Risk, Uncertainty, and Asset-Pricing ‘Antipuzzles’

Martin L. Weitzman

UC Berkeley: April 17, 2006

This talk only informal, motivational, gestalt. A lot is happening: paper has details. First pass introduction to three asset-pricing big-picture “macro puzzles.” Use Table 1.

Simplest possible “big picture” version of fruit-tree economy. Very compressed treatment of asset pricing. Growth rate is $x$, consumption is $C$. Basic stochastic equation is

$$\ln C_{t+1} = x_t + \ln C_t. \quad (1)$$

A large fixed number of identical people. Marginal utility is $U'(C) = C^{-\gamma}$, where coefficient of relative risk aversion is $\gamma$. Welfare is

$$W_t = E_t \left[ \frac{1}{1-\gamma} \sum_{j=0}^{\infty} \beta^j (C_{t+j})^{1-\gamma} \right]. \quad (2)$$

Basic idea: if asset $\alpha$ has payoff $\Pi^\alpha_{t+1}$, then return is $R^\alpha_{t+1} = \Pi^\alpha_{t+1}/P_t^\alpha$ and price $P_t^\alpha$ comes from Euler equation

$$\beta E_t[\exp(-\gamma x_{t+1}) R^\alpha_{t+1}] = 1. \quad (3)$$

Next, iid assumption: $x \sim iid$. For riskfree asset, payoff is $\Pi^\alpha_{t+1}=$constant and one-period

<table>
<thead>
<tr>
<th>Quasi-Constant Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean arithmetic return on equity</td>
<td>$\ln E[R^e] \approx 7%$</td>
</tr>
<tr>
<td>Geometric standard deviation of return on equity</td>
<td>$\sigma[R^e] \approx 17%$</td>
</tr>
<tr>
<td>Risk-free interest rate</td>
<td>$r_f \approx 1%$</td>
</tr>
<tr>
<td>Implied equity premium</td>
<td>$\ln E[R^e] - r_f \approx 6%$</td>
</tr>
<tr>
<td>Mean growth rate of per-capita consumption</td>
<td>$E[x] \approx 2%$</td>
</tr>
<tr>
<td>Standard deviation of growth rate of per-capita consumption</td>
<td>$\sigma[x] \approx 2%$</td>
</tr>
<tr>
<td>Rate of pure time preference</td>
<td>$\rho \approx 2%$</td>
</tr>
<tr>
<td>Coefficient of relative risk aversion</td>
<td>$\gamma \approx 2%$</td>
</tr>
</tbody>
</table>

Table 1: Some Stylized Economic "Facts"
rate of return on a riskfree asset is

\[ R_f = \frac{1}{\beta E[\exp(-\gamma x)]} \]  \hspace{1cm} (4)

or \( r_f = \rho - \ln E[\exp(-\gamma x)] \), where \( \rho = -\ln \beta \) and \( r_f = \ln R_f \).

Comprehensive equity as claim on future fruit. With iid assumption, essentially \( \Pi_{t+1} = C_{t+1} C_t \). Derive \( R_f = \exp(x_{t+1}) \beta E[\exp((1 - \gamma) x)] \) and equity premium in ratio form is

\[ \frac{E[R_e]}{R_f} = \frac{E[\exp(x)] E[\exp(-\gamma x)]}{E[\exp((1 - \gamma) x)]}. \]  \hspace{1cm} (5)

Next, normality assumption: \( x \sim N(\mu, V) \). Formulas simplify. Derive equity premium puzzle

\[ \ln E[R_e] - r_f = \gamma V \implies \ln E[R_e] - r_f \approx 0.08\%, \]  \hspace{1cm} (6)

and riskfree rate puzzle

\[ r_f = \rho + \gamma \mu - \frac{1}{2} \gamma^2 V \implies r_f \approx 5.9\% \]  \hspace{1cm} (7)

Can also derive

\[ x_t - E[x]. \]  \hspace{1cm} (8)

but \( \sigma[r_e] \approx 17\% \), while \( \sigma[x] \approx 2\% \). Call this variability mismatch puzzle (explain).

Thus far, classical methodology. Stationary setup. Insider agents know parameters, outsider econometricians have lots of data. It is presumed to be basically OK to have agents substitute \((\hat{x}, \hat{V})\) for \((\mu, V)\) in their Euler equations.

An important clue to what is happening here: suppose insider agents also not sure about structural parameter values. Try heuristic of using “Bayesian-t” distribution (What is relation of Student-t to normal? What do we expect should happen?)

\[ \phi(x) \propto \left(1 + \frac{(x - \mu)^2}{(n - 1)V} \right)^{-n/2}. \]  \hspace{1cm} (9)

Plug (9) into (4) and (5), which gives

\[ r_f = -\infty \]  \hspace{1cm} (10)

and

\[ \ln E[R_e] - r_f = +\infty \]  \hspace{1cm} (11)
Point Calibration Puzzles
predicted \( \ln E[R^e] - r^f \) is WAY TOO LOW
predicted \( r^f \) is WAY TOO HIGH
predicted \( \sigma[r^e] - \sigma[x] \) is WAY TOO LOW

“Bayesian-t” Antipuzzles
predicted \( \ln E[R^e] - r^f \) is WAY TOO HIGH
predicted \( r^f \) is WAY TOO LOW
predicted \( \sigma[r^e] - \sigma[x] \) is WAY TOO HIGH

Table 2: What Are We Trying to Explain?

and (explain intuition here) also gives \( \sigma[r_e] - \sigma[x] = +\infty \). Qualitative reversal of pattern but quantitative surprise. In Table 2, which set of “facts” should we be trying to explain?

Three Big Questions: 1 Is this result generic? 2 What does it mean? 3 How can we get rid of the infinities?

Part II: Try to build a formal Bayesian “hidden structure” model. Details in paper, here just sketch. Two goals: (1) show that whenever there are variance shocks Student-t does not converge to normal, so that antipuzzle pattern persists – even with unlimited data; (2) show how to “contain the t-explosion” (basic idea here – trying to find some distribution “between” normal and Student-t, hoping if lucky to fit data exactly).

Let \( \theta \equiv 1/V \) be precision of a probability distribution. Assume that, conditional on \( \theta = \theta_t \), the growth rate \( x = x_t \) is independent-normal, for convenience with known mean \( \mu \):

\[
x_t | \theta_t \sim N(\mu, 1/\theta_t). \tag{12}
\]

Can show by induction that the posterior probability density function of \( \theta \) always takes the gamma form

\[
\Gamma(a, b) \propto \theta_t^{a-1} \exp(-b \theta), \tag{13}
\]

for which \( E[\theta_t] = a/b \) and \( V[\theta] = a/b^2 \). Then can show that if prior is \( \Gamma(a_0, b_0) \) for any \( a_0, b_0 \), and have \( n \) observations, for any \( n \), then posterior at time \( t \) must be of form \( \Gamma(a_t, b_t) \), which induces a Student-t distribution for \( x \), causing \( EMU = \infty \). In paper, show that the startling asset-pricing antipuzzle pattern (from Student-t-distributed growth rates) persists whenever there are variance shocks – even with unlimited data. Tell as if story of running a regression, always with \( n \) data points, where lose oldest but gain newest in each period.

Next task: how to contain “Student-t-explosion” in order to obtain workable theory. Any number of conceivable ad hoc “dampening specifications” might be used, all of which impose arbitrary a priori restrictions on probability distributions or utility functions – and they all give essentially the same final message. Let \( \delta \) be a hyperparameter representing an imposed a priori lower-bound support for the distribution of \( \theta \). Assume that the Bayesian
distribution of the precision is now truncated gamma (with truncation hyperparameter \( \delta \))

\[
\varphi_\delta(\theta | a_\tau, b_\tau) = \frac{\theta^{\mu_\tau - 1} \exp(-b_\tau \theta)}{\int_\delta^\infty \theta^{\mu_\tau - 1} \exp(-b_\tau \theta) d\theta}
\]

(14)

for \( \theta_\tau \geq \delta \), while \( \varphi_\delta(\theta | a_\tau, b_\tau) = 0 \) for \( \theta_\tau < \delta \). (What is \( \delta \)? Nobody knows!) Because conjugate-recursive normal-gamma family structure is preserved for \( \delta > 0 \), get relatively simple expression with \( \theta_t \) truncated at \( \theta_t = \delta \).

To summarize:

\[
\psi(\theta | \nu, n, \delta) = \frac{\theta^{-n_\tau - \frac{1}{2}} \exp(-\frac{n_{\tau 1} - \nu}{2} \theta)}{\int_\delta^\infty \theta^{-n_\tau - \frac{1}{2}} \exp(-\frac{n_{\tau 1} - \nu}{2} \theta) d\theta}
\]

(15)

is posterior of \( \theta \geq \delta \), while unconditional posterior of \( x \) is

\[
g(x | \nu, n, \delta) = \frac{\int_\delta^\infty \exp(-\frac{(x-\mu)^2}{2}) \theta_{n_\tau}^{\nu - \frac{1}{2}} \exp(-\frac{n_{\tau 1} - \nu}{2} \theta) d\theta}{\int_{-\infty}^\delta \{ \int_\delta^\infty \exp(-\frac{(x-\mu)^2}{2}) \theta_{n_\tau}^{\nu - \frac{1}{2}} \exp(-\frac{n_{\tau 1} - \nu}{2} \theta) d\theta \} \, dx}
\]

(16)

Note \( n = \infty \) implies \( x \sim N(\mu, \nu) \). The case \( \delta \rightarrow 0 \) is Student-t “as if” a regression had been run on a sample of size \( n \). Expected marginal utility or stochastic discount factor or pricing kernel is moment generating function of \( x \) and therefore depends critically on \( \delta \). For asset pricing cannot get rid of influence of prior beliefs. Why this non-ergodicity in space of \( MU = \exp(-\gamma x) \) undermines rational expectations approach to asset pricing via fitting empirically Euler equation. (No prior in rational expectations!) Program of prior-belief-free “objective” asset macro-pricing essentially doomed. Explain classical and Bayesian approaches to: “let data speak for themselves.” Next: three partial-equilibrium theorems showing how \( \delta \) can match observed data.

**Hidden-structure equity premium.** *Theorem 1* (simplified, free translation): First, for all \( n < \infty \), some value of \( \delta \) matches essentially any given one period ahead asset-price ratio

\[
\frac{P^f(\nu, n, \delta)}{P^{1e}(\nu, n, \delta)} = \frac{\int_{-\infty}^\infty e^{-\gamma x} g(x | \nu, n, \delta) \, dx}{\int_{-\infty}^\infty e^{(1-\gamma)x} g(x | \nu, n, \delta) \, dx},
\]

(17)

no matter how large. Second, by choosing carefully \( \delta(n, \nu) \) and then going to the limit \( n \rightarrow \infty \), any desired equity premium \( \pi > \gamma \nu \) can be replicated in the data as if it came from the super-simple i.i.d.-normal model of Section 2 in the sense that \( R^{1e}/R^f \rightarrow \exp(\pi - \)
\[
\mu - \nu/2 \exp(\sqrt{\gamma} z), \text{ where } z \text{ is i.i.d. } N(0, 1). \quad \text{“As if” simple iid model but with } x \text{ between normal and } t. \quad \text{(Note that equity premium is matched but not equity variability.)}
\]

Discussion of what drives result. Intuition: what do businesses and individuals fear more – risk or uncertainty? Economic interpretation of fact that m.g.f. of \( t \) is \(+\infty\). “Bayesian peso problem.” Translation into classical-frequentist language. How general is result?

**Hidden-structure riskfree rate.** Use (4) and (16) to define equity premium as function of \( \delta \):

\[
f(\delta) \equiv \rho - \ln \int_{-\infty}^{\infty} \exp(-\gamma x) g(x \mid \nu, n, \delta) \, dx. \tag{18}
\]

**Theorem 2:** Let \( r^f \) be any given continuous function of \( \nu \) satisfying \( r^f < \rho + \gamma \mu - \frac{1}{2} \gamma^2 \nu \) for all \( \nu > 0 \). Then for every \( n > 0, \nu > 0 \), there exists a \( \delta_f(n, \nu) > 0 \) such that

\[
r^f = f(\delta_f). \tag{19}
\]

Discussion & intuition. By choosing carefully \( \delta(n, \nu) \) and then going to the limit \( n \to 0 \), can make the data simulate over time essentially any pre-selected constant value of \( r^f \).

**Hidden-structure variability.** Actual evolving trajectories very complicated, extremely dependent upon priors. Suppose try to force into mold of rational-expectations-like as-if-iid-normal. From (8), should have \( \sigma[r_x] = \sigma[x] \), but in the data \( \hat{\sigma}[r_x] \gg \hat{\sigma}[x] \). Question: which value (\( \hat{\sigma}[x] \) or \( \hat{\sigma}[r_x] \)) comes closer to describing true welfare of risk-averse agent uncertain about future growth prospects? Certainly not \( \hat{\sigma}[x] \). Let’s try \( \hat{\sigma}[r_x] \). Approach is heuristic, intuitive. Telling an “as if” story. Examine welfare consequences on rest of GE system of forcing the model to match standard assumption \( r_x \) is iid-normal with known variance.

Putting aside the “rationality” of such beliefs, suppose

\[
x^N(x \mid \nu, n, \delta) \sim N(E[x^N], \sigma^2[x^N]) \tag{20}
\]

represents agent’s subjective probability belief that future growth rates are i.i.d.-normal with known parameters \( E[x^N] \) and \( \sigma[x^N] \). Let agent also have subjective probability belief in a stock-market payoff implicitly representing a unit claim on the lognormally-i.i.d. future aggregate output corresponding to (20). Such payoff claim gives rise to the subjective probability belief of a return on comprehensive economy-wide equity \( r^N(x^N) \) satisfying

\[
r^N(x^N(x)) - E[r^N] = x^N(x) - E[x^N], \tag{21}
\]

which is exactly the normal counterpart here of equation (8).
Theorem 3: Let \( \sigma > 0 \) be any given continuous function of \( \nu \) satisfying \( \sigma^2 > \nu \) for all \( \nu > 0 \). Let \( r^N(x^N(x)) \) be a solution of (20), (21). Then for every \( n < \infty, \nu > 0 \), there exists a \( \delta_\nu(n, \nu) > 0 \) such that the following four conditions are simultaneously satisfied:

\[
E[r^N(x^N(x))] = E[r^{1e}(x)], \tag{22}
\]

\[
\sigma[r^N(x^N(x))] = \sigma[x^N(x)] = \sigma, \tag{23}
\]

\[
E[x^N(x)] = E[x] = \mu, \tag{24}
\]

\[
\forall C > 0: E[U(C \exp(x^N(x)))] = E[U(C \exp(x))], \tag{25}
\]

A free translation of Theorem 3 says that if you must compress the complicated reality of future evolutionary uncertainty with super-sensitivity to prior beliefs into a simple stationary-ergodic prior-free as-if-i.i.d.-normal-growth story, then \( \sigma[x^N] = \tilde{\sigma}[r^e] \) (equity-wealth variability) tells the better welfare parable than \( \sigma[x^N] = \tilde{\sigma}[x] \) (growth rate variability).

Some Bayesian calibration exercises. Theorem 3 suggests thinking in terms of \( x_N \sim N(\mu, \tilde{\sigma}[r^e]) \) as welfare equivalent of \( x \sim g(x \mid \nu, n, \delta) \). Suppose the welfare-equivalent growth variability \( \sigma[x_N] = \tilde{\sigma}[r^e] \) replaces the realized-sample growth variability \( \tilde{\sigma}[x] \). Perform some Bayesian calibration exercises to test consistency with lognormal formulas (6) and (7). Two equations, three unknowns. Plug numerical value of any one of the trio of (1) \( \ln E[R^c] - r^f \); (2) \( r^f \); (3) \( \sigma[x_N] = \tilde{\sigma}[r^e] \) from Table 1 into lognormal formulas (6) and (7) and then back out the other two numerical values. Result is a near-perfect fit with values of Table 1. How likely is this to be coincidence?

With i.i.d. lognormality, imaginary deterministic path having same mean consumption as stochastic trajectory (1) is \( \bar{C}_{t+1} = \exp(\mu - \frac{1}{2}\sigma^2) \bar{C}_t \). Following Lucas (2003), can show welfare gain from a mean-preserving shrinkage that compresses stochastic trajectory (1) into the as-if-deterministic path is equivalent to a change in each period’s consumption of

\[
\Delta C_t = (\exp(\frac{1}{2}\gamma\sigma^2) - 1) C_t. \tag{26}
\]

When \( \gamma \approx 2 \) and the historical value of \( \sigma = \sigma[x] \approx 2\% \) is used in (26), then \( \Delta C_t/C_t \approx 0.04\% \). When \( \sigma = \sigma[x_N] = \tilde{\sigma}[r^e] \approx 17\% \) is used, then \( \Delta C_t/C_t \approx 3\% \). (Ask self intuitively, which number better captures welfare gain from eliminating all uncertainty about future growth?) In this sense, structural uncertainty about the future growth process is empirically far more significant than stationary-ergodic growth-rate risk.