CHAPTER III

METHODOLOGY AND DATA SOURCES

3.1 Methodology

Part of our objective in performing a principal components analysis (herinafter PCA) on twenty-five hedge fund strategies from two hedge fund data providers for the period 1996-2005 was to confirm and isolate at least some of the most significant market-related factors that influence returns at the strategy level and determine if they are predictable.

This thesis departs from Amenc et al (2002) in several important respects to overcome some potential issues with that research.

Specifically:

1) We derive factors from a Principal Components Analysis rather than develop separate return prediction models for each hedge fund index. This explicitly recognizes that common factors influence different hedge fund strategies and reduces the dimensionality of predicting individual strategies.

2) We lengthen the forecast window to permit less frequent reallocations. It may be that much of the active return of Amenc (2002) may come from the predictability that others have attributed to "performance smoothing" from managers trading in less-liquid securities such as convertible bonds and distressed securities. Although Amenc argued that the ability to shift assets between strategies is increasing with the creation of investible strategy indices (and thus the ability to capture these smoothing effects may also

increase) is increasing, the gains from "performance smoothing" may be unobtainable in a "real world" context..

3) Rather than start with an equally weighted benchmark of hedge fund indices, we develop a strategic asset allocation using only historical risk and return information. It may be that it is possible to improve on an equally weighted benchmark that relies only on historical information. For this reason, this may be the most appropriate benchmark for a TAA model, as otherwise the information ratio will combine the gains of both the best strategic asset allocation and any benefits that can be gained from tactical asset allocation.

3.1.1 Principal Components Analysis

We first performed a principal components analysis on data for 25 hedge fund indices from two providers to extract the most important factors.

Principal component analysis (PCA) is a classical statistical method that transforms a number of possibly correlated variables into a smaller number of linear variables that are uncorrelated with each other. These factors are called principal components. .PCA has the distinction of being the optimal linear transformation for keeping the subspace that has largest variance. The first principal component extracted accounts for as much of the variability in the data as possible, with each succeeding component accounting for as much of the remaining variability as possible.

In the context of this study, *reducing the number of relevant* variables allows concentrating on a few variables that may impact strategies differently, while retaining as much of the variability as possible.

หอสมุดกลาง สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

3.1.2 Development of Factor Prediction Equations

We retain the factor scores from the PCA and find, through stepwise linear regression, the best prediction equation for both three month and six month horizons from market-based variables identified by other researchers including equity returns, the Fama French SMB and HML factors, default and yield spreads.

These regressions yielded predictions of the factor scores.

3.1.3 Application of Arbitrage Pricing Theory

The Arbitrage Pricing Theory contends that the price of a security is driven by a number of factors. We use the predicted three-month and six month factor realizations developed in the previous section with estimates of the factor exposures for each strategy based on a PCA reestimated every three months beginning December 1998 (allowing 36 months data for both indices for the December 1998 prediction) to derive estimates of expected returns using equation (1) below. These estimates of expected returns and factor weights then entered our asset allocation optimizations.

$$\text{Predicted Re}turn_s = Intercept + B_{S1}Factor_1 + B_{S2}Factor_2 + B_{S3}Factor_3 \qquad (1)$$

where s=Strategy
and the factors are the first three factors identified by the PCA.

In this formulation and by construction, the factors have mean of 0 and standard deviation of 1. In other words, the intercept includes the return from the anticipated portion of the factors, in addition to any return not explained by the factors. This would include the impact of other factors (possibly including what Waring and Siegel (2005) characterize as "exotic beta" and alpha. Thus used, the model is in essence consistent with the Arbitrage Pricing Theory (APT).

3.1.4 Asset Allocation Optimizations

Finally, using the regression equations estimated in the previous section, we estimate factor realizations for the first three factors beginning January 1999 and every subsequent three- (for the three month model) and six- month (for the six month prediction model) period until December 2005. Using these factor realization estimates and the cumulative factor loadings for each of the CSFB Tremont and EACM indices, we estimate returns for each of the performed two sets of portfolio optimizations using three different maximum level of loadings to the factors as constraints. Furthermore, we constrained the strategy weights to be non-negative (no shorting of hedge funds) but less than 15%, and sum to 100%.

The absolute value of the factor betas for the first three factors prove to be excellent representations of the standard deviation of strategy returns, with adjusted r-squared of approximately 0.98 and an F score of 220.7.

3.2 Data and Data Sources

3.2.1 Index Level Data

We limit ourselves to hedge fund index return data since 1996, a common period for two hedge fund index providers accessed: CSFB/Tremont and Evaluation Associates. (See Table III). Although the CSFB/Tremont indices extend to 1994, we do not use the early returns here in the interest of the cross-checking of conclusions across index families using a common ten year period (seven years of which were available for building portfolios) .

Table V
Hedge Fund Index Data Providers

Index	**CSFB/Tremont**	**Evaluation Associates**
Description	**Offshore**	**Onshore**
Weighting	**Capitalization-weighted**	**Equal-weighted**
Data Series	**January 1994-June 2006**	**January 1996-December 2005**
Number of Indices	**14 (12)**	**18 (13)**

Although Table V above suggests a total of 32 indices, we eliminated two CSFB/Tremont and five Evaluation Associates indices which appear to be (and confirmed by on multiple regression) linear combinations of other indices. So the principal components analysis was performed on 25 hedge fund indices.

3.2.3 Market Variables

Similar to Amenc et al (2002), variables are chosen based on previous evidence to predict asset returns. While Table VI is somewhat expansive, we reduced the number of variables to a manageable number to make the problem more tractable.

หอสมุดกลาง สถาบันวิทยบริการ จุฬาลงกรณ์มหาวิทยาลัย

Table VI

Market Variables and Sources

Variable	Description	Source/Rationale
US Equity Market Return	S & P 500 and Small Cap Return Index	BARRA Amenc et al (2002)
3-Month Treasury Bill Rate	Short term government bond rate	Federal Reserve Amenc et al (2002) Fama (1981) Fama and Schwert (1977)
VIX	Levels and Changes in the Implied Volatility of the S & P 500 index	Chicago Board Options Exchange Amenc et al (2002)\ Schneeweis and Spurgin (1999)
Default Premium	Difference in Yield Between BAA and 10 year Constant Maturity	Federal Reserve Brandt (1999)
Fama and French Factors	Fama and French factors explaining size, value, lagged excess returns effects	Fama and French (1994)
Term Premium	Difference in annualized returns between 10-year constant maturity government bonds and 3 –month constant maturity Treasury Bill	Federal Reserve Brandt (1999)
Dividend Yield	Dividend yield paid on the NYSE Composite index over the last 12 months	Brandt (1999) Harvey (1995)
Price Earnings Ratio	Price Divided by 10 year Earnings	Schiller