Minimum Variance ETF Portfolios and Optimal Rebalancing Frequency

Student: Jaume Bargalló Alemany
Mentor: José Suárez-Lledó
Master Thesis – Msc in Finance and Banking
UPF – Barcelona School of Management
July 2015
There is a common belief that an average of 30 stocks are needed to achieve a well-diversified portfolio. However, after analyzing data from 2008 – 2015, we show that this statement does not hold for ETFs; for this type of investment vehicle, an average of just 5 assets is enough to obtain all the benefits from diversification in terms of risk reduction. After introducing transaction costs, we find that quarterly rebalancing approaches to a break-even point where costs start to offset diversification benefits.

Introduction

One of the most important references in portfolio selection and management is the mean-variance efficiency method for optimal portfolio construction and asset allocation (Markowitz (1952)). Markowitz’s research also provided the basis for other important advances in financial economics, including the Sharpe-Lintner Capital Asset Pricing Model or the recognition of the dichotomy between systematic and diversifiable risk.

Since the publication of the first Markowitz papers, many authors have written articles in favor or against his theories. DeMiguel, Garlappi and Uppal (2009) for instance argue that the mean-variance Markowitz model does not perform better than a naïve 1/N portfolio in terms of Sharpe Ratio. Some other asset allocation models such as the one developed by Black and Litterman (1990) have tried to overcome the problems of highly-concentrated portfolios, input-sensitivity and estimation error maximization.

Although different investment strategies or approaches can be discussed, what is common among them is the fact that investors seek to maximize their return while minimizing risk. As discussed by Sharpe, Alexander, and Bailey (1999), all investors are risk averse; they prefer less risk to more for the same level of expected return. Because of this reason, this thesis will focus on risk reduction and Markowitz’s Minimum Variance Portfolio.

At the time of achieving this risk reduction objective, diversification plays an important role. In order to diversify risk measured through standard deviation, investors can allocate their budget to more than one asset so as to obtain diversification benefits.

Portfolio variance (for two assets) is given by the formula:

Portfolio Variance = \( w_A^2 \sigma^2(R_A) + w_B^2 \sigma^2(R_B) + 2w_Aw_B \text{Corr}(R_A, R_B) \sigma(R_A) \sigma(R_B) \)
Or the generalized expression for many assets:

\[ \sigma_p^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} w_i w_j \text{Cov}(R_i, R_j) \]

On the other hand, the portfolio expected return is given by:

\[ E(R) = w_1 R_1 + w_2 R_2 + \ldots + w_n R_n \]

As it can be seen, as long as correlation among two or more assets is less than 1, building a portfolio of assets provides diversification and risk reduction benefits without giving up expected returns.

At this point, it seems that the more assets are added in a portfolio, the more diversification benefits investors can obtain. However, one common belief among investors is that almost all of the benefits of diversification can be achieved with a portfolio of only 30 stocks. In fact, Fisher and Lorie (1970) showed that 95 percent of the benefits of diversification among NYSE-traded stocks were achieved with a portfolio of 32 stocks. Other studies conducted by Campbell, Lettau, Malkiel and Xu (2001) conclude that during the 1963-85 period, a portfolio of 20 stocks reduced annualized portfolio excess standard deviation to about five percent, while in the 1986-97 subsample, this level of excess standard deviation required almost 50 stocks (due to increase of stocks volatility over time).

Nevertheless, building a portfolio of an average of 30 stocks can be difficult to attain for several reasons. One of them can be monitoring and rebalancing, as new information will be received and thus weights will change. On the other hand, transaction costs can offset all those diversification benefits. Brokers usually charge two types of fees, one fixed and the other one variable. Even with a low variable fee, fixed costs will generate important transaction costs each time the 30-stock portfolio needs to be rebalanced.

In order to overcome some of these issues, Exchange-Traded Funds can be introduced.
Exchange-Traded Funds

ETFs (ETFs from now on) are investment funds traded on stock exchanges, in a way that investors can buy and sell them like stocks. Most of them track a stock or bond index, allowing investors to efficiently invest in assets that represent diversified portfolios (such as the S&P500, Eurostoxx 50, industry indices, etc.) in a much simpler way; this way, investors incur in far lower transaction costs compared to investing in all those assets that are part of an index.

On the other hand, ETFs allow a passive investment style (as opposed to active investment) where investors do not have to worry about monitoring a universe of stocks, forecasting, taking individual investment decisions, rebalancing the portfolio if they want to track a specific index, etc. Moreover, the management fees they may pay are much lower than the ones they would pay if investing through a typical mutual fund with a manager that actively manages a portfolio that tracks an index, industry, or market.

To have a glance of the importance of the ETF market, the Deutsche Bank ETF Annual Review & Outlook (January 2015) states that ETF assets reached $2,64 trillion in 2014 and are expected to pass $3 trillion in 2015, with a trading activity of $18,7 trillion and no sign of slowing down. Of this amount, around two thirds correspond to the American market. According to Bloomberg data, in 2013 the global market for Mortgage Backed Securities was around $7.5 trillion, what gives a good reference of the importance of the growing ETF market.

Many sources state that the first forms of ETF appeared in the early nineties (see Carrel, Lawrence (2008)) with some of them, SPDRs or "Spiders" reaching a considerable size, even being nowadays some of the biggest ones in the world. However, as the chart below shows (Global Exchange Traded Products regional asset growth), it has not been until 2007-2008 when these investment vehicles have started to grow at a faster path and have achieved a considerable importance in the investment industry.
The nature of these investment vehicles have allowed both particular and institutional investors to profit from them. On the one hand, particular investors have gained access to a much wider range of strategies and assets; for instance they have been able to invest in fixed income portfolios, what was difficult some years ago due to the minimum capital requirements to invest in bonds. On the other hand, institutional investors have taken advantage of being able to invest in more diversified portfolios at a lower cost and expose their portfolios to specific industries or risks in a much more efficient way. Moreover, beating indices used to be difficult due to the constant monitoring and rebalancing; nowadays, it is far easier as for instance transaction costs have been reduced considerably thanks to ETFs.

**Motivation**

As it was shown in the first section of the paper, investing through a portfolio of stocks provide risk reduction benefits versus allocating all the capital in one single asset, and the less correlated those assets are, the more diversification the portfolio manager will obtain. ETFs are themselves a (wider or narrower) selection of stocks that form a portfolio. So just because of this, theoretically they should represent a less risky asset class than any other single asset. Some of them invest in industry-tracking indices formed by dozens of stocks;
others invest in even wider universes of stocks like the S&P 500, formed by hundreds of stocks of very different industries. Others invest in portfolios of different types of bonds: corporate, government, investment grade, high yield, etc.

So, considering all this, should investors still need a portfolio of around 30 assets to have a diversified portfolio?

In relation to this question, this paper tries to provide two answers:

Question 1. How many ETFs are needed to achieve the maximum level of diversification?

Question 2. In the case of ETFs, what is the optimal rebalancing frequency considering Transaction Costs and Returns?

The first question will try to shed light to whether diversification rules for stocks are still valid for ETFs. On the other hand, the second question will analyze whether rebalancing to keep the optimal portfolio with minimum variance is always worth it in presence of transaction costs.

Data

With the aim to replicate the process that any investor (retail or institutional) could follow, the online broker platform of the ING Bank has been selected as a reference. This way, it represents the ETFs to which any investor could potentially have access to as well as allowing to model data with real transaction fees that banks and brokers are charging. This bank has been selected because it offers a good range of ETFs to invest in and charges low fees compared to other competitors (either other banks or independent brokers) so it is a good mix of how a retail and an institutional investor could operate.

In order to replicate as close as possible a retail/institutional investor operating with the ING broker, only long operations have been considered, with no short selling allowed.

So as to consider a broad universe of assets and sectors in the economy, 13 ETFs that represent different industries, geographies and asset classes have been selected. Most broker platforms have an offer of 150-200 ETFs; however, many of them track the same (or very
similar) indices, with the difference being just the provider or manager of the ETF. Because of this, for this paper it has been selected a much narrower sample of ETFs that represent fairly well all the categories of ETF available, what allows to work with an amount of data far easier to handle.

*Sectors:*

- Utilities: Lyxor Ucits Etf Stoxx Europe 600 Utilities
- Technology: Comstage Stoxx Europe 600 Tech Nr Ucits Etf
- Healthcare: Comstage Stoxx Europe 600 HC Nr Ucits Etf
- Oil & Gas: Comstage Stoxx Europe 600 O&G Nr Ucits Etf
- Banking: Lyxor Ucits Etf Stoxx Europe 600 banks
- Energy: Lyxor Ucits Etf New Energy A

*Geographic regions:*

- Dax index – European economy: Lyxor Ucits Etf DAX
- Emerging Markets: Vanguard FTSE Emerging Markets ETF
- Asia: SPDR Index Shares Funds S&P Emerging Asia Pacific
- S&P 500 – USA: iShares Core S&P 500 ETF

*Fixed Income:*

- High Yield: iShares iBoxx $ High Yield Corporate Bond ETF
- Investment Grade: iShares iBoxx $ Investment Grade Corporate Bond ETF

*Investment strategies:*

- Value: Lyxor Ucits Etf Msci Emu Value

Some ETFs could have been included in order to broaden the scope of the analysis (like “Growth” index trackers, versus the included “Value” ETF); however, data was not available
for the chosen period and reducing the time span would have probably been counterproductive for the quality of the analysis.

As explained in previous sections, ETFs’ popularity started gaining traction around 2007-2008. Because of this, data available on them are relatively limited. Although for some of the ETFs explained above data are available for periods before 2007, most of them start to have reliable data in 2008; so in order to work with comparable data, all prices and returns have been considered from October 2008 until April 2015.

All the mentioned ETFs have been checked for availability in ING and specific price and return data have been obtained from Thomson One.

**Part 1: ETFs needed to achieve the maximum level of diversification**

In order to arrive at the maximum level of diversification, several steps have been taken.

*Step 0:* for the period October 2008 – April 2015, the monthly and annual returns have been calculated for the thirteen ETFs. Then, the variance-covariance matrix for those returns has been calculated. In this part of the paper these data have been used to compute the variance for each portfolio.

*Step 1:* for each number of assets (ETFs), all the possible portfolio combinations have been generated. For example:

- Portfolios of one asset:
  
  ETF 1
  ETF 2
  ...
  ETF 13

- Portfolios of two assets:
ETF 1 & ETF 2
ETF 1 & ETF 3
...
ETF 13 & ETF 12
And so on until portfolios of thirteen assets, of which only one combination was possible.

**Step 2**: at this point, for each portfolio category (each category is a portfolio with a specific number of ETFs) and within them, for each combination of assets obtained in Step 1, the minimum possible portfolio variance has been calculated. This process has been done with the variance-covariance data calculated in Step 0 and with weights obtained through an optimizer (Excel Solver). This way, considering the variances and covariances for each asset, the target formula to minimize has been the portfolio variance subject to the restrictions that the sum of the weights has to add up to 1 and that all those weights have to be positive (no short selling allowed).

This quite computational intensive process (bear in mind that in total there are 8191 combinations of assets to optimize) lead to find the weights that minimize portfolio variance for each of the portfolio combinations of each category (number of assets).

**Step 3**: the next step is to rank all the portfolio combinations within each category so as to find the minimum variance that can be achieved with a specific number of assets.

**Results**

The table below shows the minimum variance achievable for each number of assets for the period October 2008 – April 2015 for monthly returns. It also displays the decrease in variance in percentage terms when increasing the number of ETFs by one. The chart shows the same information but graphically.
Interestingly, it seems that with just two ETFs a great amount of risk is already reduced and with five ETFs all the potential diversification benefits are achieved. This contrasts with the average thirty stocks that are needed to arrive at a level where adding further assets is not worth it, according to Statman (1987) and other authors cited in the introduction of the paper.

When repeating the same procedure but for annual returns (from January to January, February to February, etc.), the results achieved are the following:
As it can be seen, the results for annual returns are very similar to the monthly ones, with a large portion of diversification achieved with just two ETFs and no further risk reduction benefits possible with a portfolio of more than five assets.

**Portfolio composition and weights**

The following tables show the composition of the five-asset portfolios (level at which no further risk reduction is possible by adding more ETFs) for monthly and annual returns.

<table>
<thead>
<tr>
<th>5 ETF Portfolio - Monthly Returns</th>
<th>5 ETF Portfolio - Annual Returns</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tech</td>
<td>Health Care</td>
</tr>
<tr>
<td>3,9%</td>
<td>10,2%</td>
</tr>
</tbody>
</table>

In both cases the composition of the portfolio is quite similar, with the only remarkable change being Oil & Gas instead of S&P 500. Unsurprisingly, the ETF with more weight is the iShares iBoxx $ Investment Grade Corporate Bond ETF, an Exchange-Traded fund that seeks to track the investment results of an index composed of U.S. dollar-denominated, investment grade corporate bonds (the Markit iBoxx USD Liquid Investment Grade Index, both from iShares by Blackrock).

The below table shows the variance-covariance matrix for the thirteen ETFs computed with monthly returns (the one for annual returns has an equivalent interpretation). As it can be seen, the Investment Grade ETF has the lowest variance among all the ETFs. Not only this, but it also has the lowest average covariance with the other assets, even having a negative covariance with the Oil & Gas ETF. It is because of this that its weight in the minimum variance portfolios is so high for both monthly and annual returns (above 70%).
Principal Component Analysis and Portfolio Optimization

For an N asset universe, PCA provides with N uncorrelated “eigen-portfolios” and tells us how much of the universe’s variance they each explain. Portfolio Optimization gives the portfolio with the highest risk adjusted return (Sharpe Ratio), scaled to the investor’s desired risk. It turns out that the optimal portfolio from an optimization allocates risk to each of the “eigen-portfolios” in proportion to their Sharpe Ratio.

Markowitz Portfolio Optimization

Informally, the standard Markowitz mean-variance optimization problem is as follows:

**Inputs**

N assets, each with a known expected return $E[ri]=\alpha_i$ and returns volatility $vol[ri]=\sigma_i$

Known covariances between asset returns, $cov(ri,rj)=\sigma_{i,j}$

Covariance and variances are combined to form the covariance matrix $\Sigma$, where $\Sigma_{i,j}=\sigma_{i,j}$

**Output**

The portfolio optimization finds a weights vector $\vec{w}$ which maximizes risk adjusted return (Sharpe Ratio, or excess expected return divided by expected volatility), and sizes the portfolio according to the investor’s “risk aversion.”
The solution, the optimal portfolio, is \( w = k\Sigma^{-1}\alpha \), where \( \Sigma^{-1} \) is the matrix inverse of the covariance matrix, and \( k \) is a constant that depends on investor’s "risk aversion" and determines how much risk she wants to take in this portfolio.

Note that the expected return of a portfolio given a weights vector and expected returns vector is \( \vec{\omega}'\vec{\alpha} \). The volatility of that portfolio given the covariance matrix is \( \sqrt{w'\Sigma w} \). Therefore the Sharpe Ratio is \( \frac{\vec{\omega}'\vec{\alpha}}{\sqrt{w'\Sigma w}} \).

**Principal Component Analysis**

Given a covariance matrix \( \Sigma \), PCA is simply an eigenvalue decomposition of that matrix. It gives \( N \) eigenvalues \( \lambda_i \) which each correspond to an eigenvector \( \vec{v}_i \). If the columns of matrix \( V \) are the eigenvectors \( \vec{v}_i \), and the diagonal matrix \( \Lambda \) has all the eigenvalues \( \lambda \) on its diagonal (in the same order as the eigenvectors in \( V \)), then the covariance matrix: \( \Sigma = V\Lambda V' \) has been decomposed.

The eigenvectors each represent a portfolio, a set of weights on the \( N \) assets. The eigenvalues represent the variances of those eigenportfolios (this follows from the fact that \( w'\Sigma w \) is the variance of a portfolio, and \( A\vec{x} = \lambda\vec{x} \) is the property of any eigenvector \( x \) with eigenvalue \( \lambda \) of matrix \( A \). All of the eigenportfolios are uncorrelated.

Principal Component Analysis decomposes the universe of assets into \( N \) uncorrelated "risk factors," with a ranking in order of "importance" (as measured by the risk of that factor, i.e. the eigenvalue of that eigenportfolio). The first risk factor is often a portfolio that is long every asset - it represents the "market." Other portfolios might represent things like the outperformance of a given industry. Or a portfolio might represent the outperformance of one asset over another very highly correlated asset.

**Relationship between PCA and Portfolio Optimization**

Recall that the solution to the optimization problem is \( \vec{w} = k\Sigma^{-1}\vec{\alpha} \). When all assets are uncorrelated, this reduces to \( w_i = k\frac{E[\alpha]}{\text{Var}[\alpha]} = k\frac{E[r_i]}{\sigma^2[r_i]} \). This implies \( \sigma[r_i]w_i = k\frac{E[r_i]}{\sigma[r_i]} \). \( w \) represents the portion of the total portfolio allocated to an asset. When the weight in an asset \( i \) is multiplied by the volatility of that asset, it represents the volatility of a portfolio with \( w_i \) allocated to asset \( i \). It’s the amount of volatility that the position is expected to have, ignoring
the rest of the portfolio. So what the above equation means in words is that the notional weight \(w\) times volatility, or the amount of volatility expected in a given asset, is proportional to the expected Sharpe Ratio of that asset.

PCA provides with \(N\) uncorrelated portfolios (the eigenvectors \(\vec{v}\)). The expected return and risk of any arbitrary portfolio can be figured out, as we showed above. The optimization problem has been now reduced to one that has been already solved. Just pretend that the \(N\) eigenportfolios are \(N\) uncorrelated assets, and now the solution to any optimization problem becomes: the amount of volatility targeted in a given eigen-portfolio is proportional to the expected Sharpe ratio of that portfolio. When those portfolios are added up at those proper weights the optimal solution is found.

**Technical Derivation**

Recalling that \(\vec{w} = k\Sigma^{-1}\vec{a}\), the covariance matrix can be rewritten in terms of its eigenvalues and eigenvectors:

\[
k\Sigma^{-1}\vec{a} = k(\Lambda\Lambda^{-1})\vec{a} = k(\Lambda^{-1}\Lambda')\vec{a}
\]

Where \(\Lambda^{-1}\) is simply a diagonal matrix, with diagonals \(\frac{1}{\lambda_i}\).

Reading this expression \(k(\Lambda^{-1}\Lambda')\vec{a}\) from right to left:

a) \(\Lambda^{-1}\Lambda'\vec{a}\) is simply a vector, where the \(i\)'th element represents the expected return of the \(i\)'th eigen-portfolio:

\[
\vec{\gamma} = \vec{\gamma}_i = \Lambda^{-1}\Lambda'\vec{a}_i
\]

b) Pre-multiplying that vector by the inverse eigenvalue matrix yields a vector whose \(i\)'th element represents the expected returns of the \(i\)'th eigenportfolio divided by the variance of that portfolio (recall that the eigenvalues represent the variances of the eigenportfolios):

\[
\Lambda^{-1}\Lambda'\vec{a} = \vec{\gamma}_i = \xi_i = \frac{\gamma_i}{\lambda_i}
\]

c) Pre-multiplying that vector by the eigenvector matrix \(V\) is simply doing a weighted sum of your eigenvectors, where the weight on each eigenvector \(i\) is the \(i\)'th element of the zeta matrix from (b).

\[
V\Lambda^{-1}\Lambda'\vec{a} = \vec{\zeta} = \sum_i^N \vec{v}_i \xi_i = \sum_i^N \vec{v}_i \frac{\vec{\gamma}_i}{\lambda_i}
\]
**PCA: practical application**

When carrying out a Principal Component Analysis with the returns used for this paper (for the whole period October 2008 – April 2015), the results are in line with what was shown in previous sections:

```
.pca var2- var14
```

<table>
<thead>
<tr>
<th>Component</th>
<th>Eigenvalue</th>
<th>Difference</th>
<th>Proportion</th>
<th>Cumulative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comp1</td>
<td>7.94174</td>
<td>4.97269</td>
<td>0.6109</td>
<td>0.6109</td>
</tr>
<tr>
<td>Comp2</td>
<td>2.96905</td>
<td>1.58919</td>
<td>0.2284</td>
<td>0.8393</td>
</tr>
<tr>
<td>Comp3</td>
<td>1.37986</td>
<td>1.1563</td>
<td>0.1061</td>
<td>0.9454</td>
</tr>
<tr>
<td>Comp4</td>
<td>0.22555</td>
<td>0.448095</td>
<td>0.0172</td>
<td>0.9626</td>
</tr>
<tr>
<td>Comp5</td>
<td>0.178745</td>
<td>0.0516545</td>
<td>0.0137</td>
<td>0.9764</td>
</tr>
<tr>
<td>Comp6</td>
<td>0.127091</td>
<td>0.0368716</td>
<td>0.0098</td>
<td>0.9862</td>
</tr>
<tr>
<td>Comp7</td>
<td>0.0902189</td>
<td>0.054301</td>
<td>0.0069</td>
<td>0.9931</td>
</tr>
<tr>
<td>Comp8</td>
<td>0.0359179</td>
<td>0.00866485</td>
<td>0.0028</td>
<td>0.9959</td>
</tr>
<tr>
<td>Comp9</td>
<td>0.0272531</td>
<td>0.0185218</td>
<td>0.0021</td>
<td>0.9980</td>
</tr>
<tr>
<td>Comp10</td>
<td>0.00873128</td>
<td>0.00138648</td>
<td>0.0007</td>
<td>0.9986</td>
</tr>
<tr>
<td>Comp11</td>
<td>0.0073448</td>
<td>0.00100356</td>
<td>0.0006</td>
<td>0.9992</td>
</tr>
<tr>
<td>Comp12</td>
<td>0.00634124</td>
<td>0.00218039</td>
<td>0.0005</td>
<td>0.9997</td>
</tr>
<tr>
<td>Comp13</td>
<td>0.00416086</td>
<td>.</td>
<td>0.0003</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

As it can be seen the number of eigenvectors with minimum variance ranges between 4 and 7. So PCA confirms the results obtained in the first part of the paper that a number of ETFs around 5 minimizes the portfolio variance.

**Transaction Costs**

One of the most obvious advantages of having to invest just five assets (ETFs) instead of an average of thirty (stocks) is the lower transaction costs implied in the rebalancing. As new information is received, the portfolio needs to be rebalanced so as to reflect the new views and try to achieve the minimum variance possible (if it is the objective, as it is in this paper). At that point, the costs of buying and selling five ETFs will probably be lower than the ones related to correcting weights in a thirty-stock portfolio.
This leads to the second part of the paper, where an optimal portfolio rebalancing frequency is sought such that transaction costs do not overweight diversification benefits.

**Part 2: Optimal rebalancing frequency considering transaction costs and returns**

The first part of the paper shows the portfolio that would have achieved minimum variance during the period October 2008 – April 2015. In this second part the time span is divided in shorter periods so as to introduce rebalancing.

Three different rebalancing frequencies have been considered to analyze how transaction costs evolve and how they interact with returns: annual, semiannual and quarterly. So as to have enough meaningful data, the minimum time frame considered from which information can be extracted is five years.

So starting in October 2008, moving time frames with five years of data have been used to calculate the target portfolio that minimizes variance. As time goes by and new information is incorporated, the time span moves and a new target portfolio is computed; with this new information, along with how the portfolio has evolved (and its weights) during the period, the rebalancing is carried out in order to assign the new target weights.

As an example, the following table shows the process followed for the semiannual time span. After using five years of data (October 2008 – October 2013) to find the portfolio that minimizes variance, the portfolio is rebalanced in April 2014 with information obtained in the time period April 2009 – April 2014. The same process is repeated until April 2015, when the data was gathered.

For all those moving time slots, the minimum variance portfolio has been calculated for each category (number of assets) following the same methodology used in part 1 of the paper:

**Step 0:** calculate returns and variance-covariance matrix for each period

**Step 1:** for each categories (number of ETFs), generate all the possible portfolio combinations
Step 2: for each category and within it, for each combination of assets, calculate the minimum possible portfolio variance finding the optimal weights with an optimizer.

Step 3: rank all the portfolio combinations within each category so as to find the minimum variance that can be achieved with a specific number of assets.

Then the fourth step in this part has been to select, for each time span, the minimum number of ETFs that can achieve a portfolio with variance that cannot be further lowered.

Interestingly, although that number of ETFs is similar to the one obtained in the first part of the paper, it presents slight differences depending on the selected period. The following table shows, for each period, the portfolio that minimizes variance and adding another asset doesn’t add diversification benefits:

<table>
<thead>
<tr>
<th>Period</th>
<th>Assets and Weigths</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>10/2008 - 10/2013</td>
<td>Inv. Grade 68,3%, Utilities 7,3%, Health Care 15,0%, Oil &amp; Gas 9,4%</td>
<td>0,0004577</td>
</tr>
<tr>
<td>1/2009 - 1/2014</td>
<td>Inv. Grade 92,0%, Utilities 5,4%, Tech 1,7%, Oil &amp; Gas 0,9%</td>
<td>0,00033216</td>
</tr>
<tr>
<td>4/2009 - 4/2014</td>
<td>Inv. Grade 84,0%, Utilities 6,9%, Tech 4,8%, Health Care 1,9%, New Energy 1,9%, S&amp;P500 High Yield 0,4%</td>
<td>0,00024523</td>
</tr>
<tr>
<td>7/2009 - 7/2014</td>
<td>Inv. Grade 81,0%, Utilities 5,4%, Tech 2,4%, Health Care 4,4%, New Energy 0,6%, S&amp;P500 4,1%, High Yield 2,0%</td>
<td>0,00021447</td>
</tr>
<tr>
<td>10/2009 - 10/2014</td>
<td>Inv. Grade 80,2%, Utilities 7,4%, Tech 4,0%, Health Care 5,1%, New Energy 0,9%, S&amp;P500 2,4%</td>
<td>0,00021399</td>
</tr>
<tr>
<td>1/2010 - 1/2015</td>
<td>Inv. Grade 83,6%, Utilities 4,1%, Health Care 1,4%, S&amp;P500 11,0%</td>
<td>0,00022456</td>
</tr>
<tr>
<td>4/2010 - 4/2015</td>
<td>Inv. Grade 82,3%, Utilities 3,5%, Tech 1,2%, Health Care 2,0%, S&amp;P500 11,0%</td>
<td>0,00022476</td>
</tr>
</tbody>
</table>

Although the average number of ETFs needed to build a minimum variance portfolio is still five, the exact number fluctuates depending on the specific time frame used to compute results. From a minimum of four assets in three of the periods considered, it jumps until seven ETFs in one period, even including a small percentage of the High Yield Bond ETF which is not known for having a low variance.
Similar to the first five-ETF portfolio computed for the whole period, these more specific portfolios all contain a large proportion of the Investment Grade ETF (Fixed Income is considered to have a lower risk than Equity stocks and within this category, Investment Grade is safer than High Yield) and a smaller percentage of the Utilities ETF (a quite stable and mature sector with low volatility).

Two other common ETFS in the portfolios are Tech and Health Care, which appear in almost all the combinations; as it can be seen the global variance-covariance matrix shown in page 11, Health Care has a good combination of average variance and below average covariance with the other assets. In the case of the Tech ETF, although it doesn’t have a very low variance, its covariance with other assets and especially with the Investment Grade ETF is quite low.

Other ETFs with high variance and not very low covariances such as the one representing the Asian geographic region or the Value investment strategy are not included in any portfolio since they do not contribute to minimizing portfolio variance. However, if the aim of the portfolios wasn’t minimizing variance but maximizing return or Sharpe Ratio (Sharpe 1966), these ETF could potentially have a more important role, even more important than for instance the Investment Grade ETF which offers low risk but at the expense of giving up return.

**Transaction Costs calculation**

Banks and brokers charge different type brokerage fees. On the one hand, fixed fees are usually charged when investing in specific exchange markets (Spanish, US, UK, etc.). On the other hand, variable fees are charged depending on the amount traded; their most usual forms of being expressed are either as a percentage of the amount invested or as a spread in the bid-ask prices. Other forms of expenses include custody fees and depending on the broker agent they can be a fixed amount or a percentage of the assets.

In order to simplify the analysis and taking into account that variable fees represent the largest cost when a minimum capital is reached, they are the ones that have been considered as
transaction costs for the calculations. In the case of ING Bank broker, it is currently charging an average of 0.2% of the amount traded (as of the date of this paper).

There is no complete agreement in research on how to incorporate transaction costs to the portfolio optimizing problems. Empirical literature generally finds transaction costs to be convex (e.g., Engle, Ferstenberg, and Russell (2008)), with some researchers actually estimating quadratic trading costs (e.g., Breen, Hodrick, and Korajczyk (2002)). In the case of this paper, the impact of transaction costs on returns has been computed as the percentage fee (0.2%) times the changes in the portfolio needed for the required rebalancing.

In order to know the transaction cost in percentage terms that these fees represent, the percentage of the portfolio to be rebalanced each period needs to be computed; then, this 0.2% can be applied to know the cost that the rebalancing transaction has generated. As a matter of example, and using again the semiannual frequency of rebalancing, the transaction costs are calculated in the following way:

<table>
<thead>
<tr>
<th>Inv. Grade</th>
<th>Utilities</th>
<th>Health Care</th>
<th>Oil &amp; Gas</th>
</tr>
</thead>
<tbody>
<tr>
<td>10/2013</td>
<td>68.3%</td>
<td>7.3%</td>
<td>15.0%</td>
</tr>
<tr>
<td>Ret. on that period</td>
<td>2.6%</td>
<td>11.3%</td>
<td></td>
</tr>
<tr>
<td>New weight 4/2014</td>
<td>66.6%</td>
<td>7.7%</td>
<td>16.0%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Inv. Grade</th>
<th>Utilities</th>
<th>Health Care</th>
<th>New Energy</th>
<th>S&amp;P500</th>
<th>Oil &amp; Gas</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rebalancing to</td>
<td>84.0%</td>
<td>6.9%</td>
<td>4.8%</td>
<td>1.9%</td>
<td>0.4%</td>
</tr>
<tr>
<td>Change</td>
<td>17.4%</td>
<td>-0.7%</td>
<td>4.8%</td>
<td>-14.1%</td>
<td>1.9%</td>
</tr>
<tr>
<td>Transaction Cost</td>
<td>0.10%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

First the portfolio with the minimum number of assets possible to achieve minimum variance is built in October 2013 (using five year data since October 2008). In the following half year, each ETF has its specific return that leads the weights of the portfolio to deviate from the initial target. As of April 2014, new information has been received (the five years’ time span moves from October 2018 – October 2013 to April 2009 – April 2014) so a new portfolio is the target to achieve; this represents changing the weights of some assets, buying new ETFs (such as New Energy in this case) or removing some of them from the portfolio (not in this case). Because selling an asset generates the same transaction costs than buying it (they do not compensate), all changes are considered in absolute value so as to arrive at the percentage
of the total portfolio that needs rebalancing. Then this amount is multiplied by 0.2% to arrive at the final Transaction Cost to be considered (0.10% in the above example).

Transaction Costs vs. Returns

When the transaction costs for all the rebalancing frequencies have been calculated, they are compared to the returns in those periods to have an objective reference of the importance they have. The results are summarized in the following table:

<table>
<thead>
<tr>
<th>Annual</th>
<th>Nº of rebal</th>
<th>Date</th>
<th>Return of the period</th>
<th>Trans. Costs</th>
<th>TC as % of returns</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10/2014</td>
<td>7,2%</td>
<td>0,09%</td>
<td>1,2%</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Semiannual</th>
<th>Nº of rebal</th>
<th>Date</th>
<th>Return of the period</th>
<th>Trans. Costs</th>
<th>TC as % of returns</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4/2014</td>
<td>5,2%</td>
<td>0,10%</td>
<td>1,9%</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>10/2014</td>
<td>1,7%</td>
<td>0,02%</td>
<td>1,3%</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>4/2015</td>
<td>3,5%</td>
<td>0,05%</td>
<td>1,4%</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Quarterly</th>
<th>Nº of rebal</th>
<th>Date</th>
<th>Return of the period</th>
<th>Trans. Costs</th>
<th>TC as % of returns</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1/2014</td>
<td>0,8%</td>
<td>0,10%</td>
<td>12,6%</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>4/2014</td>
<td>2,1%</td>
<td>0,03%</td>
<td>1,6%</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>7/2014</td>
<td>1,1%</td>
<td>0,03%</td>
<td>2,9%</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>10/2014</td>
<td>1,0%</td>
<td>0,02%</td>
<td>1,8%</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1/2015</td>
<td>4,5%</td>
<td>0,05%</td>
<td>1,1%</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>4/2015</td>
<td>0,1%</td>
<td>0,01%</td>
<td>10,0%</td>
<td></td>
</tr>
</tbody>
</table>

* Return of the period makes reference to the period up to the rebalancing date. i.e.: return of the first semiannual period (5.2%) is obtained from 10/2013-4/2014 (rebalancing date)

For annual and semiannual rebalancing frequencies, transaction costs do not seem to offset returns for the respective periods. In the case of quarterly, however, transaction costs start to be more important, as for instance in number 1 and 6 of rebalancing they represent 10% or more of the returns for the period.
When the transaction costs are annualized, it can be seen that they increase at a fast pace when rebalancing increases its frequency. Although increasing the frequency of rebalancing decreases the percentage of the portfolio to be rebalanced (with shorter the time spans, less information is added and returns have not modified weights so much so the target portfolio needs less changes to be achieved), it is not enough to prevent transaction costs to start increasing at a faster path.

Two more important factors need to be taken into account when analyzing these results. The first one is the flat 0.2% fee considered; ING Bank is offering a quite competitive brokerage fee, especially for retail investors. Not only this, but the considered fees are the lowest in the last ten years. This means that if average brokerage fees among the universe of investors as well as the higher fees some years ago had been considered, the transaction costs would have been even more significant. Moreover, other banks do not charge a single amount for all their ETFs but instead they have different fees depending on the asset.

The above arguments relate to the second important consideration, which is that the calculations only take into account variable costs. If all the other costs were considered, transaction costs would also increase notably.

As a sensitivity analysis, increasing transaction costs to 0.5% to reflect the above mentioned factors, the annualized results would be:

<table>
<thead>
<tr>
<th>Annual TC</th>
<th>Increase</th>
</tr>
</thead>
<tbody>
<tr>
<td>Annual</td>
<td>0.09%</td>
</tr>
<tr>
<td>Semiannual</td>
<td>0.11%</td>
</tr>
<tr>
<td>Quarterly</td>
<td>0.16%</td>
</tr>
<tr>
<td></td>
<td>30.4%</td>
</tr>
<tr>
<td></td>
<td>45.5%</td>
</tr>
</tbody>
</table>

In this case transaction costs represent a much higher percentage of returns, being as high as 31.4% of the returns in some periods of the quarterly rebalancing.

One important metric that investors look at when investing in a fund, is the quartile or decile that it occupies. Although past returns do not guarantee future returns, being consistently in a particular decile/quartile is a very good performance sign (Soe (2014)).

Data for global mutual funds for the categories of long equity and fixed income have the decile distribution shown in the following table; deciles consider annualized 5 year returns from June 2010 until June 2015.
As it can be seen, transaction costs can play an important role when including funds in a specific decile. In the case of the base case considered in the paper (brokerage fees of 0.2%), transaction costs in quarterly rebalancing can represent up to a 13% of the differences among deciles. In the case of 0.5% fees, this amount can reach a 32%, what means that transaction costs can represent around one third of the differences among some percentiles, what gives a glance of how important transaction costs can be in investment decisions.

### Conclusions

The first issue that this paper wanted to address was the number of Exchange-Traded Funds that are needed to build a completely diversified portfolio. After challenging the assumption that, like stocks, around thirty ETFs were needed to reduce portfolio variance to a minimum, the results seem clear.

For the analyzed period (October 2008 – April 2015), only five ETFs were needed to achieve a portfolio with minimum variance (considering both monthly and annual returns). When the study used shorter time frames and different moments in time, the results were quite similar, with a number of ETFs needed to build a minimum variance portfolio that ranges between four and seven.

These results have several implications. On the one hand, it allows investors (both retail and institutional) to lower not only rebalancing and transaction costs but also allocate less resources to research and monitoring. On the other hand, and together with the fact that the ETFs used for the analysis are a sample that represents broad sectors of the economy,
markets, economic regions and types of assets, it confirms that this type of investment vehicle is far more diversified by itself than a single stock.

Other research on this matter could go in the direction of studying whether these results also hold with less diversified ETFs, or with a portfolio of only equity ETFs without including fixed income ETFs. It also would be interesting to study how these results interact with the fact that assets seem to all go down and increase their correlation during turmoil periods; so, do the results from this paper still hold in downturn periods? And what about purely growth periods? Is the number of ETFs needed to build a completely diversified portfolio still an average of five during those specific periods? As it was seen in the paper, five is the average, but depending on the chosen period the range goes from four to seven ETFs.

The second central question that the thesis wanted to answer was the optimal rebalancing frequency that doesn’t allow for transaction costs to offset diversification benefits. After analyzing annual, semiannual and quarterly rebalancing, results have shown that it is in quarterly frequencies when transaction costs start to have a considerable weight. At this frequency, transaction costs have a notable size considering the returns achievable during the corresponding period.

As the results have shown, rebalancing the portfolio with new target weights every quarter generates transaction costs that can significantly offset returns to the extent of potentially leading a mutual fund to lower its decile in comparison to the rest of mutual funds universe.

Further research on this issue could go one step further and analyze the impact that more frequent rebalancing (monthly, weekly, etc.) could have on the returns of the portfolio. As it has been seen, a portfolio needs less total changes with more frequent rebalancing but this is offset by the transaction costs increasing at a faster path. So probably, and considering that quarterly rebalancing approaches to a break-even point, monthly or weekly rebalancing will present more costs than diversification benefits, at least when portfolios are optimized for minimum variance.

Other issues that could be addressed in future research could be answering similar questions but focusing not on minimum variance but on other measures such as the Sharpe ratio. Giving
more weight to returns when calculating optimal portfolios would probably lead to different outcomes to the ones achieved in this paper.

**Bibliography**


Carrel, L. (2008), *ETFs for the Long Run*, John Wiley & Sons


