Empirical Asset Pricing

Part 1: Asset Pricing Puzzles and Possible Resolutions

UCLA | Fall 2016

by Mikhail Chernov

This version: September 1, 2016

1. Motivation: Asset pricing

Main Focus

- Measuring and understanding risk premiums

- The Law of One Price implies

$$1 = E_t\left[M_{t+1}R^i_{t+1}\right]$$

where M_{t+1} is the stochastic discount factor (sdf), or pricing kernel, R^i_{t+1} is a gross return on asset i. Note that the gross risk free rate is $R_{f,t} = 1/E_t\left[M_{t+1}\right]$.

- Therefore,

$$1 = \frac{1}{R_{f,t}}E_t\left[R^i_{t+1}\right] + \underbrace{cov_t\left(M_{t+1}, R^i_{t+1}\right)}_{\text{Risk Premium}}$$

and expected net excess returns are then given by

$$E_t\left[NRX^i_{t+1}\right] \equiv \frac{E_t\left[R^i_{t+1}\right]}{R_{f,t}} - 1 = -cov_t\left(M_{t+1}, R^i_{t+1}\right) \tag{1}$$

Measuring risk premiums

- We get to observe realized net excess returns,

$$R^i_{t+1}/R_{f,t} - 1 = E_t\left[NRX^i_{t+1}\right] + \text{shocks}$$

- If risk premiums are constant and we have a very long sample, we could compute expected excess returns by averaging the realized ones

- The realized excess returns are strongly time-varying leading researchers to believe that shocks cannot be the sole explanation of this time-variation

- If risk premiums are time-varying we have to measure conditional expectations, a much more difficult task

Understanding risk premiums

- We can construct highly sophisticated empirical models that would help us measuring risk premiums
- But where are they coming from?
- A structural model allows addressing such questions
- Structural models are much less flexible than empirical models so they have difficulty matching the measured risk premiums despite providing interesting stories
- Over time, especially in the last 10 years, we see convergence of the two

2. Valuation

Valuation

- We have to specify the pricing kernel M, then

$$P_t = E_t(M_{t,t+1}CF_{t+1})$$

- So, again, the knowledge of conditional distributions is really helpful here

- Note, that the knowledge of CGF could be enough

$$P_t = E_t(e^{m_{t,t+1}+cf_{t+1}})$$

- Equilibrium models produce M by interacting preferences with quantities that investors care about (consumption, labor, etc.)

- In empirical models we have a lot of leeway

Change of probabilities

- Recall that valuation with pricing kernel is equivalent to risk-neutral valuation:

$$P_t = E_t\left(E_t(M_{t,t+1})\frac{M_{t,t+1}}{E_t(M_{t,t+1})}CF_{t+1}\right) = E_t^{\mathbb{Q}}\left(e^{-r_t}CF_{t+1}\right),$$

  where

$$\frac{M_{t,t+1}}{E_t(M_{t,t+1})} \equiv L_{t,t+1} = \frac{p_X^{\mathbb{Q}}(X_{t+1}|X_t)}{p_X^{\mathbb{P}}(X_{t+1}|X_t)}$$

- The change of measure could be exceptionally flexible – too many degrees of freedom
- Example: $M_{t,t+1} = e^{\alpha_t' X_{t+1} + \beta_t}$ (Esscher transform)

$$\begin{aligned} \log L_{t,t+1} &= \alpha_t' X_{t+1} - k_t(X_{t+1}, \alpha_t), \\ k_t^{\mathbb{Q}}(X_{t+1}, s) &= k_t(X_{t+1}, s+\alpha_t) - k_t(X_{t+1}, \alpha_t) \end{aligned}$$

Limitations of the Esscher transform

- Consider the Merton model

$$X_{t+1} = \mu + \sigma\varepsilon_{t+1} + jZ_{t+1}$$

with

$$k_t(X_{t+1}, s) = s\mu + (s\sigma)^2/2 + \omega[e^{s\theta+(s\delta)^2/2} - 1]$$

- Then

$$\begin{aligned} k_t^{\mathbb{Q}}(X_{t+1}, s) &= s(\mu + \alpha\sigma^2) + (s\sigma)^2/2 \\ &+ \omega e^{\alpha\theta+(\alpha\delta^2)/2}[e^{s(\theta+\alpha\delta^2)+(s\delta)^2/2} - 1] \end{aligned}$$

- Not flexible: changes in Poisson and normal components are controlled by one parameter (α)
- Monfort and Pegoraro (2012) propose a generic approach via second-order Esscher transform

$$M_{t,t+1} = e^{X'_{t+1}\alpha_2 X_{t+1} + \alpha'_1 X_{t+1} + \beta}$$

Alternative approach

- Our equations imply

$$M_{t,t+1} = e^{-r_t} L_{t,t+1}, \quad \text{or} \quad m_{t,t+1} = -r_t + \ell_{t,t+1}$$

- So, specify $\ell_{t,t+1}$ any way you want, just make sure that $k_t(\ell_{t,t+1}, 1) = 0$

- For example, if $\ell_{t,t+1}$ is multivariate normal, then $\ell_{t,t+1} = -\lambda_t'\lambda_t/2 - \lambda_t'\varepsilon_{t+1}$

- If $\ell_{t,t+1}$ is jumpy, then $\ell_{t,t+1} = -\omega[e^{-\theta_\lambda + \delta_\lambda^2/2} - 1] - j\tilde{Z}_{t+1}$, where $\tilde{Z}_{t+1} \sim \mathcal{N}(\theta_\lambda, \delta_\lambda^2)$

A valuation example: Nominal Bonds

- From Chernov and Mueller (2012)
- The log real pricing kernel is

$$m_{t,t+1} = -r_t - \lambda_t'\lambda_t/2 - \lambda_t'\varepsilon_{t+1}$$

- The inflation rate is

$$\pi_{t,t+1} = \pi_t^e + \sigma\varepsilon_{t+1}^{\pi},$$

  where π_t^e is expected inflation.
- The inflation shock ε^{π} is correlated with real shocks ε:

$$\varepsilon_t^{\pi} = \rho' \cdot \varepsilon_t + \sqrt{1-\rho'\rho} \cdot \varepsilon_t^{\perp}$$

- The log nominal pricing kernel is

$$m_{t,t+1}^{\$} = m_{t,t+1} - \pi_{t,t+1} = -r_t - \pi_t^e - \lambda_t'\lambda_t/2 - \lambda_t'\varepsilon_{t+1} - \sigma\varepsilon_{t+1}^{\pi}.$$

- The nominal bond is $E_t e^{m_{t,t+1}^{\$}}$

The Fisher hypothesis

- The nominal (log) yield is

$$
\begin{aligned}
r_t^{\$} &= -\log E_t e^{m^{\$}_{t,t+1}} = r_t + \pi_t^e + \lambda_t'\lambda_t/2 \\
&- \log E_t e^{-(\lambda_t+\sigma\rho)'\varepsilon_{t+1} - \sigma\sqrt{1-\rho'\rho}\varepsilon^{\perp}_{t+1}} \\
&= r_t + \pi_t^e + \lambda_t'\lambda_t/2 - (\lambda_t+\sigma\rho)'(\lambda_t+\sigma\rho)/2 - \sigma^2(1-\rho'\rho)/2 \\
&= r_t + \pi_t^e - \sigma\lambda_t'\rho - \sigma^2/2
\end{aligned}
$$

- The Fisher hypothesis posits: $r_t^{\$} = r_t + \pi_t^e$
- The third term is referred to as inflation risk premium:

$$\sigma\lambda_t'\rho = -cov_t(m_{t,t+1}, \pi_{t,t+1})$$

("price of risk" λ_t times "quantity of risk" $\sigma\rho$)

- If λ_t is a constant, Fisher is not far off

A valuation example: Jumps

- Based loosely on Backus, Chernov, and Martin (2011)
- Consider asset's log excess return

$$rx_{t+1} = \mu + \sigma\varepsilon_{t+1} + jZ_{t+1}$$

- Consider the following change of probabilities:

$$\ell_{t,t+1} = -\sigma_\lambda^2/2 - \omega[e^{-\theta_\lambda+\delta_\lambda^2/2} - 1] - \sigma_\lambda\varepsilon_{t+1} - j\tilde{Z}_{t+1}$$

- The risk-neutral CGF of returns is:

$$\begin{aligned} & k_t^{\mathbb{Q}}(rx_{t+1}, s) = \log E_t^{\mathbb{Q}} e^{s\cdot rx_{t+1}} = \log E_t e^{\ell_{t,t+1}+s\cdot rx_{t+1}} \\ = \quad & -\sigma_\lambda^2/2 - \omega[e^{-\theta_\lambda+\delta_\lambda^2/2} - 1] + s\mu + \log E_t e^{(s\sigma-\sigma_\lambda)\varepsilon_{t+1}+(sZ_{t+1}-\tilde{Z}_{t+1})j} \\ = \quad & -\sigma_\lambda^2/2 - \omega[e^{-\theta_\lambda+\delta_\lambda^2/2} - 1] + s\mu + (s\sigma-\sigma_\lambda)^2/2 \\ + \quad & \omega[e^{s\theta-\theta_\lambda+(s\delta-\delta_\lambda)^2/2} - 1] = \ldots \end{aligned}$$

A valuation example: Jumps (2)

- Continuing

$$\begin{aligned} k_t^{\mathbb{Q}}(rx_{t+1}, s) &= -\sigma_\lambda^2/2 - \omega[e^{-\theta_\lambda+\delta_\lambda^2/2} - 1] + s\mu + (s\sigma - \sigma_\lambda)^2/2 \\ &+ \omega[e^{s\theta-\theta_\lambda+(s\delta-\delta_\lambda)^2/2} - 1] \\ &= s(\mu - \sigma_\lambda\sigma) + s^2\sigma^2/2 + \omega e^{-\theta_\lambda+\delta_\lambda^2/2}[e^{s(\theta-\delta_\lambda\delta)+s^2\delta^2/2} - 1] \end{aligned}$$

- This tells you immediately the $\mathbb{Q}$ distribution of rx_{t+1}:
  - The "normal component" has mean $\mu - \sigma_\lambda\sigma$ and variance σ^2
  - Jumps arrive at the rate of $\omega e^{-\theta_\lambda+\delta_\lambda^2/2}$, their mean is $\theta - \delta_\lambda\delta$ and variance δ^2
- "Risk premiums" σ_λ, θ_λ, and δ_λ must satisfy $k_t^{\mathbb{Q}}(rx_{t+1}, 1) = 0$
- For example, in the absence of jumps, $\sigma_\lambda = (\mu + \sigma^2/2)/\sigma$ – the Sharpe ratio
- Here we see risk premiums again: $\sigma_\lambda\sigma$, $\delta_\lambda\delta$, and a more tricky one $e^{-\theta_\lambda+\delta_\lambda^2/2}$

3. Preliminary evidence

Measuring returns

We have different types of assets, and each his certain specifics associated with its own type.

- Equities: $R^e_{t+1} = (P^e_{t+1} + D_{t+1})/P^e_t$, where P^e_t and D_t are price and dividend
- (Zero-coupon) bonds: $R^b_{t+1}(T) = P^b_{t+1}(T)/P^b_t(T)$, where T is time **of** maturity, and $P^b_t(T)$ is the time t price. Time **to** maturity is different in the numerator and denominator. In particular, $R_{f,t} \equiv R^b_{t+1}(t+1) = 1/P^b_t(t+1)$.
- Currencies: A naive approach would be $R^{fx}_{t+1} = S_{t+1}/S_t$, where S_t is the domestic price of the foreign currency (for example if US is home, then it is a \$ price of £, i.e., \$1.6). However, this is wrong. If you hold domestic currency, you are not making any returns. You have to explicitly convert into foreign currency. Convert \$1 into £$1/S_t$, invest £at the foreign (UK) risk-free rate $R^*_{f,t}$. In one period, receive £$1/S_t \cdot R^*_{f,t}$ and convert into \$$S_{t+1}/S_t \cdot R^*_{f,t}$. So, $R^{fx}_{t+1} = S_{t+1}/S_t \cdot R^*_{f,t}$.
- Derivatives: we will have to discuss on a case-by-case basis. It becomes more involved, and, at times, not clear. For example, an interest rate swap is very similar to bonds in terms of cash flows. However, it is unfunded (zero initial investment). To be absolutely precise, one has to take into account margin requirement which differ from one party to another. Another example: equity options. Long positions are easy. What about shorts?

Measuring excess returns

- In practice, researchers rarely use $E_t[R^i_{t+1}]/R_{f,t} - 1$ to measure excess returns. They use either

$$E_t\left[RX^i_{t+1}\right] \equiv E_t\left[R^i_{t+1} - R_{f,t}\right] = -R_{f,t}cov_t\left(M_{t+1}, R^i_{t+1}\right)$$

or

$$E_t\left[rx^i_{t+1}\right] \equiv E_t\left[\log(R^i_{t+1}/R_{f,t})\right] = ???$$

Interpreting excess returns

- Start with the Law of One Price

$$\begin{aligned} 0 &= \log E_t(M_{t+1} R^i_{t+1}) \\ &= \log E_t(M_{t+1} R^i_{t+1}) - \log E_t M_{t+1} - \log E_t R^i_{t+1} + \log E_t M_{t+1} + \log E_t R^i_{t+1} \\ &= C_t(M_{t+1}, R^i_{t+1}) - \log R_{f,t} + \log E_t R^i_{t+1} \end{aligned}$$

- This leads to two results:
  1. (Log) risk premium:

$$\log E_t(R^i_{t+1}/R_{f,t}) = -C_t(M_{t+1}, R^i_{t+1})$$

  2. (Log) excess returns:

$$E_t\left[rx^i_{t+1}\right] = -C_t(M_{t+1}, R^i_{t+1}) - L_t(R^i_{t+1})$$

Excess returns: the case of log-normals

- The last expression simplifies for normally distributed variables:

$$E_t\left[r^i_{t+1} - r_{f,t}\right] = -cov_t(m_{t+1}, r^i_{t+1}) - var_t(r^i_{t+1})/2$$

- Therefore, you will often see risk premia defined as

$$E_t \tilde{rx}_{t+1} = E_t r^i_{t+1} - r_{f,t} + 0.5 var_t(r^i_{t+1}).$$

In this case,

$$E_t \tilde{rx}_{t+1} = -cov_t(m_{t+1}, \tilde{rx}_{t+1}) \quad (2)$$

- In practice people use *rx* for bonds and currencies, and *RX* for stocks and options (if an option expires worthless, it is impossible to compute log-returns).

Evidence

Asset	Mean	Standard Deviation	Skewness	Excess Kurtosis
Equity				
S&P 500	0.0040	0.0556	−0.40	7.90
Fama-French (small, low)	−0.0030	0.1140	0.28	9.40
Fama-French (small, high)	0.0090	0.0894	1.00	12.80
Fama-French (large, low)	0.0040	0.0548	−0.58	5.37
Fama-French (large, high)	0.0060	0.0775	−0.64	11.57
Equity options				
S&P 500 6% OTM puts (delta-hedged)	−0.0184	0.0538	2.77	16.64
S&P 500 ATM straddles	−0.6215	1.1940	−1.61	6.52
Currencies				
CAD	0.0013	0.0173	−0.80	4.70
JPY	0.0001	0.0346	0.50	1.90
AUD	−0.0015	0.0332	−0.90	2.50
GBP	0.0035	0.0316	−0.50	1.50
Nominal bonds				
1 year	0.0008	0.0049	0.98	14.48
2 years	0.0011	0.0086	0.52	9.55
3 years	0.0013	0.0119	−0.01	6.77
4 years	0.0014	0.0155	0.11	4.78
5 years	0.0015	0.0190	0.10	4.87

Source: Backus, Chernov, and Zin (2014)

Basic theories about excess returns

- Equities: CAPM. $E_t\left[RX^i_{t+1}\right] = \beta E\left[RX^M_{t+1}\right]$
- (Zero-coupon) bonds: Expectation hypothesis. $E_t\left[rx^b_{t+1}(T)\right] = const$
- Currencies: Uncovered Interest Rate Parity (UIP). $E_t\left[rx^{fx}_{t+1}\right] = 0$
- Derivatives: CAPM – this is how Black and Scholes derived their famous formula.

You may see this already (and if not, we will discuss this in the class in detail), all of these “theories” are one and the same assumption: M_t is constant. Why would this be a reasonable assumption? We have to look at the evidence to answer this question. However, it is worth mentioning, that these theories are very old. They come from the dark ages where no one knew about pricing kernels, and each asset class got its own separate treatment.

4. AP puzzles

Excess returns revisited

- Equation (2)

$$E_t \tilde{rx}_{t+1} = -cov_t(m_{t+1}, \tilde{rx}_{t+1})$$

  implies

$$\frac{E_t \tilde{rx}_{t+1}}{\sigma_t(\tilde{rx}_{t+1})} = -corr_t(m_{t+1}, \tilde{rx}_{t+1})\sigma_t(m_{t+1})$$

- If an asset is perfectly correlated with the pricing kernel, its Sharpe ratio is equal to volatility of the pricing kernel

- This case corresponds to the maximal Sharpe ratio

CRRA preferences

- Utility function

$$\begin{aligned} U_t &= (1-\beta)u(C_t) + \beta E_t U_{t+1}, \\ u(C_t) &= C_t^{\alpha}/\alpha, \end{aligned}$$

  where $1-\alpha$ is the CRRA$= -\partial \log U_t / \partial \log C_t$

- The pricing kernel

$$m_{t+1} = \log \beta - (1-\alpha) g_{t+1}, \quad g_{t+1} = \log C_{t+1}/C_t$$

- The Sharpe ratio

$$\frac{E_t \tilde{rx}_{t+1}}{\sigma_t(\tilde{rx}_{t+1})} = (1-\alpha) corr_t(g_{t+1}, \tilde{rx}_{t+1}) \sigma_t(g_{t+1})$$

- The max Sharpe ratio is $(1-\alpha)\sigma_t(g_{t+1})$

The failure of CRRA preferences

- $\sigma(g_{t+1}) \approx 3.5\%$ per year (1889-2009)
- The Sharpe ratio of the market is about 0.5 per year
- So, even if the market is perfectly correlated with consumption growth, we have

  $$0.5 = (1-\alpha) \times 0.035$$

  So, CRRA must be at least 14.5.
- In the data, $corr(g_{t+1}, \tilde{rx}_{t+1}) \approx 0.5$, so $1-\alpha > 29$
- Macro is comfortable with the range from 2 to 5; finance “allows” up to 10

Risk-free rate puzzle

- Risk-free rate:

  $$r_{f,t} = -\log E_t e^{m_{t+1}} = -\log \beta + (1-\alpha) E_t g_{t+1} - (1-\alpha)^2/2 \cdot var_t g_{t+1}$$

- The terms are: rate of time preference, consumption smoothing over time, and precautionary savings
- $Eg_{t+1} \approx 2\%$, so the second term dominates the third for large $1-\alpha$
- Take $1-\alpha = 29$: r_f =7.5% – way too high for a real rate
- Intuition: high CRRA means strong desire to smooth consumption; positive expected consumption growth means investors want to consume today by borrowing; in equilibrium, net borrowing is zero, so risk-free rate must be extremely high to discourage borrowing

A generalisation: The HJ bound

- The Hansen and Jagannathan (1991) bound gives minimum variance frontier for admissible stochastic discount factors.
- It is a visual interpretation of a model performance (not a formal test)
- The unconditional law of one price for N assets

$$Eq_t = E(x_{t+1}M_{t+1})$$

  where q_t is $N \times 1$ vector of asset prices, x_{t+1} is $N \times 1$ vector of payoffs and M_{t+1} is the 1×1 pricing kernel implied by a model

- Assume unobserved risk free asset, i.e., cannot observe EM_{t+1}
- We want to characterise how $var(M_{t+1})$ changes with EM_{t+1} (in log-normal world there is nothing else)

The minimum variance pricing kernel

- Find the MV PK by "regressing" M_{t+1} on x_{t+1}

$$M_{t+1} = x'_{t+1}\alpha_0 + \varepsilon_{t+1}$$

- Coefficients α_0 are then

$$\alpha_0 = \left(E\left[x_{t+1}x'_{t+1}\right]\right)^{-1} E\left[x_{t+1}M_{t+1}\right]$$

- Denote the fitted M_{t+1} as M^*_{t+1}. It also satisfies the law of one price.

$$E\left(x_{t+1}M^*_{t+1}\right) = E\left(x_{t+1}x'_{t+1}\right)\alpha_0 = E\left(x_{t+1}M_{t+1}\right) = E\left(q_t\right)$$

The variance of the MV PK

- This is the minimum variance a model's PK must have to be able to price assets correctly.
- Deviations of M^*_{t+1} from its mean

$$M^*_{t+1} - E\left(M^*_{t+1}\right) = \left(x_{t+1} - E\left(x_{t+1}\right)\right)' \alpha_0$$

- Rewrite the law of one price

$$\begin{aligned} E\left(q_t\right) &= E\left(x_{t+1}\right) E\left(M^*_{t+1}\right) + E\left[\left(x_{t+1} - E\left(x_{t+1}\right)\right)\left(M^*_{t+1} - E\left(M^*_{t+1}\right)\right)\right] \\ &= E\left(x_{t+1}\right) E\left(M^*_{t+1}\right) + E\left[\left(x_{t+1} - E\left(x_{t+1}\right)\right)\left(x_{t+1} - E\left(x_{t+1}\right)\right)' \alpha_0\right] \\ &\equiv E\left(x_{t+1}\right) E\left(M^*_{t+1}\right) + \Sigma \alpha_0 \end{aligned}$$

- Therefore,

$$\alpha_0 = \Sigma^{-1}\left(E\left(q_t\right) - E\left(M^*_{t+1}\right) E\left(x_{t+1}\right)\right)$$

- Then

$$\begin{aligned} M^*_{t+1} &= x'_{t+1}\alpha_0 = E\left(M^*_{t+1}\right) + \left(x_{t+1} - E\left(x_{t+1}\right)\right)' \alpha_0 \\ &= E\left(M^*_{t+1}\right) + \left(x_{t+1} - E\left(x_{t+1}\right)\right)' \Sigma^{-1}\left(E\left(q_t\right) - E\left(M^*_{t+1}\right) E\left(x_{t+1}\right)\right) \end{aligned}$$

The HJ cup

- Use $M^*_{t+1} = x'_{t+1}\alpha_0$

$$
\begin{aligned}
var\left(M^*_{t+1}\right) &= \alpha_0'\Sigma\alpha_0 \\
&= \left[E\left(q_t\right) - E\left(x_{t+1}\right)E\left(M^*_{t+1}\right)\right]'\Sigma^{-1}\left[E\left(q_t\right) - E\left(x_{t+1}\right)E\left(M^*_{t+1}\right)\right] \\
&= \left[E\left(q_t\right) - E\left(x_{t+1}\right)E\left(M_{t+1}\right)\right]'\Sigma^{-1}\left[E\left(q_t\right) - E\left(x_{t+1}\right)E\left(M_{t+1}\right)\right]
\end{aligned}
$$

- The locus of points $\left\{EM_{t+1}, \sqrt{var\left(M^*_{t+1}\right)}\right\}$, which are the MV PK's for each given expected value of the PK, are then called the *HJ bound*.

- How do we use these?
  - Take N assets. Normalize prices to zero or one, cash flows to gross returns or excess returns, correspondingly
  - Estimate Σ and average payoffs $E\left(R_{t+1}\right)$
  - Plot the above locus
  - Compare the locus to the first two moments for your PK (e.g, in the CRRA case, $var(M_{t+1}) \approx (1-\alpha)^2\sigma^2\left(g_{t+1}\right)$)

The HJ cup

Hansen-Jagannathan do this for the CRRA using data on aggregate stocks, bonds, and plot the power utility PK mean and variance for given parameters. Shows that you need very high level of risk aversion to enter the admissible region

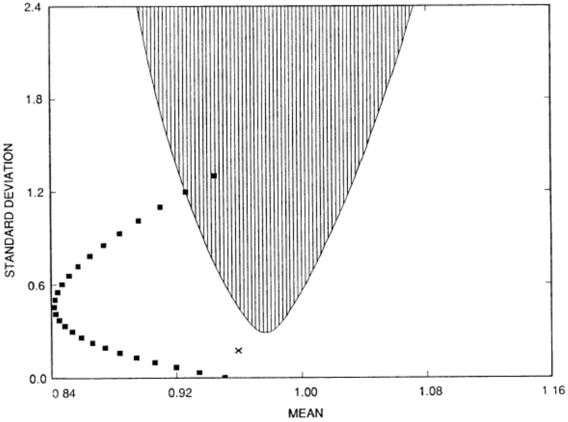


FIG. 1.—IMRS frontier computed using annual data

A generalisation: The entropy bound

- Dispersion: entropy bound

$$EL_t(M_{t+1}) \quad \geq \quad E(r_{t+1} - r_t^1)$$

- Disasters: high-order cumulants

$$\begin{aligned} L_t(M_{t+1}) &= k_t(m_{t+1}, 1) - \kappa_{1t}(m_{t+1}) \\ &= \underbrace{\kappa_{2t}(m_{t+1})/2!}_{\text{normal term}} + \underbrace{\kappa_3(m_{t+1})/3! + \kappa_4(m_{t+1})/4! + \cdots}_{\text{high-order cumulants}} \end{aligned}$$

Derivation of the Entropy Bound

- The LOOP

$$
\begin{aligned}
E_t(M_{t+1}R_{t+1}) &= 1, \\
E_t m_{t+1} + E_t r_{t+1} \le \log(1) &= 0, \text{ with equality iff } M_{t+1}R_{t+1} = 1
\end{aligned}
$$

  The growth-optimal portfolio has the highest return $r_{t+1} = -m_{t+1}$

- Risk-free rate

$$
r^1_{t+1} = -\log E_t(M_{t+1}) = -L_t(M_{t+1}) - E_t m_{t+1}
$$

- Subtract from above:

$$
L_t(M_{t+1}) \ge E_t(r_{t+1} - r^1_{t+1})
$$

Example: Macro disasters

- Consumption growth iid

$$
\begin{aligned}
g_{t+1} &= w_{t+1} + z_{t+1} \\
w_{t+1} &\sim \mathcal{N}(\mu, \sigma^2) \\
z_{t+1}|j &\sim \mathcal{N}(j\theta, j\delta^2) \\
j &\geq 0 \text{ has probability } e^{-\omega}\omega^j/j!
\end{aligned}
$$

- Parameter values
  - Match mean and variance of log consumption growth
  - Average number of disasters ($\omega = 0.01$), mean ($\theta = -0.3$) and variance ($\delta^2 = 0.15^2$)
  - Similar to Barro (2006)

Macro disasters: Deviations from normality

- Pricing kernel

$$\begin{aligned} m_{t+1} &= \log \beta - (1-\alpha) g_{t+1} \\ L(M) &= \log E e^{m} - E m = k(g, -(1-\alpha)) - (1-\alpha)\kappa_1(g) \end{aligned}$$

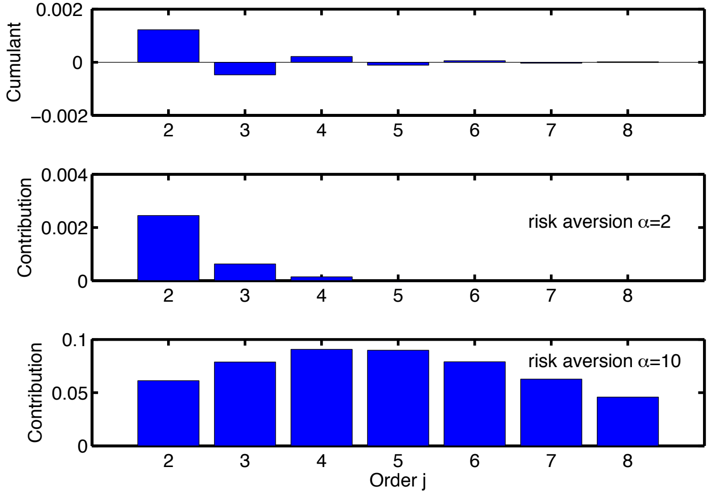
- Yaron's "bazooka"

$$\kappa_j(m)/j! \quad = \quad \kappa_j(g)(-(1-\alpha))^j/j!$$

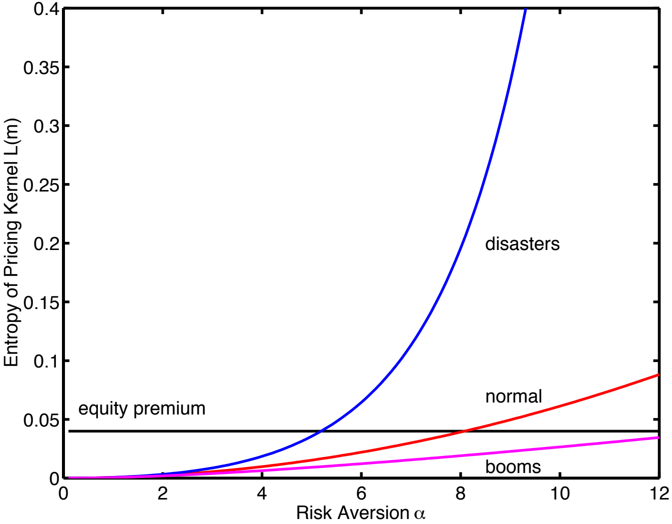
- The contribution of higher-order cumulants peaks at $j = \alpha - 1$

$$\frac{(\alpha-1)^j}{j!} = \frac{\alpha-1}{1} \cdot \frac{\alpha-1}{2} \cdot \ldots \cdot \frac{\alpha-1}{j-1} \cdot \frac{\alpha-1}{j}$$

Macro disasters: Cumulants



Macro disasters: Entropy



Predictability

- In basic terms, can one find an x_t such that

$$r_{t+1} = \gamma_0 + \gamma_x x_t + \varepsilon_{t+1}$$

with $\gamma_x \neq 0$?

- The significance of this for risk premiums is

$$E_t \tilde{rx}_{t+1} \propto x_t$$

- Before discussing what x_t is and existing evidence, let's think how this can be captured by a model

Predictability and risk premiums

- We must have that $cov_t(m_{t+1}, \tilde{rx}_{t+1}) \propto x_t$
- Assume that

$$\begin{aligned} m_{t+1} &= \mu_{mt} + \sigma_{mt}\varepsilon_{t+1} \\ \tilde{rx}_{t+1} &= \mu_{rt} + \sigma_{rt}\varepsilon_{t+1} \end{aligned}$$

- Then $cov_t(m_{t+1}, \tilde{rx}_{t+1}) = \sigma_{mt}\sigma_{rt}$
- If $\sigma^2_{jt} \propto x_t$ ($j = m, r$), then we get the predictability result
- Predictability regressions give us clues about the drivers of the pricing kernel

Evidence on predictability

- The profession has engaged in massive data-snooping exercise using past returns, interest rates, default spreads, and financial ratios (dividend yield, book-to-market, and earnings-price ratio)
- The financial ratios turn out to be the most promising ones, after all they all have the current price in the denominator
- Then the profession has engaged in a statistical analysis of the predictability evidence (OLS of a persistent variable on a persistent variable, effect of Bayesian priors, specifics of the null hypothesis).
- One of the more recent iterations is Lewellen (2004) who concludes that the evidence is strong in favor of predictability.

Summary of the evidence

Regression	b	t	$R^2(\%)$	$\sigma(bx)(\%)$
$R_{t+1} = a + b(D_t/P_t) + \varepsilon_{t+1}$	3.39	2.28	5.8	4.9
$R_{t+1} - R_t^f = a + b(D_t/P_t) + \varepsilon_{t+1}$	3.83	2.61	7.4	5.6
$D_{t+1}/D_t = a + b(D_t/P_t) + \varepsilon_{t+1}$	0.07	0.06	0.0001	0.001
$r_{t+1} = a_r + b_r(d_t - p_t) + \varepsilon_{t+1}^r$	0.097	1.92	4.0	4.0
$\Delta d_{t+1} = a_d + b_d(d_t - p_t) + \varepsilon_{t+1}^{dp}$	0.008	0.18	0.00	0.003

R_{t+1} is the real return, deflated by the CPI, D_{t+1}/D_t is real dividend growth, and D_t/P_t is the dividend-price ratio of the CRSP value-weighted portfolio. R_{t+1}^f is the real return on 3-month Treasury-Bills. Small letters are logs of corresponding capital letters. Annual data, 1926–2004. $\sigma(bx)$ gives the standard deviation of the fitted value of the regression.

Source: Cochrane (2007)

Summary of the evidence

- The statistical evidence is weak
- The economic evidence is very strong, as Cochrane (2007) argues
- Why? Because dividend growth is not forecastable – “the dog that did not bark”

Detour: Campbell-Shiller linearization

- Definition of one-period gross return to stock market

$$R_{t+1} = \frac{P_{t+1} + D_{t+1}}{P_t} = \frac{(1 + P_{t+1}/D_{t+1})D_{t+1}/D_t}{P_t/D_t}.$$

- Taking logs and linearizing around $E(p_t - d_t) = \log P/D$

$$\begin{aligned} r_{t+1} &= \log(1 + e^{p_{t+1} - d_{t+1}}) + \Delta d_{t+1} - (p_t - d_t) \\ &\approx k + \rho(p_{t+1} - d_{t+1}) + \Delta d_{t+1} - (p_t - d_t), \end{aligned}$$

  where $\rho = P/D(1 + P/D)^{-1} \approx 0.96$ in the data.

- Subsequently, ignore constant terms, that is, work with de-meaned logs of variables

(Constrained) VAR dynamics

- Joint behaviour of the variables of interest r_t, Δd_t, and $d_t - p_t$:

$$
\begin{aligned}
r_{t+1} &= a_r + b_r(d_t - p_t) + \varepsilon^r_{t+1} \\
\Delta d_{t+1} &= a_d + b_d(d_t - p_t) + \varepsilon^d_{t+1} \\
d_{t+1} - p_{t+1} &= a_{dp} + \phi(d_t - p_t) + \varepsilon^{dp}_{t+1}
\end{aligned}
$$

- The CS identity

$$r_{t+1} = \rho(p_{t+1} - d_{t+1}) + \Delta d_{t+1} - (p_t - d_t)$$

connects the three equations:

$$
\begin{aligned}
b_r &= 1 - \rho\phi + b_d, \\
\varepsilon^r_{t+1} &= \varepsilon^d_{t+1} - \rho\varepsilon^{dp}_{t+1}
\end{aligned}
$$

- Thus, one of the three equations is redundant

Implication for predictability tests

- Because $b_r = 1 - \rho\phi + b_d$, the null $b_r = 0$ implies something for the null on b_d and ϕ

- Specifically, if $\phi < 1/\rho \approx 1.04$, then b_r and b_d cannot be zero simultaneously

- To set a proper null, Cochrane assumes $b_r = 0$ and uses estimated ϕ

Estimated VAR

	Estimates			ε s. d. (diagonal) and correlation.			Null 1	Null 2
	$\hat{b}, \hat{\phi}$	$\sigma(\hat{b})$	implied	r	Δd	dp	b, ϕ	b, ϕ
r	0.097	0.050	0.101	19.6	66	−70	0	0
Δd	0.008	0.044	0.004	66	14.0	7.5	−0.0931	−0.046
dp	0.941	0.047	0.945	−70	7.5	15.3	0.941	0.99

Each row represents an OLS forecasting regression on the log dividend yield in annual CRSP data 1927–2004. For example, the first row presents the regression $r_{t+1} = a_r + b_r(d_t - p_t) + \varepsilon^r_{t+1}$. Standard errors $\sigma(\hat{b})$ include a GMM correction for heteroskedasticity. The "implied" column calculates each coefficient based on the other two coefficients and the identity $b_r = 1 - \rho\phi + b_d$, using $\rho = 0.9638$. The diagonals of the "ε s. d." matrix give the standard deviation of the regression errors in percent; the off-diagonals give the correlation between errors in percent. The "Null" columns describes coefficients used to simulate data under the null hypothesis that returns are not predictable.

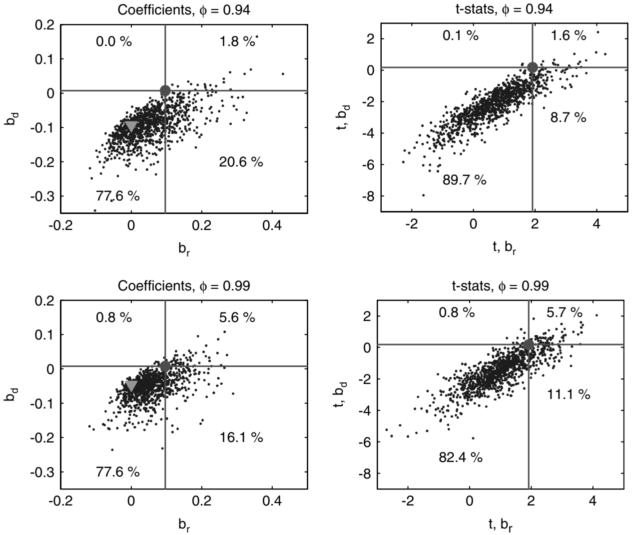
Source: Cochrane (2007)

Simulated (null) VAR dynamics

- For inference, he simulates Δd_{t+1} and $p_t - d_t$ from the null model (r_{t+1} is implied by CS)

$$\begin{aligned} d_{t+1} - p_{t+1} &= \phi(d_t - p_t) + \varepsilon^{dp}_{t+1} \\ \Delta d_{t+1} &= (\rho\phi - 1)(d_t - p_t) + \varepsilon^{d}_{t+1} \\ r_{t+1} &= \varepsilon^{d}_{t+1} - \rho\varepsilon^{dp}_{t+1} \end{aligned}$$

Joint test



Source: Cochrane (2007)

Detour: Campbell-Shiller decomposition

- Apply the no-bubble condition ($\lim_{j\to\infty} \rho^j (p_{t+j} - d_{t+j}) = 0$) and iterate the CS identity forward:

$$
\begin{aligned}
p_t - d_t &= \rho(p_{t+1} - d_{t+1}) + \Delta d_{t+1} - r_{t+1} \\
&= \rho^2(p_{t+2} - d_{t+2}) + \rho(\Delta d_{t+2} - r_{t+2}) + \Delta d_{t+1} - r_{t+1} \\
&= \sum_{j=1}^{\infty} \rho^{j-1} \Delta d_{t+j} - \sum_{j=1}^{\infty} \rho^{j-1} r_{t+j}
\end{aligned}
$$

- This implies

$$
\begin{aligned}
d_t - p_t &= E_t \sum_{j=1}^{\infty} \rho^{j-1} r_{t+j} - E_t \sum_{j=1}^{\infty} \rho^{j-1} \Delta d_{t+j}, \\
var(d_t - p_t) &= cov(\sum_{j=1}^{\infty} \rho^{j-1} r_{t+j}, d_t - p_t) - cov(\sum_{j=1}^{\infty} \rho^{j-1} \Delta d_{t+j}, d_t - p_t)
\end{aligned}
$$

Long-run implications of the tests

- The CS decomposition leads to a relationship between regression betas, if divided by $var(d_t - p_t)$
- Denote a beta form regressing y on x by $\beta(y, x)$
- Then

$$\beta(\sum_{j=1}^{\infty} \rho^{j-1} r_{t+j}, d_t - p_t) - \beta(\sum_{j=1}^{\infty} \rho^{j-1} \Delta d_{t+j}, d_t - p_t) = 1$$

- Computing "long-run" betas:

$$\begin{aligned} \beta(\sum_{j=1}^{\infty} \rho^{j-1} r_{t+j}, d_t - p_t) &= \sum_{j=1}^{\infty} \rho^{j-1} \beta(r_{t+j}, d_t - p_t) \\ &= \sum_{j=1}^{\infty} \rho^{j-1} \phi^{j-1} b_r = b_r (1 - \rho\phi)^{-1} \equiv b_r^{lr} \end{aligned}$$

- The second one is, similarly, $b_d (1 - \rho\phi)^{-1} \equiv b_d^{lr}$

Long-run betas

Variable	$\hat{b}^{lr}$	s. e.	t	% p value
r	1.09	0.44	2.48	1.39–1.83
Δd	0.09	0.44	2.48	1.39–1.83
Excess r	1.23	0.47	2.62	0.47–0.69

The long-run return forecast coefficient $\hat{b}_r^{lr}$ is computed as $\hat{b}_r^{lr} = \hat{b}_r/(1 - \rho\hat{\phi})$, where $\hat{b}_r$ is the regression coefficient of one-year returns r_{t+1} on $d_t - p_t$, $\hat{\phi}$ is the autocorrelation of $d_t - p_t$, $\rho = 0.961$, and similarly for the long-run dividend-growth forecast coefficient $\hat{b}_d^{lr}$. The standard error is calculated from standard errors for $\hat{b}_r$ and $\hat{\phi}$ by the delta method. The t-statistic for Δd is the statistic for the hypothesis $\hat{b}_d^{lr} = -1$. Percent probability values (% p value) are generated by Monte Carlo under the $\phi = 0.941$ null. The range of probability values is given over the three choices of which coefficient ($\hat{b}_r$, $\hat{\phi}$, $\hat{b}_d$) is implied from the other two.

Source: Cochrane (2007)

Predictability and valuation

- Is this exercise consistent with an asset-pricing framework?

- Price-dividend ratio:

$$
\begin{aligned}
P_t &= E_t\left[M_{t+1}(P_{t+1}+D_{t+1})\right] \Leftrightarrow \\
P_t/D_t &= E_t\left[M_{t+1}D_{t+1}/D_t(1+P_{t+1}/D_{t+1})\right]
\end{aligned}
$$

- With the same approximation as before, we have

$$\log\left[M_{t+1}D_{t+1}/D_t(1+P_{t+1}/D_{t+1})\right] = m_{t+1} + \Delta d_{t+1} + \rho(p_{t+1} - d_{t+1})$$

- Need to assume *m* and *d* to progress further

Vasicek example

- Suppose m and d have the loglinear form

$$\begin{aligned} m_{t+1} &= x_t + \lambda w_{t+1} \\ \Delta d_{t+1} &= \delta x_t + \gamma w_{t+1} \\ x_{t+1} &= \varphi x_t + \sigma w_{t+1} \end{aligned}$$

- Guess $p_t - d_t = A + Bx_t$ for coefficients to be determined.
- The (approximate) pricing relation gives us

$$\begin{aligned} p_t - d_t &= \log E_t e^{m_{t+1} + \Delta d_{t+1} + \rho A + \rho B x_{t+1}} \\ &= (1 + \delta + \rho B \varphi) x_t + (\lambda + \gamma + \rho B \sigma)^2/2 + \rho A \\ &= A + Bx_t. \end{aligned}$$

- That gives us

$$\begin{aligned} B &= 1 + \delta + \rho B \varphi \;=\; (1+\delta)/(1 - \varphi\rho) \\ A &= (\lambda + \gamma + \rho B \sigma)^2/2 + \rho A \end{aligned}$$

Implications

- Here, $p_t - d_t$ inherits the dynamics of x_t, so consistent with Cochrane's AR(1) specification
- No predictability in Δd_{t+1} if $\delta = 0$
- The (appproximate) log return is

$$
\begin{aligned}
r_{t+1} &= \Delta d_{t+1} - (p_t - d_t) + \rho(p_{t+1} - d_{t+1}) \\
&= (\rho - 1)A + [\delta - B(1 - \varphi\rho)]x_t + (\gamma + \rho B\sigma)w_{t+1} \\
&= -(\lambda + \gamma + \rho B\sigma)^2/2 - x_t + (\gamma + \rho B\sigma)w_{t+1} \\
&= -(\lambda + \gamma + \rho B\sigma)^2/2 + A/B - B^{-1}(p_t - d_t) + (\gamma + \rho B\sigma)w_{t+1}
\end{aligned}
$$

- The x will drop out if we use excess returns
- Loading on $p_t - d_t$ depends on δ

5. Resolving AP puzzles with recursive preferences

Recursive preferences

- Recursive utility is an asset-pricing-friendly way to introduce nonseparabilities across states of nature
- Separates the coefficient of relative risk aversion from inverse of the elasticity of intertemporal substitution
- Equations (Kreps and Porteus, 1978; Epstein and Zin, 1989; Weil, 1989)

$$
\begin{aligned}
U_t &= \left[(1-\beta)C_t^{\rho} + \beta\mu_t(U_{t+1})^{\rho}\right]^{1/\rho} \\
\mu_t(U_{t+1}) &= \left(E_t U_{t+1}^{\alpha}\right)^{1/\alpha} \\
EIS &= \psi = 1/(1-\rho) \\
CRRA &= \gamma = 1-\alpha \\
\alpha &= \rho \quad \Rightarrow \quad \text{additive preferences}
\end{aligned}
$$

A CES utility function

- Define

$$U = F(c_1, c_2) = \left[(1-\beta)c_1^{\rho} + \beta c_2^{\rho}\right]^{1/\rho}$$

- Marginal rate of substitution

$$MRS = \frac{\partial U/\partial c_1}{\partial U/\partial c_2} = \frac{1-\beta}{\beta}\left(\frac{c_1}{c_2}\right)^{\rho-1}$$

- Elasticity of substitution

$$ES = \frac{\partial \log(c_2/c_1)}{\partial \log MRS} = -\left(\frac{\partial \log MRS}{\partial \log(c_1/c_2)}\right)^{-1} = 1/(1-\rho)$$

CES across time (non-stochastic)

- Lifetime utility at t depends on current consumption and lifetime utility at $t+1$

$$U_t = F(C_t, U_{t+1})$$

- Recursively substitute to show CES between consumption at different dates:

$$\begin{aligned}
U_t^\rho &= (1-\beta)C_t^\rho + \beta U_{t+1}^\rho && \text{(make linear)} \\
U_t^\rho &= (1-\beta)C_t^\rho + \beta(1-\beta)C_{t+1}^\rho + \beta^2(1-\beta)C_{t+2}^\rho + \ldots \\
& && \text{(substitute)} \\
U_t &= (1-\beta)^{1/\rho}(C_t^\rho + \beta C_{t+1}^\rho + \beta^2 C_{t+2}^\rho + \ldots)^{1/\rho} && \text{(rewrite)}
\end{aligned}$$

- EIS is $1/(1-\rho)$

CES across time (stochastic)

- Utilize risk aversion by applying CRRA to future lifetime utility, not to each of C_{t+i} separately
- That's why we use certainty equivalent $\left(E_t U_{t+1}^{\alpha}\right)^{1/\alpha}$. It answers the question: which non-stochastic $\bar{U}$ provides the same time t lifetime utility as stochastic U_{t+1}?
- The recursive form of lifetime utility is

$$U_t = F(C_t, \mu_t(U_{t+1}))$$

A special case

- In general, no-closed form expression for U_t
- Consider the case when $\alpha = \rho$:

$$U_t = \left[(1-\beta)C_t^{\rho} + \beta E_t U_{t+1}^{\rho}\right]^{1/\rho}$$

- Put it in linear form

$$U_t^{\rho} = (1-\beta)C_t^{\rho} + \beta E_t U_{t+1}^{\rho}$$

- Recursively substitute

$$U_t^{\rho} = (1-\beta)\sum_{i=0}^{\infty}\beta^i E_t C_{t+i}^{\rho}$$

- We get CRRA

The pricing kernel

- Scale problem by C_t ($u_t = U_t / C_t$, $G_{t+1} = C_{t+1}/C_t$)

$$u_t \quad = \quad [(1-\beta) + \beta \mu_t (G_{t+1} u_{t+1})^\rho]^{1/\rho}$$

- Pricing kernel (mrs)

$$\begin{aligned} M_{t+1} \quad &= \quad \beta \left(\frac{C_{t+1}}{C_t} \right)^{\rho - 1} \left(\frac{U_{t+1}}{\mu_t (U_{t+1})} \right)^{\alpha - \rho} \\ &= \quad \beta \underbrace{G_{t+1}^{\rho-1}}_{\text{short-run risk}} \underbrace{\left(\frac{G_{t+1} u_{t+1}}{\mu_t (G_{t+1} u_{t+1})} \right)^{\alpha - \rho}}_{\text{long-run risk}} \end{aligned}$$

- Note the role of recursive preferences: if $\alpha = \rho$
  - Second term disappears
  - No role for predictable consumption growth or volatility (later)

An "observable" pricing kernel

- The pricing kernel is unobservable because U_{t+1} is unobserved
- Epstein and Zin (1989) show how to make it observable
- Wealth of a rep agent is the value of discounted future aggregate consumption

$$W_t = E_t(M_{t+1}(C_{t+1} + W_{t+1})) \tag{3}$$

- Next, we want to show that

$$U_t = (W_t + C_t)\frac{\partial F(C_t, \mu_t(U_{t+1}))}{\partial C_t}$$

- Use the above as a guess and obtain the l.h.s. of (3):

$$\begin{aligned} W_t &= U_t\left(\frac{\partial F(C_t, \mu_t(U_{t+1}))}{\partial C_t}\right)^{-1} - C_t \\ &= U_t((1-\beta)C_t^{\rho-1}F(C_t, \mu_t(U_{t+1}))^{1-\rho})^{-1} - C_t \\ &= \frac{C_t^{1-\rho}}{1-\beta}(U_t^{\rho} - (1-\beta)C_t^{\rho}) = \frac{C_t^{1-\rho}}{1-\beta}\beta\mu_t(U_{t+1})^{\rho} \end{aligned}$$

An "observable" pricing kernel (cont'd)

- Use the guess and obtain the r.h.s. of (3):

$$
\begin{aligned}
& E_t(M_{t+1}(C_{t+1}+W_{t+1})) \\
= \quad & E_t\left(\beta\left(\frac{C_{t+1}}{C_t}\right)^{\rho-1}\left(\frac{U_{t+1}}{\mu_t(U_{t+1})}\right)^{\alpha-\rho} U_{t+1}\left(\frac{\partial F(C_{t+1},\mu_{t+1}(U_{t+2}))}{\partial C_{t+1}}\right)^{-1}\right) \\
= \quad & E_t\left(\beta\left(\frac{C_{t+1}}{C_t}\right)^{\rho-1}\left(\frac{U_{t+1}}{\mu_t(U_{t+1})}\right)^{\alpha-\rho} U_{t+1}((1-\beta)C_{t+1}^{\rho-1}U_{t+1}^{1-\rho})^{-1}\right) \\
= \quad & \frac{C_t^{1-\rho}}{1-\beta}E_t\left(\beta U_{t+1}^{\alpha}\mu_t(U_{t+1})^{\rho-\alpha}\right) = \frac{C_t^{1-\rho}}{1-\beta}\beta\mu_t(U_{t+1})^{\rho-\alpha}E_t U_{t+1}^{\alpha}
\end{aligned}
$$

- Do the l.h.s. and r.h.s. match?

$$
\begin{aligned}
\frac{C_t^{1-\rho}}{1-\beta}\beta\mu_t(U_{t+1})^{\rho} \quad &= \quad \frac{C_t^{1-\rho}}{1-\beta}\beta\mu_t(U_{t+1})^{\rho-\alpha}E_t U_{t+1}^{\alpha} \\
\mu_t(U_{t+1})^{\alpha} \quad &= \quad E_t U_{t+1}^{\alpha}
\end{aligned}
$$

An “observable” pricing kernel (cont’d)

- Return to total wealth

$$R_{w,t+1} = \frac{C_{t+1} + W_{t+1}}{W_t}$$

- Plug this into the derived utility:

$$\begin{aligned}
R_{w,t+1} &= U_{t+1}\left(\frac{\partial F(C_{t+1},\mu_{t+1}(U_{t+2}))}{\partial C_{t+1}}\right)^{-1}\left(U_t\left(\frac{\partial F(C_t,\mu_t(U_{t+1}))}{\partial C_t}\right)^{-1} - C_t\right)^{-1} \\
&= \frac{U_{t+1}}{(1-\beta)C_{t+1}^{\rho-1}U_{t+1}^{1-\rho}}\,\frac{(1-\beta)c_t^{\rho-1}U_t^{1-\rho}}{U_t - C_t(1-\beta)c_t^{\rho-1}U_t^{1-\rho}} \\
&= \left(\frac{C_{t+1}}{C_t}\right)^{1-\rho} U_{t+1}^{\rho}\left(U_t^{\rho} - (1-\beta)C_t^{\rho}\right)^{-1} \\
&= \left(\frac{C_{t+1}}{C_t}\right)^{1-\rho}\frac{U_{t+1}^{\rho}}{\beta\mu_t(U_{t+1})^{\rho}} = \beta^{-1}\left(\frac{C_{t+1}}{C_t}\right)^{1-\rho}\left(\frac{U_{t+1}}{\mu_t(U_{t+1})}\right)^{\rho}
\end{aligned}$$

- Substitute this into the pricing kernel

$$M_{t+1} = \beta^{\alpha/\rho}\left(\frac{C_{t+1}}{C_t}\right)^{(\rho-1)\alpha/\rho} R_{w,t+1}^{(\alpha-\rho)/\rho}$$

A resolution of the AP puzzles?

- Another source of risk premium: not just covariance with contemporaneous consumption growth, but covariance with return to total wealth
- Under log-normality the risk premium is:

$$
\begin{aligned}
& E_t \tilde{rx}_{t+1} = -cov_t(m_{t+1}, \tilde{rx}_{i,t+1}) \\
= \quad & \alpha/\rho(1-\rho) cov_t(g_{t+1}, r_{i,t+1}) + (1-\alpha/\rho) cov_t(r_{w,t+1}, r_{i,t+1})
\end{aligned}
$$

  Recall that $\alpha = \rho$ recovers the CRRA case
- The risk-free rate is:

$$
\begin{aligned}
& r_{f,t} = -E_t m_{t+1} - var_t(m_{t+1})/2 \\
= \quad & \alpha/\rho(-\log\beta + (1-\rho) E_t g_{t+1}) + (1-\alpha/\rho) E_t r_{w,t+1} \\
- \quad & [\alpha/\rho(1-\rho)]^2 var_t(g_{t+1})/2 - (1-\alpha/\rho)^2 var_t(r_{w,t+1})/2 \\
- \quad & \alpha/\rho(1-\alpha/\rho)(1-\rho) cov_t(g_{t+1}, r_{w,t+1})
\end{aligned}
$$

Manipulating the risk-free rate

- Put expected excess return to W on the r.h.s.

$$\begin{aligned} & r_{f,t} = \alpha/\rho(-\log\beta + (1-\rho)E_t g_{t+1}) \\ + \quad & (1-\alpha/\rho)(E_t r_{w,t+1} - r_{f,t} + var_t(r_{w,t+1})/2) \\ + \quad & (1-\alpha/\rho) r_{f,t} - (1-\alpha/\rho) var_t(r_{w,t+1})/2 \\ - \quad & [\alpha/\rho(1-\rho)]^2 var_t(g_{t+1})/2 - (1-\alpha/\rho)^2 var_t(r_{w,t+1})/2 \\ - \quad & \alpha/\rho(1-\alpha/\rho)(1-\rho) cov_t(g_{t+1}, r_{w,t+1}) \end{aligned}$$

- "Solve" for $r_{f,t}$

$$\begin{aligned} r_{f,t} \quad = \quad & -\log\beta + (1-\rho)E_t g_{t+1} + (\rho/\alpha - 1)E_t(\tilde{rx}_{w,t+1}) \\ - \quad & \alpha/\rho(1-\rho)^2 var_t(g_{t+1})/2 - (\rho/\alpha - 1)(2-\alpha/\rho) var_t(r_{w,t+1})/2 \\ - \quad & (1-\alpha/\rho)(1-\rho) cov_t(g_{t+1}, r_{w,t+1}) \end{aligned}$$

Manipulating the risk-free rate (cont'd)

- Expected excess return to W :

$$E_t(\tilde{rx}_{w,t+1}) = \alpha/\rho(1-\rho)cov_t(g_{t+1}, r_{w,t+1}) + (1-\alpha/\rho)var_t(r_{w,t+1})$$

- Plug this into the $r_{f,t}$ equation and simplify (the covariance term cancels out)

$$\begin{aligned} r_{f,t} &= -\log\beta + (1-\rho)E_t g_{t+1} \\ &- (1-\alpha/\rho)var_t(r_{w,t+1})/2 - \alpha/\rho(1-\rho)^2 var_t(g_{t+1})/2 \end{aligned}$$

- As is the case with risk premiums, we gain an extra term that allows to lower the risk-free rate

Assesement

- The issue is that $r_{w,t+1}$ is endogenous, that is, we are not free to choose its mean and variance
- John Campbell proposes a procedure for endogenizing $r_{w,t+1}$ (see Campbell, 1999 for an excellent survey)
- However, the procedure is lengthy; works for the log-normal case only (constant volatility is required)
- The alternative is to avoid the “observable” representation of the PK, assume a specific form of g_{t+1} and solve for U_t and, therefore, M_t (e.g., Backus, Chernov, and Zin, 2014)

6. Solving Bellmans

Log-linear approximation

- Start with the scaled utility

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

- A loglinear approximation is

$$\begin{aligned} \log u_t &= \rho^{-1}\log\left[(1-\beta) + \beta\mu_t(G_{t+1}u_{t+1})^{\rho}\right] \\ &= \rho^{-1}\log\left[(1-\beta) + \beta e^{\rho\log\mu_t(G_{t+1}u_{t+1})}\right] \\ &\approx b_0 + b_1\log\mu_t(G_{t+1}u_{t+1}). \end{aligned} \tag{4}$$

- 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 $b_1 = \beta$ (verify with the L'Hopital rule).
- Assume g_t, then use guess and verify to solve u_t

Example: MA consumption growth

- Start with

$$g_t = g + \sum_{j=0}^{\infty} \gamma_j v^{1/2} w_{t-j} \;=\; g + \gamma(B) v^{1/2} w_t,$$

where $\{w_t\} \sim \text{NID}(0, 1)$.

- A popular case: $\gamma_0 = 1$, $\gamma_{j+1} = \varphi_g \gamma_j = \varphi_g^j \gamma_1$, $j \geq 1$

$$g_t = g(1 - \varphi_g) + \varphi_g g_{t-1} + v^{1/2} w_t + (\gamma_1 - \varphi_g) v^{1/2} w_{t-1}$$

- In state-space form this can be expressed as

$$\begin{aligned} g_t &= g + x_{t-1} + v^{1/2} w_t \\ x_t &= \varphi_g x_{t-1} + \gamma_1 v^{1/2} w_t \end{aligned}$$

- In Bansal and Yaron (2004) the two shocks are independent from each other

The value function

- Guess a loglinear value function:

$$\log u_t = u + \sum_{j=0}^{\infty} \omega_j v^{1/2} w_{t-j}$$

- Given the guess, compute the log of $G_{t+1} u_{t+1}$ and $\mu_t(G_{t+1} u_{t+1})$:

$$\begin{aligned}
\log(G_{t+1} u_{t+1}) &= g + u + (\gamma_0 + \omega_0) v^{1/2} w_{t+1} + \sum_{j=0}^{\infty} (\gamma_{j+1} + \omega_{j+1}) v^{1/2} w_{t-j} \\
\log \mu_t(G_{t+1} u_{t+1}) &= g + u + \sum_{j=0}^{\infty} (\gamma_{j+1} + \omega_{j+1}) v^{1/2} w_{t-j} + \alpha(\gamma_0 + \omega_0)^2 v/2.
\end{aligned}$$

The Bellman equation solution

- If we substitute the parameters into (4) and collect terms, we get

$$\begin{aligned} \text{constant}: \quad u &= b_0 + b_1 [g + u] \\ w_{1t-j}: \quad \omega_j &= b_1(\gamma_{j+1} + \omega_{j+1}). \end{aligned}$$

- 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. Iterating forward, we have (for each $j \geq 0$)

$$\begin{aligned} \omega_j &= \sum_{i=1}^{\infty} b_1^i \gamma_{j+i} \equiv \Gamma_{j+1} \\ \gamma_j + \omega_j &= \gamma_j + \Gamma_{j+1} = \sum_{i=0}^{\infty} b_1^i \gamma_{j+i} = \Gamma_j / b_1. \end{aligned}$$

- Γ_js are geometric sums reflecting the impact of innovations to current consumption growth on future utility. They summarize the "predictable component" in the sense that if g_t is iid, $\Gamma_j = 0$ for $j \geq 1$.
- We denote $\gamma_0 + \omega_0 = \gamma_0 + \Gamma_1 = \gamma(b_1)$.

The pricing kernel

- One component is

$$\log(G_{t+1}u_{t+1}) - \log\mu_t(G_{t+1}u_{t+1}) \quad = \quad \gamma(b_1)v^{1/2}w_{t+1} - (\alpha/2)\gamma(b_1)^2 v$$

- We get the pricing kernel

$$\begin{aligned} m_{t+1} \quad = \quad & \log\beta + (\rho-1)g - (\alpha-\rho)(\alpha/2)\gamma(b_1)^2 v \\ & + \underbrace{[(\rho-1)\gamma_0 + (\alpha-\rho)\gamma(b_1)]v^{1/2}}_{a_0} w_{t+1} \\ & + \sum_{j=0}^{\infty} \underbrace{(\rho-1)\gamma_{j+1}v^{1/2}}_{a_j} w_{1t-j} \end{aligned}$$

- In the iid case, $\gamma(b_1) = \gamma_0$. Otherwise, there's an additional role for the "persistent component."

- More examples in Backus, Chernov, and Zin (2014)

7. Assessing models with recursive preferences

Recursive preferences: consumption

- Consumption growth

$$\begin{aligned} g_t &= g + \gamma(B) v^{1/2} w_t \\ \{w_t\} &\sim \text{NID}(0,1) \end{aligned}$$

- Pricing kernel

$$\begin{aligned} m_{t+1} &= \text{constants} \\ &+ \underbrace{[(\rho - 1)\gamma_0 + (\alpha - \rho)\gamma(b_1)] v^{1/2}}_{a_0} w_{t+1} \\ &+ a_1 w_t + a_2 w_{t-1} + \cdots \end{aligned}$$

- Entropy

$$EL_t(M_{t+1}) = [(\rho - 1)\gamma_0 + (\alpha - \rho)\gamma(b_1)]^2 v/2 = 0.0214 \ (0.0049 \text{ PU})$$

- Critical term: $\gamma(b_1) = \gamma_0 + b_1\gamma_1 + b_1^2\gamma_2 + \cdots$ (no effect if iid or PU)
- No return predictability!

Recursive preferences: ... and volatility

- Consumption growth

$$\begin{aligned} g_t &= g + \gamma(B) v_{t-1}^{1/2} w_{gt}, \; \gamma_{j+1} = \varphi_g \gamma_j, \; j \geq 1 \\ v_t &= v + \nu(B) w_{vt}, \; \nu_{j+1} = \varphi_v \nu_j, \; j \geq 0 \\ \{w_{gt}, w_{vt}\} &\sim \text{NID}(0, I) \end{aligned}$$

- Pricing kernel

$$m_{t+1} = \log m + \underbrace{a_g(B)(v_t/v)^{1/2} w_{gt+1}}_{\text{non-linearity}} + a_v(B) w_{vt+1}$$

- Entropy

$$\begin{aligned} EL_t(M_{t+1}) &= [(\rho - 1)\gamma_0 + (\alpha - \rho)\gamma(b_1)]^2 v/2 \\ &+ (\alpha - \rho)^2 ((\alpha/2)\gamma(b_1)^2 b_1 \nu(b_1))^2 / 2 = 0.0218 \end{aligned}$$

- Predictability of returns can be driven by v_t only, so $d_t - p_t$ must be a function of v_t

Disasters

- Consumption growth

$$\begin{aligned} g_t &= g + v^{1/2} w_{gt} + z_t \\ h_t &= h + \eta(B) w_{ht} \\ (w_{gt}, w_{ht}) &\sim \text{NID}(0, I) \\ z_t | j &\sim \mathcal{N}(j\theta, j\delta^2) \\ j &\geq 0 \text{ has jump intensity } h_t \end{aligned}$$

- Entropy

$$\begin{aligned} EL_t(M_{t+1}) &= (\alpha - 1)^2 v/2 \\ &+ (e^{(\alpha-1)\theta + (\alpha-1)^2\delta^2/2} - (\alpha-1)\theta - 1)h \\ &+ (\alpha - \rho)^2 [b_1 \eta(b_1)(e^{\alpha\theta + \alpha^2\delta^2/2} - 1)/\alpha]^2/2 \\ &= 0.0512 \text{ (0.0485 if IID)} \end{aligned}$$

Numerical summary

	Power Utility	Recursive Utility
Preferences		
ρ	-9	1/3
α	-9	-9
β	0.9980	0.9980
Consumption		
g	0.0015	0.0015
γ_0	1	1
γ_1	0.0271	0.0271
φ_g	0.9790	0.9790
$v^{1/2}$	0.0099	0.0099
Derived		
b_1		0.9978
$\gamma(b_1)$		2.165
$\gamma(1)$		2.290
$EL_t(m_{t,t+1})$	0.0049	0.0214

	Recursive Utility with SV
Preferences	
ρ	$1/3$
α	-9
β	0.9980
Consumption	
γ_0	1
γ_1	0.0271
φ_g	0.9790
$v^{1/2}$	0.0099
ν_0	0.2×10^{-5}
φ_v	0.9870
Derived	
b_1	0.9977
$\gamma(b_1)$	2.164
$\nu(b_1)$	0.0002
$EL_t(m_{t,t+1})$	0.0218

	IID w/ Jumps	Stochastic Intensity
Preferences		
ρ	$1/3$	1/3
α	-9	-9
β	0.9980	0.9980
Consumption		
$v^{1/2}$	0.0025	0.0025
h	0.0008	0.0008
θ	-0.3000	-0.3000
δ	0.1500	0.1500
η_0	0	0.0001
φ_h		0.9500
γ_0	1	1
ψ_0	1	1
Derived		
b_1	0.9974	0.9973
$\gamma(b_1)$	1	1
$\psi(b_1)$	1	1
$\eta(b_1)$	0	0.0016
$EL_t(m_{t,t+1})$	0.0485	0.0512

8. Resolving AP puzzles with habits

Habits

- Preferences

$$
\begin{aligned}
U_t &= f(C_t, X_t) + \beta E_t U_{t+1} \\
X_t &= \text{“external habit”}
\end{aligned}
$$

- Standard inputs

$$
\begin{aligned}
g_t &= g + \gamma(B) v^{1/2} w_t \\
x_t &= x + \chi(B) c_{t-1} \quad [\chi(1) = 1]
\end{aligned}
$$

- Examples
  - Ratio habit: C_t / X_t
  - Difference habit: $C_t - X_t$
  - Campbell and Cochrane (1999): P2C2E

Ratio habit

- Preferences

$$f(C_t, X_t) \quad = \quad (C_t/X_t)^{\rho}/\rho, \quad \rho \leq 1$$

- Pricing kernel

$$\begin{aligned} M_{t+1} &= \beta\,(C_{t+1}/C_t)^{\rho-1}\,(X_{t+1}/X_t)^{-\rho} \\ m_{t+1} &= \text{constants} \\ &\quad + [(\rho-1) - \rho B\chi(B)]\gamma(B)v^{1/2}w_{t+1} \end{aligned}$$

- Entropy is the same as that of PU
- No return predictability

Difference habit

- Preferences

$$f(C_t, X_t) \quad = \quad (C_t - X_t)^\rho/\rho, \quad \rho \leq 1$$

- Define surplus

$$S_t \quad = \quad (C_t - X_t)/C_t \quad = \quad 1 - X_t/C_t$$

- Pricing kernel

$$\begin{aligned} M_{t+1} \quad &= \quad \beta G_{t+1}^{\rho-1} (S_{t+1}/S_t)^{\rho-1} \\ m_{t+1} \quad &= \quad \text{constants} \\ &\quad + (\rho - 1)(1/s)[1 - (1-s)B\chi(B)]\gamma(B)v^{1/2}w_{t+1} \end{aligned}$$

- Entropy with $s = 1/2$

$$EL_t(M_{t+1}) = (\rho - 1)^2 \gamma_0^2 v/(2s^2) = 0.0197$$

- No return predictability

Difficulties with difference habit

- The pricing kernel is

$$m_{t+1} = \log\beta + (\rho - 1)g_{t+1} + (\rho - 1)(s_{t+1} - s_t)$$

- Under log-normality

$$\begin{aligned} r_{f,t} &= -\log\beta - (\rho - 1)E_t(g_{t+1} + s_{t+1} - s_t) \\ &- (\rho - 1)^2/2 \cdot (var_t g_{t+1} + var_t s_{t+1} + 2 \cdot covar_t(g_{t+1}, s_{t+1})) \end{aligned}$$

- Issue: If s_{t+1} is high, it is expected to revert to its average (if specified as mean-reverting process, which it usually is). Thus, there is substantial variation in the expected growth rate of s_t, which leads to large variation in the risk free rate (intertemporal substitution effect).

Campbell-Cochrane model

- Specifying the process for s_t, such that the just mentioned intertemporal substitution effect, $(1-\rho)E_t(g_{t+1}+s_{t+1}-s_t)$, is balanced by the *precautionary savings* effect, $(\rho-1)^2/2\cdot(var_t g_{t+1}+var_t s_{t+1}+2covar_t(g_{t+1},s_{t+1}))$.
- Assume that surplus follows

$$s_{t+1}-s_t = (\varphi_s-1)(s_t-s)+\lambda(s_t)v^{1/2}w_{t+1}$$

- Therefore

$$\begin{aligned} r_{f,t} &= -\log\beta-(\rho-1)(g+(\varphi_s-1)(s_t-s)) \\ &- (\rho-1)^2 v/2\cdot(1+\lambda(s_t))^2 \end{aligned}$$

- Select

$$1+\lambda(s_t)=v^{-1/2}\left(\frac{(1-\rho)(1-\varphi_s)-b}{(1-\rho)^2}\right)^{1/2}(1-2[s_t-s])^{1/2}$$

CC entropy

- The pricing kernel is

$$
\begin{aligned}
m_{t,t+1} &= \text{constant} + (\rho - 1)(\varphi_s - 1)(s_t - s) \\
&\quad + (\rho - 1)\left[1 + \lambda(s_t)\right] v^{1/2} w_{t+1} .
\end{aligned}
$$

- Entropy

$$
\begin{aligned}
L_t(M_{t+1}) &= (\rho - 1)^2 [1 + \lambda(s_t)]^2 v \\
&= ((\rho - 1)(\varphi_s - 1) - b)/2 + b(s_t - s)
\end{aligned}
$$

- If $b = 0$ (flat term structure)

$$
L_t(M_{t+1}) = EL_t(M_{t+1})
$$

- Return predictability is controlled by $\lambda(s_t)$

Habits: numerical examples

Parameter	Power Utility (1)	Ratio Habit (2)	Difference Habit (3)	Campbell-Cochrane (4)
Preferences				
ρ	-9	-9	-9	-1
Consumption growth				
$v^{1/2}$	0.0099	0.0099	0.0099	NA
Habit				
χ_0		0.25	0.25	
φ_x or φ_s		0.9000	0.9000	0.9885
s			0.5	
Entropy				
$EL_t(m_{t+1})$	0.0049	0.0049	0.0197	0.0230

Bibliography I

Backus, David, Mikhail Chernov, and Ian Martin, 2011, Disasters implied by equity index options, *Journal of Finance* 66, 1967–2010.

Backus, David, Mikhail Chernov, and Stanley Zin, 2014, Sources of entropy in representative agent models, *Journal of Finance* 69, 51–99.

Bansal, Ravi, and Amir Yaron, 2004, Risks for the long run: A potential resolution of asset pricing puzzles, *Journal of Finance* 59, 1481–1509.

Barro, Robert, 2006, Rare disasters and asset markets in the twentieth century, *Quarterly Journal of Economics* 121, 823867.

Campbell, John, 1999, Asset prices, consumption, and the business cycle, in J.B. Taylor, and M. Woodford, ed.: *Handbook of Macroeconomics, Volume I* (Elsvier).

Bibliography II

——— , and John Cochrane, 1999, By force of habit: a consumption-based explanation of aggregate stock market behavior, *Journal of Political Economy* 107, 205–251.

Chernov, Mikhail, and Philippe Mueller, 2012, The term structure of inflation expectations, *Journal of Financial Economics* 106, 367–394.

Cochrane, John, 2007, The dog that did not bark: A defense of return predictability, *Review of Financial Studies* 21, 1533–1575.

Epstein, Larry G., and Stanley E. Zin, 1989, Substitution, risk aversion, and the temporal behavior of consumption and asset returns: a theoretical framework, *Econometrica* 57, 937–969.

Hansen, Lars, and Ravi Jagannathan, 1991, Implications of security market data for models of dynamic economies, *Journal of Political Economy* 99, 225–262.

Bibliography III

Kreps, David M., and Evan L. Porteus, 1978, Temporal resolution of uncertainty and dynamic choice theory, *Econometrica* 46, 185–200.

Lewellen, Jonathan, 2004, Predicting returns with financial ratios, *Journal of Financial Economics* 74, 209–235.

Monfort, Alain, and Fulvio Pegoraro, 2012, Asset pricing with second-order esscher transforms, *The Journal of Banking and Finance* 36, 1678–1687.

Weil, Philippe, 1989, The equity premium puzzle and the risk-free rate puzzle, *Journal of Monetary Economics* 24, 4201–4219.