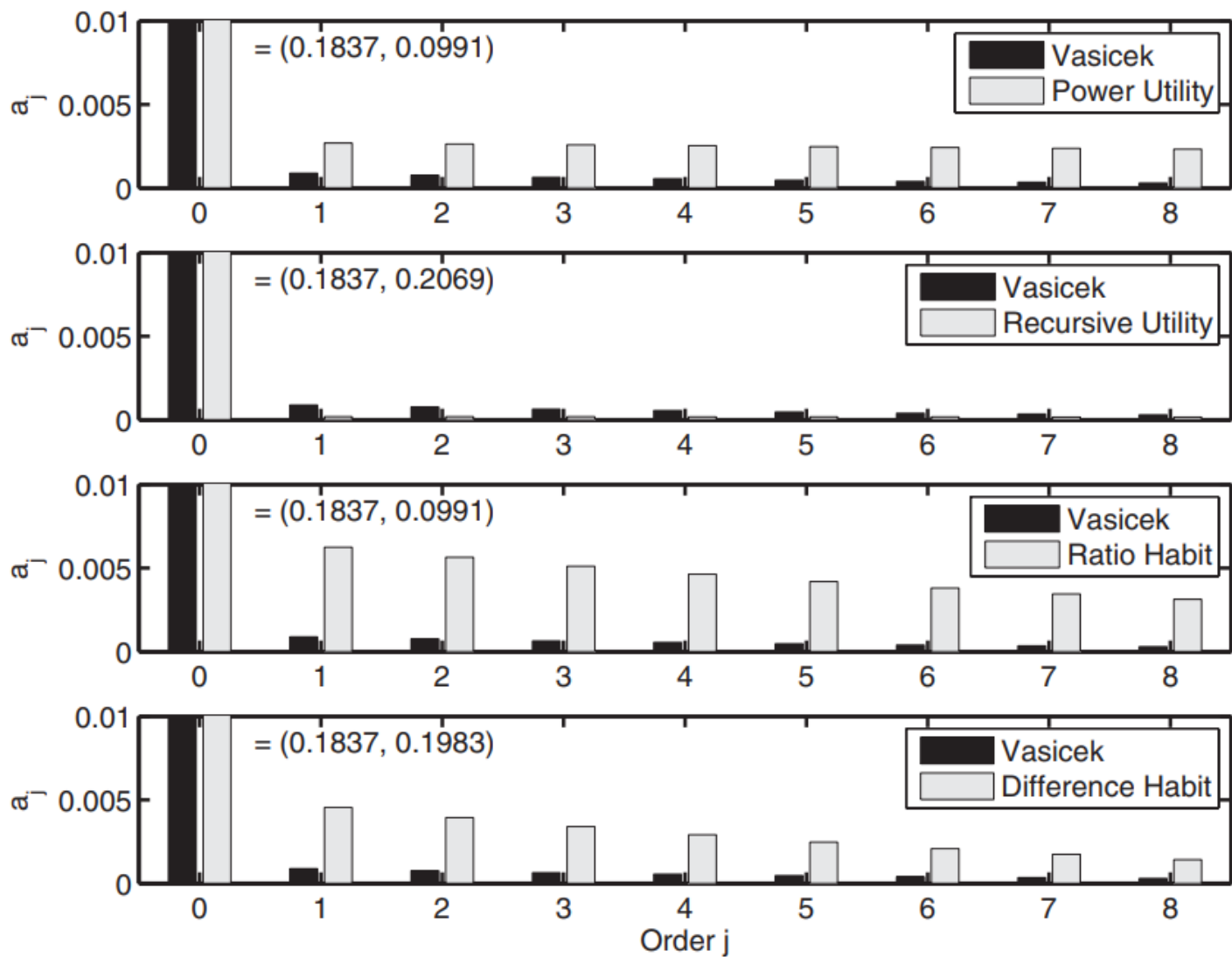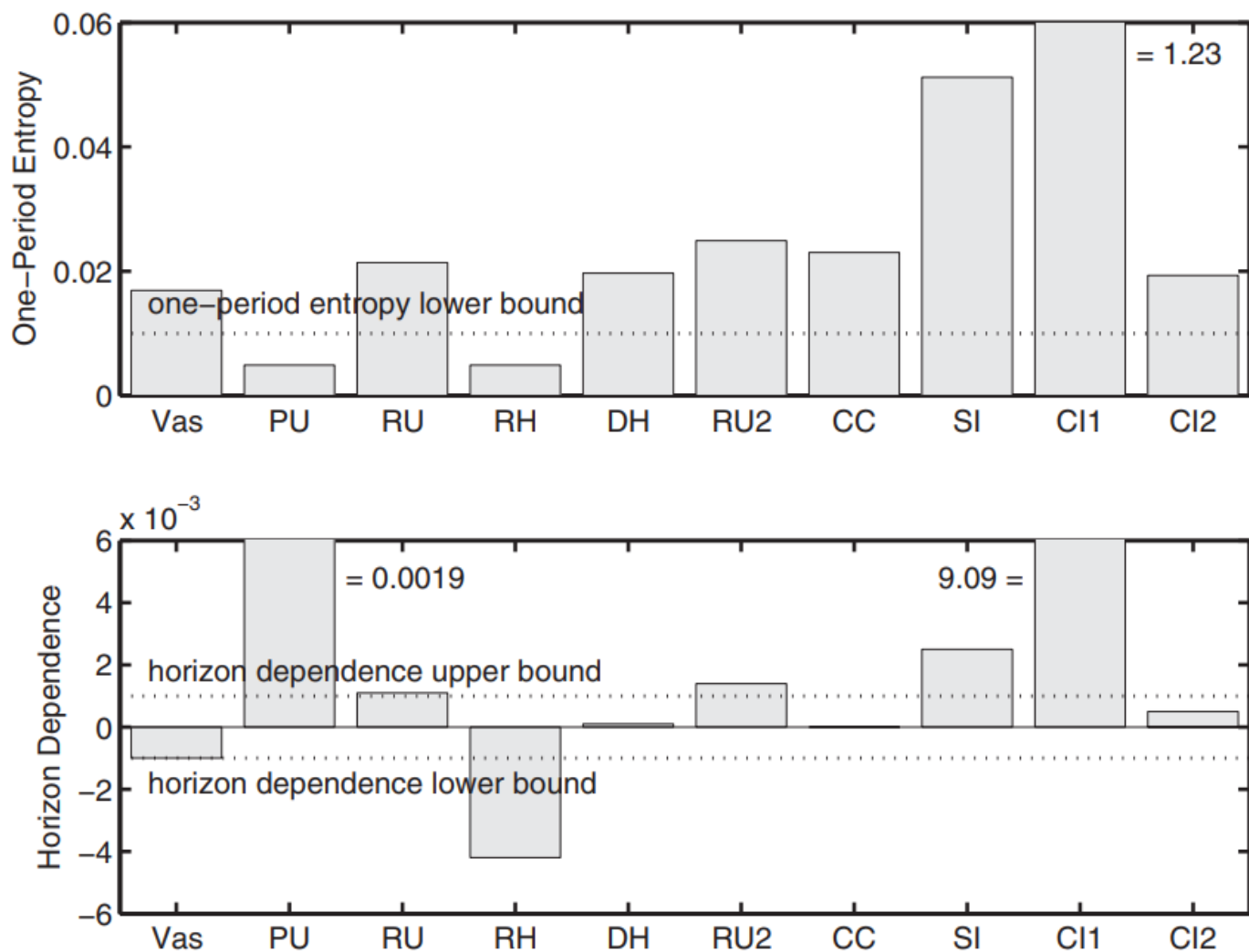
4. Empirical Results for New Restrictions

4.1 Rationale for $L(m^2)$

From above, $L[m^2]$ assigns bigger weights to each central moment of $log(m)$ than $L[m]$:

$$L[m^n] = log(E[m^n]) - E[log(m^n)]$$

$$= log(1 + \sum_{j=2}^{n} \frac{n^j}{j!} \mu_{log(m)}^{[j]})$$

$$where \quad \mu_{log(m)}^{[j]} = E[(log(m) - E[log(m)])^j]$$

When m is **lognormal**, $L[m^2] = 4L[m]$.

The **link** among the entropy of m, the entropy of m^2 and variance of m:

$$L[m^2] = 2L[m] + log(1 + \frac{var[m]}{(E[m])^2})$$

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.

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

4.1.1 Lower entropy bound

Under the SDF set:

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

We have

$$L[m^2] \geq LB_{m^2} \equiv$$
$$2\left(E\left[log\left(\frac{(1 - E[q_t]E[R_{t,t+1}])'\Sigma^{-1}}{1'\Sigma^{-1}(1 - E[q_t]E[R_{t,t+1}])}R_{t,t+1}\right)\right] - log((E[q_t])^{-1})\right)$$
$$+ log(1 + (1 - E[q_t]E[R_{t,t+1}])'\Sigma^{-1}(1 - E[q_t]E[R_{t,t+1}])/(E[q_t])^2)$$

with :

$$a \equiv \underbrace{\frac{\Sigma^{-1}(1 - E[q_t]E[R_{t,t+1}])}{1'\Sigma^{-1}(1 - E[q_t]E[R_{t,t+1}])}}_{\equiv y}$$

and through Taylor Expansion, finally we get:

$$LB_{m^2} = 2log(E[a'R_{t,t+1}]) - log((E[q_t])^{-1})$$
$$+ log(1 + y'\Sigma y/(E[q_t]^2))$$
$$+ 2\Sigma_{i=1}^{\infty} \frac{(-1)^{1+i}}{i} \frac{E[(a'R_{t,t+1} - E[a'R_{t,t+1}])^i]}{(E[a'R_{t,t+1}])^i}$$

> ✏️ ***which is computable from the time-series of asset returns***

For LB_m :

$$L[m] \geq LB_m \equiv E[log(a'R_{t,t+1})] - log((E[q_t])^{-1})$$

4.1.2 Economic Interpretation

Focusing on the 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 the return(**Cochrane**):

$$E[mx] = p$$

$$when \ x = m \quad E[m^2] = p = \frac{m}{1 + r_{SDF}}$$

$$r_{SDF} = \frac{m}{E[m^2]} - 1$$

We can establish that the expected excess log return of a security with SDF payoff is related to $L[m^2]$ as follows:

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

> ✎ *Not practically use*

4.2 Empirical Results

4.2.1 class one

We have Two **hypothesis**:

$$L[m] - LB_m \geq 0 \tag{1}$$

$$L[m^2] - LB_{m^2} \geq 0 \tag{2}$$

One method: **bootstrap**. During each bootstrap of each test, we generate time series of $m_{t,t+1}$, and then calculate the $L[m^2]$ and $L[m]$. And we calculate LB_m and LB_{m^2} via the asset returns.

One statistics: **empirical p-value**. The bootstrap procedure is repeated 100,000 times, and we count **the proportion of bootstrap draws that satisfy the two hypothesis**. We record this proportion as the empirical p-value.

We consider **models of SDFs** that are based on the **excess return of out-of-the-money index puts and calls**, as well those that incorporate **the excess return of return-based equity characteristics**. Each considered model of the SDF accommodates *positivity*.

Table 1

$$m_{t,t+1} = \eta'[R_{t+1,f}Q_{t,t+1}]$$

$$Q_{t,t+1} = \begin{pmatrix} \dfrac{S_{t+1}}{S_t} \\ \dfrac{max(S_t e^{-0.03} - S_{t+1}, 0)}{P_t [S_t e^{-0.03}]} \\ \dfrac{max(S_t e^{-0.01} - S_{t+1}, 0)}{P_t [S_t e^{-0.01}]} \\ \dfrac{max(S_{t+1} - S_t e^{0.01}, 0)}{C_t [S_t e^{0.01}]} \\ \dfrac{max(S_{t+1} - S_t e^{0.03}, 0)}{C_t [S_t e^{0.03}]} \end{pmatrix}$$

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 the exercise presented in Table 1 is 1990:01–2015:12 (311 observations). The model is

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

The 5×1 vector of gross returns $\mathbf{Q}_{t,t+1}$ contains the following:

i) Gross return of the market (Standard & Poor's (S&P) 500 index)

ii) Gross return of a 3% out-of-the-money put on the S&P 500 index

iii) Gross return of a 1% out-of-the-money put on the S&P 500 index

iv) Gross return of a 1% out-of-the-money call on the S&P 500 index

v) 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), pp. 65–66)

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

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 1990:01–2015:12. The reported p-values are based on a bootstrap and test 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 the risk-free return plus a 26×1 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. Dev.	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 26×1			
	$\mathbb{LB}_m$	Bootstrap p-Value	$\mathbb{LB}_{m^2}$	Bootstrap p-Value
25 size × book to market	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

In table 2, **SDF Is Exponentially Linear in Returns based on Fama and French(2015).**

$$m_{t,t+1} = e^{\lambda_0 - 1 + \lambda' er_{t,t+1}}$$

$$er_{t,t+1} \equiv Q_{t,t+1} - 1R_{t+1,f}$$

where $er_{t,t+1}$ is a 9×1 vector of excess returns, where $Q_{t,t+1}$ is based on market, plus the low and high extreme portfolios of Fama and French(2016).

The sample period for the exercise presented in Table 2 is 1963:07–2017:06 (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 9×1 vector of excess returns. The excess returns $\mathbf{er}_{t,t+1}$ are based on $\mathbf{Q}_{t,t+1}$ from the 5-factor model of Fama and French (2015):

   i)   Gross return of the market

   ii) Gross returns of the two extreme low, and high, size portfolios

   iii) Gross returns of the two extreme low, and high, book-to-market portfolios

   iv) Gross returns of the two extreme low, and high, operating-profitability portfolios

   v)   Gross returns of the two extreme low, and high, investment portfolios.

The constants $(\lambda_0, \boldsymbol{\lambda})$ are a solution to the minimization problem $\inf_{(\lambda_0, \boldsymbol{\lambda})} -\mathbb{E}[q_t]\lambda_0 + \mathbb{E}[\exp(\lambda_0 - 1 + \boldsymbol{\lambda}' \mathbf{er}_{t,t+1})]$. This form of the stochastic discount factor (SDF) obtains by minimizing the objective $\mathbb{E}[m \log(m)]$ subject to $\mathbb{E}[m\,\mathbf{er}] = \mathbf{0}$ and $\mathbb{E}[m - \mathbb{E}[q_t]] = 0$. The Lagrange multipliers on the equality constraints are $\boldsymbol{\lambda}$ and λ_0, respectively. Our procedure implies that

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

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 1963:07–2017:06. 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 we 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.

			Properties of the SDF (monthly)					
	$L[m]$	$L[m^2]$	Mean	Std. Dev.	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 $\mathbb{S}$	1963:07–2017:06
$R_{t+1,f}$ Plus Market Plus	$\mathbf{R}_{t,t+1}$ is of Dimension 26×1

	$\mathbb{LB}_m$	Bootstrap p-Value	$\mathbb{LB}_{m^2}$	Bootstrap p-Value
25 size × book to market	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

In table 3, **SDF Is Augmented Fama and French(2015) with momentum portfolio.**

TABLE 3

Model Where log SDF Is Linear in a Set of Baseline Excess Returns, Augmented with Momentum Portfolios

The sample period for the exercise presented in Table 3 is 1963:07–2017:06 (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 an 11×1 vector of excess returns. The excess returns $\mathbf{er}^{\text{aug}}_{t,t+1}$ are based on $\mathbf{Q}_{t,t+1}$ from the baseline 5-factor model of Fama and French (2015) augmented with two momentum portfolios (reflecting the gross returns of past losers and winners):

- i) Gross return of the market
- ii) Gross returns of the two extreme low, and high, size portfolios
- iii) Gross returns of the two extreme low, and high, book-to-market portfolios
- iv) Gross returns of the two extreme low, and high, operating-profitability portfolios
- v) Gross returns of the two extreme low, and high, investment portfolios
- vi) Gross returns of the two extreme low, and high, momentum portfolios.

The constants $(\lambda_0, \boldsymbol{\lambda})$ are a solution to the minimization problem $\inf_{(\lambda_0, \boldsymbol{\lambda})} -\mathbb{E}[q_t]\lambda_0 + \mathbb{E}[\exp(\lambda_0 - 1 + \boldsymbol{\lambda}' \mathbf{er}^{\text{aug}}_{t,t+1})]$. This form of the stochastic discount factor (SDF) obtains by minimizing the objective $\mathbb{E}[m\log(m)]$ subject to $\mathbb{E}[m\,\mathbf{er}^{\text{aug}}] = \mathbf{0}$ and $\mathbb{E}[m - \mathbb{E}[q_t]] = 0$. The Lagrange multipliers on the equality constraints are $\boldsymbol{\lambda}$ and λ_0, respectively. Our procedure implies that

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

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 1963:07–2017:06. 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 we 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.

			Properties of the SDF (monthly)					
	$L[m]$	$L[m^2]$	Mean	Std. Dev.	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 $\mathbb{S}$	1963:07–2017:06			
$R_{t+1,f}$ Plus Market Plus	$\mathbf{R}_{t,t+1}$ is of Dimension 26×1			
	$\mathbb{LB}_m$	Bootstrap p-Value	$\mathbb{LB}_{m^2}$	Bootstrap p-Value
25 size × book to market	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**

4.2.2 class two

Table 4

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

$$we\ set\quad L[(m^{\bullet})^2] = LB_{m^2}$$
$$we\ test\quad L[m^{\bullet}] - LB_m \geq 0$$
$$m^{\bullet}_{t,t+1} = \frac{\Psi}{(a'R_{t,t+1})^{\nu}}$$

TABLE 4
Model That Passes the Bound on $L[m^2]$ also Passes the Bound on $L[m]$

In Table 4, the lower bound on $L[m^2]$, denoted by $\mathbb{LB}_{m^2}$, is based on equation (12), whereas 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 stochastic discount factor (SDF) that passes the bound $\mathbb{LB}_{m^2}$ is of the following type (see Problem 1):

$$m^{\bullet}_{t,t+1} = \frac{\Psi}{(a'R_{t,t+1})^{\nu}} \quad \text{(where } m^{\bullet} \text{ is consistent with } L[(m^{\bullet})^2]=\mathbb{LB}_{m^2}\text{).}$$

The parameters ν and Ψ solve equations (C-13) and (C-14), 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; and vii) gross returns of the two extreme low, and high, momentum portfolios. The sample period is 1963:07–2017:06. 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 book to market

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, shown in bold).

			Properties of $m^{\bullet}$						
ν	Ψ	$L[m^{\bullet}]$	$L[(m^{\bullet})^2]$	Mean	Std. Dev.	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

4.2.3 class three

Table 5

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

<div align="center">

TABLE 5

Model That Passes the Bound on L[m] but May or May Not Pass the Bound on L[m²]

</div>

In Table 5, the lower bound on $L[m^2]$, denoted by $\mathbb{LB}_{m^2}$, is based on equation (12), whereas 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 stochastic discount factor (SDF) that passes the bound $\mathbb{LB}_m$ is of the following type (see Problem 2):

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

The parameters α and β solve $0 = \mathbb{E}[q_t] - \exp((\alpha-1)\mathbb{E}[\log(\mathbf{a}'\mathbf{R}_{t,t+1})])\mathbb{E}[(\mathbf{a}'\mathbf{R}_{t,t+1})^{-\alpha}]$, and $0 = \beta - \exp((\alpha - 1)\mathbb{E}[\log(\mathbf{a}'\mathbf{R}_{t,t+1})])$. We use the following returns (a total of 12) to construct the SDF $m^G_{t,t+1}$: 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; and vii) gross returns of the two extreme low, and high, momentum portfolios. The sample period is 1963:07–2017:06. 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 book to market

   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 $L[(m^G)^2] \geq \mathbb{LB}_{m^2}$ (low p-values imply rejection, shown in bold).

<div align="center">Properties of m^G</div>

α	β	$L[m^G]$	$L[(m^G)^2]$	Mean	Std. Dev.	Skewness	Kurtosis	Min.	Max.
2.920	1.051	0.0220	0.0963	0.996	0.23	2.9	25.0	0.33	3.16

<div align="center">p-Values for the Null Hypothesis $L[(m^G)^2] \geq \mathbb{LB}_{m^2}$</div>

	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

> ✎ **What could be a possible rationale for findings of class two and three?**

- $L[m^2] - L[m]$ ***captures the tail behavior of the SDF distribution.***
- The returns used to build $m^\bullet$ are ***not the same as*** those used to construct LB_m. Our approach aligns with Hansen Jagannanthan(1991), who propose a minimum-variance SDF consistent with

a lower bound on the variance of the SDFs when