Performance Attribution for Equity Portfolios

Yang Lu and David Kane

Introduction

Many portfolio managers measure performance with reference to a benchmark. The difference in return between a portfolio and its benchmark is the active return of the portfolio. Portfolio managers and their clients want to know what caused this active return. Performance attribution decomposes the active return. The two most common approaches are the Brinson-Hood-Beebower (hereafter referred to as the Brinson model) and a regression-based analysis.¹

Portfolio managers use different variations of the two models to assess the performance of their portfolios. Managers of fixed income portfolios include yield-curve movements in the model Lord (1997) while equity managers who focus on the effect of currency movements use variations of the Brinson model to incorporate “local risk premium” Singer and Karnosky (1995). In contrast, in this paper we focus on attribution models for equity portfolios without considering any currency effect.

The pa package provides tools for conducting both methods for equity portfolios. The Brinson model takes an ANOVA-type approach and decomposes the active return of any portfolio into asset allocation, stock selection, and interaction effects. The regression-based analysis utilizes estimated coefficients from a linear model to attribute active return to different factors. After describing the Brinson and regression approaches and demonstrating their use via the pa package, we show that the Brinson model is just a special case of the regression approach.

Data

We demonstrate the use of the pa package with a series of examples based on real-world data sets from MSCI Barra’s Global Equity Model II (GEM2).² MSCI Barra is a leading provider of investment decision support tools to investment institutions worldwide. According to the company:

GEM2 is the latest Barra global multi-factor equity model. It provides a foundation for investment decision support tools via a broad range of insightful analytics for developed and emerging market portfolios. The latest model version provides:

- Improved accuracy of risk forecasts and increased explanatory power.
- An intuitive structure that accommodates different investment processes in developed vs. emerging markets.
- Greater responsiveness to market dynamics.
- Comprehensive market coverage.

GEM2 leverages the decades of experience that MSCI Barra has in developing and maintaining global equity multi-factor models and indices, and offers important enhancements over GEM, which is utilized by hundreds of institutional fund managers worldwide.

The original data set contains selected attributes such as industry, size, country, and various style factors for a universe of approximately 48,000 securities on a monthly basis.

For illustrative purposes, this article uses three modified versions of the original data set, containing 3000 securities, namely year, quarter, and jan. The data frame, quarter, is a subset of year, containing the data of the first quarter. The data frame, jan, is a subset of quarter with the data from January, 2010.

> data(year)
> names(year)

[1] "barrid" "name" "return"
[4] "date" "sector" "momentum"
[7] "value" "size" "growth"
[10] "cap.usd" "yield" "country"
[13] "currency" "portfolio" "benchmark"

- barrid: security identifier by Barra.
- name: name of a security.
- return: monthly total return in trading currency.
- date: the starting date of the month to which the data belong.
- sector: consolidated sector categories based on the GICS.³
- momentum: capture sustained relative performance.
- value: capture the extent to which a stock is priced inexpensively in the market.
- size: differentiate between large and small cap companies.

¹See Brinson et al. (1986) and Grinold (2006) for more information.
²See www.msci.com and Menchero et al. (2008) for more information.
³Global Industry Classification Standard
• growth: capture stock’s growth prospects.
• cap.usd: capitalization in model base currency USD.
• yield: dividend of a security.
• country: the country in which the company is traded.
• currency: currency of exposure.
• portfolio: top 200 securities based on value scores in January are selected as portfolio holdings and are held through December 2010. This is to avoid the complexity of trading in the analyses.
• benchmark: top 1000 securities based on size each month. The benchmark is cap-weighted.

Here is a sample of rows and columns from the data frame year:

<table>
<thead>
<tr>
<th>name</th>
<th>return</th>
<th>date</th>
<th>sector</th>
<th>size</th>
</tr>
</thead>
<tbody>
<tr>
<td>44557 BLUE STAR OPPORTUNITIES CORP</td>
<td>0.0000</td>
<td>2010-01-01</td>
<td>Energy</td>
<td>0.00</td>
</tr>
<tr>
<td>25345 SEADRILL</td>
<td>-0.07905</td>
<td>2010-01-01</td>
<td>Energy</td>
<td>-0.27</td>
</tr>
<tr>
<td>264017 BUXLY PAINTS (PKR10)</td>
<td>-0.01754</td>
<td>2010-05-01</td>
<td>Materials</td>
<td>0.00</td>
</tr>
<tr>
<td>380927 CDN IMPERIAL BK OF COMMERCE</td>
<td>0.02613</td>
<td>2010-08-01</td>
<td>Financials</td>
<td>0.52</td>
</tr>
<tr>
<td>388340 CDN IMPERIAL BK OF COMMERCE</td>
<td>-0.00079</td>
<td>2010-11-01</td>
<td>Financials</td>
<td>0.55</td>
</tr>
<tr>
<td>country portfolio benchmark</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>44557 USA</td>
<td>0.0000</td>
<td></td>
<td></td>
<td>0.00</td>
</tr>
<tr>
<td>25345 NGR</td>
<td>0.0000</td>
<td></td>
<td></td>
<td>0.00427</td>
</tr>
<tr>
<td>264017 PAK</td>
<td>0.005</td>
<td></td>
<td></td>
<td>0.00000</td>
</tr>
<tr>
<td>380927 CAN</td>
<td>0.005</td>
<td></td>
<td></td>
<td>0.000012</td>
</tr>
<tr>
<td>388340 CAN</td>
<td>0.005</td>
<td></td>
<td></td>
<td>0.000012</td>
</tr>
</tbody>
</table>

The portfolio has 200 equal-weighted holdings. The row for Canadian Imperial Bank of Commerce indicates that it is one of the 200 portfolio holdings with a weight of 0.5% in 2010. Its return was 2.61% in August, and almost flat in November.

The Brinson Model

Single-Period Brinson Model

Consider an equity portfolio manager who uses the S&P 500 as the benchmark. In a given month, she outperformed the S&P by 3%. Part of that performance was due to the fact that she allocated more weight of the portfolio to certain sectors that performed well. Call this the allocation effect. Part of her outperformance was due to the fact that some of the stocks she selected did better than their sector as a whole. Call this the selection effect. The residual can then be attributed to an interaction between allocation and selection – the interaction effect. The Brinson model provides mathematical definitions for these terms and methods for calculating them.

The example above uses sector as the classification scheme when calculating the allocation effect. But the same approach can work with any other variable which places each security into one, and only one, discrete category: country, industry, and so on. In fact, a similar approach can work with continuous variables that are split into discrete ranges: the highest quintile of market cap, the second highest quintile and so forth. For generality, we will use the term “category” to describe any classification scheme which places each security into one, and only one, category.

Notations:

- \( w_i^B \) is the weight of security \( i \) in the benchmark.
- \( w_i^P \) is the weight of security \( i \) in the portfolio.
- \( W_j^B \) is the weight of category \( j \) in the benchmark.
- \( W_j^B = \sum w_i^B \), \( i \in j \).
- \( W_j^P \) is the weight of a category \( j \) in the portfolio.
- \( W_j^P = \sum w_i^P \), \( i \in j \).
- The sum of the weight \( w_i^j \), \( w_i^P \), \( W_j^B \), and \( W_j^P \) is 1, respectively.
- \( r_i \) is the return of security \( i \).
- \( R_j^B \) is the return of a category \( j \) in the benchmark.
- \( R_j^B = \sum w_i^j r_i \), \( i \in j \).
- \( R_j^P \) is the return of a category \( j \) in the portfolio.
- \( R_j^P = \sum w_i^P r_i \), \( i \in j \).

The return of a portfolio, \( R_P \), can be calculated in two ways:

- On an individual security level by summing over \( n \) stocks: \( R_P = \sum_{i=1}^n w_i^P r_i \).
- On a category level by summing over \( N \) categories: \( R_P = \sum_{j=1}^N W_j^P R_j^P \).

Similar definitions apply to the return of the benchmark, \( R_B \):

- \( R_B = \sum_{i=1}^n w_i^B r_i \).
- \( R_B = \sum_{j=1}^N W_j^B R_j^B \).
Active return of a portfolio, $R_{active}$, is a performance measure of a portfolio relative to its benchmark. The two conventional measures of active return are arithmetic and geometric. The package implements the arithmetic measure of the active return for a single-period Brinson model because an arithmetic difference is more intuitive than a ratio over a single period.

The arithmetic active return of a portfolio, $R_{active}$, is the portfolio return $R_P$ less the benchmark return $R_B$:

$$R_{active} = R_P - R_B.$$ 

Since the category allocation of the portfolio is generally different from that of the benchmark, allocation plays a role in the active return, $R_{active}$. The same applies to stock selection where assuming that the portfolio has the exact same categorical exposures as the benchmark does, equities within each category are different. This contributes to $R_{active}$ as well. Allocation effect $R_{allocation}$ and selection effect $R_{selection}$ over $N$ categories are defined as:

$$R_{allocation} = \sum_{j=1}^{N} W_j^P R_j^P - \sum_{j=1}^{N} W_j^B R_j^B,$$

$$R_{selection} = \sum_{j=1}^{N} W_j^P R_j^P - \sum_{j=1}^{N} W_j^B R_j^B.$$ 

The intuition behind the allocation effect is that a portfolio would produce different returns with different allocation schemes ($W_j^P$ vs. $W_j^B$) while having the same stock selection and thus the same return ($R_j^B$) for each category. The difference between the two returns, caused by the allocation scheme, is called the allocation effect ($R_{allocation}$). Similarly, two different returns can be produced when two portfolios have the same allocation ($W_j^B$) yet dissimilar returns due to differences in stock selection within each category ($R_j^P$ vs. $R_j^B$). This difference is the selection effect ($R_{selection}$).

Interaction effect, $R_{interaction}$, is the result of subtracting return due to allocation $R_{allocation}$ and return due to selection $R_{selection}$ from the active return $R_{active}$:

$$R_{interaction} = R_{active} - R_{allocation} - R_{selection}.$$ 

Weakness of the Brinson Model

The Brinson model allows portfolio managers to analyze the relative return of a portfolio using any attribute of a security, such as country or sector. One weakness of the model is to expand the analysis beyond two categories. As the number of categories increases, this procedure is subject to the curse of dimensionality.

Suppose an equity portfolio manager wants to find out the contributions of any two categories (for instance, country and sector) to her portfolio based on the Brinson model. She can decompose the active return into three broad terms – $R_{allocation}$, $R_{selection}$, and $R_{interaction}$. The allocation effect can be further split into country allocation effect, sector allocation effect and the product of country and sector allocation effects:

$$R_{allocation} = R_{countryallocation} + R_{sectorallocation} + R_{countryallocation} R_{sectorallocation}.$$ 

Specifically, the country allocation effect is the return caused by the difference between the actual country allocation and the benchmark country allocation while assuming the same benchmark return within each level of the category country, that is,

$$R_{countryallocation} = \sum_{j=1}^{N_1} cW_j^PC_j - \sum_{j=1}^{N_1} cW_j^BC_j,$$

where

- $cW_j^P$ and $cW_j^B$ refer to the weight of each country $j$ ($N_C$ countries in total) in the portfolio and that in the benchmark, respectively.
- $cR_j^B$ refers to the benchmark return of any country $j$.

Similarly, the sector allocation effect is the difference in return between a portfolio’s sector allocation and the benchmark’s sector allocation while having the same benchmark returns:

$$R_{sectorallocation} = \sum_{j=1}^{N_2} sW_j^PS_j - \sum_{j=1}^{N_2} sW_j^BS_j,$$

$sW_j^P$ and $sW_j^B$ refer to the weight of the sector $j$ in the portfolio and the weight of the sector $j$ in the benchmark, respectively. $sR_j^B$ is the benchmark return of any given sector $j$ of all $N_S$ sectors.

In the same vein, the return as a result of the selection effect $R_{selection}$ is the sum of country selection effect, sector selection effect, and the product of country and sector selection effects:

$$R_{selection} = R_{countryselection} + R_{sectorselection} + R_{countryselection} R_{sectorselection}$$

$$= \sum_{j=1}^{N_1} cW_j^PB_j - \sum_{j=1}^{N_1} cW_j^BC_j$$

$$+ \sum_{j=1}^{N_2} sW_j^PS_j - \sum_{j=1}^{N_2} sW_j^BS_j$$

$$+ (\sum_{j=1}^{N_1} cW_j^PC_j - \sum_{j=1}^{N_1} cW_j^BC_j)$$

$$* (\sum_{j=1}^{N_2} sW_j^PS_j - \sum_{j=1}^{N_2} sW_j^BS_j).$$

\cite{Brinson1991} proposed a framework to include two variables in the Brinson analysis.
The interaction effect, $R_{interaction}$, includes the interaction between country allocation and sector selection and that between country selection and sector allocation.

Therefore, in the case of $Q$ categories where $Q > 1$, the Brinson model becomes very complex (assume $Q \geq 3$):

$$R_{allocation} = \sum_{j=1}^{Q} R_{allocation_j} + \sum_{j=1}^{Q} \sum_{k=1}^{Q} R_{allocation_j} R_{allocation_k} + \sum_{j=1}^{Q} \sum_{k=1}^{Q} \sum_{p=1}^{Q} R_{allocation_j} R_{allocation_k} R_{allocation_p} = \ldots,$$

$$R_{selection} = \sum_{j=1}^{Q} R_{selection_j} + \sum_{j=1}^{Q} \sum_{k=1}^{Q} R_{selection_j} R_{selection_k} + \sum_{j=1}^{Q} \sum_{k=1}^{Q} \sum_{p=1}^{Q} R_{selection_j} R_{selection_k} R_{selection_p} = \ldots,$$

where $R_{allocation_j}$ is the allocation effect of any given category $j$, $j \in Q$, and $R_{selection_j}$ is the selection effect of any given category $j$, $j \in Q$. $i$, $j$, $k$ have different values.

As the number of categories grows, the numbers of terms for the allocation and the selection effects grow exponentially. $Q$ categories will result in $2^Q - 1$ terms for each of the allocation and selection effect.

Due to the interaction between allocation and selection of each of the $Q$ categories (it could be interaction between 2, 3 or even all $Q$ categories), the number of terms included in the interaction effect grows exponentially to take into all the interaction effects among all categories.

$$R_{interaction} = \sum_{j=1}^{Q} \sum_{k=1}^{Q} R_{allocation_j} R_{selection_k} + \sum_{j=1}^{Q} \sum_{k=1}^{Q} \sum_{p=1}^{Q} R_{allocation_j} R_{selection_k} R_{allocation_p} + \ldots,$$

$Q$ categories has $2^{2n} - 2^{n+1} + 1$ terms of interaction effects.

For instance, when there are 3 categories, the allocation effect and the selection effect each have $2^3 - 1 = 7$ terms. The interaction effect has $2^6 - 2^4 + 1 = 49$ terms. When there are 4 categories, $2^4 - 1 = 15$ terms have to be estimated for the allocation effect as well as the selection effect, respectively. $2^8 - 2^5 + 1 = 225$ terms have to be calculated for the interaction effect of 4 categories. This poses a significant computational challenge when a portfolio manager performs a multivariate Brinson analysis.

To some extent, the regression-based model detailed later solves the problem of multivariate attribution.
The br.single summary shows that the active return of the portfolio, in January, 2010 was 1.47%. This return can be decomposed into allocation effect (-0.14%), selection effect (1.42%), and interaction effect (0.19%).

> plot(br.single, var = "sector", type = "return")

![Figure 1: Sector Return.](image)

Figure 1 is a visual representation of the return of both the portfolio and the benchmark sector by sector in January, 2010. This plot shows that in absolute terms, Utilities performed the best with a gain of more than 5% and Consumer Discretionary, the worst performing sector, lost more than 10%. Utilities was also the sector with the highest active return in the portfolio.

Multi-Period Brinson Model

To obtain Brinson attribution on a multi-period data set, one calculates allocation, selection and interaction within each period and aggregates them across time. There are five methods for this – arithmetic, geometric, optimized linking by Menchero (2004), linking by Davies and Laker (2001), and linking by Frongello (2002). We focus on the first three methods in this paper. Arithmetic measure calculates relative performance of a portfolio and its benchmark by a difference; geometric measure does so by a ratio. Arithmetic measure is more intuitive but a well-known challenge in arithmetic attribution is that active returns do not add up over multiple periods due to geometric compounding. Geometric is able to circumvent the adding-up problem. Menchero (2004) discussed various linking algorithms to connect arithmetic return with geometric return and argued that the optimized linking algorithm is the best way to link attribution over time.

Arithmetic Attribution. The arithmetic attribution model calculates active return and contributions due to allocation, selection, and interaction in each period and sums them over multiple periods. The arithmetic active return over T periods \( R_{\text{arithmetic}} \) is expressed as:

\[
R_{\text{arithmetic}} = \sum_{t=1}^{T} R_t^{\text{active}},
\]

and \( R_t^{\text{active}} \) is the active return in a single period \( t \).

Geometric Attribution. The geometric attribution is to compound various returns over \( T \) periods where,

\[
1 + R_P = \prod_{t=1}^{T} (1 + R_P^t),
\]

\[
1 + R_B = \prod_{t=1}^{T} (1 + R_B^t),
\]

and \( R_P^t \) and \( R_B^t \) are portfolio and benchmark returns in a single period \( t \), respectively. Geometric return \( R_{\text{geometric}} \) is thus the difference between \( R_P \) and \( R_B \):

\[
R_{\text{geometric}} = R_P - R_B.
\]

Optimized Linking Algorithm. The well-known challenge faced in arithmetic attribution is that the actual active return over time is not equal to the geometric summation of single-period active returns,

\[
R_{\text{geometric}} \neq R_{\text{arithmetic}},
\]

i.e.,

\[
R_P - R_B \neq \sum_{t=1}^{T} R_t^{\text{active}}.
\]

Menchero (2004) proposed an optimized linking coefficient \( b_t^{\text{opt}} \) to link arithmetic returns of individual periods with geometric returns over time,

\[
R_P - R_B = \sum_{t=1}^{T} b_t^{\text{opt}} R_t^{\text{active}},
\]

where \( b_t^{\text{opt}} \) is the optimized linking coefficient in a single period \( t \).

The optimized linking coefficient \( b_t^{\text{opt}} \) is the summation of a natural scaling \( A \) and an adjustment \( a_t \) specific to a time period \( t \),

\[
b_t^{\text{opt}} = A + a_t,
\]

---

5See Bacon (2008) for a complete discussion of the complexity involved.
where $A$ is a coefficient for linking from the single-period to the multi-period return and $a_t$ is an adjustment to eliminate residuals$^6$.

Since active return over time $R_P - R_B$ is a summation of active return in each period adjusted to the optimized linking algorithm, the following is true:

$$R_P - R_B = \sum_{t=1}^{T} b^\text{opt}_t (R^\text{allocation}_t + R^\text{selection}_t + R^\text{interaction}_t),$$

where $R^\text{allocation}_t$, $R^\text{selection}_t$, and $R^\text{interaction}_t$ represent allocation, selection and interaction in each period $t$, respectively.

Within each period $t$, the adjusted attribution is thus expressed as

$$\hat{R}^\text{allocation}_t = b^\text{opt}_t R^\text{allocation}_t,$$
$$\hat{R}^\text{selection}_t = b^\text{opt}_t R^\text{selection}_t,$$
$$\hat{R}^\text{interaction}_t = b^\text{opt}_t R^\text{interaction}_t.$$

Therefore, across $T$ periods, active return $R_{\text{active}}$, the difference between portfolio return $R_P$ and benchmark return $R_B$, can be written as

$$R_{\text{active}} = \sum_{t=1}^{T} (\hat{R}^\text{allocation}_t + \hat{R}^\text{selection}_t + \hat{R}^\text{interaction}_t),$$

where $R_{\text{active}} = R_P - R_B$.

**Multi-Period Brinson Tools**

In practice, analyzing a single-period portfolio is meaningless as portfolio managers and their clients are more interested in the performance of a portfolio over multiple periods. To apply the Brinson model over time, we can use the function `brinson` and input a multi-period data set (for instance, `quarter`) as shown below.

```r
> data(quarter)
> br.multi <- brinson(quarter, date.var = "date",
+                    cat.var = "sector",
+                    bench.weight = "benchmark",
+                    portfolio.weight = "portfolio",
+                    ret.var = "return")
```

The object `br.multi` of class `brinsonMulti` is an example of a multi-period Brinson analysis.

```r
> exposure(br.multi, var = "size")
```

The `exposure` method on the class `br.multi` object shows the exposure of the portfolio and the benchmark based on a user-defined category. Here, it shows the exposure on `size`. We can see that the portfolio overweighted the benchmark in the lowest quintile in `size` and underweighted in the highest quintile.

```r
> returns(br.multi, type = "linking")
```

The `returns` method shows the results of the Brinson analysis applied to the data from January, 2010 through March, 2010. The optimized linking algorithm is applied here by setting the type to `linking`. The first portion of the `returns` output shows the Brinson attribution in individual periods. The second portion shows the aggregate attribution results. The portfolio formed by top 200 value securities in January had an active return of 12.7% over the first quarter of 2010. The allocation and the selection effects contributed 0.95% and 1.73% respectively; the interaction effect made a loss of 1.42%.

---

$^6$See Menchero (2000) for more information on the optimized linking coefficients.
Regression

Single-Period Regression Model

One advantage of a regression-based approach is that such analysis allows one to define their own attribution model by easily incorporating multiple variables in the regression formula. These variables can be either discrete or continuous.

Suppose a portfolio manager wants to find out how much each of the value, growth, and momentum scores of her holdings contributes to the overall performance of the portfolio. Consider the following linear regression without the intercept term based on a single-period portfolio of \( n \) securities with \( k \) different variables:

\[
\mathbf{r}_n = \mathbf{X}_{n,k} \mathbf{f}_k + \mathbf{u}_n
\]

where

- \( \mathbf{r}_n \) is a column vector of length \( n \). Each element in \( \mathbf{r}_n \) represents the return of a security in the portfolio.
- \( \mathbf{X}_{n,k} \) is an \( n \times k \) matrix. Each row represents \( k \) attributes of a security. There are \( n \) securities in the portfolio.
- \( \mathbf{f}_k \) is a column vector of length \( k \). The elements are the estimated coefficients from the regression. Each element represents the factor return of an attribute.
- \( \mathbf{u}_n \) is a column vector of length \( n \) with residuals from the regression.

In the case of this portfolio manager, suppose that she only has three holdings in her portfolio. \( \mathbf{r}_3 \) is thus a 3 by 1 matrix with returns of all her three holdings. The matrix \( \mathbf{X}_{3,3} \) records the score for each of the three factors (value, growth, and momentum) in each row. \( \mathbf{f}_3 \) contains the estimated coefficients of a regression on \( \mathbf{X}_{3,3} \).

The active exposure of each of the \( k \) variables, \( X_i \), \( i \in k \), is expressed as

\[
X_i = \mathbf{w}_{\text{active}} X_{n,i}
\]

where \( X_i \) is the value representing the active exposure of the attribute \( i \) in the portfolio, \( \mathbf{w}_{\text{active}} \) is a column vector of length \( n \) containing the active weight of every security in the portfolio, and \( x_{n,i} \) is a column vector of length \( n \) with attribute \( i \) for all securities in the portfolio. Active weight of a security is defined as the difference between the portfolio weight of the security and its benchmark weight.

Using the example mentioned above, the active exposure of the attribute value, \( X_{\text{value}} \), is the product of \( \mathbf{w}_{\text{active}} \) (containing active weight of each of the three holdings) and \( x_3 \) (containing value scores of the three holdings).

The contribution of a variable \( i \), \( R_i \), is thus the product of the factor returns for the variable \( i \), \( f_i \) and the active exposure of the variable \( i \), \( X_i \). That is,

\[
R_i = f_i X_i
\]

Continuing the example, the contribution of value is the product of \( f_{\text{value}} \) (the estimated coefficient for value from the linear regression) and \( X_{\text{value}} \) (the active exposure of value as shown above).

Therefore, the active return of the portfolio \( R_{\text{active}} \) is the sum of contributions of all \( k \) variables and the residual \( u \) (a.k.a. the interaction effect),

\[
R_{\text{active}} = \sum_{i=1}^{k} R_i + u.
\]

For instance, a hypothetical portfolio has three holdings (A, B, and C), each of which has two attributes – size and value.

<table>
<thead>
<tr>
<th>Return</th>
<th>Name</th>
<th>Size</th>
<th>Value</th>
<th>Active_Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A</td>
<td>1.2</td>
<td>3.0</td>
<td>0.5</td>
</tr>
<tr>
<td>2</td>
<td>B</td>
<td>2.0</td>
<td>2.0</td>
<td>0.1</td>
</tr>
<tr>
<td>3</td>
<td>C</td>
<td>0.8</td>
<td>1.5</td>
<td>-0.6</td>
</tr>
</tbody>
</table>

Following the procedure as mentioned, the factor returns for size and value are -0.0313 and -0.1250. The active exposure of size is 0.32 and that of value is 0.80. The active return of the portfolio is -11% which can be
decomposed into the contribution of size and that of value based on the regression model. Size contributes 1% of the negative active return of the portfolio and value causes the portfolio to lose the other 10.0%.

**Single-Period Regression Tools**

Another conventional attribution methodology is the regression-based analysis. As mentioned, the pa package provides tools to analyze both single-period and multi-period data frames.

```r
> rb.single <- regress(jan, date.var = "date", +    ret.var = "return", +    reg.var = c("sector", "growth", +                "size"), +    benchmark.weight = "benchmark", +    portfolio.weight = "portfolio")
> exposure(rb.single, var = "growth")
```

<table>
<thead>
<tr>
<th>Portfolio</th>
<th>Benchmark Diff</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>0.305</td>
</tr>
<tr>
<td></td>
<td>0.2032</td>
</tr>
<tr>
<td></td>
<td>0.1018</td>
</tr>
<tr>
<td>2</td>
<td>0.395</td>
</tr>
<tr>
<td></td>
<td>0.4225</td>
</tr>
<tr>
<td></td>
<td>-0.0275</td>
</tr>
<tr>
<td>3</td>
<td>0.095</td>
</tr>
<tr>
<td></td>
<td>0.1297</td>
</tr>
<tr>
<td></td>
<td>-0.0347</td>
</tr>
<tr>
<td>4</td>
<td>0.075</td>
</tr>
<tr>
<td></td>
<td>0.1664</td>
</tr>
<tr>
<td></td>
<td>-0.0914</td>
</tr>
<tr>
<td>High</td>
<td>0.130</td>
</tr>
<tr>
<td></td>
<td>0.0783</td>
</tr>
<tr>
<td></td>
<td>0.0517</td>
</tr>
</tbody>
</table>

reg.var specifies the columns containing variables whose contributions are to be analyzed. Calling exposure with a specified var yields information on the exposure of both the portfolio and the benchmark by that variable. If var is a continuous variable, for instance, growth, the exposure will be shown in 5 quantiles. Majority of the high value securities in the portfolio in January have relatively low growth scores.

```r
> summary(rb.single)
```

<table>
<thead>
<tr>
<th>Period: 2010-01-01</th>
</tr>
</thead>
<tbody>
<tr>
<td>Methodology: Regression</td>
</tr>
<tr>
<td>Securities in the portfolio: 200</td>
</tr>
<tr>
<td>Securities in the benchmark: 1000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Returns</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010-01-01</td>
</tr>
<tr>
<td>sector 0.003189</td>
</tr>
<tr>
<td>growth 0.000504</td>
</tr>
<tr>
<td>size 0.002905</td>
</tr>
<tr>
<td>Residual 0.008092</td>
</tr>
<tr>
<td>Portfolio Return -0.029064</td>
</tr>
<tr>
<td>Benchmark Return -0.043753</td>
</tr>
<tr>
<td>Active Return 0.014689</td>
</tr>
</tbody>
</table>

The summary method shows the number of securities in the portfolio and the benchmark, and the contribution of each input variable according to the regression-based analysis. In this case, the portfolio made a loss of 2.91% and the benchmark lost 4.38%. Therefore, the portfolio outperformed the benchmark by 1.47%. Sector, growth, and size contributed 0.32%, 0.05%, and 0.29%, respectively.

**Multi-Period Regression Model**

The same challenge of linking arithmetic and geometric returns is present in multi-period regression model. We apply the optimized linking algorithm proposed by Menchero (2000) in the regression attribution as well.

Within each period \( t \),

\[
R_{\text{active}}^i = \sum_{i=1}^k R_{i,t} + u_t,
\]

where \( R_{i,t} \) represents the contribution of a variable \( i \) of the time period \( t \) and \( u_t \) is the residual in that period. Across \( T \) periods, the active return can be expressed by a product of the optimized linking coefficient \( b_{opt}^t \) and the individual contribution of each of the \( k \) attributes. The adjusted contribution of each of the \( k \) variables \( i, \hat{R}_{i,t} \), is expressed by

\[
\hat{R}_{i,t} = b_{opt}^t R_{i,t}.
\]

Thus, the overall active return \( R_{\text{active}} \) can be decomposed into

\[
R_{\text{active}} = \sum_{t=1}^T \sum_{i=1}^k \hat{R}_{i,t} + U,
\]

where \( U \) is the residual across \( T \) periods.

**Multi-Period Regression Tools**

```r
> rb.multi <- regress(quarter, date.var = "date", +    ret.var = "return", +    reg.var = c("sector", "growth", +                "size"), +    benchmark.weight = "benchmark", +    portfolio.weight = "portfolio")
> rb.multi
```

<table>
<thead>
<tr>
<th>Period starts: 2010-01-01</th>
</tr>
</thead>
<tbody>
<tr>
<td>Period ends: 2010-03-01</td>
</tr>
<tr>
<td>Methodology: Regression</td>
</tr>
<tr>
<td>Securities in the portfolio: 200</td>
</tr>
<tr>
<td>Securities in the benchmark: 1000</td>
</tr>
</tbody>
</table>

Regression-based analysis can be applied to a multi-period data frame by calling the same method regress. By typing the name of the class object rb.multi directly, a short summary of the analysis is provided, showing the starting and ending period of the analysis, the methodology, and the average number of securities in both the portfolio and the benchmark.

```r
> summary(rb.multi)
```

<table>
<thead>
<tr>
<th>Period starts: 2010-01-01</th>
</tr>
</thead>
<tbody>
<tr>
<td>Period ends: 2010-03-01</td>
</tr>
<tr>
<td>Methodology: Regression</td>
</tr>
<tr>
<td>Avg securities in the portfolio: 200</td>
</tr>
<tr>
<td>Avg securities in the benchmark: 1000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Returns $Raw</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010-01-01 2010-02-01 2010-03-01</td>
</tr>
</tbody>
</table>
The regression-based summary shows that the contribution of each input variable in addition to the basic information on the portfolio. The summary suggests that the active return of the portfolio in year 2010 is 1.27%. The Residual number indicates the contribution of the interaction among various variables including sector, growth, and size.

Visual representation of relative performance of a portfolio against its benchmark is best viewed across a longer time span. Here, we use the data frame year for illustrative purposes.

We obtained an object rb.multi2 of class regress-Multi based on the data set from January, 2010 through December, 2010. The portfolio beat the benchmark by 10.1% over this period. Based on the regression model, size contributed to the lion share of the active return.

Figure 3: Performance Attribution.

Figure 3 displays both the cumulative portfolio and benchmark returns from January, 2010 through December, 2010. It suggests that the portfolio, consisted of high value securities in January, consistently outperformed the benchmark in 2010. Outperformance in May and June helped the overall positive active return in 2010 to a large extent.

Brinson as Regression

Another way to think about the analysis as Brinson et al. (1986) have done is to consider it in the context of a regression model. Conducting a Brinson attribution is similar to running a linear regression without the intercept term. Estimated coefficients will then be the mean return of each category of the attributed specified in the universe, a.k.a. the factor return of each category. The mean return of each category also appears in the Brinson analysis. The equivalent to the allocation effect for the universe in the Brinson model is the sum of the product of the estimated coefficient and the active weight of each category.
Using the same regression model as before,
\[
R_{\text{allocation}} = \sum_{j=1}^{N} W_j^P R_j^B - \sum_{j=1}^{N} W_j^B R_j^P = (W^P - W^B)'f + U,
\]
where \(W^P\) is a column vector indicating the portfolio weight of each category within the attributed specified by the manager; \(W^B\), a column vector indicating the benchmark weight of each category, and \(f\) is the column vector which has benchmark return of all the categories. Assuming that in this case, the benchmark is the universe and the portfolio holdings are all categories. Assuming that in this case, the benchmark weights to calculate the selection effect.

The Brinson model is a special case of the regression approach. In order to estimate the selection effect in the Brinson model can be calculated by using linear regression.

An numerical example of showing that the Brinson model can be special case of the regression approach is as follows.

Suppose that an equity portfolio manager has a portfolio named \texttt{test} with the universe as the benchmark.

We can also show the allocation effect by running a regression model based on sector only.

When we apply the standard single-period Brinson analysis, we obtain an active return of -35.9 bps which can be further decomposed into allocation (-3.4 bps), selection (-42.5 bps), and interaction (10.1 bps).
The contribution from sector based on the regression approach (-3.4 bps) matches the allocation effect from the Brinson model as shown above.

However, in order to calculate the selection effect from the regression approach, we need to apply another regression model to a universe limited to the securities held in the portfolio. Using the factor returns from the regress class object, test.reg, and those from the linear regression, we can obtain the selection effect (-42.5 bps) via the regression approach.

```r
> lm.test <- lm(return ~ sector - 1, +
   data = test[test$portfolio != 0, ])
> lm.test$coefficients

sectorEnergy sectorMaterials
-0.03561 -0.05146
sectorIndustrials sectorConDiscre
0.00194 -0.00533
sectorConStaples sectorHealthCare
-0.02514 0.04327
sectorFinancials sectorInfoTech
-0.02376 -0.02376
sectorTeleSvcs sectorUtilities
0.00916 -0.03878
```

`exposure(br.single, var = “sector”)[ ,2] %*
+ (lm.test$coefficients - test.reg$coefficients)

[,1]
[1, ] 0.00653

Conclusion

In this paper, we describe two widely-used methods for performance attribution – the Brinson model and the regression-based approach, and provide a simple collection of tools to implement these two methods in R with the pa package. We also show that the Brinson model is a special case of the regression method. A comprehensive package, portfolio Enos and Kane (2006), provides facilities to calculate exposures and returns for equity portfolios. It is possible to use the pa package based on the output from the portfolio package. Further, the flexibility of R itself allows users to extend and modify these packages to suit their own needs and/or execute their preferred attribution methodology. Before reaching that level of complexity, however, pa provides a good starting point for basic performance attribution.

Yang Lu and David Kane
yang.lu@williams.edu and
dave.kane@gmail.com

Bibliography


