Portfolio Performance Measurement

(Text reference: Chapter 20)

- background
- some common measures
- potential problems
- market timing
- the effects of skewness on risk measurement
- Morningstar’s risk-adjusted ratings
- style analysis
- persistence in mutual fund performance
- performance by individual investors

Background

- the simplest (and most popular) way is to compare returns with other similar investment funds, but the criteria are often vague: e.g. “resource sector”
- there are other more formal ways to measure performance on a risk-adjusted basis; this is usually done using the single index model:

\[ r_i - r_f = \alpha_i + \beta_i (r_M - r_f) + e_i \]

- there are two significant problems:
  - to get an accurate statistical assessment, we need a lot of observations
  - active management leads to shifts in parameters
Some common measures

- Jensen’s performance measure:

\[ \alpha_p = \bar{r}_p - \bar{r}_f - \beta_p (\bar{r}_M - \bar{r}_f). \]

This is simply the excess return over that predicted by CAPM.

- Appraisal ratio:

\[ \frac{\alpha_p}{\sigma(e_p)}. \]

This is the extra return over CAPM per unit of risk that could be diversified away by holding an index portfolio.

- Treynor’s measure:

\[ \frac{\bar{r}_p - \bar{r}_f}{\beta_p}. \]

This is the ratio of excess return to systematic risk.

- Sharpe’s measure:

\[ \frac{\bar{r}_p - \bar{r}_f}{\sigma_p}. \]

This is the ratio of excess return to total risk (a.k.a. the reward to variability ratio).
the $M^2$ measure:

$$M^2 = r_P - r_M.$$ 

Suppose that a portfolio $P$ is combined with a position in risk free assets such that the standard deviation of the total portfolio $P^*$ equals that of the market. Take the difference between returns of $P^*$ and $M$.

---

**text question p. 779:**

<table>
<thead>
<tr>
<th>Portfolio $P$</th>
<th>Market $M$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\bar{r}$</td>
<td>0.35</td>
</tr>
<tr>
<td>$\beta$</td>
<td>1.20</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>0.42</td>
</tr>
<tr>
<td>$\sigma(e)$</td>
<td>0.18</td>
</tr>
</tbody>
</table>

Calculate the various performance measures assuming $r_f = 0.06$. By which measures did $P$ outperform $M$?

- Jensen:

$$\alpha_P = \bar{r}_P - r_f - \beta_P (\bar{r}_M - r_f)$$

$$= 0.35 - 0.06 - 1.2(0.28 - 0.06) = 0.026$$

$$\alpha_M = 0$$
- Appraisal ratio:

\[
\frac{\alpha_P}{\sigma(e_P)} = \frac{.026}{.18} = .144
\]
\[
\frac{\alpha_M}{\sigma(e_M)} = \frac{0}{0} = ?
\]

- Treynor:

\[
\frac{\bar{r}_P - \bar{r}_I}{\beta_P} = \frac{.35 - .06}{1.2} = .2417
\]
\[
\frac{\bar{r}_M - \bar{r}_I}{\beta_M} = \frac{.28 - .06}{1} = .22
\]

- Sharpe:

\[
\frac{\bar{r}_P - \bar{r}_I}{\sigma_P} = \frac{.35 - .06}{.42} = .6905
\]
\[
\frac{\bar{r}_M - \bar{r}_I}{\sigma_M} = \frac{.28 - .06}{.30} = .7330
\]

- \( M^2 \):

\[
M_P^2 = \left( \frac{.30}{.42} \right) \times .35 + \left( 1 - \frac{.30}{.42} \right) \times .06 - .28 = -.013
\]
\[
M_M^2 = 0
\]

- Performance rankings can differ depending on which measure is used, so it is important to know which is appropriate under various circumstances.

- Under MV utility \( U = E(r_p) - \frac{1}{2} \lambda \sigma_P^2 \), the investor wants the highest possible Sharpe measure of his/her complete portfolio.

- Scenarios:

1. The portfolio under consideration is the complete portfolio of risky assets. In this case (assuming MV utility), the appropriate criterion is Sharpe’s measure. It should be compared to a broad market index (in which case \( M^2 \) is also appropriate), or to other similar portfolios.
2. The portfolio under consideration is one of several combined into a single large fund. Since there are several portfolios, nonsystematic risk should not matter, so we use $\beta$.

$$\begin{align*}
\tau - \tau_f \\
\text{SML}
\end{align*}$$

3. The portfolio under consideration is an actively managed portfolio and will be combined with the passive index portfolio $M$. Let the active portfolio be $A$ and the total portfolio $P$. Let $w$ be the % invested in $A$. Under MV utility we want to maximize the Sharpe measure of $P$:

$$S_P = \frac{w\tau_A + (1-w)\tau_M - r_f}{\left[w^2\sigma_A^2 + (1 - w)^2\sigma_M^2 + 2w(1-w)\sigma_{AM}\right]^{1/2}}.$$ 

It can be shown that

$$w^* = \frac{w_0}{1 + (1-\beta_A)w_0},$$

where

$$w_0 = \frac{\alpha_A/\sigma^2(e_A)}{\left(\tau_M - r_f\right)/\sigma_M^2}.$$

Moreover, at $w^*$:

$$S_P = \left[\left(\frac{\tau_M - r_f}{\sigma_M}\right)^2 + \left(\frac{\alpha_A}{\sigma(e_A)}\right)^2\right]^{1/2}.$$
Potential problems

- changing return distribution over time
  - standard statistical inference assumes means and variances are constant, but the whole point of active management is to change means and variances over time

- use of a market proxy
  - different rankings can occur depending on which proxy is used

- hard to distinguish luck from skill
  - need long data samples (text e.g. pp. 786-787)

- use of the T-bill rate
  - if you have to borrow at a higher rate, using the T-bill rate may set the standard too high
- CAPM validity
  - doesn’t apply to Sharpe’s measure, but if this is a concern, we can use APT instead to determine αs

- measurement of returns
  - time-weighted returns:
    * suppose you buy 1 share at $t = 0$ and a second share at $t = 1$. You sell both shares at $t = 2$. Let $P_0 = 60, P_1 = 65, P_2 = 68, D_1 = D_2 = 2$.
    Return in year 1:
    $$P_0 = \frac{D_1 + P_1}{1 + r_1} \Rightarrow 60 = \frac{2 + 65}{1 + r_1} \Rightarrow r_1 = 11.67\%$$
    Return in year 2:
    $$P_1 = \frac{D_2 + P_2}{1 + r_2} \Rightarrow 65 = \frac{2 + 68}{1 + r_2} \Rightarrow r_2 = 7.69\%$$
    Your time-weighted return is $(11.67\% + 7.69\%)/2 = 9.68\%$.

- dollar-weighted returns:
  * your dollar-weighted return is found by solving:
    $$P_0 + \frac{P_1}{1 + r} = \frac{D_1}{1 + r} + \frac{2 \times (D_2 + P_2)}{(1 + r)^2}$$
    $$\Rightarrow 60 + \frac{65}{1 + r} = \frac{2}{1 + r} + \frac{2 \times (2 + 68)}{(1 + r)^2}$$
    $$\Rightarrow r = 9.02\%$$
    - dollar-weighted returns are appropriate as a measure of the investor’s return over time, but not as a measure of a fund manager’s performance
– time-weighted returns are arithmetic averages:

\[ r_A = \frac{r_1 + r_2 + \ldots + r_N}{N} \]

– to incorporate compounding, use geometric averages:

\[ 1 + r_G = [(1 + r_1)(1 + r_2)\ldots(1 + r_N)]^{1/N} \]

– using the example above:

\[ 1 + r_G = [(1.1167)(1.0769)]^{1/2} \Rightarrow r_G = 9.66\% \]

– in general \( r_G \approx r_A - \frac{1}{2} \sigma^2 \)

– use \( r_G \) as a measure of past performance, \( r_A \) for a future performance estimate (see text e.g. pp. 809-810)

---

**Market timing**

• market timing may be defined as the ability to tell whether (over the short-term) T-bills will outperform the market

<table>
<thead>
<tr>
<th>U.S. 1927-78</th>
</tr>
</thead>
<tbody>
<tr>
<td>T-bills</td>
</tr>
<tr>
<td>NYSE</td>
</tr>
<tr>
<td>Perfect timing (1 month forecast)</td>
</tr>
</tbody>
</table>

• the value of being able to time the market can be estimated from the prices of index call options
• in practice fund managers don’t shift funds completely, nor do they invest exactly in $M$, but we should see an increase (decrease) in portfolio $\beta$s if $r_M - r_f > (\prec) 0$ if managers have timing ability

• can be tested for using:
  – quadratic regression:

\[
r_p - r_f = a + b(r_M - r_f) + c(r_M - r_f)^2 + e_p
\]

\[
\begin{align*}
 & r_p - r_f \\
 & r_M - r_f
\end{align*}
\]

– regression with a dummy variable:

\[
r_p - r_f = a + b(r_M - r_f) + c(r_M - r_f)D + e_p
\]

where

\[
D = \begin{cases} 
1 & \text{if } r_M > r_f \\
0 & \text{if } r_M \leq r_f
\end{cases}
\]

\[
\begin{align*}
 & r_p - r_f \\
 & r_M - r_f
\end{align*}
\]
The effects of skewness on risk measurement

- text example p. 793:

<table>
<thead>
<tr>
<th></th>
<th>Strategy 1</th>
<th>Strategy 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expected return</td>
<td>5%</td>
<td>7.5%</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>0%</td>
<td>2.5%</td>
</tr>
<tr>
<td>Maximum return</td>
<td>5%</td>
<td>10%</td>
</tr>
<tr>
<td>Minimum return</td>
<td>5%</td>
<td>5%</td>
</tr>
</tbody>
</table>

- another example (from “The Peculiar Persistence of Catastrophic Trading Strategies” by D. Emanuel):

Let me manage $1 million. I’ll get you a return of $2,000 every day. Am I a good trader?

Morningstar’s risk-adjusted ratings


- Morningstar Inc.’s RAR is the single most popular measure of U.S. mutual fund performance

- each fund is assigned to a peer group (star: domestic equity, international equity, taxable bond, municipal bond; category: 20 domestic equity categories, 9 international equity categories, 10 taxable bond categories, and 5 municipal bond categories) and rated relative to other funds in the peer group

- the RAR for a fund is calculated as:

\[
RAR_i = RRet_i - RRisk_i
\]
• letting $g(i)$ denote the peer group to which fund $i$ is assigned, then:

$$R_{Ret_i} = \frac{Ret_i}{B_{Ret_{g(i)}}}$$

and

$$R_{Risk_i} = \frac{Risk_i}{B_{Risk_{g(i)}}}$$

where $B_{Ret_{g(i)}}$ and $B_{Risk_{g(i)}}$ are the bases used for the relative return and risk for all funds in the group.

• furthermore:

$$Ret_i = VR_i - VR_b$$

$$B_{Ret_{g(i)}} = \max \left[ \text{mean}_{i \in g(i)}(Ret_i), VR_b - 1 \right]$$

$$Risk_i = -\text{mean}_{i \in g(i)}(\text{min}(ER_{it}, 0))$$

$$B_{Risk_{g(i)}} = \text{mean}_{i \in g(i)}(Risk_i)$$

• all funds in a peer group are ranked and the top 10% are given 5 stars (or a category 5 rating), next 22.5% get 4, next 35% get 3, next 22.5% get 2, and bottom 10% get 1.

• recall:

$$RAR_i = \frac{Ret_i}{B_{Ret_{g(i)}}} - \frac{Risk_i}{B_{Risk_{g(i)}}},$$

and rearrange to get:

$$RAR_i = \left( \frac{1}{B_{Ret_{g(i)}}} \right) \left( Ret_i - \frac{B_{Ret_{g(i)}}}{B_{Risk_{g(i)}}} \times Risk_i \right)$$

• this can be rescaled as:

$$RRAR_i = Ret_i - k_{g(i)} \times Risk_i$$
note the resemblance to the MV utility function, but the definition of risk used here implies that an investor who is picking a mutual fund using Morningstar’s ratings acts as if he/she has expected utility of:

\[ E(U) = E(R_i - R_b) + aE(L_i) \]

this utility function has linear indifference curves ⇒ extreme investment choices

however, rankings may still be consistent with the Sharpe measure:
Style analysis

- use an asset class factor model:

\[
\tilde{R}_i = b_{i1} \tilde{F}_1 + b_{i2} \tilde{F}_2 + \ldots + b_{in} \tilde{F}_n + \tilde{e}_i
\]

where \( \sum_{j=1}^{n} b_{ij} = 1 \)
- represents return on a fund as a portfolio invested in the \( n \) asset classes plus a residual component \( \tilde{e}_i \)
- note that:

\[
R^2 = 1 - \frac{\text{Var}(\tilde{e}_i)}{\text{Var}(\tilde{R}_i)}
\]

measures the proportion of the variance of the fund’s return “explained” by the \( n \) asset classes

---

- consider a 12 asset class model (note that the return of each class is represented by some type of market index):

<table>
<thead>
<tr>
<th>Class</th>
<th>Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bills</td>
<td>Salomon Brothers’ 90 day T-Bill index</td>
</tr>
<tr>
<td>Intermediate bonds</td>
<td>Lehman Brothers’ intermediate-term gov’t. bond index</td>
</tr>
<tr>
<td>Long-term bonds</td>
<td>Lehman Brothers’ long-term gov’t. bond index</td>
</tr>
<tr>
<td>Corporate bonds</td>
<td>Lehman Brothers’ corporate bond index</td>
</tr>
<tr>
<td>Mortgages</td>
<td>Lehman Brothers’ mortgage-backed securities index</td>
</tr>
<tr>
<td>Large capitalization value stocks</td>
<td>Sharpe/BARRA value stock index</td>
</tr>
<tr>
<td>Large capitalization growth stocks</td>
<td>Sharpe/BARRA growth stock index</td>
</tr>
<tr>
<td>Medium capitalization stocks</td>
<td>Sharpe/BARRA medium stock index</td>
</tr>
<tr>
<td>Small capitalization stocks</td>
<td>Sharpe/BARRA small stock index</td>
</tr>
<tr>
<td>Foreign bonds</td>
<td>Salomon Brothers’ non-U.S. government bond index</td>
</tr>
<tr>
<td>European stocks</td>
<td>FTA Euro-Pacific ex Japan index</td>
</tr>
<tr>
<td>Japanese stocks</td>
<td>FTA Japan index</td>
</tr>
</tbody>
</table>
• example: Trustees’ Commingled Fund U.S. Portfolio 1/85-12/89

<table>
<thead>
<tr>
<th></th>
<th>Unconstrained Regression</th>
<th>Constrained Regression</th>
<th>Quadratic Programming</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bills</td>
<td>14.69</td>
<td>42.65</td>
<td>0.00</td>
</tr>
<tr>
<td>Intermediate bonds</td>
<td>-69.51</td>
<td>-68.64</td>
<td>0.00</td>
</tr>
<tr>
<td>Long-term bonds</td>
<td>-2.54</td>
<td>-2.38</td>
<td>0.00</td>
</tr>
<tr>
<td>Corporate bonds</td>
<td>16.57</td>
<td>15.29</td>
<td>0.00</td>
</tr>
<tr>
<td>Mortgages</td>
<td>5.19</td>
<td>4.58</td>
<td>0.00</td>
</tr>
<tr>
<td>Large capitalization value stocks</td>
<td>109.52</td>
<td>110.35</td>
<td>69.81</td>
</tr>
<tr>
<td>Large capitalization growth stocks</td>
<td>-7.86</td>
<td>-8.02</td>
<td>0.00</td>
</tr>
<tr>
<td>Medium capitalization stocks</td>
<td>-41.83</td>
<td>-43.62</td>
<td>0.00</td>
</tr>
<tr>
<td>Small capitalization stocks</td>
<td>45.65</td>
<td>47.17</td>
<td>30.04</td>
</tr>
<tr>
<td>Foreign bonds</td>
<td>-1.85</td>
<td>-1.38</td>
<td>0.00</td>
</tr>
<tr>
<td>European stocks</td>
<td>6.15</td>
<td>5.77</td>
<td>0.15</td>
</tr>
<tr>
<td>Japanese stocks</td>
<td>-1.46</td>
<td>-1.79</td>
<td>0.00</td>
</tr>
<tr>
<td>Total</td>
<td>72.71</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td>$R^2$</td>
<td>95.20</td>
<td>95.16</td>
<td>92.22</td>
</tr>
</tbody>
</table>

• performance measurement
  - a passive manager provides style, an active manager provides both style and selection
  - define a fund’s selection return as the difference between the fund’s return and that of a passive mix with the same style, e.g.
    1. estimate fund’s style using months $t - 60, \ldots, t - 1$
    2. calculate return on resultant style for month $t$
    3. calculate selection return for month $t$ as difference between fund’s return in month $t$ and that of the style benchmark
    4. repeat for each month during some time period and examine selection returns throughout it
  - Sharpe examined 636 stock, bond, and balanced funds and found mean underperformance of -7.4 basis points per month (0.89% per year)
Persistence in mutual fund performance


- examines diversified U.S. equity funds 1/62-12/93

- no survivorship bias problem since all equity funds are included in the sample (as of 12/93, 1,892 funds of which 1,310 were “alive” and 582 were “dead”)

- the average fund had total net assets of $218 million

- on average a fund trades 77% of its assets in a year

- 64.5% of the funds charge load fees, which average 7.33%

- performance benchmarks
  - CAPM:
    \[ r_{it} = \alpha_i + \beta_i V W R F_t + e_{it} \]
  - 4 factor model:
    \[ r_{it} = \alpha_i + \beta_i R M R F_t + \gamma_i S M B_t + h_i H M L_t + p_i P R 1 Y R_t + e_{it} \]

- analysis
  - form 10 equally-weighted portfolios at start of each year sorted by excess returns in previous year
  - basic results are in Table III (CAPM does not perform very well, but 4 factor model explains most of the cross-sectional variation)
there is a 67 basis point spread between deciles 1 and 10; of this 38 basis points are due to the momentum factor and of the remaining 29 basis points 20 occur between deciles 9 and 10

decile 10 has higher than average expenses and relatively high turnover (Table IV); this accounts for some of its relatively poor performance

portfolio managers claim that expenses and turnover do not reduce performance, but the evidence in Table V (for individual funds) suggests this is not the case:
  – for every 100 basis point increase in expense ratios annual return drops 154 basis points
  – for every 100 basis point increase in turnover, annual return decreases 95 basis points

the composition of the top decile changes a lot each year (by roughly 80%)
⇒ positive persistence is short-lived

the average load fund underperforms the average no-load fund by 80 basis points per year (excluding the bottom 20% of funds)

short run mutual fund returns have strong persistence: the net gain from buying the decile of last year’s winners and selling the decile of last year’s losers is 8% per year (about 4.6% due to the 4 factors and 1.7% due to expense ratios and transactions costs; around 2/3 of the remainder is due to the spread between deciles 9 and 10)

3 rules-of-thumb for investors:
1. avoid funds with persistently poor performance
2. funds with high returns last year have higher than average expected returns this year but not in subsequent years
3. expense ratios, transactions costs, and load fees have a direct and negative impact on performance
Performance by individual investors


- use account data for 78,000 households from a discount brokerage from 1991-1996; focus on the subsample of 66,465 which owned common stock at some point during the period

- the average household had 2 accounts, the value of which was 60% stocks (an average of 4.3 stocks worth around $47,000)

- the average household turned over more than 75% of its stock portfolio each year

- the average round-trip trade cost 3% in commissions and 1% in spread

- over the sample period, an investment in a value-weighted index of all NYSE, AMEX, and NASDAQ stocks earned an average compound return of 17.9% per year

- the average household earned a gross return of 18.7%, but a net return of 16.4%

- in terms of risk-adjusted performance, the average household underperformed by almost 20 basis points per month (Jensen’s $\alpha$)

- investors tend to have small stocks with relatively high $\beta$s (so non-risk-adjusted performance looks better than it really is)

- there is a lot of variation across households (Table IV) (1/4 of households beat the market on a net basis by more than 6% per year; 1/4 underperform by more than 8% per year)

- the 20% of households which trade the most turn over an average of more than 20% of their stock portfolios per month, and they do very poorly (no better in gross terms, far worse on a net basis)