

Leakage in Diversity Weighted Index Portfolios

The fundamental theorem on mathematically generated portfolios provides a relationship between changes in the value of the generating function of a portfolio and the portfolio's performance relative to the capitalization weighted market portfolio on which it is based (see equation (3.1) in R. Fernholz (1996), *Dynamic equity indices*). It follows that if the generating function is stable over time, we gain insight into the expected long-term relative performance of the generated portfolio. Diversity weighted index portfolios are mathematically generated portfolios whose generating functions measure the diversity a capitalization weighted index. The theoretical work on mathematically generated portfolios generally assumes that the market portfolio is passive with a fixed list of names, and in such a setting it can be argued that the market diversity is likely to remain fairly stable over time. However, applications will usually be based on some subset of the market such as the S&P 500 or the Russell 1000, and these indices are not passive—stocks whose capitalization has declined are systematically replaced by those whose capitalization has grown. While it can still be argued that the diversity will be stable for these subset indices, the systematic replacement of the stocks in the index calls for an adjustment in the fundamental relationship between the generating function and the portfolio performance.

Let $X_1 \geq X_2 \geq \dots \geq X_n$ represent the capitalizations of the stocks in the market in descending order. Suppose we have an index Z in which the selection of stocks depends on the stocks' capitalizations; the S&P 500 and the Russell 1000 are such indices because they are both large capitalization indices, while the S&P MidCap 400 is also because it holds mid-sized stocks. In this case we have

$$Z = X_{i_1} + \dots + X_{i_m},$$

for some selection of indices $i_1 < i_2 < \dots < i_m$, and the weights of Z are

$$\pi_{i_k} = X_{i_k}/Z,$$

for $k = 1, \dots, m$. Let \mathbf{S} be the generating function which is a measurement of diversity and hence generates a diversity weighted index portfolio, and let

$$\mathbf{S}_Z = \mathbf{S}(\pi_{i_1}, \dots, \pi_{i_m})$$

be the value of the generating function evaluated at the weights of Z .

Suppose a “unit” of time passes and we have returns r_1, \dots, r_n on the stocks. Suppose that these returns are net of dividends, so the stocks’ capitalizations are now

$$X'_i = X_i(1 + r_i),$$

for $i = 1, \dots, n$. Hence the new weights are

$$\pi'_{i_k} = X'_{i_k} / (X'_{i_1} + \dots + X'_{i_m}),$$

for $k = 1, \dots, m$, and the value of the generating function is now

$$\mathbf{S}'_Z = \mathbf{S}(\pi'_{i_1}, \dots, \pi'_{i_m}).$$

The new capitalizations X'_i of the stocks are no longer in descending order, so let us reorder them with a permutation p such that

$$X'_{p(1)} \geq X'_{p(2)} \geq \dots \geq X'_{p(n)}. \quad (1)$$

Now consider the index

$$Z' = X'_{p(i_1)} + \dots + X'_{p(i_m)}.$$

Z' is composed of the stocks which, under the new ordering (1), represent the same relative positions as the stocks in Z with the original ordering. If we let

$$\pi'_{p(i_k)} = X'_{p(i_k)} / Z',$$

then the value of the generating function \mathbf{S} for Z' is

$$\mathbf{S}_{Z'} = \mathbf{S}(\pi'_{p(i_1)}, \dots, \pi'_{p(i_m)}).$$

Suppose we have reason to believe that the generating function \mathbf{S} is likely to remain stable over time for the index Z . What we mean is that \mathbf{S}_Z and $\mathbf{S}_{Z'}$ should be about the same, not \mathbf{S}_Z and \mathbf{S}'_Z . However the theory relates the performance of the diversity

weighted portfolio to the latter values rather than the former, so we must correct for the discrepancy. The correction,

$$\lambda = \log(\mathbf{S}_{Z'}/\mathbf{S}'_Z),$$

is called *leakage*, and must be subtracted from the relative performance of the diversity weighted portfolio. With this correction, equation (3.1) cited above becomes

$$d \log(Y/Z) = d \log \mathbf{S}(\pi) - \lambda dt + \Theta(\pi)dt.$$

where Y is the total value of the diversity weighted portfolio, $d \log \mathbf{S}(\pi)$ represents the change in diversity from \mathbf{S}_Z to $\mathbf{S}_{Z'}$, and Θ is called the drift function corresponding to Y . In practice for the INTECH 500SM Index which is based on the S&P 500 with generating function \mathbf{D}_p where $p = .76$, the value of λ is about 15 basis points a year. For larger indices like the Wilshire 5000 which approximate the whole market, the leakage would be negligible.

Robert Fernholz

11/27/96