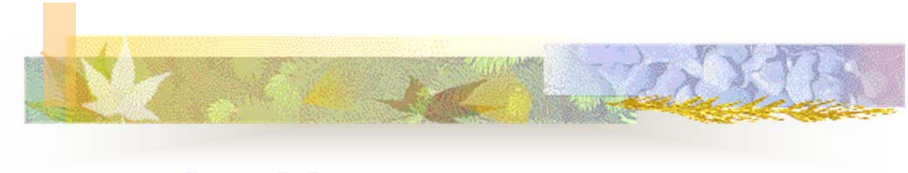
PEVA

Day 1 Slides

Performance Evaluation and Attribution (PEVA)



Russ Wermers

University of Maryland

January 2010

Performance Evaluation

- Why should we care?
  - As long as an asset manager provides returns that are market-beating it doesn't matter, right?
  - Or, if the manager can beat his/her peer funds, that is all we need to know, right?

Famous U.S. Asset Managers

Asset Manager #1: Peter Lynch
Fund: Fidelity Magellan (1977-1990)
Size: $20 Million (in 1977), $14 Billion (in 1990)
Strategy: U.S. Equity, Aggressive-Growth

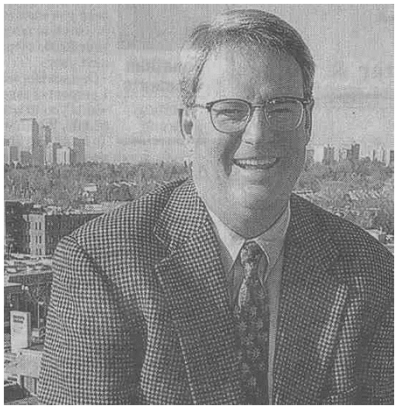
Asset Manager #2: Bill Miller
Fund: Legg-Mason Value Trust
Size: $10-billion
Strategy: "Follows a value discipline of investing by purchasing primarily large-capitalization stocks at large discounts to the manager's assessment of their intrinsic value."

Asset Manager #3: Scott Schoelzel

Fund: Janus 20

Size: $10-billion

Strategy: U.S. Equity, Concentrated Large-Cap Growth



Asset Manager #4: Gus Sauter

Fund: Vanguard 500 Index

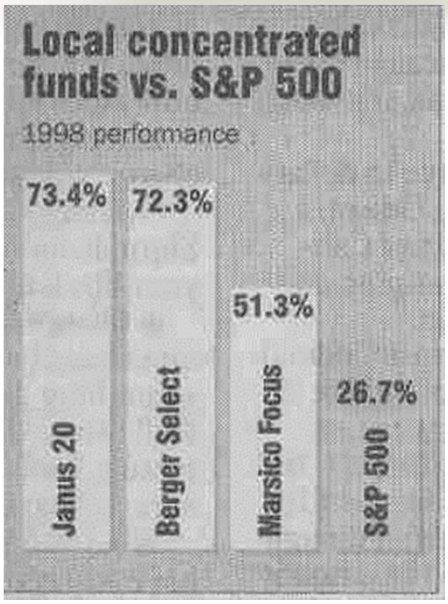
Size: $88.2 Billion (Dec. 2000)

Strategy: U.S. S&P Index

Asset Manager #5: Bill Gross (since 1987)
Fund: PIMCO Total Return
Size: $16-billion (At December 2002)
Strategy: Fixed-Income

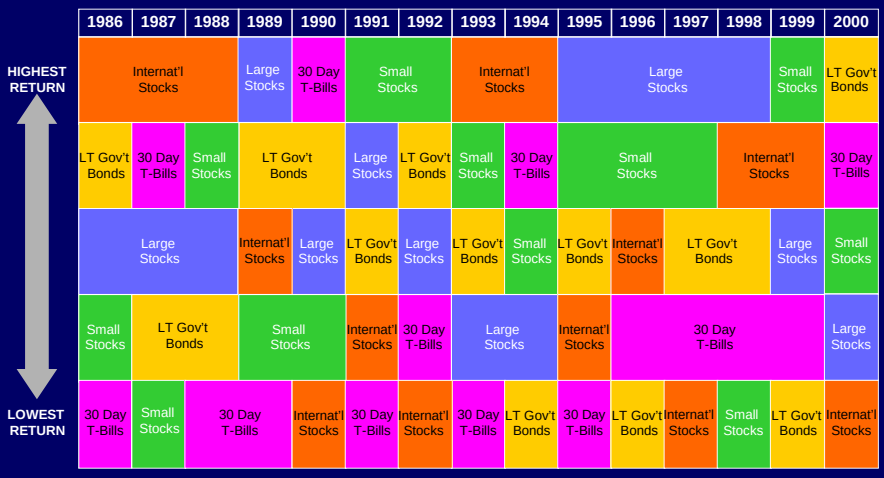
Let's Look at Manager #3

- Scott Schoelzel of the Janus 20 Fund, ranked the #1 money manager in the world during 1998 by *Mutual Funds* magazine



- Yet, from July 2000 to July 2001, the Janus 20 fund lost almost 50% of its assets!
- In any given year, or even calendar quarter, one asset class can outperform another
  - But, the relative performance reverses quite frequently!

Asset Class Winners and Losers



This is for illustrative purposes only and not indicative of any investment.
Past performance is no guarantee of future results. 1/1/2001.

Ibbotson Associates

- But, is past performance any guide to future performance?
  - Most studies say “no”
  - My new study, “Is Money Really ‘Smart’,” suggests “yes”
  - However, by no means is this issue settled, even for U.S. domestic equity funds, where plenty of research has been conducted
- In any case, it is important to set fair goals and incentives for managers, even if performance cannot be repeated

Basic Issues in Performance Evaluation

- 1. Minimize Type I Error: separate luck from skill (difficult without a long history of returns and good model and benchmarks)
- 2. Minimize Type II Error: maximize power of test
  - Related: choose proper benchmark(s): separate controllable results from outcomes that the manager cannot control (e.g., is currency risk a choice available to the manager?)
  - Related: provide a measure of performance that can be appropriately compared with peer performance
    - Not simply using unadjusted returns!

Basic Issues in Performance Attribution

- 1. Precisely determine the source of risk-adjusted performance (market timing (asset allocation), sector timing, security selection, currency bets)
- 2. Decompose performance separately for each class of assets to increase the understanding of what went right or wrong, e.g., stocks vs. currencies
  - Also, many funds have separate managers for each asset class

Momentum Investing Strategies

Originally Presented to:
Denver Roundtable
January 19, 2000

Outline

- Are momentum profits a statistical "fluke"?
- Yearly variations in momentum profits
- Estimated profits after transaction costs
- Reducing turnover with momentum strategies/trading only large stocks
- Momentum without short-selling
- Momentum as an industry effect
- Avoiding January

Papers Used for this Talk

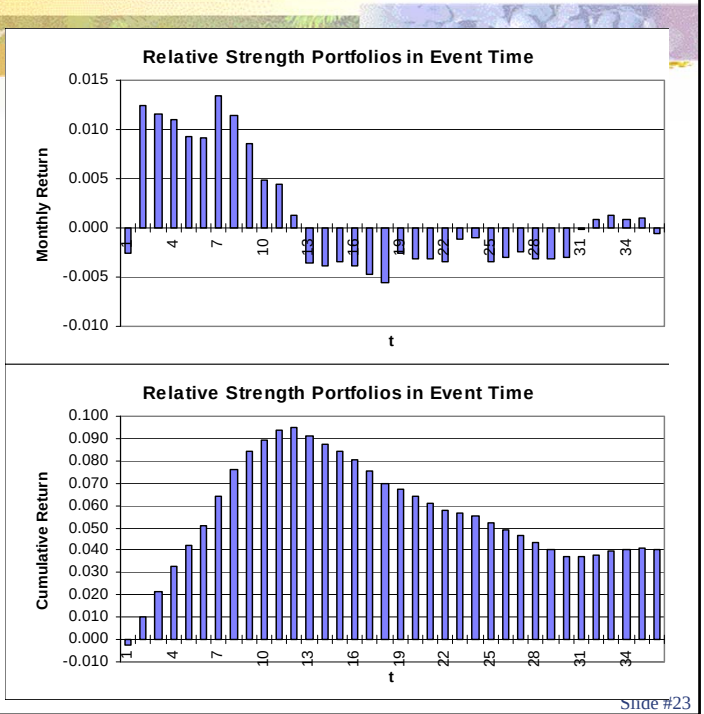
- "JT": Jegadeesh and Titman (1993; Journal of Finance)--Baseline Momentum Strategy Paper
- "CJL": Chan, Jegadeesh, and Lakonishok (1996; Journal of Finance)--Price Momentum and Earnings Momentum Paper
- "Mos": Moskowitz (1999; Journal of Finance)--Momentum as an Industry Effect
- "Car": Carhart (1997; Journal of Finance)--Mutual Fund Returns After Expenses

- "W1": Wermers (1997; unpublished)--Mutual Fund Returns Before Expenses
- "W2": Wermers (1999; Journal of Finance)--Mutual Fund Herding
- "GTW": Grinblatt, Titman, and Wermers (1995; American Economic Review)--Mutual Fund Momentum Strategies Before Expenses

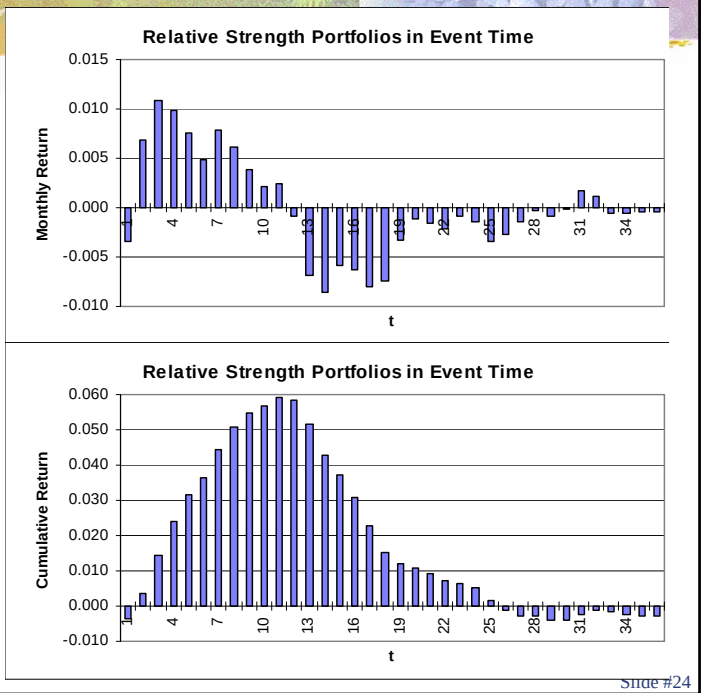
Momentum Profits: A Statistical Fluke?

- Look at profits to 6-month lagged return sort over three different time periods (JT)
  - 1965-1989
  - 1941-1964
  - 1927-1940

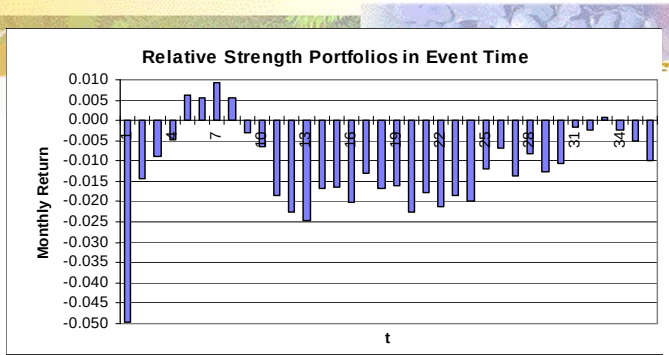
1965-1989

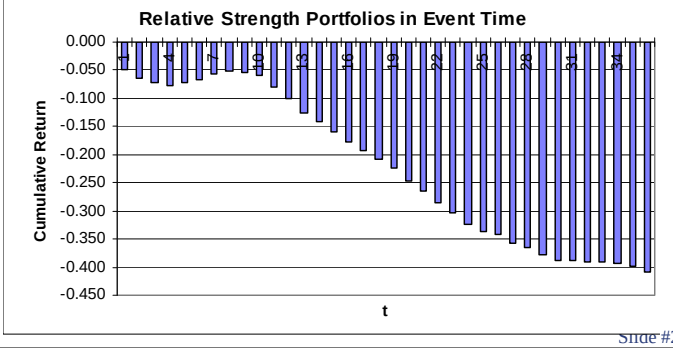


1941-1964



1927-1940

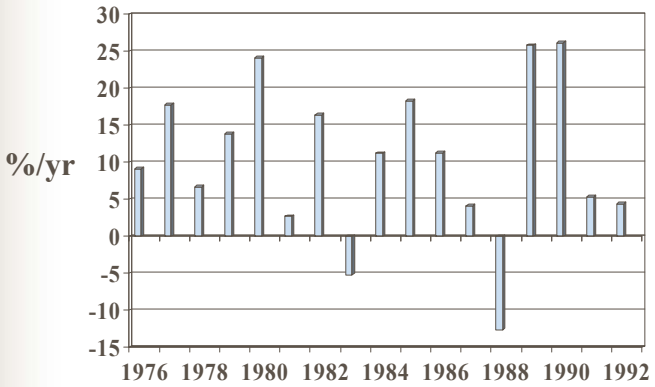




Other Evidence

- International Momentum Results—momentum in Europe, but not Asia
- Examine Trading Strategies of Mutual Funds
  - Momentum investing strategies widespread (GTW)
  - Buy-side herding in high past-return stocks, sell-side herding in low past-return stocks (W2)

Yearly Variations in Momentum Profits

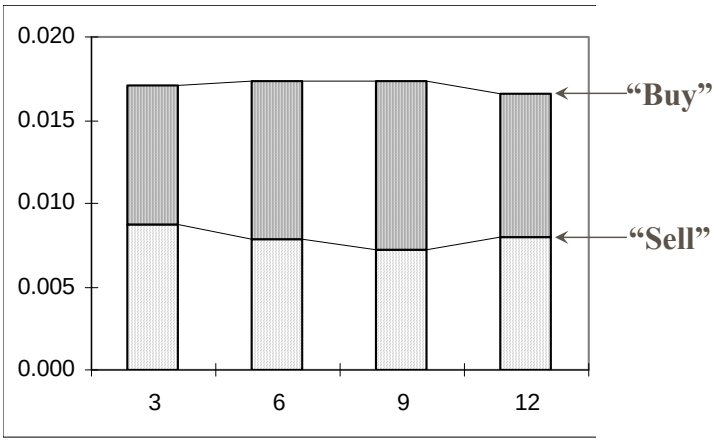


Profits After Transaction Costs

- Strategy: buy last-year's best mutual funds, sell last-year's worst
- Hold for one year
  - Compare hypothetical gross returns (W1) with actual net returns (Car)
  - Problem: sample periods are different!

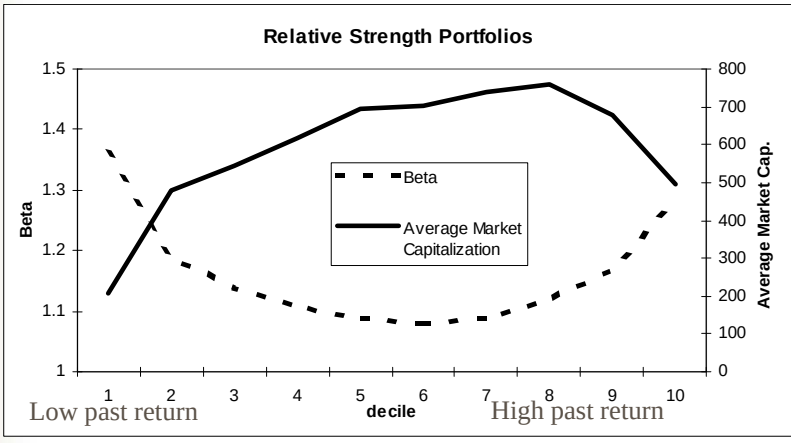
I am currently working on this problem.

Reducing Turnover: Use Longer Holding Period



6-mo lookback (JT)

Market Caps of Momentum Portfolios



Trading Only Large Stocks

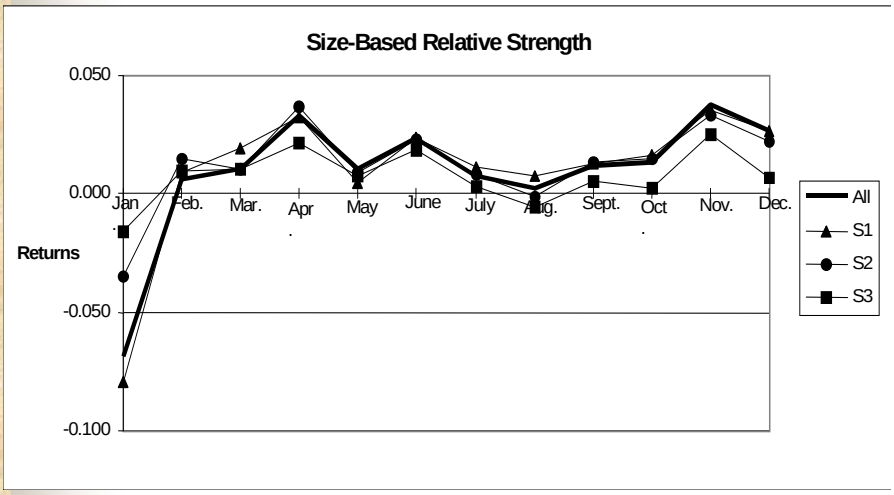
Momentum Profits Without Short-Selling

- Instead of buying high past-return, short-selling low past-return:
  - Just buy high past-return
  - Profits reduced

Momentum as an Industry Effect

- Industry momentum strategies are profitable, even after controlling for size, book-to-market, and individual stock momentum effects (Mos)
- However, this entails holding portfolios concentrated in a few "hot" industries

Avoiding January



Conclusions

- Momentum profits appear to be reliable
- Transactions costs a large drag (how much???)
  - Can reduce turnover/trade only large stocks
- Momentum profits can be enhanced by refinements (but, how feasible?)
  - Trade on industry momentum, not stock momentum
  - Avoid Januaries ????

Basic Performance Evaluation Models

(Using Net Returns Data Only)

Definitions

- Performance Evaluation: Measuring the "skill" of an asset manager
- Performance Attribution: Measuring all sources of return, including skill-based and non-skill based (luck- and risk-based)

Basic Issues in Performance Evaluation

- 1. Minimize Type I Error: separate luck from skill (difficult without a long history of returns and good model and benchmarks)
- 2. Minimize Type II Error: maximize power of test
  - Related: choose proper benchmark(s): separate controllable results from outcomes that the manager cannot control (e.g., is currency risk a choice available to the manager?)
  - Related: provide a measure of performance that can be appropriately compared with peer performance
    - Not simply using unadjusted returns!

Basic Issues in Performance Attribution

- 1. Precisely determine the source of risk-adjusted performance (market timing (asset allocation), sector timing, security selection, currency bets)
- 2. Decompose performance separately for each class of assets to increase the understanding of what went right or wrong, e.g., stocks vs. currencies
  - Also, many funds have separate managers for each asset class

The Ideal Goal of Performance Measures

- To rank managers by the accuracy of their private information on future asset returns
- The main problem: managerial risk-aversion may differ in the cross-section of managers
- A good manager who is less risk-averse may "outperform" a great manager who is more risk-averse

Two Non-Regression Approaches

- 1. Sharpe Ratio
- 2. Tracking Error

Two Ratios Based on Regressions

- 1. Treynor's Ratio
- 2. Treynor-Black Information Ratio

Various Regression Specifications

- 1. Simple, single-benchmark regression—Jensen model
- 2. Multiple-benchmark regression models, such as the Carhart 4-factor model, or the APT-based models
- 3. Conditional regressions (we'll discuss later in the week)
- 4. Regressions with both selectivity and timing regressors (we'll discuss later in the week)

Non-Regression Approaches

- Sharpe Ratio
- Tracking Error

Sharpe Ratio

- Defined:

$$SR = \frac{E[R_p] - R_F}{\sigma_p}, \quad \text{or}$$

$$Estimated\ SR = \frac{\bar{R}_p - R_F}{\hat{\sigma}_p}$$

- The Sharpe Ratio, geometrically, is the slope of the ray from the risk-free asset through the expected return of the managed portfolio
  - On a mean return/standard deviation of return plot
  - See Figure 2 of Grinblatt and Titman chapter

Benefits and Drawbacks of SR

- Benefits:
  - Extremely simple to compute
  - Provides a reward-to-variability trade-off
- Drawbacks:
  - Today's specialized funds (e.g., growth funds) may have a poor SR, but great skills (due to a good deal of diversifiable risk)

- Drawbacks (continued):
  - Ignores the agency problem inherent in asset management: a manager may have aversion to idiosyncratic risk, as he/she may be fired if poor results occur (even simply due to bad luck)
  - Thus, manager may forego an investment that increases $E[R_p]$ if it increases σ_p
  - In fact, most performance measures suffer from this agency problem

Tracking Error

- Definitions:

"Tracking-error gain" = $\bar{r}_p - \bar{r}_b$

"Tracking–error variance" = $\text{var}(r_{p,t} - r_{b,t})$, where

$r_{p,t}$ = managed portfolio's return during a given period

$r_{b,t}$ = benchmark portfolio's return during a given period

- "Tracking error" often is used to refer to either tracking-error variance or (equivalently) the standard-deviation of tracking error
- Usually, tracking-error methods involve ranking managers that have a desired level of tracking-error "gain" by their tracking-error variance—lowest is best

- We'll discuss tracking-error methods in more detail later

Simple Performance Ratios Based on Regressions

- Treynor ratio (TR)
- Information ratio (IR)

Treynor's Ratio

- Defined:

$$TR = \frac{E[R_p] - R_F}{\beta_p}, \quad \text{or}$$

$$Estimated\ TR = \frac{\bar{R}_p - R_F}{\hat{\beta}_p}$$

- The Treynor Ratio, geometrically, is the slope of the ray from the risk-free asset through the expected return of the managed portfolio
  - On a mean return/BETA plot
  - See Figure 1 of Grinblatt and Titman chapter

- Benefits:
  - Extremely simple to compute
  - Provides a reward-to-variability trade-off that is based only on market-based risk
- Drawbacks:
  - Relies on the choice of a model with a single benchmark for computation of beta, which recent work by Fama and French suggests will be misspecified

- Drawbacks (continued):
  - Ignores the agency problem inherent in asset management: a manager may have aversion to systematic risk, as he/she may be fired if poor results occur (even simply due to bad luck)
  - Thus, manager may forego an investment that increases $E[R_p]$ if it increases β_p

Treynor-Black Information Ratio

- Definition: Take, for example, the Jensen model (which we will discuss shortly):

$$R_{p,t} - R_{F,t} = \alpha_p + \beta_p (R_{B,t} - R_{F,t}) + \varepsilon_{p,t}$$

$$IR = \frac{\alpha_p}{\sigma_{\varepsilon_p}}$$

Levels of the IR

- In actual practice, an IR of 0.5 is "good" and 1.0 is "exceptional" [reported by Grinold and Kahn (1995)]
  - Only 10% of managers have an IR > 1.0

- Benefits:
  - Provides a measure of the signal-to-noise of a manager's private information on stock values
  - The IR can be computed using any type of regression-based model!

- Benefits (continued)
  - Fairly recent Grinold and Kahn book really hypes this measure, because of the following (a long proof):

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

Where IC = "information coefficient"=average correlation between a manager's forecasts and the return outcomes, and

BR = "breadth" = number of independent forecasts made by the manager

- Drawbacks:
  - Again, a manager's risk-aversion can distort this measure of performance
  - And, as mentioned this morning, the IR forces a choice of a regression model, which can having timing-related biases in alpha, as well as survival bias

Relation Between the IR and the T-Statistic of a Regression "Alpha"

- For example, assume that alpha is computed using the Jensen measure (which we'll discuss shortly):

$$R_{p,t} - R_{F,t} = \alpha_p + \beta_p (R_{B,t} - R_{F,t}) + \varepsilon_{p,t}$$

Then, We Can Note the Following:

$$IR = \frac{\alpha}{\sigma_\varepsilon}, \text{ and}$$

$$t-statistic = \frac{\alpha}{\sigma_\alpha}, \text{ thus}$$

$$\frac{IR}{t-stat} = \frac{\sigma_\alpha}{\sigma_\varepsilon}$$

Finally:

$$\sigma_\alpha \propto \frac{\sigma_\varepsilon}{\sqrt{T}}, \text{ where T =\# observations, so}$$

$$\frac{IR}{t-stat} = \frac{\sigma_\alpha}{\sigma_\varepsilon} \propto \frac{1}{\sqrt{T}}, \text{ or}$$

$$IR \propto \frac{t-stat}{\sqrt{T}}$$

Comments

- Thus, the t-statistic of the alpha can be used as a rough ranking measure across several funds, as long as the funds have an equal number of observations

Also:

- The IR seems like a sensible performance measure, as it captures:
  - The signal/noise ratio of a manager's private information
- Also, the IR seeks to summarize, in a single measure, the Mean-Variance properties of a portfolio
- So, if only returns are available, then the IR (using an appropriate model to generate alpha) is a reasonable approach, subject to the drawbacks we mentioned

Regression Specifications

- Simple Jensen model
- Multi-benchmark models
- Conditional performance models
- Models with timing and selectivity

First, We Need to Understand the Problems with All Regression-Based Models of Performance

- Market timers exhibit downward biased alphas
- Style timers—a similar problem in multifactor regressions
- Even models with timing and selectivity measures require taking a stand on the form of the manager's timing function and the correlation of timing and selectivity abilities across managers
- Regressions require a long time-series, which induces survival biases in studying individual funds
- Regressions lack power due to the need to extract factor loadings from NOISY return time-series

Univariate Regressions: Jensen's Alpha

$$R_{p,t} - R_{F,t} = \alpha_p + \beta_p (R_{B,t} - R_{F,t}) + \varepsilon_{p,t}$$

- Regress time-series of monthly mutual fund excess returns on excess return to the benchmark portfolio:
  - Benchmark is usually the value-weighted market index less T-bills

Why Has the Jensen Measure Largely Been Abandoned in Favor of Multi-benchmark Models?

- As the DFA Appendix indicates, multi-benchmark specifications of asset returns are now solidly in place (in the U.S., as well as worldwide)
- For example, it was easy to generate a large Jensen measure simply by passively holding small-cap stocks during 1975-1983
- Also, the Jensen measure is difficult to support in theory
  - Grinblatt and Titman (1989) show that it implies a quadratic utility function (decreasing marginal utility as a function of wealth)

Results for Our "Star Managers" (Jensen Alpha)

		Jensen Alpha (%/yr) (1985-2002)
Magellan		2.0
LM VT		1.6
Janus 20		1.0
Vanguard		1.0
PIMCO		3.3

Multivariate Regressions: A. Carhart's Alpha

$$R_{p,t} - R_{F,t} = \alpha_p + \beta_p RMRF_t + h_p HML_t + s_p SMB_t + y_p PR1YR + \varepsilon_{p,t}$$

- Regress time-series of monthly mutual fund excess returns on portfolio returns accruing to four zero-investment factor-mimicking portfolios:
  - High book-to-market minus low book-to-market **(HML)**
  - Small size minus big size **(SMB)**
  - High prior-year return less low prior-year return **(PR1YR)**
  - CRSP value-weighted index less T-bills **(RMRF)**

Results for Our "Star Managers" (Carhart 4-Factor Alpha)

		Carhart Alpha, %/yr (1985-2002)
Magellan		1.7
LM VT		1.4
Janus 20		3.7
Vanguard		0.6
PIMCO		Not Applicable

B. APT-Based Regression Models

- Approach: use factor-analysis or principal-components analysis to extract the "priced factors" from the universe of securities in question
- At one time, these were popular
- However, they have lost popularity, because they violate the "investable" and other characteristics of good benchmarks

Models That Separate Timing From Selectivity

- A. Treynor-Mazuy regression:

$$R_{p,t} - R_{F,t} = \alpha_p + \beta_p (R_{B,t} - R_{F,t}) + \gamma_p (R_{B,t} - R_{F,t})^2 + \varepsilon_{p,t}$$

  - Assumes the manager has a linear response of chosen beta to forecasted market return

Models That Separate Timing From Selectivity

- B. Merton and Henriksson regression:

$$R_{p,t} - R_{F,t} = \alpha_p + \beta_p (R_{B,t} - R_{F,t}) + \gamma_p \max(0, (R_{B,t} - R_{F,t})) + \varepsilon_{p,t}$$

  - Assumes the manager chooses between two betas (high or low) depending on the forecasted market return

Conditional Models of Performance

- Conditional-beta models
- Conditional-alpha-and-beta models

- The idea is that you should be able to separate the alpha that could be obtained during period t that is based on publicly available information at the end of period t-1

A. Conditional β Models

- Conditional Jensen:
  - Regressors are RMRF, z_1*RMRF, z_2*RMRF, z_3*RMRF, z_4*RMRF, z_5*RMRF, where z_1, etc. are instruments for publicly available economic information
- Conditional Carhart
  - Regressors are RMRF, SMB, HML, PR1YR, z_1*RMRF, z_2*RMRF, z_3*RMRF, z_4*RMRF, z_5*RMRF, z_1*SML, z_2*SML, etc. (24 regressors plus the intercept)

B. Conditional α and β Models

- Conditional Jensen (z_1, z_2, z_3, z_4, z_5, RMRF, z_1*RMRF, z_2*RMRF, z_3*RMRF, z_4*RMRF, z_5*RMRF)
- Conditional Carhart (z_1, z_2, z_3, z_4, z_5, RMRF, SMB, HML, PR1YR, z_1*RMRF, z_2*RMRF, z_3*RMRF, z_4*RMRF, z_5*RMRF, z_1*SML, z_2*SML, etc. (29 regressors plus the intercept)

Some Technical Points About Performance Regressions

- Heteroskedasticity (time-series) could be present.
  - Use White (1980) corrected standard errors
- Heteroskedasticity and autocorrelation (time-series) could be present
  - E.g., if using overlapping observations with heteroskedasticity
  - Use Newey-West (1983) corrected standard errors

To Conclude…

- Although the conditional regressions and maybe the Carhart regression seems to have some nice features, all regression models (and ratios computed from regression alphas) have common problems
  - Survival bias in alpha estimates, because we need a long time-series
  - Timing-related bias in alpha estimates for regressions not including a timing benchmark
  - Noise in extracting style or market loadings from returns (noisy alpha estimates)

To Conclude (continued)

- Regressions that include a market-timing benchmark generally have problems as well, as we will see later
- However, regression alphas are still, by far, the most widely used by performance evaluation services
- My research has developed security holding-based performance measures, which is where performance evaluation is headed in the future!