Backus, Chernov, & Zin

Notes on Time-Dependence and Entropy*

(Started: August 8, 2009; Revised: August 17, 2009)

We assess the role of time-dependence in generating entropy in the pricing kernel. Examples: stochastic volatility, jumps, and so on. All of this is done in the familiar exchange economy with Kreps-Porteus/Epstein-Zin/Weil preferences. Solutions are based on loglinear approximations, whose mechanics build on a long line of earlier research.

Facts about entropy

We focus on the entropy of the pricing kernel. The entropy of a positive random variable x is

$$L(x) \quad = \quad \log Ex - E \log x.$$

Our interest stems from the Alvarez-Jermann bound on the entropy of the pricing kernel:

$$L(m) \quad \geq \quad E(\log r^j - \log r^1) + L(q^1),$$

where r^1 is the (gross) return on a one-period bond, $q^1 = 1/r^1$ is its price, and r^j is the return on any other asset.

Some facts about the entropy of the pricing kernel:

- Entropy. The equity premium gives us a lower bound on $L(m)$ of around 0.05. Alvarez and Jermann (Econometrica, 2005, Table I, column 1) report 0.0664 for the postwar period. The prewar period is a little lower, but other assets can give larger numbers.

- Conditional entropy. Entropy can be decomposed into

  $$L(m) \quad = \quad EL_t(m_{t+1}) + L(E_t m_{t+1}) \quad = \quad EL_t(m_{t+1}) + L(q^1).$$

  With annual data, AJ estimate $L(q^1)$ to be around 0.0005, with monthly data slightly less (Table I, column 3). This is about two orders of magnitude smaller than the equity premium, so it's small change. Evidently most of $L(m)$ comes from $EL_t(m_{t+1})$. The result is similar to Cochrane and Hansen (Macro Annual, 1992, Section 2.7), who show that the variance of the conditional mean of the pricing kernel is much smaller than the mean of the conditional variance.

  Note, too, that if we combine this with the AJ bound, we get

  $$EL_t(m_{t+1}) \quad \geq \quad E(\log r^j - \log r^1),$$

  which is easier to compute in many models. The bottom line in any case is that the action is in conditional entropy.

*Working notes, no guarantee of accuracy or sense.

- Dynamics. We're looking for a result that says the pricing kernel is negatively autocorrelated if the mean yield curve is upward sloping. Here's a sketch of a version based on entropy, more work needed. Let $y^n_t = -n^{-1} \log q^n_t$ be the yield on an n-period bond. Then we're looking for something like: If $Ey^2 > Ey^1$ (upward-sloping mean yield curve), then $L(m_t m_{t+1}) < 2L(m)$ (negative autocorrelation). A one-period bond implies

  $$\begin{aligned} L_t(m_{t+1}) &= \log E_t m_{t+1} - E_t \log m_{t+1} \;=\; -y^1_t - E_t \log m_{t+1} \\ L(m) &= -Ey^1 - E \log m + L(q^1). \end{aligned}$$

  Similarly, a two-period bond leads to

  $$\begin{aligned} L_t(m_{t+1}m_{t+2}) &= \log E_t(m_{t+1}m_{t+2}) - E_t \log(m_{t+1}m_{t+2}) \\ L(m_{t+1}m_{t+2}) &= -2Ey^2 - 2E \log m + L(q^2). \end{aligned}$$

  That gives us

  $$2L(m) - L(m_{t+1}m_{t+2}) \;=\; 2E(y^2 - y^1) + [2L(q^1) - L(q^2)].$$

  Both terms are positive in the data (entropy of bond prices should increase less than proportionately with maturity). Second term small in any case. Also: can kill off last term using mean conditional entropy; extends naturally to other maturities.

  [extend to bond returns?]

Theory

Consumption (= output). We specify a stationary process for the growth rate $x_t = c_t/c_{t-1}$. Asset prices then depend on whatever state variables show up in the x process.

Preferences. Define "utility from data t on" recursively with the time aggregator,

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

and certainty equivalent function,

$$\mu_t(U_{t+1}) \;=\; \left[E_t(U_{t+1}^\alpha)\right]^{1/\alpha}.$$

Here $\rho < 1$ captures time preference (the intertemporal elasticity of substitution is $1/(1-\rho)$) and $\alpha < 1$ captures risk aversion (the coefficient of relative risk aversion is $1-\alpha$). Additive utility is a special case with $\alpha = \rho$.

Both the time aggregator and certainty equivalent function are homogeneous of degree one, which allows us to scale everything by current consumption. If we define scaled utility $u_t = U_t/c_t$, equation (1) becomes

$$u_t \;=\; \left[(1-\beta) + \beta\mu_t(x_{t+1}u_{t+1})^\rho\right]^{1/\rho}.$$

A loglinear approximation is

$$\begin{aligned}\log u_t &= \rho^{-1}\log\left[(1-\beta)+\beta\mu_t(x_{t+1}u_{t+1})^{\rho}\right] \\ &= \rho^{-1}\log\left[(1-\beta)+\beta e^{\rho\log\mu_t(x_{t+1}u_{t+1})}\right] \\ &\approx b_0+b_1\log\mu_t(x_{t+1}u_{t+1}). \end{aligned} \tag{2}$$

The last line is a first-order approximation of $\log u_t$ in $\log\mu_t$ around the point $\log\mu_t=\log\mu$, with

$$\begin{aligned} b_1 &= \beta e^{\rho\log\mu}/[(1-\beta)+\beta e^{\rho\log\mu}] \\ b_0 &= \rho^{-1}\log[(1-\beta)+\beta e^{\rho\log\mu}]-b_1\log\mu. \end{aligned}$$

The approximation is exact when $\rho=0$, in which case $b_0=0$ and $b_1=\beta$. Note: this is the only source of approximation in what follows, everything else is exact. Note, too: b_0 and b_1 are not free parameters.

Pricing kernel. With these preferences, the pricing kernel (marginal rate of substitution) is

$$\begin{aligned} m_{t+1} &= \beta(c_{t+1}/c_t)^{\rho-1}\left[U_{t+1}/\mu_t(U_{t+1})\right]^{\alpha-\rho} \\ &= \beta x_{t+1}^{\rho-1}\left[x_{t+1}u_{t+1}/\mu_t(x_{t+1}u_{t+1})\right]^{\alpha-\rho}. \end{aligned} \tag{3}$$

This has a nice loglinear structure, too, as long as x and u do.

Example: Bansal-Yaron

Consider asset pricing with a univariate loglinear consumption growth process plus stochastic volatility:

$$\begin{aligned} \log x_t &= x+\sum_{j=0}^{\infty}\chi_j v_{t-1}^{1/2}w_{1t-j} = x+\chi(L)v_{t-1}^{1/2}w_{1t} \\ v_t &= (1-\varphi_v)v+\varphi_v v_{t-1}+\sigma_v w_{2t}, \end{aligned}$$

where $\{(w_{1t},w_{2t})\}\sim \text{NID}(0,I)$.

Value function. Given this process, we can derive an approximate value function that is loglinear in the state: volatility v and the history of shocks w_1.

- Guess a loglinear value function:

$$\log u_t = u+\sum_{j=0}^{\infty}\omega_j w_{1t-j}+\omega_v v_t$$

  with parameters to be determined.

- Given the guess, compute the log of $x_{t+1}u_{t+1}$ and its certainty equivalent:

$$\begin{aligned}
\log(x_{t+1}u_{t+1}) &= x + u + (\chi_0 + \omega_0)v_t^{1/2}w_{1t+1} + \sum_{j=0}^{\infty}(\chi_{j+1} + \omega_{j+1})v_{t-j-1}^{1/2}w_{1t-j} \\
&\quad + \omega_v[(1-\varphi_v)v + \varphi_v v_t + \sigma_v w_{2t+1}] \\
\log \mu_t(x_{t+1}u_{t+1}) &= x + u + \omega_v(1-\varphi_v)v + \alpha(\omega_v\sigma_v)^2/2 \\
&\quad + \sum_{j=0}^{\infty}(\chi_{j+1} + \omega_{j+1})v_{t-j-1}^{1/2}w_{1t-j} + [\alpha(\chi_0+\omega_0)^2/2 + \omega_v\varphi_v]v_t.
\end{aligned}$$

- Solve the "Bellman equation" for the parameters. If we substitute the parameters into (2) and collect terms, we get

$$\begin{aligned}
\text{constant}: \quad u &= b_0 + b_1\left[x + u + \omega_v(1-\varphi_v)v + \alpha(\omega_v\sigma_v)^2/2\right] \\
w_{1t-j}: \quad \omega_j &= b_1(\chi_{j+1} + \omega_{j+1}) \\
v_t: \quad \omega_v &= b_1[\alpha(\chi_0+\omega_0)^2/2 + \omega_v\varphi_v].
\end{aligned}$$

  The first equation defines u; we'll ignore it, although ultimately it's needed to compute b_1. The second leads to a forward-looking geometric sum familiar to readers of Sargent's books. Iterating forward, we have (for each $j \geq 0$)

$$\begin{aligned}
\omega_j &= \sum_{i=1}^{\infty} b_1^i \chi_{j+i} \equiv X_{j+1} \\
\chi_j + \omega_j &= \chi_j + X_{j+1} = \sum_{i=0}^{\infty} b_1^i \chi_{j+i} = X_j/b_1.
\end{aligned}$$

  The upper case X_js are geometric sums reflecting the impact of innovations to current consumption growth on future utility. They summarize the Bansal-Yaron "predictable component" in the sense that if x_t is iid, $X_j = 0$ for $j \geq 1$. Hansen would write $\chi_0 + \omega_0 = \chi_0 + X_1 = \chi(b_1)$. That implies

$$\omega_v = (1 - b_1\varphi_v)^{-1} b_1 \alpha \chi(b_1)^2/2.$$

  Note that the dynamics of consumption growth (represented here by $\chi(b_1)$) affect the impact of volatility on utility. So does the persistence of volatility.

Pricing kernel. One component is

$$\begin{aligned}
\log(x_{t+1}u_{t+1}) - \log\mu_t(x_{t+1}u_{t+1}) &= \chi(b_1)v_t^{1/2}w_{1t+1} + \omega_v\sigma_v w_{2t+1} \\
&\quad - (\alpha/2)(\omega_v\sigma_v)^2 - (\alpha/2)\chi(b_1)^2 v_t.
\end{aligned}$$

If we substitute into (3), we get the pricing kernel

$$\begin{aligned}
\log m_{t+1} &= \log\beta + (\rho-1)x - (\alpha-\rho)(\alpha/2)(\omega_v\sigma_v)^2 \\
&\quad + [(\rho-1)\chi_0 + (\alpha-\rho)\chi(b_1)]v_t^{1/2}w_{1t+1} + (\alpha-\rho)(\omega_v\sigma_v)w_{2t+1} \\
&\quad + (\rho-1)\sum_{j=0}^{\infty}\chi_{j+1}v_{t-j-1}^{1/2}w_{1t-j} - (\alpha-\rho)(\alpha/2)\chi(b_1)^2 v_t.
\end{aligned}$$

In the iid case, $\chi(b_1) = \chi_0$. Otherwise, there's an additional role for the "persistent component."

Its conditional entropy is (since everything is conditionally normal) one-half the conditional variance:

$$
\begin{aligned}
L_t(m_{t+1}) &= [(\rho-1)\chi_0 + (\alpha-\rho)\chi(b_1)]^2 v_t/2 + (\alpha-\rho)^2(\omega_v \sigma_v)^2/2 \\
&= [(\rho-1)\chi_0 + (\alpha-\rho)\chi(b_1)]^2 v_t/2 + (\alpha-\rho)^2 \chi(b_1)^4 [\alpha\sigma_v/(1-b_1\varphi_v)]^2/2.
\end{aligned}
$$

Mean conditional entropy follows from substituting v for v_t.

Calibration. Here's a quick and dirty set of monthly numbers, based loosely on Bansal and Yaron. There are two pieces: the dynamics of consumption growth (the χs in our notation) and the process for volatility. We streamline the first into a univariate process, because it's easier to illustrate the role of dynamics that way.

Consumption growth. The Bansal-Yaron growth rate process is the sum of an AR(1) and white noise. It implies, using their notation,

$$
\begin{aligned}
\text{Var}(x) &= \sigma_v^2 + (\varphi_e \sigma_v)^2/(1-\rho^2) \\
\text{Cov}(x_t, x_{t-1}) &= \rho(\varphi_e \sigma_v)^2/(1-\rho^2) \\
\text{Corr}(x_t, x_{t-1}) &= \text{Cov}(x_t, x_{t-1})/\text{Var}(x).
\end{aligned}
$$

With input from their Table I ($\rho = 0.979$, $\sigma_v = 0.0078$, $\varphi_e = 0.044$), the unconditional standard deviation is 0.0080 and the first autocorrelation is $\rho(1) = 0.0436$.

We construct an ARMA(1,1) with the same autocovariances. The essential parameters are $(\chi_0, \chi_1, \varphi)$, with the rest of the MA coefficients defined by $\chi_{j+1} = \varphi\chi_j = \varphi^j \chi_1$ for $j \geq 1$. If we define $b = \chi_1/\chi_0$, this implies

$$
\begin{aligned}
\text{Var}(x) &= \chi_0^2 + \chi_1^2/(1-\varphi^2) &&= \chi_0^2\left[1 + b^2/(1-\varphi^2)\right] \\
\text{Cov}(x_t, x_{t-1}) &= \chi_0\chi_1 + \varphi\chi_1^2/(1-\varphi^2) &&= \chi_0^2\left[b + \varphi b^2/(1-\varphi^2)\right] \\
\text{Corr}(x_t, x_{t-1}) &= \frac{b + \varphi b^2/(1-\varphi^2)}{1 + b^2/(1-\varphi^2)} .
\end{aligned}
$$

We set $\varphi = 0.979$ (BY's ρ). We choose b to match the autocorrelation $\rho(1)$, which gives us a quadratic in b:

$$
b^2[\varphi - \rho(1)] + (1-\varphi^2)b - \rho(1)(1-\varphi^2) \;=\; 0.
$$

We choose the "+" root (needed to get invertible MA?):

$$
b \;=\; \frac{-(1-\varphi^2)^2 + \left\{(1-\varphi^2) + 4[\varphi-\rho(1)](1-\varphi^2)\rho(1)\right\}^{1/2}}{2[\varphi - \rho(1)]} \;=\; 0.0271.
$$

To match the first autocorrelation, we set $\chi_0 = 1$ (a normalization) and $\chi_1 = 0.0271$. These imply $\chi(b_1) = 2.1311$ (using the BY value $b_1 = 0.997$).

Volatility. The mean conditional variance of x is $v = 0.0079^2$, the autocorrelation $\varphi_v = 0.987$ (ν_1 in Table IV), and the innovation volatility $\sigma_v = 0.23 \times 10^{-5}$ (σ_w in the table).

Preferences. We set $\alpha = -9$ (so that the CRRA is 10) and $\rho = 1/3$ (so that the IES is 1.5). With the other parameters, that implies $\omega_v = -1280$.

Entropy. With these parameters, the two components of entropy are

$$\begin{aligned} [(\rho-1)\chi_0 + (\alpha-\rho)\chi(b_1)]^2 v/2 &= 0.0112 \\ (\alpha-\rho)^2(\omega_v\sigma_v)^2/2 &= 0.0004. \end{aligned}$$

Evidently volatility has a minor role here. Total entropy is 0.0116. (Remember, this is monthly.)

The Bansal, Kiku, and Yaron (2007, "A note...") numbers make volatility more important. See BBK (Table IV) or Beeler and Campbell (2009, NBER 14788, Table I). With $b_1 = 0.9989$ (the discount factor, the first-order approximation), we get entropy of

$$EL_t(m_{t+1}) = 0.0065 + 0.0153 = 0.0218.$$

This is pretty sensitive to the choice of b_1, but the basic idea must be right: by making volatility more persistent, we've increased its impact on utility and therefore on asset pricing. This despite shrinking $\chi(b_1)$ to 1.77. How much depends on the dynamics and recursive preferences? If we set $\rho = \alpha$, the second component is zero and the first falls to 0.0026. So we're getting a lot of action from the dynamics.

[All these computations in `entropy_BansalYaron.m`.]

Example: Wachter

Here's a similar approach to Wachter (2009, "Time-varying risk of rare disasters"). Let log consumption growth be

$$\log x_t = x + \sigma w_t + z_t,$$

where $\{w_t\} \sim \text{NID}(0,1)$ and $\{z_t\}$ is a Poisson mixture of normals with intensity λ_{t-1} and distribution $\text{N}(\theta, \delta^2)$. The jump intensity is AR(1):

$$\lambda_t = (1-\varphi)\lambda + \varphi\lambda_{t-1} + \tau v_t$$

with $\{v_t\} \sim \text{NID}(0,1)$. [Comment. We use a linear process rather than a square root process for the same reason as above: with the square root, we can still find an approximate value function of the same form, but the value function parameter for that state variable then solves a quadratic equation, which is a little less transparent than the linear equation we get this way.]

Value function. Look for one that's loglinear in λ_t.

- Guess a value function of the form

$$\log u_t \;=\; u + \omega\lambda_t.$$

- Given the guess, compute the log of $x_{t+1}u_{t+1}$ and its certainty equivalent:

$$\begin{aligned}
\log(x_{t+1}u_{t+1}) &= x + u + \omega(1-\varphi)\lambda + \omega\varphi\lambda_t + \sigma w_{t+1} + z_{t+1} + \omega\tau v_{t+1} \\
\log \mu_t(x_{t+1}u_{t+1}) &= x + u + \omega(1-\varphi)\lambda + \omega\varphi\lambda_t + (\alpha/2)[\sigma^2 + (\omega\tau)^2] + \lambda_t(e^{\alpha\theta+(\alpha\delta)^2/2} - 1)/\alpha.
\end{aligned}$$

  [Check last term.]

- Solve the Bellman equation for ω:

$$\omega \;=\; b_1[\omega\varphi + (e^{\alpha\theta+(\alpha\delta)^2/2} - 1)/\alpha] \;=\; (1-b_1\varphi)^{-1} b_1(e^{\alpha\theta+(\alpha\delta)^2/2} - 1)/\alpha.$$

Pricing kernel and entropy. The "long-run risk" component is

$$\log(x_{t+1}u_{t+1}) - \log \mu_t(x_{t+1}u_{t+1}) \;=\; \sigma w_{t+1} + z_{t+1} + \omega\tau v_{t+1} - (\alpha/2)[\sigma^2 + (\omega\tau)^2] - \lambda_t(e^{\alpha\theta+(\alpha\delta)^2/2} - 1)/\alpha$$

If we substitute into (3), we get the pricing kernel

$$\begin{aligned}
\log m_{t+1} &= \log\beta + (\rho-1)x - (\alpha-\rho)(\alpha/2)[\sigma^2 + (\omega\tau)^2] \\
&\quad + (\alpha-1)(\sigma w_{t+1} + z_{t+1}) + (\alpha-\rho)[\omega\tau v_{t+1} - \lambda_t(e^{\alpha\theta+(\alpha\delta)^2/2} - 1)/\alpha.
\end{aligned}$$

Its entropy is

$$L_t(m_{t+1}) \;=\; (\alpha-1)^2\sigma^2/2 + \lambda_t\left\{[e^{(\alpha-1)\theta+(\alpha-1)^2\delta^2/2} - 1] - (\alpha-1)\theta\right\} + (\alpha-\rho)^2(\omega\tau)^2/2.$$

The mean: replace λ_t with λ. Note that the last term is the only one in which recursive preferences or the dynamics of λ_t matter. The other terms would be the same in an iid setting with power utility.

Calibration. The mapping between our parameters and Wachter's with numbers from her Table I, adapted to a time interval $h = 1/12$ (monthly):

$$\begin{aligned}
\alpha &= 1-\gamma \;=\; -2 \\
\rho &= 1 - 1/\text{IES} \;=\; 0 \\
\sigma &= 0.02h^{1/2} \;=\; 0.0058 \\
\lambda &= 0.017h \;=\; 0.0014 \\
\tau &= \sigma_\lambda \lambda^{1/2} h^{1/2} \;=\; 0.978 \times 10^{-3} \\
\varphi &= e^{-\kappa h} \;=\; 0.9882 \\
b_1 &= e^{-\beta h} \;=\; 0.9983.
\end{aligned}$$

Remarks: (i) IES is one ($\rho = 0$), in which case the Bellman equation is exact and b_1 is the discount factor. (ii) Jump parameters are $\theta = -0.3$ and $\delta = 0.15$ (Backus, Chernov, and Martin, "Disasters in options"). (Wachter bases the jump distribution "on the empirical distribution" documented by Barro. This is similar, probably a little more modest.) (iii) Together, these parameters imply $\omega = 67.45$.

Entropy. The three components (same order as above) are 0.0001, 0.0012, and 0.0087, for a total of 0.0100. Most of the entropy comes from the dynamics via recursive preferences (last term). If we set $\rho = \alpha$, the contributions are 0.0001, 0.0012, and 0.0000 (only the last one changes).

[All these computations in `entropy_Wachter.m`.]

This suggests some variations. Namely:

- Wachter with dynamics in consumption growth. Suppose we use

$$\log x_t \quad = \quad x + \chi(L)(\sigma w_t + z_t).$$

  Conjecture: dynamics affect the impact of jumps, just as they did with volatility in Bansal-Yaron. This has a similar flavor to Drechsler and Yaron.

- Wachter with jumps in the λ process. Note that we have control (through $\alpha - \rho$) on the weight on v_{t+1} in the pricing kernel. Conjecture: this gives us another route for high-order cumulants in the pricing kernel.

- Santa-Clara and Yan (2008, "Crashes"). Different process for jumps.