4. ACTIVE PORTFOLIO MANAGEMENT

Reading in Chincarini and Kim (C&K):

Chap. 1 (optional)
Chap. 2: Focus mainly on Sections 2.5 and 2.6 but do read the other sections
Chap. 9: Sections 9.8.1 and 9.8.4 (we will return to 9.8.2 and 9.8.3 later)

References

Bertrand, P. (2010). "Another Look at Portfolio Optimization under Tracking Error Constraints", *Financial Analysts Journal*, 66, 3.

Grinold, R. C. and Kahn, R. N. (2000). *Active Portfolio Management*, 2nd edition, McGraw-Hill. (the "bible" for years, but difficult to read)

Jorion, P. (2003). "Portfolio Optimization with Tracking-Error Constraints", *Financial Analysts Journal*, Sept./Oct., 70-82.

Lee, W. (2000). *Theory and Methodology of Tactical Asset Allocation*, Frank J. Fabozzi Associates.

Qian, E.E., Hua, R. H. and Sorensen, E.H. (2007). *Quantitative Equity Portfolio Management*, Chapman and Hall.

Roll, R. (1992). "A Mean-Variance Analysis of Tracking Error", *The Journal of Portfolio Management*, Summer 1992, 13-22.

Scherer, B. (2007). *Portfolio Construction and Risk Budgeting*, 4th edition, Risk Books

4.1 Basic Concepts and Terms

r_B : Benchmark portfolio returns, typically an index

r_P : Active manager's portfolio returns

$$r_A = r_P - r_B$$ **Active Returns**

$$\sigma_A^2 = \mathrm{var}(r_A)$$ **Tracking error variance (TEV)**

$$\sigma_A = \mathrm{var}^{1/2}(r_A)$$ **Tracking error (TE)**

Example

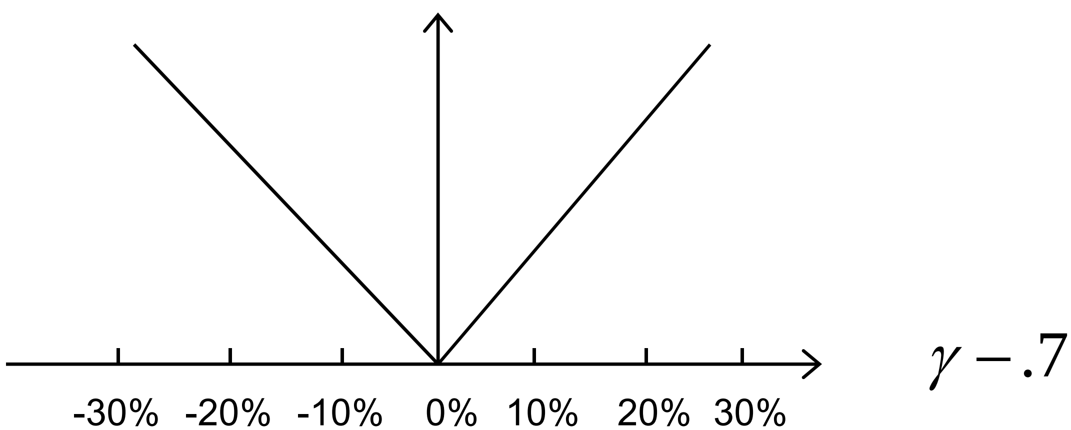
$$r_B = .3 \cdot r_{MSFT} + .7 \cdot r_{ORCL}$$

$$r_P = (1 - \gamma) \cdot r_{MSFT} + \gamma \cdot r_{ORCL}$$

$$r_A = (.7 - \gamma) \cdot r_{MSFT} + (\gamma - .7) \cdot r_{ORCL}$$
$$= \underbrace{(\gamma - .7)} \cdot (r_{ORCL} - r_{MSFT})$$

Active holding → $(\gamma - .7)$

$$\sigma_A = |\gamma - .7| \cdot \text{var}^{1/2}(r_{ORCL} - r_{MSFT})$$



Active Portfolios

Benchmark weights: $\mathbf{w}_B$

Portfolio weights (active manager): $\mathbf{w}_P$

Active weights: $\mathbf{w}_A = \mathbf{w}_P - \mathbf{w}_B$

Self-financing: $\mathbf{1}'\mathbf{w}_A = 0$

Active returns:

$$
\begin{aligned}
r_A &= r_P - r_B \\
&= \mathbf{w}_P' \cdot \mathbf{r} - \mathbf{w}_B' \cdot \mathbf{r} \\
&= \mathbf{w}_A' \cdot \mathbf{r}
\end{aligned}
$$

Active Portfolios

Expected active returns:

$$\mu_A = E(r_A) = E(\mathbf{w}'_A \cdot \mathbf{r}) = \mathbf{w}'_A \cdot \boldsymbol{\mu}$$

Tracking error variance (TEV):

$$\sigma_A^2 = \text{var}(r_A) = \text{var}(\mathbf{w}'_A \cdot \mathbf{r}) = \mathbf{w}'_A \boldsymbol{\Sigma} \mathbf{w}_A$$

Alternatively:

$$\sigma_A^2 = \text{var}(r_P - r_B) = \sigma_P^2 + \sigma_B^2 - 2\,\text{cov}(r_P, r_B)$$

Alphas and Residual Returns

To better understand stock alphas, assume that for a benchmark B (that could be a market proxy) assume that portfolio returns are adequately modeled by a linear single factor model (SFM) without intercept:

Alternative representations :

$$r_P = \beta_P r_B + \varepsilon \qquad \text{where} \qquad E(\varepsilon) = \alpha_P$$

$$r_P = \alpha_P + \beta_P r_B + \varepsilon \qquad \text{where} \qquad E(\varepsilon) = 0$$

$$\beta_P = \frac{\mathrm{cov}(r_P, r_B)}{\sigma_B^2}$$

Either way: $\mu_P = E\left(r_P\right) = \alpha_P + \beta_P \mu_B$

The active portfolio manager wants to maximize: $\alpha_P > 0$

Residuals are Uncorrelated with BM Returns

It turns out that the residuals are uncorrelated with the benchmark returns:

$$\mathrm{cov}(\varepsilon, r_B) = 0$$

This important property follows from the fact that $\beta_P = \dfrac{\mathrm{cov}(r_P, r_B)}{\sigma_B^2}.$

Proof: Easy exercise.

NOTE: There is an important connection between "betas" of the above form and minimum mean-squared-error (MSE) prediction of one random variable with another random variable. For details see appendix 4A.1.

Active Beta

The **active beta** is computed as the beta of the active portfolio with respect to the benchmark:

$$\beta_A = \frac{\mathrm{cov}(r_A, r_B)}{\sigma_B^2}$$

$$= \frac{\mathrm{cov}(r_P - \mathrm{r}_B, r_B)}{\sigma_B^2},$$

$$= \frac{\mathrm{cov}(r_P, \mathrm{r}_B)}{\sigma_B^2} - \frac{\mathrm{cov}(\mathrm{r}_B, r_B)}{\sigma_B^2}$$

$$= \beta_P - 1.$$

Active Returns with Single Factor Model

$$
\begin{aligned}
r_A &= r_P - r_B \\
&= \alpha_P + \beta_P \cdot r_B + \varepsilon_P - r_B, \qquad E(\varepsilon_P) = 0 \\
&= \alpha_P + \beta_A \cdot r_B + \varepsilon_P, \qquad \beta_A = \beta_P - 1
\end{aligned}
$$

With the SFM the mean and variance of the active returns are:

$$\mu_A = \alpha_P + \beta_A \cdot \mu_B \qquad \mu_A = \alpha_P \text{ only when } \beta_A = 0$$

$$\sigma_A^2 = \beta_A^2 \cdot \sigma_B^2 + \sigma_{\varepsilon_P}^2$$

Active risk (TE): σ_A

Residual risk: σ_{ε_P}

These are the same only when $\beta_A = 0$, Equivalently $\beta_P = 1$, and in that case they are the same as the TE σ_A.

Active Managers Performance

- Performance of portfolio manager is monitored with TEV (low = good)
- TEV is minimized w.r.t. to β_A by the choice $\beta_A = 0 \quad \Rightarrow \quad \beta_P = 1$
- Active managers like $\beta_P \approx 1$ but often choose $\beta_P > 1$ in order to get larger returns
- Managers will be ranked by their **information ratios**, to be defined next

The Information Ratio

The most important performance measure of an active (benchmark tracking) portfolio manager is the **Information Ratio**:

$$IR = \frac{\alpha_P}{\sigma_A} = \frac{E(r_A)}{\sigma_A}$$

Goal of a portfolio manager is to maximize the information ratio.

Example (Grinold, Kahn, p.113):

A new manager: $$\alpha_P = 3.5\%,\ \sigma_A = 5.5\% \Rightarrow IR = \frac{\alpha_P}{\sigma_A} = 0.64\%.$$

How good is her IR?

Compare to others:

% Tile	90	75	50	25	10
IR	1.0	0.5	0	-0.5	-1.0

hire ⟶ fire

Appraisal Ratio versus IR

The term **Appraisal Ratio** is used to define a quantity very similar to the IR, namely:

$$A_P = \frac{E(r_A)}{\sigma_{\varepsilon_P}}$$

Where $\sigma_{\varepsilon_P}^2 = \text{var}(\varepsilon_P)$ is the variance of the SFM residuals, and

$$\sigma_A^2 = \beta_A^2 \cdot \sigma_B^2 + \sigma_{\varepsilon_P}^2$$

So we have $A_P = IR_P$ if and only if the active beta is zero, equivalently the portfolio beta is one.

Since actively managed long-only portfolios often have a beta close to one there often is little difference between the appraisal ratio and IR.

The IR Estimate as Scaled t-Statistic

$$\widehat{IR} = \frac{\bar{r}_A}{\hat{\sigma}_A} = \frac{\frac{1}{T}\sum_{t=1}^{T} r_{A,t}}{\sqrt{\frac{1}{T-1}\sum_{t=1}^{T}(r_{A,t}-\bar{r}_A)^2}}$$

For independent and identically distributed (i.i.d.) $X_1, X_2, \cdots, X_n$ with mean μ, the t-statistic for hypothesis tests and confidence intervals is:

$$\frac{\sqrt{n}\cdot\left(\bar{X}_n - \mu\right)}{\hat{\sigma}_n} \qquad \begin{aligned} \bar{X}_n &= \text{ sample mean} \\ \hat{\sigma}_n &= \text{ sample std. dev.} \end{aligned}$$

So the following quantity closely related to $\widehat{IR}$ is a t-statistic:

$$\frac{\sqrt{T}\cdot\left(\bar{r}_A - \mu\right)}{\hat{\sigma}_A}$$

For $NIID(\mu, \sigma^2)$ random variables the t-statistic has a t-distribution with n-1 degrees of freedom (dof). So with that assumption about returns and with $t_{T-1}(\alpha)$ the α-quantile of a t-distribution with T-1 d.o.f., we have the following.

Level γ t-test that mean of active returns is greater or equal to μ_0:

If $\bar{r}_A > \mu_0 + T^{-1/2}\hat{\sigma}_A \cdot t_{T-1}(1-\gamma)$ reject null hypothesis that $\mu \le \mu_0$, else accept the alternative hypothesis that $\mu > \mu_0$. Note that for $\mu_0 = 0$, the test rejects and decides that IR is positive if:

$$\sqrt{T} \cdot \widehat{IR} > t_{T-1}(1-\gamma)$$

Important point: Test may not be very reliable if the returns have a fat-tailed and/or skewed non-normal distribution, or if there is serial correlation!

4.2 MV Optimal Active Portfolios

Two approaches:

1. Quadratic utility (QU) maximization
   - Will discuss this one next
   - It is useful for not only active portfolio management but also for hedge fund long-short zero investment portfolios
2. Minimize TEV
   - Roll (1992)
   - Jorion (2003)
   - Bertrand (2012)

These are mathematically tedious but they provide guidance on how to think about the performance of active management, and intuition on what to do when minimizing TEV with constraints that require numerical optimization. Details are provided in appendix 4A.2

Maximum QU Active Portfolio

Maximize $\quad \mathbf{w}'_A \boldsymbol{\mu} - \frac{1}{2}\lambda \mathbf{w}'_A \boldsymbol{\Sigma} \mathbf{w}_A \quad$ subject to: $\quad \mathbf{w}'_A \mathbf{1} = 0$

N.B. This formulation applies not only to active portfolios but also to any dollar neutral portfolio, e.g., a hedge fund long dollar neutral portfolio.

Result:

$$\mathbf{w}_{A,opt} = \lambda^{-1} \frac{(\mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1})\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} - (\mathbf{1}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu})\boldsymbol{\Sigma}^{-1}\mathbf{1}}{\mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1}}$$

NOTE 1: The weights are independent of the benchmark weights! This provides a basis for "portable alpha" .

NOTE 2: The above expression is the same as the second term in the two-fund separation theorem of LSS-1 appendix 1A.1.

QHS 2.2.4: Uses $\mathbf{a},\ \mathbf{f}$ where we use $\mathbf{w}_A,\ \boldsymbol{\mu}$.

Warning about notation in QHS

They use $\alpha = \mathbf{w}'_A \mathbf{r}$ which is a random variable not a fixed unknown parameter. And this tendency persists throughout the book, often causing confusion about what are the parameters to be estimated and what are the data

Equivalent Representation

$$\mathbf{w}_{A,opt} = \lambda^{-1}\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - l\mathbf{1}) \qquad l = \frac{\mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{f}}{\mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1}}$$

Note that without the term involving l we have the standard non-active MVO weight vector expression. So what is the role of l aside from insuring that the active weights sum to one? Consider the following single factor model:

$$\boldsymbol{\mu} = \mathbf{1} \cdot \mu + \mathbf{v}, \quad \mathbf{v} \sim (\mathbf{0}, \boldsymbol{\Sigma})$$

The GLS estimate of μ is: $\hat{\mu}_{GLS} = \dfrac{\mathbf{1}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}}{\mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1}} = l$

So the adjustment is to remove the BLUE for a model in which all asset mean returns are the same.

Optimal Active Mean Return, TE and IR

$$\mu_{A,opt} = \boldsymbol{\mu}'\mathbf{w}_{A,opt} = \lambda^{-1}\frac{(\mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1})\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} - (\mathbf{1}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu})^2}{\mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1}}$$

$$\sigma_{A,opt} = \sqrt{\mathbf{w}'_{A,opt}\boldsymbol{\Sigma}\mathbf{w}_{A,opt}} = \lambda^{-1}\sqrt{\frac{(\mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1})\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} - (\mathbf{1}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu})^2}{\mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1}}}$$

Take the ratio of the above to get the maximized information ratio;

$$IR_{\max} = \frac{\mu_{A,opt}}{\sigma_{A,opt}} = \sqrt{\frac{(\mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1})(\mathbf{f}'\boldsymbol{\Sigma}^{-1}\mathbf{f}) - (\mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{f})^2}{\mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1}}}$$

NOTE: $\mu_{A,opt} = \alpha_{opt}$ in QHH (somewhat an abuse of term "alpha")

Two Types of Efficient Frontiers

We see that the "efficient frontier" for the active portfolio is a straight line with slope equal to the information ratio, as the efficient frontier for full investment in risky assets is hyperbolic. The following two slides compare the two efficient frontiers with the following simple example of mean returns, volatilities and correlation matrix (see QHH, pp. 33-37):

$$\boldsymbol{\mu}' = (10\%, 0\%, 10\%)$$

$$\boldsymbol{\sigma}' = (30\%, 30\%, 30\%)$$

$$\mathbf{corr} = \begin{pmatrix} 1 & .5 & .5 \\ .5 & 1 & .5 \\ .5 & .5 & 1 \end{pmatrix}$$

Two efronts for previous means & covariances

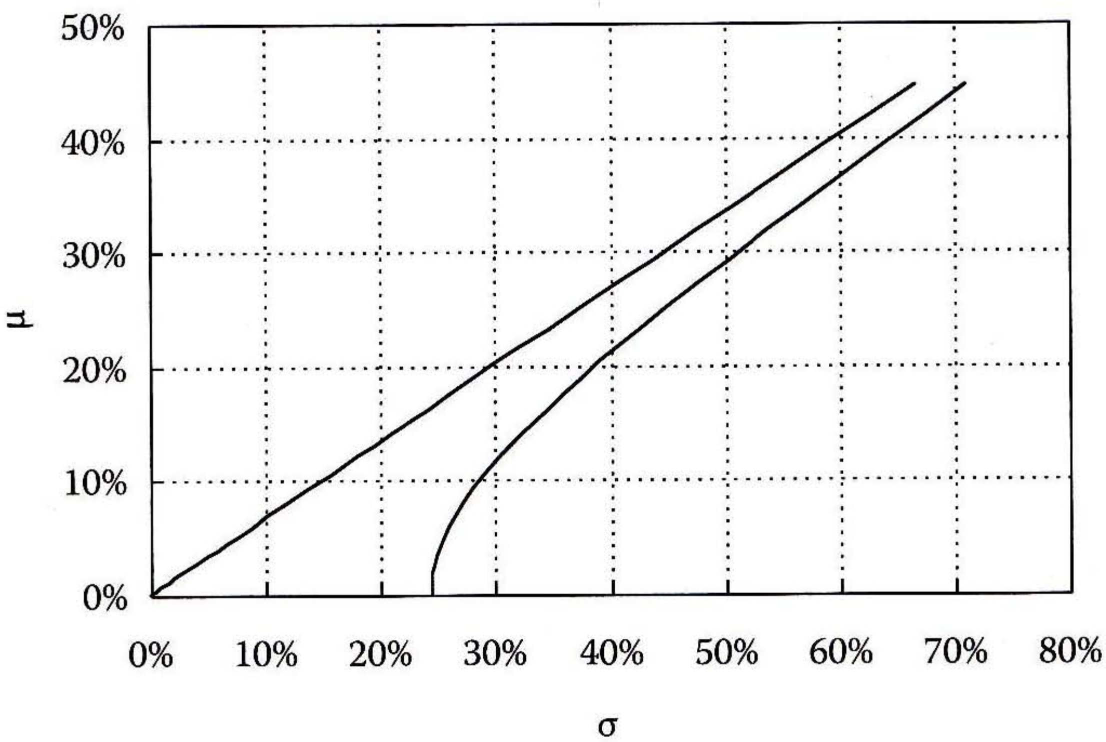


FIGURE 2.2. Efficient frontiers: the curved line is the efficient frontier of a fully invested equity portfolio, and the straight line is the efficient frontier of a long-short dollar neutral portfolio.

The two efronts for 10% higher mean returns

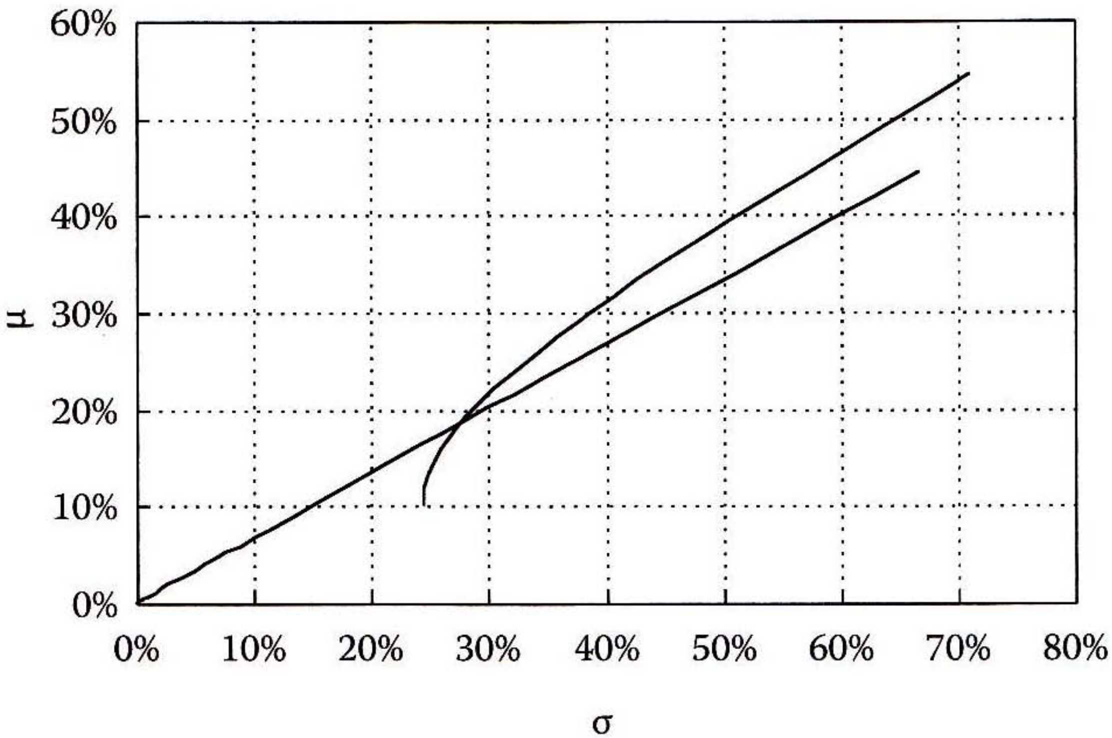


FIGURE 2.3. Efficient frontiers similar to those in Figure 2.2, except for the change in the expected returns, which are 10% higher for each stock.

4.3 Value Added Active Management

Expected excess (rate of) return on asset *i* : $\mu_i = E(R_i) - 1 - r_f$

gross return

$$\mu_i = \beta_i \cdot \mu_B + \alpha_i$$
$$= \beta_i \cdot \bar{\mu}_B + \underbrace{\beta_i \cdot \Delta\mu_B + \alpha_i}$$

μ_B { Benchmark mean return

Exceptional excess return

$\beta_i \cdot \bar{\mu}_B$ **The risk premium**: Very long-term benchmark returns, e.g., 70+ years (??). 3% - 7% for most equity markets.

$\beta_i \cdot \Delta\mu_B$ **Benchmark timing**: Difference between expected benchmark return in the near future and in the long-term. Long term average expected value of $\Delta\mu_B$ is 0.

α_i **Stock selection**: Manager skill in picking stocks.

Portfolio Mean Returns Decomposition

Follows Grinold and Kahn (2000). We have the usual representation

$$r_P = \beta_P r_B + \varepsilon_P$$

The portfolio expected (excess) returns are

$$\mu_P = \beta_P \mu_B + \alpha_P = (\beta_{PA} + 1)\mu_B + \alpha_P .$$

We substitute

$$\mu_B = \bar{\mu}_B + \Delta\mu_B ,$$

where $\bar{\mu}_B$ is the long-term average (consensus return) and $\Delta\mu_B$ is the "local" variation of benchmark mean (benchmark timing), to get:

$$\boxed{\mu_P = \mu_B + \beta_{PA}\bar{\mu}_B + \beta_{PA}\Delta\mu_B + \alpha_P}$$

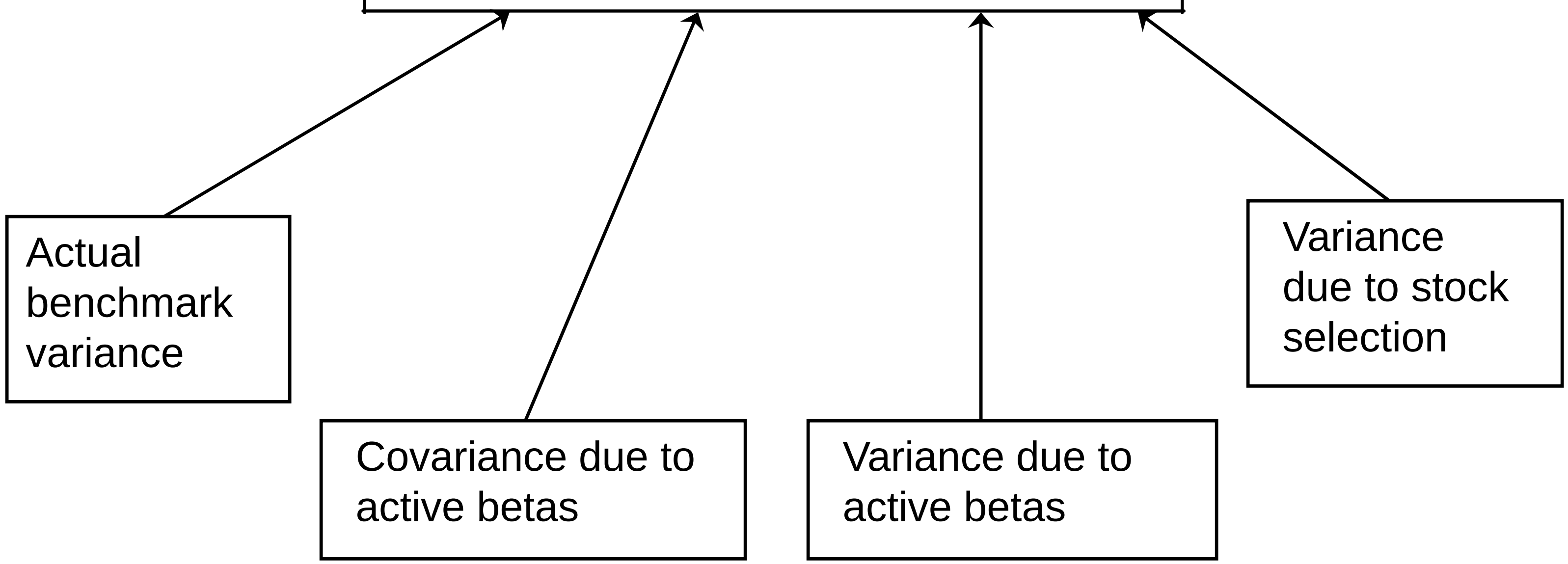
Portfolio Mean Returns Decomposition

$$\mu_P = \mu_B + \beta_{PA}\bar{\mu}_B + \beta_{PA}\Delta\mu_B + \alpha_P$$

- μ_B: Actual benchmark expected excess return
- $\beta_{PA}\bar{\mu}_B$: Return due to active management and consensus return
- $\beta_{PA}\Delta\mu_B$: Return due to active management and benchmark timing
- α_P: Return due to stock alphas and stock selection

Portfolio Variance Decomposition

$$\sigma_P^2 = (1+\beta_{PA})^2\sigma_B^2 + \sigma_\varepsilon^2$$
$$= \sigma_B^2 + 2\beta_{PA}\sigma_B^2 + \beta_{PA}^2\sigma_B^2 + \sigma_\varepsilon^2$$



Now express quadratic utility for the portfolio *P* in terms of the previous quantities, and generalize the risk aversion by using different risk aversions for different components below:

$$\mu_P - \lambda_T \sigma_P^2 = \mu_B - \lambda_T \sigma_B^2$$ Forecast, no action } Can't change

$$+\beta_{PA}(\bar{\mu}_B - 2\lambda_T \sigma_B^2)$$ Curious small term } Ignore

$$+\beta_{PA}(\Delta\mu_B - \lambda_{BT}\beta_{PA}\sigma_B^2)$$ Forecast + action

$$+\left(\alpha_P - \lambda_R \sigma_\varepsilon^2\right)$$ Forecast + action

} Value added

Value added: $$VA = \beta_{PA}(\Delta\mu_B - \lambda_{BT}\beta_{PA}\sigma_B^2) + \left(\alpha_P - \lambda_R \sigma_\varepsilon^2\right)$$

Requires exceptional benchmark timing and active beta

Focus on this!

4.4 Fundamental Law of Active Mgm't.

A well-known "law" whose proofs are involved and often confusing, and whose usefulness is somewhat questionable. None-the-less we need to know something about. The basic idea is due to Richard Grinold and rationalized and extended by others. See the book by Grinold and Kahn (2000) and the some of the many papers:

- Clarke, R., Silva, H. and Thorley, S. (2002). "Portfolio Constraints and the Fundamental Law of Active Management." *Financial Analysts Journal*, vol. 58, no. 5 (September/October):48–66.
- Clarke, R., de Silva, H. and Thorley, S. (2006). "The Fundamental Law of Active Management." *Journal of Investment Management*, vol. 4, no. 3: 54–72.
- Grinold, R. C. (1989). "The Fundamental Law of Active Management." *The Journal of Portfolio Management*, vol. 15, no. 3 (Spring): 30–38.
- Grinold, R. C. (1994). "Alpha is Volatility Times IC times Score." *The Journal of Portfolio Management*, vol. 20, no. 4 (Summer): 9–16.

- Qian, E, and Hua, R. “Active Risk and Information Ratio.” *The Journal of nvestment Management*, vol 2, no. 3 (2004), pp. 20-34.
- Qian, E., Hua, R., and Sorensen, E.H. (2007). *Quantitative Equity Portfolio Management: Modern Techniques and Applications,* London: CRC Press.
- Sorensen, E. H., Hua, R., Qian, E. and Schoen, R. (2004). “Multiple Alpha Sources and Active Management.” *The Journal of Portfolio Management*, vol. 30, no. 2 pp. 39-45.
- And other references to be added ……..

The Initial Grinold Version (1989)

The "information coefficient" IC is defined as the sample cross-section correlation coefficient between a forecast of returns and the returns them selves. As such is interpreted as a measures of skill of the portfolio manager. Then assuming a quadratic utility portfolio optimization with unconstrained weights except for full investment, Grinold derives the following fundamental law of active management (FLAM) formula for the information ratio

$$IR = \sqrt{BR} \cdot IC$$

where BR is defined as the number of independent forecast of exceptional return the manager makes (per year). If there are N assets in the portfolio and an independent forecast is made for each return then the formula becomes

$$IR = \sqrt{N} \cdot IC$$

Qian et. al. 2007 Extension

In a more thorough treatment the above authors use risk adjusted forecasts and returns, and taking into account that the as a cross-section correlation coefficient the IC is a random variable, derive the following version of FLAM:

$$IR = \frac{E(IC)}{\sigma(IC)}$$

Since IC is a sample correlation coefficient $\hat{\rho}$ one can use the following result from statistics under normality:

$$\sigma(\hat{\rho}) \approx \frac{(1-\rho^2)}{\sqrt{N}}$$

Since in practice values of IC are small, e.g., .1 to .2, applying the above approximation to IR gives:

$$IR \approx \sqrt{N} \cdot E(IC)$$

Clarke et. al. 2002 Extension

The Grinold original formula is limited by the assumption of no constraints in the optimization except full-investment. This is not realistic in practice, so Clarke et. al. (2002) extended the result to the case of constrained optimization by introducing the a "transfer coefficient" TC defined as the cross-section correlation between the weights under a constrained optimization and the forecasts. Then the FLAM becomes

$$IR = \sqrt{N} \cdot IC \cdot TC$$

Since TC is a correlation coefficient it is bounded above by one and can be substantially smaller for constrained optimization. Clarke et. al. give some interesting examples of the reduction in IR that occurs for various types of constraints.

Some FLAM Theory Building Blocks

There are many papers on this topic and there does not yet exist a totally clear and satisfactory derivation of the FLAM. None-the-less the next two conceptual building blocks that are worth knowing about.

Active Return and Initial Simplistic *IC*

The key intuition is that if a portfolio manager has effective "signals" for forecasting asset returns, these signals can be transformed to a specification of portfolio weights. So let's assume that with $\mathbf{g}$ the transformed signal we have active weights:

$$\mathbf{w} = c\mathbf{g} \qquad \mathbf{g} = (g_1, g_2, \cdots, g_N)' \qquad \mathbf{w}'\mathbf{1} = c\mathbf{g}'\mathbf{w} = 0$$

Then we can express active returns r_A as:

$$r_A = \sum_{i=1}^{N} w_i r_i$$

NOTE: $r_A \rightarrow \alpha$ in QHS (bad choice!)

$$= \sum_{i=1}^{N} w_i (r_i - \bar{r}_{cs})$$

$$\bar{r}_{cs} = \frac{1}{N} \sum_{i=1}^{N} r_i$$

Now substituting $\mathbf{w} = c\mathbf{g}$ we have:

$$
\begin{aligned}
r_A &= c \cdot \sum_{i=1}^{N} g_i \cdot (r_i - \overline{r}_{cs}) \\
&= cN \cdot \frac{1}{N} \sum_{i=1}^{N} g_i (r_i - \overline{r}_{cs}) \\
&= cN \cdot \widehat{\text{cov}}_{cs}(\mathbf{g}, \mathbf{r}) \\
&= cN \cdot \widehat{\text{corr}}_{cs}(\mathbf{g}, \mathbf{r}) \cdot \hat{\sigma}_{g,cs} \cdot \hat{\sigma}_{r,cs}
\end{aligned}
$$

$IC \triangleq$ cross-section sample correlation coefficient

cross-section sample std. deviations

$\hat{\sigma}_{r,cs} =$ "opportunity" (cross-section volatility)

$\hat{\sigma}_{g,cs} =$ "conviction" (quality of forecast)

Alpha Equals IC Times Volatility Times Score (Grinold, 1994)

Appendix 4A.1 shows that among all linear predictors $\hat{Y} = a + b \cdot X$ of the random variable Y , the minimum MSE predictor is:

$$\hat{Y} = \mu_Y + \frac{\operatorname{cov}(X,Y)}{\sigma_X^2} \cdot X, \qquad \mu_Y = E(Y), \ \ \sigma_X^2 = \operatorname{var}(X)$$

More general result that is fairly easy to prove: Among all non-linear predictors $\hat{Y} = h(X)$ of Y , the minimum MSE predictor is the conditional mean:

$$\hat{Y} = E(Y \mid X)$$

This conditional mean is in general nonlinear. However for the special case of bivariate normal random variables X and Y the conditional mean turns out to be linear and is given by the minimum MSE linear predictor above.

Summarizing: In the case of normality $\hat{Y} = \mu_Y + \frac{\operatorname{cov}(X,Y)}{\sigma_X^2} \cdot X = E(Y \mid X)$.

Now let $Y = r_A$, $X = g$, with $\mu_A = E(r_A)$, $\mu_g = E(r_g)$, and consider the model:

$$r_A = a + b \cdot g + \varepsilon, \qquad E(\varepsilon \mid g) = 0$$

$$\mu_A = a + b \cdot \mu_g \quad \rightarrow \quad a = \mu_A - b \cdot \mu_g$$

<u>Conditional</u> expected value predictor of r_A given g for the above model :

$$\begin{aligned} \hat{r}_A &= E(r_A \mid g) \\ &= a + b \cdot g \\ &= \mu_A + b \cdot (g - \mu_g) \end{aligned}$$

Since this is a linear predictor we can plug in the optimal b that yields minimum MSE to get:

$$\hat{r}_A = \mu_A + \frac{\operatorname{cov}(r_A, g)}{\sigma_g^2} \cdot (g - \mu_g)$$

In active portfolio management one assumes that the unconditional mean is zero: $\mu_A = 0$ (alpha generation is a zero-sum game). In that case:

$$\begin{aligned}
\hat{r}_A &= \frac{\operatorname{cov}(r_A, g)}{\sigma_g^2} \cdot (g - \mu_g) \\
&= \frac{\operatorname{cov}(r_A, g)}{\sigma_A \sigma_g} \cdot \sigma_A \cdot \frac{g - \mu_g}{\sigma_g} \\
&= \rho(r_A, g) \cdot \sigma_A \cdot score_g \\
&= IC_{r_A, g} \cdot volatility_{r_A} \cdot score_g
\end{aligned}$$

$$\sigma_A^2 = \operatorname{var}(r_A)$$

Note that: $E\left(score_g\right) = 0, \quad \operatorname{var}\left(score_g\right) = 1$

N.B. In the above expression $\hat{r}_A$ is fixed conditioned on g but since g is a random variable that will vary with different market information conditions used to construct $\hat{r}_A$ is unconditionally a random variable.

4.5 Numerical Active Optimization

Two Methods (as usual, see C&K Sections 9.8.1 and 9.8.4):

- Active MVO (AMVO): Minimize tracking error subject to portfolio mean return constraint and other constraints, or maximize portfolio mean return subject to ***TE*** constraint and other constraints
- Active Quadratic Utility (AQU)

AMVO Primal Method

$$TEV = \sigma_A^2 = \mathrm{var}(r_P - r_B) \qquad \rightarrow \qquad TE = \sigma_A$$

$$\sigma_A^2 = \sigma_P^2 - 2\,\mathrm{cov}(r_P, r_B) + \sigma_B^2$$

So just need to minimize: $\quad \sigma_P^2 - 2\,\mathrm{cov}(r_P, r_B)$

$$\text{cov}(r_P, r_B) = \sum_{i=1}^{N} w_{P,i} \cdot \text{cov}(r_i, r_B)$$
$$= \sum_{i=1}^{N} w_i \cdot \gamma_i \qquad \gamma_i = \text{cov}(r_i, r_B)$$
$$= \mathbf{w}'_P \boldsymbol{\gamma}$$

Thus we minimize TE subject to constraints by minimizing:

$$\mathbf{w}'_P \boldsymbol{\Sigma} \mathbf{w}_P - 2 \cdot \mathbf{w}'_P \boldsymbol{\gamma}$$

subject to $\mathbf{w}'_P \boldsymbol{\mu} \geq \mu_{P,0} = \mu_B + \mu_{A,0}$ and other constraints (e.g., box, etc.)

↖ portfolio manager determines

AMVO Dual Method

Maximize $\mathbf{w}'\boldsymbol{\mu}$ subject to $TEV = \sigma_A^2(\mathbf{w}) = \sigma_0^2$ and other constraints.

Usually don't take this approach to achieve a desirable TEV, since must check for feasibility of σ_0^2. Just do primal problem for various μ_0 to achieve desired TEV.

Active Quadratic Utility Maximization Method

At first it may seem natural to formulate the active quadratic utility maximization problem as:

maximize $$AQU(\mathbf{w}) = \mathbf{w}'_A\boldsymbol{\mu} - \frac{\lambda}{2} \cdot \mathbf{w}'_A\boldsymbol{\Sigma}\mathbf{w}_A$$

subject to $\mathbf{w}'_A\mathbf{1} = 0$ and additional constraints on the portfolio weights $\mathbf{w}_P$, such as long-only, box, etc.

Example: Portfolio long-only constraint

$$\mathbf{0} \le \mathbf{w}_P = \mathbf{w}_A + \mathbf{w}_B \le \mathbf{1} \quad \Rightarrow \quad -\mathbf{w}_B \le \mathbf{w}_A \le \mathbf{1} - \mathbf{w}_B$$

This works if you have the benchmark weights (can afford the data subscription price). But if you don’t have the benchmark weights, you can use an equivalent formulation described next.

First note that: $\mathbf{w}'_A\boldsymbol{\mu} = \mathbf{w}'_P\boldsymbol{\mu} - \mathbf{w}'_B\boldsymbol{\mu}$

Note carefully that the benchmark weight vector $\mathbf{w}_B$ may have a very large dimension, e.g., for the R1000 benchmark it will be a 1,000 dimensional vector, as will $\boldsymbol{\mu}$. On the other hand $\mathbf{w}_P$ may have a relatively small number of non-zero weights, e.g., 100 to 200.

Now recall from the previous slides that:

$$\sigma_A^2 = \mathbf{w}'_P\boldsymbol{\Sigma}\mathbf{w}_P - 2\mathbf{w}'_P\boldsymbol{\gamma} + \sigma_B^2 \quad \text{where} \quad \gamma_i = \mathrm{cov}(r_i, r_B)$$

Thus:

$$AQU(\mathbf{w}_A) = \mathbf{w}'_P\boldsymbol{\mu} - \frac{\lambda}{2}\cdot\left(\mathbf{w}'_P\boldsymbol{\Sigma}\mathbf{w}_P - 2\mathbf{w}'_P\boldsymbol{\gamma}\right) - \underbrace{\left(\mathbf{w}'_B\boldsymbol{\mu} + \frac{\lambda}{2}\sigma_B^2\right)}_{\text{a constant}}$$

So subject to constraints we just maximize:

$$AQU(\mathbf{w}_P) = \mathbf{w}'_P\boldsymbol{\mu} - \frac{\lambda}{2}\cdot\left(\mathbf{w}'_P\boldsymbol{\Sigma}\mathbf{w}_P - 2\mathbf{w}'_P\boldsymbol{\gamma}\right)$$

Collecting the linear terms in $AQU(\mathbf{w}_P)$ on the previous slide into a single linear term gives the following, which you maximize subject to constraints:

$$AQU(\mathbf{w}_P) = \mathbf{w}'_P\left(\boldsymbol{\mu} + \lambda\boldsymbol{\gamma}\right) - \frac{\lambda}{2}\cdot\mathbf{w}'_P\boldsymbol{\Sigma}\mathbf{w}_P$$

As usual when $\lambda \to 0$ you are maximizing the portfolio mean return subject to the constraints. But when $\lambda \to \infty$ you are minimizing

$$\mathbf{w}'_P\boldsymbol{\Sigma}\mathbf{w}_P - 2\mathbf{w}'_P\boldsymbol{\gamma}$$

Setting the derivative of the above expression equal to zero and solving for the optimal weights gives:

$$\mathbf{w}_P = \boldsymbol{\Sigma}^{-1}\boldsymbol{\gamma}$$

Let’s examine the structure of $\boldsymbol{\gamma}$.

We know that $\boldsymbol{\gamma}$ has elements $\gamma_i = \text{cov}(r_i, r_B), \; i = 1, \cdots, N$ where $r_i, \; i = 1, \cdots, N$ are the asset returns used to calculate the portfolio returns r_P. However the benchmark returns r_B are given by

$$r_B = \sum_{i=1}^{N} w_{B,i} r_i + \sum_{j=1,\, j \neq \text{ any } i}^{M} w_{B,j} r_j$$

Where M is a number generally much larger than N. For example an enhanced index portfolio might be constructed with N = 150 and M = 1,000. However, even when we don't have the weights we can still compute an estimates of $\gamma_i = \text{cov}(r_i, r_B), \; i = 1, \cdots, N$ and σ_B^2 , and thereby estimate the tracking error $TE = \sigma_A$ by estimating the tracking error variance:

$$TEV = \sigma_A^2 = \mathbf{w}_P' \boldsymbol{\Sigma} \mathbf{w}_P - 2\mathbf{w}_P' \boldsymbol{\gamma} + \sigma_B^2$$

Note that when there are many more assets in the benchmark than in your portfolio, you can not obtain $TE = 0$ as $\lambda \to \infty$. However, in practice you can obtain very small values of TE with 100-200 assets.

In the very special (and unrealistic) case where the assets in your portfolio are the same as the assets in the benchmark, you will get $TE = 0$ as $\lambda \to \infty$. To see that this is the case, just note that in this special case

$$\boldsymbol{\gamma} = \boldsymbol{\Sigma} \mathbf{w}_B$$

and so as $\lambda \to \infty$ you have

$$\begin{aligned} \mathbf{w}_P &= \boldsymbol{\Sigma}^{-1} \boldsymbol{\gamma} \\ &= \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma} \mathbf{w}_B \\ &= \mathbf{w}_B \end{aligned}$$

and the tracking error will (obviously) be zero.

AQU R-Code Example

```
library(xts)
library(quadprog)
load("C:/Doug/AMath Courses/AMATH 543/Data/crsp.short.Rdata")
ret = largecap.ts[,1:20]
p = ncol(ret)
ret.B = ret%*%rep(1/p,p)
#ret.B = largecap.ts[,"market"]

constraints = function(A,b,meq)
{
    list(A=A,b=b,meq=meq)
}
# Long-Only Constraints
p = ncol(ret)
A = cbind(rep(1,p), diag(rep(1,p))) # Constraint matrix
b = c(1,rep(0,p))                   # Constraint bound
cset.lo = constraints(A,b,1)

# For back-test use default wts.only = T
# For efficient frontier use wts.only = F
# For  printout of results use digits = 3 or 4
# Active Quadratic Utility Optimal Portfolio
```

```
aqu = function(ret,ret.B,cset=NULL,lambda,wts.only=T,digits =
NULL)
{
    require(quadprog)
    V = var(ret)
    Vlambda = lambda*var(ret)
    g = as.numeric(cov(ret,ret.B))
    mu = apply(ret,2,mean)
    p = ncol(ret)
    d = mu + lambda*g
    if(is.null(cset))
    {A = cbind(rep(1,p))
    b = 1
    meq = 1}
    else
    {A = cset$A
    b = cset$b
    meq = cset$meq}
    port.aqu = solve.QP(Vlambda,d,A,b,meq)
    wts = port.aqu$solution      # Get optimal weights
```

```
    mu = sum(wts*mu)
    sigma.sq = as.numeric(t(wts)%*%V%*%wts)
    sigma = sqrt(sigma.sq)
    wts = as.numeric(wts)
    if(!is.null(digits))
    {wg = t(wts)%*%g
    tev = as.numeric(sigma.sq - 2*wg + var(ret.B))
    te = sqrt(tev)
    mu.B = mean(ret.B)
    sd.B = sqrt(var(as.numeric(ret.B)))
    names(wts)= dimnames(ret)[[2]]
    out = list(WTS = wts,MU.PORT = mu,SD.PORT = sigma,MU.BM =
               mu.B,SD.BM = sd.B,TE = te)
    lapply(out,round,digits)}
    else
    {if(wts.only) wts else c(mu,sd,wts)}
}

# Test qu.constrained
cset = cset.lo
aqu(ret,ret.B,cset=cset,500,digits = 5)
aqu(ret,ret.B,cset=cset,.01,digits = 3)
```

```
$WTS
  AMAT   AMGN   CAT   DD    G     GENZ     GM      HON       KR     LLTC
0.10154 0.10177 0.10264 0.09617 0.09780 0.10181 0.09777 0.09992
0.09992 0.10066

$MU.PORT
[1] 0.01964

$SD.PORT
[1] 0.06497

$MU.BM
[1] 0.01945

$SD.BM
[1] 0.06477

$TE
[1] 0.00062
```

Note that with $\lambda = 500$ the tracking error is less than .1%. Run the code yourself to see what you get with $\lambda = .01$ and try different λ values. Also see what you get with $\lambda = 500$ and the market as benchmark instead of the equal weighted large-cap portfolio.

AQU R-Code Efficient Frontier Example

To be added in next version of slide deck

4A.1 Minimum MSE Prediction

1-D random variables only. Will treat vector case later.

You have random variables X and Y and wish to "predict" Y with a linear (strictly speaking "affine") function of X :

$$\hat{Y} = a + b \cdot X$$

a: Intercept; b: Slope

and want to minimize the prediction mean-squared-error

$$\text{MSE}(a,b) = E(Y - a - b \cdot X)^2$$

Take derivatives of MSE(a,b) with respect to a and b, and set equal to zero:

$$\frac{\partial}{\partial a} MSE(a,b) = -2 \cdot E(Y - a - b \cdot X) = 0$$

$$\frac{\partial}{\partial b} MSE(a,b) = -2 \cdot E\left(X \cdot (Y - a - b \cdot X)\right) = 0$$

The solutions $\tilde{a}$ and $\tilde{b}$ are:

$$\tilde{a} = E(Y) - \tilde{b} \cdot E(X)$$ (generally non-zero!)

$$\tilde{b} = \frac{E\left(X \cdot (Y - E(Y))\right)}{E\left(X \cdot (X - E(X))\right)} = \frac{E\left((X - E(X)) \cdot (Y - E(Y))\right)}{E(X - E(X))^2} = \frac{\mathrm{cov}(X,Y)}{\mathrm{var}(X)}$$

So the MMSE predictor is:

$$\begin{aligned}\hat{Y} &= \tilde{a} + \tilde{b} \cdot X \\ &= E(Y) - \tilde{b} \cdot E(X) + \tilde{b} \cdot X \\ &= E(Y) + \tilde{b} \cdot (X - E(X)) \\ &= E(Y) + \frac{\mathrm{cov}(X,Y)}{\mathrm{var}(X)} \cdot (X - E(X))\end{aligned}$$

It is easy to check that the minimum MSE is:

$$\begin{aligned}\sigma_{\hat{Y}}^2 &= E(Y - \hat{Y})^2 \\ &= \sigma_Y^2 \cdot (1 - \rho_{X,Y}^2)\end{aligned}$$

Prediction Error Properties

$$\begin{aligned}\hat{\varepsilon} &= Y - \hat{Y} \\ &= Y - (\tilde{a} + \tilde{b} \cdot X) \\ &= Y - E(Y) - \tilde{b} \cdot \left(X - E(X) \right)\end{aligned}$$

Zero Mean:

$$\begin{aligned}E(\hat{\varepsilon}) &= E\left[Y - E(Y) - \tilde{b} \cdot \left(X - E(X) \right) \right] \\ &= E\left(Y - E(Y) \right) - \tilde{b} \cdot E\left(X - E(X) \right) \\ &= 0\end{aligned}$$

Errors are Orthogonal* to Predictor:

$$E\left(X \cdot \hat{\varepsilon}\right) = E\left(X \cdot (Y - \tilde{a} - \tilde{b} \cdot X)\right) = 0$$

Errors are Uncorrelated with Predictor

$$\operatorname{cov}\left(X, \hat{\varepsilon}\right) = E\left(X \cdot \hat{\varepsilon}\right) - E(X) \cdot E(\hat{\varepsilon}) = 0$$

Model with Intercept in Error Term

$$Y = \tilde{b} \cdot X + \hat{\varepsilon}_0 \qquad \text{where} \qquad \hat{\varepsilon}_0 = \tilde{a} + \hat{\varepsilon}$$

X and $\hat{\varepsilon}_0$ are uncorrelated, but in general $E\left(\hat{\varepsilon}_0\right) \neq 0$!

* Two random variables X and Y are said to be orthogonal if $E(XY) = 0$

Model with Intercept in the Error Term (continued)

Let $\hat{\varepsilon}_* = Y - \tilde{b} \cdot X$

Optimal $\tilde{b}$ for predictor with intercept, but dropping intercept $\tilde{a}$

Then

$$
\begin{aligned}
\operatorname{cov}\left(X, \hat{\varepsilon}_*\right) &= \operatorname{cov}\left(X \cdot (Y - \tilde{b}X)\right) \\
&= \operatorname{cov}(X, Y) - \tilde{b} \cdot \operatorname{cov}(X, X) \\
&= \operatorname{cov}(X, Y) - \frac{\operatorname{cov}(X, Y)}{\sigma_X^2} \cdot \sigma_X^2 \\
&= 0
\end{aligned}
$$

So X and $\hat{\varepsilon}_*$ are also uncorrelated !

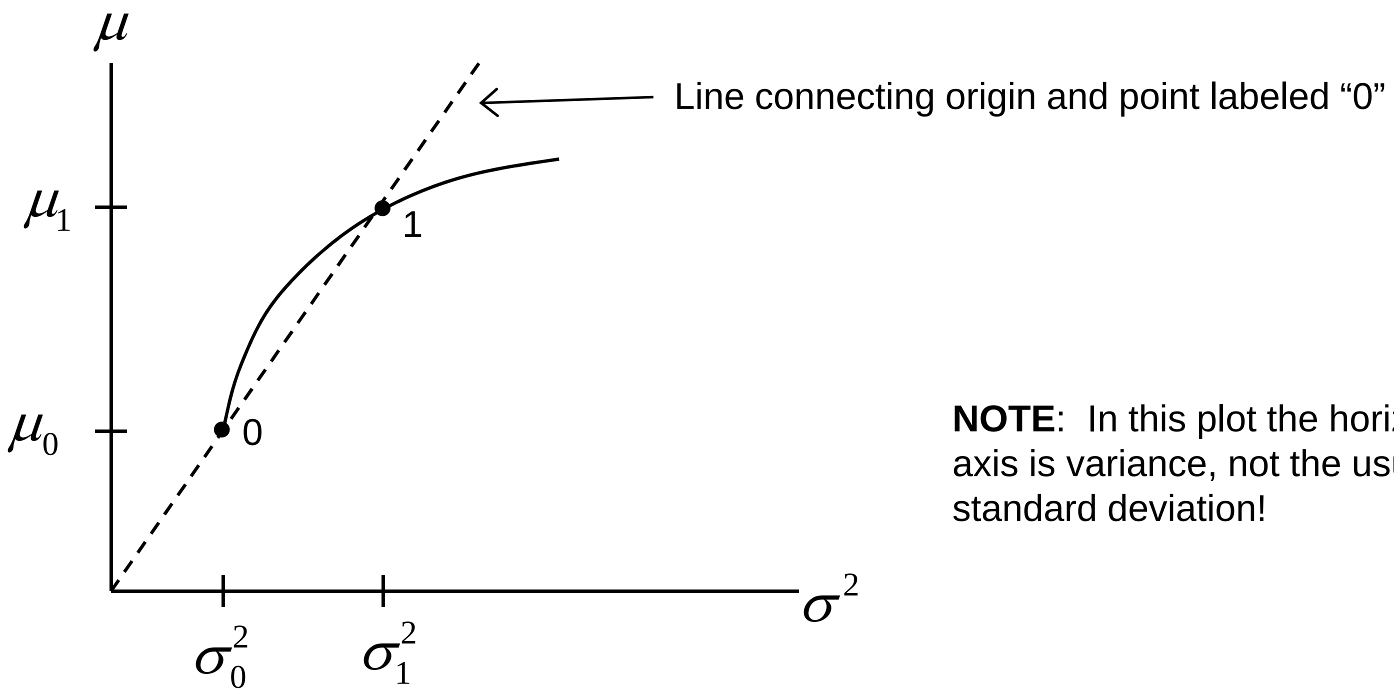
4A.2 TEV Optimal Portfolios

The R. Roll (1992) Solution

Assumption of unlimited short-selling allows for an analytic solution (R. Roll, 1992). Not of direct practical value, but interesting insights.

Minimize	$\mathbf{w}'_A \mathbf{\Sigma} \mathbf{w}_A$	$\left(= \sigma_A^2\right)$	(Minimize TEV)
Subject to	$\mathbf{w}'_A \cdot \boldsymbol{\mu} = \mu_A$	$\left(= \mu_P - \mu_B\right)$	(Expected active return)
	$\mathbf{w}'_A \cdot \mathbf{1} = 0$		(Self-financing)

The solution is in terms of the weight vectors $\mathbf{w}_0$ and $\mathbf{w}_1$, and associated portfolio means μ_0, μ_1 and variances σ_0^2, σ_1^2 shown in the picture below.



NOTE: In this plot the horizontal axis is variance, not the usual standard deviation!

Portfolio 0 is the familiar global minimum variance portfolio, and portfolio 1 is the portfolio given in the two-fund separation theorem obtained by maximizing quadratic utility. The expressions for the weight vectors, means and variances are given on the next slide for convenience.

Recall the quantities: $a = \mathbf{1}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}, \quad b = \boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}, \quad c = \mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1}$

Portfolio 0:

$$\mathbf{w}_0 = \boldsymbol{\Sigma}^{-1}\mathbf{1} / c$$

$$\mu_0 = a / c$$

$$\sigma_0^2 = 1 / c$$

Portfolio 1:

$$\mathbf{w}_1 = \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} / a$$

$$\mu_1 = b / a$$

$$\sigma_1^2 = b / a^2$$

Optimal Active Weights

$$\mathbf{w}_A = \frac{\mu_A}{\mu_1 - \mu_0} \cdot (\mathbf{w}_1 - \mathbf{w}_0) \qquad \left(\mu_A = \mu_P - \mu_B\right)$$

The active weights do not depend upon the benchmark portfolio!
All managers make the same alteration to the benchmark:

$$\mathbf{w}_P = \mathbf{w}_B + \mathbf{w}_A$$

Optimal Tracking Error Variance

$$\sigma_A^2 = TEV = \left(\frac{\mu_A}{\mu_1 - \mu_0}\right)^2 \cdot (\sigma_1^2 - \sigma_0^2)$$

Note that $\mu_A = 0 \Rightarrow \sigma_A = 0$!

Total Variance

$$\sigma_P^2 = \sigma_B^2 + TEV + \frac{2\mu_A}{\mu_1 - \mu_0} \cdot \sigma_0^2 \cdot \left(\frac{\mu_B}{\mu_0} - 1 \right)$$

Note that $\mu_A = 0 \;\Rightarrow\; \sigma_P = \sigma_B$

Beta of Portfolio P (on B)

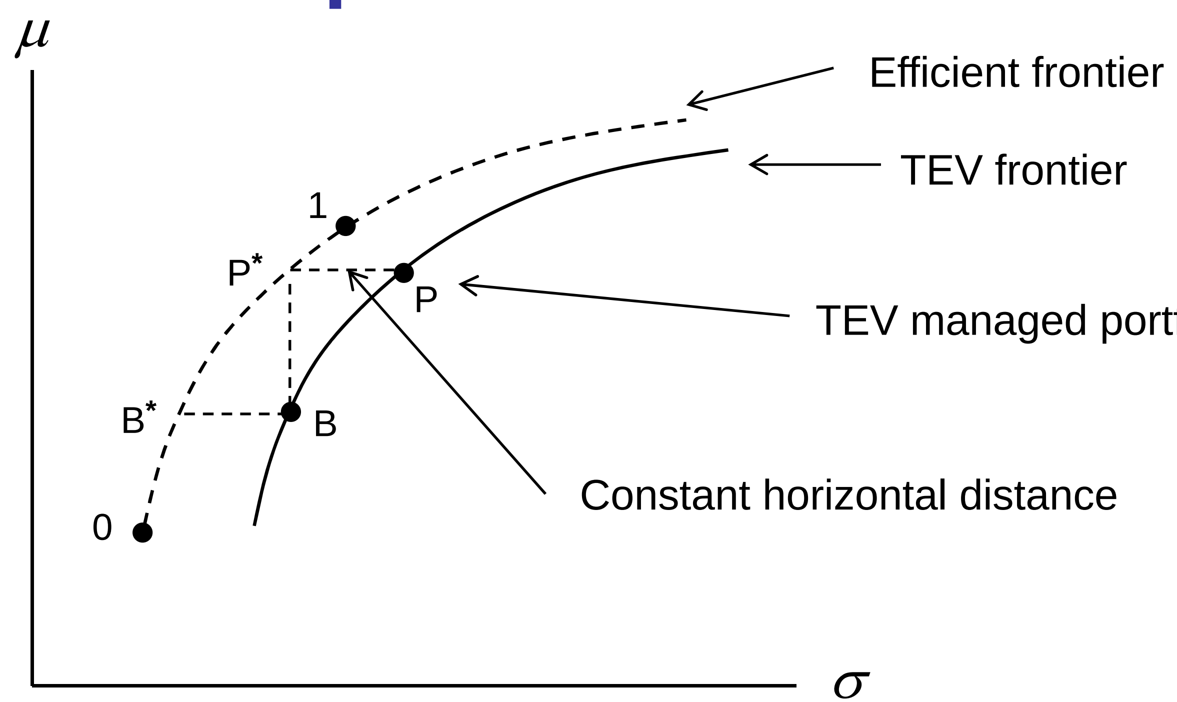
$$\beta_P = 1 + \frac{\mu_A}{\mu_1 - \mu_0} \cdot \frac{\sigma_0^2}{\sigma_B^2} \cdot \left(\frac{\mu_B}{\mu_0} - 1 \right)$$

Note that $\mu_A = 0 \Rightarrow \beta_P = 1$ and $\mu_A > 0 \Rightarrow \beta_P > 1$ (assuming $\mu_B > \mu_0$)

Proofs of the Above Results

Straightforward constrained quadratic minimization as in the efficient frontier with short-selling, plus straightforward (a little tedious) algebraic manipulations (see appendices in Roll, 1992)

TEV Optimal Portfolios are Inefficient



See Roll (1992) for details on these properties

- As P $\longrightarrow$ B, $TEV \longrightarrow 0$
- Typically P has higher μ_P and σ_P than B
- $\mu_B > \mu_0 \quad \Rightarrow \quad \beta_P > 1$
- It can be shown that all portfolio's P′ that dominate B with respect to both mean and standard deviation have $\beta_{P'} < 1$

The Jorion (2003) TEV Constrained Solution

Jorion (2003) cited the inefficiency of TEV optimal portfolios pointed out by R. Roll (1992), and proposed an alternate approach that is attractive, namely: Maximize active return subject to a constraint on tracking error variance as well as absolute variance, along with the usual self-financing condition on the active weights. That is:

Maximize $$\mathbf{w}'_A \cdot \boldsymbol{\mu} = \mu_A \qquad (= \mu_P - \mu_B)$$

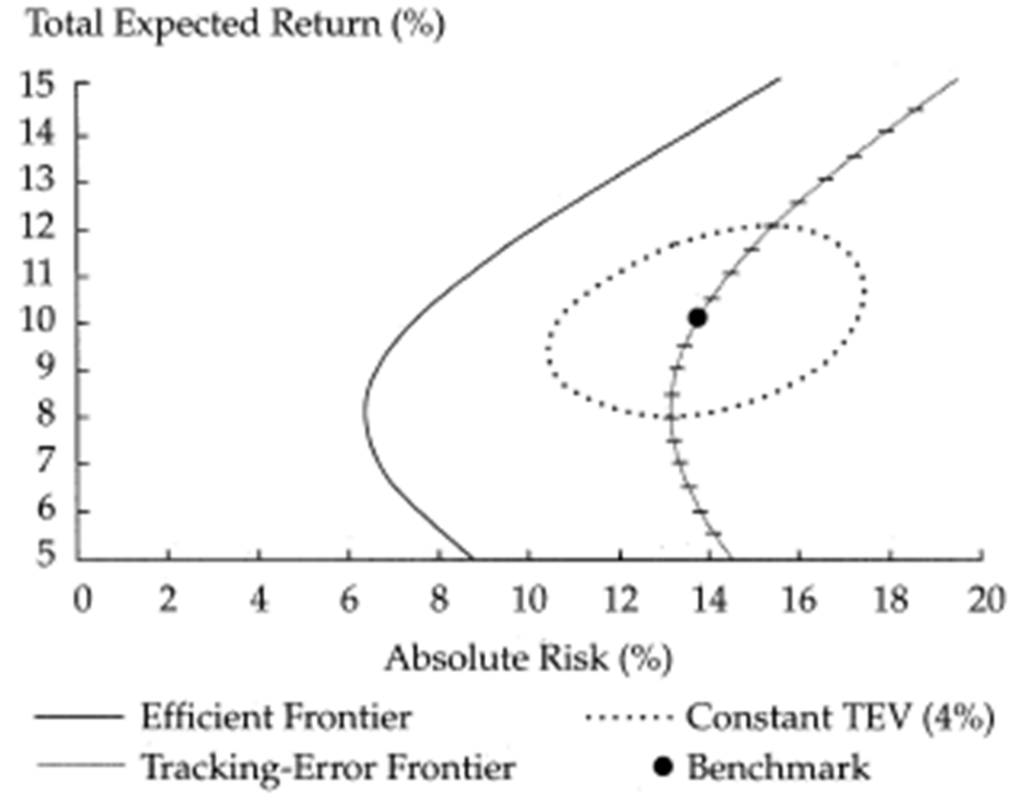
Subject to $$\mathbf{w}'_A \cdot \mathbf{1} = 0$$ (Self-financing)

$$\mathbf{w}'_A \boldsymbol{\Omega} \mathbf{w}_A = \tilde{\sigma}_A^2$$ (R.H.S. = TEV constraint)

$$\mathbf{w}'_P \boldsymbol{\Omega} \mathbf{w}_P = \tilde{\sigma}_P^2$$ (R.H.S. = absolute variance constraint)

Jorion (2003): Solution is ellipse in mean return versus absolute variance coordinates, a bit distorted in mean return versus absolute standard deviation coordinates.

Figure 4. Frontier with Constant Tracking-Error Volatility



With TEV constraint of about 11.5% and total variance constraint set at benchmark variance (about 14% std. dev.) the return improvement is approximately 3+ %.

Figure 5. Constant-TEV Frontiers for Various Steps