Sources of entropy in representative agent models

David Backus, Mikhail Chernov and Stanley Zin

September 13, 2010

Abstract
We use properties of asset returns to characterize the pricing kernel, including measures of its dispersion (implied by excess returns) and time-dependence (implied by bond yields). The same measures are then applied to representative agent models with long-run risk, stochastic volatility, several versions of habits, and jumps (departures from conditional lognormality). Using loglinear approximations, we show how each of these models generates dispersion and time-dependence of the pricing kernel. The exercise clarifies the mechanisms underlying these models and reveals their similarities and differences.

JEL Classification Codes: E44, G12.

Keywords: asset returns, pricing kernel, entropy, equity, bonds.

*Preliminary and incomplete: no guarantees of accuracy or sense. We welcome comments, including references to related papers we inadvertently overlooked.
† Stern School of Business, New York University, and NBER; dbackus@stern.nyu.edu.
‡ London School of Economics and CEPR; mchernov@lse.ac.uk.
§ Stern School of Business, New York University, and NBER; stan.zin@nyu.edu.
1 Introduction

Our objective is to describe the mechanisms underlying many of the popular models of asset pricing: to show how they generate the features of asset returns typically used to assess them.

In modern asset pricing theory, a pricing kernel accounts for differences in returns across assets. The reverse is also true: asset returns tell us about the pricing kernel. We use some of the salient properties of asset returns to establish properties of the pricing kernel and compare them to the same properties of some popular models of asset pricing.

2 Properties of asset returns

Summary: excess returns, nonnormal...

3 Properties of the pricing kernel

Modern asset pricing starts with this result: there is a positive random variable $m$ that satisfies the pricing relation

$$E_t \left( m_{t+1} r^j_{t+1} \right) = 1. \quad (1)$$

for gross returns $r^j$ on all traded assets. Here $E_t$ denotes the expectation conditional on information available at date $t$.

Stationary ergodic environment...

3.1 Entropy

The first result concerns the dispersion of the pricing kernel. Similar to HJ.

Define the conditional entropy of positive random variable $x$ by

$$L_t(x_{t+1}) \equiv \log E_t x_{t+1} - E_t \log x_{t+1}.$$
Measure of dispersion: always nonnegative, zero only if \( x \) is constant. Unconditional entropy, or simply entropy, is the same based on stationary distribution.

We can now state the entropy bound: In a stationary ergodic setting (wording?), the mean excess return is bounded above by the mean conditional entropy:

\[
EL_t(m_{t+1}) \geq E \left( \log r_{t+1} - \log r^1_{t+1} \right).
\]  

(2)

There’s a similar result for entropy (appendix).

Proof. 1. Since log is a concave function, Jensen’s inequality and the unconditional version of the pricing relation (1) imply that for any positive return \( r \),

\[
E_t \log m_{t+1} + E_t \log r_{t+1} \leq \log(1) = 0.
\]

2. Since the price of a one-period bond is \( q^1_t = E_t m_{t+1} \), the log short rate is

\[
\log r^1_{t+1} = -\log q^1_t = -\log E_t m_{t+1}.
\]

3. Combining the two, we have

\[
\log E_t m_{t+1} - E_t \log m_{t+1} = L_t(m_{t+1}) \geq E_t \left( \log r_{t+1} - \log r^1_{t+1} \right).
\]

4. Taking \( E \) of both sides, we have (2).

Like the HJ bound, expected excess returns place a lower bound on the dispersion of the pricing kernel. Here the measure of dispersion of entropy, and expected excess returns are in logs.

### 3.2 Time-dependence

Bond prices contain information about dynamics... The (gross) return on an \( n \)-period bond is \( r^1_{t+1} = q^{n-1}_t / q^n_t \). Applying (1) gives us the recursive pricing of bonds: \( q^n_t = E_t(m_{t+1}q^{n-1}_{t+1}) \) starting with \( q^0_t = 1 \). Repeated substitution yields

\[
q^n_t = E_t(m_{t+1}m_{t+2} \cdots m_{t+n}).
\]

Bond yields defined by : \( q^n_t = \exp(-ny^n) \) or \( ny^n = -\log q^n_t \).

Bond prices and yields tell us something about the dynamics of the pricing kernel. [Start with iid, then generalize.] Obviously, if \( ms \) iid, price is \( (E_t m_{t+1})^n \). Departures from iid evident in difference.

Define the “multiperiod pricing kernel” \( m_{t,t+n} = m_{t+1}m_{t+2} \cdots m_{t+n} \). Its entropy is

\[
L_t(m_{t,t+n}) = \log E_t m_{t,t+n} - E_t \log m_{t,t+n} = \log q^n_t - E_t \log m_{t,t+n}.
\]
Taking $E$ of both sides:

$$EL_t(m_{t,t+n}) = E \log q^n - nE \log m.$$ 

Difference from $n$ times $n = 1$ case:

$$nEL_t(m_{t+1}) - EL_t(m_{t,t+n}) = nE \log q^1 - E \log q^n = n(y^n - y^1).$$

Set up so difference is small? Numbers: suppose $EL_t(m_{t+1}) = 0.20$. For $n = 10$ years, we have $EL_t(m_{t+1}) = 2.00$. What about $EL_t(m_{t,t+n})$? Difference must be something like $n(y^n - y^1) = 10 \times 0.02 = 0.2$, which is small relative to the total.

Look at ratio? Something like

$$Ratio(n) = \frac{EL_t(m_{t,t+n})}{nEL_t(m_{t+1})}.$$ 

## 4 An example

Vasicek example: $\log m \sim \text{ARMA}(1,1) + \text{noise for equity}$.  

Start table with columns: model, $L(m)$, $EL_t(m_{t+1})$, $L(q^1)$, $E(y^{120} - y^1)$  

Talk about non-normal innovations... [??] Eg, $L(e^w)$.  

Power utility? Show how this works... Talk about preference shocks as a solution... State-dependent risk aversion?  

Stoch vol: show how nonlinearity works.

## 5 Recursive preferences

Recursive setup: prefs, scaling, mrs 

iid case  

The usual loglin approx. Show how we get the extra parameter back, jack up initial term.  

Bansal and Yaron...  

Drechsler and Yaron...  

Eraker and Shaliastovich
5.1 Recursive preferences

5.2 Persistent consumption growth

5.3 Stochastic volatility

6 Habits

Talk about habit/durability setup. Read Deaton and Deaton-Muelbauer on the general theory. Show how foc has (in general) multiple terms...

Heaton: loglinear approx... higher-order approx...

Abel: nada!

Recursive Abel...

Campbell and Cochrane...

7 Jumps

Add jumps...

Power utility

Wachter...

BCDG/Drechsler-Yaron

CC

8 Extensions

“Jumps”

This affects dispersion: mean conditional entropy...
Long memory

This affects dynamics...

9 Conclusions

Punchline: habits and recursive preferences are complements.
References


A Appendix

A.1 The entropy bound

OLD VERSION, CHANGE AS NEEDED

The entropy bound ((??) is derived by Alvarez and Jermann (2005) as a byproduct of their Proposition 2. Bansal and Lehmann (1997, Section 2.3) have a similar result that treats variation in the short rate differently [the term $L(q^1)$ in (5) below]. We derive the bound like this:

- **Bound on mean log return.** Since log is a concave function, Jensen’s inequality and the unconditional version of the pricing relation (1) imply that for any positive return $r$,

  \[ E \log m + E \log r \leq \log(1) = 0, \]

  with equality if and only if $mr = 1$. Therefore no asset has higher expected (log) return than the inverse of the pricing kernel:

  \[ E \log r \leq -E \log m. \quad (3) \]

  The asset with this return is sometimes called the “growth optimal portfolio.” We call it the “high-return asset.”

- **Short rate.** A one-period (risk-free) bond has price $q^1_t = E_t m_{t+1}$, so its return is $r^1_{t+1} = 1/E_t m_{t+1}$.

- **Entropy of the one-period bond price.** With the bound in mind, our next step is to express $E \log r^1$ in terms of unconditional moments. The entropy of the one-period bond price does the trick:

  \[ L(q^1) = \log E q^1 - E \log q^1 = \log Em + E \log r^1. \quad (4) \]

- **Entropy bound.** (3) and (4) imply

  \[ L(m) \geq E (\log r^j - \log r^1) + L(q^1). \quad (5) \]

  Inequality (??) follows from $L(q^1) \geq 0$ (entropy is nonnegative). In practice, $L(q^1)$ is small; in the iid case, it’s zero.

A.2 Log returns, Sharpe ratios, and leverage

It’s common to report Sharpe ratios for investment strategies (the mean of the excess returns divided by its standard deviation). The Hansen-Jagannathan bound then uses it to generate a lower bound on the ratio of the standard deviation of the pricing kernel to its mean. Our
entropy bound has a similar flavor but defies direct comparison. We work through a couple examples here.

Sharpe ratio. The simplest example is lognormal. Suppose an asset has return \( r^j \) with \( \log r^j \sim \mathcal{N}(\log r^1 + \kappa_1, \kappa_2) \). The entropy bound is based on the mean log excess return:

\[
E(\log r^j - \log r^1) = \kappa_1.
\]

The excess return in levels, \( x = r^j - r^1 \), has mean and variance

\[
E(x) = r^1 \left( e^{\kappa_1 + \kappa_2/2} - 1 \right)
\]

\[
\text{Var}(x) = \left( r^1 e^{\kappa_1 + \kappa_2/2} \right)^2 \left( e^{\kappa_2} - 1 \right).
\]

The mean has the “Ito term” \( \kappa_2/2 \), but is otherwise similar to what we had above. The Sharpe ratio is

\[
\text{SR} = \frac{E(x)}{\text{Var}(x)^{1/2}} = \frac{e^{\kappa_1 + \kappa_2/2} - 1}{e^{\kappa_1 + \kappa_2/2} (e^{\kappa_2} - 1)^{1/2}}.
\]

For a small time interval, this is approximately

\[
\text{SR} \approx \frac{\kappa_1 + \kappa_2/2}{\kappa_2^{1/2}}.
\]

If the return isn’t lognormal, high-order terms show up in both numerator and denominator.

Leverage. For the same example, consider the same asset levered \( \lambda \) times. Its return is \( r = \lambda(r^j - r^1) + r^1 \) and its excess return is \( r - r^1 = \lambda(r^j - r^1) \). The Sharpe ratio doesn’t change. What about the difference in log returns? Consider the function

\[
f(\lambda) = E \log [\lambda(r^j - r^1) + r^1].
\]

We know \( f(0) = \log r^1 \) and \( f(1) = E \log r^j = \log r^1 + \kappa_1 \). If \( \kappa_1 > 0 \), \( f \) is increasing between zero and one. At some point at or above one, it starts decreasing. ...

A.3 Bond premiums

Basics. Let \( q(n,t) \) be the price of an \( n \)-period zero at \( t \). Then yields and forward rates are defined by

\[
-\log q^n_t = ny^n_t = f^0_t + f^1_t + ... + f^{n-1}_t.
\]

where \( f^0_t \) is the short rate.

Premiums. The one-period return on an \( n \)-period bond is

\[
r^n_{t+1} = \log(q^n_{t+1}/q^n_t).
\]
We define the bond risk premium as \( bp(n, t) = E_t r(n, t + 1) - f(0, t) \). and the (forward) term premium as

\[
f^n_t = E_t f^0_{t+n} + tp^n_t.
\]

It’s now just algebra to express the bond premium in terms of forward rates and term premiums. The excess return on an \((n+1)\)-period bond is

\[
r^n_{t+1} - f^0_t = (f^1_t - f^0_{t+1}) + \ldots + (f^n_t - f^{n-1}_{t+1}).
\]

The typical term is

\[
f^j_t - f^{j-1}_{t+1} = (E_t f^0_{t+j} - E_{t+1} f^0_{t+j}) + tp^j_t - tp^{j-1}_{t+1}.
\]

That is, you get innovations in forward rates and changes in term premiums. Since the former have (conditional) mean zero by construction, only the latter show up in bond risk premiums:

\[
bp_{t}^{n+1} = E_t (tp^1_{t+1} - tp^0_t) + E_t (tp^2_{t+1} - tp^1_t) + \ldots + E_t (tp^n_{t+1} - tp^{n-1}_t)
\]

\[
= E_t (tp^1_{t+1} - tp^1_t) + E_t (tp^2_{t+1} - tp^2_t) + \ldots + E_t (tp^n_{t+1} - tp^n_t) + tp^n_t.
\]

Kind of a mess, but it shows that term premiums are the ingredients of bond premiums. The unconditional means are the same: \( E(bp^{n+1}) = E(tp^n) \).

Yield curve...

### A.4 The pricing kernel for recursive preferences

The pricing kernel in a representative agent model is the marginal rate of substitution between (say) consumption at date \( t \) \([c_t]\) and consumption in state \( s \) at \( t + 1 \) \([c_{t+1}(s)]\). Here’s how that works with recursive preferences. With this notation, the certainty equivalent (??) might be expressed less compactly as

\[
\mu_t(U_{t+1}) = \left[ \sum_s \pi(s) U_{t+1}(s)^\alpha \right]^{1/\alpha},
\]

where \( \pi(s) \) is the conditional probability of state \( s \) and \( U_{t+1}(s) \) is continuation utility. Some derivatives of (??) and (??):

\[
\frac{\partial U_t}{\partial c_t} = U_t^{1-\rho} (1 - \beta) c_t^\rho - 1
\]

\[
\frac{\partial U_t}{\partial \mu_t(U_{t+1})} = U_t^{1-\rho} \beta \mu_t(U_{t+1})^{\rho - 1}
\]

\[
\frac{\partial \mu_t(U_{t+1})}{\partial U_{t+1}(s)} = \mu_t(U_{t+1})^{1-\alpha} \pi(s) U_{t+1}(s)^\alpha - 1.
\]

The marginal rate of substitution between consumption at date \( t \) and consumption in state \( s \) at \( t + 1 \) is

\[
\frac{\partial U_t}{\partial c_{t+1}(s)} = \frac{[\partial U_t/\partial \mu_t(U_{t+1})][\partial \mu_t(U_{t+1})/\partial U_{t+1}(s)][\partial U_{t+1}(s)/\partial c_{t+1}(s)]}{\partial U_t/\partial c_t}
\]

\[
= \pi(s) \beta \left( \frac{c_{t+1}(s)}{c_t} \right)^{\rho - 1} \left( \frac{U_{t+1}(s)}{\mu_t(U_{t+1})} \right)^{\alpha - \rho}.
\]
The pricing kernel is the same with the probability $\pi(s)$ left out and the state left implicit.

A.5 The Campbell-Shiller approximation

We define equity at $t$ as a claim to consumption from $t + 1$ on. The return is the ratio of its value at $t + 1$, measured in units of $t + 1$ consumption, to the value at $t$, measured in units of $t$ consumption. The value at $t + 1$ is $U_{t+1}$ expressed in $c_{t+1}$ units:

$$U_{t+1}/(\partial U_{t+1}/\partial c_{t+1}) = U_{t+1}/[(1-\beta)U_{t+1}^{1-\rho}c_{t+1}^{-\rho}] = (1-\beta)^{-1}u_{t+1}^\rho c_{t+1}.$$  

The value at $t$ is the certainty equivalent expressed in $c_t$ units:

$$q_t c_t = \partial U_t/\partial \mu_t(U_{t+1}) = \beta \mu_t(U_{t+1})^\rho/(1-\beta)c_t^\rho c_t.$$  

The return is the ratio:

$$r^{c}_{t+1} = \beta^{-1}[u_{t+1}/\mu_t(x_{t+1}u_{t+1})]^\rho x_{t+1} = \beta^{-1}[x_{t+1}u_{t+1}/\mu_t(x_{t+1}u_{t+1})]^\rho x_{t+1}^{1-\rho}.$$  

Check to see if this satisfies the Euler equation:

$$E_t(m_{t+1}r^{c}_{t+1}) = E_t[x_{t+1}u_{t+1}/\mu_t(x_{t+1}u_{t+1})]^\alpha = \mu_t(x_{t+1}u_{t+1})^\alpha/\mu_t(x_{t+1}u_{t+1})^\alpha = 1.$$  

[To do: Connect to Campbell-Shiller approx... Details needed, but we should be able to show that the approx in $u$ is equiv to that in $q$ with the same $\kappa_1$.]

Levered equity...
### Table 1
Properties of excess returns

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>Autocorr</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Equity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Valued-weighted market</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fama-French (small, low)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fama-French (small, high)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fama-French (large, low)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fama-French (large, high)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Nominal bonds</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 months</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 year</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 years</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 years</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 years</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Other assets</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equity option straddles</td>
<td>249.xx</td>
<td>15.xx</td>
<td>1.2</td>
<td>2.2</td>
<td></td>
</tr>
</tbody>
</table>

Notes. Entries are sample moments of monthly observations of log excess returns: $\log r^j - \log r^1$. The mean and standard deviation have been multiplied by 1200 to convert them to annual percentages. The other statistics are invariant to scale. Sample periods: ...