Lecture 2: The Consumption CAPM and the Equity Premium Puzzle
1 Overview

This lecture derives the “consumption-based capital asset pricing model”, or consumption CAPM. By applying this model we can understand the basis for a major empirical puzzle on the borderline between macro and finance: the “Equity Premium Puzzle”.
2 Optimal Consumption and Portfolio Choice in a 2 Period Model

We shall first assume that our consumer lives only 2 periods; but it turns out that the solution generalises very easily to multiple periods. Problem is:

\[
\max_{C_t, X_{1t} \ldots X_{Nt}} u(C_t) + \frac{1}{1 + \Theta} E_t(u(C_{t+1}))
\]

subject to

\[
A_t = C_t + \sum_{j=1}^{N} X_{jt}
\]

\[
C_{t+1} \equiv A_{t+1} = \sum_{j=1}^{N} (1 + R_{jt+1})X_{jt}
\]
This looks tricky to solve because the second constraint is stochastic in the second period: the consumer can’t *choose* second-period consumption (unless they only invest in a risk-free asset); they can only choose assets, the stochastic returns on which determine second-period consumption. But we can make it soluble by substituting from the 2nd constraint into the maximand, and rewrite the problem as a Lagrangian

$$\max_{C_0, X_{1t} \ldots X_{Nt}} u(C_t) + \frac{1}{1 + \Theta} E_t \left[ u \left( \sum_{j=1}^{N} (1 + R_{jt+1}) X_{jt} \right) \right]$$

$$- \lambda \left[ C_t + \sum_{j=1}^{N} X_{jt} - A_t \right]$$

where everything is now chosen in the first period. The first order conditions
are

\[ u'(C_t) - \lambda = 0 \]

\[ \frac{1}{1 + \Theta} E_t \left[u'(C_{t+1})(1 + R_{jt+1})\right] - \lambda = 0 \text{ for all } j \]

\[ C_t + \sum_{j=1}^{N} X_{jt} - A_t = 0 \]

which, by substituting from the first into the second, yields, for every asset \( j \), an equivalent of the Euler equation for the single asset in Lecture 1:

\[ u'(C_t) = E_t \left[ \left( \frac{1 + R_{jt+1}}{1 + \Theta} \right) u'(C_{t+1}) \right], \quad j = 1...J \]  \hspace{1cm} (1)

The left-hand side of (1) is the marginal utility cost of one less unit of consumption today, which, at the optimum, must equal the expected marginal utility
gain from investing that extra unit of saving in asset $j$, earning the return $R_{jt+1}$, and consuming it in period $t + 1$. 
3 The Stochastic Discount Factor

There is a close link between the Euler Equation and what is these days the dominant approach to asset pricing in finance. By dividing through, and exploiting the fact that $t-$dated variables can be taken in and out of the expectation at will, we can rewrite (1) as:

$$1 = E_t \left[ (1 + R_{jt+1}) M_{t+1} \right]$$

(2)

where

$$M_{t+1} = \left( \frac{1}{1 + \Theta} \right) \frac{u'(C_{t+1})}{u'(C_t)}$$

(3)

In finance $M_{t+1}$ is referred to as a “Stochastic Discount Factor”. Investors can be thought of as doing a present value calculation to compare expected
returns on all assets, where the same discount factor is applied to all assets. ie, usually we would expect $M_{t+1} < 1$.

In the absence of capital markets, every individual investor would have their own stochastic discount factor. But with complete, and frictionless capital markets and homogeneous expectations across investors, it is not too hard to see out that they will all end up sharing a *common* stochastic discount factor. To prove this formally is outside the scope of this course, but the intuition is actually quite straightforward. To see why, note that:

- Under these assumptions each individual investor faces the same set of asset prices;

- They also share expectations about the distribution of asset prices
• Hence in market equilibrium each investor must have the same stochastic discount factor

An alternative way of writing (2) that brings out the implications for asset pricing is to note that, if asset $j$ pays income flow $Y_{jt+1}$, then

$$1 + R_{jt+1} = \frac{Y_{jt+1}}{X_{jt}}$$

where $X_{jt}$ is the current market value of the asset. Hence

$$1 = E_t \left[ \left(1 + R_{jt+1}\right) M_{t+1} \right]$$

$$= E_t \left[ \frac{Y_{jt+1}}{X_{jt}} M_{t+1} \right]$$

$$X_{jt} = E_t \left[ Y_{jt+1} M_{t+1} \right]$$
so the market value of the asset is effectively its present value, but using a
discount factor that is itself stochastic, reflecting different valuations of an
additional pound of income in different states of nature.
Exercise: 1) Is the stochastic discount factor positively or negatively correlated with consumption?

2) Show that in a risk free general equilibrium where aggregate consumption is constant

\[ M_{t+1} = \frac{1}{1 + \Theta} = \frac{1}{1 + R_j} \quad \forall j, \forall t \]
In a stochastic world we can use the rule for the expectation of a product and write

\[ X_{jt} = E_t Y_{jt+1} E_t M_{t+1} + \text{cov}_t \left( Y_{jt+1}, M_{t+1} \right) \]

so assets will be valued more, the more correlated their payoffs are with the stochastic discount factor - ie, if they have a higher probability of higher payoffs when the marginal valuation of an extra pound of income is high: will this be when consumption is high or low?
4 Generalising optimal choice to multiple periods

The objective now becomes

\[
\max_{C_t, X_{1t} \ldots X_{Nt}} \quad U_t = E_t \sum_{i=0}^{T-t} \frac{1}{(1 + \Theta)^i} u(C_{t+i})
\]

subject to

\[
A_t = C_t + \sum_{j=1}^{N} X_{jt}
\]

\[
A_{t+1} = \sum_{j=1}^{N} (1 + R_{jt+1}) X_{jt}
\]

\[
C_T = A_T
\]
where now the first two constraints must hold in every single period and there is a terminal condition. But just as in the simple consumption problem the consumer cannot commit to a choice on any variable in the future, so the choice variables are exactly as in the two period problem. We can again formulate as a recursive value function. NB rewrite multiperiod objective as

\[
U_t = \frac{1}{(1 + \Theta)^t} (\mathbb{E}_t (u(C_{t+i})))
\]

\[
= u(C_t) + \frac{1}{(1 + \Theta)^t} \mathbb{E}_t \sum_{i=1}^{T-t} \frac{1}{(1 + \Theta)^i} (u(C_{t+i}))
\]

\[
= u(C_t) + \mathbb{E}_t \sum_{i=0}^{T-t} \frac{1}{(1 + \Theta)^i} (u(C_{t+1+i}))
\]

\[
= u(C_t) + \frac{1}{(1 + \Theta)^t} \mathbb{E}_t U_{t+1}
\]
and then reformulate the problem in terms of the value function as

\[ V_t = \max_{C_t, X_{1t} \ldots X_{Nt}} u(C_t) + \frac{1}{(1 + \Theta)} E_t V_{t+1} \]

subject to

\[ A_t = C_t + \sum_{j=1}^{N} X_{jt} \]

\[ A_{t+1} = \sum_{j=1}^{N} (1 + R_{jt+1}) X_{jt} \]

\[ C_T = A_T \]

Solving the two period problem and applying the envelope theorem on the assumption that the multiperiod problem has already been solved gives exactly the same condition for optimal asset choices as in (1) for \( j = 1 \ldots N \). But note that, as in the case of optimal consumption with a single asset these are again necessary but definitely not sufficient conditions for the optimal choice.
Exercise: 1) Clarify why $U_t \neq V_t$; 2) Derive first order condition for asset $j$ for two period problem by substituting out for $C_t$, giving

$$u'(C_t) = E_t \left[ \left( \frac{1 + R_{jt+1}}{1 + \Theta} \right) V_{t+1}' \right], \quad j = 1...J \tag{4}$$

then apply the envelope theorem to get (1) for $j = 1...N$; 3) explain why these are necessary but not sufficient conditions

For further background on multiperiod choice see Deaton pp 24-25.
5 An Apparent Digression: the Lognormal Distribution

5.1 A Useful Property of the Lognormal Distribution

If

\[ \log X \equiv x \sim N(\bar{x}, \sigma^2_x) \]

\[ E(X) \equiv E(e^x) = e^{\bar{x} + \frac{\sigma^2_x}{2}} \]

(which you should be able to see is actually just an example of Jensen’s Inequality). You may if you wish categorise this result as "mathematical magic"
something you don’t have to prove unless you like that sort of thing. For now focus on key features:

Exercise: 1) Show the link with Jensen’s Inequality for a strictly convex function ie

\[ E(f(x)) > f(E(x)) \text{ for } f \text{ strictly convex} \]

2) Give a geometric demonstration of this inequality if \( x \) can take only two possible values with equal probability;

3) Show that with lognormality \( X \) is bounded below at zero.
5.2 A simple application of lognormality: the link between risk aversion and intertemporal substitution

We saw last week that in the risk-free world, with power utility

\[ u(C) = \frac{C^{1-\gamma}}{1-\gamma} \]

\[ = \ln(C) \text{ for } \gamma = 1 \]

the optimal consumption path implied

\[ \frac{C_{t+i+1}}{C_{t+i}} = \left( \frac{1 + R}{1 + \Theta} \right)^{\frac{1}{\gamma}} \]

where \(1/\gamma\) is typically referred to as the elasticity of intertemporal substitution: it measures how sensitive the slope of the optimal consumption path is to the interest rate: a higher value of \(R\) gives a more upward-sloping path over time.
(ie, more intertemporal substitution) but the higher is $\gamma$, the less consumers will engage in intertemporal substitution (ie, the stronger their preference for stable consumption over time). We also saw that in general equilibrium this implies

$$r = \theta + \gamma g$$  \hspace{1cm} (7)

and hence

$$R \approx \Theta + \gamma G$$  \hspace{1cm} (8)

We can now show how $\gamma$ relates to risk aversion. Assume a two period model without assets, but with the possibility of insurance against consumption fluctuations. Suppose that in the absence of insurance consumption in the next period is lognormal, ie

$$\ln C_{t+1} = c_t \sim N(\bar{c}, \sigma^2_c)$$

Then it can be shown that, if a consumer is prepared to give up a proportion $\lambda$ of their expected income in the next period, $E_t C_{t+1}$ if this fully insures them
against consumption risk, and has power utility, then
\[ \lambda \approx -\ln(1 - \lambda) = \gamma \frac{\sigma_c^2}{2} \]

**Exercise:** 1) Show this! Some hints:

a) First write down an expression in terms of expected utility, i.e. \( \lambda \) is defined implicitly by
\[ U((1 - \lambda)E_tC_{t+1}) = E(U(C_{t+1})) \]

b) Substitute for utility, \( E_tC_{t+1} \) and \( E_tU_{t+1} \) writing everything as an exponential.

c) Simplify.
2) Hence show that if $\sigma = .2$ (implying approximately 20% standard deviation of consumption in the absence of insurance) and $\gamma = 2$ then $\lambda \approx .4\%$; for $\gamma = 10$, $\lambda \approx 33\%$.

3) NB: What is the approximate probability of uninsured consumption falling below $\exp(-.4) = 67\%$ of its expected value?

4) In the light of this answer, do values of $\gamma$ as high as 10 seem plausible?
6 Risk Premia in the Log-Normal CCAPM

Let

\[ r_{jt+1} = \log(1 + R_{jt+1}) \]
\[ r_t = \log(1 + R_t) \text{ (the safe return)} \]
\[ \theta = \log(1 + \Theta) \]
\[ c = \log C \]

And assume that \( c_{t+1} \) and \( r_{jt+1} \) are jointly normally distributed with constant conditional variances \( \sigma^2_c \) and \( \sigma^2_j \), and constant conditional covariance \( \sigma_{cj} \).

Now write the Euler Equation giving the explicit expression for the stochastic discount factor as

\[ 1 = E_t \left[ (1 + R_{jt+1}) \left( \frac{1}{1 + \Theta} \right) \frac{u'(C_{t+1})}{u'(C_t)} \right] \]  \hspace{1cm} (10)
Then, by applying these assumptions to the Euler Equation it can be shown, first, that:

$$0 = E_t(r_{jt+1}) - \theta - \gamma E_t \Delta c_{t+1} + \frac{1}{2} \left( \sigma_j^2 + \gamma^2 \sigma_c^2 - 2\gamma \sigma_{cj} \right)$$ \quad (11)

$$r_t = \theta + \gamma E_t \Delta c_{t+1} - \gamma^2 \frac{\sigma_c^2}{2}$$ \quad (12)

$$\approx \Theta + \gamma E_t \frac{\Delta C_{t+1}}{C_t} - \gamma^2 \frac{\sigma_c^2}{2}$$ \quad (13)

$$E_t(r_{jt+1} - r_t) + \frac{\sigma_j^2}{2} = \gamma \sigma_{cj}$$ \quad (14)
or, equivalently, and more compactly,

$$\log E_t \left( \frac{1 + R_{jt+1}}{1 + r} \right) \equiv \rho_j = \gamma \sigma_{cj}$$  \hspace{1cm} (15)
Exercise: Derive these expressions! (hints: a) use the properties of lognormality (so it wasn’t a digression after all....); b) use the property that if $a$ and $b$ are constants,

$$
var(ax + by) = a^2 var(x) + b^2 var(y) + 2ab \text{cov}(x, y)
$$

and c) take logs of both sides of (10)).
Equation (12), ie,

\[ r_t = \theta + \gamma E_t c_{t+1} - \gamma^2 \sigma_c^2 \]

is a stochastic equivalent of equation (11) in the first handout. Since \( r \) is the return on the safe asset, it is non-stochastic. Apart from the terms we had before, we now have an additional term, whereby the safe return is decreasing in the variance of log consumption growth. This is often called the “precautionary” effect: risky consumption raises demand for the safe asset, thus depressing its return.

Equation (15), ie

\[ \log E_t \left( \frac{1 + R_{jt+1}}{1 + r} \right) \equiv \rho_j = \gamma \sigma_{cj} \]

says that \( \rho_j \), the risk premium on asset \( j \) is given by \( \gamma \) times the covariance of \( r_j \) with log consumption growth. Assets that are strongly correlated with
consumption growth (hence with consumption in period $t+1$) will have a high risk premium. A higher degree of risk aversion (higher $\gamma$) will imply higher risk premia.
7 The “Equity Premium Puzzle”

This puzzle, identified by Mehra and Prescott is that, on the basis of observed covariances of stock returns with consumption, the implied degree of risk aversion is far too high to be consistent with estimates from other sources. To identify the puzzle in the data, they and subsequent authors apply the Law of Iterated Expectations working backwards in time to (12) and (14), repeatedly, eg applying it to the latter,

\[ E_{t-1}E_t(r_{jt+1} - r_t) + E_{t-1}\frac{\sigma_j^2}{2} = E_{t-1}\gamma\sigma_{cj} \]

but this simplifies to

\[ E_{t-1}(r_{jt+1} - r_t) + \frac{\sigma_j^2}{2} = \gamma\sigma_{cj} \]
since the expectation of a constant is a constant, and $E_{t-1}E_t = E_t$. If we do this over and over again, back to the dawn of time, we get

$$E(r_{jt+1} - r_t) + \frac{\sigma^2_j}{2} = \gamma \sigma c_j \tag{18}$$

where $Ex$ is the unconditional expectation of $x$. If $x$ has a stationary distribution, the best estimate of $Ex$ is its sample mean, $\bar{x}$. So the mean equity premium can be compared with observed covariances of consumption and equity returns to see the implied degree of risk aversion (which can effectively be treated as the only unknown in the above expression).

What they showed is that while risk premia are qualitatively consistent with what theory would predict, they are quantitatively way out. Evidence from elsewhere shows that faced with risky gambles, most people have a value of $\gamma$ in a range something like 2 to 4. In the problem set below you’ll be asked to use US data to show that the only way to reconcile the observed equity
premium with the data is to assume a value of $\gamma$ of 19! This finding has been confirmed by a range of datasets, using different statistical approaches. Some implied figures are even higher.

**Exercise:** With this value of $\gamma$, what proportion of their expected consumption would the consumer give up in the exercise in Section 5.2?
So maybe people are just more risk-averse than we thought? Unfortunately, if we assume this, we open up another puzzle. If we plug a value of $\gamma = 19$ into (12), the only way we can make it match the observed mean return on a safe asset (proxied by short-term risk-free paper) is to assume a negative value for $\Theta$ - implying that, far from discounting the future, investors would, other things being equal, actively prefer future consumption to current consumption. This seems massively counter-intuitive, since in a world of certainty it would imply negative real interest rates. This is the “Risk-Free Rate Puzzle”.

There is a massive literature on both of these puzzles. Both remain puzzles.

**Exercise:** Campbell *et al* provide the following annual data (see P308 for sources and definitions):
<table>
<thead>
<tr>
<th>$\bar{r}_{\text{stocks}}$</th>
<th>0.0601</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\bar{r}_{\text{safe}}$</td>
<td>0.0183</td>
</tr>
<tr>
<td>$\sigma_{\text{stocks}}$</td>
<td>0.1674</td>
</tr>
<tr>
<td>$\sigma_c$</td>
<td>0.0328</td>
</tr>
<tr>
<td>$\text{corr}(\Delta c_{t+1}, r_{\text{stocks}, t+1})$</td>
<td>0.4902</td>
</tr>
<tr>
<td>$\Delta c$</td>
<td>0.0172</td>
</tr>
</tbody>
</table>

Use this data and the definition of the correlation coefficient

$$\text{corr}(x, y) = \frac{\text{cov}(x, y)}{\sigma_x \sigma_y}$$

to derive a) an implied value for $\gamma$ from (18), and hence b) an implied value of $\theta$ (and hence $\Theta$) from implied by the unconditional version of (12). As the main text should have indicated, both of these implied figures will be silly.
We shall now move on to look at the stochastic growth model (though we have not heard the last of the equity premium puzzle)

For introductory coverage, read (in order of ease) relevant chapters of Williamson, Romer, or, if you are reasonably familiar with the basic ideas, see Campbell (1994), “Inspecting the Mechanism”, *Journal of Monetary Economics*