Towards the Design of Better Equity Benchmarks

Lionel Martellini
Professor of Finance, EDHEC Business School
Scientific Director, EDHEC Risk Institute

lionel.martellini@edhec.edu
www.edhec-risk.com
The EDHEC Risk Institute is dedicated to the production and international diffusion of academic research relevant to the investment community, at a time when the industry is affected by a number of profound paradigm shifts and when academic guidance can be of some usefulness.

The goal of this particular presentation is to provide an overview of the latest results of our research program on “Indices and Benchmarks”.

Other research programs: ALM and Asset Management; Asset Allocation and Alternative Diversification; Asset Management and Derivatives Instruments; Performance and Style Analysis; Best Execution and Operational Performance.
Outline

- Problems with existing equity indices
- Rehabilitating the tangency portfolio
- Implementation and empirical results
- Problems with existing equity indices
- Rehabilitating the tangency portfolio
- Implementation and empirical results
Problems with Existing Indices

Paradigm Changes in Asset Management

- Main paradigm changes in asset management:
  - 1980s: passive money management
  - 1990s: multi-management
  - 2000s: core-satellite investing (combination of the two).

- Whatever the paradigm, market cap weighted indices are central stage.

- A new paradigm change is under way, only accelerated by recent market conditions.

- It stems from the recognition that market cap indices are poorly diversified, inefficient, portfolios.
A benchmark is a reference portfolio that should reflect the risks taken by the portfolio manager, as well as the normal reward to be expected in exchange for given risk exposures.

This portfolio can be used ex-ante (to generate the performance) or ex-post (to measure the performance).

An index, on the other hand, is a portfolio that is supposed to be representative of a given segment of the market (e.g., large cap stocks in the US).

Standard market cap weighted indices can be good indices (?) but poor benchmarks (lack of efficiency).

Security & manager selection tasks are second order with respect to benchmark selection.
The standard practice of constructing stock market indices based on cap weighting schemes has faced severe criticism. More than 15 years ago, a number of papers (e.g., Haugen and Baker (1991) and Grinold (1992)) offered empirical evidence that market-cap weighted indices provide an inefficient risk-return trade-off.

“Cap-weighted stock portfolios are inefficient investments. [...] Even the most comprehensive cap-weighted portfolios occupy positions inside the efficient set.” (Haugen and Baker (1991))

“Market indices [...] are if anything inside that [mean-variance] frontier” (John Cochrane (2001))
Problems with Existing Indices

Inefficiency - Empirical Arguments

- Cap-weighted index lies deep inside the ex-post efficient frontier.

Based on data for the period 1979-1998. The efficient frontier assumes a perfect forecast of the future covariance matrix and of the future mean return. Figure taken from Schwartz (2000), Figure 3, page 19.
Problems with Existing Indices

Market Cap Indices are Poor Proxies for MSR Portfolio

The true GMV and tangency portfolio are functions of the (unknown) true parameter values:

\[ w_{MSR} = f(\mu_i, \sigma_i, \rho_{ij}) \]

\[ w_{GMV} = g(\sigma_i, \rho_{ij}) \]

Problems with Existing Indices

Market Cap Indices are Poor Proxies for MSR Portfolio

- Volatility
- Expected Return
- True Tangency (MSR) Portfolio
- True GMV
- Cap-weighted index

Expected Return

Volatility
Problems with Existing Indices

Inefficiency - Theoretical Arguments

- The poor risk-adjusted performance of cap-weighted indices should not come as a surprise given that the efficiency of the market portfolio is based on unrealistic assumptions:
  - Unlimited risk-free borrowing and short selling
  - Homogenous preferences, expectations and horizons
  - No frictions (taxes, transaction costs)
  - Non-tradable assets (social security claims, housing, human capital) do not exist

*Sharpe (1991) and Markowitz (2005) state that under real-world conditions the market portfolio may not be efficient.*
Problems with Existing Indices

Concentration - Effective Number of Stocks

- Cap-weighting leads to high concentration

<table>
<thead>
<tr>
<th>Index</th>
<th>Nominal number</th>
<th>Effective number</th>
</tr>
</thead>
<tbody>
<tr>
<td>S&amp;P</td>
<td>500</td>
<td>94</td>
</tr>
<tr>
<td>NASDAQ</td>
<td>100</td>
<td>37</td>
</tr>
<tr>
<td>Eurostoxx</td>
<td>300</td>
<td>98</td>
</tr>
<tr>
<td>Topix</td>
<td>500</td>
<td>94</td>
</tr>
</tbody>
</table>

\[ \tilde{n} = \frac{1}{\sum_{i=1}^{n} w_i^2} \]

\(\tilde{n}\) is the reciprocal of the Herfindhal index, a commonly used measure of portfolio concentration

\[ \tilde{n} = 1 \text{ if } \exists \ i \text{ such that } w_i = 100\% \]

\[ \tilde{n} = n \text{ if } w_i = \frac{1}{n} \quad \forall \ i = 1, \ldots, n \]

\(\Rightarrow\) hence the interpretation as effective number of stocks
Overall, commercial indices are not efficient or well-diversified portfolios because they have never meant to be efficient or well-diversified.

The main objective of these indices is to represent the stock market, thus neglecting investors’ need for the most efficient risk-return trade-off.

Alternative weighting schemes have been suggested to generate indices that better represent the economy: characteristics-based indices that weight stocks according to their economic footprint, as opposed to their market cap.

These indices focus on representativity, and do not explicitly aim at improving the risk-reward ratio (efficiency).
Various weighting schemes have been considered in an attempt to generate more efficient equity portfolios, including:

1. Equally-weighted (EW) portfolios: naïve diversification.
2. Global minimum variance (GMV) portfolios: scientific diversification.

Investing in EW portfolios amounts to giving up entirely on economics and statistics.

Investing in GMV portfolios involves a focus on risk parameter estimates; the out-of-sample Sharpe ratio of GMV benchmarks is, however, typically lower than that of EW benchmarks (DeMiguel, Garlappi and Uppal (2007)): the reduction in risk comes at the cost of a (higher-than-proportional) reduction in performance.
Problems with Existing Indices

GMV Portfolios have Low Risk ... But also Low Performance

The true GMV and tangency portfolio are functions of the (unknown) true parameter values

\[ w_{MSR} = f(\mu_i, \sigma_i, \rho_{ij}) \]

\[ w_{GMV} = g(\sigma_i, \rho_{ij}) \]

Implementable proxies depend on estimated parameter values

\[ \hat{w}_{GMV} = g(\hat{\sigma}_i, \hat{\rho}_{ij}) \]
- Problems with existing equity indices
- Rehabilitating the tangency portfolio
- Implementation and empirical results
For a rational investor, the goal is not to have the portfolio with the lowest risk or the highest representativity.

The goal is instead to obtain the best risk-adjusted performance.

In the end, if one cares for a high reward-to-risk ratio, one should aim at maximizing the reward-to-risk ratio, which requires:

- estimate of risk parameters
- estimate of expected return parameters

Are we really ready to believe that absolutely nothing meaningful can be said about expected returns?
Rehabilitating the Tangency Portfolio

Real-World Implementation of Portfolio Theory

The true GMV and tangency portfolio are functions of the (unknown) true parameter values

\[ w_{MSR} = f(\mu_i, \sigma_i, \rho_{ij}) \]

\[ w_{GMV} = g(\sigma_i, \rho_{ij}) \]

Implementable proxies depend on estimated parameter values

\[ \hat{w}_{MSR} = f(\hat{\mu}_i, \hat{\sigma}_i, \hat{\rho}_{ij}) \]

\[ \hat{w}_{GMV} = g(\hat{\sigma}_i, \hat{\rho}_{ij}) \]
To get a decent proxy for the true tangency portfolio, we need to rely on an expected return estimate of reasonable quality.

Sample-based expected return estimates, however, are of little help (Merton (1980)).

If statistics is close to useless in terms of expected return estimation, one should turn to economic intuition or (perhaps better) to common sense.

If there is a risk-return tradeoff, expected return parameters should be related to risk parameters.
Rehabilitating the Tangency Portfolio

Estimating Expected Returns

- That there is a risk-return trade-off is the most basic principle of finance.

- There are different possible risk measures.
  - CAPM beta: the problem is that the model performs poorly in practice (differences in expected returns are not well-explained by differences in betas).
  - Multifactor models: the problem is that we need to reliable estimates for factor risk premia!

- Recent literature suggests that we should take into account:
  - Not only systematic risk, but also stock-specific risk;
  - Not only covariance risk, but also extreme risk.
Rehabilitating the Tangency Portfolio

*Total Volatility: Theoretical Arguments*

- In the presence of market imperfections and/or irrational behaviour, it has been argued that stocks with high total (systematic + specific) volatility should earn higher returns.
  - Barberis and Huang (2001) study an equilibrium model of *loss-averse investors* and find that firms with higher stock-specific volatility should earn higher returns.
Rehabilitating the Tangency Portfolio

Total Volatility: Empirical Results

- Tinic and West (1986) find that adding idiosyncratic volatility to CAPM beta leads to better return prediction.

- Malkiel and Xu (2002), Spiegel and Wang (2006), Chua, Goh, and Zhang (2008), and Fu (2009) find a positive relation (firm or portfolio level) between (conditional) specific volatility and expected returns.

- See also Bali and Cakici (2008), as well as Huang, Liu, Rhee, and Zhang (2009).

- Martellini (2008) analyzes portfolio implications of these findings.
Rehabilitating the Tangency Portfolio

Beyond Volatility: Theoretical Arguments

- Investors express non-trivial preferences over higher-order moments of asset returns (see Martellini and Ziemann (2009) for a recent reference).

- As a result, total downside risk measures matter for expected returns:
  - Mitton and Vorkink (2007) argue that some investors may remain undiversified because diversification erodes skewness.
  - Such investors dislike stocks with negative skewness (and like stocks with positive skewness).
  - Thus low skewness stocks will have a low price and high expected returns (and high skewness stocks will have a high price and low expected returns).
  - Barberis and Huang (2007) generate similar findings with cumulative prospect theory preferences.
Rehabilitating the Tangency Portfolio

*Beyond Volatility: Empirical Evidence*

Evidence that stock downside risk is related to expected returns:

<table>
<thead>
<tr>
<th>Authors</th>
<th>Risk Measure</th>
<th>Relation</th>
<th>Moments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zhang (2005)</td>
<td>Skewness</td>
<td>+</td>
<td>Skew</td>
</tr>
<tr>
<td>Zhang (2005)</td>
<td>Skewness</td>
<td>+</td>
<td>Skew</td>
</tr>
<tr>
<td>Boyer, Mitton and Vorkink (2009)</td>
<td>Skewness</td>
<td>+</td>
<td>Skew</td>
</tr>
<tr>
<td>Tang and Shum (2003)</td>
<td>Skewness (but not kurtosis)</td>
<td>+</td>
<td>Skew</td>
</tr>
<tr>
<td>Connrad, Dittmar and Ghysels (2009)</td>
<td>Skewness (but not kurtosis)</td>
<td>+</td>
<td>Skew</td>
</tr>
<tr>
<td>Ang et al. (2006)</td>
<td>Downside correlation</td>
<td>+</td>
<td>Vol, Skew, Kurt</td>
</tr>
<tr>
<td>Chen et al. (2009)</td>
<td>Semi-deviation</td>
<td>+</td>
<td>Vol, Skew</td>
</tr>
</tbody>
</table>
- Problems with existing equity indices
- Rehabilitating the tangency portfolio
- Implementation and empirical results
Our objective is to go back to the basics of Modern Portfolio Theory to generate a proxy for the tangency portfolio. Such a portfolio may provide investors with a more efficient way of extracting the equity risk premium from the stock market.

We use the link between expected stock returns and total risk to estimate expected returns.

For practical reasons, we also wish to control portfolio turnover (optimal control techniques: only want to incur the turnover from updating the optimal weights if the weight changes are substantial).
Empirical Tests

Methodology – Portfolio Optimisation

- Input estimation
  - Covariance matrix: we compute the covariance matrix based on an implicit factor model (keeping a number of factors based on a conservative interpretation of statistical tests inspired by random matrix theory).
  - Expected returns: we use downside risk as a robust proxy for excess expected returns; for increased robustness, we sort stocks into decile portfolios and use the median risk of stocks in each portfolio as the expected return estimate for those stocks.

- Optimisation
  - Objective: we use these two set of inputs to obtain the weights of the portfolio with the highest Sharpe ratio.
  - Constraints: we use the following weight constraints (for an index with N constituents): min: 1/(2N), max: 2/N.
Empirical Tests

Methodology – Data

- We use weekly returns data on S&P 500 constituents (CRSP) from 01/1957 to 12/2008 to construct an out of sample backtest for the 50 year period 01/1959 to 12/2008.

- We use a calibration period of 2 years and rebalance the portfolio every three months (at the beginning of January, April, July and October).

- We assign optimal weights to all current S&P 500 constituents that have returns data for the calibration period.

- For stocks where sufficient data is not available in the calibration period, we attribute a weight that corresponds to the minimum weight constraint for stocks that enter the optimisation.
Empirical Tests

Results – Risk & Return

<table>
<thead>
<tr>
<th>Index</th>
<th>Ann. average return</th>
<th>Ann. std. Deviation</th>
<th>Sharpe Ratio</th>
<th>Information Ratio</th>
<th>Tracking Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Efficient Index</td>
<td>11.63%</td>
<td>14.65%</td>
<td>0.41</td>
<td>0.52</td>
<td>4.65%</td>
</tr>
<tr>
<td>Cap-weighted</td>
<td>9.23%</td>
<td>15.20%</td>
<td>0.24</td>
<td>0.00</td>
<td>0.00%</td>
</tr>
<tr>
<td>Difference (Efficient minus Cap-weighted)</td>
<td>2.40%</td>
<td>-0.55%</td>
<td>0.17</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>p-value for difference</td>
<td>0.14%</td>
<td>6.04%</td>
<td>0.04%</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

The table shows risk and return statistics portfolios constructed with using the same set of constituents as the cap-weighted S&P 500 index. Rebalancing is quarterly subject to an optimal control of portfolio turnover (by setting the reoptimisation threshold to 50%). Portfolios are constructed by maximising the Sharpe ratio given an expected return estimate and a covariance estimate. The expected return estimate is set to the median total risk of stocks in the same decile when sorting on total risk. The covariance matrix is estimated using an implicit factor model for stock returns. Weight constraints are set so that each stock’s weight is between 1/2N and 2/N, where N is the number of index constituents. P-values for differences are computed using the paired t-test for the average, the F-test for volatility, and a Jobson-Korkie test for the Sharpe ratio. The results are based on weekly return data from 01/1959.
Empirical Tests

Results – Turnover and Concentration

<table>
<thead>
<tr>
<th>Index</th>
<th>Annual one-way turnover</th>
<th>Excess turnover vs. Cap-weighted</th>
<th>Average Effective constituents</th>
<th>Effective constituents to nominal constituents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Efficient Index</td>
<td>23.10%</td>
<td>18.41%</td>
<td>382</td>
<td>76%</td>
</tr>
<tr>
<td>Cap-weighted</td>
<td>4.69%</td>
<td>0.00%</td>
<td>94</td>
<td>19%</td>
</tr>
</tbody>
</table>

The table shows the resulting turnover measures for Efficient Indexation portfolios that have been implemented using the controlled reoptimisation with a threshold value of 50%. The table indicates the effective number of constituents in the efficient index and in the cap-weighted index, computed as the inverse of the sum of squared constituent weights. This measure is computed at the start of each quarter and averaged over the entire period. The results are based on weekly return data from 01/1959 to 12/2008.
Empirical Tests

Results – Evolution of Wealth

- Prolonged lower returns occurred in the bull market of the late 1990s.

- This underperformance happened as the cap-weighted index returned in excess of 20% annual.

- Even in this period, efficient indexation had lower volatility than cap-weighting.
Empirical Tests

Results – Robustness Checks

Risk and return in different decades

<table>
<thead>
<tr>
<th>&quot;Decade&quot;</th>
<th>Ann. average return</th>
<th>Ann. Volatility</th>
<th>Sharpe ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cap-weighting</td>
<td>Efficient</td>
<td>Cap-weighting</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Indexation</td>
<td></td>
</tr>
<tr>
<td>1999-2008</td>
<td>-1.22%</td>
<td>3.47%</td>
<td>18.98%</td>
</tr>
<tr>
<td>1989-1998</td>
<td>19.16%</td>
<td>16.43%</td>
<td>12.84%</td>
</tr>
<tr>
<td>1979-1988</td>
<td>16.32%</td>
<td>20.82%</td>
<td>16.02%</td>
</tr>
<tr>
<td>1959-1978</td>
<td>2.96%</td>
<td>4.24%</td>
<td>16.02%</td>
</tr>
<tr>
<td>1959-1968</td>
<td>10.33%</td>
<td>14.29%</td>
<td>10.65%</td>
</tr>
</tbody>
</table>

The table shows risk and return statistics when dividing the sample in periods of ten years. The results are based on weekly return data from 01/1959 to 12/2008.

Risk and return in recessions and expansions

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cap-weighting</td>
<td>Efficient</td>
<td>Cap-weighting</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Indexation</td>
<td></td>
</tr>
<tr>
<td>Recessions</td>
<td>-1.64%</td>
<td>2.26%</td>
<td>22.85%</td>
</tr>
<tr>
<td>Expansions</td>
<td>11.19%</td>
<td>13.30%</td>
<td>13.47%</td>
</tr>
</tbody>
</table>

The table shows risk and return statistics computed for 2 subsamples. The subsamples are obtained by sorting the weekly observations based on a recession indicator for that week. The recession indicator is obtained from NBER dates for peaks and troughs of the business cycle. The results are based on weekly return data from 01/1959 to 12/2008.
Conclusion

- We extract robust expected return estimates and a robust covariance matrix estimate from the cross-section of stock returns and use these as inputs in a maximisation of the Sharpe ratio.

- Our empirical tests show that this procedure allows us to generate efficient indices with out-of-sample Sharpe ratios that are significantly higher than that of their value-weighted counterparts: economically significant increase in efficiency for investors that seek exposure to the equity risk premium.

- Performance is consistent across different time periods (also across regions – unreported results); the only case of lower risk-return efficiency occurs in the extreme bull market of the late 1990s.
Designing better equity benchmarks is only one piece of a greater puzzle related to better servicing investors’ needs.

Investor needs are typically expressed in terms of long-term consumption/liability objectives subject to short-term risk constraints.

To meet these needs, the asset management industry should focus on the following sources of added-value (with alpha only being the icing on the cake):

- Design of enhanced performance-seeking building blocks;
- Design of enhanced liability-hedging building blocks;
- Design of optimal strategies based on these building blocks.
Extensions

- Design of enhanced performance-seeking building portfolios (PSP)
  - The focus in PSP should be on maximizing risk-reward ratio.
  - PSP is constructed from investable proxies for asset class MSR portfolios.

- Design of enhanced liability-hedging building portfolios (LHP)
  - The focus in LHP should be on maximizing liability hedging potential.
  - Key challenge on enhanced inflation-hedging portfolios: selection of securities/assets for their horizon inflation-hedging properties.

- Design of optimal long-term allocation strategies
  - Accounting for horizon effects (mean-reverting interest rates & risk premia).
  - Accounting for short-term constraints & goals (risk-controlled strategies).

- When PSP has a higher risk-reward ratio, upside potential generated by LT dynamic allocation strategies is much higher.
References

- Barberis, N. and M. Huang, Stocks as lotteries: The implications of probability weighting for security prices, 2007, working paper.
Schwartz, T., 2000, How to Beat the S&P500 with Portfolio Optimization, DePaul University, working paper.
Zhang, Y., 2005, Individual Skewness and the Cross-Section of Average Stock Returns, Yale University, working paper.