On Diversification

November 15, 2012

Ben Jacobsen and Frans de Roon

Abstract

Undiversified - or stock picking - portfolios may dominate well diversified benchmarks, when these benchmarks are not mean-variance efficient. Starting from Markowitz's Modern Portfolio Theory we derive simple (linear regression) tests to separate stock picking from diversification. Over 60% of the time we cannot reject our null hypothesis of stock picking in favor of well diversified benchmarks, even for individual stocks. Stock picking dominates during recessions, diversification during expansions. 'Stockpicking'-stocks tend to be stocks of large size companies, stocks with high B/M, high E/P or Momentum stocks. Our new tests also explicitly relate diversification and return predictability.

Key words: Diversification, Stock Picking, Modern Portfolio Theory, Return Predictability.

Ben Jacobsen, Massey University and New Zealand Institute for Advanced Study, New Zealand (b.jacobsen@massey.ac.nz). Frans de Roon, Tilburg University, The Netherlands (f.a.de.roon@uvt.nl).

Acknowledgements: The authors would like to thank Henk Berkman, Frank de Jong, John Stroomer and Christine Vos for detailed comments. This paper has also benefitted from comments made by participants at seminars at Massey University, Auckland and UNSW, Sydney. Of course the usual disclaimer applies.

"The expected return variance of return rule, on the other hand, implies diversification for a wide range of μ_i,σ_{ij}. This does not mean that the E-V rule never implies the superiority of an undiversified portfolio. It is conceivable that one security might have an extremely higher yield and lower variance than all other securities; so much so that one particular undiversified portfolio would give maximum E and minimum V." Markowitz, Portfolio Selection, 1952, pg. 89.

Introduction

Markowitz's quote raises an interesting point. When are undiversified portfolios superior? With mean-variance efficient portfolios diversification is optimal. But mean-variance efficient portfolios can only be determined with hindsight. In reality, both practitioners and academics rely on well diversified benchmark portfolios and market indices like the S&P 500. We consider a very simple practical problem. Given expectations about an asset and a well diversified benchmark when would an investor choose to invest fully (or more) in the asset and not in (or even short) the benchmark? Starting from the Modern Portfolio Theory we derive simple conditions and linear regression tests that allow us to test a null hypothesis of stock picking (which we define as fully investing in a small undiversified portfolio of one or only a few stocks or assets) against the alternative hypothesis of diversification (which we define as holding a broad, diversified - but not mean-variance efficient - benchmark portfolio).

When we take our tests to the CRSP data, we often cannot reject stock picking in favor of diversification. For instance, based on historical five year moving averages - as proxy for future expectations - the hypothesis that investors would prefer even individual stocks over a well-diversified value weighted benchmark cannot be rejected roughly 30% of the time based on point estimates. When we replace historical means with analyst expectations, 40% of individual stocks dominate a value weighted benchmark. Results favor stock picking even more if we take estimation uncertainty into account. For instance, 70% of the time we cannot reject the null hypothesis that individual stocks dominate a value weighted benchmark based on a conservative ten percent level. Not surprisingly, with portfolios of a limited number of stocks (or more commonly used confidence levels) these percentages increase even further. While our results vary - as we show - with different assumptions regarding benchmarks and portfolio formation, the overall conclusion based on the empirical evidence is that surprisingly often it is difficult to reject stock picking in favor of holding a diversified benchmark portfolio.

We can interpret these results as new tests for mean-variance efficiency of well-diversified benchmarks. Clearly, if a benchmark portfolio is mean-variance efficient, rejecting the hypothesis that it is optimal to be fully invested in one or a only a few individual stocks is a small hurdle to take. Yet, our empirical results show that many commonly used benchmarks cannot overcome this hurdle, indicating that they are far from being mean-variance efficient. A slightly more daring interpretation might be that these results offer little evidence for the benefits of diversification over stock picking. Or in other words, that benchmarks that we use for diversification puposes are mean variance inefficient to the extent that stock picking becomes a viable alternative. Of course this assumes that the power of our tests is sufficient. However, our tests bear a close resemblance to the standard Jensen's regression thus 'lack of power'-concerns do not seem a major issue. Moreover, our results do not support diversification strongly even if we rely on the point estimates and analyst forecasts only and simply compare stocks and diversified benchmarks deriving optimal portfolio weights from the Modern Portfolio Theory. A more aggressive interpretation of these results is that investors might as well focus on stock picking rather than investing in a diversified benchmark portfolio. At least, our results do suggest that stock picking as an alternative to diversification may deserve more attention.

As our results open the door to stock picking as a possible alternative to diversification, we derive some simple rules of thumb which may help investors in deciding whether or not they should diversify. For instance, if the Sharpe ratio of the average stock is double the Sharpe ratio of the S&P 500, an investor should prefer that stock over the S&P 500. In the case of multiple stocks an investor who expects the S&P 500 to go up with ten percent, but believes that four stocks will outperform the market and generate returns of twenty percent or more, should not diversify but hold an equally weighted portfolio of those four stocks. The 'Stockpicking'- versus 'Diversification'-decision also fluctuates over time. Stock picking works well during recessions, diversification is a better strategy during expansions (based on NBER business cycle data). The Fama and French data allows us to test what characterizes typical 'Stockpicking'- and 'Diversification'-stocks. 'Stockpicking'-stocks tend to be stocks of large size companies, with high 'Book to Market'-ratios, high 'Earnings Price'-ratios or Momentum stocks. Stocks with low 'Book to Market'-ratios, low 'Earnings Price'-ratios or low 'Cash Flow'-ratios showing short term reversals tend to be better suited for diversification.

Except for Size, our 'Stockpicking'-stocks tend to be similar to stocks known to outperform the market. Indeed, while our theoretical results offer a new perspective on diversification, they have a strong link with tests for return predictability. We find an exact (inverse) relation between Jensen's alpha and the new 'Stockpicking'-alpha we derive from our regression tests. If one holds an undiversified portfolio it generally implies a high Jensen's alpha relative to the benchmark. But, as we will show, the degree in which stock picking and return predictability are the same strongly depends on the benchmark used.

Since the introduction of the Modern Portfolio Theory in 1952, academia has paid little attention to stock picking. Brealey and Meyers' quote in their textbook[1] summarizes the general perception quite clearly: "Wise investors don't put all their eggs into just one basket: They reduce their risk by diversification." For generations of finance students and investors the Modern Portfolio Theory implies that diversification reduces risk and therefore holding a portfolio of a larger number of assets is always optimal.[2] In fact, this view on diversification has been so strong that the Modern Portfolio Theory is one of the few economic theories which implications have obtained a stronghold in law: "A trustee shall diversify the investments of the trust (...)" as the Uniform Prudent Investor Act puts it bluntly.[3]

This does not mean there are no opposing views. Boyle, Garlappi, Uppal and Wang (2012) refer to a quote by Keynes: "The right method in investment is to put fairly large sums into enterprises which one thinks one knows something about .... It is a mistake to think that one limits one's risk by spreading too much between enterprises about which one knows little and has no reason for special confidence. ... One's knowledge and experience are definitely limited and there are seldom more than two or three enterprises at any given time in which I personally feel myself entitled to put full confidence." And Keynes is not the only one. Many individual investors only hold one or two stocks and almost ninety percent of all investors hold less than ten stocks. This finding already dates back to Blume and Friend (1975) and has been considered a puzzle ever since. Similarly, studies in the last years have found little evidence of diversification. Many retail investors and even professional investors like Warren Buffett hold underdiversified portfolios (see for instance, Kelly, 1995 and Goetzmann and Kumar, 2008). This has led researchers to suggest that investors are not rational or may have higher moment preferences (Mitton and Vorkink, 2007). Our results suggest that investor expectations do not need to be exceptional to prefer stock picking over diversification. We assume that investors extrapolate past stock

[1] Brealey and Myers, Principles of Corporate Finance, 5th international edition (1996) pg. 160.

[2] Two other examples: Kelly (1995) summarizes the widely held academic view for instance as follows: "If finance text books are to be believed, investors hold an equally weighted portfolio of twenty or more stocks. By doing so they eliminate the idiosyncratic risk of individual stocks and face only undiversifiable risk, equal to the average covariance of returns of the portfolio. Such mean-variance efficiency represents one of the most basic forms of economic rationality and is the basis of asset return models such as CAPM and APT." Goetzmann and Kumar (2008) state: "(...) As a consequence, most rational models of investor choice suggest that investors hold diversified portfolios to reduce or eliminate non-compensated risk".

[3] This holds for modern trust law in general. As Sterk (2010) points out modern trust law (the Restatement (Third) of Trusts, the Uniform Prudent Investor Act and the Uniform Trust Code) has implemented the modern portfolio in a number of ways. One of these ways is that it has imposed on trustees a duty to diversify.

(market) behavior or that they might have expectations similar to analyst expectations. Greenwood and Shleifer (2012) provide evidence that measures of investor expectations tend to be extrapolative and analyst expectations are frequently used in the literature (for instance Brav, Lehavy and Michaely, 2005). For our empirical study and tests it is not relevant what drives these expectations. This might simply be a result of tastes as in Fama and French (2007), an informational advantage as in Nieuwerburgh and Veldkamp (2009) or ambiguity and ambiguity aversion as in Boyle, Garlappi, Uppal and Wang (2012). While their study is theoretical in nature Boyle, Garlappi, Uppal and Wang (2012) comes closest to our paper in terms of the topic we are trying to address. Boyle, Garlappi, Uppal and Wang (2012) also start from the Modern Portfolio Theory and develop a model to see whether, when and how stock picking could be a viable alternative.

Our paper contributes to the literature in several ways. Firstly, our results show that the Modern Portfolio Theory, contrary to popular belief, does not imply that holding well-diversified benchmarks is always optimal. When portfolios are not mean-variance efficient the Modern Portfolio Theory covers the full spectrum from diversification to holding just one asset. This seemingly ignored aspect of the Modern Portfolio Theory offers - as we will show - a wealth of new perspectives on diversification. Secondly, we derive simple regression tests to distinguish between these different options. These cannot only be used to test for stock picking versus diversification, as we do here, but might also be used in measuring the relative mean variance efficiency of benchmarks. Clearly, one would expect close to mean variance efficient benchmarks should generate a low degree of stock picking. Thirdly, we offer practically useful rules of thumb to assess whether and when stock picking or diversification is a better strategy and we document characteristics that set 'Diversification'- and 'Stock picking'-stocks apart. Our simulations on portfolios with a limited number of stocks may offer investors guidance in their decision to diversify or to focus on stock picking. Fourthly, to the best of our knowledge we are the first to show an exact link between the optimal level of diversification and return predictability. This is particularly interesting as the evidence that financial markets are to some extent predictable is growing (for instance, Rapach and Zhou, 2012). And last but not least, our results can explain why so many investors hold underdiversified portfolios and why it may be prudent to do so.

1 Diversification or Stock picking

1.1 Deriving the conditions

We are interested in the case where diversification is not optimal, i.e., where an investor would prefer to invest in a single asset or a small (undiversified) portfolio of assets only, rather than a broad diversified benchmark portfolio. The single asset or small portfolio has expected excess return $E\left[r_{A,t}\right]$,about which the investor may have high expectations, more informed beliefs, or is better able to predict returns. whereas the diversified or benchmark portfolio has expected excess return $E\left[r_{B,t}\right]$. If we know expected (possibly idiosyncratic) returns, standard devations of both the asset and the benchmark and the correlation between the returns of the two, the necessary condition can be easily derived from the Modern Portfolio Theory with two risky assets (see Appendix A). In that case imposing that the optimal weight in the asset should be larger than or equal to one hundred percent leads to:

$$\frac{E\left[r_{B,t}\right]}{\sigma_B} < \rho_{AB}\frac{E\left[r_{A,t}\right]}{\sigma_A} \tag{1}$$

Or, in terms of Sharpe ratios Sh, $Sh_B < \rho_{AB} Sh_A \alpha_B$. This condition says that - for $\rho_{AB} \geq 0$ - in order for a full investment in the single stock to be optimal, the Sharpe ratio Sh_A should exceed Sh_B and the difference should be bigger, the lower the correlation between the single stock and the benchmark. If the correlation is negative, the Sharpe ratio Sh_A should be smaller than the Sharpe ratio Sh_B. Thus, given expectations about future returns, standard deviations and the correlation between asset and benchmark an investor can determine whether to invest fully in the asset and leave the benchmark alone. In case an investor feels comfortable estimating the necessary parameters from the past we get:

$$\begin{aligned}
& Sh_B < \rho_{AB} Sh_A \alpha_B && (2)\\
\Leftrightarrow\quad & \frac{E\left[r_{B,t}\right]}{\sigma_B} < \frac{\sigma_{AB}}{\sigma_A^2}\frac{E\left[r_{A,t}\right]}{\sigma_B} && (3)\\
\Leftrightarrow\quad & E\left[r_{B,t}\right] < \beta_{BA} E\left[r_{A,t}\right] && (4)\\
\Leftrightarrow\quad & \alpha_B < 0 && (5)
\end{aligned}$$

where α_B is nothing else than a constant in the regression of the benchmark on the asset:

$$r_{B,t} = \alpha_B + \beta_{BA} r_{A,t} + \varepsilon_{B,t}. \tag{6}$$

and we can test the null hypothesis $H_0 : \alpha_B \leq 0$. We will refer to α_B as the 'Stock-picking'-alpha. The regression in (6) at first sight may seem unusual. Rather than comparing a single asset to a diversified benchmark, as in a Jensen-regression, we measure the performance of the benchmark to a single asset. However, this regression gives us precisely the information we need, as $\alpha_B > 0$ would imply it is not optimal for the investor to be fully invested in the single asset, whereas $\alpha_B \leq 0$ implies it would be optimal to hold the single asset only. In fact, a strict inequality $\alpha_B < 0$ would imply (s)he would even want to go short in the benchmark and use the proceeds to invest more than 100 percent in the single asset A (provided $\beta_{BA} > 0$). While one might consider investors who would hold substantial weights in one or a few stocks but also invest partially in a benchmark, stock pickers, these are not considered under our null hypothesis. In that sense our tests err on the side of caution.

For clarity, it may be good to compare our regression to the standard Jensen regression. If the benchmark portfolio would be mean-variance efficient, i.e., if it would be optimal for the investor to be fully invested in this diversified portfolio, than standard spanning tests[4] imply that Jensen's alpha of the asset A with respect to the benchmark would be zero. Thus, using the standard Jensen regression

$$r_{A,t} = \alpha_A + \beta_{AB} r_{B,t} + \varepsilon_{A,t}, \tag{7}$$

mean-variance efficiency of the benchmark implies $\alpha_A = 0$. Furthermore, it is well known that $\alpha_A > 0$ implies it is optimal for a mean-variance investor to take a long position in asset A and invest less in the benchmark B (if $\beta_A > 0$), whereas $\alpha_A < 0$ would imply that the investor wants to go short in asset A and use the proceeds to invest even more in the benchmark B.[5]

Our null-hypothesis refers to the opposite situation where it is optimal for the investor to be fully invested in asset A and not to invest anything in the benchmark B. Using the above results from mean-variance spanning tests and ruling out short selling, this means we can derive the reverse Jensen regression:

$$r_{B,t} = \alpha_B + \beta_{BA} r_{A,t} + \varepsilon_{B,t}, \tag{8}$$

If $r_{B,t}$ would indeed be mean-variance efficient, then rejecting this null-hypothesis should be a small hurdle to take especially because contrary to Jensen's alpha which is zero under the null, our stock picking alpha α_B should be positive as it is given by:

$$\alpha_B = E\left[r_{B,t}\right] - \frac{\sigma_{AB}}{\sigma_A^2} E\left[r_{A,t}\right] \tag{9}$$

Next suppose we want to look at the situation where a group of assets dominate the benchmark. Denote the excess return on the benchmark again as $r_{B,t}$ and let the excess returns on a subset of K assets that we think may

[4]See, e.g., Huberman & Kandel (1987)

[5]This follows directly by noting that for an investor with risk aversion γ, the optimal portfolio weight in the asset A is $w_A = \gamma^{-1} \frac{\alpha_A}{\sigma_{\varepsilon A}^2}$.

be superior, be denoted by $r_{A,t} = \begin{pmatrix} r_{1,t} & r_{2,t} & ... & r_{K,t} \end{pmatrix}$. In this case we may be interested in two different hypotheses: 1) that *each* of the K assets individually are superior to the benchmark, or 2) that a (efficient) *portfolio* of the K assets is superior to the benchmark. The first hypothesis can be tested by using K separate simple regressions:

$$\begin{aligned} r_{B,t} &= \alpha_1 + \beta_{B1} r_{1,t} + \varepsilon_{1,t} \\ r_{B,t} &= \alpha_2 + \beta_{B2} r_{2,t} + \varepsilon_{2,t} \\ &\vdots \\ r_{B,t} &= \alpha_K + \beta_{BK} r_{K,t} + \varepsilon_{Kt}, \end{aligned}$$

and then test whether $\alpha_i < 0$ for $i = 1...K$ (i.e., test K restrictions). The second hypothesis can be tested using one multiple regression:

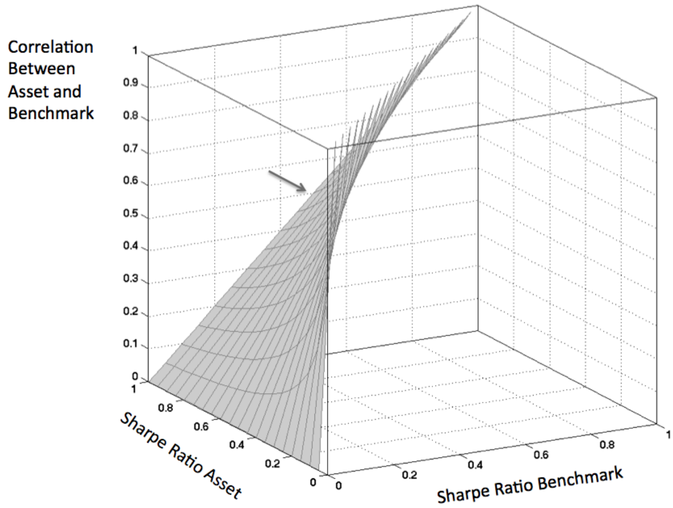
$$r_{B,t} = \alpha_B + \beta_1 r_{1,t} + ... + \beta_K r_{K,t} + \varepsilon_{B,t},$$

and then test the single restriction whether $\alpha_B < 0$. Notice that this regression gives the same intercept as when we would first create the mean-variance efficient portfolio from the K assets and use the return on that portfolio as the right hand side variable. Thus, this means that in the multiple regression we basically test the benchmark against the efficient portfolio of the K assets, which has the highest Sharpe ratio. In other words, this is the portfolio that is most difficult for the benchmark to beat.

1.2 Rules of Thumb

To see the implications of our results in the previous section consider Figure 1, where we plot the relation in (positive) Sharpe-Correlation space, a useful tool to visualize the concepts we introduce in this paper. We plot the Sharpe ratio of the asset on the horizontal axis from middle to left, the Sharpe ratio of the benchmark on the horizontal axis from the middle to the right and the correlation between A and the benchmark on the vertical axis. The surface in the figure marks where the equality of (2) holds. (Without loss of generality we restrict this for ease of exposure to Sharpe ratios between 0 and 1 and positive correlations). Assets on the right hand side of the surface fall in the area where diversification is optimal. The benchmark dominates the asset. However, assets on the left hand side of the surface dominate the benchmark. Investors would prefer these over the benchmark in the sense that they would invest 100% or more in those assets. While this non-diversification space seems large, to be more than a theoretical artifact, it should be inhabited by a large enough number of assets. The main point of our paper is that such a space exists and that based on reasonable data on future market expectations it is inhabited by a large number of assets. In many cases, this might be individual stocks.

Figure 1: The Sharpe Correlation Space



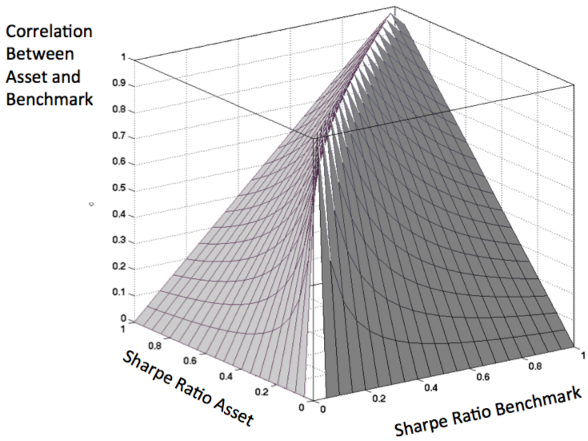
It may be good to illustrate this with some realistic numbers. The average annual standard deviation of an individual stock is about twice as high as the average standard deviation of a market index (40% versus 20% are good proxies). Moreover, the average stock has a correlation of 0.5 with the benchmark. A range between 0.6-0.4 for a historical Sharpe ratio for the US market would also be a good indication. We use 0.5 here in our illustration. Based on these numbers we can derive some simple rules of thumb.

To prefer a one stock investment over a diversified benchmark, the expected Sharpe ratio of that stock should be double the Sharpe ratio of the benchmark. For instance, with a benchmark Sharpe ratio of 0.5 and a correlation for the average of 0.5 for the average stock it is easy to see in figure 1, that the Sharpe ratio for the asset should be one or larger to fully invest in that stock (point indicated by the arrow).

In terms of expected returns this means that for the average stock the expected excess return should be four times the expected excess return of the benchmark. With annual excess returns on indices between 5 and 10 percent, one would need to expect that this stock would generate excess returns between 20 and 40 percent in a given year. For a stock that co-moves strongly with the benchmark and for which the correlation goes towards one, the expected return must be at least twice as high as the expected return on the benchmark. Between 10 and 20 percent. This is also clear from Figure 1 for higher correlations the space on the left hand side gets larger.

If an investor has high expectations for more than one stock, it is easier to satisfy the condition above. With more than one stock the correlation goes up and the standard deviation of the portfolio will decrease. If we consider

Figure 2: The Sharpe Correlation Space



an investor who will invest in a few stocks and will hold equally weighted positions in these different stocks and holds stocks in larger companies the required Sharpe ratios and expected returns go down fast.[6] For limited portfolios consisting of four or five stocks the Sharpe ratio of the portfolio only needs to be 50 percent higher than for the benchmark, which translates in expected returns roughly twice the expected returns on the benchmark. With eight to ten stocks a Sharpe ratio 25% higher suffices, which translates to expected returns being some seventy percent higher than the benchmark. Suppose that investor expects the S&P 500 to go up with ten percent, but believes that four stocks will outperform the market and generate returns of twenty percent or more. This investor should according to the Modern Portfolio Theory not diversify at all but just hold an equally weighted portfolio of these four stocks.

On a final note it may be good to point out that the diversification space on the right hand side of the surface in Figure 1 can be further subdivided in two parts. In Figure 2 we add this subdivision to Figure 1 by adding a new surface (dark grey). One part (the left hand side of this new surface but on the right hand side of the old surface) indicates the combinations between correlations and Sharpe ratios where investors would opt to invest both in the benchmark B and asset A (for instance, 50 percent in the benchmark and 50 percent in asset A). The space on the right hand side of the new surface is the 'complete diversification'-area, the area where the benchmark truly dominates other assets. An investor would only hold the benchmark.

We have established the theoretical part. However, its practical relevance still needs to be determined. While there may be an asset picking space, we

[6]We base these required Sharpe ratios and required expected returns on results on the Top 500 stocks for the equally weighted portfolios we report in Table 11.

still have to determine whether it is inhabited.

2 Empirical Results

2.1 Single stocks

To illustrate the main concepts we use monthly returns taken from the CRSP data for the largest 500 non financial companies (measured by market capitalization in December 2011), with a history of at least five years. The data cover the years 1926-2011 which allows us to study our approach in many different and interesting historical time periods. We do not include companies that have SIC code starting with a 6 (banks, financial and investment companies, Reits, iShares trusts, index funds). We remove these from our analysis as these may by themselves be well diversified portfolios of assets. The risk free rate is taken from Ken French's website[7] and for the first few months of 1926 we use a similar T-bill rate taken from Global Financial Data.

We use individual stock returns, and construct a value weighted index from these Top 500 stocks to use as a benchmark and later construct (value weighted) portfolios of a limited number of stocks. This we use as our base case. While there may be some drawbacks to this choice, the main reason we focus on these results is because these seem to be most representative (the 'median result' if one likes) of our overall results. Of course, the Top 500 stocks contain survivorship bias but as we also use a value weighted index created from these stocks, we have a similar bias on both sides of the equation.

In our robustness section we consider what happens if we change indices (we compare our results later on with an equally weighted index derived from these stocks, the CRSP value-weighted index and the S&P 500 total return index). We also compare results when we construct investor portfolios (with a limited number of shares) in different ways (we will use mimicking portfolios, efficient portfolios and value and equally weighted portfolios). And, we consider the impact if we use a random sample of 500 stocks rather than the Top 500 to prevent selection bias. Overall results vary but our general conclusion remains unaffected.

To give a preview of results we discuss more extensively below: results tend to favor stock picking more compared to our base case if we use the CRSP index or the S&P 500 total return. However, if we use the equally weighted index diversification dominates. Results for the 500 randomly selected stocks

[7]http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html. We appreciate he makes these data available.

favor diversification more strongly as well. Mimicking, efficient or equally weighted portfolios of a limited number of stocks tend to favor stock picking more strongly.

Table 1 compares the value weighted index with the equally weighted index, the CRSP index and the S&P 500 total return index over the sample period. In the last two columns we report the average results for the individual stocks and the standard deviations of these estimates. The value weighted index is highly correlated with the other indices and has a similar slightly lower standard deviation. The mean (excess) returns tend to be higher for the value weighted index and particularly for the equally weighted index. Mean returns for the individual stocks are high and vary substantially across stocks as the standard deviation of 8.12% indicates.

The average excess return on our value weighted benchmark index is 8.78% annually with a standard deviation of 18.25% leading to a Sharpe ratio of 0.48. The average excess return for the individual stocks is higher with 17.24% annually and a standard deviation 39.05%, yielding an average Sharpe ratio of 0.128 (Note that this is the average Sharpe ratio measured over all stocks and contrary to the indices cannot be derived from dividing excess returns by the standard deviation in Table 1).

Just looking at the standard deviation of individual stocks suggests that diversification would reduce risk substantially. For instance, the average correlation of a stock with the value weighted benchmark equals 0.45. These numbers are close to the rounded numbers we used in our rules of thumb: a stock generally needs a Sharpe ratio of over 0.48/0.45=1.07 or slightly higher than 1 to make an investor strictly prefer this stock over the fully diversified value weighted benchmark.

2.1.1 Single stock results: estimates based on historical data

We now test for the individual stocks how often we can reject the null hypothesis that the diversification alpha is equal to zero or negative in the regression:

$$r_{B,t} = \alpha_B + \beta_{BA} r_{A,t} + \varepsilon_{B,t}$$

We assume that investors use an estimation window of five years of historical returns to as an indication of future returns. (Of course, the choice of five years is arbitrary but switching to for instance, 1 or 10 years does not affect our main results). In other words, this means we assume that an investor in January 2011 would use estimates over 60 observations for each of the largest 500 non financial stocks as a proxy for future expectations for these stocks to determine whether this investor should hold a diversified value weighted

	Value Weighted Index	Equally Weighted Index	S&P 500 Total Return	Fama & French Index	Individual Stocks	Standard Deviation Individual Stocks
Annual Mean Return	12.31%	17.93%	9.27%	10.93%	21.00%	8.12%
Annual Mean Excess Return	8.78%	14.40%	5.75%	7.41%	17.24%	8.49%
Annual Standard Deviation	18.25%	20.98%	19.11%	18.87%	39.05%	15.52%
Annual Sharpe Ratio	0.480	0.686	0.300	0.392	0.128	0.063
Correlations between indices/stocks						
Value Weighted Index	1	0.937	0.975	0.977	0.451	0.125
Equally Weighted Index		1	0.944	0.970	0.405	0.117
S&P 500			1	0.988	0.380	0.124
Fama & French Index				1	0.406	0.122
Average Individual Stock					0.222	0.116

Table 1: Basic characteristics of indices and the Top 500 stocks

Basic characteristics for indices and Top 500 stocks (largest market capitalization in December 2011) over the period 1926-2011 using data from CRSP, Global Financial Data and Ken French's website. The last column reports the standard deviation for the mean individual stocks results in the previous column.

benchmark or prefer an investment in that single stock. The second column of Table 2 contains our estimation results. We report the percentage of stocks with negative alphas and the percentage of stocks for which we cannot reject the null hypothesis of stock picking i.e. we cannot reject that alpha positively deviates from zero at different significance levels.

Based on stocks we find 30.88% negative 'Stockpicking'-alphas. In other words, for almost a third of all stocks we find that based on the full sample data their Sharpe ratio multiplied with the correlation of the benchmark exceeds the Sharpe ratio of our benchmark (as per equation 2). This means that in our Sharpe Correlation space these stocks are in the stock picking space. Looking at the second row of Table 2 we find that for almost 70 percent (69.92%) of all stocks we cannot (at the 10 percent level) reject the null that an investors should hold these individual stocks (in favor of the well diversified benchmark). Or in other words, based on this confidence level statistically speaking 70 percent of all stocks are so close to the stock picking space, they might as well be in it (or should be substantially overweighted relative to the benchmark). And, if we consider generally more widely used confidence levels these percentages increase even further. This first direct test of the stock picking versus diversification decision does not support diversification as a favorable option very strongly. Particularly, if we take into account that we are talking about individual stocks.

A five year rolling window also allows us to measure the percentage of stocks that dominate the different benchmarks in different months. We report these results in last row of Table 2. In any given month almost a quarter (23.21%) of the individual stocks dominate the benchmark. So in an average month, based on five year estimation periods, we would expect to find 23.21% stocks with negative alphas in comparison with a value weighted index. And again this percentage increases dramatically even at the 10% level. At the 10% level we cannot reject the null hypothesis that alpha is larger than or equal to zero for two thirds (65.21%) of all stocks in our sample.

It seems that if we are willing to assume that investors might have expectations based on historical estimates, they would choose individual stocks over a well diversified benchmark surprisingly often. If one feels that these historical estimates based on five year intervals might be a good proxy for investor expectations, we cannot reject the null that the benchmark dominates individual stocks for more than half of the stocks in our sample. Of course, one could argue that this assumption may be too strong.

2.1.2 Single stock results: analyst expectations

Do the stock picking versus diversification results hold for realistic expectations about future stock returns? An ideal data set to answer this question is

Significance level:	stocks	months
Percentage negative alphas:	30.88%	23.21%
Percentage alpha not significantly larger than zero:		
10% level	69.92%	65.21%
5% level	76.46%	74.12%
1% level	85.50%	85.73%
0.1% level	94.74%	95.02%

Table 2: Stock Picking Versus Diversification

This table reports the percentage negative 'Stockpicking'-alphas and the percentage 'Stockpicking'-alphas not significantly larger than zero for given confidence levels based on one sided tests. Alphas are derived from the regression $r_{B,t} = \alpha_B + \beta_{BA} r_{A,t} + \varepsilon_{B,t}$ based on five year rolling windows and indicate the percentage of stocks for which we cannot reject the null hypothesis that we should choose the benchmark over a 100% investment in the stock.

the expected return data set compiled by Brav, Lehavy and Michaely (2005) (downloadable from Reuven Lehavy's Web page).[8] The database provides annualized expected returns for individual stocks based on Value Line data and is available on a monthly basis for the period from January 1975 to December 2001. Value Line is an independent research provider with no affiliation to investment banking. It analyzes each company on a quarterly cycle such that a typical firm receives four reports per year. For the remainder of this section we assume that an investor has Value Line expectations and we match our selection of stocks with the Value Line expectations.

We use five year estimates of historical correlations and standard deviations as proxies for future risk and comovements between these stocks as before. But we replace the historical mean by the analyst forecast. As expected market return we assume a weighted average of expected returns of the individual stocks available in that month. In an average month in the 1975-2001 period we have 208 stocks in our database. We have on average monthly expectations for 50 stocks and because these are revised every three months, the Value Line expectations cover 72% of the stocks in our sample. Median and average expected annual excess returns are pretty close to each other with 16.54% and 18.10%, respectively. The maximum is 68.53% and minimum expected return equals -8.05%. This maximum may seem high but the maximum actual annual return for a stock in this selection was over 200%. For the expected excess return on the index we find an average of 3.19% (median 2.39%) and these expectations vary between -0.55% and 10.4%.

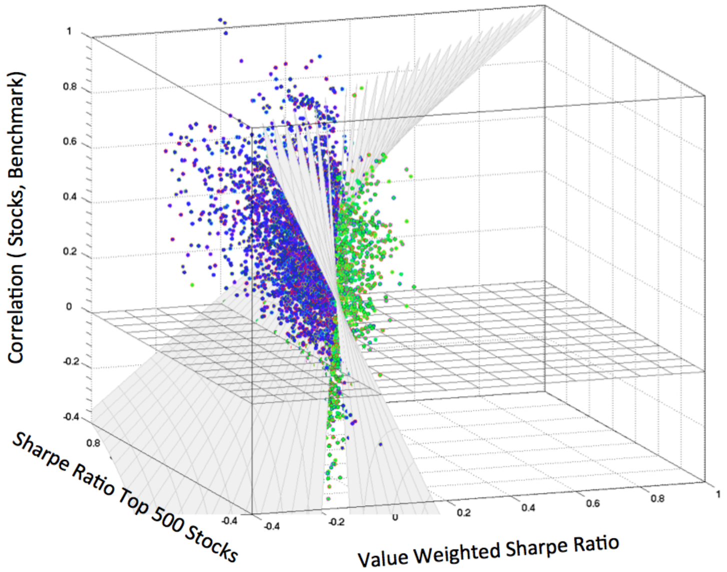
Many financial websites, like CNN Money report analyst expectations for individual stocks and these Value Line expectations seem fairly representative for analyst expectations reported on those websites. If we are willing to assume that investors may hold expectations like these, should they diversify? Figure 3 depicts all these analyst forecasts in the Sharpe Correlation space.

A large percentage of stocks inhabit the stock picking space (blue (dark grey) dots on the left, green (light grey) stocks are in the diversification space). In any given month investors would prefer 40% of all stocks over the value weighted index if they had expectations similar to the Value Line expectations we use here. This is not caused by a limited number of months. In all months in our sample 1975-2001 one can find stocks that dominate a diversified portfolio.

It seems that with reasonable expectations stock picking can be a preferred strategy even if we only consider one stock. One would expect results to improve if we combine more - but a limited number of - stocks to obtain at least some diversification benefits.

[8] http://webuser.bus.umich.edu/rlehavy/VLdata.htm. We thank the authors for making the data available.

Figure 3: Analyst Expectations in Sharpe Correlation Space



2.2 Portfolios of a limited number of stocks

We construct 10,000 value weighted portfolios of a limited number of stocks ranging from one to ten using the Top 500 stocks and compare these with a value weighted index of all 500 stocks. For all portfolios we use the regression $r_{B,t} = \alpha_B + \beta_{BA} r_{A,t} + \varepsilon_{B,t}$ as we did above. For each portfolio we consider each five year interval possible. As noted before we focus on these value weighted portfolios in relation to the value weighted index as they are a good representation of the average results we find. In our robustness tests we consider mimicking, efficient and equally weighted portfolios of a limited number of stocks. Table 3 contains the results for the value weighted portfolios compared with the results for our value weighted benchmark.

The results for the one stock portfolio are - not surprisingly - close to the results for the individual stock results in Table 2. If the number of stocks in our portfolio goes up, the percentage of negative alphas increases. For ten stock portfolios we find 50% negative alpha's so half of the ten stock portfolios are in the stock picking space and for most ten stock portfolios (86% and 91%) we cannot reject the null hypothesis (at the 10 and 5 percent level, respectively) that they are not. Interestingly, the marginal benefits of diversification decrease fast. For instance, for the five stock portfolio we already find 45% negative alphas and we cannot reject the null for 82% (10 percent level) and 87% (5 percent level) of all portfolios.

It would be interesting to see what is driving these results. From the in-

Number of stocks	1	2	3	4	5	6	7	8	9	10
% 'Stockpicking'-alpha < 0	30.78%	35.71%	39.65%	42.21%	44.73%	46.51%	47.77%	49.22%	50.11%	50.85%
% based on t-value at 10% level	68.61%	72.96%	76.47%	79.32%	81.55%	83.21%	84.31%	85.38%	85.97%	86.38%
% based on t-value at 5% level	75.07%	78.79%	81.95%	84.53%	86.65%	88.09%	89.11%	90.06%	90.49%	90.79%
Mean alpha	0.0033	0.0025	0.0018	0.0013	0.0009	0.0007	0.0004	0.0002	0.0001	0.0000
Mean t-value	0.7891	0.6275	0.4914	0.3836	0.2933	0.2267	0.1729	0.1233	0.0882	0.0616
Mean s.e. of alpha	0.0049	0.0046	0.0044	0.0042	0.0041	0.0039	0.0038	0.0037	0.0036	0.0035
Mean of benchmark	0.59%	0.52%	0.48%	0.45%	0.43%	0.42%	0.41%	0.41%	0.40%	0.40%
Standard deviation of benchmark	4.27%	4.24%	4.23%	4.22%	4.20%	4.19%	4.17%	4.15%	4.14%	4.12%
Sharpe ratio benchmark	0.1552	0.1410	0.1316	0.1244	0.1197	0.1166	0.1147	0.1144	0.1144	0.115
Mean of portfolio	1.45%	1.08%	0.98%	0.92%	0.89%	0.86%	0.84%	0.84%	0.83%	0.82%
Standard deviation of portfolio	10.61%	8.44%	7.63%	7.12%	6.80%	6.54%	6.34%	6.20%	6.04%	5.92%
Sharpe ratio of portfolio	0.1411	0.1362	0.1382	0.1393	0.1416	0.1430	0.1442	0.1477	0.1493	0.1514
Correlation benchmark & portfolio	0.4496	0.5272	0.5748	0.6114	0.6412	0.6654	0.6877	0.7041	0.7197	0.7321
Number of portfolios	10,000	10,000	10,000	10,000	10,000	10,000	10,000	10,000	10,000	10,000
Rescaled limited portfolio values	1	2	3	4	5	6	7	8	9	10
Mean of portfolio	1.46%	1.21%	1.19%	1.18%	1.20%	1.20%	1.20%	1.21%	1.20%	1.19%
Standard deviation of portfolio	10.61%	8.51%	7.71%	7.22%	6.92%	6.68%	6.49%	6.38%	6.24%	6.14%
Sharpe ratio of portfolio	0.1411	0.1466	0.1586	0.1691	0.1786	0.1853	0.1901	0.1958	0.1980	0.2001

Table 3: Characteristics of the Portfolios with a limited number of stocks

We report simulation results for value weighted portfolios with a limited number of (Top 500) stocks based on the regression $r_{B,t} = \alpha_B + \beta_{BA} r_{A,t} + \varepsilon_{B,t}$. The benchmark is a value weighted benchmark created from the same stocks.

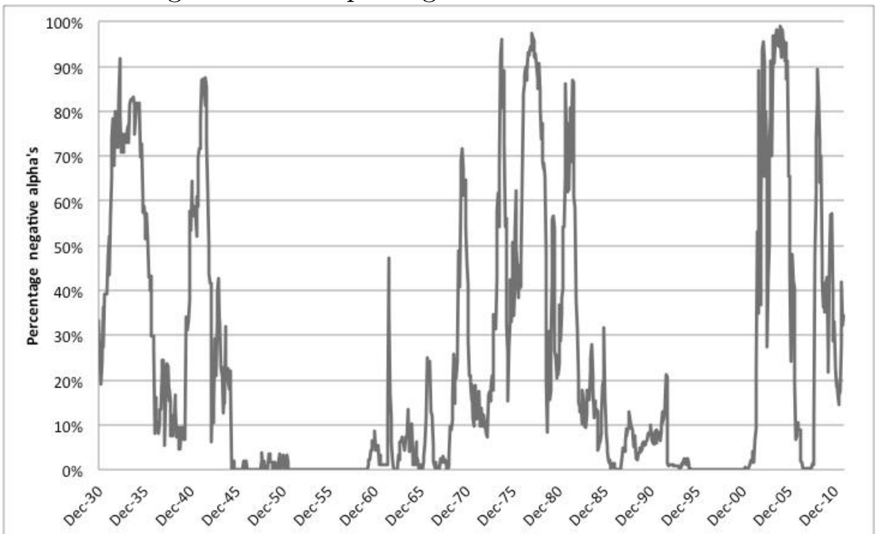
equality $Sh_B < \rho_{AB} Sh_A$ we know this depends on the Sharpe ratio of the benchmark, the Sharpe ratio of the portfolio and the correlation between benchmark and portfolio returns. Firstly, it seems safe to say, looking at the last row, that the increasing correlation between portfolios and the benchmark is an important factor. And, as with the result for alpha, the marginal increase in correlation seems to decrease fast when the number of stocks gets larger.

Analyzing the Sharpe ratios is a bit more complicated. There is no reason why the Sharpe ratio of the benchmark should change across portfolios if all stocks had an equally long history (and indeed if we run a simulation with equal observations for stocks it does not). However, in our sample here, it seems that the more recent past when we have observations for many stocks, stock returns have been lower on average. In those periods possible combinations of stocks into larger portfolios are more likely, hence in our simulations these periods will be overweighted. The decreasing Sharpe ratio of the benchmark reflects this. While the standard deviation of the value weighted benchmark portfolio is hardly affected, the benchmark mean goes down from 0.59% to 0.40% monthly excess returns. The portfolios of limited stocks should exhibit a similar bias. However, to see this more clearly we added some rows where we correct the limited stock portfolio measures for the sample selection bias. We do the rescaling based on the fact that the benchmark mean and standard deviation should be constant in the absence of this bias.

This reveals some interesting phenomena. Firstly, diversification reduces risk we see an increase in Sharpe ratios when the number of stocks in our portfolios increases. This means an increase in Sharpe ratio also drives the increase in percentage of negative alpha's when the number of stocks in a portfolio increases. Secondly, the mean excess returns remains unchanged, with the exception when we move from one to two value weighted stock portfolios the mean return decreases substantially as well. A single stock portfolio performs on average better than a two stock value weighted portfolio. An interesting phenomenon which seems similar to the results generally found for equally weighted versus value weighted indices (see Plyakha, Uppal and Vilkov, 2012). (Indeed in our section where we consider alternative portfolios we see this does not happen for equally weighted portfolios of a limited number of stocks). Note that the numbers in this table also allow the rules of thumb we gave before. We derive them from the small equally weighted portfolios we use in the robustness tests, but the principle is the same. Once we know the correlation of these limited stock portfolios with the benchmark the required Sharpe ratio follows and if we know the standard deviations we can get the required rates of expected returns.

Generally speaking it seems that correlations of limited stock portfolios go up quite fast when we increase the number of stocks and at the same time

Figure 4: Stock picking and Diversification over Time



standard deviations of these portfolios drop quite rapidly. Based on the empirical evidence our tests suggests that stock picking may not be such a bad strategy after all if we cannot reject the null hypothesis in favor of diversification most of the time. Even if we consider only one or a very limited number of stocks in a portfolio. But is stock picking always a good strategy or does an investor benefit more in some periods than in others?

2.3 Stock picking versus diversification over time

Figure 4 illustrates how the percentage negative alphas (from Table 2) drastically fluctuates over time.

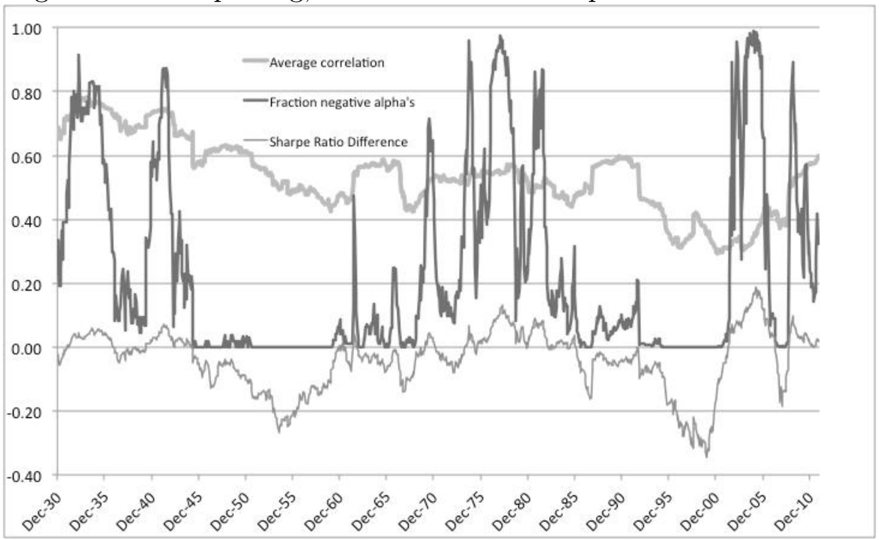
There are periods when stock picking dominates diversification (high percentage negative alphas) and the other way around (low percentage of negative alphas). Once more the inequality $Sh_B < \rho_{AB} Sh_A$ implies that either changes in Sharpe ratios of the market and stocks should be driving these fluctuations or changes in correlations between stocks and market should cause these fluctuations. In Figure 5 we compare the negative alphas with the average correlation and with the difference of the average Sharpe for the stocks and the Sharpe ratio of the market.

Interestingly, the difference in Sharpe ratios (Average Sharpe ratio of stocks in a month minus Sharpe ratio of the value weighted index) seems to be the main driver of the fluctuations in stock picking versus diversification decisions over time. This is confirmed in Table 4 where we report the correlations between these variables. The correlation between the average correlation of the stocks and the value weighted index and the percentage of negative alphas

	% negative alphas	Average correlation	Sharpe difference	Std. dev. difference	Mean difference	NBER 5yr MA
% negative alphas	1	0.25	0.74	0.48	0.41	0.45
Average correlation	0.25	1	0.33	-0.06	0.05	0.52
Sharpe difference	0.74	0.33	1	0.32	0.49	0.35
Std. dev. difference	0.48	-0.06	0.32	1	0.51	0.22
Mean difference	0.41	0.05	0.49	0.51	1	0.10
NBER 5yr MA	0.45	0.52	0.35	0.22	0.10	1

Table 4: Correlations between different variables

Figure 5: Stock picking, Correlation and Sharpe ratio differences over time



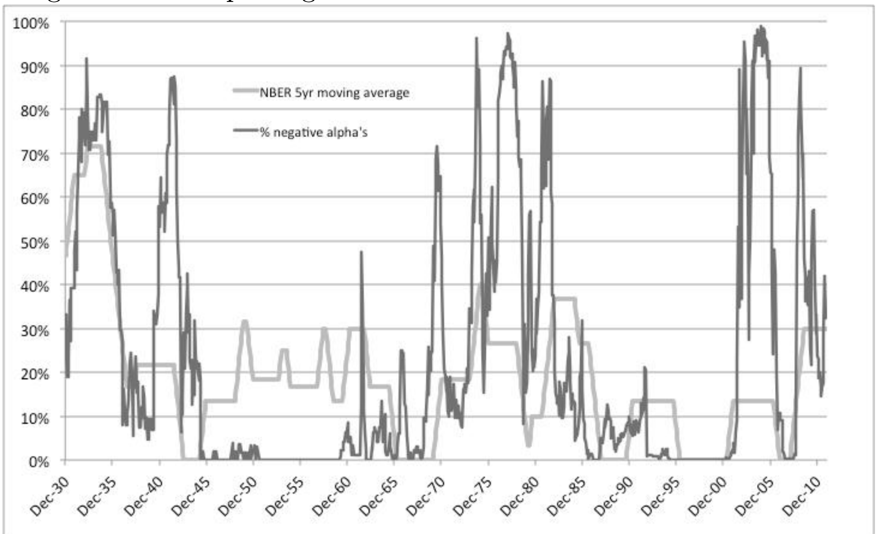
equals 0.25 whereas the correlation of the percentage of negative alphas and the difference between Sharpe ratios of the stocks minus the Sharpe ratio of the market equals 0.74. This difference in Sharpe ratio's in turn seems to correlate highly with both the difference in mean returns (0.49) and the difference in standard deviation (0.32). It seems a natural question to ask whether the attractiveness of diversification versus stock picking depends on the business cycle. Therefore we also report the correlation with a NBER Business Cycle indicator. (We construct a five year moving average from NBER dummy, which has the value 1 in contractions and the value 0 in expansions). This indicator measures the percentage recession states in our estimates. The overall correlation between these two variables is 0.45. We compare how the percentage of negative alpha's and this indicator change over time in in Figure 6.

Generally, diversification seems a better strategy in expansions and stock picking dominates in recessions.

2.4 'Stockpicking' and "Diversification'-stocks

What type of stocks are best suited for stock picking and which ones are better suited for diversification? The Fama and French data - again taken from Ken French's website - seem ideally suited to answer these questions. We consider their decile portfolio data sorted on different criteria (Size, Book to Market, Earnings/Price, Cash Flow/Price, Dividend Yield, Momentum, Short Term Reversals, Long Term Reversals) and portfolios sorted into ten industries. For all these portfolios we calculate the percentage of negative

Figure 6: Stock picking and Diversification related to the Business Cycle



alphas to get a feel for what characterizes 'Diversification' and 'Stockpicking'-stocks. We use the CRSP index as benchmark. Table 5 contains our results. The 35% in the highest decile portfolio of the 'Size'-column indicates that based on the five year rolling window estimates we find 35% negative alphas for this portfolio. The numbers in italics mark the 'Diversification'-stocks (the three deciles with the lowest percentages of negative 'Stockpicking'-alphas) and the bold values indicate 'Stockpicking'-stocks (the three deciles with the highest percentages of negative alphas).

'Stockpicking'-stocks tend to be large but not the largest decile stocks (which are more likely 'Diversification'-stocks). They have high 'Book to Market'-ratios, high 'Earnings Price'-ratios, high 'Cash Flow Earnings'-ratios and high Momentum. They show little evidence of Short Term Reversals, exhibit average Long Term Reversal effects and pay average dividend yields. 'Diversification'-stocks are either large or small firms, with low 'Book to Market'-ratios, low 'Earnings Price'-ratios, low 'Cash Flow Earnings' -ratios and no Momentum. Dividends yields can be high or low, and they show strong Short Term Reversals and either very little or very strong Long Term Reversals. The last column of Table 5 distinguishes between the different industries where we can find these stocks. 'Stockpicking'-stocks prevail in Non Durables, Energy, Shops and Healthcare industries. "Diversification' stocks dominate in Durables, Telecom and Other industries.

	Size	Book/Market	Earnings/Price	Cash Flow/Price	Div.Yield	Momentum	Short Term Reversals	Long Term Reversals	10 industries	
Low 10%	*43%*	*27%*	*36%*	*31%*	*40%*	*27%*	**65%**	*47%*	**54%**	NoDur
2	*40%*	*35%*	*44%*	*43%*	*58%*	*39%*	**66%**	52%	*41%*	Durbl
3	49%	*48%*	*46%*	*47%*	*59%*	*45%*	**60%**	56%	50%	Manuf
4	49%	58%	59%	60%	**72%**	53%	**60%**	61%	**53%**	Enrgy
5	46%	62%	66%	70%	**72%**	61%	55%	**71%**	46%	HiTec
6	**58%**	**66%**	72%	75%	**72%**	**76%**	55%	**70%**	*43%*	Telcm
7	**56%**	59%	**82%**	**86%**	**72%**	**76%**	43%	**66%**	**51%**	Shops
8	53%	**63%**	**81%**	**81%**	70%	**78%**	*34%*	58%	**51%**	Hlth
9	**64%**	**71%**	**85%**	**81%**	65%	**80%**	*27%*	*49%*	47%	Utils
High 10%	*35%*	50%	78%	76%	53%	71%	*8%*	*24%*	*38%*	Other

Table 5: Stock Picking, Diversification for Fama and French Decile Portfolio's

Percentage of cases when 'Stockpicking'-alphas from the regression $r_{B,t} = \alpha_B + \beta_{BA} r_{A,t} + \varepsilon_{B,t}$ are negative for the different decile portfolios sorted on Size, Book to Market, Earnings/Price, Cash Flow/Price, Dividend Yield, Momentum, Short Term Reversals, Long Term Reversals and portfolios sorted into ten industries.

3 Relation with Return Predictability

3.1 The 'Stockpicking'-alpha and Jensen's alpha

Higher expectations on certain assets or portfolios may come from returns being partially predictable i.e. a positive alpha. A natural question to consider is how our 'Stockpicking'-alpha α_B and Jensen's alpha α_A are related? Obviously, for a given benchmark this follows directly from standard regression analysis, but it is useful to restate it in terms of the current analysis

$$\begin{aligned}
\alpha_B &= E\left[r_{B,t}\right] - \frac{\sigma_{AB}}{\sigma_A^2} E\left[r_{A,t}\right] & (10)\\
&= E\left[r_{B,t}\right]\left(1 - \frac{\sigma_{AB}^2}{\sigma_A^2\sigma_B^2}\right) - \frac{\sigma_{AB}}{\sigma_A^2}\left(E\left[r_{A,t}\right] - \frac{\sigma_{AB}}{\sigma_B^2} E\left[r_{B,t}\right]\right) & (11)\\
&= E\left[r_{B,t}\right]\left(1 - R^2\right) - \beta_{BA}\alpha_A & (12)
\end{aligned}$$

equation (12) relates the level of return predictability (measured by Jensen's alpha) to the level of stock picking (measured by our 'Stockpicking'-alpha). Generally, the higher Jensen's alpha, the lower the 'Stockpicking' alpha as all the other terms and variables will under normal circumstance not cause a sign switch.

Note that if $R^2 = 0$, then $\beta_B = 0$ also, and we have $\alpha_B = E\left[r_{B,t}\right]$ (obviously). If $R^2 = 1$, then $\beta_B = 1$, and $\alpha_B = \alpha_A$ (obviously as well). The restriction on the intercept that $\alpha_B < 0$ is translated into α_A in the previous subsection. equation (12) also tells us what to expect for the reverse Jensen's alpha if the benchmark B actually is mean-variance efficient. Since mean-variance efficiency of the benchmark implies that $\alpha_A = 0$, we get that in the reverse Jensen regression $\alpha_B = E\left[r_{B,t}\right](1 - R^2)$. In general then, with positive expected returns on the benchmark portfolio and unless the single asset A is perfectly correlated with the benchmark, the reverse Jensen measure $\alpha_B > 0$ if the benchmark is mean-variance efficient. Our test of the null hypothesis $H_0 : \alpha_B \leq 0$ would therefore be a small hurdle to take for efficient benchmarks.

A related question is to express the restriction (2) in terms of the Jensen regression (7). Starting from the inequality condition for Sharpe ratios we can rewrite:

$$\begin{aligned}
Sh_B &< \rho_{AB} Sh_A \\
&\Leftrightarrow \frac{E\left[r_{B,t}\right]}{\sigma_B} < \frac{\sigma_{AB}}{\sigma_A\sigma_B}\frac{E\left[r_{A,t}\right]}{\sigma_A} \\
&\Leftrightarrow \frac{E\left[r_{B,t}\right]}{\sigma_B} - \frac{\sigma_{AB}^2}{\sigma_A^2\sigma_B^2}\frac{E\left[r_{B,t}\right]}{\sigma_B} < \frac{\sigma_{AB}}{\sigma_A^2\sigma_B}\left(E\left[r_A, t\right] - \frac{\sigma_{AB}}{\sigma_B^2}E\left[r_{B,t}\right]\right)
\end{aligned}$$

$$\Leftrightarrow \quad Sh_B\left(1-R^2\right) < \rho_{AB}\frac{\alpha_A}{\sigma_A}$$
$$\Leftrightarrow \quad \alpha_A > Sh_B\left(1-R^2\right)\frac{\sigma_A}{\rho_{AB}}.$$

Recognizing that $(1-R^2)\,\sigma_A^2 = \sigma_{\varepsilon A}^2$, i.e., the residual variance of the regression, and by noting that $R^2 = \rho_{AB}^2$ we can rewrite the inequality condition as:

$$\frac{\alpha_A}{\sigma_{\varepsilon A}} > Sh_B\sqrt{\frac{1-\rho_{AB}^2}{\rho_{AB}^2}}. \tag{13}$$

This inequality says that the appraisal ratio or information ratio of the single stock should exceed the Sharpe ratio of the benchmark, adjusted by a function of the correlation. If the correlation is perfect (± 1), the information ratio should exceed zero, which makes sense. If the the correlation gets close to zero, the right-hand-side of (13) goes to infinity, implying that we will not want to invest 100% in the single stock, but always invest part of the portfolio in the benchmark as well for diversification reasons. This also makes sense. If $\rho_{AB} = \sqrt{1/2}$, the square root term becomes one, and we need that the information ratio should exceed the Sharpe ratio of the benchmark.

3.2 Empirical link between both alphas

Based on the Fama and French decile portfolios we compare Jensen's alphas with our 'Stockpicking'-alphas.[9] In Table 6 we compare the alphas for both the Size and Book to Market Decile Portfolios.[10] Note that for the 'Size'-portfolios Jensen's alpha decreases consistently with Size reconfirming the effect that small firms tend to perform better on average. However, this does not imply that from a stock picking perspective it makes sense for an investor to focus solely on small firms. At least not for an investor who uses the CRSP index as a benchmark. The lowest decile portfolio - even though it shows a very high Jensen's alpha (0.84%) has a relatively low R^2 (0.55) and this makes it uninteresting from a stock picking perspective. The most interesting portfolios from a stock picking perspective are the 6th and 7th decile portfolio with high R^2's (0.88 and 0.92, respectively) even though these portfolios have substantially lower Jensen's alphas. Note that the Size portfolios tend to be the exception to the more general rule that there is a close link between Jensen's alpha and our 'Stockpicking'-alpha.

[9] The numbers in the table do not match the equation exactly as we report averages. However, for the individual subsamples the relation does hold exactly.

[10] To save space we do not report the results for the other decile portfolios. These are however available on request from the authors.

The Book to Market portfolios show this more general trend (which we observe for the (non-reported) other decile portfolios as well). If Jensen's alpha increases the 'Stockpicking'-alpha decreases and this means that these portfolios are more interesting from a stock picking perspective as well. However, also in this case the 10th decile portfolio has a lower R^2 which makes it less interesting from a stock picking perspective (if an investor uses the CRSP index as a benchmark).

4 Alternative data sets, assumptions and tests

4.1 500 Random Stocks

The Top 500 stocks are likely biased and while it may be a reasonable selection of stocks an investor might consider, it may be good to verify what happens if we use a completely random sample of stocks. Here we discuss the results of a random selection of 500 stocks. Table 7 contains the basic characteristics of the random sample and the value weighted and equally weighted indices created from this sample. For ease of comparison we also include the other indices. Note, that these stocks have a substantial lower return than the Top 500 (16.58% versus 21.00%). They also tend to be more risky with a standard deviation of 56.61% (versus 39.05%), resulting in a substantially lower Sharpe ratio of 0.069 (versus 0.128).

These stocks also show a higher variation in outcomes as indicated by the higher standard deviation for all these estimates. Last but not least, the correlations between stocks and between stocks and indices are lower compared to the Top 500 stocks. While there are differences, these seem to hardly affect our results. In Table 8 we report the results for the random 500 stocks similar to our earlier results in Table 2 for the Top 500. These results are quite similar.

4.2 Other indices

Our results so far have relied on a value weighted index. Here we compare these results with several other indices. We choose the CRSP index given its importance in academic research, the S&P 500 (total return) because of its importance as a benchmark for mutual funds. Cremer, Petajisto and Zitzewitz (2008) show how the S&P 500 is the primary benchmark for most mutual funds. Moreover, they also show that results between the CRSP index and the S&P 500 can be different. The equally weighted index is interesting as Plyakha, Uppal and Vilkov (2012) show that an equally weighted index substantially outperforms value weighted indices.

Size Decile Portfolios	1	2	3	4	5	6	7	8	9	10
Jensen's alpha (α_A)	0.84%	0.32%	0.28%	0.22%	0.19%	0.17%	0.13%	0.09%	0.08%	-0.03%
β_{AB}	1.26	1.29	1.26	1.22	1.20	1.17	1.14	1.11	1.05	0.98
Percentage positive α_A	73%	60%	64%	61%	64%	68%	69%	65%	77%	42%
'Stockpicking'-alpha (α_B)	0.00%	0.04%	-0.02%	-0.01%	-0.03%	-0.05%	-0.05%	-0.03%	-0.04%	0.05%
β_{BA}	0.46	0.58	0.63	0.67	0.71	0.76	0.8	0.84	0.89	0.97
Percentage negative α_B	43%	40%	49%	49%	46%	58%	56%	53%	64%	35%
average R^2	0.55	0.73	0.79	0.82	0.85	0.88	0.92	0.93	0.95	0.96
average mean of benchmark	0.60%	0.60%	0.60%	0.60%	0.60%	0.60%	0.60%	0.60%	0.60%	0.60%

B/M decile portfolios	1	2	3	4	5	6	7	8	9	10
Jensen's alpha (α_A)	-0.27%	0.01%	0.13%	0.31%	0.35%	0.41%	0.49%	0.56%	0.78%	0.91%
β_{AB}	1.25	1.16	1.13	1.13	1.09	1.09	1.08	1.10	1.15	1.27
Percentage positive α_A	32%	47%	63%	79%	78%	79%	78%	81%	87%	83%
'Stockpicking'-alpha (α_B)	0.27%	0.09%	0.00%	-0.10%	-0.13%	-0.18%	-0.19%	-0.21%	-0.28%	-0.14%
β_{BA}	0.68	0.74	0.74	0.72	0.74	0.75	0.72	0.70	0.63	0.49
Percentage negative α_B	27%	35%	48%	58%	62%	66%	59%	63%	71%	50%
average R^2	0.82	0.84	0.84	0.81	0.81	0.81	0.77	0.75	0.71	0.59
average mean of benchmark	0.60%	0.60%	0.60%	0.60%	0.60%	0.60%	0.60%	0.60%	0.60%	0.60%

Table 6: Relation between 'Stockpicking'-alphas and Jensen's alphas
Based on equation 6: $\alpha_B = E\left[r_{B,t}\right]\left(1-R^2\right) - \beta_{BA}\alpha_A$ we report the link between the 'Stockpicking'-alpha and Jensen's alpha.

	Value weighted index	Equally weighted index	S&P 500 total return	CRSP index	Individual Stocks	Standard Deviation Individual Stocks
Annual Mean Return	11.76%	16.52%	9.27%	10.93%	16.58%	13.19%
Annual Mean Excess Return	8.24%	12.99%	5.75%	7.41%	11.80%	13.23%
Annual Standard Deviation	18.78%	25.00%	19.11%	18.87%	56.61%	25.85%
Annual Sharpe Ratio	0.438	0.519	0.3	0.392	0.069	0.063
Correlations between indices/stocks						
Value weighted index	1	0.826	0.917	0.925	0.326	0.153
Equally weighted index		1	0.851	0.897	0.416	0.135
S&P 500 tot return index			1	0.988	0.336	0.154
Fama & French index				1	0.361	0.152
Individual Stocks					0.157	0.159

Table 7: Basic characteristics indices and Random 500 stocks

Basic characteristics for indices and 500 random stocks over the period 1926-2011 using data from CRSP, Global Financial Data and Ken French's website. The last column reports the standard deviation for the mean individual stocks results in the previous column.

Significance level:	stocks	months
Percentage negative alphas:	25.46%	24.28%
Percentage alpha not significantly larger than zero:		
10% level	65.27%	62.61%
5% level	75.13%	72.96%
1% level	85.70%	86.35%
0.1% level	93.40%	95.35%

Table 8: Percentage of negative 'Stockpicking' alphas for random sample

This table reports the percentage negative 'Stockpicking'-alphas and the percentage 'Stockpicking'-alphas not significantly larger than zero for given confidence levels based on one sided tests. Alphas are derived from the regression $r_{B,t} = \alpha_B + \beta_{BA} r_{A,t} + \varepsilon_{B,t}$ using rolling windows of five years of historical data and indicate the percentage of stocks for which we cannot reject the null hypothesis that we should choose the benchmark over a 100% investment in the stock.

Equally weighted index		
Significance level:	stocks	months
Percentage negative alphas:	2.82%	6.52%
Percentage alpha not significantly larger than zero:		
10% level	31.45%	38.30%
5% level	43.41%	50.91%
1% level	66.87%	71.09%
0.1% level	84.60%	87.17%
CRSP index		
Significance level:	stocks	months
Percentage negative alphas:	41.32%	29.72%
Percentage alpha not significantly larger than zero:		
10% level	76.25%	70.32%
5% level	82.27%	78.43%
1% level	91.80%	89.98%
0.1% level	98.86%	97.90%
S&P 500 total return		
Significance level:	stocks	months
Percentage negative alphas:	50.97%	38.90%
Percentage alpha not significantly larger than zero:		
10% level	79.62%	75.39%
5% level	84.65%	82.11%
1% level	92.60%	91.96%
0.1% level	98.75%	98.12%

Table 9: Percentage of negative 'Stockpicking'-alphas for different indices

This table reports the percentage negative 'Stockpicking'-alphas and the percentage 'Stockpicking'-alphas not significantly larger than zero for given confidence levels based on one sided tests. Alphas are derived from the regression $r_{B,t} = \alpha_B + \beta_{BA} r_{A,t} + \varepsilon_{B,t}$ using rolling windows of five years of historical data and indicate the percentage of stocks for which we cannot reject the null hypothesis that we should choose the benchmark over a 100% investment in the stock.

Table 9 contains our estimation results. Results improve in favor of stock picking if we use the CRSP index or the S&P 500. For instance, for more than half of the Top 500 stocks we cannot say we prefer the well diversified S&P 500 (based on point estimates if we use a five year history to estimate the relevant variables as before) and 1/3 of all stocks in the random sample (result not reported in table). However, for individual stocks an equally weighted index is hard to beat. Based on point estimates again using five years of data these percentages are 2.82% of all stocks and 6.52% of stocks in an average month. While these percentages increase rapidly if we use realistic confidence bounds, this does suggest that the difference in the performance of equally weighted and value weighted indices is an interesting issue that warrants attention as also shown by Plyakha, Uppal and Vilkov (2012).

4.3 Limited portfolios

For the portfolios with a limited number of stocks we relied on value weighted portfolios in our base case. Here we consider three alternative portfolios. From a practical point of view the equally weighted portfolio is interesting, but from an academic point of view it also interesting to see how results relate to (ex post) mimicking portfolios and efficient portfolios (The latter two are the same only when the null hypothesis holds).

In Table 10, 11 and 12 we report these results. Not surprisingly, the efficient and mimicking portfolios perform well in comparison with the value weighted limited portfolios. More interesting is the relative performance of the equally weighted limited portfolios. For instance, equally weighted portfolios of three stocks dominate a value weighted benchmark more than 50% of the time just based on point estimates. Note also that the drop off in returns we see for the value weighted portfolio when we move from one to two stocks does not happen here. Mean returns for these equally weighted limited portfolios remains high which explains their relative attractiveness.

5 Conclusions

We started this paper with a quote from one of our favorite textbooks by Brealey and Myers: "Wise investors don't put all their eggs into just one basket: They reduce their risk by diversification," Our tests and empirical results suggests that this may be a strong statement. It seems that the Modern Portfolio Theory covers the full spectrum from diversification to holding just one asset. If we use this insight to derive stock picking and diversification tests the empirical evidence does not favor diversification very often. This means that either commonly used benchmarks are very inefficient in mean variance terms, or that stock picking may not be such a bad alternative after

Number of stocks	1	2	3	4	5	6	7	8	9	10
% 'Stockpicking'-alpha < 0	30.78%	43.85%	53.77%	61.08%	66.63%	71.18%	74.76%	77.82%	79.79%	81.80%
% based on t-value at 10% level	68.61%	75.42%	81.55%	86.12%	89.40%	91.72%	93.33%	94.56%	95.10%	95.66%
% based on t-value at 5% level	75.07%	80.57%	85.96%	89.82%	92.48%	94.25%	95.41%	96.28%	96.58%	96.92%
Mean alpha	0.0033	0.0015	0.0002	-0.0007	-0.0013	-0.0018	-0.0021	-0.0024	-0.0025	-0.0027
Mean t-value	0.789	0.408	0.110	-0.141	-0.347	-0.525	-0.673	-0.805	-0.904	-1.011
Mean s.e. of alpha	0.005	0.004	0.004	0.004	0.003	0.003	0.003	0.003	0.003	0.003
Mean of benchmark	0.59%	0.53%	0.49%	0.45%	0.43%	0.42%	0.41%	0.41%	0.40%	0.40%
Standard deviation of benchmark	4.27%	4.23%	4.23%	4.21%	4.20%	4.18%	4.17%	4.15%	4.14%	4.12%
Sharpe ratio benchmark	0.155	0.142	0.132	0.125	0.120	0.117	0.115	0.114	0.114	0.115
Mean of portfolio	1.45%	1.28%	1.21%	1.17%	1.13%	1.11%	1.08%	1.07%	1.05%	1.04%
Standard deviation of portfolio	10.61%	8.32%	7.45%	6.89%	6.50%	6.23%	6.01%	5.85%	5.70%	5.57%
Sharpe ratio of portfolio	0.141	0.163	0.176	0.184	0.189	0.194	0.196	0.200	0.200	0.203
Correlation benchmark & portfolio	0.450	0.595	0.678	0.737	0.78	0.812	0.838	0.857	0.874	0.888
Number of portfolios	10,000	10,000	10,000	10,000	10,000	10,000	10,000	10,000	10,000	10,000

Table 10: Simulation results mimicking portfolios

We report simulation results for mimicking portfolios with a limited number of (Top 500) stocks based on the regression $r_{B,t} = \alpha_B + \beta_{BA} r_{A,t} + \varepsilon_{B,t}$. The benchmark is a value weighted benchmark created from the same stocks.

Number of stocks	1	2	3	4	5	6	7	8	9	10
% 'Stockpicking'-alpha < 0	30.78%	42.33%	51.94%	59.00%	64.34%	68.72%	72.15%	75.00%	76.85%	78.76%
% based on t-value at 10% level	68.61%	72.43%	79.91%	85.15%	88.80%	91.19%	92.74%	93.69%	94.06%	94.35%
% based on t-value at 5% level	75.07%	77.02%	83.95%	88.54%	91.49%	93.19%	94.16%	94.67%	94.75%	94.82%
Mean alpha	0.0033	0.0015	0.0002	-0.0007	-0.0013	-0.0018	-0.0021	-0.0024	-0.0025	-0.0027
Mean t-value	0.789	0.384	0.121	-0.067	-0.197	-0.290	-0.350	-0.397	-0.419	-0.444
Mean s.e. of alpha	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.006	0.006	0.006
Mean of benchmark	0.59%	0.55%	0.50%	0.47%	0.45%	0.43%	0.42%	0.42%	0.42%	0.42%
Standard deviation of benchmark	4.27%	4.22%	4.21%	4.20%	4.19%	4.17%	4.16%	4.14%	4.12%	4.11%
Sharpe ratio benchmark	0.155	0.148	0.137	0.129	0.124	0.121	0.119	0.119	0.119	0.119
Mean of portfolio	1.45%	2.56%	3.03%	3.37%	3.73%	4.05%	4.38%	4.64%	4.96%	5.24%
Standard deviation of portfolio	10.61%	12.20%	12.20%	11.95%	11.89%	11.84%	11.88%	11.79%	11.81%	11.82%
Sharpe ratio of portfolio	0.141	0.226	0.272	0.310	0.345	0.376	0.405	0.435	0.462	0.488
Correlation benchmark & portfolio	0.450	0.444	0.444	0.438	0.428	0.418	0.404	0.393	0.378	0.368
Number of portfolios	10,000	10,000	10,000	10,000	10,000	10,000	10,000	10,000	10,000	10,000

Table 11: Simulation results efficient portfolio

We report simulation results for efficient portfolios with a limited number of (Top 500) stocks based on the regression $r_{B,t} = \alpha_B + \beta_{BA} r_{A,t} + \varepsilon_{B,t}$. The benchmark is a value weighted benchmark created from the same stocks.

Number of stocks	1	2	3	4	5	6	7	8	9	10
% 'Stockpicking'-alpha <0	30.78%	44.48%	54.29%	61.58%	67.36%	71.95%	75.55%	78.64%	81.02%	83.18%
% based on t-value at 10% level	68.61%	76.18%	81.72%	86.11%	89.40%	91.83%	93.59%	94.85%	95.58%	96.26%
% based on t-value at 5% level	75.07%	81.29%	86.10%	89.79%	92.50%	94.38%	95.63%	96.54%	96.95%	97.33%
Mean alpha	0.0033	0.0015	0.0003	-0.0007	-0.0015	-0.002	-0.0024	-0.0028	-0.003	-0.0032
Mean t-value	0.789	0.413	0.113	-0.141	-0.35	-0.53	-0.678	-0.809	-0.92	-1.023
Mean s.e. of alpha	0.005	0.005	0.004	0.004	0.004	0.004	0.003	0.003	0.003	0.003
Mean of benchmark	0.59%	0.52%	0.48%	0.45%	0.43%	0.42%	0.41%	0.41%	0.40%	0.40%
Standard deviation of benchmark	4.27%	4.24%	4.23%	4.22%	4.20%	4.19%	4.17%	4.15%	4.14%	4.12%
Sharpe ratio benchmark	0.155	0.141	0.132	0.124	0.12	0.117	0.115	0.114	0.114	0.115
Mean of portfolio	1.45%	1.37%	1.32%	1.29%	1.27%	1.25%	1.24%	1.23%	1.22%	1.22%
Standard deviation of portfolio	10.61%	8.31%	7.28%	6.69%	6.33%	6.06%	5.88%	5.73%	5.60%	5.51%
Sharpe ratio of portfolio	0.141	0.172	0.191	0.204	0.213	0.22	0.224	0.23	0.233	0.238
Correlation benchmark & portfolio	0.45	0.554	0.624	0.676	0.715	0.746	0.771	0.79	0.807	0.82
Number of portfolios	10000	10000	10000	10000	10000	10000	10000	10000	10000	10000

Table 12: Simulation results equally weighted portfolio

We report simulation results for equally weighted portfolios with a limited number of (Top 500) stocks based on the regression $r_{B,t} = \alpha_B + \beta_{BA} r_{A,t} + \varepsilon_{B,t}$. The benchmark is a value weighted benchmark created from the same stocks.

all at least not for commonly used benchmarks. While in times of expansion diversification may be a good strategy our results indicate stock picking may be a viable alternative during recessions. These new insights also allow us to explicitly link two pillars of investment theory: diversification and predictability. Whether and how much investors should move away from the benchmark and invest in strategies that generate positive (Jensen's) alpha is related to the benchmark they use.

While we focused on the stock picking versus diversification our tests might also be used to rank the relative mean variance efficiency of benchmarks. Our tests seem to confirm the well known result that equally weighted benchmarks are hard to beat as documented by Plyakha, Uppal and Vilkov (2012), Tu and Zhou (2011), DeMiguel, Garlappi and Uppal (2009) and Ang (2012). It would however be interesting to see how we can rank other benchmarks and indices (for instance as in DeMiguel, Garlappi and Uppal (2009) and Ang (2012)) in terms of relative efficiency. The poor performance of the ex ante mean variance portfolio is well known, but is it possible to create a benchmark that is (almost) mean variance efficient from a practical perspective? This is a topic we are currently investigating.

"Put all your eggs in one basket – and watch that basket!" is a famous quote by Mark Twain. Our results indicate that this may not be such a bad strategy after all. Many (professional) investors including Warren Buffet regard diversification as a strategy for the uninformed investor. In his words:

"On the other hand, if you are a know-something investor, able to understand business economics and to find five to ten sensibly-priced companies that possess important long-term competitive advantages, conventional diversification makes no sense for you. It is apt simply to hurt your results and increase your risk. I cannot understand why an investor of that sort elects to put money into a business that is his 20th favorite rather than simply adding that money to his top choices - the businesses he understands best and that present the least risk, along with the greatest profit potential. In the words of the prophet Mae West: 'Too much of a good thing can be wonderful."'[11]

Our results suggest that the stock picking approach and these arguments may deserve substantially more academic attention. Because, as we mentioned in the introduction, the Uniform Prudent Investor Act states: "A trustee shall diversify the investments of the trust (...)". The Act will not change overnight. So no doubt trustees 'shall' diversify for time to come. But whether they 'should'?

[11]http://www.berkshirehathaway.com/letters/1993.html

6 References

Ang, Andrew (2012) "Mean-Variance Investing", Working Paper, SSRN.

Blume, Marshall E. and Irwin Friend (1975), "The Asset Structure of Individual Portfolios and Some Implications for Utility Functions" *The Journal of Finance* Vol. 30, No. 2, pp. 585-603.

Boyle, Phelim, Lorenzo Garlappi, Raman Uppal and Tan Wang Wang (2012) "Keynes Meets Markowitz: The Trade-Off Between Familiarity and Diversification." *Management Science,* Vol. 58, No. 2, pp. 253-272.

Brav, Alon, Reuven Lehavy, and Roni Michaely (2005), "Using Expectations to Test Asset Pricing Models" *Financial Management*, 5 – 37.

Cremers, Martijn, Antti Petajisto, and Eric Zitzewitz (2008), "Should Benchmark Indices Have Alpha? Revisiting Performance Evaluation", Working Paper on SSRN.

DeMiguel, Victor, Lorenzo Garlappi, and Raman Uppal (2009), "Optimal Versus Naive Diversification: How Inefficient is the 1/N Portfolio Strategy?" *The Review of Financial Studies*, Vol. 22, no. 55, pp. 1915-1953.

Fama, E.F. and Kenneth R. French (2007) "Disagreement, tastes and asset prices", Journal of Financial Economics, 667-689.

Greenwood, Robin, and Andrei Shleifer (2012) "Expectations of Returns and Expected Returns", Working Paper on SSRN.

Goetzman William N., Alok Kumar,(2008), "Equity Portfolio Diversification", *Review of Finance*, Volume12, Issue3, pg. 433-463.

Huberman, Gur, and Schmuel Kandel (1987), "Mean-Variance Spanning", *The Journal of Finance*, Vol. 42 (Sep., 1987), pg.873-888.

Kelly, Morgan (1995), "All their eggs in one basket: Portfolio diversification of US households", *Journal of Economic Behavior and Organization*, Vol. 27 87-96.

Markowitz, Harry, (1952), "Portfolio Selection", *The Journal of Finance*, Vol. 7, No. 1 (Mar., 1952), pp. 77-91.

Nieuwerburgh, Van, Stijn and Laura Veldkamp (2009) "Information immobility and the home bias puzzle", *The Journal of Finance,* Vol 64, No.3 , 1187-1215.

Plyakha, Yuliya, Raman Uppal and Grigory Vilkov (2012), Why Does and Equal-Weighted Portfolio Outperform Value- and Price-Weighted Portfolios? Working Paper on SSRN.

Rapach, David E. and Guofu Zhou (2012), "Forecasting Stock returns" Forthcoming in the "Handbook of Economic Forecasting" Volume 2.

Sterk, Stewart E. (2010), "Rethinking Trust Law Reform: How Prudent is Modern Prudent Investor Doctrine?" Vol. 95, No. 5, *Cornell Law Review*, pp. 851-904.

Tu, Jun and Guofu Zhou (2011), "Markowitz meets Talmud: A combination of sophisticated and naive dversification strategies" *Journal of Financial Economics*, Vol. 99, pp. 204-215.

7 Appendix A

We consider an investor who is considering two investment opportunities: one is a well-diversified (passive) benchmark, say the S&P 500 index, with return $r_{A,t}$, the other a stock-picking portfolio with return $r_{A,t}$. We think of $r_{A,t}$ as a portfolio of a few (perhaps even one) stock(s) only, about which the investor may have high expectations, more informed beliefs, or is better able to predict returns. The investor's expectations (possibly idiosyncratic) for the returns are reflected in μ and Σ, with elements μ_i and σ_{ij}, $i = A, B$.

For this investor, the optimal mean variance efficient portfolio equals

$$\begin{aligned}\begin{pmatrix} w_A \\ w_B \end{pmatrix} &= \gamma^{-1} \begin{pmatrix} \sigma_{AA} & \sigma_{AB} \\ \sigma_{BA} & \sigma_{BB} \end{pmatrix}^{-1} \begin{pmatrix} \mu_A \\ \mu_B \end{pmatrix} \\ &= \gamma^{-1} \frac{1}{\sigma_{AA}\sigma_{BB} - \sigma_{AB}\sigma_{BA}} \begin{pmatrix} \sigma_{BB} & -\sigma_{AB} \\ -\sigma_{BA} & \sigma_{AA} \end{pmatrix} \begin{pmatrix} \mu_A \\ \mu_B \end{pmatrix}.\end{aligned}$$

Our null-hypothesis is that the investor's beliefs are such that he only wants to invest in $r_{A,t}$ nothing, or even going short in $r_{B,t}$, i.e., $w_B \leq 0$. From the portfolio equation this means

$$\begin{aligned}\frac{-\sigma_{BA}\mu_A + \sigma_{AA}\mu_B}{\sigma_{AA}\sigma_{BB} - \sigma_{AB}\sigma_{BA}} &\leq 0 \\ \sigma_{AA}\mu_B &\leq \sigma_{AB}\mu_A\end{aligned}$$

where the second inequality uses that the determinant is positive, and from this equation (1) follows directly by dividing the l.h.s. and r.h.s. by σ_B and by σ_{AA}.