A Model for Estimating Daily Hedge Fund Net Asset Value

David Lee¹

¹Affiliation not available

January 25, 2024

Abstract

This paper presents a model to use market indices to predict the daily net asset value of hedge funds that are not market observable. The model allows financial market participants to produce a daily mark-to-model for their hedge fund positions, and eventually options derived thereon. We use the historical data of indices to generate a robust estimate of the required index weighting parameters. Empirical study shows that the model makes reasonable predictions when appropriate index choices have been made and produces diagnostic information that can indicate the relative reliability of predicted daily hedge fund returns.
A Model for Estimating Daily Hedge Fund Net Asset Value

David Lee

ABSTRACT
This paper presents a model to use market indices to predict the daily net asset value of hedge funds that are not market observable. The model allows financial market participants to produce a daily mark-to-model for their hedge fund positions, and eventually options derived thereon. We use the historical data of indices to generate a robust estimate of the required index weighting parameters. Empirical study shows that the model makes reasonable predictions when appropriate index choices have been made and produces diagnostic information that can indicate the relative reliability of predicted daily hedge fund returns.

Key words: hedge fund, daily return, net asset value, cash flow, market index.

JEL Classification: E44, G21, G12, G24, G32, G33, G18, G28
Most hedge fund companies require to mark-to-model their transactions in order to maintain their trading book status. As most hedge fund businesses are involved in investments on hedge funds or baskets of hedge funds, that requires daily estimation of the net asset value (NAV) of the fund assets, which is typically only available in monthly intervals.

Hedge fund managers generally use some trading strategies that hedge long positions in one sector with long positions in another. The strategies include convertible bond arbitrage, spot-future arbitrage, currency arbitrage, large-small cap arbitrage, etc.

It appears reasonable to describe daily hedge fund performance based on “portfolios” or linear combinations of returns on indices, selected to provide a comprehensive spectrum of possible market exposures. Indices are designed to capture the direction and size of market movements and are generally accepted as good proxies for overall market behavior of a given sector.

These “portfolios” should be allowed to have positive or negative weights, depending on the style of the trading strategy of each fund. The problem therefore revolves around finding suitable weights (coefficients).

We have taken an approach that relies on the assumption that if a fund has a certain exposure to some combination of market factors on a monthly basis, that will also be the case for daily returns, and that daily exposures will reflect monthly ones. Note that we
use the term “portfolios” only in a figurative sense here, since from the mathematical point of view there is no requirement that the weights add up to 100%.

There is rich literature on hedge fund performance. Bollen et al (2021) document a decline in aggregate hedge fund performance over the past decade and test whether a set of prediction models can select subsets of individual funds that buck the trend and subsequently outperform. Harvey et al (2016) introduce a new multiple testing framework and provides historical cutoffs from the first empirical test in 1967 to 2016.

Dimmock and Gerken (2016) and Honigsberg (2019) show that various measures of misreporting decline after increases in regulation, and this could worsen observed performance. If fund managers smooth returns less intensively, then reported volatility would increase and Sharpe ratios would decrease.

Aragon and Nanda (2017) study the timeliness of hedge fund monthly performance disclosures and conclude that timely disclosure is an important consideration for hedge fund managers and investors. Barth et al (2021) estimate that the worldwide net assets under hedge fund management is larger than the most generous estimate and show that the total returns earned by funds that report to the public databases are significantly lower than the returns of funds that report only on regulatory filings.

Sullivan (2021) seeks to demystify hedge fund strategies by evaluating fund performance that can be attributed to the markets and other well-known systematic factors. Jackwerth
and Slavutskaya (2016) assess the addition of alternative assets to pension fund portfolios in terms of the total benefit derived from diversification, addition of positive skewness, and the elimination of left tails of returns.

Stafylas et al (2017) provide an integrated view of the implicit factors and statistical factor models that are largely able to explain the hedge fund return-generating process. Joenväärä et al (2019) re-examine the fundamental questions regarding hedge fund performance and find a significant association between fund-characteristics related to share restrictions as well as compensation structure and risk-adjusted returns.

Jorion and Schwarz (2019) show that truncation largely preserves backfilled returns and document that either of these backfill treatments can lead to biased empirical findings, including cross-sectional results. McLean and Pontiff (2016) study the out-of-sample and post-publication return predictability of variables shown to predict cross-sectional stock returns.

This paper presents a method to calculate the daily returns of hedge funds when only monthly data for the funds is available. The basic idea of this model is to infer inter-monthly hedge fund returns from indices that are only available on a monthly basis.

The calculation of monthly returns for indices is based on the difference between the last piece of data available for each index on two consecutive months. It is assumed that there
will be no gaps in the data stream, that is, after inception of any given fund, at least one daily value must be available on any given month.

Empirically, we use the historical hedge fund returns in a rolling 36-month window over 12-months to predict the next month returns, and then compare the theoretical returns with the actual observed returns.

Empirical results show that the model predictions match the market observables very well indicating the model provides good estimates and predictive power. The market index selection represents the “best” estimate of the systematic return of the fund in the past.

The rest of this paper is organized as follows: The model is presented in Section 1; Section 2 elaborates index selection. Empirical results are discussed in Section 3; the conclusions are given in Section 4.

1. Model

Assume that we have \( N \) known monthly hedge fund returns: \( \{x_i\} \ i=1...N \) and returns on a collection of indices: \( \{y^a_i\} \). The fundamental assumption is that there is a ‘reasonably’ reliable estimate of future (daily) hedge fund returns:

\[
x_t = \sum_a W_a y^a_t + B t + I S_t,
\]

(1)
where we have written here $B$ as a constant monthly drift term, $t$ is the time since the last known return in months (the $N$-th return corresponds to $t = 0$), and $ISt$ is a drift-free idiosyncratic piece not captured by the indices. If a model determines the weights (and $B$) accurately enough and the volatility of the idiosyncratic piece is small, then the inter-monthly hedge fund returns can be ‘predicted’ from those of the indices.

If we assume that the idiosyncratic piece is normally distributed with zero drift and constant volatility, then we can perform a fit to the observed returns for the weights using:

$$x_i = \sum_a W_a y_i^a + B \quad (2)$$

and determine the idiosyncratic volatility $\sigma$ as the volatility of the distribution of the residuals.

In a maximum likelihood scenario, this would lead to minimizing the $\chi^2$ function:

$$\chi^2 = \sum_i \frac{(x_i - \sum_a W_a y_i^a - B)^2}{\sigma^2} \quad (3)$$

with respect to the weights $W_a$ and the intercept $B$, to find:

$$B = \langle x \rangle - \sum_a W_a \langle y^a \rangle, \quad \langle x \tilde{y}^a \rangle = \sum_b W_b \langle \tilde{y}^a \tilde{y}^b \rangle \quad (4)$$
\[
\sigma^2 = \frac{N}{N - 1} \left[ \langle x^2 \rangle - 2B \langle x \rangle + B^2 + \sum_{a,b} W_a W_b \langle y^a y^b \rangle - 2 \sum_a W_a \langle x y^a \rangle + 2B \sum_a W_a \langle y^a \rangle \right],
\]

(5)

where for some quantity \( Q_i \) determined in terms of the returns \( x_i \) and \( y_i \):

\[
\langle Q \rangle = \frac{1}{N} \sum_i Q_i,
\]

(6)

and deviations are defined as, for example:

\[
\bar{x}_i = x_i - \langle x \rangle, \quad \rightarrow \quad \langle \bar{x} \rangle = 0, \quad \sigma_x^2 = \langle \bar{x}^2 \rangle.
\]

(7)

We can maximize the \( R^2 \) statistic (equivalent to minimizing \( \chi^2 \)):

\[
R^2 = 1 - \frac{\langle (x_i - \sum_a W_a y_i^a - B)^2 \rangle}{\langle \bar{x}^2 \rangle} = 1 - \frac{\sigma_x^2}{\sigma_x^2},
\]

(8)

where \( \sigma_x^2 \) is the variance of the residuals: \( x_i - \sum_a W_a y_i^a - B \). The result of either procedure is a parameter set with residuals having the smallest possible variance with respect to the original variance of the fund returns. The \( \bar{R}^2 \) or ‘adjusted \( R^2 \)’ statistic is a weighted ratio:

\[
\bar{R}^2 = 1 - \frac{\langle (x_i - \sum_a W_a y_i^a - B)^2 \rangle / \text{NDF}}{\langle \bar{x}^2 \rangle / (N - 1)}
\]

(9)

where \( N \) is the number of fund returns fit to, and the number of degrees of freedom in the fit is:

\[
\text{NDF} = N - (N_t - 1), \quad N_t = \# \text{indices in fit}.
\]

(10)
This adjusted $R^2$ penalizes fits which may have small residuals merely because a large number of indices has been fit to.

The “$t$-statistic” is determined from a fitted parameter $v$ and its standard error (uncertainty) $\Delta v$ as:

$$t_v = \frac{v}{\Delta v}. \quad (11)$$

A small $t$-statistic indicates that the parameter is not significantly far from zero to be distinguishable from zero, whereas a large one indicates a significant variable. The (2-tailed) $T$-distribution is used to turn this statistic into a probability:

$$P_t = T(|t|, \text{NDF}), \quad (12)$$

which is used as an indicator as to whether the parameter should be excluded or not.

The basic idea is that more degrees of freedom should determine all variables more accurately, which the $T$-distribution attempts to account for. Note that it assumes that the uncertainties in the variables are gaussian-distributed and uncorrelated—probably neither of these are quite the case here.

A typical use of the $t$-statistic in this context would be to examine the index weights and their standard errors and use the $T$-distribution to decide whether the index in question is significant or can be deleted from the fit. To do so requires choosing a ‘cutoff’ $P^c_t$, so that $P_t < P^c_t$ the weight is set equal to zero.
The model (https://finpricing.com/lib/EqBarrier.html) made use of this, but chose a different cutoff depending on the number of indices in the fit:

\[ 1 - P^c_t = \frac{1}{N_t} (1 - P^1_t) , \]  

(13)

where \( P^1_t \) is the cutoff when fitting to a single index.

The idea is that while the \( T - \)distribution accounts for differing numbers of degrees of freedom in the fit, we also want to ‘encourage’ the fits to use as few funds as possible, and this lowers the cutoff as the number of indices in the fit increases.

An \( F - \)statistic for a fit is defined as:

\[ f_v = \frac{\sum_i (x_i - \langle x \rangle)^2 / N_i}{\sum_i (x_i - \sum_a W_{ai} y_i^a - B)^2 / \text{NDF}} \]  

(14)

which is a weighted ratio of the fund variance and the residual variance.

Large values of \( F \) indicate that the fit represents the data fairly well since the residuals are small, whereas small values indicate that the residuals did not decrease ‘enough’ for the fit to mean anything. This again is quantified further by using the \( F - \)distribution to determine a significance:

\[ P_f = F(f_v, N_1, \text{NDF}) . \]  

(15)
The resulting probability is small for a good fit, and large for a poor one. This test is essentially an attempt to decide how successfully the fit has reduced the variance by comparing the original fund variance to that of the residuals.

In a similar way to the adjusted threshold for the t-distribution, an adjusted F-statistic can be introduced:

\[ 1 - \text{Padj}_F = (1 - P_f)^C, \quad C = \sum_{m=1}^{N_i} \frac{N_s!(N_s - m)!}{m!}, \]  

so that \( C \) is the total number of combinations of index choices with \( \leq N_i \) indices out of the set of \( N_s \) indices. This results in \( \text{Padj}_F \approx CP_F > P_F \), and if a ‘goodness of fit’ is based on this adjusted statistic, it will again bias the fits against using a large number of indices.

The second of (4) indicates where some problems with performing such fits may lie: In order to solve for the weights, the index covariance matrix must be inverted. In the case of a single index, the result is straightforward:

\[ W = \frac{\langle \tilde{x}\tilde{y} \rangle}{\langle \tilde{y}^2 \rangle} = \frac{\langle \tilde{x}\tilde{y} \rangle}{\sigma_y^2} = \rho \frac{\sigma_x}{\sigma_y}. \]  

(17)

where \( \rho \) here is the correlation between the single index and the fund (the ratio of volatilities essentially normalizes the returns distribution to have unit width).
For more than one index, small eigenvalues of the covariance matrix could easily lead to inflated weights. In the case of two indices correlated by $\rho_{12}$ and correlated to the fund through $\rho_1$ and $\rho_2$, we find that

$$
\begin{pmatrix}
W_1 \\
W_2
\end{pmatrix} = \frac{\sigma_x}{1 - \rho_{12}^2} \begin{pmatrix}
(\rho_1 - \rho_{12}\rho_2)/\sigma_1 \\
(\rho_2 - \rho_{12}\rho_1)/\sigma_2
\end{pmatrix},
$$

(18)

which leads to returns:

$$
W_1 y_1 + W_2 y_2 = \frac{\sigma_x}{1 - \rho_{12}^2} \left[ (\rho_1 - \rho_{12}\rho_2) \frac{y_1}{\sigma_1} + (\rho_2 - \rho_{12}\rho_1) \frac{y_2}{\sigma_2} \right] \sim \rho_1 + \rho_2 \frac{\sigma_x}{\sigma_1} y_1,
$$

(19)

where in the latter case we have assumed that the two returns are perfectly correlated: $y_1/\sigma_1 = y_2/\sigma_2$.

The point of this is that the large weights that result from high index correlations ($\rho_{12} \sim 1$) can nevertheless lead to reasonable/important returns, and so neglecting highly correlated indices is not necessarily a good idea. Worse, the cancellation of a small factor from the numerator and denominator of (19) will inevitably lead to numerical difficulties if not handled very carefully.

Ideally one would like to avoid strongly correlated funds altogether to avoid numerical issues, but in that case a different index should be included that in some sense ‘captures’ the difference in the two indexes.

The model partially deals with the above issues by employing a hybrid approach to determining the weights: first selecting ‘forward’ to select a small index set to act as a
‘good candidate’, then attempting to remove indices that have an insignificant contribution to the fit.

2. Index Selection

If one selecting all available indices for the model may be problematic. Because 1) the number of indices cannot exceed the number of dates for which returns are available; 2) the accuracy of forecast cannot increase (and usually decreases) when indices that are not relevant for the particular fund are included in the model; 3) inclusion of highly correlated indices in the model leads to large errors in estimation of $b$-coefficients and unstable forecasts.

We propose a forward selection approach for selecting indices from an index set. First, we calculate the correlations between the fund and all the indices. Then we choose the index with the largest correlation and determine its individual weight. Next, we calculate the residuals ($x_i - W y_i$). After that, the correlations of the remaining indices to these residuals are used to add more indices.

The procedure continues until there are a preset number of indices selected for the fund. Note that fitting to a single index using equation (4) will lead to residuals that are uncorrelated to the index, so the selection by correlation and performing the fitting are consistent.
This selection based on maximum correlation should result in residuals that are as uncorrelated as possible with the index set. If the index set contains indices that can represent the fund, the resulting candidate set of weights should contain enough information to predict the fund behaviour. There is no guarantee that: a) the indices in the index set are appropriate, b) the number of indices selected is sufficient (more than 6 may be required), c) minimizing correlation will sufficiently reduce fund drift and volatility.

The model then proceeds to examine the selected indices to remove ones with insignificant contributions:

A regression analysis is performed on the full set of indices to determine the weights to the fund. The index that has a weight with the highest \( P_t \) (12) is removed, and the regression is performed again on the smaller index set.

The adjusted F-statistic \( \text{PadjF} \) (15) is used to compute:

\[
\text{Var}(\hat{R}^2) = \frac{\hat{R}^2}{\text{Ninv}(\text{Padjf})^2},
\]

(20)

where \( \text{Ninv} \) is the inverse cumulative normal distribution (mean zero, unit width) function, effectively mapping the probability \( \text{Padjf} \) into a gaussian deviate.
This is a rather *ad hoc* quantity that combines the *F*-statistic with the quality for the fit—an *F*-significant reduction in variance increases the magnitude of the denominator, and a better (larger) adjusted $R^2$ will likewise increase the numerator.

Using this, find the further adjusted quantity:

$$
\hat{R}^2 = \frac{\bar{R}^2}{(1 + \frac{\text{Var}(R^2)}{\text{VarP}})^2}, \quad \text{VarP} = \frac{1}{N_s},
$$

(21)

where $\text{VarP}$ is the “variance of the prior distribution of the multiple correlation coefficient”.

If one fits to $N_s$ indices, then variances should scale as $\sim 1/N_s$, but it is far from clear to me given the definition of $\text{Padjf}$ under what circumstances the ‘turnover point’ is reached, that is, when $\text{Var}(\bar{R}^2) \sim \text{VarP}$. When $\text{Var}(\bar{R}^2) \leq \text{VarP}$ then $\bar{R}^2 \sim \bar{R}^2$ and the *F*-statistic has no effect.

The value of $\bar{R}^2$ for this index set is compared to the value of $\bar{R}^2$ computed from the same number of indices during forward selection, and the set with the largest value is kept. We repeat the above steps, beginning from this smaller set of indices.

Once the index selection is complete, we have 6 index sets, one with one index, one with two indices, etc., each with a value of $\bar{R}^2$. The index set with the largest value of $\bar{R}^2$ is considered the ‘best index set’ and used to predict hedge fund returns.
Once the ‘best index set’ has been chosen, the ability of the set to predict the fund returns is estimated by a ‘Jacknife’ procedure: sequentially exclude each return, performing a regression analysis for each case, and use the regression to predict what the excluded hedge fund return should be.

If we write $\Delta i = x_i - X_i$ as the resulting error ($X_i$ is the return predicted by the weights determined with return $x_i$ removed), then a measure of how well the index set does is the provided by the Jacknife procedure:

$$J^2 = 1 - \frac{\sigma_{\Delta}^2}{\sigma_x^2},$$  \hspace{1cm} (22)

which is adjusted as before to find the “Predicted $R^2$”:

$$PR^2 = \frac{J^2}{(1 + Var(\bar{R}^2)/VarP)^2}$$  \hspace{1cm} (23)

This is expected to be a better measure of the ability of the index set to predict the hedge fund returns.

3. **Empirical Study**

We use historical hedge fund returns in a rolling 36-month window over 12-months to predict the next month’s returns, and then compare the theoretical returns with the actual observed returns.
To assess the performance of the model, we use three indicators:

First, the correlation $\rho$ between the 12 actual monthly returns and the model predicted returns. A high correlation would indicate that at least to the predicted fund returns move in the correct direction.

Second, the root of the average $PR^2$ statistic ($\sqrt{PR^2}$) for the 12 returns. This is the model’s internal estimate of how well it should be able to predict the returns.

Third, a relative performance $R^2$ indicator:

$$P^2 = 1 - \frac{\text{Var}(\text{residuals})}{\text{Var}(\text{returns})},$$

(24)

which compares the spread of the residuals (observed model) to the variance of the returns themselves.

A negative result indicates that the model has performed worse than merely assuming that all returns are zero, and a result close to one that the model results are tracking the observed returns fairly well.

None of these quantities should be compared quantitatively, but all should be strongly correlated if they are good indicators of the model performance.
In Figure 1 we show these quantities for the convertible arbitrage funds provided, first for
the entire spectrum of such funds, and next for the sub-sample with only complete fund
return data in the 12-month window. We see that there are a few cases where the model
performance is rather poor, but the worst offenders vanish when those funds with missing
data are removed.

**FIGURE 1. Performance indicators—from top to bottom: ρ, R² and PR²—derived from 12
months of convertible arbitrage fund returns, using 36 months of data. The first is for the
entire sample of funds, the second only for those funds with complete historical data (48
months) in the provided testing database.**
More importantly, the three indicators are fairly well correlated, so that the ‘predicted $R^2$’ appears to reflect the ability of the model to predict returns fairly well. For example, considering only those fits with $PR^2 \geq 0.22$ selects out most of the funds on which the model performs reasonably well. Whether this statistic alone will prove sufficient will only become clear with further experience, but the figure suggests that it is perhaps not sensitive enough. Results (not shown) for equity funds were not nearly as good, but such funds are expected to be the hardest to fit.

Here we will consider two sample fits—one fairly good, one fairly poor. For the former we consider one of the funds appearing in Figure 1. From Table 1 we see that there is a very high correlation between the actual fund returns and the model, and good performance in both other indicators. Also included in this table are (in order) the behaviour of simpler models where the returns are based on a 36-month rolling average return, 12 month rolling average return, and assuming zero returns.

<table>
<thead>
<tr>
<th></th>
<th>AAG model</th>
<th>36 month mean</th>
<th>12 month mean</th>
<th>zero return</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho$</td>
<td>0.91</td>
<td>0.65</td>
<td>-0.84</td>
<td>0.00</td>
</tr>
<tr>
<td>$PR^2$</td>
<td>0.73</td>
<td>0.12</td>
<td>-0.44</td>
<td>0.64</td>
</tr>
<tr>
<td>$\sqrt{(PR^2)}$</td>
<td>0.37</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

TABLE 1. The quantitative measures of performance for the fund returns in Figure 2.

Recall that while the rolling averages may (or may not) be correlated fairly strongly with the actual fund returns, a strong correlation does not guarantee that the returns will have the correct magnitude. In contrast, the performance indicator $R^2$ captures the magnitude
only, without caring whether the predicted and actual returns are correlated or not. In this example, only the model is strongly correlated and has a large $R^2$, and from Figure 2 it does a much better job of predicting fund returns.

For this fund we also kept track of the index weighting to FIRSTCBX–there were no other indices that appeared in the model fits. We see that the weight is relatively stable, with an upward drift over time. This is likely also a sign of a good fit–the weights should not change dramatically from month to month.

**FIGURE 2.** The observed fund returns, the predicted AAG model returns, a rolling 12-month average return, and a rolling 36-month average return for the SAMACBF convertible arbitrage fund.
For a relatively poor fit, we consider the equity fund AMICEHD, which we see from Table 2 has a fairly large value of the mean predicted $R^2$ reported by the model (note that the $PR^2$ remains roughly constant for all of the 13 fits). This is bad since, as we can see from the table, the historical performance of the model is rather poor—in fact, assuming zero returns would appear to have the best performance of all. We also give the historical versus predicted returns in Figure 4.

<table>
<thead>
<tr>
<th></th>
<th>AAG model</th>
<th>36 month mean</th>
<th>12 month mean</th>
<th>zero return</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho$</td>
<td>0.40</td>
<td>-0.39</td>
<td>-0.25</td>
<td>0.00</td>
</tr>
<tr>
<td>$P^2$</td>
<td>-0.12</td>
<td>-0.08</td>
<td>-0.26</td>
<td>0.14</td>
</tr>
<tr>
<td>$\sqrt{(PR^2)}$</td>
<td>0.86</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*TABLE 2. The quantitative measures of performance for the AMICEHD fund returns in Figure 3.*
FIGURE 3. As in Figure 2 for the AMICEHD fund.

Part of the problem here is that it is an equity fund, some of which have management strategies that can cause fund returns characteristics to change rapidly. The index weights are not particularly stable, and two highly correlated funds (RAY and RUMCI) have relatively large weights (with magnitude > 1 in some cases) in the latter half of the year. This is precisely the case discussed in Equation 19.

4. Conclusion

This article presents a model for calculating daily returns and corresponding net asset value (NAV) changes of hedge funds. NAV values of hedge funds are typically available on monthly basis. The model to estimate the daily NAV for a hedge fund is based on modeling daily returns of the hedge fund as a weighted sum of returns of a combination of several market indices and factors.

Empirical study shows that the model matches the market observations very well and provides reasonable estimates of the predictive power of a fit. For each hedge fund, the Model is calibrated based on historical monthly returns of the fund and the market indices and factors representing the “best” estimate of the systematic return of the fund in the past.
Generating a robust estimate of the required index weighting parameters using historical data is a nontrivial exercise and is complicated by the need to choose appropriate indices to weight a fund to. The model appears to make reasonable predictions when an appropriate index choice has been made and produces diagnostic information that can indicate the relative reliability of predicted inter-month hedge fund returns.

To avoid numerical inaccuracies when highly correlated indices are used, would it be better to perform a principal component analysis on the index set? Then one would be fitting to the least correlated linear combination of indices, and large ‘cancelling’ weights may not appear.

References:


