

Do value stocks outperform in a prolonged downturn?

Introduction

There is substantial empirical evidence suggesting that value stocks outperform growth stocks, for example Fama and French (1998) and Brandes Institute (2007). Fundamental justifications for superior performance of value stocks are also ample. Frequently, arguments are rooted in behavioural aspects, such as the fact that growth stocks tend to be over-researched by analysts. Growth stocks even tend to be labelled as glamour stocks. Value stocks on the other hand, tend to be neglected stocks (so-called dogs) that have fallen out of favour by analysts and investors. High earnings growth that (some) glamorous stocks may have realized in the past years, tends to be extrapolated to (many) coming years. Moreover, growth expectations for winners within the peer group are used as anchor or reference for other (potential) growth stories within the group.

Aggregation of these elevated expectations can unveil unrealistic optimism, whereas on the opposite side the neglected stocks become over-pessimistically valued, by sheer lack of attention (and information).

Despite the evidence in favour of value stocks, there are – as always – a number of caveats. For one, neglected stocks tend to be small cap stocks and successful stocks tend to have a larger market capitalization. Hence it stands to reason that there is a significant correlation between size and performance on one hand, and value (versus growth) performance on the other hand. This begs the question whether value versus growth is the ‘true’ explanatory factor. Perhaps a value strategy is only a derivative of a small cap bias, the latter being a superior selection factor, or rather not even an alpha, but a small cap

Drs. Ivan D. Moen RBA (l)

Senior Fund Manager at Optimix Vermogensbeheer NV and Chairman of the TAA committee at the VBA Association

Drs. Jelte de Boer CFA (r)

Investment Analyst at Optimix Vermogensbeheer NV



beta, as small caps are riskier than large caps and demand a higher risk premium. There is extensive literature on value and size explaining stock returns, see for example Fama and French (1992) and Asness et al (1997).

A second issue could be that value strategies may falter when major developed economies will go through a prolonged structural economic slow-down. It is conceivable that we are at the start of such a regime shift, as it is becoming clear that we are currently undergoing a significant structural change in the macro-environment. Perhaps on aggregate value opportunities will turn out to be value traps when earnings turn structurally scarce in a phase of ever shrinking leverage possibilities. Even absent such a structural shift, growth stocks sometimes outperform over several consecutive years. The late nineties are a good example of such a period.

Because of the abovementioned caveats we (as active fund managers) are uncomfortable with a 'buy and hold approach' towards the value proposition as such. We require more reassurance on its robustness. For one, we all but rule out the possibility of a prolonged downturn. Secondly, in our profession we cannot afford the unconditional luxury of 'the long run'. We fear that Keynes' classic and poignant remark that in the long run we are all dead, may be all too relevant for us as fund managers (professionally at least, anyway...).

In the remainder of this article we will address how we dealt with the abovementioned issues. We broadly looked at the following three questions:

- How robust is a value strategy for shorter time-frames?
- Is a value strategy 'just' a derivative of a small cap bias?
- Is a value strategy a good idea during a prolonged downturn?

Selecting value stocks

We first need to address how we will screen for value stocks. There have been many (comparable) approaches for such screens. Different valuation metrics are used/combined, such as Price/Book, Price/Earnings, Dividend/Price, Enterprise Value/EBITDA, etc. It is beyond the scope of this article to dwell extensively

on the many different options. We will describe the screening we opted for, and briefly touch upon the rationale for preferring this particular method.

We used a (value) screening methodology first suggested in a book by Joel Greenblatt (2006), a book with the disputable title 'The little book that beats the market'. In his book Greenblatt screens stocks on two factors:

Return on Capital (ROC):	Earnings Before Interest and Tax (EBIT)/ Tangible Capital Employed (TCE)
Earnings Yield (EY):	EBIT/Enterprise Value (EV)

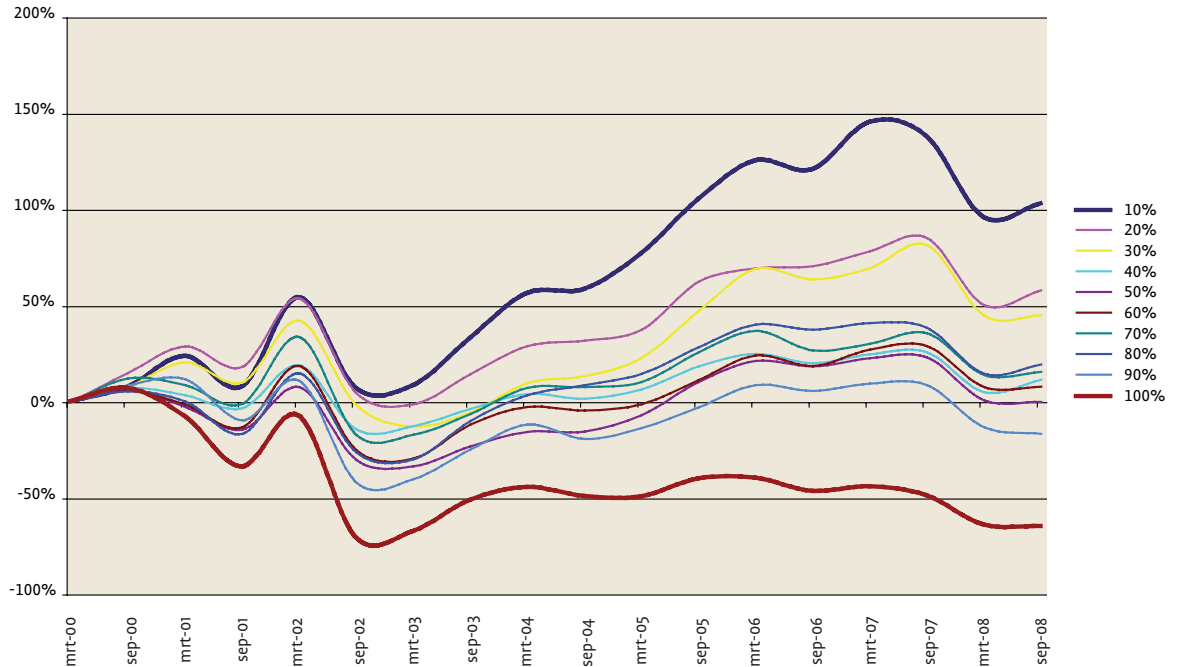
where,

TCE:	Net Working Capital + Net Fixed Assets
EV:	Equity Value + Net Debt.

We choose the selection criteria used by Greenblatt, as we believe they offer more information than other more commonly used value factors. For instance, as defined above, Greenblatt uses Return On Capital (ROC), rather than Return On Total Assets (ROA), or ROE (Earnings/Book value). ROC gives a better idea on the total amount of capital actually required to generate operational profits. This is also the main reason to exclude goodwill from this calculation, which is more a historical cost and needs no replacement to maintain future business. Also, Greenblatt makes use of a definition of Earnings Yield (EY) with EBIT (instead of operating earnings) and Enterprise Value (instead of equity value). These metrics are purposely employed to avoid any possible distorting effect from different tax and/or debt levels (i.e. leverage)¹.

Greenblatt ranks all available stocks on ROC and EY and proposes to select the top 30 (or so) stocks from the summed ranking procedure. We exclude financials and utilities from the screening. For these sectors the required data is either less relevant, or not obtainable/measurable from the regular accounting data we use in this method.

Figure 1: Greenblatt Decile Cumulative Returns



Backtest on the S&P 500

We used the selection criteria from Greenblatt and applied it to the S&P 500 (ex-utilities and financials), using Bloomberg data from March 2000 to September 2008. We purposely chose this period and the S&P 500 as a test period. For one, we wanted to evaluate the screen performance on the most researched universe we could think of. Secondly, the 2000-2008 period provides both a clear bear market period and a recovery period, which we liked for this test.

We ranked and divided stocks of the S&P 500 index into deciles. We used historic members and quarterly data and ranked/rebalanced every six months, starting in the year 2000 ultimo Q1. Figure 1 shows the results on the ranked deciles' cumulative performances (total returns) over the test period².

Figure 1 demonstrates that top-ranked stocks outperform (the 10% decile, the upper blue line), whereas the bottom-ranked stocks are the biggest underperformer (the 100% decile, the lower red line). Moreover, the deciles in between are fairly consistently scattered in terms of cumulative relative performance. This indicates that the screening metho-

dology is indeed a consistent explanatory factor for the relative performance of the deciles.

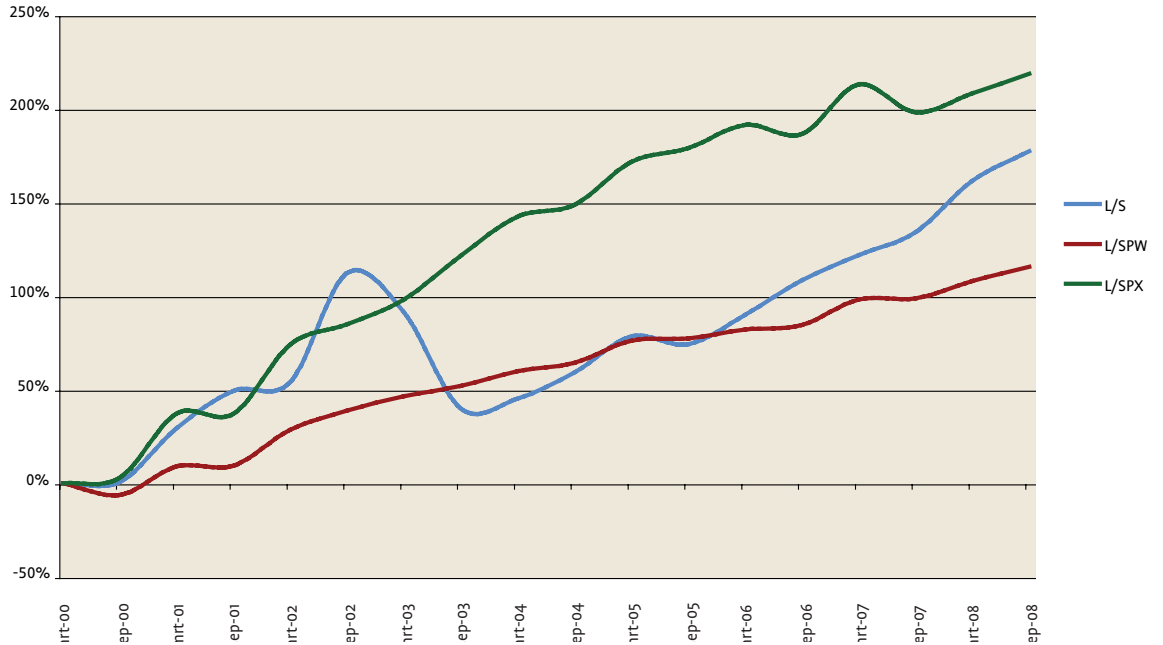
Figure 2 depicts how much a Long/Short Strategy with top and bottom deciles (L/S), would have actually outperformed, cumulatively. In similar fashion, the graph shows the performance of the top-decile versus both the Market Cap and Equally Weighted S&P Index (SPX and SPW, respectively).

What is reassuring is the extent of the consistency and stability in outperformance, especially from buying the top-decile versus the market (SPX and SPW). On the other hand, what is also clearly visible is the temporary underperformance of a L/S strategy during the recessionary years 2002 and 2003.

As a side note, SPW clearly outperformed SPX over the test period, indicating that small(er) cap did relatively well. Size mattered, at least for the overall performance of the S&P 500 during our test period.

To make sure the Greenblatt methodology isn't just a small cap screen, but is indeed a true value strategy, we wanted to know how much size mattered in

Figure 2: Greenblatt Decile Cumulative Returns



ranking our stocks. Therefore, we have calculated the rank-correlations between our stock-rankings and a ranking in stock market capitalizations. The results are in figure 3. A positive number indicates positive correlation between (small cap) size and (high value) ranking.

As can be seen in figure 4, there is little correlation between stock size and our screening, in any direction. In fact, most of the test period there is even a slight negative correlation, indicating that highly-ranked stocks tend to have a large cap bias. Hence, we can safely assume that our ranking for value, was not a small cap selection in disguise.

Figure 3: Correlation Greenblatt Ranking & Market Cap Ranking

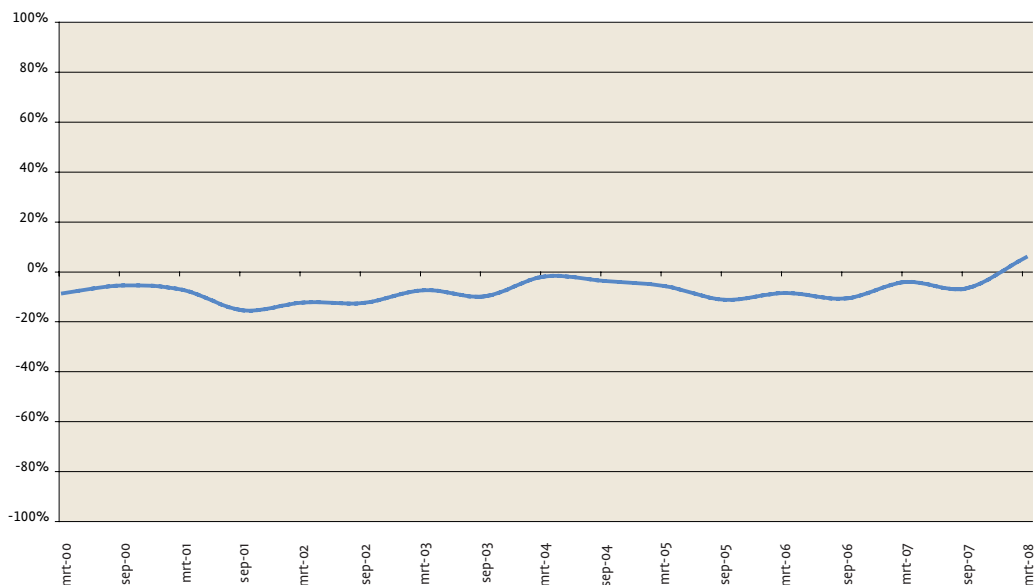
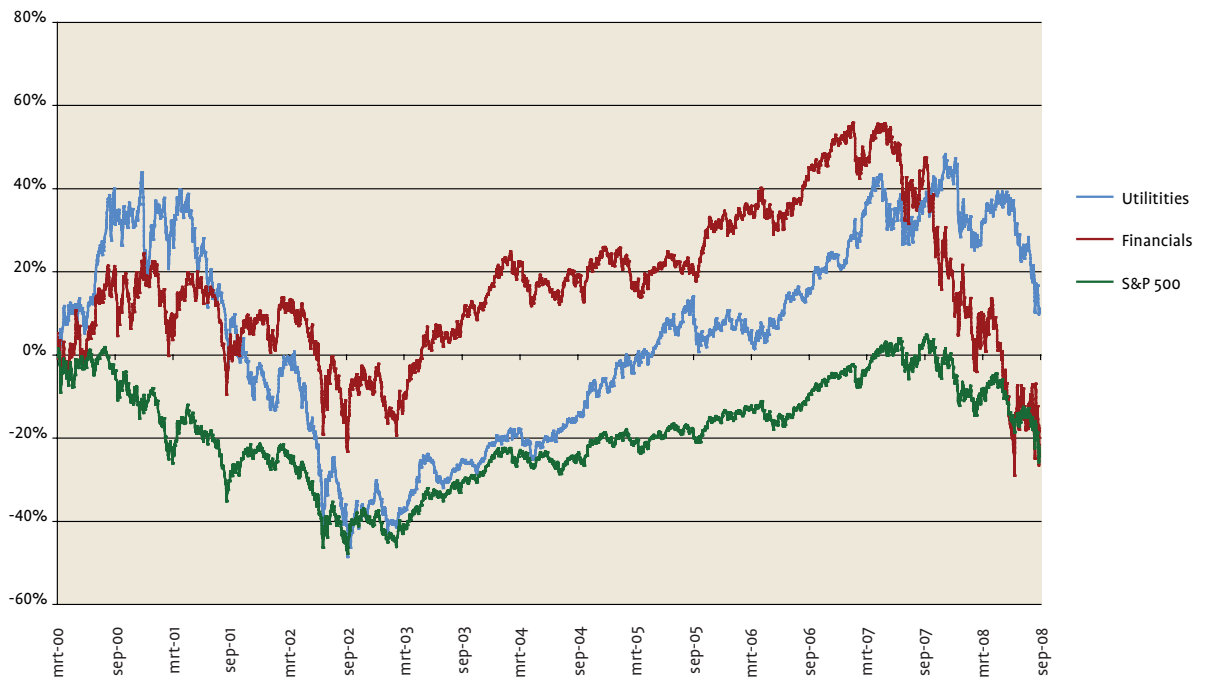


Figure 4: Performance Utilities & Financials Vs. S&P 500



Could our value outperformance be explained by the deletion of utilities and financials from our screens? We think not, because financials and utilities either outperformed or performed in line with the S&P 500 Index over the test period (see figure 4). So, if anything, omitting these stocks from our selection has probably been a drag on the measured outperformance.

We also checked sector weights during our test period. Figure 5, 6 and 7 depict the sector allocations over time for the S&P 500, the SPW (equal weight S&P 500) and for the top 10% picks from our screening (also equal weight) during the test period. We checked to see how big a sector tilt our selection had over time. The graphs show that (apart from the omission of utilities and financials) sector tilts were not structural, though sometimes substantial. For example, the weight of information technology in the screen is at times relatively low, healthcare is relatively big as is the sector consumer discretionary. We think the graphs show that there is no clear structural sector tilt, at least not in such a way that it could significantly undermine the validity of the screening method. We would argue that the dynamic sector tilts are a second order result from the screening, hence the screening method is the real

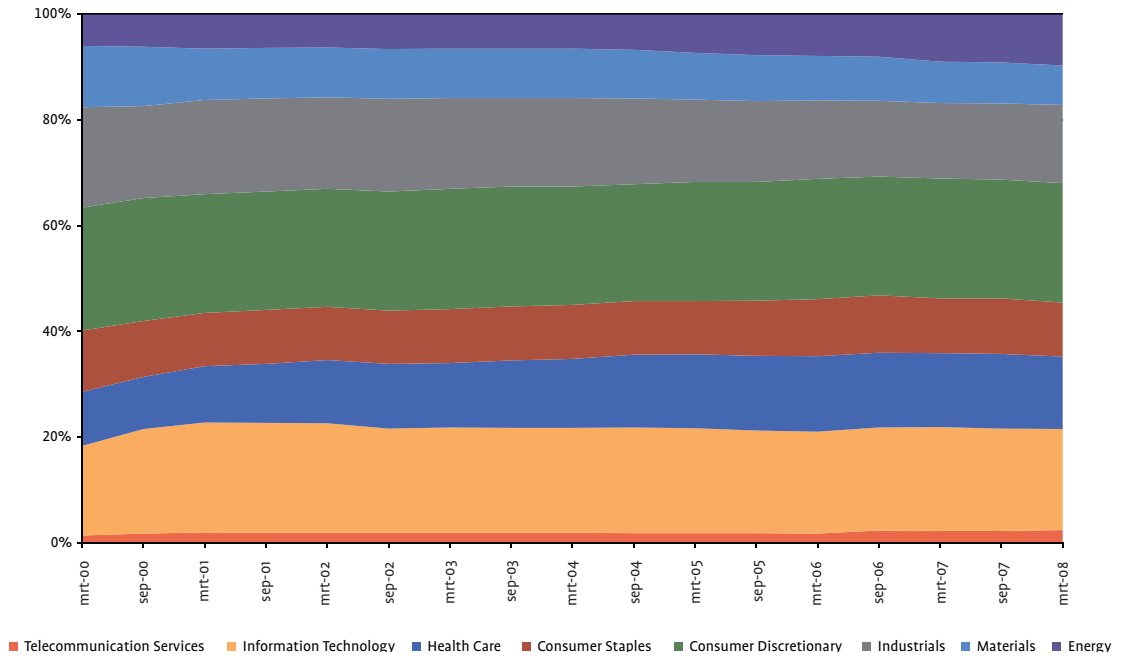
theme, and resulting sector weights are a derivative from the screening.

To sum up so far, we have seen evidence for consistent outperformance of value stocks over the test period for the S&P 500 Index. This provides us with at least some reassurance that a value strategy –under certain conditions– can perform satisfactory, even over shorter timeframes. We have also seen that in this period, small cap stocks have outperformed large cap stocks, so to be sure, we established that there is no significant link between our value selection procedure (Greenblatt) and a stocks’ market capitalization (i.e. no significant small cap bias).

This leaves us with the third and final question, is our value strategy viable in a prolonged downturn? In order to test this, we needed a good proxy for ‘a prolonged downturn’. Given the much debated potential parallel between Japan and the current (global) macro-economic circumstances, we figured that it would be appropriate to test our strategy with data on Japanese stocks during the 1990s.

We will not dwell on the matter of legitimacy of this parallel. Suffice to state that we think that it is not

Figure 5: Sector Composition S&P 500 (excl. Financials & Utilities; capital weight)



implausible that future events could fold out in a broadly similar fashion to what we have witnessed in Japan. We assume the parallel will at least function as a robustness check on our strategy, even if matters turn out differently.

Japan equity in the 1990s as a robustness check

We performed the same backtest as we did on the S&P 500 on Nikkei data (from most part of the nineties). We collected constituent data starting in 1992³

Figure 6: Sector Composition S&P 500 (excl. Financials & Utilities; equal weight)

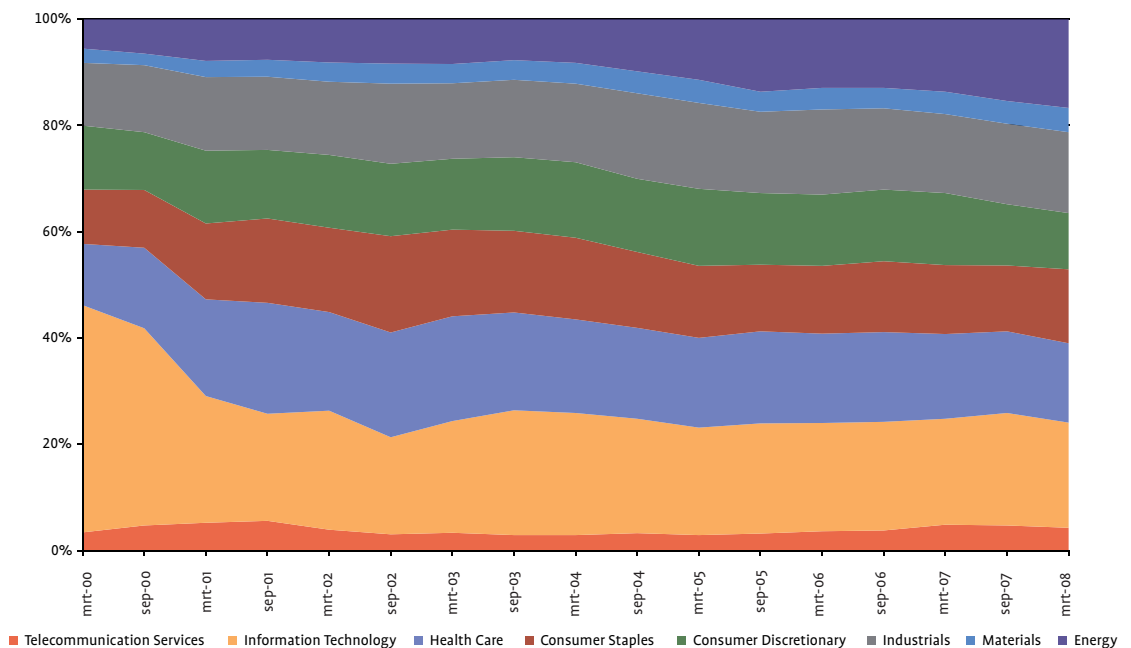
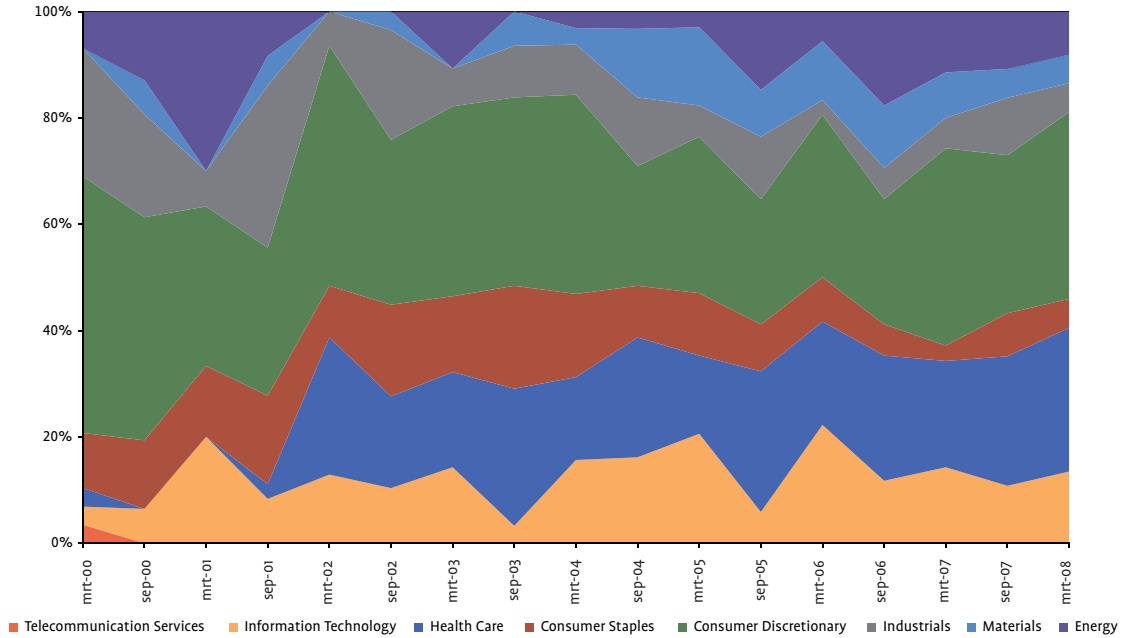


Figure 7: Sector Composition Top Decile Greenblatt Basket



(using calendar year data, rebalancing and ranking every 12 months) until 2001. Figure 8 depicts the resulting decile performances for the Nikkei index.

What can be seen is that the value ranking also works for this data set. The top 20% of stocks according to this ranking method clearly outperformed

and although the bottom decile (the red line) was not the worst performer, the bottom-ranked 30% in our value-ranking was on aggregate the weakest performing part.

Figure 9 on Nikkei data, depicts how much a Long/Short Strategy with top and bottom deciles (L/S),

Figure 8: Greenblatt Decile Cummulative Returns

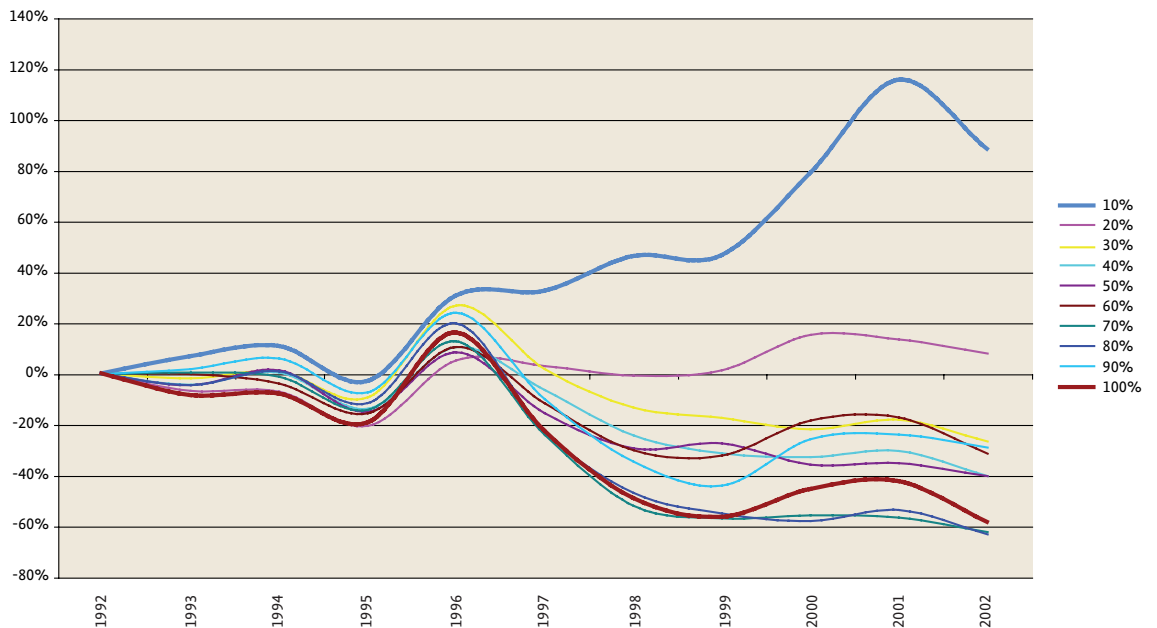
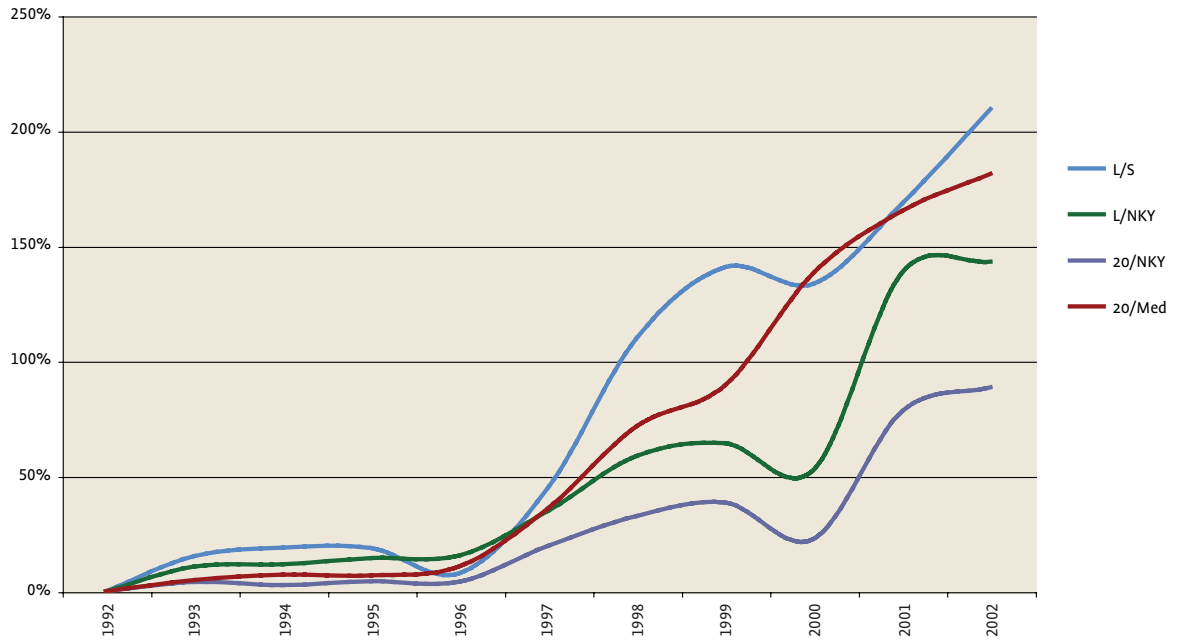


Figure 9: Greenblatt Decile Cumulative Returns



would have actually outperformed, cumulatively. In similar fashion, the graph shows the performance of the top-decile and the top 20% versus the Nikkei Index. Also the top 20% versus the median stock is depicted. The graph makes clear that in the 1990s a

value strategy in Japan would have made significant money, but that there were more 'rough patches' compared to our results with the S&P 500. Still, especially the L/S Strategy result remains impressive.

Figure 10: Performance Electric Power & Gas, Insurance, Banks & Small Caps Vs. Nikkei 225

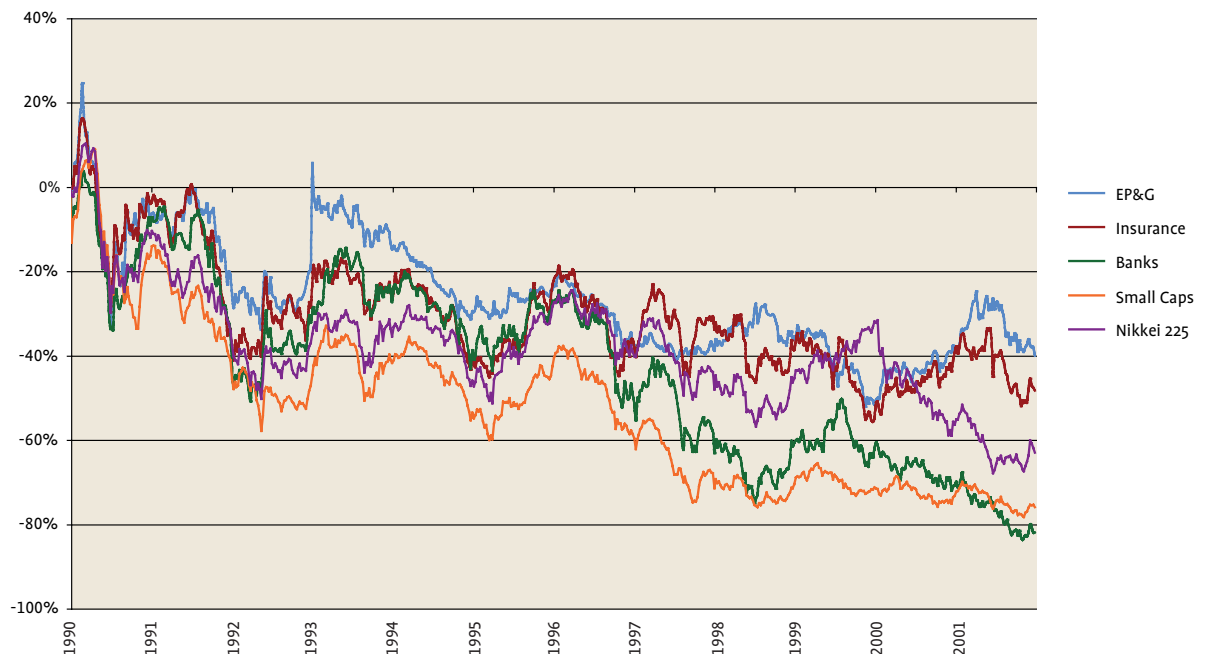
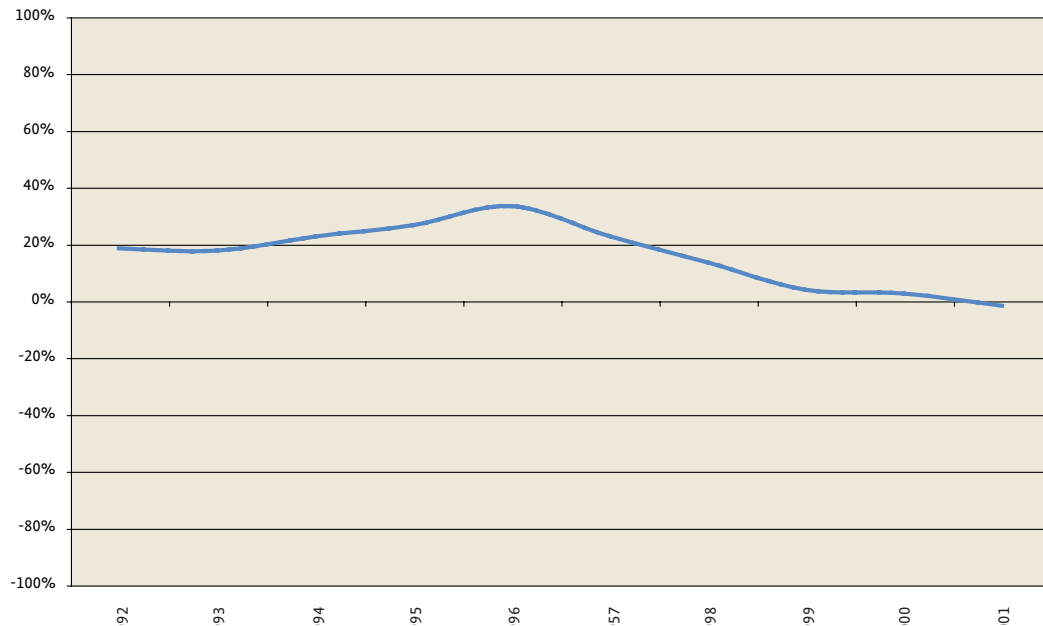


Figure 11: Correlation Greenblatt Ranking & Market Cap Ranking



Like with our S&P 500 backtest, we want to make sure that our value-basket outperformance for Japan is not dominated by the structural omission of utilities and financials from our selections. From the same perspective we looked at small cap performance over the test period. Figure 10 illustrates the relative performances of utilities (proxy-ed by the Topix Electric Power & Gas sector, TPELEC), financials (proxy-ed by the Topix Insurance and Banks sector, TPINSU and TPNBNK), and small caps (proxy-ed by the Topix Small Caps Index, TPXSM).

The graph (perhaps unsurprisingly) shows small caps underperformance to the Nikkei during these economically challenging years. Hence any bias towards small caps in our screening method is -if anything- a drag on the back test performance. Given that utilities and insurance outperformed the Nikkei, its omissions from our selection were also a drag on the (out)performance of the value screen. The opposite is true for banks. Banks underperformed the Nikkei by about 22%. So the omission of the sector helped the performance of our value screen. All in all we think that these results on aggregate do not undermine the validity of the outperformance of our Value screening methodology.

Like with the S&P 500 data, we calculated the rank correlations between our Nikkei stock-ranking and stock market capitalization ranking. The results are in figure 11. A positive number indicates positive correlation between (small cap) size and (high value) ranking. As can be seen, there is some correlation between stock size and our ranking procedure, but not in any disturbing quantity.

Conclusion

We were looking to answer three questions. First, how (un)stable, or robust is a value strategy? Using the S&P 500 data from 2000 onwards, we found a notable consistency in value performance, applying the Greenblatt screening method. Second, we wanted to make sure that our performance was not the result of a small cap bias in our selection process. Given the low correlation we found between our ranking results and market capitalizations, this could be dismissed as well. Finally, we wanted to check for robustness during a prolonged downturn by using data on Japan equity during the nineties as a (worst case) template for our value strategy. The consistency in the outperformance of the value screens did suffer somewhat (compared to our S&P 500 outcomes). Still, the remaining outperformance and

overall results provided us with sufficient comfort to allocate risk budget from our own Funds to this type of strategy. We applied (a proprietary variant of) this screening method, to the Optimix America Fund, starting in September this year. This generated outstanding results thus far in relative performance. The Fund is outperforming *all* peers across Europe this year, to date⁴.

Notes

- 1 For more elaborate details, see Greenblatt, J. (2006), "The Little Book That Beats The Market", John Wiley & Sons Inc., New Jersey.
- 2 We also tested with a 1-month time lag, to avoid any backfill bias in the data. The results were however very similar. This more or less confirmed the reaction we got from Bloomberg, that the backfill issue on the quarterly data is a matter of a couple of days most of the time, and should therefore not affect the results much with six-months' rotation cycles. To apply a 1-month lag in the data, would therefore be too conservative.
- 3 We started in 1992 because we were not able to find data on historic members prior to this date. We only had yearly data as well, hence the yearly rebalancing and ranking.
- 4 See www.Morningstar.nl, Optimix America Fund in the Category US Large Cap Blend Equity.

Literature

- Asness, C, Liew, J.M. and R.L. Stevens (1997), "Parallels Between the Cross-Sectional Predictability of Stock and Country Returns", *Journal of Portfolio Management*, Spring, pp. 79-87.
- The Brandes Institute (2007), "Value vs Glamour: A Study of the Indices", The Brandes Institute Research Paper.
- Fama, E.F. and K.R. French (1992), "The Cross-Section of Expected Returns", *Journal of Finance*, Volume 45, pp. 427-465.
- Fama, E.F. and K.R. French (1998), "Value versus Growth: The International Evidence", *Journal of Finance*, Volume 53, pp. 1975-1999.
- Greenblatt, J. (2006), "The Little Book That Beats The Market", John Wiley & Sons Inc., New Jersey.