

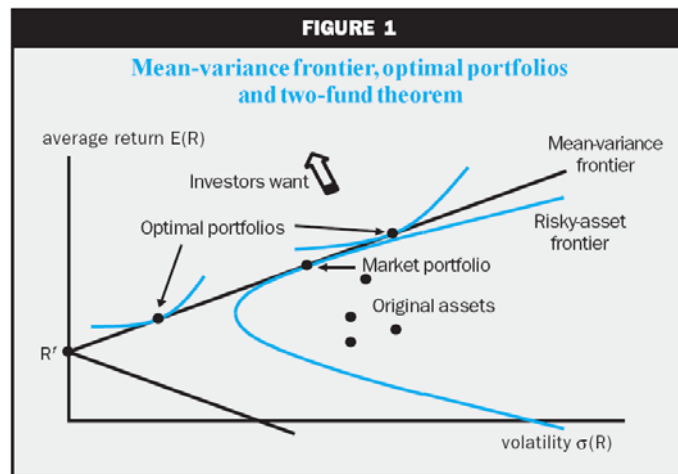
## 31 Week 10. Portfolio theory – detailed notes

1. We've learned a lot of new facts. How do these facts affect portfolios?
  - (a) D/P and related predictability. Should you market-time – buy more when DP is high, less when low?
  - (b) FF: Should you tilt towards value?
  - (c) Anomalies: how much should you allocate to momentum, hedge funds?
  - (d) You are a hedge fund. How should you exploit a new idea, with minimum variance?
2. Outline
  - (a) Mean-variance
  - (b) Multifactor portfolios (value etc.)
  - (c) Outside income, labor income
  - (d) Taking advantage of predictability
  - (e) Options
  - (f) Doubts
  - (g) (Contingent claims approach)
  - (h) (Taming wild weights, Bayesian portfolio theory)
3. Notes:
  - (a) There is a bit of math here. In one sense it's unavoidable – we need to actually end up with numbers, how much do you allocate to each security. But many people don't understand the formulas they're using and wildly misuse portfolio theory. So you have to understand “what question is this the answer to?” and “does that question make any sense?”
  - (b) In general here's the attitude I'd like to advocate. We do not hope for a black box that will magically tell you exactly what portfolio to hold. We can calculate optimal portfolios in some very simplified versions of the real world. Those calculations can be very surprising, and yield intuition and stories you had not thought about. Those stories are a good guideline for the much tougher real world job; they help build judgment.
  - (c) Where we are: Classic portfolio theory assumes that returns are independent over time and the CAPM holds. We know that isn't true. Yet even highly sophisticated quant funds – funds whose whole reason for being is that returns are *not* independent over time and there *are* many non-market premiums to chase – use one-period mean-variance optimization! Implementing modern portfolio theory in a usable way is low-hanging fruit.
4. *The average investor holds the market portfolio.*

- (a) No portfolio advice can apply to everybody. For everyone who overweights  $x$ , someone else must underweight  $x$ . Everyone seems to forget this fact! Portfolio theory is usually calculated as “how much do you want to buy given prices,” a classic demand curve question, and as such ignoring the question “why doesn’t everyone buy more and drive prices up?” Like “tomatoes are on sale, let’s buy a lot of them.”
- i. For example, *we can’t all rebalance*. Yet it is standard advice to rebalance, so you hold (say) 60/40. If the market goes up so market weights are 70/30, who buys if you sell?
  - ii. For everyone who overweights stock  $i$ , and profits, someone must underweight stock  $i$  and lose relative to just holding the value-weighted index. In that sense, everything but holding the index is a zero sum game.
- (b) In an important sense, then, the issue is “What is the *question* to which ‘hold the market’ is the *answer*, despite all the new facts we have learned?”
- (c) With that in mind, it’s useful to phrase portfolio advice in terms of how you are *different* from the average, to defend weights *different* from market weights, to understand not just the apparently good prices, but what equilibrium generates those prices. If you think you’re different because you’re smarter, remember that everyone thinks that, and half are by definition deluded! Not just “tomatoes seem cheap, I’ll buy some,” but “I like tomatoes more than the average person, so it makes sense that the price is low and other people aren’t buying.”

### 31.1 Mean variance portfolio reminder

Reminder from your previous classes



1. The optimal portfolio is the best combination of what investors want (utility, indifference curve) and what they can get (budget constraint, mean-variance frontier.) In this model “what they want” is more mean, and less variance of their portfolio return.
2. Result: “Two fund theorem.” Optimal portfolios are split between the risk free rate and the market (tangency) portfolio.

3. People more risk averse than average hold more in the risk free rate, people less risk averse than average hold more in the stock market portfolio. People are “different from average” by one dimension, risk aversion. So portfolio theory is not totally boring, though it’s pretty close.
4. This theorem is astounding for what it does not say: *The composition of the stock (tangency) portfolio is the same for all investors.* It says tailored stock portfolios (and high fees) are pointless and we’re all out of business. Risk aversion is the *only* reason people don’t do exactly the same thing. Preview: we’ll end up with a way to revive tailored portfolios and all our salaries. There are lots of *other* dimensions on which people differ.
5. It stems from a deep insight: Investors care about the *portfolio* mean and standard deviation, not the behavior of *individual stocks in the portfolio.* It is a very common mistake to think about investments in isolation (should I buy google?) not as part of a portfolio.
6. It reminds us that *names, styles, asset classes, etc. don’t matter.* All that matters is betas, means and covariances! That’s still true!
7. “Active” management. If you think there is alpha, you start with the index portfolio, then shade in the direction of more alpha.
  - (a) But not too much – don’t take nondiversified risks.
  - (b) One way to quantify this tradeoff: As you increase exposure, alpha with respect to *your* portfolio will decline. Keep going until alpha is zero with respect to your portfolio.
  - (c) Alpha – here it represents ways in which you or a manager are “smarter” or “more informed” than others. But don’t forget the average alpha across all investors must be zero!

### 31.2 Real portfolio problems : Review of the classic two-period approach

1. Where does mean-variance come from? *Portfolio maximization problems.*
2. A simple warmup problem: “two-period” portfolios (not in lecture):
  - (a) *Objective:* Start with a simple two-period model,

$$\max E [u(c_{t+1})]; c_{t+1} = W_{t+1};$$

- (b) *Constraint:*  $R^p$  is the portfolio return.

$$\begin{aligned} W_{t+1} &= R_{t+1}^p W_t; \\ R_{t+1}^p &= R_t^f + w' R_{t+1}^e \end{aligned}$$

where

$$w = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \end{bmatrix}; R_{t+1}^e = \begin{bmatrix} R_{t+1}^1 \\ R_{t+1}^2 \\ \vdots \end{bmatrix}$$

i.e.  $R^e$  is a vector of individual asset returns,  $w$  is a vector of portfolio weights.

- i. Why? Let  $w_1 =$  weight (fraction of wealth) in asset 1. Since they all have to add up to one,

$$\begin{aligned} R_{t+1}^p &= w_1 R_{t+1}^1 + w_2 R_{t+1}^2 + \dots + (1 - w_1 - w_2 \dots) R_t^f \\ R_{t+1}^p &= R_t^f + w_1 (R_{t+1}^1 - R_t^f) + w_2 (R_{t+1}^2 - R_t^f) + \dots \\ R_{t+1}^p &= R_t^f + w_1 R_{t+1}^{e1} + w_2 R_{t+1}^{e2} + \dots \end{aligned}$$

- ii. The trick: It's easier to choose among excess returns with no constraint on weights rather than impose weights that must sum up to one,  $wR^f + (1 - w)R$

- (c) Problem: choose portfolio weights to maximize utility

$$\begin{aligned} &\max_{\{w_1, w_2, \dots, w_N\}} E \left\{ u \left[ W_t (R^f + w_1 R_{t+1}^{e1} + w_2 R_{t+1}^{e2} + \dots) \right] \right\} \\ &\max_{\{w\}} E \left\{ u \left[ W_t (R^f + w' R_{t+1}^e) \right] \right\} \end{aligned}$$

- (d) Find the optimal portfolio

$$\frac{\partial}{\partial w_i} : E \left[ u'(W_{t+1}) R_{t+1}^{ei} \right] = 0 \quad \text{for each } i$$

$$E \left\{ u' \left[ W_t (R^f + w_1 R_{t+1}^{e1} + w_2 R_{t+1}^{e2} + \dots) \right] R_{t+1}^{ei} \right\} = 0$$

$N$  equations (i) in  $N$  unknowns ( $w_i$ ) .... solve for  $w_i$

- (e) *This is the same as our old friend*

$$0 = E_t \left[ \beta \frac{u'(c_{t+1})}{u'(c_t)} R_{t+1}^e \right]$$

- (f) When we did asset pricing, we took  $c_t, c_{t+1}$ , as fixed, and we found asset pricing implications  $E(R^{ei})$ . We asked, What must  $E(R^{ei})$  be for answer of a portfolio problem to be “hold the market portfolio?” In doing portfolio theory, we take  $E(R^{ei})$  etc. as fixed, and we find the optimal portfolio, hence the optimal  $W_{t+1}$  or  $c_{t+1}$ . (Isn't optimal  $W_{t+1}$  always more you ask? No,  $W_{t+1}$  is random, so by “optimal  $W_{t+1}$ ” we're asking how large its variance should be in tradeoff for more mean, etc.) This is like a “demand curve.” Our asset pricing question was like “find the market equilibrium price where supply = demand,” or “find the prices (expected returns and covariances) so that the *answer* to the portfolio problem is ‘hold the market portfolio.’ ”

3. The difference between the demand for tomatoes and the demand for stocks, and the average investor theorem again. If you do a demand curve for salad (maximize utility of tomatoes, lettuce, etc. given the prices) you are not at all concerned that your answer might be heavy on tomatoes, i.e. you're not buying the market portfolio of salad. Why do I care so much here? There is an easy answer for salads: tastes differ. I might like tomatoes a lot more than you. I don't learn much from knowing the average salad uses fewer tomatoes than I do.

But where is the difference in “tastes” here? We all want the same thing, more money! And stocks have information in them. So if I'm buying something different from everybody else, and I know we all have the same tastes, now we really should worry. Did I solve the demand

problem right? Does everybody else know something I don't know (e coli in the tomatoes?) Also, the "prices" (expected return and covariance) are a lot less clear. Maybe I'm not reading the price sticker right.

Now we do differ along risk aversion. But that only tells us to go up and down the security market line, it doesn't justify all the other stuff.

Well maybe we differ in other ways. Anyway, you can see the issue: In vegetables it's easy to see why people buy all sorts of different portfolios. There is a lot of "heterogeneity" in tastes. In portfolios, it's less clear, the "prices" are less clear, and I think understanding where you stand relative to all the other shoppers, understand the "heterogeneity" in tastes, is useful to buying a better salad cheaper.

### 31.3 The real portfolio problem

Here is a real portfolio problem. People live a long time. They have jobs or outside income  $y_t$ . The infinite horizon portfolio problem.

$$\max_{\{c_t, w_t\}} E \sum_{j=0}^{\infty} \delta^j u(c_t) \text{ or } E \int_0^{\infty} e^{-\rho t} u(c_t) dt; \quad (36)$$

$$\begin{aligned} w_t &= N \times 1 \text{ portfolio} \\ R_{t+1}^e &= N \times 1 \text{ asset excess returns} \end{aligned}$$

$$W_{t+1} = R_{t+1}^p (W_t + y_t - c_t); \quad (37)$$

$$R_{t+1}^p = R_t^f + w_t' R_{t+1}^e \quad (38)$$

$$\begin{aligned} E_t(R_{t+1}^e) &= \mu_t, \\ cov_t(R_{t+1}^e) &= \Sigma_t \end{aligned} \quad (39)$$

In words:

1. (36): Maximize the lifetime expected utility of consumption. Not  $EU(W_T)$ . Consumption makes you happy, wealth is just a means to consumption. The choice variables are how much to consume  $c_t$  vs. save at each date and the weights  $w_t$  by which you invest wealth in the available assets. So we solve for the optimal consumption/saving as well as portfolio. (It's easy to have finite horizons, death and bequests.)
2. (37): Consumption is limited by invested wealth and income. Wealth tomorrow equals the invested portfolio return  $R_{t+1}^p$  times wealth  $W_t$  plus saving/dissaving, income  $y_t$  minus consumption  $c_t$ .
3. (38): In turn the portfolio return depends on the investment decision. It is the risk free rate  $R_t^f$  plus the portfolio weights  $w_t$  times the excess returns of risky assets  $R_{t+1}^e$ . (I set this up as risk free plus excess returns, so the weights  $w$  do not have to add up to 1).
4. (39),(??): The inputs to the portfolio problem are the mean  $\mu_t$  and covariance matrix  $\Sigma_t$  of asset returns. These potentially vary over time.

## 31.4 Mean-variance portfolios

1. We look at a special case. Why? It's easy to solve and easy to interpret. Finance is about good quantitative parables, not big hairy black-box answers.
2. These are solutions to the basic problem (36). Deriving them takes more math than I want to do here (see "Portfolio theory" on my webpage) but let's look at the answers. *If the investor has no job  $y_t = 0$  and returns are iid,  $\mu_t, \Sigma_t$  constant over time, and with power utility, then*

(a) *Portfolio formula.* The *weights for the optimal portfolio* are

$$\begin{aligned} w &= \frac{1}{\gamma} \Sigma^{-1} E(R^e) \\ (N \times 1) &= (1 \times 1)(N \times N)(N \times 1) \end{aligned}$$

where  $\Sigma = \text{cov}(R^e R^{e'})$ ;  $\gamma = \text{risk aversion}$

- i. This answer has two parts: The *composition* of the portfolio is  $\Sigma^{-1} E(R^e)$  this is an  $N \times 1$  vector, which gives how much to allocate to each asset. The *composition of the risky asset portfolio is the same for everyone.*
- ii. The *scale* is  $1/\gamma$ . This says that less risk averse people should scale it all up more, and more risk averse people less. In this application, it's pretty boring – everyone holds "two funds." But we have broken the tyranny of "the average investor holds the market." People do, justifiably, do different things! It's just pretty boring for now, but that will change.
- iii. This is the "standard formula" in portfolio optimizers all over the world.

(b) The *return* on the optimal portfolio is

$$R_{t+1}^p = R_t^f + w' R_{t+1}^e = R_t^f + \frac{1}{\gamma} E(R_{t+1}^e)' \Sigma^{-1} R_{t+1}^e$$

3. Why is this "mean-variance efficient?" Consider the solution to another problem

$$\begin{aligned} \min \text{var}(R^p) \text{ s.t. } E(R^p) &= \mu \\ \min w' \Sigma w \text{ s.t. } w' E &= \mu \end{aligned}$$

The first order conditions are

$$\Sigma w = \lambda E, \rightarrow w = \lambda \Sigma^{-1} E$$

Thus, the weights given by the answer to our utility maximization problem are the same as these weights which derive from a simple mean-variance problem.)

4. Intuition/special cases.

- (a) This math underlies the mean-variance frontier and intuition above.
- (b) For a single risky asset ("stocks") the formula specializes to

$$w = \frac{1}{\gamma} \frac{E(R) - R^f}{\sigma^2(R)}$$

Invest more if the mean return is higher, variance is lower, or risk aversion is lower.

(c) What counts is the ratio of mean to *variance*, not Sharpe ratio. (The units need to be independent of horizon, and Sharpe ratios go up with the square root of horizon.) The resulting *portfolio* does have the maximum Sharpe ratio even though the thing on the right hand side here is not the Sharpe ratio.

(d) Numbers:

$$\frac{E(R^e)}{\sigma^2(R^e)} = \frac{.06}{.18^2} = 1.85$$

$\gamma = 1 : w = 1.85$

$\gamma = 3.08 : w = 0.60$ , i.e. 60/40. This all seems sensible.

(e)  $\Sigma$  takes account of correlation/diversification opportunities.  $\Sigma^{-1}$  unwinds the correlation of returns to tell you how to take advantage of diversification.

i. Compare three cases, high positive correlation, zero correlation and high negative correlation.

ii. No correlation

$$\Sigma = \begin{bmatrix} 1 & 0 \\ 0 & 0.5 \end{bmatrix}$$

$$\Sigma^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 2.0 \end{bmatrix}$$

If two assets have the same mean,

$$\Sigma^{-1}E(R^e) = \begin{bmatrix} 1 & 0 \\ 0 & 2.0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2.0 \end{bmatrix}$$

*put more of your money in the less variable asset.* But not all of it of course – you want diversification.

iii. Positive correlation.

$$\Sigma = \begin{bmatrix} 1 & 0.9 \\ 0.9 & 1 \end{bmatrix}$$

$$\Sigma^{-1} = \begin{bmatrix} 5.3 & -4.7 \\ -4.7 & 5.3 \end{bmatrix}$$

If the two assets have the same mean, say 1%, then

$$\Sigma^{-1} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 5.3 & -4.7 \\ -4.7 & 5.3 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.6 \\ 0.6 \end{bmatrix}$$

duh, put the same amount in them. But if one has greater mean than the other,

$$\Sigma^{-1} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 5.3 & -4.7 \\ -4.7 & 5.3 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 5.3 \\ -4.7 \end{bmatrix}$$

it advocates a massive long-short position. Well, of course – high correlation means these are nearly the same thing selling for different prices. This is either good advice, or a warning that portfolio optimization programs are very sensitive to mean return assumptions.

iv. Negative correlation ( hml and umd)

$$\Sigma = \begin{bmatrix} 1 & -0.9 \\ -0.9 & 1 \end{bmatrix}$$

$$\Sigma^{-1} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 5.3 & 4.7 \\ 4.7 & 5.3 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 5.3 \\ 4.7 \end{bmatrix}$$

Notice the dramatic difference. Now you put your money *in* both assets, even the one with poorer return. The reason is, you get huge risk-reduction benefits from diversification. You might even put money in an asset with negative return, if it had a strong negative covariance with a positive return asset.

5. You vs. the market, understanding portfolios by how “different” you are than average. *If the investor lives in an economy where everyone else is the same except risk aversion, the optimal portfolio is*

$$w^i = \frac{\gamma^m}{\gamma^i} w^m$$

with return

$$R^{pi} = R^f + \frac{\gamma^m}{\gamma^i} R^{em}$$

where  $w^m$  and  $R^{em}$  are the market portfolio weights and return. You invest differently from others (more in the market portfolio) only if you are more or less risk averse than the average investor. This is the math behind the “Two fund” theorem.

- (a) This is useful! “What’s your risk aversion coefficient  $\gamma^i$ ” is hard to answer! “Are you more or less risk averse than the average investor?” is a question you (or a client) might be able to answer.
- (b) (For nerds: Where does this come from? We just add people up. The market portfolio is the wealth-weighted average of individual portfolios,

$$w^m = \frac{\sum_{i=1}^N W_t^i w^i}{\sum_{i=1}^N W_t^i} = \frac{\sum_{i=1}^N W_t^i \frac{1}{\gamma^i}}{\sum_{i=1}^N W_t^i} \Sigma^{-1} E(R^e) = \frac{1}{\gamma^m} \Sigma^{-1} E(R^e)$$

$$w^m \gamma^m = \Sigma^{-1} E(R^e)$$

$$w^i = \frac{1}{\gamma^i} \Sigma^{-1} E(R^e) = \frac{\gamma^m}{\gamma^i} w^m$$

6. *If everyone is like this* (i.e. with different risk aversion), *the CAPM holds*. (And if the CAPM doesn’t hold, that means we are living in a world in which some people are doing something different!)

- (a) Proof 1: Market is mean-variance efficient. MVF  $\rightarrow$  CAPM (Theory week)
- (b) Proof 2:

$$\gamma^m w^m = \Sigma^{-1} E(R^e)$$

$$E(R^e) = \gamma^m \Sigma w^m = \gamma^m \text{cov}(R^e, R^{em} w^m) = \frac{\text{cov}(R^e, R^{em})}{\text{var}(R^{em})} [\text{var}(R^{em}) \gamma^m]$$



### 31.5 Horizon effects and active/passive portfolio

1. **Horizon.** Should older people move to bonds, and younger people hold more stocks, since they have more time to ride out market ups and downs?
2. **Classic result.** *If returns are lognormal iid and the investor has power utility and no job, the allocation to risky assets is independent of investment horizon.*

(a) To derive this result we write (36) with a “horizon”,

$$\max E \sum_{j=0}^T \beta^j u(c_j) \text{ or } E \int_0^T e^{-\rho t} u(c_t) dt; \text{ or } E_0 [u(W_T)]$$

The fact here is that *you get the same result with any T.*

(b) Intuition: risk is *not* less at long horizons. The “fallacy of time diversification.”

$$\begin{aligned} E(r_1 + r_2 + r_3) &= 3E(r) \\ \sigma^2(r_1 + r_2 + r_3) &= 3\sigma^2(r) \\ \frac{3E(r)}{3\sigma^2(r)} &= \frac{E(r)}{\sigma^2(r)} \end{aligned}$$

and our formula

$$w = \frac{1}{\gamma} \frac{E(R^e)}{\sigma^2(R^e)}$$

said the ratio of mean to *variance* is what counts. Ok, this isn’t quite right since I’m looking at geometric not arithmetic returns, but the point is correct when you do it right.

(c) The “fallacy of time-diversification” (oft-repeated). The variance of *annualized* returns is lower for longer horizon, but not the variance of *total* returns.

$$\begin{aligned} \sigma^2(r_1 + r_2 + \dots + r_T) &= T\sigma^2(r) \\ \sigma^2 \left[ \frac{1}{T}(r_1 + r_2 + \dots + r_T) \right] &= \frac{1}{T}\sigma^2(r) \end{aligned}$$

Your risk is the annualized variance *multiplied by the horizon*, so does not decrease even though annualized variance does.

(d) If you don’t like the result, either the model is wrong – maybe stocks aren’t random walks; mean reversion makes  $\sigma^2$  rise slower with horizon; maybe people have jobs (!) – or maybe the standard advice is wrong. There’s a lot of work on “what is there about the real portfolio problem that means people should shift to bonds as they get older.” One view, for example, is that people’s wages are like a bond (coupon), so they should be more invested in stocks when younger, but when they retire and aren’t getting wages any more they should shift to bonds. Except maybe in the modern world, wages are more like stock dividends!

3. **Factor Models and alphas.** Or splitting the portfolio decision into “passive” “policy” and “systematic” and “alpha” components. If returns follow a factor model, such as the CAPM or FF3F,

$$R_t^i = \alpha^i + \beta^i F_t + \varepsilon_t^i;$$

Write the portfolio problem as “portable alpha,”

$$\begin{aligned} R_t^p &= R^f + w'_F F_t + w'_\varepsilon (R_t^{ei} - \beta^i F_t) \\ &= R^f + w'_F F_t + w'_\varepsilon (\alpha^i + \varepsilon_t^i) \end{aligned}$$

$w_F$  tells you how much factor exposure to take – market, hml, smb, etc.  $w_\varepsilon$  tells you how much “alpha” to take. The answer is

$$w_F = \frac{1}{\gamma} \Sigma_f^{-1} E(F) \quad w_\varepsilon = \frac{1}{\gamma} \Sigma_\varepsilon^{-1} \alpha$$

Thus, *separately think about your factor exposures and your alpha exposures*. Using the same formula!

- First invest in  $F$  as if they were the only assets. (How much rmr, value, smb?) Then add alphas, as if they were the only assets with mean  $\alpha$  and covariance given by the residual covariance. (Recall the Sharpe ratio discussion with FF3F model and problem set.)
- Note  $w_{\varepsilon_i}$  is not the weight in  $R^{ei}$ . It is the weight in  $\alpha_i + \varepsilon_i = R^{ei} - \beta_i F$ . The weight in  $R^{ei}$  directly is a) its weight in  $R^{em}$  and then b) this weight. If your “active” managers come with “style betas,” you have to offset this in the “passive portfolio.”
- The active allocation depends on  $\alpha/\sigma^2$  not Sharpe  $\alpha/\sigma$ . Thus, small alphas with modest tracking error can give huge allocations to the active portfolio. Example:  $\alpha = 0.4\%$ ,  $\sigma_\varepsilon = 2\%$ ,  $Sharpe = 0.25$ . But

$$w_\alpha = \frac{1}{\gamma} \Sigma^{-1} \alpha = \frac{1}{\gamma} \frac{.004}{.02 \times .02} = \frac{1}{\gamma} \frac{0.004}{0.0004} = \frac{1}{\gamma} \frac{4}{0.4} = \frac{10}{\gamma}$$

- You need to consider a *portfolio* of your hedge funds or other active investments!  $\Sigma_\varepsilon$  captures correlation of the (non-market) parts of your various active managers. (Remember hedge-fund lectures)
- (Math. I’ll show the mean-variance case in which  $\lambda$  is a Lagrange multiplier. Getting risk aversion out of it is a little tougher.

$$\min w'_f \Sigma_f w + w'_\varepsilon \Sigma_\varepsilon w_\varepsilon$$

$$\text{s.t. } w'_f E(f) + w'_\varepsilon \alpha = \mu;$$

taking derivatives with respect to  $w$ ,

$$\Sigma_f w = \lambda E(f). \quad \Sigma_\varepsilon w_\varepsilon = \lambda \alpha.$$

- Recipe 2 for incorporating alpha:** Recall the basic theorem, a return  $R^{ep}$  is on the mean variance frontier  $\leftrightarrow E(R^{ei}) = 0 + \beta_{i,p} E(R^{ep})$ . Thus you can find a mean-variance efficient return by *chasing alpha calculated relative to your own portfolio until the alpha is gone*. This rule holds if **you** care about mean and variance, even if the CAPM does not hold. (Don’t get confused – the investor can’t change the character of  $R^{ei}$ . The only thing he/she can change here is the character of  $R^p$ , by putting more/less weight in different assets.)

- (a) *Example:* Start holding the market. Value looks like it has a low  $\beta$  and a good  $\alpha$ . Start buying value stocks. As you buy more value stocks, the portfolio  $R^p$  starts to look more value-like. The  $\beta$  of value stocks *relative to the portfolio*  $R^p$  starts to rise, and the  $\alpha$  starts to decline. Keep going until the  $\alpha$  is zero. We're adjusting the *portfolio*, the right hand side of the CAPM-like regression, until the alpha goes away.
- (b) The CAPM is alive and well! A mean-variance investor should evaluate assets using the CAPM, relative to *his or her* portfolio, whether or not the *market* portfolio satisfies the CAPM. He/she should buy positive alpha assets until they are no longer positive alpha with respect to his/her portfolio.
- (c) If there are true alphas relative to the *market* portfolio (say, hml is right), then the investor's  $R^p$  will not be the market portfolio – it will have more weight in the high alpha investments (e.g. value).

### 31.6 Assumptions / what's wrong

1. This model made strong assumptions: People have no job! Returns are iid,  $E_t R_{t+1}$  is the same always. What about  $R_{t+1} = a + b \times DP_t + \varepsilon_{t+1}$ , i.e.  $E_t(R_{t+1}) = a + b \times DP_t$  not a constant?
2. People often use it anyway in a time-varying return environment,

$$w_t = \frac{1}{\gamma} \Sigma_t^{-1} E_t(R_{t+1}^e)$$

That includes practically all hedge funds or other quantitative portfolios. But it make no sense: the model is derived using iid returns, yet if returns were iid you would never actively trade!

3. Technically, mean-variance portfolios are also true in a one-period environment: *If the investor has no job, a one period horizon, and (quadratic utility, or if returns are normally distributed,) then he wants to hold a mean-variance efficient portfolio.* But most of us want to live more than a year! It's very worth knowing this result. When your hedge fund says “we use the latest mean-variance optimizer to optimally diversify our alpha bets ( $E_t R_{t+1}^e$ )” you can answer “why do you have a one-period horizon?”
4. What's wrong with the standard advice?

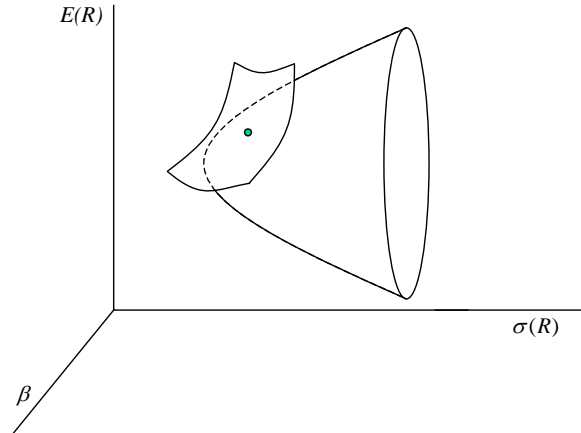
- (a) Assumptions are wrong – people have jobs, and  $E_t(R_{t+1}^e)$ ,  $\Sigma_t$  vary a lot over time.
- (b) Predictions 1: Returns follow the CAPM. What about value, momentum, etc.? Is everyone else really just to dumb to notice?
- (c) Predictions 2: The portfolio problem also tells you to consume  $c_t = kW_t$  proportionally to wealth. That's sensible in this model. But it tells you to follow a policy in which  $\sigma(\Delta \ln c_t) = \sigma(\Delta \ln W_t) = 16\%$  per year or more.

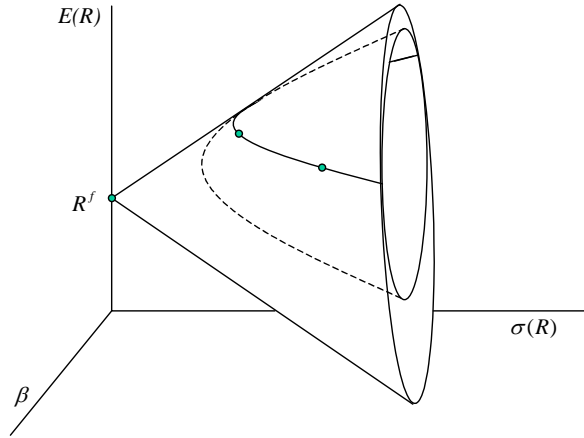
In reality, people ignore “temporary” fluctuations in market value, so consumption declines much less than market values. Well, this model ruled out “temporary” fluctuations. Maybe this is just an indication of how silly the model is! If you don't like the consumption advice, why do you like the portfolio advice?

Note: This consideration is *totally* ignored in practice. People take the portfolio advice, ignore the consumption advice.

### 31.7 Multifactor/Merton.

1. How do value/growth/momentum effects, jobs, etc. change this picture?
  - (a) A1: Not at all. The average investor holds the market. If you're average, and this is a "risk premium" you ignore the value premium, because it's risky.
  - (b) A2: Maybe you're not average. Value, other priced factors adds another dimension to the graph, and much more interest (before we start fishing for alpha).
  - (c) As before, let's do intuition first, and then look at the classic math.
  
2. Example for intuition: Suppose hml represents a "recession" factor, compensating investors for the risk that stocks do especially badly in a recession or financial crisis in which lots of people are afraid of losing their jobs. What do investors want? Before: More mean, less variance. Now: More mean, less variance *and* less tendency for assets to fall in a recession. They are willing to accept less mean or higher variance to have stocks that do better in recessions. Axes: mean  $E(R^p)$ , standard deviation  $\sigma(R^p)$ , covariance or  $\beta$  of portfolio with recessions.
  
3. What can they get? The MVF becomes a *multifactor efficient frontier*





Negative betas come out of the page; investors want portfolios that covary negatively with recessions. The cone is not necessarily centered on  $\beta = 0$ , in fact it is most likely centered on a positive  $\beta$  since most stocks do badly in recessions.

- (a) The optimal portfolio is now on the cone, the *multifactor efficient frontier*.
  - (b) The average (and hence market) portfolio *is no longer on the mean-variance frontier*. (The MVF is on the top of the nosecone).
  - (c) How did we change the prediction? We removed the assumption from Mean-Variance theory that “the investor has no job or outside income.” Most investors have labor, business income, illiquid portfolio, real estate or other asset that cannot be sold, and those outside incomes are correlated with stock returns! Investors want portfolios that “hedge against risks to these state variables,” i.e. don’t tank when all these other things tank.
  - (d) “Intertemporal portfolio theory” works the same way. If you are going to invest for a long time, a rise in expected returns is good news, and a decline in expected returns is bad news. Hence, you want assets that go up when expected returns go down – you want assets that go up when dp goes down. This is “another dimension” just like “recession.” You want assets that “hedge state variables for the investment opportunity set.” Another example (Fall 2008) volatility. You want assets that do well in times of huge volatility.
4. Math: Real portfolio problems have a long-lived investor, with outside (job) income  $y_t$  and recognize time-varying means (and potentially variances too), for example.

$$\max_{\{c_t, w_t\}} E \sum_{j=0}^{\infty} \delta^j u(c_t) \text{ or } E \int_0^{\infty} e^{-\rho t} u(c_t) dt;$$

$$\begin{aligned} W_{t+1} &= R_{t+1}^p (W_t + y_t - c_t); \\ R_{t+1}^p &= R_t^f + w_t' R_{t+1}^e \end{aligned} \tag{40}$$

$$\begin{aligned} R_{t+1}^e &= a_r + b_r' z_t + \varepsilon_{t+1}^r \\ \text{cov}_t(R_{t+1}^e) &= \Sigma \end{aligned} \tag{41}$$

$$\begin{aligned} y_{t+1} &= a_y + b'_y z_t + \varepsilon_{t+1}^y \\ z_{t+1} &= a_z + b'_z z_t + \varepsilon_{t+1}^z \end{aligned} \tag{42}$$

This is just as before with tiny but vital differences. In words:

- (a) (40): Labor or business income  $y_t$  now appears.
- (b) (41): The mean asset return  $E_t(R_{t+1}^e) = a_r + b'_r z_t$  can vary over time.
- (c) (42). We'll model that variation over time by a vector of *state variables*  $z_t$ . These variables carry all the information you use to forecast the future of asset returns and of income  $y_t$ . For example, when we forecast returns from dividend yields,

$$\begin{aligned} R_{t+1} &= a + b \times DP_t + \varepsilon_{t+1} \\ E_t(R_{t+1}) &= a + b \times DP_t \end{aligned}$$

then  $DP_t$  is one element of the vector  $z_t$ .  $z_t$  describes how the future looks at any moment in time. I wrote down linear versions, but you can quickly see how the idea generalizes to bigger  $z$ , nonlinear forecasts,  $\Sigma_t$  varying over time, etc.

- (d) Language:  $z$  are “state variables” as decisions will depend on  $W_t$  and  $z_t$ . All information you need to make a decision is wrapped up in these.  $\varepsilon_{t+1}^r$  are “shocks to asset returns.”  $\varepsilon_{t+1}^z$  are “shocks to state variables” (especially “state variables for the investor’s intertemporal opportunity set.”) Changes in  $z$  are good or bad news for your utility, so you will like/dislike assets according to the correlation of return shocks and  $z$  shocks, i.e. how they “hedge shocks to state variables.”

5. *Weights* The weights of the optimal portfolio are

$$w = \frac{1}{\gamma} \Sigma^{-1} E(R^e) + \beta_{R^e, z'} \frac{\eta}{\gamma} \tag{43}$$

(Everything has a  $t$  subscript, and can vary over time.)

- (a) “Market timing”  $w_t = \frac{1}{\gamma} \Sigma_t^{-1} E_t(R_{t+1}^e)$ . This term now has  $t$  on it. Other things equal, you will invest more when expected returns are higher (e.g. DP is higher).
- (b) Portfolio weights now add an extra term, “*Hedging Demand*” to control state-variable risk.
  - i.  $\eta$  = “aversion to state variable risk.” The more you care about the state variables, the more you hedge. ( $\eta$  is related to, “if  $z$  goes up, are you happier or unhappy.” Really,  $\eta(W_t, z_t)$  can change over time)
  - ii.  $\beta_{R^e, z'}$  = (Number of assets)  $\times$  (Number of state variables) matrix of regression coefficients, describing regressions of each “state variable,” on returns, i.e. “does this stock tend to fall extra special bad in a recession?” These are hml or smb betas if hml or smb are the “state variables.” (Really  $\beta$  also can vary over time)

6. (Highly Optional. What is this  $\eta$  business?) A little more deeply, we define the “value function” as the maximum level of happiness you can get given wealth  $W$  and state variables  $z$ ,

$$V(W_t, z_t) = \max_{\{c_t, w_t\}} E \sum_{j=0}^{\infty} \delta^j u(c_j) \text{ s.t.} \dots$$

Now,  $V$  becomes basically your “utility function” defined over wealth and the forecast variable  $z$ . (Technically, it’s called a “value function of the dynamic program.”) Risk aversion and state variable aversion are defined by the curvature of this new “utility function”,

$$\begin{aligned}\gamma &= -\frac{W\partial^2V(W, z)/\partial W^2}{\partial V(W, z)/\partial W} = -\frac{WV_{WW}(W, z)}{V_W(W, z)} \\ \eta &= \frac{\partial V(W, z)/\partial z}{\partial V(W, z)/\partial W} = \frac{V_z(W, z)}{V_W(W, z)}\end{aligned}$$

The two versions of the expression are the same, just different notation. In the first equation, you recognize risk aversion. The second one tells us that  $\eta$  measures how much your utility is impacted by a change in the state variable  $z$ . For example, if  $z$  forecasts your future income, then a high  $z$  is good news, just like a good shock to wealth today.

$\eta$  has to be computed numerically by solving the whole problem, and it’s not easy. It’s not a “preference” like  $\gamma$ . It depends on the nature of your outside income or your portfolio vs. the return process. For example, if you own a steel company, then you will have a high  $\eta$  on steel stocks, since you care about that exposure. If you are a long horizon investor, then you have a high  $\eta$  on state variables that describe long run returns.

For example, suppose you’re a bond investor with a 10 year horizon, power utility, and you have decided to invest all your money in the 10 year bond. Your utility is then just the utility of wealth in 10 years.

$$U(W_T) = \frac{W_T^{1-\gamma}}{1-\gamma}$$

The *value function* is your expected utility at time  $t$  of this outcome, in terms of time  $t$  variables. If your zero coupon bond maturing at time  $T$  has a yield  $Y_t$ , then wealth  $W_t$  means final wealth  $W_t Y_t^{T-t}$ . Thus, the value function in terms of time  $t$  state variables is

$$V(W_t, Y_t) = \frac{(W_t Y_t^{T-t})^{1-\gamma}}{1-\gamma}$$

We can treat this as your “utility function.”

$$\begin{aligned}V_W &= (W_t Y_t^{T-t})^{-\gamma} Y_t^{T-t} \\ V_Y &= (W_t Y_t^{T-t})^{-\gamma} (T-t) Y_t^{T-t-1} W_t \\ V_{WW} &= -\gamma (W_t Y_t^{T-t})^{-\gamma-1} Y_t^{2(T-t)}\end{aligned}$$

so

$$-\frac{WV_{WW}(W, Y)}{V_W(W, Y)} = \gamma \frac{W_t (W_t Y_t^{T-t})^{-\gamma-1} Y_t^{2(T-t)}}{(W_t Y_t^{T-t})^{-\gamma} Y_t^{(T-t)}} = \gamma$$

and

$$\eta = \frac{V_Y}{V_W} = \frac{(W_t Y_t^{T-t})^{-\gamma} (T-t) Y_t^{T-t-1} W_t}{(W_t Y_t^{T-t})^{-\gamma} Y_t^{T-t}} = \frac{(T-t)W_t}{Y_t}$$

That’s the point – if you look at it from a one-period portfolio point of view, this investor seems to “care” about yield changes, just as he cares about getting a new ipad. Thus,  $\eta$  gives

him a demand for securities that go up when yield goes up (prices go down), and the long term bond is an excellent such security. Viewed from a one-period point of view that's why he holds long term bonds.

7. (Optional. Nosecone math.) How do we expand the notion of mean-variance efficient? Add a new constraint to control exposure to the “state variable”  $z$ ,

$$\begin{aligned} \min var(R^p) \text{ s.t. } E(R^p) &= E, cov(R^p, z) = c \\ \min w' \Sigma w \text{ s.t. } w' E(R^e) &= \mu, w' cov(R^e, z) = c \end{aligned}$$

Introduce Lagrange multipliers,

$$\min_w w' \Sigma w - \lambda (w' E(R^e) - \mu) - (w' cov(R^e, z) - c) \delta$$

First order conditions, from  $\partial/\partial w$  :

$$\begin{aligned} \Sigma w &= \lambda E(R^e) + cov(R^e, z') \delta \\ w &= \lambda \Sigma^{-1} E(R^e) + \Sigma^{-1} cov(R^e, z) \delta \\ w &= \lambda \Sigma^{-1} E(R^e) + \beta_{R^e, z'} \delta \end{aligned}$$

you pick  $\lambda$  and  $\delta$  to attain the constraints  $E, c$ . You can sweep out the frontier by varying  $\lambda, \delta$ . Notice this “nosecone” answer is the same as the answer (43) I gave to the portfolio problem.  $\lambda$  and  $\delta$  correspond to the investor's value of  $\gamma, \eta$ .)

8. *Relative to Market* Once again, it's useful to think about your (i) portfolio *relative* to the market portfolio (m) and how you're different from others. This is the math behind our “nosecone” discussion. Investor  $i$  portfolio return is

$$\begin{aligned} R_{t+1}^{pi} &= R_t^f + \frac{\gamma^m}{\gamma^i} R_{t+1}^{em} + \frac{1}{\gamma^i} (\eta^{i'} - \eta^{m'}) R_{t+1}^{ez} \\ R_{t+1}^{ez} &\equiv \beta'_{R^e, z'} R_{t+1}^e \end{aligned}$$

- (a) The first term is familiar – invest more or less in the market, depending on how risk averse you are relative to other people.
- (b) The second term says to invest in additional special portfolios  $R^{ez}$  according to how different your *aversion to state variable risk* is than theirs.
- (c) Constructing  $R^{ez}$ : By definition,  $\beta_{R, z'}$  comes from the regression

$$z_{t+1} = \beta'_{R, z'} R_{t+1}^e + \varepsilon_{t+1}$$

$R^{ez}$  is the fitted value in a regression of  $z$  on  $R^e$ . It is the hedge portfolio or “mimicking portfolio” for state variable risk. It is the portfolio “closest” to the state variable, the one you would *short* to best hedge the state variable. For example, if  $z =$  unemployment, shorting  $R^{ez}$  is the best portfolio you can buy that provides “unemployment insurance,” rising when unemployment rises. (Note, the regression really connects the unexpected shock to  $z$  and the unexpected shock to returns, so you don't need an intercept.)

- (d) This is what Fama-French mean by ‘value is the mimicking portfolio for a state variable of concern to the average investor’



9. *If everyone is like this, a multifactor model is the equilibrium.* Prices (expected returns) must adjust until the average investor is happy to hold the market portfolio! If you add up (43) and solve for the expected return that makes weights = market weights, you get for any asset return  $R^e$

$$E(R^e) = \text{cov}(R^e, R^m)\gamma^m - \text{cov}(R^e, z')\eta^m$$

10. Implications of hedging demand

- (a) If you're like everyone else  $\eta^i = \eta^m$  you just hold the market, *despite* the value premium!
- (b) A mean-variance investor ( $\eta^i = 0$ ) living in the multifactor world can do better than hold the market. He takes on "recession risk" that others don't want, and earns a premium for doing so. (He goes to the top of nosecone. This is the sales pitch for value sales pitch.)
- (c) Investors differ by risk aversion  $\gamma^i$  and aversion to recession risk  $\eta^i$ . Different investors will want to be higher up, lower down, and further to the side or the top. They need another portfolio to move around the cone – a "recession-risk" portfolio. Thus we have a *three (or more)* fund theorem.
- (d) This analysis is not limited to "priced factors" with mean  $> 0$  ( $E(hml) > 0$ ) or  $\eta^m \neq 0$ . If the factor is not priced (smb? industry, etc.) that means exactly as many people want to go long as want to short this factor. *They should do so.* "priced" only means that the hypothetical investor with no other risks (job, house, business, etc.) wants to buy or sell. Who cares? Thus, we should think about factors such as industry portfolios too.
- (e) *Hedging outside income risk is job 1, and does not require any alphas.* It's like buying house insurance.
  - i. *Don't hold company stock*
  - ii. *Don't hold stock in your own industry*
  - iii. *If you can identify factors (oil prices, financial stocks) that covary with your job/business, short them.*
- (f) We've spent all our time looking for "alphas", "anomalies",  $\eta^m \neq 0$ . These are only interesting for the one last MV investor with  $\eta^i = 0$ . For everyone else, hedge portfolios are important even if  $\eta^m = 0$ !
- (g) Example: Suppose steel prices go down. Steelworkers are unhappy, we are happy. Steelworkers should short steel, we should go long. Each insures the other. The steel industry portfolio is not priced (no alpha), but still enters importantly into everybody's portfolio decision.

11. More recipes to make use of this insight. All of these amount to the same thing. What's the easiest way to implement these ideas?

- (a) *We can split the optimal portfolio in two parts. 1) Hedge as much of the outside income as possible, by shorting a portfolio that looks as much like outside income as possible (second term) 2) The rest of the portfolio is on the mean-variance frontier.* (first term)

$$w = \frac{1}{\gamma} \Sigma^{-1} E(R^e) + \beta_{R,zy'} \frac{\eta}{\gamma}$$

This is hard though, because what's mean variance efficient?

- (b) 1) Hedge outside income to the extent that you're different from everyone else. then 2) hold market index

$$R^i = R^f + \frac{\gamma^m}{\gamma^i} R^{em} + \frac{1}{\gamma^i} (\eta^{it} - \eta^{mt}) R^{ez}$$

- (c) Figuring out the optimum “style” allocation is not so simple as risk aversion, and the stock/bond split decision. How *much* value (other factors, other styles) should you hold? What are the corresponding risks, how are *you* affected by them? How can we find good portfolios (industry?) to hedge *your* labor, business, etc. risks. “Style coaching” *is* a reason for tailored portfolios! It needs complex understanding of markets and can charge fees! This is separate business – separate management companies can set up 1) hedge portfolios 2) passive indices 3) alphas.
- (d) JC guess at next generation challenge: set this up. Most of academia and industry is blind to this. By habit, we're focused on finding alpha (priced factors) for the last remaining mean-variance investor.

12. Summary: Two big sources of “hedging demand.”

- (a) Hedging outside income – jobs, businesses, human capital etc. – that expose you to risks correlated with some assets in your portfolio
- (b) Hedging “intertemporal opportunities” the fact that returns are not i.i.d. Bonds are a great example, as you see next....

### 31.8 An example of why extra state variables / intertemporal thinking is important.

(This is the story from “Discount rates” p. 1082)

1. Let's start with the above calibration to “normal times,” a 60/40 allocation with 6% equity premium and 18% volatility

$$w = \frac{1}{\gamma} \frac{E(R) - R^f}{\sigma^2(R)}$$

$$0.6 = \frac{1}{3.08} \frac{0.06}{0.18^2}$$

2. Now, in Oct 2008, monthly volatility rose to 70%. How should we adjust our portfolio weights?

$$\frac{1}{3.08} \frac{0.06}{0.70^2} = 0.0397$$

*Quick, sell down to 4% equity!* This seems nuts. Of course, we can't all do it. Are people still in the market just nuts?

3. Ok, you say, expected returns rose. How much? With prices down you're probably at 40% not 60% equity, so we only need to find

$$\frac{1}{3.08} \frac{E(R^e)}{0.70^2} = 0.40$$

$$E(R^e) = 3.08 \times 0.40 \times 0.70^2 = 0.60 = 60\%$$

*You'd have to believe stocks dropped so much the expected return is 60% per year to justify not selling.*

4. This consideration led many apparently sophisticated institutional managers to dump stocks. But both conclusions are completely nutty. It can't be true that the right thing to do is to hold 4% equity – or that markets needed to fall to about Dow = 2 cents before it made sense to stay in. It also seems unlikely that expected returns rose to 60%.

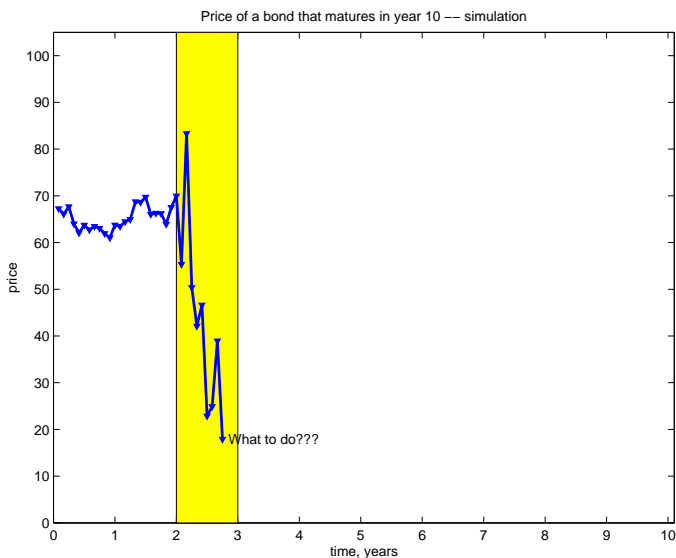
*And the average investor must hold the market.* Really, you can't justify everyone else holding 60/40 and you holding 4% because you're so much smarter, the only one who read Markowitz 1952 (yes, 1952) mean-variance paper. Really, in December 2008 was the market ridiculously “behaviorally” *optimistic, and overvalued*, so it makes sense to sell to all those morons who don't get it?

What's wrong? Well, *the model must be wrong!* Garbage in = garbage out. The model assumed returns are i.i.d., so the fact that volatility rose from 20% to 70% already invalidates the main assumption. *Returns are not i.i.d.; declines in price/dividend ratios mean rises in expected returns. Stocks are a little like bonds. This portfolio formula is wrong.* It ignores “state variables”. In particular lower prices mean higher mean returns – a “*hedging demand*” and volatility, which is also a “state variable”

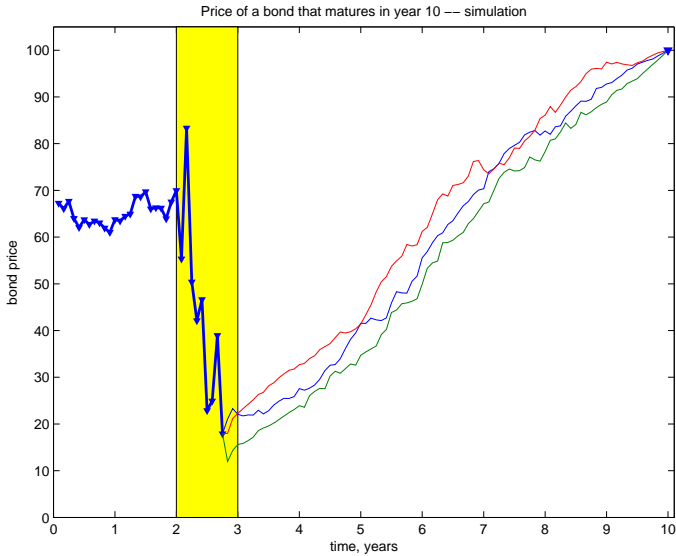
This case shows that one-period mean-variance is *not* a simple pretty good approximation, and that “state variables” are really first-order important for getting the right answer. In this case, “don't panic.”

5. The long term bond investor example

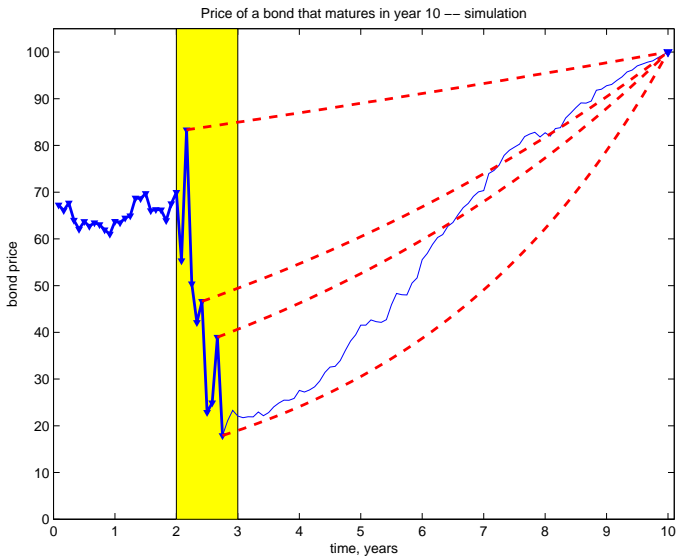
- (a) Imagine you have a liability in 10 years, 10 year zero coupon TIP. What do you do if the market crashes and volatility spikes (Fall 2008, but in bonds?)



Bonds: Answer: *Do nothing*



Bonds: *Do nothing*



(b) Viewed Merton-style

$$\text{risky share} = \frac{1}{\text{r.a.}} \frac{\text{expected return} - \text{riskfree rate}}{\text{return variance}}$$

+ “aversion to yield change” × cov.(return, yield change)

Bonds are a perfect “hedge” against the “state variable”, yield.

- (c) Looking directly at cashflows is much simpler than these one period formulas! (See Cochrane, “mean variance benchmark” extra reading)
- (d) For most investors the risk free asset is a coupon-only long-term TIP, not a money market fund, and they should ignore its price fluctuations!
- (e) Are stocks a bit like bonds, so “do nothing” is at least partially right? Does a price decline mean higher yield (return)?. Yes! DP and returns gives a “hedging demand”

for stocks for long run investors! *Stocks are in fact a better investment for long term investors, because when prices fall, expected returns rise.*

6. Bond/Hedging demand example in more detail.

- (a) If a bond yield goes down (bad news) then the bond price goes up (it hedges its own bad news). Hence long term bonds are “safe” for long term investors, despite poor  $E/\sigma^2$ . For high risk aversion, this consideration *completely reverses* the standard portfolio advice. Don’t hold “cash,” hold long term (inflation indexed) bonds. *Hedging demands can be big.*

Bond portfolio management is particularly misguided, since we know bonds are not iid. For example, suppose you have a 10 year horizon, are completely risk averse, so you buy a 10 year zero coupon indexed bond. Now, if the price goes down or volatility goes up, you don’t care since you are completely hedged.

- (b) The standard one period mean-variance formula gets this completely wrong. The expected excess return of the long term bond - short rate is small, say  $E(R_{t+1}^{e(10)}) = 1\%$ . But of course 10 year bonds are very risky, say  $\sigma(R^{e(10)}) \approx 10\%$ . If you use standard one-period mean-variance analysis

$$w = \frac{1}{\gamma} \Sigma^{-1} E(R^e) = \frac{1}{\gamma} \frac{0.01}{0.10^2} = \frac{1}{\gamma}$$

For regular  $\gamma = 3$  this says to hold only 1/3 in long term bonds. Worse, as  $\gamma$  rises, this says to hold short term bonds, where we know the right answer is to move to the long term bond.

- (c) What went wrong? Bond returns are not iid! When bond prices go down, bond yields and expected returns must rise. Bond returns are perfectly correlated with yields. Thus, there is a big  $\beta_{R,z'}$  term, and a long term bond investor cares a lot about bond yields so  $\eta$  is a big number. The right formula is

$$w = \frac{1}{\gamma} \Sigma^{-1} E(R^e) + \beta_{R,z'} \frac{\eta}{\gamma}$$

In this case we know the last term is really important! If you want to get to the right answer,  $w = 1$  as  $\gamma$  declines, you can’t leave it out.

- (d) As this reveals, thinking about bonds in a one-period framework with hedging demands is pretty silly. If you look at 10 year payoffs “the 10 year bond is riskfree at a 10 year horizon” is pretty obvious. But we don’t yet know how to do long-horizon portfolio theory directly.
- (e) Bad stock returns come with lower P/D, thus higher future expected returns. This means the “hedging demand” might also be positive, and stocks are “safer” for long term investors. This, I think, is the resolution to the puzzle I posed above, how can you possibly not sell everything when  $\sigma$  rises from 18% to 70%.

7. *It makes no sense whatsoever to carefully model time-varying expected returns and variance , and then use a one-period optimizer!* When you do, you are ignoring the hedging demands. In this case, you would make exactly the wrong decision, buying short term bonds rather than long term bonds. *If state variables vary through time, hedging demands can be big and reverse the whole answer.* (Though most hedge funds do this!) Instead, do the problem right.....

8. But, note we got the right answer when we just looked at the 10 year problem, not the one-year hedging demands. *Intertemporal hedging demands (state variables for expected returns) disappear from the long-run portfolio if you look at it directly.* State variables for outside income remain.

### 31.9 Predictability

**How much should you take advantage of stock return forecastability (d/p), bond return forecastability (f-y), exchange rate forecastability ( $r^d - r^f$ )?** Yes, invest more when expected returns are higher, but how much?

1. Questions/preview/intuition:

- (a) Market timing: The allocation to stocks vs. bonds changes over time; you should put more into stocks when the frontier is steeper

$$w_t = \frac{1}{\gamma} \Sigma^{-1} E_t(R_{t+1}^e) + \dots$$

- (b) (More subtle) The risk that  $E_t(R)$  or  $\sigma_t(R)$  might change acts like another factor. “Shifts in the investment opportunity set.” Stocks might be a bit like bonds. As another example, if the huge increase in volatility F 08 was going to make us sell, we should have been long vol (though that’s expensive).

$$w_t = \dots + \dots \beta_{R^e, z} \frac{\nu}{\gamma}$$

- (c) (Even more subtle). Are there horizon effects on the portfolio allocation? Maybe mean-reversion in stock prices means more stock allocation for long run investors? Are stocks a bit like long term bonds, studied above?

2. *Portfolio problem* The objective is just like before, and let’s simplify by leaving out outside income .

$$\max E \sum_{t=0}^{\infty} \beta^t u(c_t) \text{ s.t. } W_{t+1} = \left( R_t^f + w_t R_{t+1}^{em} \right) W_t - c_t$$

But now our description of mean and variance of returns changes over time. For example,

$$\begin{aligned} R_{t+1}^{em} &= a + b(D/P_t) + \varepsilon_{t+1} \\ E_t(R_{t+1}^{em}) &= a + b(D/P_t) \\ D/P_{t+1} &= \phi D/P_t + v_{t+1} \end{aligned}$$

More complex examples use a whole vector of “signals” to capture mean returns and the conditional variance of returns.

3. *Answer: We’ve already done it.*  $D/P_t$  i.e.  $E_t(R_{t+1}^{em})$  is a “state variable,”  $z_t$  The general formula

$$w = \frac{1}{\gamma} \Sigma^{-1} E(R^e) + \beta_{R, z'} \frac{\eta}{\gamma}$$

means, in this context,

$$w_t = \frac{1}{\gamma} \frac{E_t(R_{t+1}^e)}{\sigma_t^2(R_{t+1}^e)} + \frac{\eta}{\gamma} \frac{\text{cov}(\varepsilon_{t+1}, v_{t+1})}{\sigma_t^2(R_{t+1}^e)}$$

You see the “Market timing” demand – invest more when expected returns are higher, plus the “hedging demand” – buy assets whose value goes up when investment opportunities are awful.

- (a) The hedging demand depends on the covariance of *expected return (dp) shocks* with *actual returns*. If *expected* returns rise when *actual* returns are low the asset is a good hedge for its reinvestment risk. (If yields rise when prices decline, long-term bonds are a good hedge for their reinvestment risk)
- (b) If *expected* returns are uncorrelated with actual returns,  $\text{cov}(\varepsilon_{t+1}, v_{t+1}) = 0$  then the hedge demand is zero, *even though* expected returns vary through time. It’s possible for  $E_t(R_{t+1})$  to vary through time and give rise to no hedging demand at all.  
Intuition: like all beta zero ideas. You don’t like it when you wake up and expected returns are lower. But if no available assets can hedge that risk, that consideration does not change your attitude to those available assets.
- (c) Horizon effects? That would show up in  $\eta$ . If you actually solve the problem, you find that hedging demands are more important for long term investors, who have a higher  $\eta$ .

#### 4. A calculation

- (a) Now: given a regression model, i.e.

$$R_{t+1}^{em} = a + b \left( \frac{D_t}{P_t} \right) + \varepsilon_{t+1}$$

we have

$$\begin{aligned} E_t(R_{t+1}^{em}) &= a + b \left( \frac{D_t}{P_t} \right) \\ \sigma_t^2(R_{t+1}^{em}) &= \sigma_\varepsilon^2. \end{aligned}$$

We can calculate the optimal market timing portfolio<sup>26</sup>

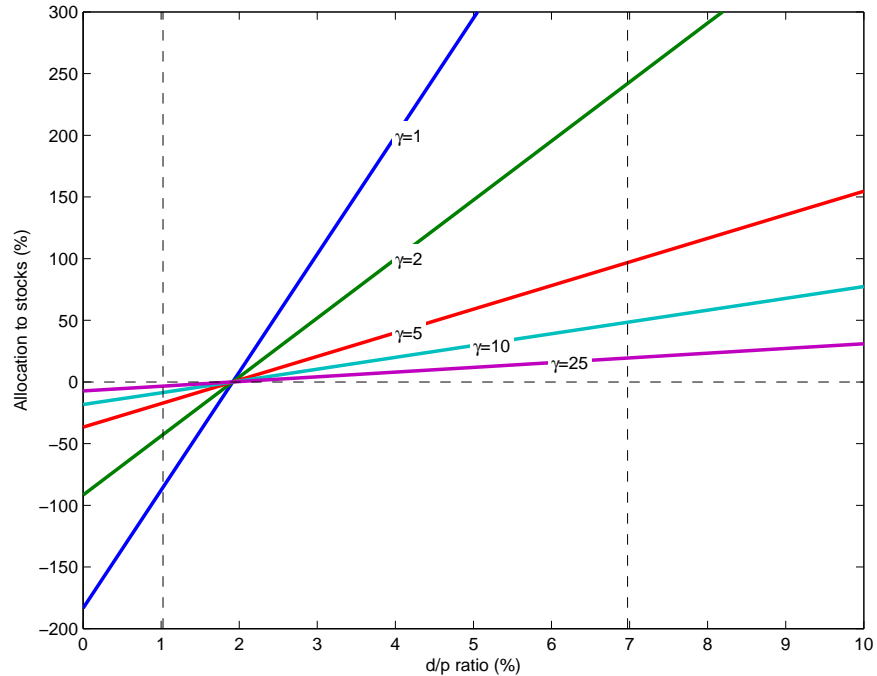
$$w_t = \frac{1}{\gamma} \frac{a + b \left( \frac{D_t}{P_t} \right)}{\sigma_\varepsilon^2}$$

- (b) Let’s use the result from problem set 1 d/p regression

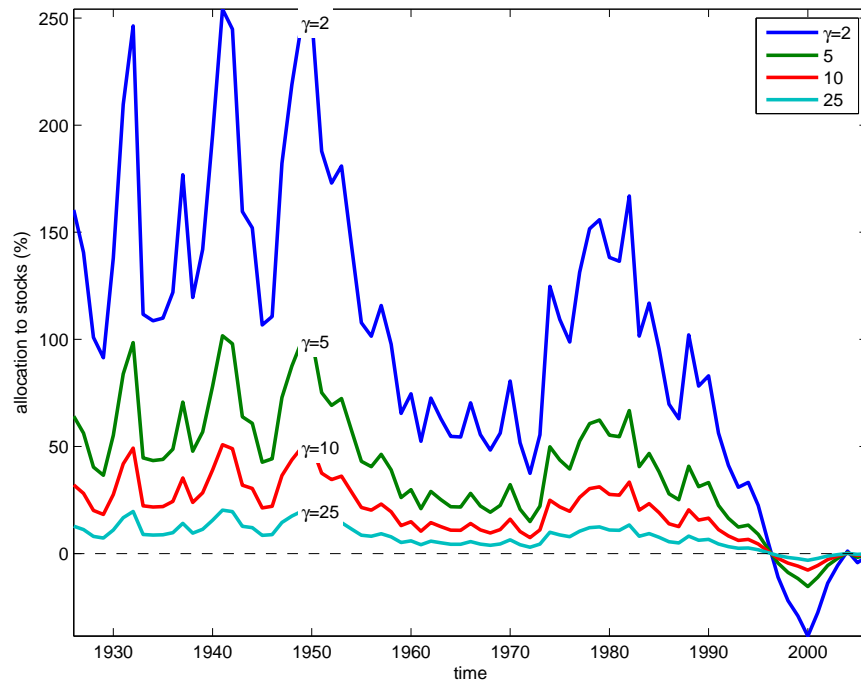
$$\begin{aligned} R_{t+1}^e &= \begin{matrix} -7.20 \\ (t) \end{matrix} + \begin{matrix} 3.75 \\ (2.66) \end{matrix} \frac{D}{P_t} + \begin{matrix} \varepsilon_{t+1} \\ \sigma_\varepsilon^2 = 19.81\% \end{matrix} \end{aligned} \quad (44)$$

---

<sup>26</sup>It isn’t really as easy as all this.  $\gamma$  is not really the power in the utility function. It’s the local curvature of the value function  $\gamma = -WV_{WW}/V_W$  where  $V(W, y)$  is the maximized value of the objective function. To really solve this problem, you have to solve a dynamic program, and there is no closed form so it has to be numerical. But this shortcut gives you results that are indistinguishable to the naked eye.



Market timing portfolio allocation. The allocation to risky stocks is  $\alpha_t = \frac{1}{\gamma} \frac{E_t(R^e)}{\sigma^2}$ . Expected excess returns come from the fitted value of a regression of returns on dividend yields,  $R_{t+1}^e = a + b(D_t/P_t) + \varepsilon_{t+1}$ . The vertical lines mark  $E(D/P) \pm 2\sigma(D/P)$ .

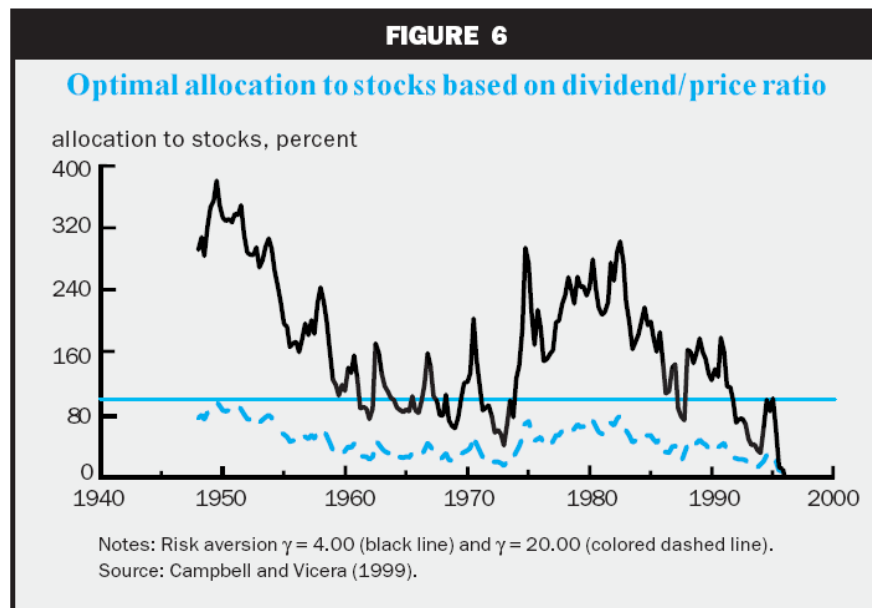
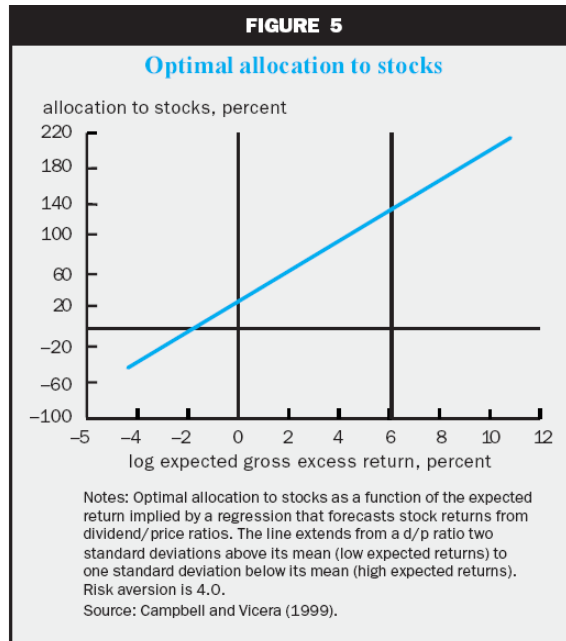


Market timing portfolio allocation over time. The allocation to risky stocks is  $\alpha_t = \frac{1}{\gamma} \frac{E_t(R^e)}{\sigma^2}$ . Expected excess returns come from the fitted value of a regression of returns on dividend yields,  $R_{t+1}^e = a + b(D_t/P_t) + \varepsilon_{t+1}$ .

*This seems to imply incredibly strong market timing based on d/p!*



- (c) A more careful calculation by Campbell and Viceira. They do it right, and include the hedging demand as well. There is no formula – they do this by computer



- (d) This seems like awfully strong market timing! Do we believe this result? Some considerations later. But for the moment, it serves as important education: *If you put the return forecast from something as simple as a dp regression in a portfolio optimizer, it will tell you to take very strong advantage of the time-varying expected returns. You can expect even stronger results from more complex return forecasting strategies.* This is part of the more general “wacky weight” problem.
- (e) Hedging demands? They are important in Campbell and Viceira, making stocks more

attractive to long-term investors. I haven't found a way of making the calculation in a simple way, i.e. without solving the whole dynamic program to find  $\eta$ .

(f) Reality? *This is what hedge funds do. Except without hedging demands! And their portfolio optimizers blow up too.* They just do much bigger regressions, many returns at the same time, models of conditional variance, high frequency, lots and lots of right hand variables.

5. Note: The danger of all portfolio analysis. *The average investor holds the market. Why doesn't everyone do this? And if they did, the phenomenon would disappear.* Well, maybe the average investor has outside income that tanks just when the DP opportunity looks good. And we left outside income  $y$  out of the problem. OK, but if so are you sure you're not in the same boat? Maybe portfolio theory should be "what is the question so that 'do nothing' is the answer!"

## 32 "Mean-Variance Benchmark"

1. The idea: Hedging demands are really important in these problems (long bond vs. short bond). But people are very reluctant to use hedging demands. Rightly so, they're black boxes. A solution?

2. Example: We didn't need hedging demands to understand long-term bonds. :

(a) "Buy long term (indexed, riskfree) bonds because their one-year returns covary negatively with shocks to their reinvestment risks"

(b) "Buy long term bonds because at a long horizon, they pay the same amount no matter what happens"

3. Why not look at long horizon portfolios directly? In a very simple example,

$$\max E[u(W_T)] \quad s.t. E(R_T), cov(R_T)$$

(a) Dynamic trading strategies (for example, exploiting D/P) become ways of generating more complex  $R_T$ . Changing weights over time is the same as changing weights between securities. It's also the same as a static investment in a fund that changes weights over time. The dynamic trading problem is not any easier, it's just brushed under the rug, or assigned to active funds, and then you make a static choice between active funds

(b) People don't want a single long horizon. They want streams of dividends,

$$\max E \sum_{t=0}^{\infty} \beta^t u(c_t)$$

$c_t$  is the dividend stream from an investment strategy

4. The paper idea

$$E \sum_{t=0}^{\infty} \beta^t u(c_t) = \sum_{t=0}^{\infty} \pi_t \beta^t u(c_t) = \tilde{E}[u(c_t)]$$

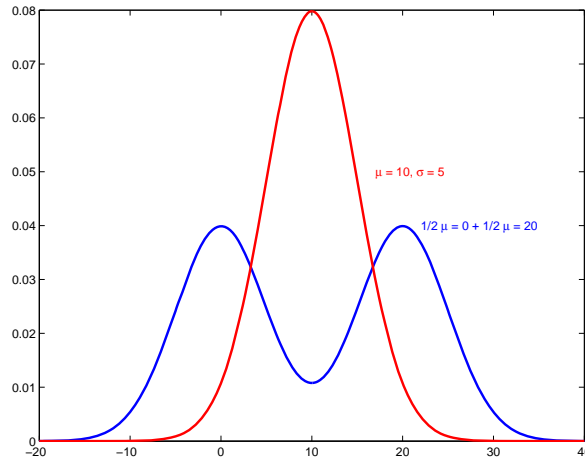
Treat time and probability symmetrically, then the long horizon problem is the same as a one-period problem, with the stream of payoffs (dividend) in place of the one-period return

5. Results: With quadratic utility, even with all sorts of dynamics  $E_t(R_{t+1})$  etc.
  - (a) Investors choose a dividend stream  $\{c_t\}$  on a “long-run” mean-variance frontier.
  - (b) A “long-run” CAPM describes long-run expected returns.
  - (c) State variables for investment opportunities drop out! (They are used to form dynamic trading strategies that expand the set of long-run opportunities, just as value or momentum expand the set of one-period opportunities.)
  - (d) Outside income does not drop out. First, short a portfolio that hedges your outside income, then invest in a long-run mean-variance efficient portfolio
  - (e) A long-run multifactor model with average outside income hedge payoffs as the second factor.
6. Limitations: Quadratic utility is an awful approximation. So it won't be that useful for quantitative analysis
7. The chat. There is a huge amount that asset management can do, as these problems are not easy.

### 32.1 Bayesian portfolios, parameter uncertainty, and taming market timing.

1. Intuition: The market timing solution has wild variation in weights over time. We don't really trust DP regressions that much, so we should tone down their advice. How can we make this intuition into an equation?
2. In D/P regression, the t stat is 2.66 in 70 years of data. We really don't know the coefficient that well. How does this affect the portfolio? In  $w = E(R^e) / (\gamma\sigma^2(R^e))$ , we don't really know the equity premium that well. Yet the portfolio calculation treats these numbers as if we really know them. How does *uncertainty about the inputs* affect the optimal portfolio?
3. A1: It's not obvious uncertainty matters. It's not as easy as “add standard errors.” Standard errors are symmetric. They say a wild portfolio is still your best guess, but point out maybe you should be 300% long or 100% long (standard errors) rather than 200% long (best guess). But 200% is still the best guess. Maybe we're not doing enough, maybe too much; maybe the best decision is to ignore the uncertainty.
4. A2: *Parameter uncertainty is a source of risk to investors, and says to be more conservative.*
  - (a) Intuition says that : Uncertainty should make us shade back to “hold the market”, “take less risk.” How do we incorporate that (excellent) intuition in a formula? *We have to capture the fact that uncertainty about parameter values is a real source of risk.*
  - (b) A very simple example. Suppose you know the standard deviation of returns is 18%. You think there is a 50% chance the equity premium  $E(R^e)$  is 6% and a 50% chance that the equity premium is 2%
    - i. It *is true* that the “equity premium” you should use in this context is the average of the two possibilities,  $E(R^e) = E(E(R^e)) = 4\%$ .

- ii. This does *not* mean that *standard deviation* of the returns you face is 18%. *Uncertainty about the mean is like more standard deviation.*
- iii. The distribution of returns is a two-hump camel that is wider than each of the humps separately.



5. How to do this in real problems? Bottom line preview: *Add the standard error of the mean to the variance of the return..*

$$\text{risky share} = \frac{1}{\text{r.a.}} \times \frac{\text{expected return} - \text{riskfree rate}}{\text{return var.} + \text{uncertainty about E return} \Leftarrow$$

$$w = \frac{1}{\gamma} \frac{E(R^e)}{\sigma^2(R^e) + \sigma^2(E(R^e))}$$

*parameter uncertainty makes you take less risk*

6. Where does this come from?

- (a) What we did before:

$$\begin{aligned} & \max E u(W_T) \\ & \max \int u(W_0 R_T) f(R_T | \theta) dR_T \end{aligned}$$

$\theta$  = parameters,  $E(R)$ ,  $\sigma(R)$ ,  $E_t R_{t+1} = a + bDP_t$ , etc. *We pretended to know the parameters for sure.*

- (b) But we don't know the parameters. We really want to do this

$$\max \int u(W_0 R_T) f(R_T) dR_T$$

where  $f(R_T)$  measures your total uncertainty about returns.

- (c) Intuition: *Try all different values of the parameters, weighted by the chance you think those parameters are right. (You can use standard errors for  $f(\theta)$ )*

- (d) What  $f(R)$  means is this: Given we describe the distribution of returns by parameters  $\theta$ ,  $f(R_T) = \int f(R_T|\theta)f(\theta)d\theta$ . Thus we want to do this

$$\max \int u(W_0 R_T) \left[ \int f(R_T|\theta)f(\theta)d\theta \right] dR_T$$

In words: *try all different values of the parameters, weighted by the chance you think those parameters are right. (Use standard errors for  $f(\theta)$ )*

7. How much should you allocate to stocks, given that you don't know the mean  $\mu$ ? if your uncertainty about the mean is also normally distributed we don't have weird two-hump distributions, we have again a nice normal but with a bit more variance.

**Fact:** *If  $R \sim N(\mu, \sigma^2)$  and  $\mu \sim N(\bar{\mu}, \sigma_\mu^2)$ , then  $f(R_T)$  is normal with mean  $\bar{\mu}$  and variance<sup>27</sup>  $\sigma^2 + \sigma_\mu^2$ . Simple rule: *Add the standard error of the mean to the variance of the return.**

8. Thus, recognizing parameter uncertainty, the optimal portfolio weight is

$$w = \frac{1}{\gamma} \frac{E(R^e)}{\sigma^2(R^e) + \sigma^2(E(R^e))}$$

*parameter uncertainty makes you take less stock market risk.*

9. Example

$$w = \frac{1}{\gamma} \frac{E(R^e)}{\sigma^2(R^e) + \frac{1}{T}\sigma^2(R^e)} = \frac{1}{\gamma} \frac{E(R^e)}{\sigma^2(R^e)(1 + \frac{1}{T})}$$

*Parameter uncertainty makes stocks riskier for long term investors. T = length of data/horizon. Intuitively, uncertainty about the mean return generates no extra risk for a day. But it generates lots of extra risk over 20 years. A calculation, after Barberis (2000).<sup>28</sup>*

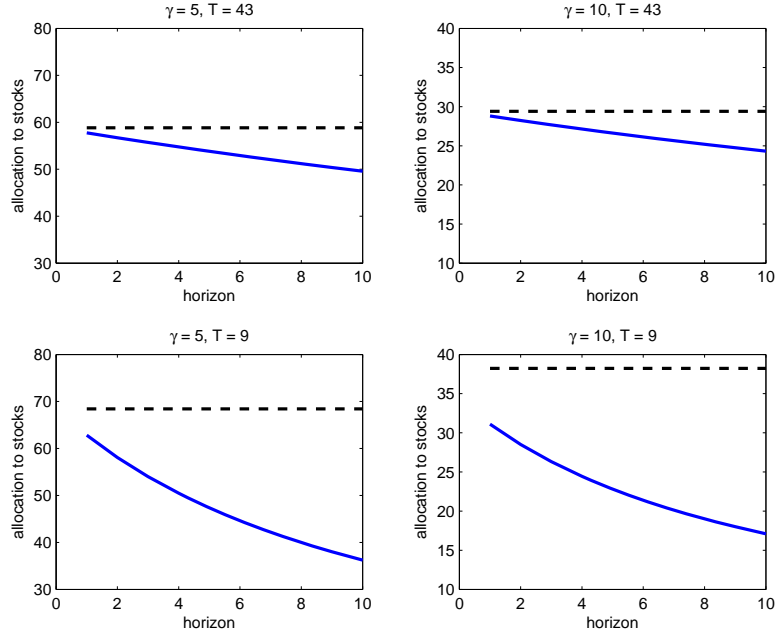
---

27

$$\begin{aligned} f(R|\mu) &= \frac{1}{\sqrt{2\pi \|\Sigma\|}} e^{-\frac{1}{2}(R-\mu)'\Sigma^{-1}(R-\mu)} \\ f(\mu) &= \frac{1}{\sqrt{2\pi \|\Sigma_\mu\|}} e^{-\frac{1}{2}(\mu-\bar{\mu})'\Sigma_\mu^{-1}(\mu-\bar{\mu})} d\mu \\ f(R) &= \int f(R|\mu)f(\mu)d\mu \\ f(R) &= \frac{1}{\sqrt{2\pi \|\Sigma + \Sigma_\mu\|}} e^{-\frac{1}{2}(R-\bar{\mu})'(\Sigma + \Sigma)^{-1}(R-\bar{\mu})} \end{aligned}$$

The last equality takes about a page of algebra to establish. You complete the square and integrate out the  $\mu$ . You don't actually do the intergral, you just express the integral of a normal distribution, which is one.

<sup>28</sup>This is a replication of his Figure 1. Barberis, Nicholas, 2000, "Investing for the Long Run when Returns Are Predictable," The Journal of Finance, 55,. 225-264, <http://www.jstor.org/stable/pdfplus/222555.pdf>

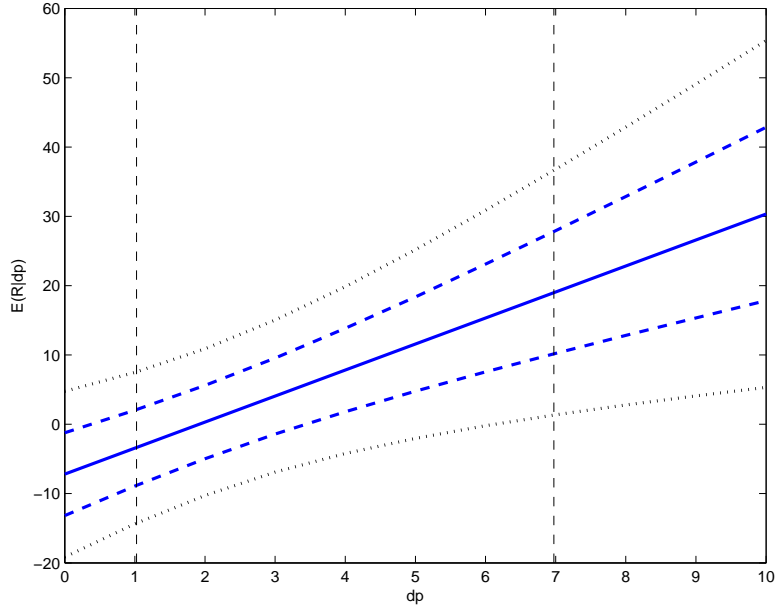


Portfolio allocation to stocks with parameter uncertainty. The solid lines present the case with parameter uncertainty, and the dashed lines ignore parameter uncertainty. The allocation to stocks is  $w = \frac{1}{\gamma} \frac{\mu - r}{\sigma^2 \left[ 1 + \frac{(\gamma - 1)h}{\gamma T} \right]}$ .  $\mu = 0.06, \sigma = 0.1428$  for  $T = 43$  and  $\mu = 0.078, \sigma = 0.151$  for  $T = 9$  with  $r = 0$ . (The extra  $(\gamma - 1)/\gamma$  allows for continuous rebalancing. See “Portfolio Theory.”)

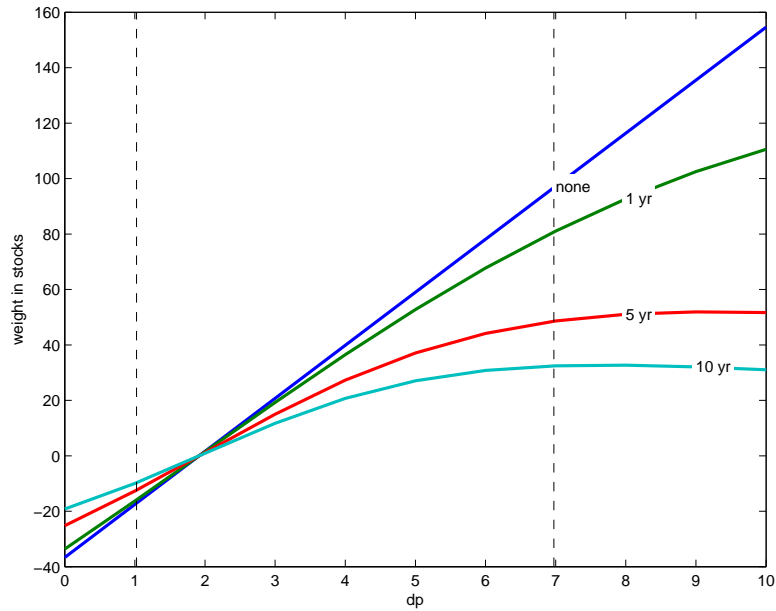
10. We started with the classic result that the horizon doesn't matter. We started thinking that mean reversion makes stocks *more* attractive to long run investors. Parameter uncertainty makes them *less* attractive. Which consideration wins? If you have a very long sample so you know the parameters, stocks are more attractive. If you have a short sample, or think the world has changed, this consideration wins. (And then there are other considerations, like the stock or bond-like nature of the outside income stream.)
11. We can use the same idea for allocation over time. *Recognizing the uncertainty about variation in  $E(R)$  over time leads you to shade portfolios back towards constant weights.* The central point is that  $E_t(R^e)$  is less well known as DP gets further from historical experience.
12. Example.

$$R_{t+1}^e = a + b(D_t/P_t) + \varepsilon_{t+1}$$

$$\sigma^2 [E_t(R_{t+1}^e)] = \sigma^2(\hat{a}) + \sigma^2(\hat{b})(D_t/P_t)^2 + 2cov(\hat{a}, \hat{b})D_t/P_t \quad (45)$$



Expected excess returns as a function of dividend yield, with one and two standard error bands. Vertical lines are the mean dividend yield plus and minus two standard deviations.

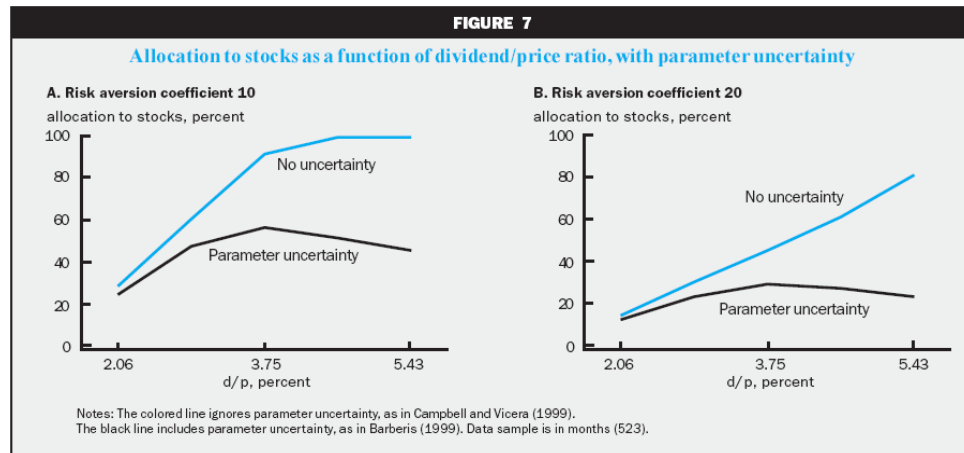


Optimal allocation to stocks given that returns are predictable from dividend yields, and including parameter uncertainty, for  $\gamma = 5$ .

$$w = \frac{1}{\gamma \sigma^2(\varepsilon) + \text{horiz} \times \sigma^2 [E_t (R_{t+1}^e)]} \hat{a} + \hat{b}(D_t/P_t)$$

Note it is the *widening* of the error bands that matters here. We are *more uncertain* about what mean corresponds to very high/low dp. *Functional form* uncertainty is the real issue. (Adding squared terms or more flexible right hand side would widen things up a lot more)

13. A much more precise (but hard) calculation by Barberis (2000)

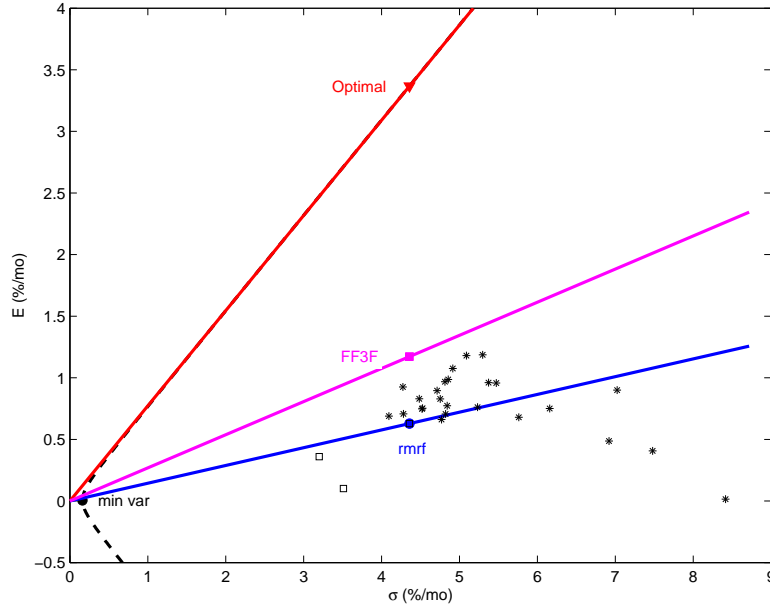


## 32.2 Wacky weights in mean-variance analysis

Similar incredibly strong advice comes from standard portfolio optimizers. Let's study this issue in a simple context.

1. Example 1. FF 25 portfolios + 3 factors. 20 years of data (a lot!)





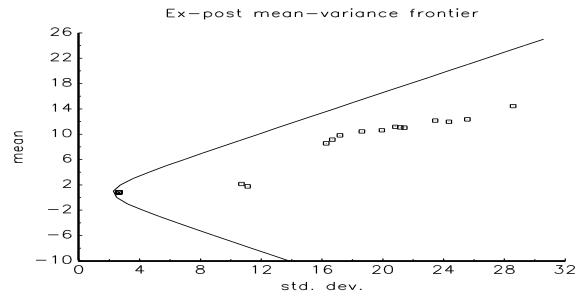
Mean-variance optimization with excess returns of the Fama French 25 size and book/market portfolios, together with the 3 Fama French factors. “Optimal” is the mean-variance optimal portfolio, at the same variance as the market return. “FF3F” is the mean-variance optimal combination of the 3 Fama-French factors. The stars are the means and standard deviations of the individual portfolio returns. The squares are the means and variances of the factors. The thin black line gives the mean-variance frontier when weights sum to one. Min var is the minimum variance excess return with weights that sum to one. Mean and covariance estimates based on 20 years of data.

$$\text{“optimal”}: w = \frac{1}{\gamma} \Sigma^{-1} \mu$$

	low	2	3	4	high
small	-149	51	69	96	52
2	-19	-57	190	-13	-60
3	29	-34	-31	-93	41
4	116	-39	-42	35	-2
large	87	-19	8	-22	2
rmrf	hml	smb			
-94	77	-69			

The weights are wild. The improvement in mean-variance performance is unbelievably large.

2. Example 2: Mean-variance frontier from size portfolios using historical return data

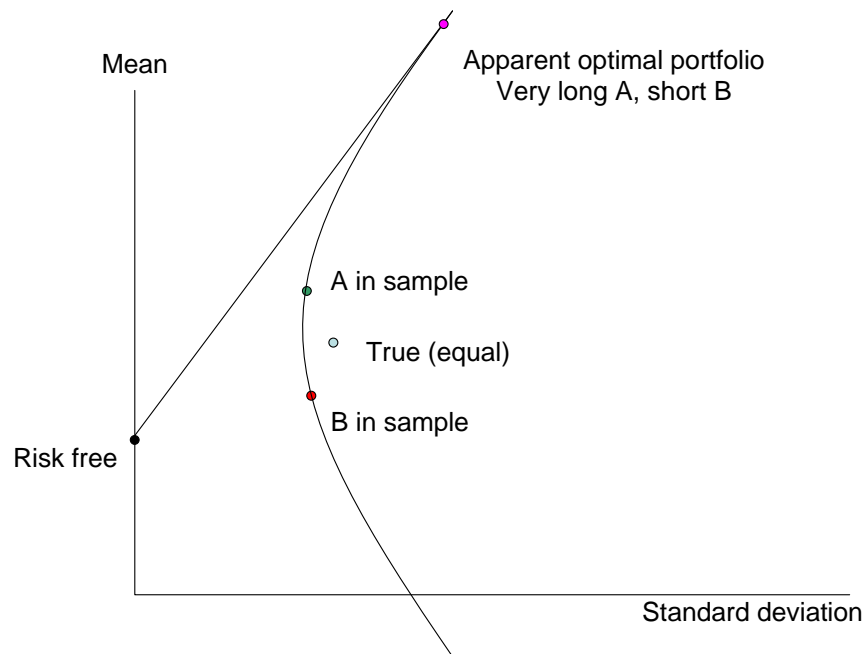


Weights to achieve “optimal portfolio” (minimum variance at VW mean)

VW	EW	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	GB	CB	TB
-2093	-577	113	94	-21	44	123	184	40	268	509	1362	-144	118	79

This is nuts.

3. What’s going on? Here is the intuition:



*Portfolio optimization problems are very sensitive to mean return inputs, especially when there are highly correlated assets.*

### 32.3 Bayesian portfolio theory to tame wacky weights

1. What do we do to tame wacky weights?
  - (a) Short constraints are often used, but they are a poor solution to the problem. In the example, just stopping at 100% A, 0% B isn't the right answer, 50/50 is the right answer.
  - (b) Lots of ad-hoc rules are used to constrain portfolios. Don't put more than x% in any single issue, insist on industry buckets, etc. Why does this make sense?
  - (c) But.. you also miss the opportunities this way! (A good recent example. HF traders hit limits in the flash crash, and got out. Thus missing the buy opportunity of a lifetime. Example 2: with short limits you turn down 3 com/ palm even if offered to you on a platter.)
  - (d) Standard rules fail when faced with hedge funds/derivatives. 10% exposure in a  $\beta = 10$  investment is 100% exposure, so limiting exposure to 10% *weight* doesn't limit the *risk*. (Many common rules of thumb implicitly assume long-only stock investments, and can fall apart when applied to "alternatives") .
2. The Bayesian idea of the last section does not really help. For iid returns, the standard error is just the covariance matrix over  $T$ , so if the covariance matrix is  $\Sigma$ , the real risk is  $\Sigma + \frac{1}{T}\Sigma$ . This consideration doesn't change or solve the wacky weight problem, it just scales down investments a bit.

$$w = \frac{1}{\gamma} \left( \left( 1 + \frac{1}{T} \right) \Sigma \right)^{-1} E(R) = \frac{1}{\gamma} \frac{T}{T+1} \Sigma^{-1} E(R).$$

We need a more involved solution – one that recognizes uncertainty in the covariance matrix not just the mean.

### 32.4 (Optional) Black-Litterman Bayesian portfolios

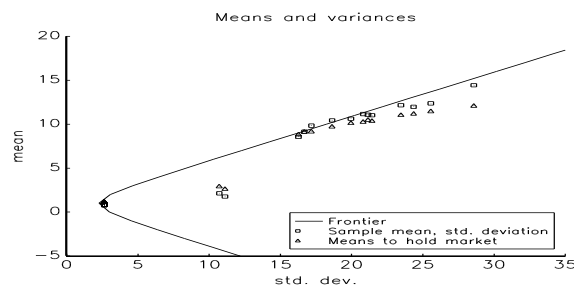
1. *Everything that follows is optional*: not covered in class, since it's popular but really doesn't work that well.
2. A better solution (Black-Litterman; Pastor-Stambaugh): Here's the insight:
  - (a) Portfolio problems are very sensitive to the input *means or alphas* (given there is a lot of positive correlation in the covariance matrix.)
  - (b) The CAPM says hold the market. If you used  $E(R^{ei}) = \beta_{im}E(R^{em})$  as the *input*, the portfolio optimizer would give market weights as the *answer*. Or, reexpressing portfolio theory in terms of CAPM and alphas, if you used  $\alpha_i = 0$  as the input, the optimizer says just to hold the market as the output.
  - (c) So modify your *inputs*  $E(R)$  or  $\alpha$  in the direction of the CAPM to tame the weights.
  - (d) There is nothing in the portfolio procedure so far that looks even vaguely at market weights as a benchmark! If the mean and variance look good, the optimizer will say to put all your money in a stock with \$1 million total capitalization. Find a way to tell the optimizer that the index is a pretty good idea and to deviate from that only with a good reason.

3. If your *inputs* to the portfolio problem are as given by the CAPM, then the *output* of the portfolio problem must be “just hold the market portfolio,” i.e., “hold each asset in its market weight.”

$$E(R^{eA}) = \beta_{A,m}E(R^{em}) \rightarrow \text{hold market}$$

Check: use CAPM (“hold market row”) mean returns  $\beta_{A,m}E(R^{em})$  in the portfolio calculation rather than the actual mean returns. The inputs are not very different!

	VW	EW	D1	D2	D3	D4	D5	
Actual	9.14	11.07	14.44	12.36	11.97	12.16	11.02	
Hold market	9.14	10.50	12.09	11.47	11.19	11.05	10.38	
	D6	D7	D8	D9	D10	GB	CB	TB
Actual	11.16	10.62	10.43	9.83	8.56	1.76	2.14	0.81
Hold market	10.27	10.14	9.72	9.17	8.82	2.62	2.92	1.15



Notice that *small changes in the mean return assumption gave huge changes in the optimal portfolio advice.* .

4. Now here’s the idea: use *inputs* that shade between sample averages, model predictions, investor views, etc. and the CAPM. Use *inputs for asset A*

$$E(R^{eA}) = \beta_{A,m}E(R^{em}) + \lambda\alpha_A$$

with  $\lambda < 1$ . Thus, in the version of portfolio theory that separates between market exposures and alpha exposures, you keep the same market exposure, but lower the alpha exposure

$$w = \frac{1}{\gamma} \frac{E(R^{em})}{\sigma^2(R^{em})}$$

$$w_\varepsilon = \frac{1}{\gamma} \Sigma_\varepsilon^{-1} (\lambda\alpha)$$

5. *How much  $\lambda$ ?* You can just do it until the portfolio makes sense. But “Bayesian portfolio theory” gives a way to think about it.

- (a) Intuition: How much to lower the weights depends on your “prior” about how good the CAPM (hold the market portfolio) is vs. sampling or other sources of uncertainty in the  $E(R^{eA})$  or  $\alpha$  etc. inputs. Are you pretty sure this is a real  $\alpha$ ? Raise  $\lambda$ . Are you pretty sure that the CAPM (or FF3F) is right and your  $\alpha$  are noise? lower  $\lambda$ .
- (b) *A Formula.* Suppose you measure  $\alpha_i$  with  $\sigma_{\alpha_i}^2$ . (This is the OLS standard error, or otherwise your confidence on this measurement or opinion.) We also measure your “confidence in the CAPM” by a measure  $\sigma_p$ ; before you started how big did you think alphas really could be? (This is your “prior distribution.”) Then, the Bayesian alpha – the one you use in the portfolio model – weights the two pieces of information (the estimate and your prior) inversely to their accuracy

$$\hat{\alpha}_i = \frac{\frac{\alpha_i}{\sigma_{\alpha_i}^2} + \frac{0}{\sigma_p^2}}{\frac{1}{\sigma_{\alpha_i}^2} + \frac{1}{\sigma_p^2}} = \frac{\frac{1}{\sigma_{\alpha_i}^2}}{\frac{1}{\sigma_{\alpha_i}^2} + \frac{1}{\sigma_p^2}} \alpha_i = \frac{\sigma_p^2}{\sigma_{\alpha_i}^2 + \sigma_p^2} \alpha_i$$

(I added 0 so you could see in general how to mix any two pieces of information.)

6. *Intuition.*

- (a) If you were dogmatically sure the CAPM was right,  $\sigma_p^2 = 0$ , and  $\hat{\alpha} = 0$ . You are so sure the CAPM is right, that no evidence can change your mind.
- (b) If you have no idea at all whether the CAPM is right, i.e you are willing to accept *any* size alpha with no quibbles, then  $\sigma_p^2 = \infty$ , and  $\hat{\alpha} = \alpha$ . With no prior idea, take the estimate and go. (And using the estimate alone is *equivalent* to having no prior at all that the CAPM might be right.)
- (c) The greater  $\sigma_{\alpha_i}^2$  – the less sure you are of the measurement – the less you use the measurement in your final  $\hat{\alpha}_i$

7. *Practical advice.* This is a famous formula and widely advocated. If you just use standard errors  $\sigma_{\alpha}^2 = \sigma^2(\varepsilon)/T$  on sample estimates with a common prior, it doesn’t work very well. The problem is, it shrinks all the  $\alpha$  by about the same amount, so you get the same wacky weights as before, just smaller. I think it’s better to set all the alphas to zero except a few that you really care about, and the formula may be more useful for mixing alphas with much different “standard errors”

8. Note: Obvious extension to multifactor models; either in the “factor bets” (how certain are you about momentum?) or the remaining alphas.

9. (Optional) *Many assets:*  $\lambda_i$  can be different for different  $\alpha_i$ .  $\lambda_i =$  “how much you trust  $\alpha_i$ ” ( $\neq 0$ ). It will be lower for high standard-error  $\alpha_i$ . You want to shade badly-measured alphas *more* back to the CAPM than the others. But using the last formula directly would be a mistake – it ignores cross-correlation, and thus the evidence about  $\alpha_i$  in a measured  $\alpha_j$ . To do this right, we use the vector-matrix version,

$$\hat{\alpha} = \left( \Sigma_{\alpha}^{-1} + \Sigma_p^{-1} \right)^{-1} \left( \Sigma_{\alpha}^{-1} \alpha + \Sigma_p^{-1} 0 \right)$$

(Again, I'm giving you the general formula. 0 is a vector of zeros.)  $\Sigma_a$  is the *covariance* matrix (standard error) of the alphas. In a time-series regression this is just  $\Sigma/T = \text{cov}(\varepsilon, \varepsilon')/T$ .  $\Sigma_p$  is the *covariance* matrix representing your confidence in the CAPM. You might be more willing to entertain larger alphas for small, illiquid stocks than for large liquid ones. Usually, you'll just use something simple like

$$\Sigma_p = \begin{bmatrix} \sigma_p^2 & 0 & 0 \\ 0 & \sigma_p^2 & 0 \\ 0 & 0 & \sigma_p^2 \end{bmatrix}$$

### 32.5 Optional: Taming covariance matrices

(Optional, not in lecture)

1. Covariance matrices also tend to be over fit – they display more correlation than is really there.
2. Example: Look at the minimum variance portfolios above. They were obviously ridiculous. There is no way to get that riskfree a portfolio by holding stocks.
3. Example: Two assets, two data points. Then the *sample* covariance matrix shows that two returns are always perfectly correlated! It suggests a perfectly riskless long-short portfolio. With 6000 assets, 600 data points., the problem is worse. There are thousands of apparently riskless portfolios.
4. Solution 1: Make the diagonals bigger

$$\Sigma = \lambda \Sigma + (1 - \lambda) \begin{bmatrix} \sigma_1^2 & & \\ & \sigma_2^2 & \\ & & \sigma_3^2 \end{bmatrix}$$

5. Solution 2: Use large factor models (including non-price factors) to structure the covariance matrix.

$$\begin{aligned} R_t &= BF_t + \varepsilon_t \\ \Sigma &= B \Sigma_f B' + \text{Diagonal} \end{aligned}$$

This avoids spurious  $\varepsilon$  correlation.

### 32.6 Summary of Bayesian insight.

Models of the mean and variance are not perfectly fit. This fact does more than add standard errors – you should optimally shy away from models based on the uncertainty in their parameters.

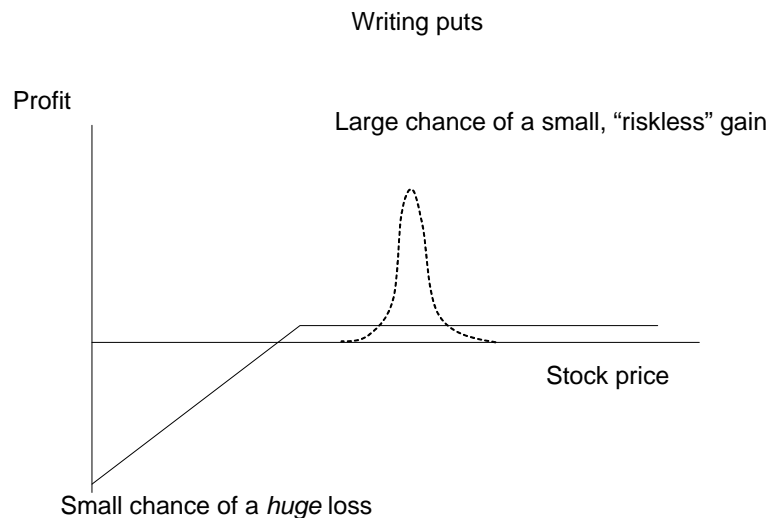
This insight constitutes a first big doubt about all the fun we've had today. Maybe “hold the market” isn't such bad advice?

It is also incredibly important if you ever go work for a hedge fund. There is a tendency to fit complex models *and then treat the estimates as perfectly known* when doing portfolio calculations.

If you find yourself with a big black box portfolio optimizer, it's generating ridiculous wacky weights, then you're patching it up with glue and duck tape (short constraints, etc.) a) welcome to the world b) step back and think a bit better about what you're doing!

(I'm not persuaded that Bayesian portfolio theory answers all these questions perfectly. For example, it only quieted down the wacky weight of very strong market timing advice if the uncertainty was *bigger* in the large and small values of  $\delta p$ . Standard error lines parallel to the estimation line made no difference. Applied to a sample estimate, Black-Litterman seems to just scale all the wacky weights down, but not change their pattern. Still, this is the best conceptual framework I know of, and it sure beats programming up wacky weights that you know are silly and then taming them with short restrictions or other ad hoc fixes. The perfect portfolio theory still lies ahead.)

## 32.7 Options



- Write put options? The premiums are high!
- But *the average investor holds the market*. Why are put option premiums so high?
- A: there is a lot of demand to *buy* put options despite the high premium. Very sensibly!

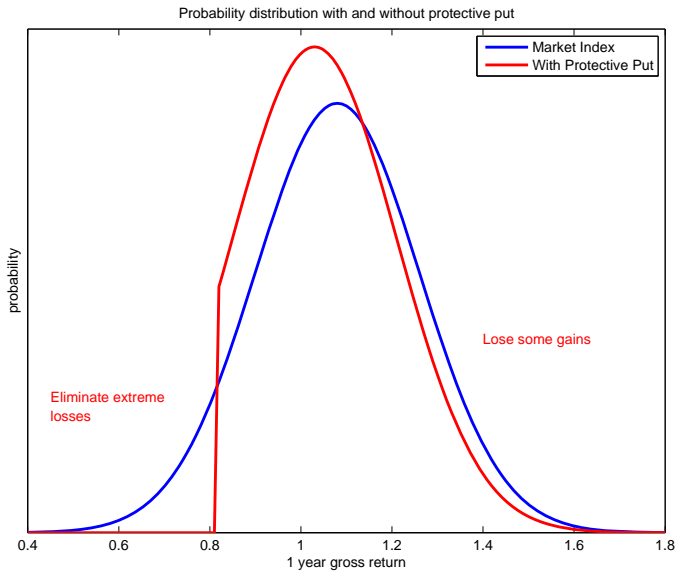
1. Some common feelings:

- (a) Leverage. “If we lose more than 20%, we default on our debt”
- (b) “If we lose more than 20% we have to cut core functions.”
- (c) “If we lose more than 20%, our sponsors will give up and fire us.”

These sentiments are investors with “habit” style utility functions, where the utility declines very steeply for losses. A market full of people like this will generate high prices for out of the money put options (and the non-normal or stochastic volatility of stock

returns that will undo the Black-Scholes formula) If you have a utility function that is less averse to big losses than these, then you will profit by selling out of the money puts – you will gain more in the good states (weighted by marginal utility) than you lose in the bad states.

- Put options are attractive to these investors – despite the cost



It’s all about whether you like the tail risk on the left vs. the slight leftward shift of the rest of the curve.

- Note: *A stop loss order is not the same as a put option.* Many investors think they can cut off the left tail by planning to sell on the way down with stop loss orders. Then, they cut off the return distribution without paying the put premium – without shifting the peak of the red curve to the left. As you can imagine, this is a classic fallacy. You don’t get something for nothing.
- Buy vs. write options? Receive premium or buy insurance despite the large premium?  
A: Are you more or less able to take crash risks than the average investor? Really, now?

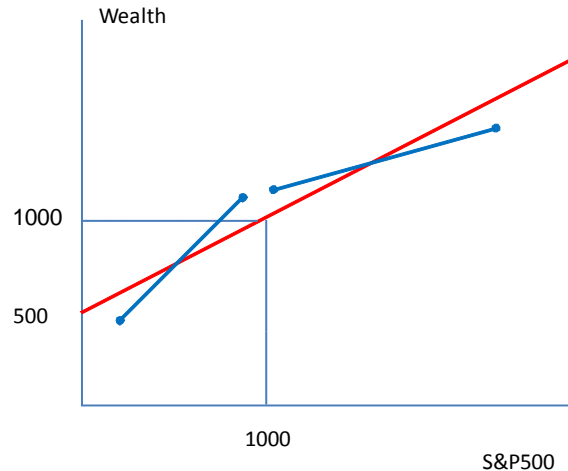
- Big picture: You can tailor the entire shape of the return distribution to take account of your ability to take risks of any size, and the premiums offered to you for doing so.

- Example 2: Should we rebalance?

- Wall Street wisdom says “always rebalance.”
- But this is a great puzzle to me, because *the average investor can’t rebalance.* If the market is 60/40, stocks go up so wealth is 80/20, and our investor wants to rebalance, who is he going to sell to? For him to hold 79/21 someone else needs to hold 81/19! We can’t all rebalance! (My “Two trees” is an academic paper aimed at this puzzle.) For now, here’s an insight:
- Rebalancing is just like writing both call and put options.* It gives a better payoff if stocks don’t move much, at the expense of worse payoff if stocks move a lot.



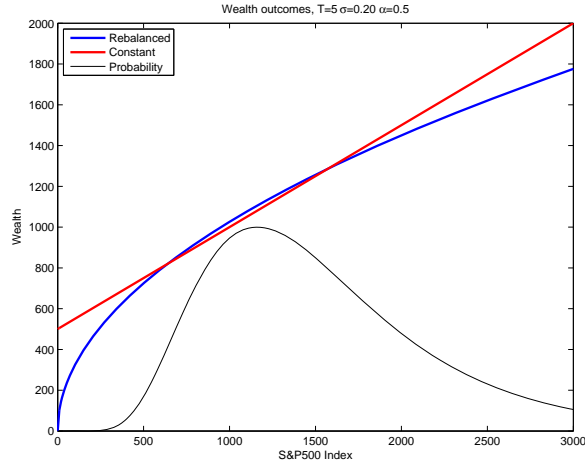
4. Here's the essence of the situation. Suppose the S&P500 index is at 1,000, you have \$1,000, and you want a 50/50 allocation. If you do not rebalance, your wealth follows the red line. If the S&P500 goes to 0, you have \$500 left; as it rises you get \$0.50 for every \$1 that the S&P500 rises.



Now suppose the S&P500 falls to 500. You now have \$500 bonds and \$250 stocks = \$750, and your weights are 33/66. If you rebalance to 50/50, shown in the blue line, you have \$325 bonds and \$325 stocks. Now, for each 1 point change in the S&P500, you gain  $325/500 = 0.65$  dollars rather than 0.50 dollars. If the S&P 500 comes back to 1,000, you have  $750 + 325/500 \times 500 = \$1075$ . However, if the S&P 500 goes down more, you lose more than you would have otherwise. The same thing happens in the other direction if the S&P500 rises and you rebalance out.

In sum, rebalancing means your wealth will end up on the blue dots rather than the red line after two moves. You're better off if the S&P 500 ends up near where it started, you're worse off if there are big moves. You're betting on low volatility. This is just like selling out of the money puts and calls.

5. Next, I programmed up a simulation of what happens if you *constantly* rebalance<sup>29</sup>. Here's a graph



- (a) Again you start with \$1000 and the S&P index is also 1000, and a 50/50 stock/bond split. Again, the red line shows what happens if you don't rebalance. If the S&P500 is still at \$1000, you have \$1000. If the S&P500 goes down to zero, you have \$500. If the S&P500 goes up to \$2000, you have \$1500. Every dollar change in S&P500 means 50¢ increase in wealth. (I assume zero interest rate and an expected stock return of 7% for all graphs.)
- (b) The blue line shows what happens in 5 years if you constantly rebalance. If the S&P declines, you can now drive the portfolio all the way to zero. As the S&P rises, you slowly get out, so the gains are slower than they were before. This looks like an option payoff diagram, and it is: This is the payoff from S&P if you also write both puts and calls.

<sup>29</sup>You need a few dz and dts to do this:

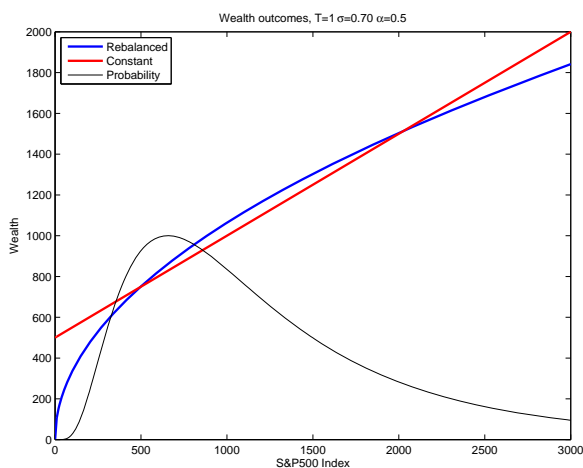
$$\begin{aligned} \frac{dS}{S} &= \mu dt + \sigma dz \\ d \log S &= \mu dt + \sigma dz - \frac{1}{2} \sigma^2 dt = \frac{dS}{S} - \frac{1}{2} \sigma^2 dt \\ \frac{S}{S_0} &= e^{\mu - \frac{1}{2} \sigma^2 dt + \sigma \sqrt{t} \varepsilon}; \quad \varepsilon \sim N(0, 1) \\ \frac{dW}{W} &= \alpha \frac{dS}{S} + r^f dt \\ d \log W &= \frac{dW}{W} - \frac{1}{2} \frac{dW^2}{W^2} = \alpha \frac{dS}{S} + r^f dt - \frac{1}{2} \alpha^2 \sigma^2 dt \\ &= \alpha d \log S + \alpha \frac{1}{2} \sigma^2 dt + r^f dt - \frac{1}{2} \alpha^2 \sigma^2 dt \\ &= \alpha d \log S + \frac{1}{2} \alpha (1 - \alpha) \sigma^2 dt + r^f dt \\ \log W - \log W_0 &= \alpha (\log S - \log S_0) + \left[ r^f + \frac{1}{2} \alpha (1 - \alpha) \sigma^2 \right] t \\ \frac{W}{W_0} &= \left( \frac{S}{S_0} \right)^\alpha e^{[r^f + \frac{1}{2} \alpha (1 - \alpha) \sigma^2] t} \end{aligned}$$

(c) This looks like an insanely bad deal – how is it you do worse if the S&P goes up and worse if it goes down? Because you do better if it goes nowhere at all. If you implemented this using options, you would get the option premium, the put and call price. It’s a little more subtle if you implement it using rebalancing, but we all know dynamic trading is the same as explicit options. The key is that the S&P500 will never get back to 1000 by sitting at 1000 the whole time. If it gets back to 1000, it did that either by first falling and coming back, or by first rising and then falling again. In either scenario, you do better by rebalancing.

It’s surprising in the graph that the option premium is so small. In this case you only make \$25 extra at S&P=1000 by rebalancing. That reflects the numbers I put in. I assumed volatility  $\sigma = 0.20$  or 20%, typical of normal times, and the same assumptions as underlie the Black-Scholes formula. The *probability* of catastrophic declines or huge rises is quite small as graphed, so you don’t make that much more money in the high probability states near S&P = 1000 than you lose in the more extreme but much less probable states.

In normal times, the gain is less at a one year horizon – \$5 at S&P=1000 instead of \$25. That’s because the chance of implosion or explosion is even less. Option values rise with time to expiration and volatility. This is the next figure – you can hardly tell that the blue line is above the red line in the middle. It is, but only by \$5.

In wild times, the value of options has increased dramatically, as has the probability of the catastrophe on the left or the boom on the right. Here’s the graph at a one year horizon and vol of 70. You can see much better performance if the S&P ends up between 500 and 2000, but the much greater possibility of disaster or boom since the probability curve is more spread out.



6. Alright already, so should you rebalance? Rebalancing is in fact optimal—*if* you have power utility and live in a world in which stock returns are iid and lognormal. *They’re not*. If price changes are accompanied by expected return changes, then all bets are off. Remember, the average investor holds the market —without rebalancing! (I explore this a bit more in “two trees” on my webpage fyi)

## 32.8 More portfolio theory doubts

1. *The average investor holds the market portfolio.* For everyone who says “wow, that’s great, I’ll buy value stocks” there should be another person who says “you must be kidding. Those stocks all tank precisely when x happens (lose my job?). I want to *short* value (buy growth). I know I’ll lose a bundle in average return of my portfolio, but it’s worth it for me to buy “portfolio insurance” that my portfolio does not go down when x happens.” If not, the value effect is not a reward for risk and it will not last. *Where are all these investors?* (Maybe once they look at value stocks, e.g. GM, the fear on their faces reveals their risk exposure.)
2. Predictability: Why do expected returns vary in the first place? Why doesn’t everyone market-time? If they did, the low price/high average return would disappear. *The average investor holds the market.* You should only do this if you have less time-varying risk aversion, or time-varying risk (like labor income) than everyone else. Do you? Really?
3. Value, size, momentum: Why is there value in the first place? Why doesn’t everyone (try to) load up on value? If they did, the value effect would disappear. You should only do this if you have less value-exposure in other income than everyone else. Do you? Really?
4. Catch 22. Anomalies only last if you can’t use them. Why is there an  $\alpha$ ?
  - (a) Data-dredging? Then it was never there in the first place.
  - (b) Mispricing/overlooked? It will be gone quickly. (One week reversal).
  - (c) Mispricing, but you can’t arbitrage it due to constraints or transactions costs? You can’t trade it. .
  - (d) Mispricing, but it needs institutional reform? It needs institutional reform, you can’t trade it. (Small stocks, catastrophe bonds)
  - (e) A risk factor? Then it will stay, but it offers no advice for the average investor. If a mass of investors changes portfolios, the effect changes.
5. All portfolio problems violate the equity premium puzzle. The implied *consumption* is incredibly volatile. If you take the *portfolio advice*, why don’t you buy the *consumption advice*  $\sigma(c) = 16\%$ ? If the problem is misstated, then how do you know the portfolio advice is right?
  - (a) Details:

$$w = \frac{1}{\gamma} \frac{E(R_{t+1}^e)}{\sigma^2(R_{t+1}^e)}.$$

Thus, the risk aversion necessary to have the investor put all weight in stocks  $w = 1$  is

$$\gamma = \frac{E(R_{t+1}^e)}{\sigma^2(R_{t+1}^e)} = \frac{0.08}{0.16^2} = 3.125.$$

Q: What’s the deal with the “equity premium puzzle?” We seem fine with a sensible 3.125 risk aversion coefficient. A: This portfolio model has  $c_t = kW_t$  so  $\sigma(\Delta c) = \sigma(\Delta W) = 16\%$ ! But real consumption has  $\sigma(\Delta c) = 2\%$ . Or is something wrong with the setup?

- (b) (JC idea: People think there is a lot more mean reversion in stocks than our statistics tell us, so their consumption ignores “short term price fluctuations” Maybe they, not our statistics, are right – but this radically changes the portfolio advice)

6. Transactions costs! If you rebalance weekly and pay 1% roundtrip, you need 52% alpha! Transactions costs make the problem much harder, since then what you hold *now* determines what you should hold in the future – and before buying something you have to think about how soon will you need to sell it. Most maximizers don't handle this now.
7. Taxes! Astonishingly, most active management ignores taxes, i.e. short term capital gains, capital gains forgiveness at death, etc.
8. Even more realistic portfolio problems take account of
  - (a) Taxes, tax treatment of different investment vehicles. (For example, put bonds in the tax-free vehicle since you avoid taxation of interest; you can defer the capital gains in the taxable portfolio)
  - (b) Transactions costs. (This is a big headache. What you hold depends on what you happen to have, and what you buy then will become “what you happen to have” next time. Thinking about transactions costs leads you to weight signals that last longer more heavily.)
  - (c) Liquidity and liquidity needs. You may think you're the “long run investor” who doesn't sell at the bottom. Maybe you're not.

### 32.9 Some good things to say

1. There is a good reason for portfolio advice again!
  - (a) Optimal portfolio advice in 1962 was, you need a tailored portfolio, pay big fees.
  - (b) Optimal portfolio advice 1972: Forget that. There is  $\beta$  on market, and the two fund theorem. The only reason for active management is to chase zero sum alpha. Too bad for my salary and your job prospects.
2. But look at optimal portfolios today!
  - (a) Hedge labor/outside income risks
  - (b) Find desired “priced factor allocation” *rmrf*, *hml*, *smb*, *umd*, etc.
  - (c) Find the right distribution – put options etc.
  - (d) Chase alpha? (alpha=beta you don't understand)
  - (e) Offset implied factor premia in active part.
  - (f) There is a lot for professional advice to offer again! A big insurance market that needs “agents.”
3. *Investors should hedge their labor or proprietary income risks.* They need help. Picking appropriate styles is almost an excuse for tailored portfolio advice (and fees). Help them!
  - (a) This advice *does not* depend *at all* on the nature of “priced factors” (*hml*, *smb*, etc.) That question only determines (a bit) the *price* of this insurance, and the (academic?) question of how much somebody who feels no risk should *sell* insurance.

4. Investors do silly things. They trade too much (costs, taxes), invest with many active managers and wind up with the index, hold undiversified portfolios, ignore risks of nonmarket income, evaluate investments in isolation, don't know the betas of the overall portfolio, pay too many fees. We can help them to stop.
5. If you're going to work for a hedge fund, forget all my doubts and learn to do complex optimal portfolios with good risk management (covariance matrices).

### 32.10 “Real” portfolio problems.

A survey of real calculations, beyond my chatty approach.

1. Long-horizon investor,

$$\max E_t \sum_j \beta^j u(c_{t+j}) \text{ or } \max E_t \int e^{-\delta s} u(c_{t+s}) ds$$

2.  $u()$  captures feelings about big losses, losses when other people are losing, etc.
3. Solving this: Two-period problems are actually a good approximation. “Dynamic programming” lets you map all the future into the value of wealth tomorrow

$$\max_{\{c_{t+j}\}} E_t \sum_{j=0}^{\infty} \beta^j u(c_{t+j}) = \max_{\{c_t\}} u(c_t) + \beta E_t \sum_{j=0}^{\infty} \beta^j u(c_{t+1+j}) = \max_{\{c_t, W_{t+1}\}} u(c_t) + \beta E_t V(W_{t+1})$$

$V(W_{t+1})$  acts like a “utility function” over wealth.

4. Add Labor, proprietary or other income
5. Add “Size”, “value”, “momentum” (?) effects in the cross section of returns
6. Add Time-varying expected returns of stocks, bonds, etc.
7. Add Time-varying  $\Sigma$  matrix.
8. Add Options for nonlinear payoffs, crash protection.
9. Add *Transactions costs*. (The weights you choose today are the givens for tomorrow.)
10. Add *Taxes*.
11. Add “Bayesian” or other ways to avoid wacky answers that load up on small estimation errors and think they're real.
12. Lots to do! – But don't forget that the average investor holds the market!
13. If you're interested “Portfolio theory” and “A mean-variance benchmark for intertemporal portfolio theory” on my website are good places to dig deeper.

## 32.11 Portfolio theory summary

1. Portfolio theory problems: find the portfolio that maximizes utility. The basic problem:

$$\max_{\{c_t, w_t\}} E \sum_{j=0}^{\infty} \beta^j u(c_t) \text{ or } E \int_0^{\infty} e^{-\rho t} c_t dt;$$

$$\begin{aligned} W_{t+1} &= R_{t+1}^p (W_t + y_t - c_t); \\ R_{t+1}^p &= R_t^f + w_t' R_{t+1}^e \\ &E_t(R_{t+1}^e), \text{ cov}_t(R_{t+1}^e) \text{ given} \end{aligned}$$

The problem is often simplified, i.e.  $\max U(W_T)$ , no income  $y$ ,  $E_t(R_{t+1}^e)$  constant, etc. We build models to build intuition, so keeping it simple is important.

2. Results:

- (a) Classic mean-variance portfolio, two fund theorem.
  - i. Comes from quadratic utility or power, normal, continuous trading; no job ( $y$ ), and iid returns ( $E_t(R_{t+1}^e)$  constant).
- (b) Three dimensional extension for multifactor worlds.
  - i. *Hedging demand – buying and selling insurance.* Portfolios are “multifactor efficient” – minimize variance for given mean *and exposure (beta) to additional factors.*
  - ii. The market portfolio may not be mean-variance efficient; the average investor may give up some mean to buy “recession insurance.”
  - iii. Source: job or varying  $E_t(R_{t+1})$  (“investment opportunity set.”)
- (c) Who buys/sells options (or synthetic options like hedge funds)? You can buy/sell any part of the return distribution with options!
- (d) Market timing based on forecasts? It’s easy to calculate from the forecasting regression. But it’s sensitive to uncertainty about just how good the regression is.
- (e) Warning: all portfolio problems are sensitive to expected return assumptions, can give Wacky Weights. “Bayesian” techniques to shade advice back towards holding the market help a bit.

3. Note the traditional CAPM / mean-variance framework survives *very much generalized*.

- (a) Market index  $\rightarrow$  multiple factors (hedges for outside income/risks, value, small, momentum, options, bond long/short, international, etc. etc.)
- (b) Two funds  $\rightarrow$  many many funds
- (c) Index based on your risk aversion vs. market  $\rightarrow$  many dimensions of you vs. market

4. Takeaway deep thoughts

- (a) *The average investor holds the market.* Understand how you are *different* to do otherwise
- (b) Understand the economic function of a class of investment

- (c) Bond investor story – a long-horizon, cashflow-based perspective
- (d) Hedge outside risks, non-priced factors! Match your liability risk profile. This is very important, very overlooked, free from an asset pricing perspective, and people should be willing to pay a lot to do it better.
- (e) This is hard, needs quantification? Yes! A reason for us to exist without winning the zero-sum alpha chase.



## 32.12 Options in portfolios – a deeper look

(I will *not* cover this material due to time constraints. I include it in case you feel like reading it later. It both introduces a powerful technique for doing portfolio theory and it gets you thinking more deeply about the question, who should buy and who should write put options? )

We saw hedge funds and Warren Buffet are writing put options. Lots of investors buy them to protect losses. Does this make sense? How do we add options to our portfolio thoughts?

1. A new ingredient: utility is not quadratic—marginal utility is not linear. Some people really worry about losses – they will buy put options. Others will worry less – they will write put options.
2. A second new modeling ingredient: *complete markets/contingent claims* approach to portfolio calculations. Index all possible outcomes by “states”  $i$ . The investor can buy securities that pay 1 in state  $i$  and zero elsewhere. (This is a second and important technique for solving portfolio problems and makes the math much much easier. Notice I didn’t do the math above, I just quoted the results. Doing that math takes continuous-time dynamic programming.)
3. The investor eats  $W_{t+1,i}$  in each state  $i$ , and thus must buy  $W_{t+1,i}$  units of the security that pays 1 in state  $i$ , at cost  $pc_i$  each.

$$\begin{aligned} \max Eu(W_{t+1}) &= \sum_{\text{states } i} \pi_i u(W_{t+1,i}) \\ \text{s.t. } W_t &= \sum_i pc_i W_{t+1,i} \end{aligned}$$

The portfolio problem now looks *exactly* like “buy a basket of apples, bananas and cherries s.t. sum of price times quantities = wealth.”

4. A little more familiar notation: Let’s write

$$m_i \equiv \frac{pc_i}{\pi_i}.$$

Then

$$W_t = \sum_i \pi_i m_i W_{t+1,i} = E(m_{t+1} W_{t+1})$$

$$W_t = \text{time } t \text{ value of the portfolio that pays } W_{t+1}$$

$m_{t+1}$  is a “stochastic discount factor” “transformation to risk-neutral probability”;  $p_i$  are the “contingent claim prices”; “state-price density”; the price of “digital options” that pay 1 unit in state  $i$ .

5. In sum, the problem is to choose the *random variable*  $W$  (choose what it does in each state of nature)

$$\max_{\{W_{t+1}\}} Eu(W_{t+1}) \text{ s.t. } W_t = E(m_{t+1} W_{t+1})$$

Let’s do it. (You introduce a Lagrangian  $\lambda$  to maximize with a constraint – see the Notes)

$$\max_{\{W_{t+1,i}\}} \sum_i \pi_i u(W_{t+1,i}) - \lambda \sum_i \pi_i m_i W_{t+1,i}$$

$$\begin{aligned} \frac{d}{dW_{t+1,i}} & : \pi_i u'(W_{t+1,i}) = \lambda \pi_i m_i \\ u'(W_{t+1}) & = \lambda m_{t+1} \end{aligned}$$

Interpretation:

- (a) Investor faces prices  $m_{t+1}$ . The investor adjusts portfolio holdings until marginal utility equals the market discount factor.
- (b)  $\lambda$  adjusts to get total value right; if you have  $W_t = \$10,000$ , you invest twice as much as someone with  $W_t = \$5,000$ .
- (c) LHS: how hungry are you in various states? RHS: how expensive is it to buy consumption in those states?
- (d) Marginal rate of substitution = price ratio. Eat less of high priced goods, more of low priced goods. “good” = consumption or wealth, indexed by state of nature, i.e. value of stock return.
- (e) Again, note how portfolio and equilibrium problems are the same thing backwards. Before we used this to learn what  $m$  is by watching  $W$  (=c in this case), *after* the consumer has optimized. Now we use this to learn what  $W$  (c) should be *given*  $m$ .

6. The solution to portfolio problem is then

$$W_{t+1} = u'^{-1}(\lambda m_{t+1})$$

$\lambda$  sets the level right to match initial wealth. *This is it* – it describes the optimal portfolio! ( $\lambda$  just adjusts scale to initial wealth.) Examples should make it less mysterious.

7. Example:

$$u'(W) = W^{-\gamma}$$

then

$$W_{t+1} = (\lambda m_{t+1})^{-\frac{1}{\gamma}}.$$

Constraint:

$$\begin{aligned} W_t & = E(m_{t+1} W_{t+1}) = E \left[ m_{t+1} (\lambda m_{t+1})^{-\frac{1}{\gamma}} \right] \\ \lambda^{-\frac{1}{\gamma}} & = \frac{W_t}{E(m_{t+1}^{1-\frac{1}{\gamma}})} \end{aligned}$$

complete solution

$$W_{t+1} = W_t \frac{m_{t+1}^{-\frac{1}{\gamma}}}{E(m_{t+1}^{1-\frac{1}{\gamma}})}.$$

8. Apparent mystery: Where does  $m$  come from? Answer:  $m$  represents the available prices and returns, in place of  $E(R)$ ,  $cov(R)$  above. You can always find an  $m$  to represent the returns at hand.

9. Example 1: Suppose the world is risk neutral. All assets are priced by

$$p = E(x)/R^f;$$

Then, we can use

$$m = \frac{1}{R^f}.$$

i.e.

$$p = E\left(\frac{1}{R^f}x\right) = \frac{E(x)}{R^f}.$$

Equivalently, risk neutral means

$$E(R^e) = -cov(m, R^e) = 0 \rightarrow E(R) = R^f$$

so we know  $m$  must be a constant, and

$$R^f = 1/E(m)$$

so again  $m = 1/R^f$ .

10. Now, suppose the investor has power utility and  $m = 1/R^f$ .

$$u'(W) = W^{-\gamma}$$

Then

$$\begin{aligned} u'(W_{t+1}) &= \lambda m_{t+1} \\ W_{t+1}^{-\gamma} &= \lambda m \\ W_{t+1} &= \text{constant} \end{aligned}$$

- (a) Constant means “the same for all states of nature” or “the same no matter what happens.” This is about how to adjust wealth *across states of nature* tomorrow; how much “randomness” to accept.
  - (b) How do you get  $W_{t+1} = \text{constant}$ ? Invest all in  $R^f$ .
  - (c) Intuition: You’re risk averse. Everyone else is risk neutral. You buy bonds.
  - (d) Intuition: you *want* the same amount of each “good”. The price of all goods is the same. Buy the same amount of each good.
  - (e) Preview: A real problem must have a higher price  $m$  for consumption when the market is down, and a lower price for consumption when the market is up. This will induce you to eat less when the market is down, i.e. hold a portfolio long the market.
11. Example 2: Binomial, with equal 1/2 probability of good and bad outcome. (Basic option pricing setup)

$$S \rightarrow \{u S, d S\}, R^f = 1$$

(i.e.  $u = 1.3, d = 0.9, 100 \rightarrow \{130, 90\}$ .) Our first task is to represent these prices with a  $m$  and  $p = E(mx)$  for every asset.

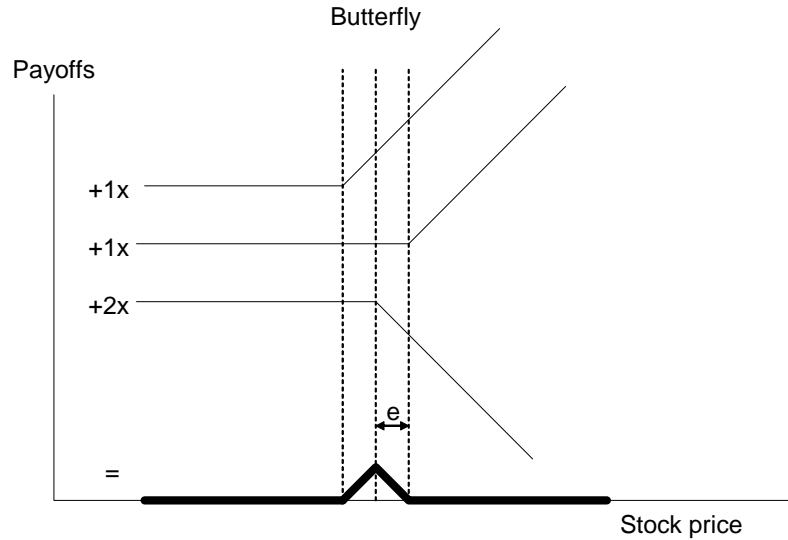
$$\begin{aligned} S &: S = \frac{1}{2}m_u(u S) + \frac{1}{2}m_d(d S) \\ R^f &: 1 = \frac{1}{2}m_u 1 + \frac{1}{2}m_d 1 \end{aligned}$$

Two equations in two unknowns gives<sup>30</sup>

$$m_u = 2 \frac{1-d}{u-d}; \quad m_d = 2 \frac{u-1}{u-d}$$

12. Example 3: (serious). The investor can buy a full set of options, and bond, and dynamically trade stock with positive risk premium.

- (a) This looks hard!
- (b) What's  $m$ ? Where do we read  $m$  prices in the newspaper? Fact: *You can recover the discount factor / contingent claim prices from the second derivative of the call option price with respect to strike price.*
- (c) Why? States  $i$  are indexed by the value of the stock at date  $T$ . Construct a payoff that is one if the stock has value  $X$ , zero elsewhere? A butterfly:



Total payoff of butterfly (area of triangle) =  $\varepsilon^2$ .

Buy  $1/\varepsilon^2$  butterflies to get a payoff of 1 if  $S_T = X$ .

Price of 1 butterfly =  $-[C(X - \varepsilon) - 2C(X) + C(X + \varepsilon)] = \varepsilon^2 \frac{\partial^2 C(X)}{\partial X^2}$

Price = value of  $1/\varepsilon^2$  butterflies =  $\frac{\partial^2 C(X)}{\partial X^2}$

$m$  = price/probability (lognormal stock distribution, for example)

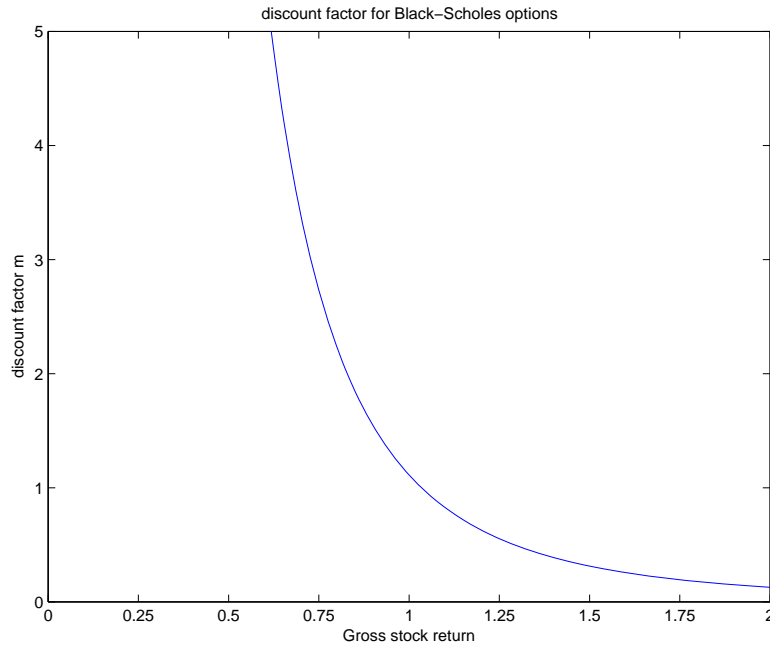
Point: yes, options really are the same as contingent claims, and here's how to read contingent claims prices from options prices!

30

$$\begin{aligned} \begin{bmatrix} u & S & d & S \\ 1 & & 1 & \end{bmatrix} \begin{bmatrix} m_u \\ m_d \end{bmatrix} &= \begin{bmatrix} 2S \\ 2 \end{bmatrix} \\ \begin{bmatrix} m_u \\ m_d \end{bmatrix} &= \frac{1}{u-d} \frac{1}{S} \begin{bmatrix} 1 & -d & S \\ -1 & u & S \end{bmatrix} \begin{bmatrix} 2S \\ 2 \end{bmatrix} \\ &= \frac{2}{u-d} \begin{bmatrix} 1-d \\ u-1 \end{bmatrix} \end{aligned}$$

- (d) (Note: From Put-Call parity  $C - P = S - X/R^f$ ,  $\partial^2 C/\partial X^2 = \partial^2 P/\partial X^2$  so you can use puts or calls)
- (e) Let's suppose option prices are given by Black-Scholes. *Fact:* In the Black-Scholes world with a lognormal stock return, with instantaneous mean  $\mu$  and volatility  $\sigma$ , and interest rate  $r$ , the discount factor  $m$  is

$$m_{t,t+T} = (\text{const}) (R_{0 \rightarrow T})^{-\frac{\mu-r}{\sigma^2}}$$



- (f) *Why?* Take the second derivative of the B-S formula. Read the long footnote.<sup>31</sup>
- (g) *Intuition.* It's easy to cheaply synthesize a payoff to good states – invest one dollar, if the market goes up you can pay a lot. Risk aversion is implicit in market prices: it's much more expensive to buy wealth in the state that stocks go down than in the state that stocks go up, because everyone wants more in the down state. This induces you to economize on down states, to accept consumption and wealth that are lower in the

<sup>31</sup>In the Black Scholes setup, the stock price follows

$$R_{0 \rightarrow T} = \frac{S_T}{S_0} = e^{\left(\mu - \frac{\sigma^2}{2}\right)T + \sigma\sqrt{T}\varepsilon}; \varepsilon \sim N(0, 1)$$

$$\frac{S_T}{S_0} = e^{\left(\mu - \frac{\sigma^2}{2}\right)T + \sigma\sqrt{T}\varepsilon}; \varepsilon \sim N(0, 1)$$

(i.e.  $\frac{dS}{S} = \mu dt + \sigma dz$ ). The discount factor follows

$$m_{t,t+T} = e^{-\left[r + \frac{1}{2}\left(\frac{\mu-r}{\sigma}\right)^2\right]T - \frac{\mu-r}{\sigma}\sqrt{T}\varepsilon}.$$

I get the formula in the text by substituting out  $e^{\sigma\sqrt{T}\varepsilon}$  in the  $m$  equation from  $S_T/S_0$ .

To show that this formula for  $m$  is right, you can take the second derivative of the Black Scholes formula. Get out a big pad of paper first. More easily, you can verify that  $m$  does what a discount factor should do,  $S_0 = E(m_{t,t+T}S_T)$ ,

expensive states. This plot of  $m$  will lead you to a portfolio with less payoff in bad states. *You're looking at the a la carte menu for consumption in each state of the world, indexed by the stock return.*

13. Now suppose the investor has power utility

$$u'(W_{t+T}) = W_{t+T}^{-\gamma}$$

Solution:

$$\begin{aligned} W_{t+1}/W_t &= (\text{const}) m_{t+1}^{-\frac{1}{\gamma}} \\ W_{t+1}/W_t &= (\text{const}) (R_{0 \rightarrow T})^{\frac{\mu-r}{\sigma^2} \frac{1}{\gamma}} \end{aligned}$$

*The investor wants a portfolio that is a nonlinear function of the stock return.*

14. The plots uses  $\mu - r = 8\%$  and  $\sigma = 16\%$  for a few values of risk aversion  $\gamma$ . For  $\gamma = \frac{0.09-0.01}{0.16^2} = 3.125$ , the function is linear – the investor just puts all his wealth in the stock. At lower  $\gamma$ , the investor exploits the strong risk-return tradeoff, taking a position that is much more sensitive to the stock return. At higher  $\gamma$ , he/she accepts lower payoffs in the

---

$1 = E(m_{t,t+T}R^f)$  as follows:

$$\begin{aligned} E(m_{t,t+T}S_T) &= E \left[ e^{-\left(r + \frac{1}{2} \left(\frac{\mu-r}{\sigma}\right)^2\right)T - \frac{\mu-r}{\sigma} \sqrt{T}\varepsilon} S_0 e^{\left(\mu - \frac{\sigma^2}{2}\right)T + \sigma \sqrt{T}\varepsilon} \right] \\ &= S_0 E \left[ e^{\left[(\mu-r) - \frac{1}{2} \left(\frac{\mu-r}{\sigma}\right)^2 - \frac{\sigma^2}{2}\right]T + \left(\sigma - \frac{\mu-r}{\sigma}\right)\sqrt{T}\varepsilon} \right] \\ &= S_0 \left[ e^{\left[(\mu-r) - \frac{1}{2} \left(\frac{\mu-r}{\sigma}\right)^2 - \frac{\sigma^2}{2}\right]T + \frac{1}{2} \left(\sigma - \frac{\mu-r}{\sigma}\right)^2 T} \right] \\ &= S_0 \left[ e^{\left[(\mu-r) - \frac{1}{2} \left(\frac{\mu-r}{\sigma}\right)^2 - \frac{\sigma^2}{2}\right]T + \frac{1}{2} \left(\sigma^2 - 2(\mu-r) + \left(\frac{\mu-r}{\sigma}\right)^2\right)T} \right] \\ &= S_0 \left[ e^{\left[(\mu-r) - \frac{1}{2} \left(\frac{\mu-r}{\sigma}\right)^2 - \frac{\sigma^2}{2} + \frac{\sigma^2}{2} - (\mu-r) + \frac{1}{2} \left(\frac{\mu-r}{\sigma}\right)^2\right]T} \right] \\ &= S_0 [e^0] = S_0 \end{aligned}$$

$$\begin{aligned} E(m_{t,t+T}e^{rT}) &= E \left( e^{-\left(r + \frac{1}{2} \left(\frac{\mu-r}{\sigma}\right)^2\right)T - \frac{\mu-r}{\sigma} \sqrt{T}\varepsilon} e^{rT} \right) \\ &= e^{-\left(r + \frac{1}{2} \left(\frac{\mu-r}{\sigma}\right)^2\right)T - \frac{1}{2} \left(\frac{\mu-r}{\sigma}\right)^2 T} e^{rT} \\ &= e^{-rT} e^{rT} = 1 \end{aligned}$$

That's nice, but how did I know that was the right  $m$ ? I constructed  $m$  to price the stock and bond instantaneously, i.e.

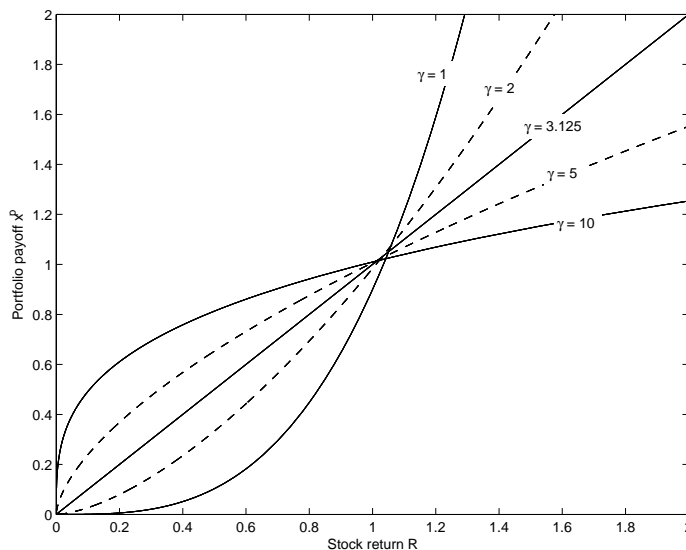
$$\frac{dm_t}{m_t} = -r dt - \frac{\mu - r}{\sigma} dz$$

satisfies

$$\begin{aligned} E \left( \frac{dm}{m} \right) &= -r dt \\ E \left( \frac{dS}{S} \right) - r dt &= (\mu - r) dt = -E \left( \frac{dm}{m} \frac{dS}{S} \right) = - \left( -\frac{\mu - r}{\sigma} \sigma \right) dt \end{aligned}$$

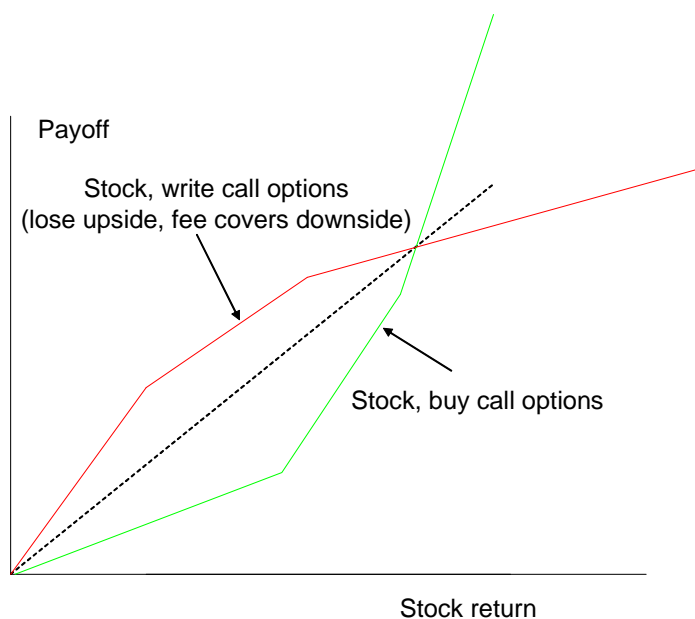
Then, the formula for  $m$  is the solution to this stochastic differential equation. Obviously, you have to know some continuous time to have any idea what this means.

good states (on the right) in order to get a better payoff in the more expensive (high  $m$ ) bad states on the left.



15. This is the *final payoff* – it describes how wealth  $W_{t+1}$  varies as a function of return or stock prices. How do you get it? – What is the actual portfolio?

- (a) The graph describes directly the number of butterfly options (hard)
- (b) You can approximate it with a few put, call options.



(c) Dynamic trading, constant rebalancing

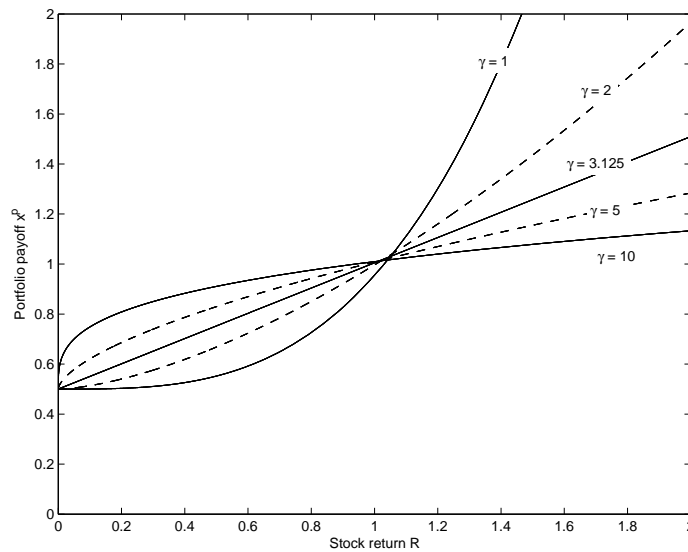
16. It's easy to think of other utility functions and other portfolios now that we have  $m$ . Many investors seem to have a special aversion to loss, “don't lose more than x dollars” One way to

model this is a minimum “subsistence level of consumption.” (“No matter what, we absolutely must be able to keep the jet. I’d rather die than fly commercial.”)

$$u'(c_{t+1}) = (c_{t+1} - h)^{-\gamma}$$

We can solve this the same way,

$$\begin{aligned} (c_{t+1} - h)^{-\gamma} &= \lambda m_{t+1} \\ c_{t+1} &= W_{t+1} = (\lambda m_{t+1})^{-\frac{1}{\gamma}} + h \end{aligned}$$

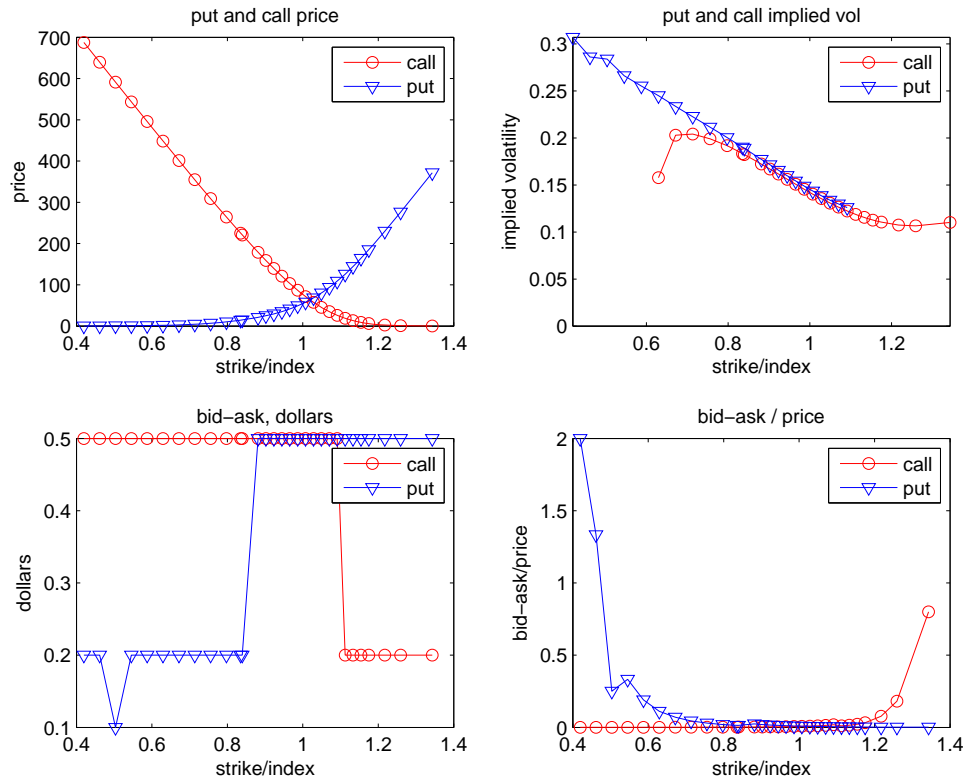


$\gamma = 3.125$  mixes a stock and bond.  $\gamma = 1$  looks a lot like a bond plus a call option or stock + protective puts.  $\gamma = 10$  looks like a bond and writing a call option.  $\gamma = 10$  sells off gains in good times in order to get more protection in bad times.

17. How would you do this with real data? I did!

- (a) Step 1: Get options price data (mine are June 30 2005, for 1 year to exp S&P index puts and calls)

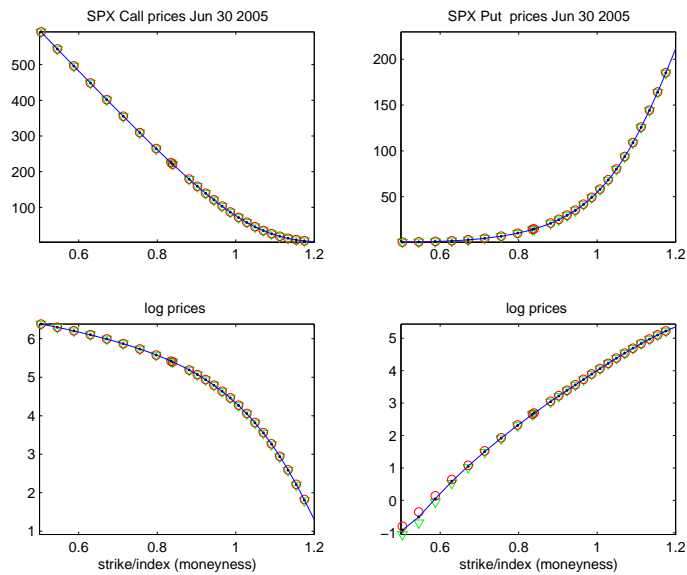




(b) Step 2: Fit a smooth function to options data so you can take second derivatives. I fit

$$p_i = e^{a+bx_i+cx_i^2+dx_i^3+\dots x_i^{12}}$$

where  $p$  = price, and  $x$  =strike price. (Some people do “nonparametric” functions)

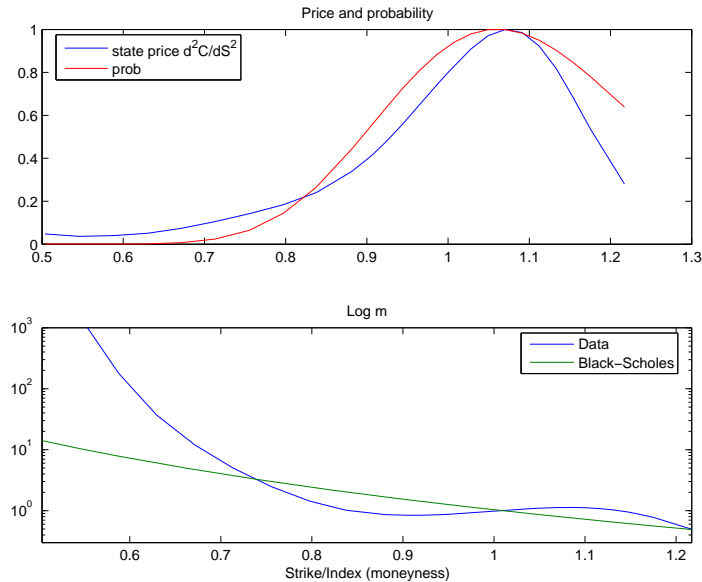


(c) Step 3: Take second derivatives to find state-prices

$$p = e^{f(x)}$$

$$\frac{d^2 p}{dx^2} = [f''(x) + f'(x)^2] p$$

(d) Step 4: Divide by probability to find  $m$ . I used lognormal probabilities,  $\mu = 0.09$ ,  $\sigma = 0.145$  (implied vol. at the money). (I use in the money puts and out of the money calls to make a single line)



Since price is greater than probability, the state-price density goes through the roof for out of the money options. (Log scale!)

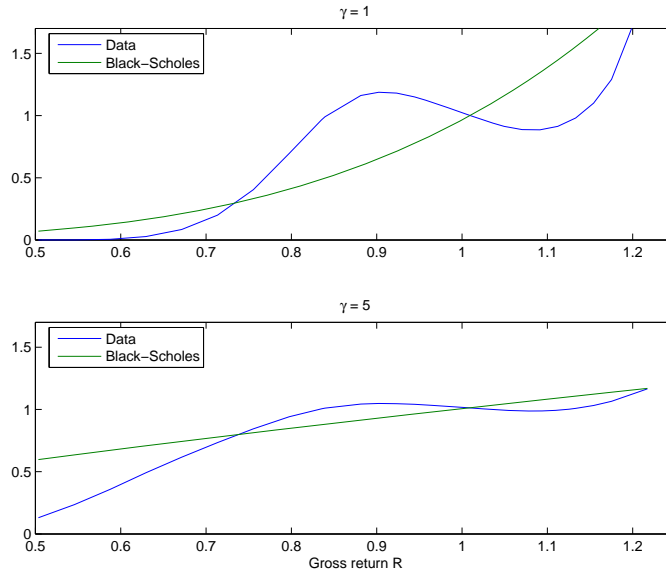
- (e) Again, this is the a la carte consumption menu. Consumption in down markets is more expensive.
- (f)  $u'(W) = \lambda m$ . You're *looking* at the average investor's (marginal) utility, one integral away from *looking* at the average investor's utility function. It has a big cliff for losses!
- (g) Compute optimal portfolios. (For an investor who is not typical, i.e. values losses by power formulas)

$$W = \lambda m^{-\frac{1}{\gamma}};$$

$$\lambda : E(mW) = W_0$$

$$E(m\lambda m^{-\frac{1}{\gamma}}) = W_0$$

$$W = W_0 \frac{m^{-\frac{1}{\gamma}}}{E\left(m^{-\frac{1}{\gamma}}\right)}$$



Notice in real data, the cost of out of the money puts is higher than BS, so our guy stays away from them more than under BS prices. If he holds stock, then “stay away from” = write lots of index puts. *Here is our hedge fund investor.* He responds to the cheap price relative to probability of at the money options by buying more of them. As usual, less risk averse investors respond more to all signals.

18. Big picture: *Think about every part of the return distribution.* Should you sell or buy wealth in each state? How valuable is it to you? How expensive is it? (Note: this can describe dynamic strategies as well as explicit options of course.) Connect to hedge funds – some people *should* be “writing put options”
19. Big picture 2: If hedge fund investors are less averse to big losses and other investors are more averse to big losses, it makes perfect sense for the HF investors to write put options and the others to buy (even if expensive) “portfolio insurance”.

### 32.13 Bayesian Portfolio Derivation

$$\begin{aligned}
f(R|\mu) &= \frac{1}{\sqrt{2\pi \|\Sigma\|}} e^{-\frac{1}{2}(R-\mu)' \Sigma^{-1} (R-\mu)} \\
f(\mu) &= \frac{1}{\sqrt{2\pi \|\Sigma_\mu\|}} e^{-\frac{1}{2}(\mu-\bar{\mu})' \Sigma_\mu^{-1} (\mu-\bar{\mu})} \\
f(R) &= \int f(R|\mu) f(\mu) d\mu \\
f(R) &= \frac{1}{\sqrt{2\pi \|\Sigma\|}} \frac{1}{\sqrt{2\pi \|\Sigma_\mu\|}} \int e^{-\frac{1}{2}(R-\mu)' \Sigma^{-1} (R-\mu)} e^{-\frac{1}{2}(\mu-\bar{\mu})' \Sigma_\mu^{-1} (\mu-\bar{\mu})} d\mu \\
&= \sqrt{\frac{\left(\frac{1}{\sigma_\mu^2} + \frac{1}{\sigma^2}\right)^{-1}}{2\pi\sigma^2\sigma_\mu^2}} e^{-\frac{1}{2} \left[ \frac{R^2 + \bar{\mu}^2}{\sigma^2 + \sigma_\mu^2} - \frac{\sigma_\mu^2}{(\sigma^2 + \sigma_\mu^2)} \frac{R^2}{\sigma^2} - \frac{\sigma^2}{(\sigma^2 + \sigma_\mu^2)} \frac{\bar{\mu}^2}{\sigma_\mu^2} - \frac{2R\bar{\mu}}{\sigma^2 + \sigma_\mu^2} \right]} \\
&= \frac{1}{\sqrt{2\pi (\sigma^2 + \sigma_\mu^2)}} e^{-\frac{1}{2} \frac{(R-\bar{\mu})^2}{\sigma^2 + \sigma_\mu^2}} \\
&= \frac{1}{\sqrt{2\pi (\sigma^2 + \sigma_\mu^2)}} e^{-\frac{1}{2} \frac{(R-\bar{\mu})^2}{\sigma^2 \sigma_\mu^2 \left(\frac{1}{\sigma^2} + \frac{1}{\sigma_\mu^2}\right)}}
\end{aligned}$$

$$\begin{aligned}
f(R) &= \int f(R|\theta) f(\theta) = \frac{1}{\sqrt{2\pi \|\Sigma\|} \sqrt{2\pi \|\Sigma_\mu\|}} \int e^{-\frac{1}{2}(R-\mu)' \Sigma^{-1} (R-\mu)} e^{-\frac{1}{2}(\mu-\bar{\mu})' \Sigma_\mu^{-1} (\mu-\bar{\mu})} d\mu \\
&= \frac{1}{\sqrt{2\pi \|\Sigma\|} \sqrt{2\pi \|\Sigma_\mu\|}} \int e^{-\frac{1}{2} [R' \Sigma^{-1} R + \mu' \Sigma^{-1} \mu - 2R' \Sigma^{-1} \mu + \mu' \Sigma_\mu^{-1} \mu + \bar{\mu}' \Sigma_\mu^{-1} \bar{\mu} - 2\bar{\mu}' \Sigma_\mu^{-1} \mu]} d\mu
\end{aligned}$$

Working on the term in brackets

$$\begin{aligned}
&\mu' (\Sigma^{-1} + \Sigma_\mu^{-1}) \mu - 2\mu' (\Sigma^{-1} R + \Sigma_\mu^{-1} \bar{\mu}) + (R' \Sigma^{-1} R + \bar{\mu}' \Sigma_\mu^{-1} \bar{\mu}) \\
(\mu - x)' \Omega (\mu - x) &= \mu' \Omega \mu + x' \Omega x - 2\mu' \Omega x; \quad x = (\Sigma^{-1} + \Sigma_\mu^{-1})^{-1} (\Sigma^{-1} R + \Sigma_\mu^{-1} \bar{\mu}) \\
&\mu' (\Sigma^{-1} + \Sigma_\mu^{-1}) \mu - 2\mu' (\Sigma^{-1} + \Sigma_\mu^{-1}) (\Sigma^{-1} + \Sigma_\mu^{-1})^{-1} (\Sigma^{-1} R + \Sigma_\mu^{-1} \bar{\mu}) + (R' \Sigma^{-1} R + \bar{\mu}' \Sigma_\mu^{-1} \bar{\mu}) \\
&\left[ \mu - (\Sigma^{-1} + \Sigma_\mu^{-1})^{-1} (\Sigma^{-1} R + \Sigma_\mu^{-1} \bar{\mu}) \right]' \left[ (\Sigma^{-1} + \Sigma_\mu^{-1})^{-1} \right]^{-1} \left[ \mu - (\Sigma^{-1} + \Sigma_\mu^{-1})^{-1} (\Sigma^{-1} R + \Sigma_\mu^{-1} \bar{\mu}) \right] \dots \\
&- (\Sigma^{-1} R + \Sigma_\mu^{-1} \bar{\mu})' (\Sigma^{-1} + \Sigma_\mu^{-1})^{-1} (\Sigma^{-1} R + \Sigma_\mu^{-1} \bar{\mu}) + (R' \Sigma^{-1} R + \bar{\mu}' \Sigma_\mu^{-1} \bar{\mu})
\end{aligned}$$

As in the scalar case, the first term integrates out to zero, leaving only the second term

$$\sqrt{\frac{\left\| \left[ (\Sigma^{-1} + \Sigma_\mu^{-1})^{-1} \right] \right\|}{2\pi \|\Sigma\| \|\Sigma_\mu\|}} e^{-\frac{1}{2} \left[ R' \Sigma^{-1} R + \bar{\mu}' \Sigma_\mu^{-1} \bar{\mu} - (\Sigma^{-1} R + \Sigma_\mu^{-1} \bar{\mu})' (\Sigma^{-1} + \Sigma_\mu^{-1})^{-1} (\Sigma^{-1} R + \Sigma_\mu^{-1} \bar{\mu}) \right]}$$

Again, work on the term in brackets:

$$\begin{aligned}
& R'\Sigma^{-1}R + \bar{\mu}'\Sigma_{\mu}^{-1}\bar{\mu} - (\Sigma^{-1}R + \Sigma_{\mu}^{-1}\bar{\mu})' (\Sigma^{-1} + \Sigma_{\mu}^{-1})^{-1} (\Sigma^{-1}R + \Sigma_{\mu}^{-1}\bar{\mu}) \\
& R'\Sigma^{-1}R + \bar{\mu}'\Sigma_{\mu}^{-1}\bar{\mu} - R'\Sigma^{-1} (\Sigma^{-1} + \Sigma_{\mu}^{-1})^{-1} \Sigma^{-1}R - \bar{\mu}'\Sigma_{\mu}^{-1} (\Sigma^{-1} + \Sigma_{\mu}^{-1})^{-1} \Sigma_{\mu}^{-1}\bar{\mu} - 2R'\Sigma^{-1} (\Sigma^{-1} + \Sigma_{\mu}^{-1})^{-1} \Sigma_{\mu}^{-1}\bar{\mu} \\
& R' \left[ \Sigma^{-1} - \Sigma^{-1} (\Sigma^{-1} + \Sigma_{\mu}^{-1})^{-1} \Sigma^{-1} \right] R + \bar{\mu}' \left[ \Sigma_{\mu}^{-1} - \Sigma_{\mu}^{-1} (\Sigma^{-1} + \Sigma_{\mu}^{-1})^{-1} \Sigma_{\mu}^{-1} \right] \bar{\mu} - 2R'\Sigma^{-1} (\Sigma^{-1} + \Sigma_{\mu}^{-1})^{-1} \Sigma_{\mu}^{-1}\bar{\mu} \\
& R'\Sigma^{-1} \left[ \Sigma - (\Sigma^{-1} + \Sigma_{\mu}^{-1})^{-1} \right] \Sigma^{-1}R + \bar{\mu}'\Sigma_{\mu}^{-1} \left[ \Sigma_{\mu} - (\Sigma^{-1} + \Sigma_{\mu}^{-1})^{-1} \right] \Sigma_{\mu}^{-1}\bar{\mu} - 2R'\Sigma^{-1} (\Sigma^{-1} + \Sigma_{\mu}^{-1})^{-1} \Sigma_{\mu}^{-1}\bar{\mu}
\end{aligned}$$

Work on the first two terms separately,

$$\begin{aligned}
& R'\Sigma^{-1} \left[ (\Sigma^{-1} + \Sigma_{\mu}^{-1})^{-1} (\Sigma^{-1} + \Sigma_{\mu}^{-1}) \Sigma - (\Sigma^{-1} + \Sigma_{\mu}^{-1})^{-1} I \right] \Sigma^{-1}R + \dots \\
& R'\Sigma^{-1} (\Sigma^{-1} + \Sigma_{\mu}^{-1})^{-1} \Sigma_{\mu}^{-1}R + \dots \\
& \dots + \bar{\mu}'\Sigma_{\mu}^{-1} \left[ \Sigma_{\mu} (\Sigma^{-1} + \Sigma_{\mu}^{-1}) (\Sigma^{-1} + \Sigma_{\mu}^{-1})^{-1} - (\Sigma^{-1} + \Sigma_{\mu}^{-1})^{-1} \right] \Sigma_{\mu}^{-1}\bar{\mu} + \dots \\
& \dots + \bar{\mu}'\Sigma^{-1} (\Sigma^{-1} + \Sigma_{\mu}^{-1})^{-1} \Sigma_{\mu}^{-1}\bar{\mu} + \dots
\end{aligned}$$

so we have

$$\begin{aligned}
& R'\Sigma^{-1} (\Sigma^{-1} + \Sigma_{\mu}^{-1})^{-1} \Sigma_{\mu}^{-1}R + \bar{\mu}'\Sigma^{-1} (\Sigma^{-1} + \Sigma_{\mu}^{-1})^{-1} \Sigma_{\mu}^{-1}\bar{\mu} - 2R'\Sigma^{-1} (\Sigma^{-1} + \Sigma_{\mu}^{-1})^{-1} \Sigma_{\mu}^{-1}\bar{\mu} \\
& (R - \bar{\mu})' \Sigma^{-1} (\Sigma^{-1} + \Sigma_{\mu}^{-1})^{-1} \Sigma_{\mu}^{-1} (R - \bar{\mu}) \\
& (R - \bar{\mu})' \left[ \Sigma (\Sigma^{-1} + \Sigma_{\mu}^{-1}) \Sigma_{\mu} \right]^{-1} (R - \bar{\mu}) \\
& (R - \bar{\mu})' (\Sigma_{\mu} + \Sigma)^{-1} (R - \bar{\mu})
\end{aligned}$$

$$\frac{1}{\sqrt{2\pi \|[(\Sigma + \Sigma_{\mu})]\|}} e^{-\frac{1}{2}(R-\bar{\mu})'(\Sigma_{\mu}+\Sigma)^{-1}(R-\bar{\mu})}$$