Quantitative Investments

Dale W.R. Rosenthal 1

1 June 2018

¹info@q36llc.com

Dale W.R. Rosenthal

Quantitative Investments

O₃₆

Last lecture we discussed structured products and PE.

- Mortgages;
- Securitization;
- Collateralized Debt Obligations;
- Credit Enhancement;
- Private Equity; and,
- Other Structured Products.

Today we will talk about active portfolio management.

Active Portfolio Management

Chapter 24, A Quantitative Primer on Investments with R

Q₃₆

- Today we will discuss active portfolio management.
- In particular, we will discuss:
 - Bottom-Up: Treynor-Black Approach;
 - Bayesian Statistics;
 - Top-Down: Black-Litterman Approach;
 - Nonparametric: Almgren-Chriss Approach;
 - Risk Parity Portfolios;
 - Practical Issues; and,
 - Valuing Active Management.

Why Active Management?

- Markowitz-Roy Modern Portfolio Theory (MPT) has a flaw:
- Jensen's Inequality! Suggests we can beat MPT:

$$\underbrace{\max_{w} E(U(w|X))}_{\text{stochastic optimization}} \ge \underbrace{\max_{w} U(w|E(X), \text{Var}(X))}_{\text{modern portfolio theory}}.$$
 (1)

- Example: if $X \sim N(0,1)$, $[E(X)]^2 = 0 < E(X^2) = 1$.
- Brinson et al suggest we focus on asset allocation.
- Grinold-Kahn: hold market + asset allocation/active management.
- Some suggest smart beta: hold market, add risk factors.²
- Stochastic optimization is hard; many approximations proposed.
- Popular: Treynor-Black, Black-Litterman, Almgren-Chriss, risk parity.

Q₃₆

² "Smart beta" is often undefined: buy-and-hold factors? factor timing?

The Treynor-Black Model

- If MPT, single index model are OK: just accommodate alpha.
- That gives us the (bottom-up) *Treynor-Black (1973) model*:

Decompose portfolio into market *M*, active *A* pieces:

$$\mathbf{r}_i = \mathbf{r}_f + \alpha_A + \beta_A (\mathbf{R}_M + \epsilon_M) + \epsilon_i \quad \mathbf{r}_i \perp \mathbf{r}_j, \mathbf{R}_M \forall i, j.$$
(2)

2 Put weights w_M , $w_A = 1 - w_M$ on market, active pieces.

3 Instrument *i* active weight $w_i^A = \frac{\alpha_i/\sigma_{\epsilon_i}^2}{\sum_{i=1}^n \alpha_i/\sigma_{\epsilon_i}^2}$.

4 Active A, market-neutral alpha A* portfolio metrics:

$$\alpha_{A^*} = \alpha_A = \sum_{i=1}^n w_i^A \alpha_i, \quad \beta_A = \sum_{i=1}^n w_i^A \beta_i, \quad \sigma_{A^*}^2 = \sum_{i=1}^n (w_i^A)^2 \sigma_{\epsilon_i}^2.$$
(3)

Regroup M, A into total market M^* , market-neutral alpha A^* .

$$M^* = w_M M + \underbrace{w_A \beta_A A}_{A} , \qquad A^* = w_A A (1 - \beta_A).$$
(4)

market exposure of A

Dale W.R. Rosenthal

- Portfolio A^* weight $w_{A^*} = \frac{\alpha_{A^*}/\sigma_{A^*}^2}{E(R_M)/\sigma_M^2 + (1-\beta_A)\alpha_{A^*}/\sigma_{A^*}^2}$.
- Other weights: $w_{M^*} = 1 w_{A^*}$, and $w_i^{A^*} = w_{A^*} w_i^A$.
- Risky portfolio metrics: $\beta_P = w_{M^*} + w_{A^*}\beta_A$, $E(R_P) = \beta_P E(R_M) + w_{A^*}\alpha_A$, and $\sigma_P^2 = \beta_P^2 \sigma_M^2 + w_{A^*}^2 \sigma_{A^*}^2$.
- Alpha portfolio information ratio: $IR_{A^*} = \sqrt{\sum_{i=1}^{n} \frac{\alpha_i^2}{\sigma_{\epsilon_i}^2}}$.
- Portfolio Sharpe ratio: $S_P = \sqrt{S_M^2 + IR_A^2}$.
- Thus alpha increases our returns, often at lower risk.

Treynor-Black: Issues

The Treynor-Black model can have some problems:

- Active and market portfolios might be large.
- Can and do we want to hold, say, +5A 4M?
- Can generate high α_P , but also high σ_P .
- Effect: leveraged bet on alpha vs market; are we so certain of alpha? *Mean-variance optimization . . . is extremely sensitive to . . . assumptions the investor must provide. — Black and Litterman (1992).*
- Can add penalties on w_A , w_M ; try shrinking alphas.

$$\mathsf{Maximize} \ S_P - \kappa_L \max(w_A, w_M, 0)^2 - \kappa_S \min(w_A, w_M, 0)^2 \qquad (5)$$

where κ_L, κ_S are capital usage penalties.

• Can also try squeezing alphas.

- We keep hitting problems due to input uncertainty.
 - Worse: We are not even certain how uncertain we are.
 - Also, we have no way to merge in beliefs or biases.
- One solution: *Bayesian statistics*.
 - Uses a prior distribution to express a priori beliefs.
 - Mix prior with data likelihood to get posterior distribution.
 - Loose ("flat" / "weak") priors let data control posterior.
 - Tighter priors let preconceived beliefs affect posterior.

- An example: What is P(coin flip = heads)?
- Recall: *n* flips, P(head) = p, # heads $k \sim \text{Binomial}(n, p)$.
 - Conjugate prior (make math easier) for Binomial is Beta.
- Flat prior: Beta $(\alpha, \beta = 1) = unif(0,1)$.
- Flip coin 50 times, get 20 heads. Unusual?
- Posterior: Beta(k + 1, n k + 1) = Beta(21, 31).
 - Posterior $E(k) = \frac{\alpha}{\alpha+\beta} = \frac{21}{52} = 0.404.$
 - Posterior $sd(k) = \sqrt{\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}} = \sqrt{\frac{21\cdot31}{52^2\cdot53}} = 0.07.$
- 95% credible interval: $0.5 \in [0.276, 0.539]$. Seems fair.
- Strong prior of fair coin: $Beta(\alpha = \beta > 1)$.

- Treynor-Black hints at difficulty: we want good inputs.
- Black and Litterman (1992): top-down, Bayesian approach.
- Allows merging historical data, equilibrium model, views;
- Can accommodate non-normality of returns, other risk measures;
- Use macro, cyclical factors for return forecasts; and,
- Use higher-frequency data to forecast covariance matrix.
- Bayesian approach squeezes parameters toward mean; sensible.
- Bayesian approach also enables easy stochastic optimization.

- The Black-Litterman model can be broken into 5 steps:
 - 4 Estimate covariance matrix Σ with recent data.
 - 2 Determine baseline (prior) forecast and precision.
 - 3 Express views quantitatively.
 - 4 Add views to get revised (posterior) forecast and precision.
 - Optimize portfolio using the posterior distribution.
- Typically done with normal dist'n + mean, variance.
- Also usually stocks and bonds; ignore commodities, RE, FX.
- Fatter-tailed dist'ns + coherent risk could be used.

- 1. Estimate Σ : can use recent, frequent data; factor models.
 - Yields $\hat{\Sigma}$ for asset classes: (σ_B^2, σ_S^2) on diagonal, Σ_{BS} off diagonal.
- 2. Assume market is in equilibrium (efficient) to start.
 - Outstanding bond, stock amounts \implies weights: $w^T = (w_B, w_S)$.
 - Prior portfolio variance: $\sigma_M^2 = w^T \Sigma w = w_B^2 \sigma_B^2 + w_S^2 \sigma_S^2 + 2w_B w_S \Sigma_{BS}$.
 - Mean risk aversion $\bar{\lambda} \stackrel{?}{=} 3 \implies$ asset class risk premia.
 - In equilibrrium, risk premia = forecast: $E(R) = \overline{\lambda} \hat{\Sigma} w$.
 - Get forecast covariance $Var(\hat{R}) = \hat{\Sigma}/n$; n = # obs. to estimate Σ .
 - Thus our baseline prior: $\hat{R}^{\text{prior}} \sim N(\bar{\lambda}\hat{\Sigma}w,\hat{\Sigma}/n)$.

- 3. Next, express quantified views Q for picks P.
 - e.g. Bonds will outperform stocks by 1%; bond excess returns=3%:

$$\underbrace{\begin{bmatrix} 0.01\\0.03\end{bmatrix}}_{E(Q)} = \underbrace{\begin{bmatrix} 1 & -1\\1 & 0\end{bmatrix}}_{P} \underbrace{\begin{bmatrix} \hat{R}_{B}\\\hat{R}_{S}\end{bmatrix}}_{\hat{R}}.$$
 (6)

- The shakiest part: guess view uncertainty $Var(Q) = \Omega$.
- Concern: inept/dishonest/overconfident analyst can bias portfolio.
- Forecast model/empirical Bayes might give better Ω.
- This give us our "data" (views) likelihood: $P\hat{R} \sim N(Q, \Omega)$.

Black-Litterman: Finding the Posterior, Optimizing

- 4. Use priors, views to get posterior forecast distribution.
 - $E(\hat{R}^{\text{post}})$ is precision-weighted mean of baseline, view:

$$\hat{R}^{\text{post}} \sim N\left(\hat{\Sigma}^{\text{post}}\left[n\hat{\Sigma}^{-1}\bar{\lambda}\hat{\Sigma}w + P^{T}\Omega^{-1}Q\right], \hat{\Sigma}^{\text{post}}\right)$$
(7)
$$\hat{\Sigma}^{\text{post}} = \left[n\hat{\Sigma}^{-1} + P^{T}\Omega^{-1}P\right]^{-1}.$$
(8)

• For posterior of returns, add forecast and return variance:

$$R^{\text{post}} \sim N\left(\hat{\Sigma}^{\text{post}}\left[n\hat{\Sigma}^{-1}\bar{\lambda}\hat{\Sigma}w + P^{T}\Omega^{-1}Q\right], \hat{\Sigma} + \hat{\Sigma}^{\text{post}}\right).$$
(9)

- 5. Then optimize portfolio to find new $w^{\text{post}} = (w_B^{\text{post}}, w_S^{\text{post}})$.
 - Could do Markowitz-Roy, but why? Suboptimal; wastes hard work.
 - Could also use non-normal distributions, coherent risk.
 - Have distributions; do stochastic optimization! Simulate/quadrature.

Q₃₆

- So far, approaches have entailed distributional assumptions.
- Instead, use distribution-agnostic *nonparametric statistics*.
 - Use rank statistics $\rho = (\rho_1, \dots, \rho_k)$ wh/order returns $r = (r_1, \dots, r_k)$.
 - i.e. $r_i \ge r_j \iff \rho_i \ge \rho_j \ \forall i \ne j.$
 - Consider set of all rank-equivalent returns Q_r = {r' : ρ(r') = ρ(r)}.
- Suppose we compare two portfolios A and B w/weights w^A , w^B .
- Weakly prefer portfolio A vs B $(A \succeq B \text{ or } w^A \succeq w^B)$ if: $w^A \rho(r') \ge w^B \rho(r') \ \forall r' \in Q_r.$
- Can then impose budget constraints: $\sum_{i} w_{i}^{A} = 1, w_{i}^{A} \ge 0$, etc.
- Efficient portfolio is $P: P \succeq P'; P, P'$ meet all constraints.
- Can find P by using centroids of returns.

- Budget, industry/sector, risk constraints are common.
- *Risk parity*: each instrument/asset contributes equal risk.
- Like a risk-weighted analog of 1/n equal-weight portfolio.
- Has been successfully used at PanAgora, Bridgewater.
- For volatility, risk parity vol between min-variance and 1/n portfolio.
 - Thus risk parity portfolio is like shrinkage of 1/n portfolio.
- Is this a good idea? Maybe low-risk instruments return more?
 - Asness, Frazzini, and Pedersen (2014): low-beta stocks outperform.
 - Pearson; Boudt+Carl+Peterson: outperforms for coherent risk.
 - Outperformance seems to be about risk, not risk measure.
- Problem: over- (under-) weight if risk under- (over-)estimated.
 - Like problem with cap weighting; could squeeze risk measures.

- A few issues we should also consider/try to fix.
- Estimated vol: need Jensen's Inequality correction.
- Bond portfolios: less liquid, hard to index. Buys often:
 - *substitute* similar higher-yield bond;
 - earn sector/country intermarket spread;
 - anticipate future interest rates; or,
 - pickup yield of illiquidity premium.
- Sell OOM options to collect premium, rebalance portfolio.
- Bayesian adjustment to alphas? (Details follow)
- Model realized alphas on predicted alphas.

- Can use a Bayesian approach to squeeze alphas.
- Prior: $\alpha_A \sim N(0, \sigma_{\epsilon_A}^2)$ using historical $\sigma_{\epsilon_A}^2 = 0.02$; empirical Bayes!
- Suppose forecast yields $E(\alpha_A) = 0.01$, $Var(\alpha_A) = 0.001$.
- Combine prior and forecast (data) to get posterior:

$$\alpha_{\mathcal{A}} \sim N\left(\frac{\frac{0}{\sigma_{\epsilon_{\mathcal{A}}}^{2}} + \frac{0.01}{0.001}}{\frac{1}{\sigma_{\epsilon_{\mathcal{A}}}^{2}} + \frac{1}{0.001}}, \frac{1}{\frac{1}{\sigma_{\epsilon_{\mathcal{A}}}^{2}} + \frac{1}{0.001}}\right) = N(0.0095, 0.00095).$$
(10)

- Result is weighted average; squeezes forecast toward prior.
 - Weight by 1/Var(forecast); lower variance = more weight.
- Could later combine results from another analysis.

• If we actively managed a portfolio, we might ask:

- How successful was that management?
- How did individual strategies/decisions perform?
- Could management have been better?
- Measuring performance is hard; easy to game.
 - Gaming-resistant metrics may be used for *external* managers.
 - Other measures may only be useful for assessing internal managers.³
- Key to performance evaluation is disentangling:
 - Risk: Did we take risk and just get lucky?
 - Noise: How much performance is unpredictable $(\downarrow 0?)$?
 - Skill: What performance seems reproducible?

³This assumes we would not lie to ourselves by gaming metrics.

- Actual investors need to consider *dollar-weighted averages*.
 - Especially proper since capital invested may change.
 - DCF-based measure: internal rate of return, IRR :=

$$\left\{r: \overbrace{i_{0}+\frac{i_{1}}{1+r}+\frac{i_{2}}{(1+r)^{2}}+\cdots}^{\text{PV(cashflows in)}} = \overbrace{\frac{o_{1}}{1+r}+\frac{o_{2}}{(1+r)^{2}}+\cdots}^{\text{PV(cashflows out)}}\right\}.$$
 (11)

• Can also scale historical returns for risk taken (as in \S 8.4):

- Sharpe ratio: $S_P = \frac{\bar{r}_P \bar{r}_f}{\sigma_P}$ adjusts for volatility.
- Sortino ratio: $So_P = \frac{\vec{r}_P \vec{r}_f}{\theta_P}$ adjusts for semideviation (better).
- Cond'l Sharpe: $CS_P = \frac{\dot{r}_P \bar{r}_f}{ES_P}$ adjusts for expected shortfall (coherent).
- Treynor ratio: $T_P = \frac{\overline{r}_P \overline{r}_f}{\beta_P}$ adjusts for systematic-relative risk (sort of).

• Returns should always include transactions costs.

Benchmark-Relative Performance Metrics

- More sensible is to compare performance to benchmark *M*.
- Modigliani-Modigliani: scale volatility to σ_M w/risk-free F.

$$M_P^2 = r_{P^*} - r_M$$
, where $P^* = \frac{\sigma_M}{\sigma_P}P + \left(1 - \frac{\sigma_M}{\sigma_P}\right)F$. (12)

- Can look at *tracking error* vs benchmark we track *T*:
 - Often, people say "tracking error" but mean *target risk* $sd(R_E)$.

$$R_E = R_P - R_T \quad \text{or} \quad sd(R_E) = sd(R_P - R_T). \tag{13}$$

• Better: Factor exposure is cheap to attain; don't pay fees on that.

- Jensen's alpha: $\alpha_P = \bar{r}_P (\bar{r}_f + \beta_P(\bar{r}_M \bar{r}_f))$ CAPM-adjusted.
- Connor-Korajczak: look at α_P from a multi-factor model.
- Info ratio: $IR_P = \frac{\alpha_P}{\sigma_{\epsilon_P}}$ for α_P from factor model; like idiosyncratic S_P .
- *IR_P* in 0.4–0.6 range is very good; above 1 is rare for long.

- T_P good for portfolio additions to LARGE diversified portfolios.⁴
- S_P popular but not great; CAPM is insufficient alpha model.
- M^2 very good for comparing strategies of varying risk.
- *IR_P* good for judging active management.
- Note: S_P , IR_P are like *t*-stats; yet values > 2 are rare.
- Never good: R^2 . Can drive $R^2 \uparrow 1$: add garbage to model.
- In general, statistical analysis of manager alpha is tough.

⁴For large portfolios, idiosyncratic variance is small.

Metrics for Alternative Investments

- How to evaluate *alternatives*: Hedge funds, CTAs, PE/VC?
- Use Hasanhodzic-Lo (6-factor), Fung-Hsieh (7-/8-factor) models?
 - Problem: many of these factors are expensive to trade
- Can also consider simple Jurek and Stafford (2015):
 - Uses Bondarenko (2014) observation that puts are expensive.
 - Puts are insurance; so, consider a put-writing factor.

$$R_{i} = \alpha_{i} + \beta_{iM}R_{M} + \beta_{i,PW}PW + \epsilon_{i}.$$
(14)

- May also include lags of factors.
- Then filter for funds with good/robust α , *IR*.
- Many caveats with returns and metrics for alternatives:
 - $\bullet\,$ Funds may report CAPM α and IRs but not clarify that.
 - Observability, Peso problems, survivorship, reporting bias.

- How to detect market timing ability?
- Add nonlinear term to model (Treynor-Mazuy, Henrikksson-Merton):
 Significant γ̂ > 0 indicates timing ability.

$$R_P = \alpha + \beta_P (r_M - r_f) + \gamma (r_M - r_f)^2 + \epsilon, \quad (T-M)$$
(15)

$$R_P = \alpha + \beta_P (r_M - r_f) + \gamma (r_M - r_f)^+ + \epsilon. \quad (H-M)$$
(16)

- Another hint: semideviation θ_P much less than volatility σ_P .
- What is value of market timing?
 - Similar to a call (or call and put) option.
 - *i.e.* Get only upside performance but pay for that.

- Based on the data, do some investors seem to have skill?
- Analysts ratings are not informative; *changes* are.
 - Analysts who forecast alpha add some value ($M^2 = 2.1\%$)
- Mutual funds are more varied in performance:
 - Small group of funds+bond funds are consistently bad.
 - Most funds are just plays on momentum, have $\alpha < 0$.
 - No evidence of market timing ability either.
 - Could be some skilled managers + return-chasing investors.
- Hedge funds and mutual funds managed side-by-side:
 - Those MFs have alpha, not taken advantage of by HF; however,
 - HF managers who add a MF: MF has typical performance; and,
 - MF managers who add a HF: HF has subpar performance.

- If returns, risk vary across time, metrics may change.
- Changing means and variances add to risk.
 - Like adding between-group variation to within-group variation.

$$P_{t_0 \to t_1} : S_P = \frac{0.2}{0.3}; \tag{17}$$

$$P_{t_1 \to t_2} : S_P = \frac{0.1}{0.15}; \tag{18}$$

$$P_{t_0 \to t_2}: S_P = \frac{(0.1 + 0.2)/2}{\sqrt{(0.3^2 + 0.15^2)/2}} = \frac{0.15}{0.24} < \frac{2}{3}.$$
 (19)

Successful market timers would adjust positions.

• Get better performance when market goes up and down.

• Ingersoll et al (2007): Measures can be manipulated.

- Static: write OOM options, put premium in risk-free bonds;
- Dynamic: add leverage after extremes to mess w/stationarity;
- Dynamic: add measurement error by smoothing, illiquid holdings.
- They propose four properties of manipulation-proof measure:
 - In Functionhood: yields one score/investment;
 - 2 Scale invariance: same score for any notional;
 - 3 Unbiasedness: only informed investor can improve score (via arb);
 - 9 Economic consistency: Must agree w/market equilibrium.

• Thus their (certainty-equivalent) manipulation-proof measure $\hat{\Theta}$:

$$\hat{\Theta} = \frac{1}{(1-\rho)\Delta t} \log \left(\frac{1}{T} \sum_{i=1}^{T} \left[\frac{1+r_t}{1+r_{f,t}}\right]^{1-\rho}\right), \quad (20)$$

where ρ is the relative risk aversion (usu. near 3).

- What is a fair fee for good active management?
- Kane, Marcus, and Trippi use Treynor-Black perspective:

% fee =
$$f_{\text{one-time}} = \frac{S_P^2 - S_M^2}{2\bar{\lambda}} = \frac{IR_A^2}{2\bar{\lambda}} = \frac{\sum_{i=1}^n IR_i^2}{2\bar{\lambda}}.$$
 (21)

• If
$$S_P = 1$$
, $S_M = 0.8$, and $\bar{\lambda} = 3$, $f_{\text{one-time}} = \frac{0.36}{6} = 6\%$.

- Berk: Manager consumes all expected alpha.
- Glode: Manager may charge *more* than alpha!
 - Why? If alpha is supplied in hard times, that has more value.

What Fee is Fair? Mean-Variance Fee

- What is a fair fee for good active management?
- What if we consider fee vs mean-variance utility?
- Investor pays fee f = utility of receiving αf with volatility $\hat{\sigma}_{\alpha}$:

$$f = \alpha - f - \frac{\bar{\lambda}}{2}\hat{\sigma}_{\alpha}^2.$$
 (22)

• Implies manager claims annual fee as fraction of α :

$$\frac{f_{\text{annual}}^*}{\alpha} = \frac{1}{2} - \frac{\bar{\lambda}\alpha}{4IR_A^2} \tag{23}$$

- Foster and Young: No manipulation-proof fee structure.
- However, we should consider risk, risk aversion.

- Evaluating management performance is tough:
 - Need many observations of portfolio returns; and,
 - Parameters (μ, σ) change constantly over time.
- However, some *decompositions* are informative.
 - Break performance into meaningful pieces⁵.
- Measure baseline performance with benchmark B
 - Benchmark allocation $\{w_{B_k}\}_{k=1}^{K}$ among K asset classes.
 - Benchmark asset class weights, returns become counterfactuals.
- Benchmark *B*, portfolio *P* performance given by:

$$r_B = \sum_{k=1}^{K} w_{B_k} r_{B_k}$$
 $r_P = \sum_{k=1}^{K} w_{P_k} r_{P_k}.$ (24)

⁵For example: What we do versus don't control.

Performance Attribution: Decomposition

- Brinson-Hood-Beebower: decompose benchmark-excess returns.
- Idea: benchmark asset-class weights, returns = counterfactuals

$$r_{P} - r_{B} = \sum_{k=1}^{K} (\underbrace{w_{P_{k}}r_{P_{k}} - w_{P_{k}}r_{B_{k}}}_{\text{security/sector}} + \underbrace{w_{P_{k}}r_{B_{k}} - w_{B_{k}}r_{B_{k}}}_{\text{asset class}}).$$
(25)

- Further: Use sector *i* sub-portfolios $P_{k,i}$, $B_{k,i}$ for asset class *k*.
- Decompose asset class k sector/security selection $r_{P_k} r_{B_k}$:

$$r_{P_k} - r_{B_k} = \underbrace{\sum_{s=1}^{S} \left(\frac{w_{P_{k,s}}}{w_{P_k}} - \frac{w_{B_{k,s}}}{w_{B_k}} \right) r_{P_{k,s}}}_{S=1} + \underbrace{\sum_{s=1}^{S} \frac{w_{B_{k,s}}}{w_{B_k}} (r_{P_{k,s}} - r_{B_{k,s}})}_{S}.$$
(26)

36

We covered active portfolio management; on to investment firms next!

• All Together Now: Investment Firms, Crises.