9 Week 3 Asset pricing theory – detailed notes

1. Facts, and Fama-French procedure motivate theory.
   
   (a) Week 1: Expected excess returns = risk premia vary over time. Why?
   
   (b) Week 2: Expected excess returns = risk premia vary across assets (“cross-section”) at any point in time. Why?
   
   i. Average returns on “value” portfolios as much as 3 times greater than on “growth” portfolios.
   
   ii. More basic: Stocks have paid 6% more than bonds for 150 years.
   
   (c) Note: Not “where do my great unexploited opportunities come from?” Yes, “why might the average investor be scared of “value risk,” or more scared at some times than others, and thus does not take these opportunities?” “In what sense might they really not be “opportunities” but “rewards exactly balanced by risks?” At least ask this question before pouncing!
   
   (d) “Explanations” we have seen
   
   i. CAPM

   \[
   E(R^i) - R^f = \beta_i \lambda_m \\
   E(R^i) - R^f = \beta_i (E(R^m) - R^f)
   \]

   ii. FF: this fails. “Explain” with FF3F model

   \[
   E(R^{ei}) = b_i \lambda_m + h_i \lambda_{hml} + s_i \lambda_{smb} \\
   E(R^{ei}) = b_i E(rmrf) + h_i E(hml) + s_i E(smb)
   \]

   iii. Why is this not an “explanation”

   \[
   E(R^{ei}) = (stock_i) \lambda_m + \log(size_i) \lambda_s + \log(beme_i) \lambda_{hml}
   \]

   A: This is a good description but not a model, or explanation. Models, explanations seem to need regression coefficients on the right hand side to “explain” the pattern of average returns. But why really?

   iv. A hint: regression coefficients have something to do with markets in equilibrium, with keeping you from making a fortune. If \( E(R^{ei}) = 16\% \) while the market yields \( E(R^{em}) = 8\% \), the observation that \( \beta_i = 2 \) means you can’t buy the asset, short the market, and earn the return difference.

   v. This is not very convincing at its surface. “I have 20% returns!” “Yeah, but you have a beta of 3” “why are you bugging me about some greek letter when I’m giving you real money.”

   (e) → Where do CAPM, FF3F come from? What constitutes an “explanation?” What is the question to which this is the answer? Persuade yourself that “beta” really does mean “risk” – and understand when (and for whom) it does not.

   (f) → Theory to understand risk premia.

   (g) Bottom line: Assets pay more on average if they tend to do badly when people are “hungry”, or more desperate for money.

   (h) Just a little algebra goes a long way!

* = Optional material.
9.1 Utility function

1. Basic question: you have a cash flow – dividends (stocks), coupons (bonds), rent (real estate), profits (build a factory), call option payoff.

   (a) What is its value? What are the effects risk and time.
   (b) How does value change if the world changes – $d\text{Value}/dz$? If you know this, you know how to do risk management.
   (c) Approach: apply apples and oranges microeconomics to finance.
   (d) *Preview:

   ![Diagram of utility function]

   ![Diagram of cash flow]

2. Simple setup:
(a) Payoff $x_{t+1}$ tomorrow. (For stocks, $x_{t+1} = p_{t+1} + d_{t+1}$)

(b) $x_{t+1}$ is a random variable, like a coin flip – we don’t know at $t$ what it will be, though we can assign probabilities to the possible outcomes.

(c) Randomness: you can think of $x_{t+1}$ (and anything else that happens at $t + 1$) as taking on different values in different states of the world. For example, $E(x_{t+1}) = \sum_s \pi(s)x(s)$

\[
\text{States } s = 1, 2, 3, \ldots
\]

(d) Our question: Price or value $p_t$ today of this payoff?

(e) (A common confusion: Payoff $x_{t+1}$ is not profit $x_{t+1} - p_t$. Remember call payoff diagrams?)


\[
U(c_t, c_{t+1}) = u(c_t) + \beta E_t [u(c_{t+1})]
\]

(a) Point of a utility function: model investor’s aversion to risk and delay.

(b) Example: Log

\[
u(c) = \ln(c); \quad u'(c) = \frac{1}{c}
\]
(c) $u(c) = \text{‘happiness’}$ $u'(c) = \text{‘hunger’}$

(d) The Level of $u(c)$ doesn’t matter for anything. It’s ok if $u(c)$ is negative. (-20° is warm in Alaska) Maximizing $u(c)$ gives the same $c$ as maximizing $\{u(c) + 10\}$.

(e) $u(c)$ rises, $u'(c) > 0$. People always want more.

(f) $u'(c)$ declines, $u(c)$ concave. Hunger declines as you eat more.

(g) $\beta$ is a number, typically 0.96 or so. People prefer money now to later, they dislike delay. ($\beta$ has nothing to do with CAPM beta – two uses of the same letter.) $\beta$ captures their impatience.

(h) $c_{t+1}$ is random; you don’t know at time $t$ how things will turn out, what $c_{t+1}$ will be.

(i) Thus utility of random consumption is $E_t u(c_{t+1})$.

   i. Example: 50/50 bet. $E[u(c)] = \frac{1}{2} u(\bar{c} + x) + \frac{1}{2} u(\bar{c} - x)$

(j) Concavity and expected utility → people dislike risk. They would prefer to give up some consumption for sure in order to avoid a 50/50 bet.
In math
\[ E[u(c)] = \frac{1}{2}u(\bar{c} + x) + \frac{1}{2}u(\bar{c} - x) < u[E(c)] = u\left(c + \frac{1}{2}x - \frac{1}{2}x\right) \]

(k) *Concave \( u(c) \) (\( c \) vs. \( u(c) \) graph) induces curved indifference curves over \( U(c_t, c_{t+1}) \) (\( c_t \) vs. \( c_{t+1} \) graph) that look like the apple-orange case. (See above \( c_t \) vs. \( c_{t+1} \) graph) People are less and less willing to give up some \( c_t \) to get more \( c_{t+1} \). (We’ll explore this on the problem set) If \( u(c) = kc \), linear, then the \( c_t \) vs. \( c_{t+1} \) indifference curves are linear too. In this case people are very willing to substitute consumption over time and take risk. As both \( u(c) \) and indifference curves become more curved, people are less and less willing to take risks / move consumption over time when prices scream at them to do so. (E.g. fall 08)

4. A More useful functional form generalizes log; it lets you have more or less curved function
\[ u(c) = \frac{c^{1-\gamma}}{1-\gamma} \rightarrow u'(c) = c^{-\gamma} \]
\[ \gamma = 1 : u(c) = \ln(c) \rightarrow u'(c) = \frac{1}{c} \]
*The coefficient of relative risk aversion

\[ \gamma = -\frac{cu''(c)}{u'(c)} \]

measures how curved the utility function is, and thus how resistant people are to taking risks and to substituting consumption over time. The power utility function is also known as constant relative risk aversion

\[ u'(c) = c^{-\gamma} \]
\[ u''(c) = -\gamma c^{-\gamma-1} \]
\[ -\frac{cu''(c)}{u'(c)} = -\frac{-\gamma c^{-\gamma-1}}{c^{-\gamma}} = \gamma! \]

5. What is the value of payoff \( x_{t+1} \) to an investor with a utility function?

(a) Bottom line: The investor consumes \( c_t, c_{t+1} \). In order to get one unit more of the payoff \( x_{t+1} \), the investor is willing to pay

\[ p_t = E_t \left[ \beta \frac{u'(c_{t+1})}{u'(c_t)} x_{t+1} \right] \]

Hence \( p_t \) is the value of \( x_{t+1} \) to the investor. Meaning, He would be willing to pay \( p_t \) for a little bit extra of the payoff \( x_{t+1} \).
(b) This formula is an answer to the question we’re asking. Given the properties of the payoff \( x_{t+1} \) and some information about the investor, you can calculate the value \( p_t \). For example, a log investor

\[
p_t = E_t \left[ \frac{\beta c_t}{c_{t+1}} x_{t+1} \right]
\]

If you know \( c_t \) and have an idea of the risks the consumer faces (\( c_{t+1} \)) and the asset payoff \( x_{t+1} \)

(c) This is IT. All asset pricing and portfolio theory flows from this one equation. Yes, everything. CAPM, FF3F, option pricing, bond pricing, and portfolio theory.

(d) Why? With a \( \xi \) more benefit

\[
U = u(c_t) + \beta E_t [u(c_{t+1} + \xi x_{t+1})] \approx u(c_t) + \beta E_t [u(c_{t+1}) + u'(c_{t+1}) \xi x_{t+1}] \\
\text{benefit: } \Delta U \approx \beta E_t [u'(c_{t+1}) x_{t+1} \xi]
\]

If the price (per unit) is \( p_t \), the investor pays \( p_t \xi \), and the loss is

\[
U = u(c_t - p_t \xi) \approx u(c_t) - u'(c_t) p_t \xi \\
\Delta U \approx u'(c_t) p_t \xi.
\]

Willing to pay (value) is thus where marginal benefit is just equal to marginal cost,

\[
p_t u'(c_t) = \beta E_t [u'(c_{t+1}) x_{t+1}]
\]

### 9.2 Discount Factor

1. It’s useful to separate

\[
m_{t+1} = \beta \frac{u'(c_{t+1})}{u'(c_t)} \\
p_t = E_t [m_{t+1} x_{t+1}] \\
p = E(mx)
\]

(When I leave off subscripts, understand \( p_t, m_{t+1}, x_{t+1} \))

2. *Why?* We’ll see many models of \( m \), (not just \( \beta u'(c_{t+1})/u'(c_t) \)) and many different expressions of \( p = E(mx) \). Example: in option pricing we find \( m \) to price stock, bond, rather than from consumption data. All of asset pricing theory and practice comes down to various tricks for finding \( m \) that are useful in specific applications. This includes stocks, bonds, options, fx, real investment valuation, etc..

3. In our Example

\[
u = \ln(c) : m_{t+1} = \beta \frac{c_t}{c_{t+1}} \\
u = c^{1-\gamma} : m_{t+1} = \beta \left( \frac{c_{t+1}}{c_t} \right)^{-\gamma}
\]
4. A good approximation,

\[ m_{t+1} = \beta \left( \frac{c_{t+1}}{c_t} \right)^{-\gamma} = e^{-\delta} e^{-\gamma \Delta c_{t+1}} \approx 1 - \delta - \gamma \Delta c_{t+1}; \]

\[ m_{t+1} \approx 1 - \delta - \gamma \Delta c_{t+1} \]

where

\[ \beta = e^{-\delta} \approx 1 - \delta \]

\[ 0.096 \approx 1 - 0.04 \]

\[ \Delta c_{t+1} = \log \left( \frac{C_{t+1}}{C_t} \right) \]

\[ 0.10 \approx \log \left( \frac{1.10}{1.00} \right) \]

i.e. \( \delta = 0.05 \) for 5\% discount rate; \( \Delta c_t = 0.01 \) for 1\% consumption growth rate. (If you know continuous time, these approximations are all much easier in that framework.)

5. \( m, u(c) \) measure “hunger.” \( m, u' \) are high when \( c \) is low. *Hunger is higher in bad times*

6. Let’s *use* this theory...

9.3 **Classic issues in finance**

1. Interest rate. Pay \$1, Get \( R^f \) (e.g. 1.03). Thus

\[ 1 = E(mR^f) = E(m)R^f \]

\[ R^f_t = \frac{1}{E(t(m_{t+1})} \]

Risk free rate and consumption.

\[ R^f_t = \frac{1}{\beta \left[ E_t \left( \frac{c_{t+1}}{c_t} \right)^{-\gamma} \right]} \approx \frac{1}{1 - \delta - \gamma E_t(\Delta c_{t+1})} \approx 1 + \delta + \gamma E_t(\Delta c_{t+1}) \]

*When/where are interest rates high?*

(a) Interest rates are higher if people are impatient (high \( \delta \), low \( \beta \)).

(b) *Graph:

(c) Interest rates are higher if expected consumption *growth* is high.

\[ R^f \approx 1 + \delta + \gamma E_t(\Delta c_{t+1}) \]

i. If people know they will be richer in the future, you must offer high rates to get them to consume less now and save.

ii. No surprise, interest rates are higher in booms than in recessions. This is often attributed to the Federal Reserve, but the Fed really has little choice in the matter.

(d) \( R^f \) is more *sensitive* to consumption growth if \( \gamma \) is high.

\[ R^f \approx 1 + \delta + \gamma E_t(\Delta c_{t+1}) \]
i. Boom: high $E_t (\Delta c_{t+1})$, high $R^f$. Bust: low or negative $E_t (\Delta c_{t+1})$, low $R^f$.

ii. How much does $R^f$ vary? How much must you offer people to postpone consumption? $1/\gamma$ = “elasticity of intertemporal substitution.” *Graph:

\[ R^f_t \approx 1 + \delta + \gamma E_t (\Delta c_{t+1}) - \frac{1}{2} \gamma^2 \sigma_t^2 (\Delta c_{t+1}) \]

The approximation is exact in continuous time or with lognormal consumption growth. Higher *volatility* of consumption growth makes interest rates *lower*. In more uncertain
times, people want to save more. More demand for “precautionary savings” drives interest rates down.

(f) Does $\Delta c$ adjust to $R^f$ or does $R^f$ adjust to $\Delta c$? Both really. For you and me, $R^f$ is given, and $\Delta c$ adjusts. For a small open country, the same. For a closed economy or the world, to some extent it is $R^f$ that adjusts.

(g) *Important point: we are studying a market after it has settled down, and everyone has bought as much as they want.

2. Risk and betas

(a) Excess returns or zero-cost portfolios have price 0, payoff = return $R^e = R^i - R^f$, so $p = E(mx)$ reads

$$0 = E_t(m_{t+1}R^e_{t+1})$$

(b) Trick: Use the definition of covariance (we want betas)

$$cov(m, x) \equiv E(mx) - E(m)E(x)$$

$$\Rightarrow E(mx) = cov(m, x) - E(m)E(x)$$

Then

$$0 = E(mR^e)$$

$$E(m)E(R^e) = -cov(m, R^e)$$

$$E(R^e) = -R^f cov(m, R^e)$$

To betas

$$E(R^e) = \frac{cov(m, R^e)}{var(m)} \left[-R^f var(m)\right]$$

$$E(R^e) = \beta_{R^e, m} \times \lambda_m$$

In terms of consumption

$$E(R^e) \approx -cov(1 - \delta - \gamma \Delta c, R^e)$$

$$E(R^e) \approx \gamma cov(R^e, \Delta c).$$

To betas

$$E(R^e) \approx \frac{cov(R^e, \Delta c)}{var(\Delta c)} \times (\gamma var(\Delta c))$$

$$E(R^e) \approx \beta_{R^e, \Delta c} \times \lambda_{\Delta c}$$

(c) INTUITION.

i. Assets that covary negatively with $m$, hence positively with consumption growth must pay a higher average return.
ii. High $E(R^e) \leftrightarrow$ low price.

(d) Given volatility (price must go up or down at some point), price (risk-discount) depends on when good/bad performance comes. *Average returns are high if beta on m or $\Delta c$ is large. Stocks must pay high returns if they tend to go down in bad times.*

i. Price is depressed if a payoff is low in bad times, when “hungry” (high $m$, low $\Delta c$) - High $E(R^e)$.

ii. Price is enhanced if a payoff is high in bad times, when “full” (Insurance “costs too much”) - Low $E(R^e)$.

(e) Higher $\gamma$ implies larger price effects. $\gamma = \text{coefficient of risk aversion}$.

(f) Variance $\sigma(R^e)$ of an individual asset does not matter, only its covariance with $m$ (e.g. consumption growth) matters. First of many totally (initially) counterintuitive theorems of finance!

(g) *Why is this counterintuitive? This holds after you have taken as much as you want, and you adjust $c$ and $m$! You think of a “marginal” change, buying one more share. Variance does matter to the prospective value of a big bite.

(h) Note $E(R^i) < R^f$ is possible! “Insurance”

(i) Note in the consumption model, the market price of risk $\lambda$ is higher if a) risk aversion is higher or b) if macro volatility is higher.

3. *Valuing risk; expressing risk as a price discount rather than a return premium.*

(a) Use the definition of covariance

$$cov(m, x) \equiv E(mx) - E(m)E(x).$$

Thus,

$$p = E(mx) = E(m)E(x) + cov(m, x)$$

$$p = \underbrace{\frac{E(x)}{R^f}_\text{present value (time)}} + \underbrace{cov(m, x)}_{\text{risk correction}}$$
(b) With the approximation

\[ m_{t+1} \approx 1 - \delta - \gamma \Delta c_{t+1} \]

we get

\[ p_t \approx \frac{E_t(x_{t+1})}{R_t} - \gamma \text{cov}(x_{t+1}, \Delta c_{t+1}) \]

Price is lower if you do well in good times and price is higher if you do well in bad times.

(c) Terminology: \( m = \text{stochastic discount factor} \). Why? Remember the old discount factor,

\[ p^i_t = \frac{E(x^i_{t+1})}{R^i_t} \]

i. \( 1/R^i_t = \text{discount factor} \). But it’s different for each asset \( i \). (Reminder: 35200: use CAPM for \( R^i_t \)).

diff

ii. Our version

\[ p^i_t = E(m_{t+1} x^i_{t+1}) \]

\( m \) is stochastic (unknown at \( t \), inside \( E \)), and the same for all assets. (Different covariance of \( m, x^i \) gives different risk adjustments for different assets.)

4. Mean-variance frontier

\[
\begin{align*}
E(m R^e) &= 0 \\
E(m) E(R^e) &= -\text{cov}(m, R^e) \\
E(m) E(R^e) &= -\sigma(m) \sigma(R^e) \rho_{m,R^e} \\
\frac{E(R^e)}{\sigma(R^e)} &= -\frac{\sigma(m)}{E(m)} \rho_{m,R^e} \\
\frac{\|E(R^e)\|}{\sigma(R^e)} &\leq \frac{\sigma(m)}{E(m)} \approx \gamma \sigma(\Delta c)
\end{align*}
\]

(a) The mean and standard deviation of asset returns must lie inside a cone-shaped region.

(b) The slope of the mean-variance frontier – the reward for taking risk – is higher if macroeconomic risk is higher or if risk aversion is higher.

(c) All assets on the mean-variance frontier are perfectly correlated with each other. They are exactly equivalent to leveraged positions in the “market portfolio”.

(d) We can express all asset pricing with respect to a return on the mean-variance efficient portfolio. \( \rho_{m, R^{mv}} = 1 \) means

\( R^{mv} = a + bm. \)

Hence,

\[
\begin{align*}
E(R^{ei}) &= \beta_{i,m} \times \lambda_m \\
&= \frac{1}{b} \beta_{i,R^{mv}} \times \lambda_m \\
&= \beta_{i,R^{mv}} \times \left( \frac{\lambda_m}{b} \right) \\
&= \beta_{i,R^{mv}} \times \lambda_{mv}
\end{align*}
\]
If you can find a return on the mean-variance efficient portfolio you can price any asset. This statement does not assume returns are normal, and applies to any asset – stocks, bonds, options, fx etc. Roll theorem:

\[ E(R^{e}) = \beta_{R^{e}, R^{mv}} \lambda_{mv} \leftrightarrow R^{emv} \text{ is on the mvf.} \]

5. Predictable returns? (Why do D/P regressions work? Why does \( E_t(R^{e}_{t+1}) \) vary over time?)

\[ E_t(R^{e}_{t+1}) \approx \gamma \text{cov}(R^{e}_{t+1}, \Delta c_{t+1}) \]
\[ \approx \gamma \sigma_t(R^{e}_{t+1}) \sigma_t(\Delta c_{t+1}) \rho_t(R, \Delta c) \]

(a) Expected returns may vary over time if risk \( \sigma_t(\Delta c_{t+1}) \), \( \sigma_t(R_{t+1}) \) or risk aversion (\( \gamma \)) vary over time.

(b) Let’s look at the Sharpe ratio

\[ \frac{E_t(R^{e}_{t+1})}{\sigma_t(R_{t+1})} \approx \gamma \text{cov}(R^{e}_{t+1}, \Delta c_{t+1}) \]
\[ \approx \gamma \sigma_t(\Delta c_{t+1}) \rho_t(R^{e}, \Delta c) \]

(c) Can \( \gamma, \sigma_t(\Delta c_t) \) vary day to day? Not plausible, which is why “efficient markets” looks dimly at high frequency trading.

(d) Can \( \gamma, \sigma_t(\Delta c_t) \) vary with business cycles and longer? Possibly! Bottom of a recession has high \( \sigma_t(\Delta c_{t+1}) \), high \( \gamma \). This is why D/P regressions do not imply “inefficiency.”

(e) Why does \( \gamma \) tend to rise in a recession? A utility function \( u(c) = \frac{1}{1-\gamma} (c - x)^{1-\gamma} \) has this property. \( x \) can represent leverage, or a habitual level of consumption below which you do not want to fall. As \( c \) falls to \( x \), the risk aversion coefficient based on \( u'' \) (not \( \gamma \)) rises. Campbell and Cochrane “by force of habit” account for lots of puzzles this way.


\[ U = u(c_t) + \beta E_t u(c_{t+1}) + \beta^2 E_t u(c_{t+2}) + ... = E_t \sum_{j=1}^{\infty} \beta^j u(c_{t+j}) \]
\[ p_t = E_t \sum_{j=1}^{\infty} \beta^j \frac{u'(c_{t+j})}{u'(c_t)} d_{t+j} = E_t \sum_{j=1}^{\infty} m_{t,t+j} d_{t+j} \]

9.4 CAPM and Multifactor models

1. Big picture: We want to understand the foundations of the CAPM and FF3F model.

\[ E(R^{e_i}) = \beta_{i, R^m} \lambda_{R^m} \]
\[ E(R^{e_i}) = b_i \lambda_{r^mrf} + h_i \lambda_{mhl} + s_i \lambda_{smb} \]

2. Math: is minimal.
(a) From
\[ E(R^{ei}) = \frac{1}{E(m)} \text{cov}(R^{ei}, m) = \beta_{R^{ei}, m} \lambda_m \]
we do
\[ m = a - b \times f \]
to get
\[ E(R^{ei}) = \beta_{R^{ei}, f} \lambda_f \]

Algebra:
\[ E(R^{ei}) = \frac{1}{E(m)} \text{cov}(R^{ei}, f \times b) = \frac{1}{E(m)} \text{cov}(R^{ei}, f) \times b = \frac{\text{cov}(R^{ei}, f)}{\text{var}(f)} \times \left[ \frac{b \text{var}(f)}{E(m)} \right] \]

(b) Intuition: "if low \( f \) indicates bad times, when people are hungry, then assets which pay off badly in times of low \( f \) must have low prices and deliver high expected returns."

(c) Example 1,
\[ m_{t+1} = 1 - \delta - \gamma \Delta c_{t+1} \iff E(R^{ei}) = \beta_{i, \Delta c} \lambda_{\Delta c}. \]

(d) Example 2, CAPM, \( f = R^{em} \)
\[ m_{t+1} = a - b R_{t+1}^{em} \iff E(R_{t+1}^{ei}) = \beta_{R_{t+1}^{ei}, R_{t+1}^{em}} \lambda \]

(e) Example 3. The same algebra goes through with multiple factors.
\[ m_{t+1} = a - b_1 f_{t+1} - b_2 f_{2t+1} \iff E_t(R_{t+1}^{ei}) = \beta_{R_{t+1}^{ei}, R_{t+1}^{em}} \lambda_1 + \beta_{R_{t+1}^{ei}, R_{t+1}^{em}} \lambda_2 \]
\[ m_{t+1} = a - b_1 f_{t+1} \iff E_t(R_{t+1}^{ei}) = \beta_{R_{t+1}^{ei}, f} \lambda_1 \]

(f) Issue: What do we get to use for \( f \)? “Derivation:” find a story for \( m = a - b f \) and you’re done. For example, CAPM: What assumptions do we make to substitute \( R_{t+1}^{m} \) for \( \Delta c_{t+1} \)?

3. Motivation: Why don’t we use Government total nondurable consumption data + power utility function?
\[ m_{t+1} = \beta \left( \frac{c_{t+1}}{c_t} \right)^{-\gamma} \]
Or the linearized version \( m_{t+1} \approx 1 - \delta - \gamma \Delta c_{t+1} \) leading to
\[ E(R^{ei}) = \gamma \text{cov}(R^{ei}, m) = \beta_{i, \Delta c} \times \lambda_{\Delta c} \]
Alas, it doesn’t work too well. It may end up being good for academic “explanation” but not for practical risk management.

(a) Equity premium puzzle (book)
(b) Book: FF 25

Figure 2.4. Mean excess returns of 10 CRSP size portfolios versus predictions of the power utility consumption-based model. The predictions are generated by \(-R' \text{cov}(m, R')\) with 
\(m = \beta (c_{t+1}/c_t)^{1-\gamma}\). \(\beta = 0.98\) and \(\gamma = 2.41\) are picked by first-stage GMM to minimize the sum of squared pricing errors (deviation from 45° line). Source: Cochrane (1996).

(c) Maybe it’s not so bad after all? Jagannathan and Wang 2005 “Consumption Risk and the Cost of Equity Capital” (NBER working paper)

Figure 1: Annual Excess Returns and Consumption Betas

Plot figure of average annual excess returns on Fama-French 25 portfolios and their consumption betas. Each two digit number represents one portfolio. The first digit refers to the size quintiles (1 smallest, 5 largest), and the second digit refers to the book-to-market quintiles (1 lowest, 5 highest). Annual excess returns and consumption betas are reported in previous table.
(d) However, this result isn’t perfect either.

i. It uses the linear approximation \( m_{t+1} = 1 - \delta - \gamma \Delta c_{t+1} \) not the real \( m = \beta \left( \frac{c_{t+1}}{c_t} \right)^{-\gamma} \)

ii. It needs a risk aversion\(^9\) of 31! It doesn’t solve the equity premium, and nonlinearities matter a lot for \( \gamma = 31 \)

iii. This only helps on the difference between stock and bond returns. It doesn’t do a good job on the level of the risk free rate.

iv. It only works in annual data, only Q4-Q4. (Christmas?)

(e) JC View: This is a very useful result, giving us some hope that consumption is the underlying explanation for returns! Yahoo! But it will not displace factor models for practice soon. If you want to risk-adjust monthly returns, you need something more practical

\(^9\)Consider the slope coefficient, \( \lambda_1 \) in the cross sectional regression equation given by:

\[
R_{t,t+4} = \lambda_0 + \lambda_1 \Delta c + \epsilon_{t,t+4}
\]

If the standard consumption-based asset pricing model holds, the intercept, \( \lambda_0 = 0 \) and the slope coefficient, \( \lambda_1 = \gamma \frac{\text{var}(\Delta c_{t+4})}{[1 - \gamma E(\Delta c_{t+4}) - 1]} \), where \( \gamma \) denotes the coefficient of relative risk aversion. The estimated slope coefficient, \( \hat{\lambda}_1 = 2.56 \), therefore corresponds to an implied coefficient of relative risk aversion of 31." (p. 10-11)
4. So, a direct measure of \( u'(c) \) is not working well. What to do?

(a) *Idea 1: ("Absolute pricing") Find other proxies, data sources for consumption, marginal utility (CAPM, ICAPM, multifactor models).

(b) *Idea 2: ("Relative pricing") Find discount factors that price one set of assets by construction. Don’t ask why those assets are priced right, but use them to price other things. (APT, Black-Scholes, Term structure).

5. Idea 1. What variables \( f \) other than consumption itself might indicate bad times – high marginal utility; hunger?

\[
m_{t+1} = \beta \frac{u'(c_{t+1})}{u'(c_t)} \approx a - b' f_{t+1}
\]

Think of the linear \( m \) as a local (Taylor) approximation, just as we already did like

\[
\beta \left( \frac{C_{t+1}}{C_t} \right)^{-\gamma} \approx 1 - \delta - \gamma \Delta c_{t+1}.
\]

(a) How about the determinants of consumption: wealth, income, news?

(b) **What can be \( f \)?** Low market return = low wealth, low consumption. CAPM

\[
m_{t+1} = a - b R^m_{t+1}
\]

Equivalently,

\[
E(R^i) - R^f = \beta_{i,R^m} \times [E(R^m) - R^f]
\]

(Notation: \( R^m \) is traditional for “market.” It has nothing to do with the discount factor \( m \).)

(c) **What can be \( f \)?** (My favorite). Labor / proprietary income. Direct: lost job? Consumption will decline. Indirect: Don’t want stocks to decline when you just lost your job.

(d) **What can be \( f \)?** News about future investment opportunities. News that \( E_{t+1}(R_{t+2}) \) is low is bad news for a long-lived investor; long run wealth will now be lower. Consumption goes down when \( E_{t}(R_{t+1}) \) goes down. Indirect: We want assets that pay well when this happens. ICAPM

\[
m_{t+1} = a - b R^m_{t+1} - b_2 f_{t+1}
\]

\[
E(R^i) - R^f = \beta_{i,R^m} E(R^m - R^f) + \beta_{i,f} \lambda_f
\]

Examples: Changes in D/P, interest rates. (Note: this story requires \( \gamma \neq 1 \), as per problem set question, and the sign of \( \lambda_f \) depends on \( \gamma \)).

(e) **What can be \( f \)?** Other macroeconomic variables. Investment, GDP, interest rates, unemployment, inflation, etc. have all been used. Motivation: They affect consumption.

(f) **What can be \( f \)?** Most of this is pretty discouraging. We wanted something more practical than consumption, and here we are with stuff even harder to measure. Where do portfolios like FF3F come from? A: Mimicking portfolios, so you don’t need data on labor income, news, macro variables. The portfolio of assets formed by a regression of any \( m \) on returns is also a discount factor.
i. Why? Think of running a regression of $m$ on all returns,

$$m = a + b'R^e + \varepsilon = a + \sum_i b_i R^{ei} + \varepsilon$$

By construction,

$$E(\varepsilon R^e) = 0$$

Thus,

$$0 = E(m R^e)$$

$$0 = E \left[ (a + b'R^e + \varepsilon) R^e \right]$$

$$0 = E \left[ (a + b'R^e) R^e \right]$$

but $b'R^e$ is of course a portfolio. Thus, an excuse for portfolios as risk factors.

ii. Bigger picture, and we’ll do it again and again: the right hand side of a regression is a portfolio. We can just run regressions to construct the “optimal hedge”!

iii. Fama and French say hml and smb are “Mimicking portfolios for state variables of concern to investors.” This is what they mean! Note most people (like FF) cite this, but do not write what fundamental $m$, variables, they have in mind nor do they run the regression.

(g) What are we doing? These are still the consumption model. They use other variables, determinants of consumption or proxies for (badly measured?) consumption. With good data, $u(\cdot)$, consumption should reveal all the information in these proxies; if you’re unhappy, the market went down, you lost your job, had bad news, etc. etc. you don’t go out to eat. In theory, given $c$, we do not need to also measure determinants of $c$, and in fact $u'(c)$ should drive out all other factors.

(h) The real question is what can’t be $f$? That’s the heart of a derivation. The CAPM says not just “the market matters” but “the market is the only thing that matters.” This is the key to “deriving the CAPM.” We need a story for

$$m = \beta \frac{u'(c_{t+1})}{u'(c_t)} = a - bR^W_{t+1}$$

and only that factor – nothing else but market returns drives consumption. The actual assumptions you need to get that are pretty unrealistic: nobody has a job, and either people live only two periods, have quadratic utility, or returns are not correlated over time.

(i) *All of these are done carefully, with lots of equations, not just my verbal summaries. A precise CAPM derivation, so you can get one glimpse at the sausage factory of asset pricing theory. We need $m = a - bR^W_{t+1}$, and only that factor. So we need to tie $c$ to wealth: suppose people live 2 periods have no job and live off their portfolio. We need a linear function: suppose utility is quadratic. Then,

$$c_{t+1} = W_{t+1}$$

$$W_{t+1} = R^W_{t+1}(W_t - c_t)$$

$$u(c) = -\frac{1}{2}(c^* - c)^2 \Rightarrow u'(c) = c^* - c$$
\[ m_{t+1} = \beta \frac{u'(c_{t+1})}{u'(c_t)} = \beta \frac{(c^* - c_{t+1})}{(c^* - c_t)} \]
\[ = \beta \frac{(c^* - R^W_{t+1}(W_t - c_t))}{(c^* - c_t)} \]
\[ = \frac{\beta c^*}{(c^* - c_t)} - \frac{\beta (W_t - c_t)}{(c^* - c_t)} R^W_{t+1} \]
\[ m_{t+1} = a_t - b_t R^W_{t+1} \]

We have “derived the CAPM.”

i. Note this is the consumption model. We just substituted a determinant of consumption (market return) for consumption itself.

ii. Artificial? Yes. Why? 1) Two periods to make the market return the only determinant of consumption, 2) Quadratic to make \( m \) linear in the market return.

(j) *Another “CAPM derivation.” Assume log utility, but infinite life. With log utility the wealth portfolio claim is

\[ p_t = E_t \sum_{j=1}^{\infty} \beta^j \left( \frac{c_{t+j}}{c_t} \right)^{-\gamma} c_{t+j} \]

\[ \frac{p_t}{c_t} = E_t \sum_{j=1}^{\infty} \beta^j \left( \frac{c_{t+j}}{c_t} \right)^{-\gamma} \left( \frac{c_{t+j}}{c_t} \right) \]

\[ \gamma = 1 : \frac{p_t}{c_t} = E_t \sum_{j=1}^{\infty} \beta^j = \frac{\beta}{1 - \beta}. \]

Then, we want to tie \( m \) to \( R^W \)

\[ R^w_{t+1} = \frac{p_{t+1} + c_{t+1}}{p_t} = \frac{(\frac{p_{t+1}}{c_{t+1}} + 1) \frac{c_{t+1}}{c_t}}{\frac{p_t}{c_t}} = \frac{\beta}{1 - \beta} + 1 \frac{c_{t+1}}{c_t} \]
\[ = \frac{\beta + 1 - \beta c_{t+1}}{\beta c_t} = \frac{1}{\beta} \frac{c_{t+1}}{c_t} \left( \frac{1}{m_{t+1}} \right) \]
\[ m_{t+1} = 1/R^w_{t+1} \approx a - b R^W_{t+1} \]

Thus, we can replace \( R^w_{t+1} \) for \( \Delta c_{t+1} \).

i. OK, derives the CAPM, but note the implicit (and soon forgotten) predictions about consumption

A. \( c_{t+1}/c_t = \beta R^W_{t+1} \) Consumption growth tracks market returns perfectly! \( \sigma(\Delta c) = 18\% \), \( corr(\Delta c, R^m) = 1?! \) It implies consumption is extremely volatile or market returns are not

B. \( p/c \) is a constant??

ii. The assumptions are extreme. OK, so we’ve been waiting for years for practical alternatives to CAPM. It was not there by deep assumptions, it was there because it worked so well. Now that we look at the real theory, we should not be at all surprised that it doesn’t work so well. What’s amazing is that it took so many years to find practical multi-factors!
iii. Multifactor model derivations follow the same idea. They find “assumptions” to substitute out for $\Delta c_{t+1}$ Since no one checks these assumptions in practice, I see no need to drag you through them.

6. I emphasized link to consumption, e.g. $f$ goes down implies consumption goes down. It’s also important to understand the Portfolio logic for multifactor models: Why should something more than market return show up as a factor?

(a) Example:

i. Stocks A,B have the same mean, variance, beta. In a recession (bad times, bad consumption), for given level of the market return, A goes up while B goes down.

ii. According to CAPM should you care?

iii. Do you want A or B?

iv. $\Rightarrow$ People want more A $\Rightarrow$ Price of A goes up $\Rightarrow$ Expected returns of A go down.

v. $\Rightarrow$ People want less B $\Rightarrow$ Price of B goes down $\Rightarrow$ Expected returns of B go up.

vi. $\Rightarrow$ Expected returns depend on recession sensitivity as well as market sensitivity. In equations,

\[ E(R^e_i) = \beta_{i,m} \lambda_m + \beta_{i,\text{recess}} \lambda_{\text{recess}} \]

vii. Equivalent to a “formal” derivation in which people have outside income, privately held businesses, or news about the future.

(b) Important facts

i. $\lambda$ can be negative. $\lambda$ typically is negative for “bad” factors like oil price rise, positive for “good” factors like rise in market, consumption growth.

ii. The size of $\lambda$ is determined by how much the average investor wants to avoid that risk; larger for more important risks and higher risk aversion. “Marginal” investor is a misnomer.

iii. It must be aggregate risk to affect prices. Risks that as many people like as dislike are merely transferred through assets. Oil sensitive stocks: Texas sells, we buy, no price effect.

7. Idea 2. APT (stocks), Black-Scholes (options), Term structure models (Bonds).

(a) Absolute vs. Relative pricing. Rather than a “theory of everything,” we extend known prices to value something else. (An idea taken to its limit in Black-Scholes, where we can price options from stock and bonds.)

(b) Surely this is the right approach for most practical applications. You don’t care why the S&P500 does what it does, you want to know “can this manager (strategy, etc.) beat the S&P500?” Analogy: What’s the value of a burger at McDonald’s? 1) Cost to raise a cow, ... 2) What’s a burger at Wendy’s? (and then adjust). This is less useful for the grand “explain” project, but much more useful in practice.

(c) APT. Suppose you have $N$ test assets $R^e_i$, excess returns ($R^e_i = R^i - R^f$, for example the FF 25) and $K$ factors $f^i$, also excess returns (for example rmrf, hml, smb). Run a time-series regression

\[ R^e_{it+1} = \alpha_i + \beta_{i1}f^1_{t+1} + \beta_{i2}f^2_{t+1} + \epsilon^i_{t+1}. \]
We want to conclude
\[ E(R^{ei}) = \beta_{i1}\lambda_1 + \beta_{i2}\lambda_2 = \beta_{i1}E(f^1) + \beta_{i2}E(f^2) \]
i.e. \( \alpha_i = 0 \). What does it take to reach this conclusion?

(d) “Exact APT.” If \( \varepsilon = 0 \), \( \alpha = 0 \) or there is arbitrage. Go long \( R^{ei}_{t+1} \), go short \( \beta_{i1}f^1_{t+1} + \beta_{i2}f^2_{t+1} \). The portfolio return is
\[ R^{ep}_{t+1} = R^{ei}_{t+1} - \left( \beta_{i1}f^1_{t+1} + \beta_{i2}f^2_{t+1} \right) = \alpha_i \]
This is not random! It has zero cost, and a positive return!
i. Example: A fund with zero tracking error, return 10 bp above S&P500 index? Short index, long fund, earn 10bp for free.
ii. (Example: Black-Scholes option pricing.)
iii. Arbitrage, let’s buy!...No, if we’re describing a market at equilibrium, then arbitrage is gone, so we expect to see \( \alpha = 0 \). (Emphasis in all of this, we’re describing what a market at equilibrium should look like.)

(e) “Approximate APT.” If \( \varepsilon \) is “small,” \( \alpha \) should be “small”. Large \( \alpha \) does not imply arbitrage, but maybe a very good deal? Let’s look at the portfolio long the stock and short its factor content. This gives the alpha but now with some error too,
\[ R^{ep}_{t+1} = R^{ei}_{t+1} - \left( \beta_{i1}f^1_{t+1} + \beta_{i2}f^2_{t+1} \right) = \alpha_i + \varepsilon_{t+1} \]
\[ E(R^{ep}_{t+1}) = \alpha_i; \quad \sigma(R^{ep}_{t+1}) = \sigma(\varepsilon_i) \]
\[ \text{Sharpe} = \frac{E(R^{ep}_{t+1})}{\sigma(R^{ep}_{t+1})} = \frac{\alpha_i}{\sigma(\varepsilon_i)} \]
(Coming: Optimal investment in that portfolio?
weight = \( \frac{\alpha_i}{\gamma \sigma^2(\varepsilon_i)} \)
but you don’t know that yet.)

Conclusion: If \( \alpha >> \sigma(\varepsilon) \) there are really good deals to be had. Ruling those out (equilibrium, remember?)
\[ \alpha_i < (\text{max surviving Sharpe ratio}) \times \sigma(\varepsilon_i) \]

(f) When is \( \sigma(\varepsilon) \) small?
\[ R^{e}_{t+1} = \alpha + \beta f_{t+1} + \varepsilon_{t+1} \]
\[ R^2 = 1 - \frac{\sigma^2(\varepsilon)}{\sigma^2(R^{e})} \]
The \( R^2 \) in the time series regression should be large. Let’s look at FF table 1...Hey, this is an APT!

(g) Thus, the APT is typically useful for portfolios (FF 25) or managers (of diversified portfolios or with tracking error constraints that give high \( R^2 \)) but not for individual stocks. The APT applies to “well diversified portfolios” (Those with high \( R^2 \))
(h) What have we done? We explained test asset portfolios given factor portfolios. Who says FF rmrf, smb, hml portfolios are priced right? That’s not the question; the question is given those, are the 25 priced right? “Behavioralists” may still be right if hml, smb, are priced wrong.

(i) Bottom line: Multifactor models again. But f are portfolios that do a good job of explaining cross-correlation of asset returns, good $R^2$ in

$$R^i_{t+1} = \alpha_i + \beta_{i1}f_{1,t+1} + \beta_{i2}f_{2,t+1} + \ldots + \varepsilon^i_{t+1}, T = 1, 2, \ldots T$$

for each $i$ not necessarily “proxies for state variables.”

(j) High $R^2$ in the time series regression implies a factor model (APT logic). High $R^2$ in the time series regression is not required for a factor model. Other stories (“state variables”) do not need high $R^2$.

(k) Example. Is the CAPM a factor model or an APT? The real CAPM is a general model that should hold for any $R^2$. The CAPM is also an APT, that under much milder assumptions should still hold for assets (index futures, say) that have very high $R^2$.

8. *Alternative representations. There are lots of ways to write any model!

(a) Every multifactor model can be written as a single factor model. Example

$$E(R^{ei}) = \beta_{i1}\lambda_1 + \beta_{i2}\lambda_2$$

$$\leftrightarrow$$

$$0 = E(mR^e); m = a - b_1f_1 - b_2f_2$$

$$0 = E(mR^e); m = a - (b_1f_1 + b_2f_2)$$

$$\leftrightarrow$$

$$E(R^{ei}) = \beta_{im}\lambda_m$$

Notice $m$ is a single special combination of $f_1$ and $f_2$.

(b) Related: Recall (35000, fun fact above class): that there is always some portfolio $R_{mv}^{emv}$ on the mean-variance frontier, and that it is always true that $E(R^{ei}) = \beta_{i,R_{mv}^{emv}}E(R_{mv}^{emv})$ – a single factor representation using a mean-variance efficient portfolio.

Thus, for example, if the FF3F model is right, there is a combination of rmrf, hml, and smb that is on the mean-variance frontier, and that new portfolio could act as a single factor model.

$$E(R^{ei}) = b_iE(rmrf) + h_iE(hml) + s_iE(smb) \leftrightarrow$$

$$R_{mv}^{emv} = a \times rmrf + b \times hml + c \times smb \leftrightarrow$$

$$E(R^{ei}) = \beta_{i,mv}E(R_{mv}^{emv})$$

(c) Aha, but how do you find a portfolio on the mean-variance frontier? We can cast all our “theories” as statements about how to find such a portfolio.

(d) The CAPM says rmrf is on the mean variance frontier. Thus, multifactor models say rmrf is not on the mean-variance frontier. The average investor gives up some mean/variance to get a portfolio that (say) does not fall so much in recessions. But there still is some other portfolio on the frontier – rmrf plus a bit of hml and smb in the FF model. Again, we get a single factor model with $R_{mv}^{emv} = rmrf + \gamma_1hml + \gamma_2smb$

This model gives the same $\alpha$ but different (less interpretable?) betas.
(e) Why one representation vs. another? Use whatever gives more intuition. For FF3F, looking at separate size and b/m betas is interesting. You could express the exact same result as a single-beta model with one combination of rmrf, smb, hml, but you’d lose intuition. You’d lose “this one is like value” “this one is like small stocks” although $\alpha$ would be the same.

9.5 Asset Pricing Models final comments

1. Notice how far popular empirical specifications (CAPM, FF3F) are from theory!

2. *All models are special cases of the consumption model, with wild implications for consumption. $\sigma^2(\Delta c) = 18\%$ , $\Delta c_{t+1} = r_{t+1}^c$ for example. We keep the implications that work well, and disregard those that don’t. Doesn’t that bother us?

3. Factor model comments

(a) What should matter in factor models are aggregate risks. If a risk only matters to you, the average investor does not shy away from it, assets that covary with it do not get “underpriced” and no expected return arises. For example, a value effect requires more people who lose if value firms go down than growth, or it all offsets.

(b) What’s the right model? It depends what you want to use it for.

i. Deep economic explanation, fighting with behavioral guys over “rational” and “irrational” markets? The factors had better be well tied to real macroeconomic risks. Even hml and smb are pretty tenuous here.

ii. Evaluation of a manager – could I have gotten the same average return with some simple style indices? Now the factors should just be the tradeable style indices you could invest in. It doesn’t matter if they are “right”, it only matters if the manager can beat them.

iii. Risk management/hedging? Models for mean are different than models for variance. The BARRA model has 67 “factors.” Many of them have no premium. You get means right if you ignore them. And you don’t care about alphas! For these purposes $R^2$ does matter, and the “state variable” nature of factors does not matter. You want tradeable, hedgeable factors.

iv. For example, the CAPM was perfectly happy to admit the existence of more “factors”.

$$R_{t+i}^e = a_i + \beta_{im}R_{t+i}^m + \gamma_{i, \text{industry}} + \varepsilon_{it}$$

the $\gamma$ are in there all right, and may increase the $R^2$ of this regression a lot. Industry factors are important for understanding the variance of returns. But if the CAPM is right,

$$E(R_{t+i}^e) = \beta_{im}\lambda_m + \gamma_i \times 0$$

The additional “factor” isn’t “priced.”

v. Seeing if a new expected return strategy is genuinely new, or just new way of getting some known anomaly (value, momentum)? Just like evaluation, it doesn’t matter really how pure the factors are.

4. The future:
(a) Beta should be endogenous — a result of the model, not an input!

(b) The future: We should price stocks like bonds. Price should be on the left, not a sorting variable. Cash flows should be on the right, not betas. Betas: what are we doing “explaining” expected returns in terms of tomorrow’s price? We should be doing this

\[ p_t = E_t \sum_j m_{t,t+j} d_{t+j} \]

not this

\[ E_t(R_{t+1}^e) = \beta_{i,m} \lambda_m = \text{cov} \left( \frac{p_{t+1}^i + d_{t+1}^i}{p_t^i}, m_{t+1} \right) \]

What’s \( p_{t+1}^i \) doing on the right hand side?

5. So is the Fama French 3 Factor model a multifactor model (“state variable of concern to investors”) or an APT?

(a) ICAPM: “hml smb mimicking portfolios for state variables of concern to investors” But what are they? Why does hml have a premium?

(b) APT: “Minimalist interpretation.” High \( R^2 \) in time series regression.

i. Given hml has a premium, high \( R^2 \) means other size and B/M portfolios will follow.

ii. Suppose there is a large spread in mean returns. Suppose the assets don’t move together. Then there are near-arbitrage opportunities. Buy a diversified portfolio of high ER assets and variance goes down as \( \sigma^2 / N \). Thus: Spread in ER \( \rightarrow \) portfolios move together. If they move together, the mean returns will be explained by a “factor” formed as HML, SMB by a coarse sort. (The converse is not true. ER spread \( \rightarrow \) move together. It is not true that move together \( \rightarrow \) spread in ER. For example, industry portfolios in the CAPM.)

iii. Thus, the central puzzle is that HML seems mispriced by CAPM. Given HML mispricing other portfolios follow by APT logic

iv. The central finding of FF is that size, B/M portfolios move together. This survives even if the value premium disappears. Value is a “risk factor” in the sense that if you buy value stocks, no matter how “diversified” you are, you will still keep a risky portfolio, since all the value stocks move down together. They do not show that value is a “risk factor” in the sense that this movement of value stocks corresponds to low consumption growth or other macroeconomic sources of risk. (It might, but they did not happen to show any evidence)

v. Will it still work on non size-B/M portfolios? APT says only if they still give high \( R^2 \).

vi. Thus, momentum is especially puzzling, since \( R^2 \) is decently high the 3 factor model. But, no surprise, momentum portfolios also move together, so again there is no pure arbitrage. If we add a “momentum factor” the \( R^2 \) becomes very high indeed, and the alphas disappear.