Response to AQR's Article Titled "Deactivating Active Share"

Antti Petajisto

June 18, 2015

Executive summary

All of the key claims of AQR's paper were already addressed in the two cited Active Share papers: Petajisto (2013) and Cremers and Petajisto (2009).

1) The fact about the level of Active Share varying across benchmarks has been widely known for many years. Its performance impact was explicitly studied and discussed in the first drafts of Petajisto (2013) back in 2010, and the performance results remained broadly similar. The reason for the apparent discrepancy is AQR's choice of summarizing results by benchmark, which effectively gives the same weight to the most popular index (S&P 500, assigned to 870 funds) and the least popular index (Russell 3000 Growth, assigned to 24 funds), which is not sensible as a statistical approach.
2) The issue about four-factor alphas varying across benchmark indices does nothing to change the fact that higher Active Share managers have been able to beat their benchmark indices. However, it does raise an interesting point about the four-factor approach to measuring performance, and in fact my coauthors and I wrote a long and detailed paper about this exact issue first in 2007 (published later as Cremers, Petajisto, and Zitzewitz (2013)).
3) AQR's researchers argue that there is no theory behind Active Share and they remain mystified by the differences between Active Share and tracking error. It is unfortunate that they have entirely missed the lengthy sections of both Active Share papers that discuss this exact topic: pages 74-77 in Petajisto (2013) and sections 1.3, 3.1, and 4.1 in Cremers and Petajisto (2009). The short answer is that Active Share is more about stock selection, whereas tracking error is more about exposure to systematic risk factors.

So clearly ignoring large and essential parts of the original Active Share papers is simply not the way to conduct impartial scientific inquiry.

Responses to AQR's Three "Main Results"

1) Large-cap funds have on average lower Active Share than small-cap funds

This is a well-known result, mentioned in Petajisto (2013) and routinely discussed in my presentations at least since 2010. Also Stahl, Thomas, and Simon (2011) and many others have made this point before.

I agree that it makes sense to compare large-cap funds against other large-cap funds when making comparisons. That is precisely what I did in my paper: I ran performance tests in two different ways to control for this. The first approach was based on sorting funds within peer groups (large, mid, small-cap; even as far as within 3x3 Morningstar style buckets). The second was based on a multivariate regression with dummy variables and separate interaction terms for large, mid, and small-cap funds.

2) Controlling for benchmark indices, Active Share has no predictive power for returns

As mentioned above, I ran my performance tests separately for different market-cap groups and even the 3x3 Morningstar style buckets. Both in portfolio sorts as well as in multivariate performance regressions we still see economically and statistically significant predictive power for Active Share. It is odd that AQR's researchers seem to be unaware of this evidence as it is discussed explicitly in Petajisto (2013).

In contrast, AQR's own finding in this area is highly misleading. In Exhibit 3, they show results indicating that Active Share has been positively related to returns for some benchmark indices and negatively related for others, thus arguing that there is no effect on average. What they do not mention is that Active Share is positively related to returns in the indices that are actually used by most of the funds in the sample: S&P 500, Russell 2000, and the value and growth versions of Russell 2000, altogether about 1,400 funds. The large negative relationship comes from S&P 500 Value, which only has 66 funds assigned to it! I believe most statisticians would prefer to weight a sample of 1,400 funds more heavily than a sample of 66 funds.

Obviously this kind of predictability tests matter only when the sample is large enough for meaningful statistical inference. This is precisely the reason why I chose to keep the buckets slightly larger and did not want to get to the level of individual indices. After all, the effect we want to control for is about broad buckets: large-cap funds have lower Active Share than small-cap funds. Statistically the best way to control for that effect is to run the tests at the broader level of large, mid, and small-caps (3 buckets), even though I also mentioned the 3x3 results just for the sake of robustness.

3) Over this sample period, small-cap benchmark indices had large negative four-factor alphas compared with large-cap indices, which is crucial to the statistical significance of the performance results

This does bring up an interesting question about performance evaluation methods. However, it is puzzling why AQR presents it as a new finding, because back in 2007 my coauthors and I wrote a very extensive academic paper about this exact topic (published later as Cremers, Petajisto, and Zitzewitz (2013)). The paper received some attention in academia, it won a few best paper awards at conferences, and it was cited already in Cremers and Petajisto (2009), so we have been very open about this issue.

However, this issue does not change the simple fact that Active Share has helped predict a manager's outperformance relative to his benchmark index, and that is presumably what matters most to an investor choosing between an active fund and an index fund with the same benchmark.

Instead, the issue is a much broader one about whether the four-factor model that has been popular in academic research assigns inappropriately large nonzero alphas to the entire large-cap segment and the entire small-cap segment. Cremers, Petajisto, and Zitzewitz (2013) argue that the four-factor model could be slightly modified to eliminate this counterintuitive result, but that is a lengthy topic worthy of a separate discussion.

Responses to Other Claims in AQR's Paper

4) Funds with even a tiny Active Share can significantly outperform a benchmark index

To illustrate their point, AQR's paper mentions an example where a fund manager picks 499 stocks out of the S&P 500 index, dropping with perfect foresight the worst performer every month from 1/1990 to 10/2014. Such a strategy would have outperformed the index by 93bp per year. If a

manager drops five worst performers every month, he would have beaten the index by a hefty 4.51%. So it looks like we should not expect almost any limit on outperformance regardless of the level of Active Share?

Not so fast. How easy would it be for a manager to accomplish such an outcome? Their time period is 298 months, so the chance that a manager would successfully exclude the worst performer every month is $(1/500)^{298}$. In other words, the chance is 1 out of $2x10^{804}$ (i.e., a number with 804 zeros at the end!). For comparison, imagine lining up individual grains of sand in a continuous line to go around the Earth (around 25 per inch), and trying randomly to pick the "right" grain of sand from this very long line. The chance of success would be 1 out of $4x10^{10}$. If you keep playing this game and pick the right grain of sand $5x10^{793}$ times in a row (793 zeros), that is the chance of replicating AQR's outcome. The case of excluding 5 worst stocks (to get higher alpha) is even less likely: it is 1 in $2x10^{3399}$ (3,399 zeros), which is even more preposterously hypothetical.

AQR can still argue that it is theoretically possible to outperform significantly with Active Share of only 1-2%, but the chance of an actual manager implementing one of AQR's illustrated strategies in our lifetime is rather slim indeed.

5) **Active Share of x does not imply that a fund's return is (1-x) times the benchmark return plus x times the return on the "active" holdings.**

According to AQR, a "related misperception is that managers with low Active Share must earn heroic returns on their small 'active' holdings to justify their fees."

In fact, this is precisely how a return decomposition works. If you split your portfolio into an index component and an active component, and your Active Share is only 10%, then you will need a 50% return on your active holdings to generate a before-fee active return of 5% for the portfolio. Actually AQR's formula does need to be corrected slightly: your portfolio return with Active Share of x is equal to the index return plus x times the return on active holdings (i.e., we should not multiply index return by 1-x because active holdings are always both long and short).

This is a simple mathematical fact, not a question of opinion.

6) **It is a great mystery, and possibly data mining, that Active Share predicts returns while tracking error does not.**

According to AQR, "It might be that Active Share happens to capture some critical feature of what it means to be active and we just do not know what it is. Theory would be helpful here, but there is none. (…) One explanation is that it may just be a spurious, data-mined result."

The short answer is that Active Share captures mostly stock selection, whereas tracking error captures also factor risk, and fundamental equity managers seem to have skill in individual stock selection but not in their bets on broader economic factors. This seems like a rather clear economic message.

The puzzle here is why AQR's researchers would make such a claim, even suggesting data mining as an explanation, after they supposedly read the two papers on Active Share. Both papers start by discussing stock selection and factor risk, and one of the contributions of the papers is to present a two-dimensional grouping of funds based on both Active Share and tracking error, explaining what the two measures capture, why they are different, and how much variation we empirically observe across mutual funds in these two dimensions. In fact this two-dimensional sorting of funds was the original motivation to write the first paper on Active Share, so that we could understand active management styles across mutual funds and identify closet indexers – return predictability was only

looked at later in the process. AQR seems to have entirely missed this original contribution of the Active Share papers.

7) **The performance difference between Stock Pickers and Closet Indexers could be easily datamined.**

To illustrate the ease of datamining this result, AQR's researchers explain that "we can use our same data sample to see how much the first letter of a fund's name "influences" performance. It turns out that 'K' funds, on average, underperform 'Q' funds by 2.3% per year, a difference similar in magnitude and statistical insignificance (t-stat = 1.64) to the difference between Stock Pickers and Closet Indexers."

However, this is a seriously misleading comparison.

The key to data mining is having a large number of predictors to choose from, testing each one of them, and then picking the "best" performer. There are 26 letters in the English alphabet, so this gives AQR's researchers 26 predictors to choose from. In reality, their dataset was much bigger: they could have picked the length of the fund name, first letter of its manager's name, last digit of its year-end AUM, or anything else to data-mine a preposterous result to use as an illustration.

In contrast, I only had one variable, Active Share, to test along with tracking error, because the original Active Share paper was going to discuss these two measures regardless of their empirical return predictability. Where are the 25 other uncorrelated measures of active management that AQR researchers suggest we could have picked just as easily? AQR struggles to mention a single such alternative measure of active management.

8) **No theoretical reason why higher Active Share should outperform.**

Actually there is, at least on a before-fee basis. Prior literature (e.g., Wermers (2000)) has found that the holdings of mutual fund managers on average have beaten the market. It would be very difficult for the managers with very low Active Share to account for this outperformance because they would have to have remarkable skill in their small active positions to generate large alpha relative to their entire portfolio, which is even less plausible because we would simultaneously have to assume zero to negative skill for the more active managers. If a manager with very low Active Share had completely superior skill, that would presumably give him incentives to increase his Active Share further (and move out of this category) to earn even more impressive returns.

Hence, it seems intuitive that in a world where active managers on average outperform before fees, there would be a positive relationship going from very low Active Share managers to higher levels of Active Share. Of course this relationship does not have to be monotonic, so the moderate Active Share funds could indeed outperform high Active Share funds.

Two caveats should be mentioned here: First, despite their historical outperformance, it is not guaranteed that active managers will continue to outperform their indices before fees – this varies from year to year and the expected value may shift as well. Second, performance after fees can be negative even if before-fee performance is positive.

9) **No explanation regarding the source of the alpha of the high Active Share managers. Who is underperforming and why?**

This is an interesting question. But as the AQR authors should know, most of the papers in the literature do not address it, so if not addressing this question is a fatal flaw, then most of the published papers in the field are fatally flawed.

One potential clarification: U.S. all-equity mutual funds are not a closed system, they only account for about 17% of the equity market in my sample, so one group of mutual funds outperforming the market index does not imply that another group of mutual funds is losing the exact same dollar amount. Some of the positive alpha might come from retail investors, as suggested by Cohen, Gompers, and Vuolteenaho (2002), but properly answering that question would be a much broader research topic probably deserving a separate paper.

References

Cohen, Randolph B., Paul A. Gompers, and Tuomo Vuolteenaho, 2002, Who underreacts to cash-flow news? Evidence from trading between individuals and institutions, *Journal of Financial Economics* 66, 409-462.

Cremers, Martijn, and Antti Petajisto, 2009, How active is your fund manager? A new measure that predicts performance, *Review of Financial Studies* 22(9), 3329–3365.

Cremers, Martijn, Antti Petajisto, and Eric Zitzewitz, 2013, Should benchmark indices have alpha? Revisiting performance evaluation, *Critical Finance Review* 2, 1-48.

Petajisto, Antti, 2013, Active Share and mutual fund performance, *Financial Analysts Journal* 69(4), 73-93.

Stahl, Kent, Gregg Thomas, and Tom Simon, 2011, Active Share: Predicting alpha and risk, Wellington Management.

Wermers, Russ, 2000, Mutual fund performance: An empirical decomposition into stock-picking talent, style, transactions costs, and expenses, *Journal of Finance* 55(4), 1655–1703.