Asset Pricing VIII

Semyon Malamud

EPFL

Table of Contents

Readings:

- Back
- Cochrane

Topics:

- Excess Volatility puzzle (Shiller)
- Equity premium puzzle (Mehra-Prescott)
- Risk-free rate puzzle
- Campbell-Shiller decomposition
- Return vs. Dividend Predictability (Cochrane)
- Habit Formation (Campbell-Cochrane)
- Heterogeneous agents with idiosyncratic labor income shocks (Constantinides-Duffie)
- Recursive (Kreps-Porteus-Epstein-Zin) utility
- Long-Run Risk (Bansal-Yaron)

Excess Volatility Puzzle

- Absence of arbitrage implies there exists a stochastic discount factor M_t such that the price of any asset S_t is the present value of future dividends D_{t+n}:

$$M_t S_t = \mathbb{E}_t \left[\sum_{n=t+1}^{T} M_n D_n + M_T S_T \right]$$

- So if we define the fundamental value $S_t^* = \sum_{n=t+1}^{T} \frac{M_n}{M_t} D_n + \frac{M_T}{M_t} S_T$ then $S_t = \mathbb{E}_t[S_t^*]$ and Shiller (1981) notes that we should have $\mathbb{V}(S_t) \leq \mathbb{V}(S_t^*)$. (why?)
- Suppose that discount rates are constant $M_t = e^{-\beta t}$ then

$$S_t = \mathbb{E}_t \left[\sum_{n=t+1}^{T} e^{-\beta(n-t)} D_n + e^{-\beta(T-t)} S_T \right]$$

- Shiller's famous plot of S_t^* versus S_t shows excess volatility!

- To illustrate further, suppose dividends are i.i.d. Define $d_t = \log D_t$ and with i.i.d. $N(0,1)$ sequence of ϵ_t.

$$d_{t+1} = d_t + \mu + \sigma\epsilon_{t+1} - \frac{1}{2}\sigma^2$$

- Then $S_0 = D_0 \sum_{n=1}^{\infty} e^{-(\beta-\mu)n} = D_0 \frac{1}{e^{\beta-\mu}-1} \approx \frac{D_0}{\beta-\mu}$ if the transversality condition $\beta > \mu$ holds.
- So, we see that the price-dividend ratio is constant in this world. The volatility of stock prices should be equal to that of dividends!
- This is the excess volatility puzzle: Stock prices are much more volatile than the fundamental cash-flows they are a claim to.
- This 'puzzle' depends on
  1. Constant discount rate (i.e., no discount factor risk)
  2. i.i.d. dividend (i.e., no 'long-run' cash-flow risk)

The Campbell-Shiller decomposition

- In discrete-time it is often helpful to use the simplifying 'log-linearization' of the Return equation (due to Campbell-Shiller (1988)).
  Consider the gross return $R_{t+1} = \frac{S_{t+1}+D_{t+1}}{S_t}$. Define the price dividend ratio: $PD_t = S_t/D_t$. Since

$$\begin{aligned} R_{t+1} &= \frac{D_{t+1}}{D_t}\frac{PD_{t+1}+1}{PD_t} \\ &= \frac{D_{t+1}}{D_t}\frac{(1+e^{pd_{t+1}})}{PD_t} \end{aligned}$$

  the log-return $r_{t+1} = \log R_{t+1}$ can be expressed in terms of log dividend growth $\Delta d_{t+1} = \log(\frac{D_{t+1}}{D_t})$ and log-price dividend ratio $pd_t = \log(\frac{S_t}{D_t})$ as:

- $$r_{t+1} = \kappa_0 + \kappa_1 pd_{t+1} + \Delta d_{t+1} - pd_t$$

  where $\kappa_1 = \frac{PD}{1+PD}$ and $\kappa_0 = \log(1+PD) - \kappa_1\, pd$ and $PD = e^{pd}$ is the value of the price dividend ratio around which we log-linearize the expression $\log(1 + e^{pd_{t+1}})$.
- This expression can be helpful to derive expressions for the equilibrium price dividend ratio. Define $m_{t+1} = \log \frac{M_{t+1}}{M_t}$. Then $\mathbb{E}_t[e^{m_{t+1}+r_{t+1}}] = 1$ implies

$$\mathbb{E}_t[e^{m_{t+1}+\kappa_0+\kappa_1 pd_{t+1}+\Delta d_{t+1}-pd_t}] = 1$$

- In a conditionally log-normal framework, this is equivalent to:

$$\mathbb{E}_t[m_{t+1} + \kappa_0 + \kappa_1 pd_{t+1} + \Delta d_{t+1} - pd_t] = \\ -\frac{1}{2}\mathbb{V}_t[m_{t+1} + \kappa_1 pd_{t+1} + \Delta d_{t+1} - pd_t]$$

- For the case of a constant discount rate, this simplifies:

$$\mathbb{E}_t[e^{-\beta+\kappa_0+\kappa_1 pd_{t+1}+\Delta d_{t+1}-pd_t}] = 1$$

Looking for a constant price dividend ratio, we obtain:

$$-\beta + \kappa_0 + (\kappa_1 - 1)pd + \mu = 0$$

or equivalently plugging in the expression for κ_0, κ_1:

$$\log(\frac{1 + PD}{PD}) = \beta - \mu$$

Of course, with constant *pd* the log-linearization is exact but not really useful (but it will be later for more general settings where the pd-ratio is time-varying).

Predictable Returns or Dividends?

- Iterating the Campbell-Shiller relation forward, we obtain:

$$pd_t = C + \sum_{n=1}^{\infty} \kappa_1^{n-1} \Delta d_{t+n} - \sum_{n=1}^{\infty} \kappa_1^{n-1} r_{t+n}$$

for some constant C. Now, taking conditional expectation, we obtain:

$$pd_t = C + \mathbb{E}_t[\sum_{n=1}^{\infty} \kappa_1^{n-1} \Delta d_{t+n}] - \mathbb{E}_t[\sum_{n=1}^{\infty} \kappa_1^{n-1} r_{t+n}]$$

which shows that all variation in price dividend ratios must be attributable to variation in expected returns and/or expected dividend growth!

- Further multiplying by $pd_t - \mathbb{E}[pd_t]$ and taking expectation, we get the identity (Cochrane (1991)):

$$\mathbb{V}(pd_t) = \mathbb{C}\mathrm{ov}[\sum_{n=1}^{\infty} \kappa_1^{n-1} \Delta d_{t+n}, pd_t] - \mathbb{C}\mathrm{ov}[\sum_{n=1}^{\infty} \kappa_1^{n-1} r_{t+n}, pd_t]$$

  This identity shows that all variability in the price-dividend ratio has to come from predictability in returns and/or dividend growth! Dividing on both sides by $\mathbb{V}(pd_t)$, we get a decomposition of the coefficients in long-run predictability regressions of returns and dividends on the price dividend:

$$\beta[\sum_{n=1}^{\infty} \kappa_1^{n-1} \Delta d_{t+n}, pd_t] - \beta[\sum_{n=1}^{\infty} \kappa_1^{n-1} r_{t+n}, pd_t] = 1$$

  See the recent paper by Cochrane (2008) and his presidential address (2010) for further discussions of the implications of this identity.

- A stark conclusion of these accounting identities (!) is that if there is variability in the price-dividend ratio (there is a lot!), then it **must be** that either returns are predictable or dividend growth is predictable (or both). Cochrane (1991, 2008) argues that while the evidence for stock-return predictability is not overwhelming, the *"consensus"* that dividend growth seems unpredictable should lead to the conclusion that stocks are, in fact, predictable. (this is *The dog that didn't bark* argument).
- The jury is actually still out on whether dividend growth is more or less predictable than returns. That said, these relations clearly show that any "excess" volatility in prices relative to dividends must be related to either volatility in (i) long-term cash flows or (ii) discount rates (or in both). The initial analysis of Shiller completely ignored the possibility of stochastic discount rates.
- Note also that the relation does rely on the CS approximation being accurate, which in turn depends on the stationarity assumption of PD ratios (see the paper by Lettau and Van Nieuwerburgh for a discussion).

The equity premium puzzle

- Grossman-Shiller (1981) point out that in the simplest representative agent equilibrium exchange economy, we would expect the discount rate to be stochastic and tied to aggregate consumption. In fact, in a time separable utility framework, $M_t = e^{-\beta t}u'(C_t)$ where C_t is aggregate consumption. And with $u'(C) = C^{-\gamma} = e^{-\gamma c}$

$$S_t = \mathbb{E}_t\left[\sum_{n=t+1}^{T} e^{-\beta(n-t)-\gamma(c_n-c_t)}D_n + e^{-\beta(T-t)-\gamma(c_T-c_t)}S_T\right]$$

$$:=\mathbb{E}_t[S_t^*]$$

- Grossman-Shiller plot the variability of their redefined fundamental value and test the relation $Var(S_t^*) > Var(S_t)$, where they use aggregate consumption data as an input to the stochastic discount factor for various levels of γ. While the "excess volatility" conclusion still holds, it is less striking and now a function of the assumed level of risk-aversion (Should $\gamma = 1$ or 10?).
- We see that any test of excess volatility or, more generally, of market efficiency will be a joint test of the model for the stochastic discount factor.
- Further, volatility-bound tests are equivalent to return predictability regressions. See, in particular, the insightful discussion in Cochrane (1991).

- Mehra-Prescott (MP 1985) proposes to calibrate a representative agent economy and look at the quantitative predictions of their model for asset price moments (specifically, risk-free rate level, PD ratio, and equity risk-premium, volatility, and sharpe ratio). See your HW assignment.
- If consumption equals dividends is such that with $c_t = \log C_t$:

$$c_{t+1} = c_t + \mu + \sigma \epsilon_{t+1} - \frac{1}{2}\sigma^2$$

- With CCRA $m_{t+1} = -\beta - \gamma \Delta c_{t+1}$ implies that the continuously compounded one period risk-free rate is

$$r_f = -\log \mathbb{E}_t[e^{m_{t+1}}] = \beta + \gamma\mu - \gamma(\gamma+1)\frac{\sigma^2}{2}$$

- Note that if the consumption stream is deterministic ($\sigma = 0$) then the elasticity of intertemporal substitution (EIS) we obtain from the Euler equation is $\psi = \frac{\partial \Delta c_{t+1}}{\partial r_f} = \frac{1}{\gamma}$. That is, the EIS is equal to the inverse of the risk-aversion coefficient for a CRRA investor.
- The return must be such that:

$$0 = \log \mathbb{E}_t[e^{m_{t+1}+r_{t+1}}]$$

In a conditionally log-normal setting, this implies:

$$\mathbb{E}_t[m_{t+1} + r_{t+1}] + \frac{1}{2}\mathbb{V}_t[m_{t+1} + r_{t+1}] = 0$$

or equivalently

$$\mathbb{E}_t[r_{t+1}] + \mathbb{E}_t[m_{t+1}] + \frac{1}{2}\mathbb{V}_t[r_{t+1}] + \frac{1}{2}\mathbb{V}_t[m_{t+1}] = - \mathbb{C}\mathrm{ov}_t(m_{t+1}, r_{t+1})$$

- Now since in conditionally log-normal setting the risk-free rate is:

$$\boxed{r_f = -\mathbb{E}_t[m_{t+1}] - \tfrac{1}{2}\mathbb{V}_t[m_{t+1}]}$$

we obtain an expression for the risk-premium:

$$\boxed{\mathbb{E}_t[r_{t+1}] - r_f + \tfrac{1}{2}\mathbb{V}_t[r_{t+1}] = -\mathbb{C}\mathrm{ov}_t(m_{t+1}, r_{t+1})}$$

- With a CRRA pricing kernel, we obtain that

$$\mathbb{E}_t[r_{t+1}] - r_f + \frac{1}{2}\mathbb{V}_t[r_{t+1}] = \gamma\mathbb{C}\mathrm{ov}_t(\Delta c_{t+1}, r_{t+1})$$

That is, only the covariance with contemporaneous log-consumption growth matters for risk-premia.

- If we use the definition of a log-return on a claim to aggregate consumption and assume that the price-consumption ratio is constant, then $r_{t+1} = \Delta c_{t+1} + \log(\frac{1+PC}{PC})$ and it follows that:

$$\log(\frac{1+PC}{PC}) = \beta + (\gamma - 1)\mu - \frac{1}{2}(\gamma - 1)\gamma\sigma^2$$

- Note the impact of risk-aversion on the PC ratio by comparing it with our previous expression. Further, note that the impact of an increase in μ and σ on the PD ratio depends on the sign of $\gamma - 1$. For $\gamma = 1$ (log-utility) PC ratio is independent of μ, σ and only depends on β.
- Further, note that the expected excess log-return

$$\mathbb{E}[r_{t+1}] - r_f + \frac{1}{2}\mathbb{V}[r_{t+1}] = \gamma\sigma^2$$

- and the Variance of log-returns: $\mathbb{V}[r_{t+1}] = \sigma^2$.
- Thus, the log-return Sharpe ratio is $\gamma\sigma$.

- Mehra-Prescott's main conclusion is that they cannot reconcile the asset pricing moments estimated in the data with the very low variability of aggregate consumption.
    - To match the equity premium requires a large risk-aversion coefficient. Too large based on "introspection" (this is the equity premium puzzle).
    - If they push the risk-aversion coefficient to such a high level (maybe introspection does not apply to the representative agent?), then the risk-free rate is too high. This is the risk-free rate puzzle.
- In the end, risks that appear to be priced in asset returns (equity volatility, Sharpe ratio, low risk-free rate) seem too high relative to the macro-economic quantity of risk (consumption volatility) we measure in the data.
- The model relies on simplifying assumptions:
    - Representative agent (and/or complete markets).
    - Time-separable CRRA utility function.
    - i.i.d. consumption growth.
    - i.i.d. dividend growth (equal to consumption growth in the simplest setting).

Habit Formation (Campbell-Cochrane)

- CC assume representative agent with external habit formation: $U(t, C, X) = e^{-\beta t} \frac{(C-X)^{1-\gamma}}{1-\gamma}$
- Assume X_t is external (i.e., the agent does not take into account that his actions affect future dynamics of the habit), then the standard Euler equation holds.

$$\frac{M_{t+1}}{M_t} = e^{-\beta} \left(\frac{C_{t+1} - X_{t+1}}{C_t - X_t}\right)^{-\gamma}$$

- Defining the *Surplus Consumption Ratio* $S_t = \frac{C_t - X_t}{C_t}$, CC assume that the dynamics of X are such that $s_t = \log S_t$ follows:

$$\begin{aligned} s_{t+1} &= (1-\phi)\bar{s} + \phi s_t + \lambda(s_t)\sigma\epsilon_{t+1} \\ \Delta c_{t+1} &= \mu + \sigma\epsilon_{t+1} - \sigma^2/2 \\ \lambda(s_t) &= e^{-\bar{s}}\sqrt{1-2(s_t-\bar{s})} - 1 \\ e^{\bar{s}} &= \sigma\sqrt{\frac{\gamma}{1-\phi}} \end{aligned}$$

- The log pricing kernel is then given by

$$m_{t+1} = -\beta - \gamma\Delta(s_{t+1} + c_{t+1})$$

- The risk-free rate satisfies $e^{-r_f} = \mathbb{E}_t[e^{m_{t+1}}]$, which in the conditionally Gaussian setting leads to a constant interest rate:

$$\begin{aligned} r_f &= -\mathbb{E}_t[m_{t+1}] - \frac{1}{2}\mathbb{V}_t[m_{t+1}] \\ &= \beta + \gamma(\mu - \sigma^2/2) - \frac{1}{2}\gamma(1-\phi) \end{aligned}$$

- For the price-consumption ratio, they need to solve it numerically. They solve using quadrature techniques the recursive equation:

$$PC(s_t) = \mathbb{E}_t[e^{m_{t+1}+\Delta c_{t+1}}(1 + PC(s_{t+1}))]$$

- Pick parameters $\mu - \sigma^2/2 = 1.89\%$, $\sigma = 1.5\%$, $r_f = 0.94\%$ (matched with $\beta = -\log(0.89)$). Then set $\gamma = 2$, $\phi = 0.87$and $e^{\bar{s}} = 0.057$.
- They also assume that dividends are riskier than aggregate consumption and price the claim to aggregate dividends.
- With these parameters, they 'claim to' be able to match the average equity premium (6.5%), the average volatility of stock returns 15%, and the average Sharpe ratio, as well as the level of the PD ratio, and the autocorrelation in PD ratios.

- How does it work?
  - It turns out that γ is not the RRA coefficient any longer. The actual RRA coefficient is $-\frac{CU_{CC}}{U_C} = \gamma e^{-s_t}$. This is highly volatile and, on average, equal to 40!
  - In bad times, consumption falls close to the habit, which raises risk-aversion and thus risk-premia
    $\Rightarrow$ countercyclical risk-premia and PD ratios.
  - Their model avoids the risk-free rate puzzle (high risk-aversion $\Rightarrow$ high desire to borrow to smooth consumption $\Rightarrow$ high equilibrium risk-free rate) by having ad-hoc $\lambda(s)$ function calibrated so that precautionary motive for saving is time-varying and exactly cancels the effect of risk-aversion on the risk-free rate.

- The numerical simulations in the original paper are incorrect. The solutions are highly sensitive to the integration bounds. It is more robust to estimate the solution by calculating the present value of future dividends discounted at the risk-free rate (constant!) under the risk-neutral measure. This approach is more stable.
- In the model the consumption surplus is perfectly correlated with aggregate consumption and with PD ratios. However, in the data, both are far from perfectly related.

Heterogeneous Agents and Idiosyncratic Labor income

- Constantinides-Duffie (CD 1996) propose to relax the notion of a representative agent to explain the equity premium puzzle.
- In their model, agents have idiosyncratic labor income shocks specified in such a way that their model supports a no-trade equilibrium given any pattern in stock return and aggregate consumption.

- Suppose there is a continuum of agents indexed by i. Each with consumption $C_{i,t} = \upsilon_{i,t} C_t$ where C_t is aggregate consumption. The idiosyncratic shock is modeled as follows:

$$\log(\frac{\upsilon_{i,t+1}}{\upsilon_{i,t}}) = \eta_{i,t+1} Y_{t+1} - \frac{1}{2} Y_{t+1}^2$$

$$Y_{t+1} = \sqrt{\frac{2}{\gamma(\gamma+1)}(m_{t+1} + \delta + \gamma \Delta c_{t+1})}$$

  where

  - $c_t = \log C_t$
  - $m_t = \log \frac{M_{t+1}}{M_t}$ for some pricing kernel M_t, i.e., such that $\mathbb{E}_t[e^{m_{t+1}} R_{t+1}] = 1$ for any traded asset with return R_t.
  - $\eta_{i,t+1}$ are i.i.d. standard normal random variables.

- Note that to be well-specified, CD assume that such a pricing kernel exists and satisfies.

$$m_{t+1} + \delta + \gamma \Delta c_{t+1} > 0$$

- Under this assumption, CD show that any aggregate consumption process C_t and any return process R_t is consistent with a buy and hold equilibrium. Indeed, the Euler condition is satisfied for any agent:

$$1 = \mathbb{E}_t \left[e^{-\delta - \gamma \Delta c_{i,t+1}} R_{t+1} \right]$$

- It thus seems that CD's argument can reconcile arbitrary smooth aggregate consumption with very volatile stock returns if individual agents' consumption is affected by persistent idiosyncratic shocks $\upsilon_{i,t}$.

- How does it work?
  Note that Individual optimality holds in that for all i the Euler condition is satisfied. With CRRA, that is

$$0 = \mathbb{E}_t[e^{-\gamma \Delta c_{i,t+1}}(R_{t+1} - R_f)]$$

  Suppose for intuition purposes that the cross-sectional distribution of log individual consumption is log-normally distributed with mean $\mu_{c,t+1}$ and variance $v_{c,t+1}$. Then, integrating the Euler equation across that distribution, we get:

$$0 = \mathbb{E}_t[e^{-\gamma \mu_{c,t+1} + \frac{\gamma^2}{2} v_{c,t+1}}(R_{t+1} - R_f)]$$

  This shows that with heterogeneous agents and cross-sectional variance in their individual consumption, if the Euler equation holds at the individual level, then both the cross-sectional average consumption ($\sim$ aggregate consumption) and its variance should drive the pricing kernel. For the latter to matter and raise the equilibrium risk-premium, we see from the equation above that $v_{c,t}$ needs to be time-varying and high when market excess return is low (why?).

- The specification of the idiosyncratic shocks in CD precisely satisfies these requirements.
- In effect, CD do not really solve the equity premium puzzle. Their model shows that if agents face very specific types of idiosyncratic shocks that have time-varying cross-sectional variance that is high when stock markets are low, and if these shocks are completely unchangeable, then agents may be stuck with a very volatile individual consumption even though aggregate consumption can be very smooth. And this can explain very high risk-premia on asset prices. However, in the model, the consumption CAPM holds at the individual level, and we explain high asset premia because individual consumption is very volatile.

- This has led to various tests of that proposition by resorting to PSID data on household consumption. The typical findings are that the equity premium puzzle basically holds at the individual level. Individuals have too smooth a consumption to explain asset returns. This empirical literature uncovers further puzzles: Individuals have too little exposure to stock returns given their consumption dynamics and also do not seem to optimally invest over the life-cycle (see Heaton and Lucas).

- In general, the consensus in the literature seems to be that the CD explanation requires labor income shocks that are entirely permanent, which seems inconsistent with empirical evidence. With less permanent shocks, agents can typically hedge these shocks by trading in existing securities, even if these provide imperfect hedging. In that case, the effects of market incompleteness due to heterogeneous labor income on risk-premia seem much less important (see Telmer).
- Empirically, the CCAPM seems not to hold at the individual agent's level. However, there is still literature focusing on specific agents (rich individuals, art and luxury goods consumption, trash, etc...) that finds more promising results.

Kreps-Porteus-Epstein-Zin Utility

- Recursive utility seeks to separate risk-aversion from EIS. The continuation utility of a consumption stream is now defined recursively via:

$$
\begin{aligned}
V_t^\rho &= (1-e^{-\beta})C_t^\rho + e^{-\beta}\mathbb{E}_t[V_{t+1}^\alpha]^{\rho/\alpha} \quad \forall\ \rho \neq 0 \\
\log V_t &= (1-e^{-\beta})\log C_t + e^{-\beta}\log \mathbb{E}_t[V_{t+1}^\alpha]^{1/\alpha} \quad \text{if } \rho = 0
\end{aligned}
$$

- One verifies that if $\rho = \alpha$ then V_t is of the time-separable CRRA form. Further, Epstein and Zin (1989) show that for gambles that are risk-free, the continuation utility is of the form constant EIS (i.e., identical to the CRRA time separable model) with an EIS coefficient $\psi = \frac{1}{1-\rho}$. Instead, if agents consider one-period risky gambles, then the risk-premium they are willing to pay to avoid such a gamble is identical to that a CRRA utility agent would pay with a CRRA coefficient equal to $\gamma = 1 - \alpha$. In that sense, this utility separates the coefficient of EIS from that of relative risk aversion. It turns out it also leads to a preference for early resolution of uncertainty, as we now show. The following is based on Epstein, Farhi, and Strzalecki (2014).

- Suppose consumption is i.i.d. as before; what is the continuation utility of such consumption stream? Let's focus on the case $\rho = 0$ where calculations are simple. Guess that $\log V_t = a + bc_t$ then plugging into the recursion we obtain

$$a + bc_t = (1 - e^{-\beta})c_t + e^{-\beta} \log \mathbb{E}_t[e^{a\alpha + b\alpha c_{t+1}}]^{1/\alpha}$$

or

$$a + bc_t = (1 - e^{-\beta})c_t + e^{-\beta}(a + b(c_t + \mu - \frac{1}{2}\sigma^2) + b^2\alpha\sigma^2/2)$$

Equating terms in c_t and constants, we find two equations that are easily solved for a, b. We find $b = 1$ and $a = e^{-\beta}\frac{(\mu - \frac{1}{2}\sigma^2 + \alpha\sigma^2/2)}{1 - e^{-\beta}}$ and thus

$$\log V_0 = c_0 + e^{-\beta}\frac{(\mu - \sigma^2/2)}{1 - e^{-\beta}} + \alpha e^{-\beta}\frac{\sigma^2/2}{(1 - e^{-\beta})}$$

- To see the effect of early resolution of uncertainty, consider the case where all consumption risk is resolved next period. So in period 1, the agents know the whole future path $c_1, \ldots, c_\infty$. This utility at time 1 is simply.

$$
\begin{aligned}
\log V_1^e &= (1 - e^{-\beta}) \log C_1 + e^{-\beta} \log V_2^e \\
&= c_0 + \Delta c_1 + e^{-\beta} \Delta c_2 + e^{-2\beta} \Delta c_3 + \ldots
\end{aligned}
$$

which is normally distributed with mean $M_1 = c_0 + \frac{(\mu - \sigma^2/2)}{1 - e^{-\beta}}$ and variance $\mathbb{V}_1 = \sum_{i=0}^{\infty} \sigma^2 e^{-i2\beta} = \frac{\sigma^2}{1 - e^{-2\beta}}$.

- We can thus find the continuation utility at date 0 from this early resolution consumption plan. It solves

$$
\begin{aligned}
\log V_0^e &= (1-e^{-\beta})c_0 + e^{-\beta}(M_1 + \frac{1}{2}\alpha\mathbb{V}_1) \\
&= c_0 + e^{-\beta}\frac{(\mu - \sigma^2/2)}{1-e^{-\beta}} + \alpha e^{-\beta}\frac{\sigma^2/2}{(1-e^{-2\beta})}
\end{aligned}
$$

If we compare both utility streams we see that $\log V_0^e - \log V_0 = -\alpha e^{-\beta}\sigma^2/2\frac{e^{-\beta}}{(1-e^{-2\beta})}$. Thus:

  - if $\sigma = 0$ they are both equal.
  - if $\alpha = 0 = \rho$ then both are equal.
  - if $\sigma > 0$ then early resolution is (i) preferred if $\alpha < 0 = \rho$, but (ii) disliked if $\alpha > 0 = \rho$.

- There is another interesting alternative consumption plan to consider. Suppose instead that the agent is proposed a consumption stream, where in period one he will get one random draw Δc_1 and that for all future dates $\Delta c_i = \Delta c_1$, that is future consumption shocks are perfectly correlated, there is no time-diversification in consumption chocks. Let's call the utility associated with this consumption plan V^c, then as before:

$$\begin{aligned} \log V_1^c &= (1 - e^{-\beta}) \log C_1 + e^{-\beta} \log V_2^c \\ &= c_0 + \Delta c_1 + e^{-\beta} \Delta c_2 + e^{-2\beta} \Delta c_3 + \dots \end{aligned}$$

which is normally distributed with mean $M_1 = c_0 + \frac{(\mu - \sigma^2/2)}{1 - e^{-\beta}}$ and variance $\mathbb{V}_1^c = \sigma^2 (\sum_{i=0}^{\infty} e^{-i\beta})^2 = \frac{\sigma^2}{(1 - e^{-\beta})^2}$.

- We can thus find the continuation utility at date 0 from this "correlated" consumption plan. It solves

$$\begin{aligned}\log V_0^c &= (1-e^{-\beta})c_0 + e^{-\beta}(M_1 + \frac{1}{2}\alpha\mathbb{V}_1^c) \\ &= c_0 + e^{-\beta}\frac{(\mu-\sigma^2/2)}{1-e^{-\beta}} + \alpha e^{-\beta}\frac{\sigma^2/2}{(1-e^{-\beta})^2}\end{aligned}$$

If we compare both utility streams we see that $\log V_0^c - \log V_0 = -\alpha e^{-2\beta}\frac{\sigma^2/2}{(1-e^{-\beta})^2}$. Thus:
  - if $\sigma = 0$ they are both equal.
  - if $\alpha = 0 = \rho$ then both are equal.
  - if $\sigma > 0$ then (i) $V_0^c \geq V_0$ if $\alpha < 0 = \rho$, but (ii) $V_0^c \geq V_0$ if $\alpha > 0 = \rho$.

- Note that the difference $V_0^c - V_0$ contains two components, one due to early resolution and one due to time-diversification. We can isolate the pure time-diversification component by comparing V_0^c to V_0^e. in particular, note:

  $\log V_0^e - \log V_0^c = -2\alpha e^{-\beta} \frac{\sigma^2/2}{(1-e^{-\beta})^2(1+e^{-\beta})}$ Thus:
  - if $\sigma = 0$ they are both equal.
  - if $\alpha = 0 = \rho$ then both are equal.
  - if $\sigma > 0$ then time-diversification is preferred if $\alpha < 0 = \rho$, but (ii) disliked if $\alpha > 0 = \rho$.
- This result can be generalized to $\rho \neq 0$. In general, investors prefer early resolution of uncertainty and time-diversification if $\alpha > \rho$.

- As we will see next, a preference for the early resolution of uncertainty is crucial to explaining asset pricing puzzles in the Bansal-Yaron paper.
- Epstein et al. (2015) extend the analysis above to discuss the magnitude of the preference for early resolution. They argue that for the preference parameters chosen by BY, agents would be willing to give up an unreasonably (based on introspection) large amount of wealth (or per-period consumption) to move from V to V^e. So, in a sense, they argue that while the RRA and EIS coefficients seem "reasonable," the combination of both within the EZ utility leads to an unreasonably high aversion to late resolution *for the parameters and endowment process chosen by BY*. So perhaps the puzzle is still alive...

The stochastic discount factor in the EZ economy

- Suppose we want to value long-term assets in a BY economy.
- Let's derive the Euler equation for the EZ agent.
- Consider an optimal feasible consumption plan C_t and a deviation whereby on an $\mathcal{F}_t$ measurable state A, we have $\tilde{C}_t = C_t - \epsilon \mathbf{1}_A$ and at $t+1$ we have $\tilde{C}_{t+1}(\omega) = C_{t+1}(\omega) + \epsilon R_{t+1}(\omega) \mathbf{1}_A$. Note that $\tilde{C}_t$ is feasible and only differs from C_t on A.
- Then $\tilde{V}_t = V_t - \frac{\partial V_t}{\partial C_t} \epsilon \mathbf{1}_A + \sum_\omega \frac{\partial V_t}{\partial C_{t+1}(\omega)} \epsilon \mathbf{1}_A R_{t+1}(\omega)$

- Setting $\lim_{\epsilon \to 0} \frac{\tilde{V}_t - V_t}{\epsilon} = 0$ we get the relation:

$$\sum_{\omega} \frac{\frac{\partial V_t}{\partial C_{t+1}(\omega)}}{\frac{\partial V_t}{\partial C_t}} \mathbf{1}_A R_{t+1}(\omega) = 1 \quad \forall A \in \mathcal{F}_t$$

  This implies:

$$\mathbb{E}_t[\frac{M_{t+1}}{M_t} R_{t+1}] = 1$$

  where the stochastic discount factor M_t is defined by

$$\pi_t(\omega) \frac{M_{t+1}}{M_t} = \frac{\frac{\partial V_t}{\partial C_{t+1}(\omega)}}{\frac{\partial V_t}{\partial C_t}}$$

- Now, we use the chain rule to rewrite

$$\pi_t(\omega)\frac{M_{t+1}}{M_t} = \frac{\partial V_t}{\partial V_{t+1}(\omega)}\frac{MC_{t+1}(\omega)}{MC_t}$$

with

$$\begin{aligned} MC_t &= \frac{\partial V_t}{\partial C_t} = (1 - e^{-\beta})(\frac{V_t}{C_t})^{1-\rho} \\ \frac{\partial V_t}{\partial V_{t+1}(\omega)} &= \beta V_t^{1-\rho}\mathbb{E}_t[V_{t+1}^{\alpha}]^{\frac{\rho-\alpha}{\alpha}}\pi_t(\omega)V_{t+1}^{\alpha-1}(\omega) \end{aligned}$$

Combining, we get a state price deflator:

$$\frac{M_{t+1}}{M_t} = e^{-\beta}(\frac{C_{t+1}}{C_t})^{(\rho-1)}(\frac{V_{t+1}}{\mathbb{E}_t[V_{t+1}^{\alpha}]^{(1/\alpha)}})^{\alpha-\rho}$$

- In the literature, the state price deflator is often re-expressed in terms of the Return to the aggregate consumption claim $R_C(t+1)$ defined via the wealth equation:

$$W_{t+1} = (W_t - C_t)R_C(t+1)$$

and where the agent's wealth is defined by

$$M_t W_t = \mathbb{E}_t[\sum_{n=1}^{\infty} M_{t+n} C_{t+n}]$$

assuming the transversality condition holds.

- To identify $R_C(t+1)$ note that V_t is homogeneous of degree one in $C_t, V_{t+1}(\omega)$ so that by Euler's theorem:

$$V_t = MC_t\, C_t + \sum_{\omega} \frac{\partial V_t}{\partial V_{t+1}(\omega)} V_{t+1}(\omega)$$

Setting

$$W_t = \frac{V_t}{MC_t}$$

we can rewrite this expression as

$$W_t - C_t = \mathbb{E}_t[\frac{M_{t+1}}{M_t} W_{t+1}]$$

- Iterating forward, we see that this relation indeed defines W_t (assuming the transversality condition holds). Thus

$$R_C(t+1) = \frac{V_{t+1} MC_t}{MC_{t+1}(V_t - MC_t C_t)} = e^{\beta}(\frac{C_{t+1}}{C_t})^{1-\rho}(\frac{V_{t+1}}{\mathbb{E}_t[V_{t+1}^{\alpha}]^{1/\alpha}})^{\rho}$$

Note in particular that for the log-EIS case ($\rho = 0$), we immediately get:

$$\log R_C(t+1) = \beta + \Delta c_{t+1}$$

Since $\log R_C(t+1) = \Delta c_{t+1} + \log \frac{WC}{(WC-1)}$, this implies that the wealth consumption ratio is constant when agents have log-EIS ($\rho = 0$):

$$\log \frac{WC}{(WC-1)} = \beta$$

- However, note that the pricing kernel still reflects the risk-aversion coefficient of the agent α:

$$\frac{M_{t+1}}{M_t} = e^{-\beta}(\frac{C_{t+1}}{C_t})^{-1}(\frac{V_{t+1}^{\alpha}}{\mathbb{E}_t[V_{t+1}^{\alpha}]^{(1/\alpha)}})$$

For the case where $\rho \neq 1$, we can get a simpler expression for the pricing kernel by substituting the expression for R_C into the pricing kernel we obtain:

$$\frac{M_{t+1}}{M_t} = e^{-\beta\theta}(\frac{C_{t+1}}{C_t})^{\theta(\rho-1)} R_C(t+1)^{\theta-1}$$

or equivalently in logs:

$$m_{t+1} = -\theta\beta + \theta(\rho - 1)\Delta c_{t+1} + (\theta - 1) r_c(t+1)$$

where $\theta = \alpha/\rho$.

i.i.d. Consumption Growth

- Using the Campbell-Shiller approximation (exact in this case) and looking for a constant PC ratio, we have

$$r_c(t+1) = \log(\frac{1+PC}{PC}) + \Delta c_{t+1}$$

Thus, noting that $\theta(\rho - 1) + (\theta - 1) = \alpha - 1 = -\gamma$ the RRA coefficient.

$$m_{t+1} = -\theta\beta - \gamma\Delta c_{t+1} + (\theta - 1)\log\frac{1+PC}{PC}$$

- This implies that the risk-premia in this economy are all identical to the CRRA economy (Why?).
  Only the risk-free rate is changed relative to the benchmark, i.i.d. CRRA economy.

- Let us first solve for the equilibrium PC. The claim that aggregate consumption solves

$$1 = \mathbb{E}_t[e^{m_{t+1}+r_c(t+1)}]$$

- in this conditionally Gaussian setting, we have equivalently:

$$\mathbb{E}_t[m_{t+1} + r_c(t+1)] = -\frac{1}{2}\mathbb{V}_t[m_{t+1} + r_c(t+1)]$$

$$-\theta\beta + \alpha(\mu - \sigma^2/2) + \theta \log \frac{1+PC}{PC} = -\frac{1}{2}\alpha^2\sigma^2$$

gives the solution for the PC ratio.

- The equilibrium risk-free rate solves

$$e^{-r_f} = \mathbb{E}_t[e^{m_{t+1}}]$$

- In this conditionally log-normal setting we can rewrite:

$$\boxed{r_f = -\mathbb{E}_t[m_{t+1}] - \tfrac{1}{2}\mathbb{V}_t[m_{t+1}]}$$

Thus,

$$r_f = \theta\beta + \gamma(\mu - \sigma^2/2) - (\theta - 1)\log\frac{1+PC}{PC} - \frac{1}{2}\gamma^2\sigma^2$$

Substituting for the PC ratio, we find

$$r_f = \beta + \mu/\psi - \frac{1}{2}(1+\gamma)/\psi\sigma^2$$

- The excess log-return on the stock satisfies:

$$\mathbb{E}_t[r_C(t+1)]+\mathbb{E}[m_{t+1}]+\frac{1}{2}\mathbb{V}_t[m_{t+1}]+\frac{1}{2}\mathbb{V}[r_C(t+1)] = -\mathbb{C}\mathrm{ov}_t[m_{t+1}, r_C(t+1)]$$

  It follows then that is identical to that obtained in the CCRA model:

$$\mathbb{E}_t[r_C(t+1)] - r_f + \frac{1}{2}\mathbb{V}[r_C(t+1)] = \gamma\sigma^2$$

- The Variance $\mathbb{V}_t[\log r_C(t+1)] = \sigma^2$ and thus the Sharpe Ratio is unchanged relative to the CRRA case at $\gamma\sigma$.
- Note, however, that the model can potentially solve the risk-free rate puzzle! By breaking the link between RRA and EIS, we can have both high RRA and high EIS, which goes in the right direction to solve the puzzles!

Long-Run Risk Model (Bansal-Yaron)

- BY specify log-consumption (Δc_{t+1}) and log-dividend (Δd_{t+1}) dynamics as driven by two persistent variables (x_t, σ_t):[1]

$$
\begin{aligned}
\Delta c_{t+1} &= \mu_c + x_t + \sigma_t \, \tilde{\epsilon}_{c,\,t+1} \\
\Delta d_{t+1} &= \mu_d + \rho_d x_t + \nu_c \sigma_t \, \tilde{\epsilon}_{c,\,t+1} + \nu_d \sigma_t \, \tilde{\epsilon}_{d,\,t+1} \\
x_{t+1} &= \rho_x x_t + \nu_x \sigma_t \tilde{\epsilon}_{x,\,t+1} \\
\sigma^2_{t+1} &= \overline{\sigma}^2 + \rho_\sigma \left(\sigma^2_t - \overline{\sigma}^2 \right) + \nu_\sigma \, \tilde{\epsilon}_{\sigma,\,t+1}.
\end{aligned}
$$

- The time interval Δt is monthly. We closely follow BKY in setting the parameters $\mu_c = 0.0015$, $\mu_d = 0.0015$ $\rho_d = 2.5$, $\nu_c = 2.6$, $\nu_d = 4.5$, $\rho_x = 0.98$, $\nu_x = 0.038$, $\overline{\sigma} = 0.0072$, $\rho_\sigma = 0.995$, $\nu_\sigma = 0.0000028$.

[1] This specification has one drawback: variance is modeled as an AR1 that can become negative. We nevertheless keep this specification since it is very tractable and allows us to follow existing literature for calibration. In principle, this issue could be avoided by specifying σ_t to follow an AR1 process or by directly modeling variance as a compound autoregressive affine process (a true discrete time square root process) as proposed by Gourieroux, Monfort and Polimenis (2007) and Le, Singleton and Dai (2010).

- Preferences are specified to be recursive as in Epstein and Zin (1989), which implies that the logarithm of the one-period (time-t conditional) pricing kernel can be written as:

$$m_{t+1} = \theta \log \delta - \frac{\theta}{\psi} \Delta c_{t+1} + (\theta - 1) r_{c,t+1},$$

where r_c is the log-return on the consumption claim, γ is relative risk aversion coefficient ($\alpha = 1 - \gamma$), ψ is the elasticity of intertemporal substitution ($\rho = 1 - 1/\psi$), and $\theta = \left(\frac{1-\gamma}{1-1/\psi}\right)$.

- For tractability, two approximations are made. The first is model-specific and assumes that log price/consumption is approximately affine in the state variables:

$$z_t \equiv \log\left(\frac{V_{c,t}}{C_t}\right) \approx A_0 + A_x x_t + A_\sigma \sigma_t^2.$$

- The second approximation (the Campbell/Shiller approximation) is mechanical and approximates the log-return $r_c \equiv \log R_c$ on the consumption claim to be linear in the log price-dividend (a similar approximation is made for log-stock-return):

$$r_c \equiv \log\left(\frac{V_{c,t+1} + C_{t+1}}{V_{c,t}}\right) \approx \kappa_0 + \kappa_1 z_{t+1} - z_t + \Delta c_{t+1},$$

where the constants κ_1, κ_0 are expressed in terms of $\bar{z} = A_0 + A_\sigma \bar{\sigma}^2$.

- To identify the pricing kernel, the six parameters $\{\bar{z}, \kappa_0, \kappa_1, A_0, A_x, A_\sigma\}$ are determined from the three equations for $\{\kappa_0, \kappa_1, \bar{z}\}$ and from the Euler equation:

$$1 = \mathrm{E}_t\left[e^{m_{t+1}+r_{c,t+1}}\right],$$

which in this conditionally normal framework can be re-expressed as

$$0 = \mathrm{E}_t\left[m_{t+1}+r_{c,t+1}\right] + \frac{1}{2}\mathrm{Var}_t\left[m_{t+1}+r_{c,t+1}\right].$$

- By collecting terms linear in x_t, linear in σ_t^2 and independent of the state vector, this equation generates three more restrictions, which in turn allows the six parameters to be identified. Moreover, the risk premia coefficients $\{\lambda_c, \lambda_x, \lambda_\sigma\}$ in the equation

$$m_{t+1} - \mathrm{E}_t \left[m_{t+1} \right] \quad = \quad -\lambda_c \sigma_t \tilde{\epsilon}_{c,\,t+1} - \lambda_x \sigma_t \tilde{\epsilon}_{x,\,t+1} - \lambda_\sigma \nu_\sigma \tilde{\epsilon}_{\sigma,\,t+1},$$

  are also identified.
  The risk-free rate at date-t is determined from the pricing kernel via:

$$\begin{aligned} e^{-r_f(x_t,\sigma_t)} &= \mathrm{E}_t \left[e^{m_{t+1}} \right] \\ &= e^{-r_0 - \frac{1}{\psi} x_t - r_\sigma \sigma_t^2}. \end{aligned}$$

- Once the pricing kernel has been identified, the same approach is used to identify the claim to dividends. In particular, we combine the proposed dividend dynamics in equation (1) with the equations

$$
\begin{aligned}
z_{d,t} &\equiv \log\left(\frac{V_{d,t}}{D_t}\right) \approx F_0 + F_x x_t + F_\sigma \sigma_t^2 \\
r_d &\approx \kappa_{0d} + \kappa_{1d} z_{d,t+1} - z_{d,t} + \Delta d_{t+1} \\
1 &= \mathrm{E}_t\left[e^{m_{t+1} + r_{d,t+1}}\right]
\end{aligned}
$$

to identify the parameters $\{\bar{z}_d, \kappa_{0d}, \kappa_{1d}, F_0, F_x, F_\sigma\}$.

- The log expected excess return for the dividend claim is

$$\begin{aligned}\overline{r}_{d,t} &\equiv \log \mathrm{E}_t \left[e^{\left(r_{d,t+1} - r_{f,t}\right)} \right] \\ &= \nu_\sigma^2 \kappa_{1d} F_\sigma \lambda_\sigma + \left(\nu_x \kappa_{1d} F_x \lambda_x + \nu_c \lambda_c\right) \sigma_t^2,\end{aligned}$$

  and the stock volatility is

$$\begin{aligned}\sigma_{d,t} &\equiv \sqrt{\log \mathrm{E}_t \left[e^{2\left(r_{d,t+1} - r_{f,t}\right)} \right] - 2 \log \mathrm{E}_t \left[e^{\left(r_{d,t+1} - r_{f,t}\right)} \right]} \\ &= \sqrt{(\kappa_{1d} F_\sigma \nu_\sigma)^2 + \sigma_t^2 \left[\nu_c^2 + \nu_d^2 + (\kappa_{1d} F_x \nu_x)^2 \right]}.\end{aligned}$$

- At the steady state values $x_t = 0, \sigma_t^2 = \overline{\sigma}^2$, the calibration generates the realistic annual expected excess return of $\overline{r}_{d,t} = 6.1\%$, and volatility $\sigma_{d,t} = 15.6\%$.

Long-Run Risk and/or Habits: some remarks

- Epstein critique: In the long-run risk model, risk-premia is large because of early resolution premium. See the table in their paper.
- Campbell-Beeler critique: In the long-run risk model, the term structure of risk-free real rates is downward sloping. Consumption growth and volatility are predictable, and stock returns predict future aggregate consumption. (the model is forward-looking.)

- Hansen-Sargent critique: If i.i.d. consumption is difficult to distinguish from the long-run risk model for consumption for the econometrician, then it should be so also for the agent in the model. If he cares much about this risk, then shouldn't it affect his behavior $\Rightarrow$ learning, uncertainty, robustness?
- Habit formation model: constant term structure of real rates. The model is backward-looking. Consumption is truly iid. Actual relative risk-aversion is at times very (unreasonably?) high.
- Both Habit and Long-Run risk models predict counterfactual upward-sloping term structure of dividend strip volatilities (see Brandt, Koijen, VanBinsbergen (2012)). In the data, more risk seems to be priced at the short end. Brandt et al. also report some results about the term structure of dividend strip Sharpe ratios that seem inconsistent with the predictions of both types of models.

- For more discussion of some of these issues, see Belo, CD, Goldstein (2014) on the term structure of dividend strips and CD, Johannes, and Lochstoer (2014) on endogenous subjective Long-run risks due to learning as a possible answer to the Campbell-Beeler critique (and also to the Epstein critique?).

- *The elephant in the room*: All this literature tries to explain the 'equity premium puzzle,' or in stochastic discount factor language, it tries to generate a high enough conditional volatility of the stochastic discount factor to explain a 'maximal' Sharpe ratio as high as that observed on equity returns. However, agents can trade many other asset classes (Currency, commodity, fixed income, real estate, private equity, options, variance swaps, risk-reversals, correlation swaps,...). The evidence is that there are many high Sharpe ratio strategies ('anomalies'?) to be traded across all asset classes that appear to not be perfectly correlated. If we take it at face value, then the maximal Sharpe ratio that we should be trying to explain is much higher, and all these models are still underperforming by a significant margin. The financial economics literature is still working in the Ether.

- One possible ‘explanation’ is that all these empirical results on ‘anomalies’ are pure data-mining, which uncover strategies that might have been but won’t be available in the future. Another possible ‘explanation’ is that there is a lot of unobserved risk (‘dark matter’) that agents are pricing. Things like tail/catastrophic risks that occur with frequencies of, say, 1 in a thousand years or less, and with perhaps time-varying intensities that agents need to hedge against... Both are not very satisfactory in my view. At least the first one we’ll be able to reject in a few years. Can you propose a better explanation (perhaps based on imperfect information, bounded rationality, limited cognitive abilities, political economy, sociology, endogenous market segmentation. . .)?