# My Factor Philippic[1]

**Clifford S. Asness**
**Managing and Founding Principal**
**AQR Capital Management, LLC**
**This Version: July 14, 2016**

**Please do not quote without permission**

### *Abstract*

*Arnott, Beck, Kalesnik, and West (2016) (ABKW) study smart beta or factor-based strategies and come to the following conclusions: (1) Aside from value, most popular factor strategies currently look expensive. (2) These expensive factor valuations portend lower future returns and a strong possibility of a future "factor crash" in which they go "horribly wrong." And (3) many of these non-value factors were never real to start with because their historical performance was due to factor richening. That is, researchers mistook the one-time returns from factor richening for truly repeatable "structural alpha." ABKW's implied bottom line (their many protestations to only making modest recommendations aside): stick with value, dump the other factors. This essay elaborates on my response in Asness (2016). In summary: (1) I find non-value factor valuations moderately expensive, but not as expensive as ABKW. (2) I argue that ABKW exaggerate the power of factor timing by improperly using long-horizon regression techniques. More proper short-horizon regressions suggest some weak factor timing ability and given this predictability, I construct value-based tactical factor timing strategies to test them. Unfortunately, these strategies add little to portfolios that are already invested in the value factor. It turns out that this "newly" discovered timing tool is, yet again, mostly just a version of regular old value investing. And (3) I examine ABKW's claim that factor richening drives much of non-value long-term factor performance and find that this very serious allegation about other researchers' work is totally without merit. Overall, these results suggest that one should be wary of aggressive factor timing. Instead, investors are better off identifying factors they believe in, and staying diversified across them, unless we see far more extreme pricing than we do today.*

---

**I. Introduction[2]**

There is no question that smart beta or factor-based strategies have become increasingly popular in recent years.[3] In light of this, Arnott, Beck, Kalesnik, and West (2016) (henceforth ABKW) examine these strategies and come to provocative conclusions which, if true, have important implications for smart beta investors.[4,5,6,7] To facilitate this discussion let's break their essay into three parts:

(1) They start by examining recent valuations of a handful of popular smart beta and/or factor strategies. Valuations are a natural starting point in understanding whether these strategies have become too popular or crowded. Presumably, if this is the case, we would see the footprint of investor

---

[2] Warning, this is a somewhat harder quant geek slog than my norm these days. In particular, if you want to enter quantitative Gehenna, please delve into these footnotes which, admittedly, set a personal record for number, length, and the rare sight of four in a row in some places! In fact, one can treat this as, I think, a fairly tight readable piece by ignoring the footnotes, or really dive in and get the "full experience"!

[3] As usual I use "factors," "styles," "alternative risk premia," and "smart beta" (when judged versus the capitalization weight index) relatively interchangeably because, you know, they are.

[4] They seem to be implying their conclusions are particularly brave (my interpretation) given their history and commercial interest in smart beta. From their paper, "This provocative statement—especially by one of the original smart beta practitioners— requires careful documentation." But you don't get to imply you're speaking the truth despite your interests when you're essentially talking up the prospects for your main strategy while talking down those of others. You're certainly allowed to take that position but it should be written with disclosure not self-congratulations. In turn, I happily disclose that my commercial interests are consistent with my recommendation here of factor diversification over factor timing.

[5] After most of this essay was written Arnott, Beck, and Kalesnik (2016) (henceforth ABK) extended the written work of ABKW, and a webinar by Rob Arnott further added to the debate. I have not rewritten this essay to completely address these two new entries but will make comments on them in the footnotes (and occasionally this slips into the body text). Now I must be a bit petty and mention that the new ABK doesn't reference Asness (2016) (in the webinar they address a question about Asness (2016) and the answer makes this lack of reference more surprising) yet are clearly attempting to address my critique but with vague pronouns not citations. Ironically when asked in the webinar (around 37 minutes) about my views Rob Arnott went on about how ABKW was just a preliminary "working paper" and he viewed my piece as a very valuable "referee report" to help them make theirs better. Killing me with kindness doesn't work! And, it's only pseudo-kindness as "referee report" is still braggadocio as it makes all other work merely a shadow of and reference to theirs. Frankly, separate from the issue of attribution/citation, you don't get to make strong statements ("horribly wrong"! "factor crash"!) and then say something like "well it's only a working paper." If I sound somewhat testier in these footnotes dealing with this latest publication and the webinar then, well, you are paying attention.

[6] In particular in their footnote 1 the new ABK says "It bears mention that the previous article in this series stirred a certain amount of controversy. While we are amused at the hyperventilating in some of the reaction, we are deeply grateful for the feedback. To us, these critics serve as surrogate journal referees. Thank you." There was no hyper-ventilating in Asness (2016) (forgive me if I continue to think I'm their unnamed uncited target) but there is, admittedly, some hyper-ventilating here. Alas, if my refutations of their work have to be cast in the subordinate role of journal referee report, and then fobbed off with half measures and closed eyes to the most important critiques, at least let me give them a real referee report. Here it is: "To the authors – we will not be publishing this. To quote, perhaps apocryphally, Samuel Johnson: 'Your manuscript is good and original, but what is original is not good; what is good is not original.'"

[7] By-the-way, if you sit through the webinar and aren't irked by the repeated feigned surprised exclamations of "good gracious!" and "oh my goodness!" you're a better person than me. Greatly exaggerated and sometimes just plain wrong analysis is not made more palatable when delivered in a Mrs. Doubtfire impersonation.

flows in the form of expensive factor valuations or narrow factor "value spreads."[8] They find that, with the exception of the book-to-price factor, that these factors are currently expensive.

(2) Given that, they turn to the question of whether expensive valuations matter to future performance. To answer this, they examine regressions of future long-term factor returns on past factor valuation levels and find a significant relation. Their regressions suggest that when factor valuations are expensive, future factor returns tend to be lower – bad news for current factor investors. Furthermore, they argue by analogy that we've seen something like this before. In the late 1990's the world fell in love with the stock market and growth/technology stocks. At that time we saw extremely expensive stock market valuations and we all know what happened from there right? Is a factor crash coming? ABKW clearly think yes.

(3) Finally, and most provocatively, they argue that some of these so-called factors never really even had true positive expected returns (or "structural alpha" as ABKW call it) to start with. Why? Because the realized return to any investment can be influenced by valuation changes. To the extent valuation changes are not expected to repeat (or even worse, could reverse), any return that is due to these valuation changes shouldn't count as part of one's assessment of a factor's true expected return going forward. They argue that once you remove the effect of these valuation changes from historical factor returns, for many of the non-value factors, there is little left over. In other words, both investors and academics have failed to realize this important insight and as a result have unknowingly flocked into zero expected return strategies that have simply richened over time.

The implications of this work are seemingly clear. If you believe (1) and (2) above, then the expected return to non-value factor strategies are lower going forward and could even crash. Clearly, investors in these strategies should reduce their allocations. If you believe (3), investors should never have invested in the first place and certainly should run with their ill-gotten gains now! In this case, many non-value factor strategies have no true positive expected return and most likely no role in one's portfolio. In the case of (3) it's really not a matter of timing or valuation or "lightening up" as these factors just are not real to begin with.

Not surprisingly, I have many thoughts on this analysis and ABKW's conclusions. I share some of these in my invited [Journal of Portfolio Management (JPM) editorial](#) (Asness (2016)). However, the editorial doesn't show many actual results, as it's a space-constrained opinion piece. Rather, it promises upcoming supporting work. This note provides some of this promised support, and also serves to further clarify and expand on my thoughts (and comment on some follow-up to ABKW, the newer ABK in particular, in the footnotes). Here's a quick summary paralleling the three points above:

(1) I perform my own analysis of factor valuations using a slightly broader set of metrics than ABKW[9] and my findings agree directionally on which strategies look expensive and cheap versus history. However, I find considerably less extreme results than do ABKW.[10]

---

[8] They also offer a provocative alternative causality that the strategies have gotten popular because they have gotten expensive! Much more on this later.

(2) On the question of whether expensive valuations matter to future performance, I have several bones to pick with ABKW. First, I argue that their regressions of future long-term factor returns on past factor valuation overstate the degree of predictability.[11] One way to see this is to construct factor timing strategies that only use backward looking information (unlike the regressions). These strategies produce weaker[12] results than what you might expect from looking at the regressions. Second, while ABKW's headlines are about "crash" risk, I believe that predicting these crashes is quite difficult, and in particular, factor valuations are not great at forecasting them.[13] Finally, one implication of their analysis is that if the expected returns to these factors vary over time (as their regressions suggest), investors should use this information to tactically time the amount of factor exposure in their portfolios. When factors look expensive/cheap, they should hold less/more. However, this seemingly obvious implication isn't really obvious as it misses an important insight. It turns out that a value-based factor timing strategy is highly correlated with the value factor. In other words, naively adding such a strategy adds little (to a portfolio of factors) and leads you to double down on your value exposure. In fact, the title of their latest paper screams this: "To Win with 'Smart Beta' Ask If the Price is Right." Price means value! If you allow value to be an important stand-alone factor <u>and</u> to decide how much of the other factors you want to bet on, you're doubling up on value versus the other factors. That doesn't mean you should do no contrarian timing, but it does mean by definition it's less attractive to an investor already significantly exposed to the value factor. Investors who ignore this correlation will likely make major portfolio construction errors.

(3) On the topic of whether these factors ever had real positive expected returns or "structural alpha," the short answer here is that I strongly disagree with ABKW and think, well, that they are simply doing it dead wrong and obviously so. That's not to say they don't bring some valuable insight to the table. I agree with their statement that people should not expect past returns due to valuation changes to repeat in the future, and I agree it's important to remind investors of this important fact.[14] ABKW present two methods of gauging how practically important this is to our long-term analysis of the major factors. One version shows a large impact and the other a virtually non-existent one. The latter method is reasonable (if only an estimate). However, the former method which shows a large impact is, frankly,

---

[9] In my JPM editorial I mentioned this but didn't discuss specific results.

[10] The new ABK extends this further from ABKW's one to my two to ABK's four valuation metrics. In fact they note that, unlike using only book-to-price when they concluded value was very cheap, when they use multiple measures "value is moderately cheap in the United States relative to historical norms." The downgrade to "moderate" is basically the same thing I find and discussed in Asness (2016). Yet, there's no mention of that, and their verdict of "moderately" has a tone more of "confirmation of expensive" than what it really represents: a downgrade in current reading and thus importance from ABKW to ABK.

[11] The new ABK thankfully includes more valid shorter horizon regressions yet still interpret them as far newer and more powerful evidence than they deserve (we knew these would "work" as value works!).

[12] Especially in light of their correlations as discussed below.

[13] We are likely interpreting the word "crash" differently and I will try to be clear about the differences below.

[14] This is a battle I have fought many times, particularly back in the [tech bubble days](#) but also [since then](#). That is, if over some period something got more expensive this likely added to its realized return (how much will become an important issue where we differ as discussed below), and this should not enter into our estimate of its future return (as it would mechanically if we just measure past average returns and project them out into the future). Rather, at least to some degree (again this "degree" will become a central issue later in this essay), it might subtract from the expected return going forward.

ludicrous and with mathematical certainty greatly overstates the effect. It should not even be considered. To ABKW's credit they point out some of the weaknesses with this rather silly first method. But the credit stops there and turns to opprobrium as, unfortunately, they repeatedly ignore the reasonable method to only quote the ridiculous one (to the extent of making this the sole method presented in their webinar).[15,16] Of course, the ridiculous method also greatly overstates the point they desire to make.

I'm going to organize this discussion along the lines of the above summary of ABKW. That is, I'm going to go through each of the above three parts, summarize ABKW's analysis and conclusions and offer my own. However, before I do that, I need to lay out some methodology and definitions. That I'll do in the next section (II). From there, Section III will address the question of whether factors are currently expensive. Section IV will address whether one can use factor valuation to predict future factor performance. Finally, Section V will address whether these factors really have positive expected returns or structural alpha or are simply the result of one time long-term factor richening? After that, I'll offer some concluding thoughts.

## II. Methodology

In what follows I will look at five long-short factors for U.S. stocks covering four themes:

- the value factor (done two different ways, using book-to-price, "B/P", and sales-to-price, "S/P");

- the momentum factor;

- the profitability factor; and

- the betting-against-beta factor (BAB).

I've done this following the methodology of Fama and French (1993) in the construction of their famed HML value factor, but only among large caps. While Fama and French average the results for large and

---

[15] In a joint panel Rob Arnott and I did together he alluded to a third forthcoming method. I'm truly open to new results.

[16] Another positive I appreciate in their work is I think they offer an excellent rejoinder to those, and there are many, who point to the recently disappointing performance of the book-to-price factor (probably the canonical "value factor") and take it as evidence that the factor doesn't work anymore as "everyone knows about it." There is a popular perception that once a factor is known it simply stops working. ABKW show us that, in fact, the exact opposite has happened to the book-to-price factor. The factor has cheapened over the last decade hurting performance. That's an indication that its problem over this time period is not being too popular or well-known but that it remains currently unpopular. Many get this exactly backwards and ABKW are exactly right, at least in direction. ABKW's bête noir is factors fooling us into thinking returns from their getting more expensive are repeatable (or from getting cheaper are damning). I will argue they are woefully mistaken about this mattering at the long run. But at shorter horizons, like a decade, they are right that it can matter a lot. In fact applying the regression techniques to be introduced later in this paper (mimicking one I think ABKW use) we see that the negative value factor return over the last decade is more than all due to the factor cheapening.

small stocks, in this initial pass at factor timing I focus on results for only large cap stocks as this is most relevant to large investors and this data is where we have the most confidence (Asness (2015, 2016) and ABKW make a similar choice). Within large cap stocks, each of these five factors is capitalization weighted long the 1/3 best stocks on the respective measure and short the 1/3 worst (e.g., for book-to-price the factor is long the 1/3 of large cap stocks with the highest B/Ps).[17]

Next I follow Asness et. al. (2000),  Cohen, Polk, and Vuolteenaho (2003), Asness (2015), and ABKW in looking at what Asness et. al. (2000) coin the "value spread."[18] This spread is the ratio of the valuation of each factor's long portfolio divided by the valuation of its short portfolio. Valuation of a factor's long or short portfolio is defined as the capitalization weighted average of its valuation characteristic (I look at this using B/P and, separately, S/P ratios).

In equations for the B/P factor (others defined analogously):

B/P factor return = return of 1/3 highest B/P stocks - return of 1/3 lowest B/P stocks

B/P value spread of B/P factor = (B/P of 1/3 highest B/P stocks) ÷ (B/P of 1/3 lowest B/P stocks)

S/P value spread of B/P factor = (S/P of 1/3 highest B/P stocks) ÷ (S/P of 1/3 lowest B/P stocks)

For the value factors themselves, measuring the spread using the corresponding valuation characteristic (e.g., the B/P value spread of the B/P factor), the ratio is always by definition over 100% (and, while not by definition, practically speaking, always well over 100% when the B/P portfolio is measured using the S/P value spread and vice versa). This is meant to be a measure of how "cheap" or "expensive" the factor is through time, or put differently the factor's value (for the value factors this occasionally entails the inelegant phrase "the value of value").[19]

---

[17] The betting against beta factor differs from the others. While the others are straight differences on $1 long versus $1 short, the betting against beta factor includes a hedge ratio based on estimated shrunk betas following Frazzini and Pedersen (2014) (though I used only 20% shrinkage here as large capitalization betas are likely more precisely estimated – results are not sensitive to this choice). ABKW also look at the low beta factor but it's not clear to me whether they adjust for the beta differences between the long and short portions (not adjusting would introduce a huge negative market beta to the long-short portfolio and greatly muddy the results). Also the valuation factors follow the method of Fama and French except they use up-to-date price as studied by Asness and Frazzini (2013). Finally, the profitability factor here follows Novy-Marx (2013). It is defined as the ratio of a firm's gross profits (revenues minus cost of goods sold) to its assets. See Appendix A for more detail.

[18] In a fairly odd passage ABK say "The green line in Panel A traces a new (as far as we know) phenomenon that has received little or no attention in the academic or practitioner literature: the trajectory of relative valuation, measured by P/B, of the value portfolio relative to the growth portfolio." Uh, no, that's just a version (the reciprocal) of the line from Asness et. al. (2000) Exhibit 5 Panel B a paper which is in ABK's references. And it's not just Asness et. al. (2000). There have been multiple papers after that using the value spread. I'm sorry if pointing out "we did it first" seems petty but I'm not sure what else to do when they have an explicit passage about their doing it first! I guess I could rise above it and ignore it completely. Maybe in my next life…

[19] Buckle up! As is probably obvious, I risk confusion as I use the word "value" to describe both the B/P and S/P long-short strategies and as an ex-ante value spread measure (measuring if these two B/P and S/P strategies, and the other non-value factors momentum, profitability, and BAB, look currently more expensive or cheaper versus

Correspondingly, these value spreads may measure how the expected return on the factor moves through time. It seems intuitively obvious that, all-else-equal, as more investors attempt to exploit a factor it could get more expensive through time and this ratio or "value spread" would fall perhaps leading to a lower expected return going forward (and vice versa if a factor becomes unpopular).

### III. Are factors expensive today?

As stated above I (mostly) agree with ABKW directionally on which strategies look expensive and cheap versus history – though by including just sales-to-price as an alternative measure I find considerably less extreme results than they do.[20]

To help visualize all this starting with the value factor, below I graph four spreads (the B/P and S/P value spreads of both the B/P and S/P factors):
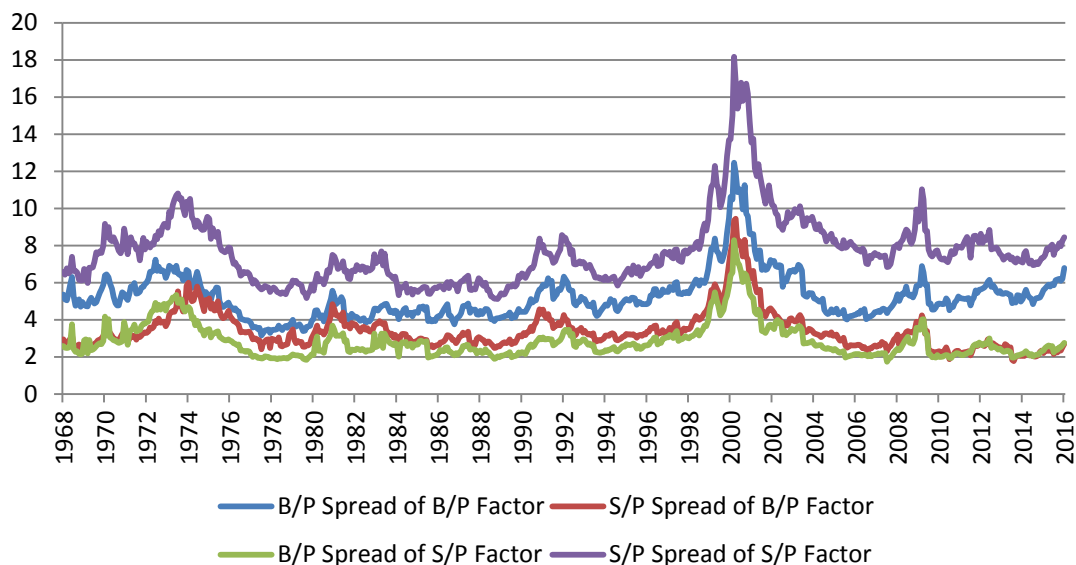


Source: AQR.  See Appendix A for more detail.

As discussed above, each is always well over 100% as expected. Each peaked at unprecedented levels at the top of the technology bubble in approximately March of 2000. The B/P value spread of the B/P factor is currently (as of January 2016) at the 92nd percentile (so it looks "cheap" as higher is more

---

history). In geek speak value strategies (long-short returns) will appear on the "left-hand side" of regressions in this paper and value spreads, or the change in value spreads, will appear on the "right-hand side" of some of these regressions. So when I say "value" I'm using it in more ways than Papa Smurf. Hopefully it will be clear in context.
[20] Where I extend ABKW from one to two value measures, the new ABK uses four valuation measures and finds similar results to my two (less extreme than the just book-to-price used in ABKW with only the low beta factor being very similar regardless of value measure or measures used).

attractive), but that level is still only 54% of the tech bubble peak.[21] The other spreads are not so wide. The S/P spread of the B/P factor actually ended not cheap but at an expensive 24th percentile (only 29% of the tech bubble high). The B/P spread of the S/P factor ended at the 59th percentile (33% of the high). Finally, the S/P spread of the S/P factor ended at the 77th percentile (47% of the high). The average percentile across the four (two value spread measures for two value factors) is 63rd. This is indicative of my view that value, while being a well-known strategy and implemented in many very popular vehicles as "smart beta," has not been arbitraged away[22] but also is not looking as extremely cheap as ABKW argue.

Using the same methodology averaging the two value spreads, profitability is 49th percentile (note there is a large difference here on the book- vs. sales-to-price spread — it looks very expensive on B/P but very cheap on S/P).[23] Only momentum and BAB look significantly expensive averaged across both valuation spreads (18th and 9th percentile respectively). However, as we'll see below these are the factors where valuations make the least sense as a predictor of returns and have the least empirical power to explain even contemporaneous returns (that is, if you were told it got more or less expensive in a year, that's far less information for these factors than for slower turnover ones). I'd also note that even 18th and 9th percentile ain't exactly "bubble" levels.[24,25,26]

---

[21] With a positively skewed series like this, percentiles can be misleading − the 92nd percentile usually makes you think of something closer to the peak (and much scarier!). ABKW don't point out this relying too much on citing percentiles.

[22] If so we'd expect to see extremely low ratios versus history not very mildly above median ones.

[23] In future work I hope to get more "economic" about what it means when, in general and in the particular case of book- and sales-to-price, we see such different readings. Clearly some theories about profit margins will enter in. I would note that over this sample sales-to-price is a slightly more effective stand-alone factor and slightly more effective timing strategy for the other factors (probably the same result in different forms); though admittedly I wouldn't draw strong conclusions from such data. ABK find that the profitability factor using their composite valuation indictor is only 40th percentile expensive (i.e., slightly cheap), so we are close to agreement here, though they can't help leading with the observation that it looks "drastically extended" on only book-to-price (as used in the original ABKW paper). If I may summarize for them more accurately: "We (ABK) find profitability mildly cheap but we still love the sound bite from the prior paper so much we're still giving it pride of place and exclamatory adjectives." They've lost their factoid that profitability is very expensive versus history but they just can't quite bear to part with it. Also in their webinar they apparently had not yet gotten to the results of ABK as they make strong statements about how expensive is the profitability factor when, by their own work in ABK, profitability is mildly cheap on the better broader measures. Never let new facts, even your own, get in the way of a good story! It's possible during the webinar they didn't yet know but ABK still undersells this downgrade of the "profitability is expensive" story rather than completely walk it back as they should have.

[24] I remind readers that the standard for calling something a "bubble" should be very high (see peeve #2). All these results are within the range of historical experience not gigantic new highs or lows. In contrast the "value of value" (the average value spread done both ways on the two value factors) in March of 2000 was 163% higher than the prior maximum of such spread up to the end of 1998 before the tech bubble took off (not all 100th percentiles are equal!). The current situation, at least when it comes to valuation, is not even vaguely comparable to the tech bubble for any factor examined and that analogy should be avoided rather than embraced by ABKW.

[25] For the record, as a vocal proselytizer for the value factor for more than a quarter of a century, I, of course, wish these two factors looked cheaper. I just don't think ABKW or ABK have come close to making the case that these valuations are damning and not already accounted for if one is already significantly exposed to the value factor (more on this overlap to come). Furthermore, if these factors are diversifying from the value factor, as is part of

Given their popularity and the amount of discussion and attention they currently generate you might expect the factors (and thus what underlies Smart Beta) to be priced to historical extremes. They're not. Frankly I was a bit pleasantly surprised. "Value" factors actually look somewhat cheaper than normal, some non-value factors are somewhat more expensive (though not profitability), and none look well outside prior norms.[27] Perhaps it's surprising that things look so normal given all the talk about smart beta and factor investing – but the prices are what they are.

**IV. Can you use valuation spreads to predict factor performance and "crashes"? And, if so, how much does it really help your portfolio?**

*5-Year Overlapping Regressions*

To answer this question, ABKW don't actually look at implementable rolling valuation-based timing strategies using only information actually available each period.[28] Instead, they look at long-horizon regressions that copy what's done for the stock market (e.g., forecasting long-horizon S&P 500 returns using starting Shiller CAPE or dividend yield). Here I argue that this approach is anywhere from somewhat inappropriate to irrelevant for these higher turnover factors and simply copying what's done for the stock market, without understanding why it's applicable there, is dangerous.

ABKW mostly look at R-squareds and t-statistics from regressing overlapping 5-year factor returns on starting valuation levels which rely on full-sample information to judge valuations. Perhaps this explains part of their relative optimism versus my relative pessimism about timing? Asness, Ilmanen, and Maloney (2015) (henceforth AIM) discuss, in the context of timing the S&P 500, various ways you can find seeming predictability for long horizon overlapping returns. Unfortunately AIM also shows that these techniques yield much less power for real-life trading strategies. I refer you to that paper for

---

their point, times when they are expensive may be when they are most important. After all, this also likely means the pure value factor is oriented more against momentum and towards less profitable companies than usual. I do not understand why ABKW's causality and worry only runs one way.

[26] In other talks Rob has referred to the sound bite that several of the famous FANGS (Facebook, Amazon, Netflix and Google) make it into low volatility and/or low beta portfolios (the low beta portfolio makes much more sense as the FANGS are far more likely to have low correlation than low volatility). I'll take their word for it. He means it as quite damning as after all everyone knows the FANGS are doomed! Of course, their presence in either portfolio would be fully reflected in the high but non-bubble like valuations of factors like BAB so this may be a fun tidbit but not new information. More importantly, these stocks would get into such a portfolio not because they've been bid up, but simply because they had the requisite low betas or low volatilities. These factors make no claim to be valuation factors. Somehow the tone here is "low beta must be broken if a FANG is in there." No, if the FANGs have low volatility, or I suspect low correlation with the market, they get in there regardless of rich or cheap valuation.

[27] And we'll soon see that valuation is, for predicting returns, or even explaining contemporaneous returns, near useless for high turnover, or more generally what I'll soon label high "friction", strategies like BAB and momentum.

[28] Nor does the new ABK though they do, welcomingly, move to shorter horizon regressions. I expect and hope their next paper will do this and I remain quite open minded about finding these things more useful – in a true implementable strategy, not an in-sample regression of any horizon, and in a portfolio context. I have tried and not yet found them very useful but there's always hope!

specifics.[29] Here I focus on the additional fact that the 5-year overlapping regression method that ABKW use is far less applicable for long-short factor timing than even for market timing.

Those of us who've worked with five year overlapping period regressions (with Newey-West t-statistics) will recognize that the results in Figure 2 of ABKW are fairly weak. T-statistics in this type of regression that are barely over two (and sometimes not even) rarely lead to economically useful strategies (ABKW find these type levels for the successful ones, with some like low vs. high beta effectively zero, and only the small vs. large factor showing truly powerful results). In particular, a centerpiece of their paper is the attractiveness of value, and yet in Figure 2 they don't show a 5-year overlapping adjusted t-stat over 2.0.[30] Casual examination of the scatterplots in ABKW's Figure 2 would not cause panic today in non-value factor investors or elation in value factor investors. So, if I had to sum up the forecasting power here, it would be "weak", and that's before assessing the basic applicability of long horizon regressions.[31]

This applicability is severely wanting. The very exercise of using long term overlapping returns is deeply flawed for the long-short factors studied in ABKW. When using this technique, you need to study a portfolio that is relatively low turnover.[32] Why would you forecast returns multiple years out on a portfolio that would be completely different from today's using information only about today's portfolio? Furthermore, you need a valuation signal that changes relatively slowly (of course these tend to be related). Under these two conditions, regressions of short-term (say monthly) returns on starting valuation understate the cumulative forecasting power over time as the portfolio and the signal are close to the same for long periods. In this case, the forecasting power piles up over long horizons and

---

[29] AIM discuss market timing (of the S&P 500) using both valuation and trend measures. We conclude in that article that, as many would tell you, market timing is mostly an investing sin, but we'd still recommend investors sin a little! That is, in a very modest way, own more of the market when it's cheaper than long-term average (and when it's heading up recently but I'll focus only on valuation here). Why only in a modest way? Well, because unlike many other places they are used, value factors don't have a lot of breadth in this context. Timing a factor, in this case the market itself, is empirically very difficult. Many studies (including many of ours) find value-based long horizon predictability for the stock market. However, realistic out-of-sample valuation-based trading strategies are weak (weak like "can be flat for 50 years" weak). Having said that, viewed long enough, based on first principles, and in particular if you can combine them with trend/momentum (again not studied here), and potentially also factor carry indicators, we'd likely recommend investors do a little of this sinful timing. I would add that I'm continually puzzled by those who'd readily concede market timing to be extremely hard but somehow think factor timing should be much easier.

[30] Not to worry it gets a small upgrade to get over the 2.0 hump in the newer ABK! Also, while not discussed the fundamental indexing strategy itself is seen to be expensive (despite the B/P factor being cheap) in ABKW Figure 3 right column middle row (this could be consistent with B/P being cheap but not the other value factors in fundamental indexing). Again, not to worry, it looks a lot better in the newer ABK! (and in the webinar where it benefits from the convenience of measuring things versus median not mean).

[31] Besides weak I'd also call it "expected" as we go in knowing value is a good strategy – it should have some power, though likely weaker for the narrow issue of factor forecasting than for diversified stock selection (just as timing the stock market using valuation is, historically and I think logically, much harder than using value effectively for diversified stock selection).

[32] The newer ABK recognizes this and produces some interesting half-life figures for the factors that are similar in spirit to the R-squareds I will soon show. See other footnotes for more on ABK's welcome (though ultimately not conclusion changing) addition of shorter horizon regressions.

the monthly noise diversifies (Fama and French (1988)). These are reasonable assumptions for the market portfolio making, say, 5- and 10-year horizon regressions economically interesting (if certainly still presenting statistical challenges), but they are not reasonable for the long-short factors.[33]

To see this, let's compare the speed in which the valuation signal changes for the market and the long-short factors, starting with the market. Regressing, over our 1968-2015 time frame, the stock market's Shiller CAPE (actually 1/CAPE or it's E/P) on the same measure lagged three years yields a coefficient of 79% (i.e., rather slow regression to the mean) and a 62% R-squared.[34] This seems to intuitively satisfy the condition above – that the valuation measure used for long-horizon forecasting is slowly changing over time.

However, doing the same for the book-to-price spread of the book-to-price factor (the "value of value" and one focused on by ABKW) only yields a corresponding coefficient of 23% and R-squared of 6% (yes, the market's R-squared of 62% drops to 6% here!). Unfortunately, for this exercise, the book-to-price value spread of the book-to-price is not nearly as slowly changing as the market's valuation.[35] Thus, it's not just the high turnover factors where ABKW are stretching this long horizon technique but even for value itself. Of course it is far worse for the higher turnover factors. For the book-to-price spread of momentum (the "value of momentum") I find a coefficient of 0% and R-squared of 0% (not typos, they round to zero).[36] That is, today's richness or cheapness of the momentum factor is completely unrelated to this same measure three years from now – not exactly a surprise for annual momentum I admit. Clearly five year regressions are very iffy for the value strategy itself, certainly compared to the overall market for which the technique was designed, and completely and utterly useless for the momentum strategy.[37] Thus, using a valuation measure now to forecast returns in the future, when the valuation measure today has no bearing on the valuation measure in the future, makes little sense.

The new ABK paper does a better job including shorter term forecasts. But it still gets overexcited at rather so-so results found in many places. A fan of value investing should expect value to have forecasting power in these new regressions and not get excited each time it's rediscovered.[38] I'm

---

[33] ABKW actually note some of the problems with higher turnover strategies (though the problems are caused by more than just turnover as described below when I discuss "frictions"), but they proceed with the regressions largely undeterred and interpret them as evidence of timing power. In ABK they then tell us they knew these were silly all along.

[34] Using the market's dividend yield produces very slightly larger figures indicating very slightly slower mean reversion.

[35] We also failed to recognize this in Asness et. al. (2000) when we introduced the value spread. We only forecasted out three years (marginally better) but I'll still take my lumps here!

[36] The pattern is born out even regressing valuations on valuations one year ago. For the long-only market CAPE the coefficient of this year on last is 91% and the R-squared is 82%. For the B/P spread of the B/P strategy the coefficient is 70% with an R-squared of 49%. And for the B/P spread of the momentum strategy the coefficient is 8% and the R-squared is 1%.

[37] Examining the other long-short factors or using sales-to-price value spreads instead of book-to-price, I find results between these extremes and none are near the stickiness of the market's valuation (Shiller CAPE or dividend yield).

[38] A relevant rather breathlessly excited section from the new ABK is "Five of eight factors exhibit a linkage between starting aggregate valuation and subsequent return over the horizon commensurate with their respective

reminded of the original debate over fundamental indexing. When I proved it was just a simple value tilt, Rob would respond, "no, you don't understand, it also works in small cap." Then I'd explain so does value, and he'd respond, "no, you don't understand, it also works internationally." And so on.[39] This time they get very excited that value has the right sign for factor forecasting in most cases and is often significant. They treat each finding as if it's brand new information. They're not. Remember, value is a good strategy. The question is not, as they seem to think, whether it "works" in the right direction, or whether "price matters" but whether or not this new version of value adds to your portfolio or are you already accounting for this relationship through the value factor. More on this coming up.

*Crash Risk*

Shifting gears and turning to my predictions, I do believe that these factors have more "crash risk" today. However, I believe it for very different reasons, and mean "crash" in a very specific way.[40] Essentially, as I discuss here crash risk seems to be less a function of being expensive vs. cheap and more a function of how well-known and popular is a strategy.[41] Of course these may be linked but they don't have to be. You can be well-known and even popular and not be expensive. As a case in point take August of 2007 (please) when many "quant" strategies crashed and then rebounded intra-month. However, going into August 2007, valuations didn't send a strong signal in either direction. Intra-month at the lows, valuations did get attractive on many of the factors we study here – so perhaps valuations had some use in encouraging people to stay the course or even tactically add to these factor bets – but they had no efficacy at all in forecasting a crash before the fact.[42]

Should investor shun investments that have potential crash risk? Unless you can time them, knowledge that future crashes will likely occur is not a reason to avoid a good strategy (please note that the strategy has to be "good" including these future crashes!). The stock market will one day crash again, I assure you. Whether you want to invest in it should be more about its long-term total return and how

---

signal half-lives significant at a 1% level, two at a 5% level, and one laggard (low beta) at a 10% level. This is powerful evidence that valuations matter. The results outside of the United States are similarly compelling." I ask again – who in finance, familiar with the power of the value effect, would doubt that valuations matter for everything? But, again, that they matter for everything, doesn't mean they should be double and triple counted in forming a portfolio.

[39] See here if you'd like an (attempted) humorous take on the original debate that applies again here.

[40] In ABKW they title their paper "horribly wrong" and, quoting them, say "Are we being alarmist? We don't believe so. If anything, we think it's reasonably likely a smart beta crash will be a consequence of the soaring popularity of factor-tilt strategies" and begin their paper with a long introductory parable about the tech bubble (which isn't even vaguely comparable to any of their or my findings). Yet in the new ABK and the webinar they equivocate that they're not predicting the chance of a real short-term crash, but rather kind of one to two standard deviation performance under the mean for a few years. OK… If that's the case I stand on my work here finding they overstate even this predictability – and separately present my thoughts on the chance of a true short sharp "crash" as new observations unrelated to ABKW/ABK.

[41] I discuss it further in my Cliff's Perspectives article including how regular old daily volatility (not just crash risk) might be somewhat higher once a factor is well-known. Putting crash risk and regular volatility together, even if a factor is not expensive, being well-known may indeed make things a somewhat wilder ride. Perhaps this is the price for factors that are truly tested and believable?

[42] The new ABK have footnote 12 where they attempt to make some link between valuations and the events of August 2007. I can't even follow it. Valuations were simply not abnormal going into that crazy month.

you measure risk. It should be less about what we might have to weather at some point (though you should be convinced you can indeed weather it). So, while ABKW and I agree on the possibility here, again assuming we agreed on the definition of a "crash", we do so for very different reasons. Also, unlike them, I would include the value factors in my crash worries (so in this sense I'm more worried than they are). They implicitly exclude them as value is "cheaper" than the other factors but if crash risk is, as I contend, more about being well-known than about valuation then the value factor is likely also at risk as it's a popular factor underlying many smart beta investment products.

*A simple factor timing strategy*

A clear implication of ABKW's work is that factor returns are to some degree predictable and that investors should take into account this predictability in managing their exposures to these factors. In other words, you can do better than just passively investing in factors by tactically timing your exposures to these factors. Of course, the regressions are suggestive, but they aren't actual timing strategies. Let's take that next step.

By combining both the factor returns and the factor value spreads I will look at a simple "style timing" or "factor timing" strategy for each of the five factors listed above done both with book-to-price and sales-to-price (so a total of ten factor timing strategies). Unlike in-sample regressions these use only the information available at the time. I use a very close cousin of the timing strategy used in AIM. Using a window that starts at 20 years and then expands, I compare the value spread today to the historical median, 95th percentile, and 5th percentile. The weight in the factor today is zero if the current spread is at historical median. If at or above the 95th percentile it's 2 x (95th percentile value spread – median value spread) / (95th percentile value spread – 5th percentile value spread) capped at 100%. If below the 5th percentile it's 2 x (5th percentile value spread – median value spread) / (95th percentile value spread – 5th percentile value spread) floored at -100%. Weights are linearly interpolated between these values. In English, essentially, if the value spread is currently at historical median, using only past data available at the time, you take no position in the factor; if above median you go long the factor (up to 100% if at or above the 95th percentile value spread) and vice versa going short up to -100% if the value spread of the factor is below median.[43] Since all five factors become available in 1968, and I require a 20 year window to begin timing, what follows discusses timing results from 1988-2016.[44]

While ABKW's regressions suggest factors are predictable and thus suggest investors should time these factors, they miss an important observation. It turns out that valuation-based factor timing mimics value exposures already commonly present in many, or even most, factor/smart beta portfolios that aren't timing factors. That is, if you're already exposed to the value factor, you'll find the valuation-based

---

[43] I describe this simple study as a common reference point for discussing value spread based factor timing using what I think is a fairly representative subset of the factors most commonly studied and implemented and using what I believe is a simple, robust and intuitive timing methodology easily replicable by other researchers. More exhaustive study of other methodologies is called for, but the below is robust to reasonably minor perturbations of this scheme (e.g., using means not medians as the center point, using 1st/99th instead of 5th/95th percentiles).
[44] Some of the factors are available earlier than that and future more exhaustive work should include investigations into the deeper past (albeit for a selective subset).

timing strategies largely mimic these returns. Below I graph the cumulative average excess return of the two simply long-short value factors (averaged over book- and sales-to-price) and the cumulative return of the average of all ten valuation-based factor timing strategies discussed above (separate axes).



**Cumulative Return Value Factors and Value Timing**
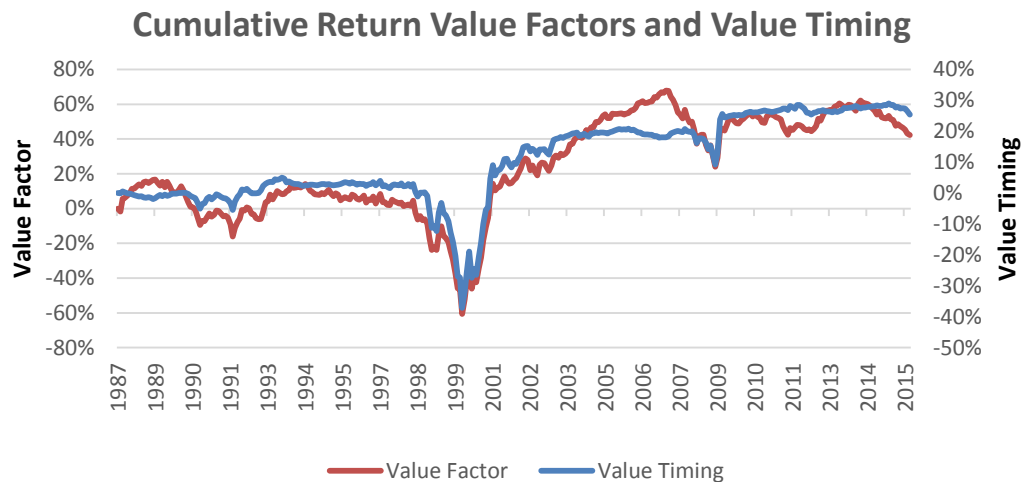
*Source: AQR.  See Appendix A for more detail.*

These are obviously highly correlated strategies, and this result is not specific to the precise timing methodology used or the precise combination of factors timed.[45,46] I've attempted various ex post optimizations (using a variety of objective functions both linear, like Sharpe ratio, and more "worst case" oriented) to see if an optimizer wants to take risk in the average of these timing strategies if already reasonably exposed to the regular old book- and sales-to-price factors. The optimizer generally wants little or none depending on the precise specification. Thus, I find that while I'm naturally inclined towards value strategies myself, for an investor already pursuing them as factors, adding any significant

---

[45] There is a level of timing implicit in any strategy and that the standard constant $1 long / $1 short strategies such as I, ABKW, and many others study, already include a fair amount of implicit value-spread style timing as the volatility of the factors do systematically vary through time. Timing factors that target constant volatility, not constant dollars, which would reduce the implicit timing in the factors is an interesting future avenue to explore, as are timing strategies that, unlike these, merely apportion risk between factors but don't allow independent timing of them. There is lots more to explore here (which I'm sure ABKW would agree is both a blessing and a curse as there's also a lot of room to data mine).

[46] Interestingly these ten timing strategies are also highly correlated to each other. That is not surprising to those of us who recognize they are all basically repeating the exercise of value investing in overlapping forms. Consider timing the value factors with valuation and separately timing the momentum factors with valuation. While the regular value and momentum factors are negatively correlated, of course valuation based timing tends to be long value and short momentum at the same time (or vice versa) so the timing strategies of these negatively correlated factors are actually quite positively correlated. This in itself is not damning but means we really have far less than ten new timing strategies (and that, again, valuation based timing is mostly just a specific and likely attenuated form of value investing itself).

amount of valuation-based factor timing is probably too much.[47] Though admittedly this is preliminary and different timing schemes over different factors might move the dial a bit (remember, we don't claim zero valuation-based timing is best, we do claim based on the evidence we've seen that you don't want to do a lot of it).[48]

Should we be surprised that these timing strategies are so correlated with value? Not really (at least for the non-value factors). If you're timing something based on its valuation you will do more of it when it's more correlated to the value factors. E.g., when profitability looks cheap the simple value factor is already more long (short) more (less) profitable companies. In other words, ABKW urge us to be factor contrarians, but the value factor itself is already contrarian including about other factors. That valuation-based timing of the value factors themselves is so correlated to value requires a bit more thought. The volatility of value factors is far greater when value spreads are wide which is also when the valuation based timing of value factors is long. Part of this is likely simply the result of some of these being relatively crazy times (e.g., tech bubble, near the bottom of the GFC, etc.). But part may be mechanical as when the value spread is wide it likely creates a "cleaner" sort than when the value spread is tight (due to noise in value's measurement). In other words, when value spreads are tight you create a more noisy, and less pure, value bet in the long and short positions of the regular value factor. Finally, I note that our (and ABKW's) timing on valuation is somewhat arbitrary. If momentum, low beta, and profitability look worse on value, value will also be looking worse on them. Meaning a pure value investor at those times is forced to be more anti momentum, lower profitability, and higher beta. If those factors are real (a separate debate soon to come) I don't see why that has less standing than the value spreads of momentum, profitability, and BAB.

In summary, five year overlapping in-sample forecasting regressions greatly exaggerate the ability to time these factors, are much less applicable to long-short factors (including value where it holds up best) than the market portfolio for which such regressions were originally developed, and are absolutely inapplicable to higher turnover factors. ABKW greatly overstate their importance and even ex ante applicability to factors and to smart beta vs. indices. None of this precludes a short-term serious correction, but such crashes are very difficult to predict, including using valuations. Finally, an investor already tilting towards any version of the value factor (like much of smart beta does) is already naturally being highly contrarian, benefitting much less, if at all, from additional valuation based factor timing not

---

[47]We will have more to say on this in future research.

[48] Sigh, yet another bone to pick with the ABKW/ABK team is they sometimes seem to agree with this statement claiming they aren't recommending draconian moves, or abandoning smart beta (they may in fact mean they don't recommend abandoning their version of smart beta). OK, according to them things might go "horribly wrong", a "factor crash" is probably coming, and our best comparison is the off-the-charts tech bubble, but when challenged, we don't think you should do much? For instance, in the webinar at about 27 minutes Rob tells us it's time to "double down" on value but then assures us soon again he's recommending nothing drastic. He also uses a straw man of "don't abandon value." I don't know if some recommend abandoning value (which would be terrible) but I'm saying stay diversified across all good factors most certainly including value. Their Fabian strategy of argument (give a little when necessary, then ignore that and make a very strong statement, then give a little, retreat, attack, retreat, never get caught) may be good debating tactic but doesn't help us get any nearer to the truth. If you aren't recommending anything beyond the trivial why are we having this debate? If you are making strong statements stand by them. You can't have your cake and aggressively time it too.

because it's bad but because it's similar to what they're already doing. The precise amount of timing to include, whether it's "none" or "a little", is admittedly still quite debatable but also not very exciting.

**V. Are these factors even real?**

I don't want to scare anyone but this is where it gets ugly.

The centerpiece of ABKW is their assertion that richening has driven these factors' long-term returns and thus the factors are not for real.[49,50] I disagree and this disagreement is not small. Rather, it's huge and important (including if I'm wrong!). I will show that their assertion is incorrect for reasons they occasionally note, but then oddly and repeatedly ignore.

Ilmanen, Nielsen, and Chandra (2015), henceforth INC, focus on the BAB strategy, but make observations widely applicable to other factors. They examine the precise issue of how much changing valuation drives contemporaneous factor performance. INC find that valuation changes generally do matter and with the expected sign. That is, when valuation spreads cheapen the factor tends to lose money over the period of the cheapening and vice versa. This is, of course, far from a shock. If someone told you your factor will richen you'd likely be pleased. But, while in the direction of common sense, INC actually find that this relation is far weaker for long-short factor strategies than for the stock market. And it's far weaker for some long-short strategies than others. Why the difference? There are things INC describe as "wedges" that I will call "frictions" that come between the change in valuation and the factor return.[51] A friction here is a reason why even if the factor gets X% cheaper (more expensive) over the period, the factor's return will be reduced (increased) by less than X%. The primary frictions are:

---

[49] Of course, any factor is still subject to the possibility it's the result of data mining and thus not "real." Strong economic stories and out-of-sample tests are our major recourse for addressing that worry. That's a completely separate issue than the targeted and, in my view, completely false charge from ABKW that this data mining is specifically over factor richening.

[50] ABKW (and the new ABK and webinar largely just reiterate them) is littered with strong statements in this regard. Often these are unqualified as to whether they mean over the short- or long-term. While still greatly exaggerated, their statements are far more egregious if about the long-term (and sometimes they are indeed quite clear it's about the long-term). At the start of the paper they write "The unsurprising reality is that many of the new factors deliver 'alpha' only because they've grown more expensive." In their introduction they say "Many of the most popular new factors and strategies have succeeded solely because they have become more and more expensive." As an aside I don't know which factors they mean are "new." But, if it's momentum or low beta they are gravely mistaken as these ain't remotely new. On page 2 they actually say regarding "many of the new factors" that "…absent rising relative relations, there's nothing left!" In another example ABKW say "Even over nearly a half-century, a shocking portion of the return for several factors comes from rising relative valuation." Clearly that's indeed about the long-term.

[51] To avoid confusion, I use the term "friction" not "wedge" as used in INC. ABKW and their other work also use the term "wedge" in a different way than INC. ABKW use it more as the equivalent of their term "structural alpha" or "inherent return." Of course, "friction" itself is a loaded term in finance often implying something like a trading cost. Here "friction" will mean natural effects that cause a change in factor valuation to be mirrored less than 1:1 (sometimes far less) in factor return. INC actually came first but I'll let ABKW have this one :).

(1) turnover (if a factor gets more expensive but ends the period as a set of mostly different stocks, this richening may be irrelevant as the factor investor didn't earn it),

(2) for strategies that must be hedged for beta differences (in our case only BAB), the beta mismatch itself can be a friction; e.g., if the market goes up and low beta stocks go up less than high beta stocks but more than implied by their ex ante betas, the BAB factor outperforms but also tends to get cheaper,

(3) changes in the fundamentals; e.g., if sales change but price doesn't, absent turnover there is no effect on factor return, but the change in sales will change the sales-to-price value spread at the end of the period.

ABKW adjust realized returns for changing valuations in two different ways. The first is to simply examine the full period change in the value spread (over either 10 or near 50 years in their table 1[52]) and subtract that result from the factor's return over that period. They declare this value spread change the unrepeatable result of valuation changes and not inherent return to the strategy (ABKW's "structural alpha"). In other words, if over 10 years the "value of value" got 10% richer they would say this 10% should not be attributed to the efficacy of the value strategy (it's a gain from the non-repeatable event of richening not inherent to the factor itself) and the real "structural alpha" of the strategy is actually about 1% less per annum than it appears to be if this effect is ignored. This method recognizes no "frictions" and applies the full change in the value spread to estimate the non-repeatable part of the factor return. We will see it's quite silly.

Their second method does recognize the "frictions" discussed by INC (ABKW only explicitly mention the one friction of turnover – I would conjecture this is indeed the most important but not the only friction[53]) by using a regression based approach I believe similar to what I'll use below to estimate the

---

[52] ABKW assert that many look at 10 or 15 year performance in judging factors. That may be true for performance chasers, or select dubious researchers, but the key factors studied by any credible researchers are looked at for far longer than 10 years. The near 50 years ABKW examine are far more relevant to the debate than the last 10 years they also study. They mention that "average factor indices have tracked returns between 14 and 17 years", but this commercial situation is irrelevant to the fact that all the main factors discussed here are studied by researchers for far longer than that. In particular, momentum, BAB, and profitability, the ones I study here along with valuation, have long been examined by researchers using data back to the 1950s or even the 1920s. They have both been studied by researchers for a longer time and over longer periods than ABKW imply. Actually, momentum studies go back to Napoleon! In addition, value, momentum, and BAB were clearly known before the last 10-15 years, making this not a period data mined over but, rather, an out-of-sample test. It would be fair to go back 10-15 years and look for factors once popular now discarded  –  an exercise neither they nor I do (can you think of many factors that had great 50 year histories we've recently abandoned?).

[53] Oddly, they get it completely right when they say, "A lesson we can draw from this is the higher a strategy's turnover, the less informative are valuation changes in understanding the strategy's performance and predicting its future performance." But this doesn't stop them from then repeatedly ignoring their own admonition applying their efforts to and making conclusions about the momentum and BAB factors. In fact, right after noting that valuation changes are only very weakly associated with even contemporaneous BAB returns (or, rather, their version of low beta), they then predict in the coming years BAB will be greatly hurt by mean reversion in value spreads. That may be true, but it's an odd statement to directly follow upon explicitly noting (and explaining why) valuation changes don't map very well to strategy returns. Similarly in ABK they come close to getting it when they

impact of valuation changes. Of course, this second method which accounts for the frictions shows a far more muted (and likely far more accurate) effect of valuation changes.

While we can debate the proper way to adjust for these frictions, what can't be debated is that the unadjusted valuation change (what I'll call "top-line changes" or ABKW's first method) overstates, in some cases massively, the effect. However, ABKW often choose to focus on these top-line changes despite acknowledging the severe problems they present (and problems in only one direction – extreme overstatement). For example, using their own regression approach[54] they find almost no effect on momentum over the last decade[55], but it's the misleading top-line results that repeatedly (and usually solely) make it into their discussion. Again, as I will demonstrate in a moment, in the presence of "frictions," the top-line method is simply and obviously wrong and patently silly for a high turnover strategy like momentum or even BAB. They outline their regression methodology in their footnote 22 but don't present the factor-by-factor results so let's do our own.

Before looking at long-short factors let's start out with the example of the market portfolio and use the Shiller P/E as the valuation factor. I think it helps with intuition. From January 1968 to January 2016 the average annual (arithmetic monthly) return of the market portfolio vs. cash was 5.9% with a t-statistic of 2.41. Now let's consider how much of that return each year comes from valuation changes. For this we use the change in the log of the Shiller P/E over the same period (the percentage change in the Shiller P/E if "continuously compounded").[56] Specifically, we regress the total excess over cash annual stock market return on this contemporaneous percentage change in the Shiller P/E. The regressions address the following questions: (1) "what do I make on stocks on average when there is no change in Shiller P/E?" (the intercept in the regression), (2) "when the Shiller P/E goes up or down by X% in a year how much of that move on average transfers to the stock market return?" (the slope in the regression) and (3) "how statistically powerful is this relationship?" (the R-squared and t-statistics in the regression). The

___

say "Strategies and factors with longer half-lives, such as small cap and profitability, are likely to have portfolios that change slowly from one year to the next, making it much easier to tease out the structural alpha." But close here is not enough. It's not easier to tease out structural alpha in low turnover strategies. It's harder. You must, as they point out, make sure they aren't the result of richening or cheapening. Higher turnover strategies don't need to worry about this issue. That doesn't make it "harder" to deal with them it makes it, all-else-equal, easier. You may or may not believe in higher turnover strategies but you generally don't worry about changing valuations driving them.

[54] The "regression coefficient-adjusted performance" located in their Table 1.

[55] This produces an 85% weaker effect than the top-line result they go on to highlight in their text when they explicitly talk about momentum being the result of richening for a decade implicitly using their flawed top-line method.

[56] It's somewhat important to work in logs here. If you look at the simple compounding average annual return to any relatively stationary ratio it will likely be positive. That's because of the well-known fact that a ratio that goes from 1.0 to 2.0 then back in the next year produces a +100% and -50% return or an average of +25%. The continuously compounded version is in turn +69% and -69% for the more intuitive average of zero. If you don't work in logs the slope may be biased and the intercept has no clear interpretation. If you work in logs the intercept has the clear interpretation as the average realized return given the value spread is unchanged from start to end of the year. It is not clear to me which ABKW use. For the left-hand side it's less important. I use summed arithmetic monthly excess returns over cash. Switching to logged returns changes the results of course but the important comparisons of the simple average to the regression intercepts are economically alike.

result over 1968-2016 looks like this (for consistency with other measures we actually regress on the logged change in the Shiller earnings-yield which is just the reciprocal of the Shiller P/E so you get negative coefficients)[57]:

$$\text{Annual stock market return} = 5.5\% - 0.94 * \text{Change in Shiller E/P}, \quad R^2 = 93\%, \quad \text{simple avg.} = 5.9\%$$
$$(7.17) \quad (-41.9)$$

In addition to the regression, I include the simple annual average return of the factor on the far right to facilitate easy comparison of this simple average and the regression intercept. The coefficient and $R^2$ are close to, but not quite, -1.00 and 100% respectively.[58] On true buy-and-hold portfolios[59] we might expect or even impose a coefficient of -1.00, but even here there may be frictions (e.g., the E in P/E changes) and we wouldn't expect 100% R-squareds due to changing carry through time. Basically, the thought experiment of "if I knew the Shiller P/E a year from now" tells you an awful lot about returns over the next year but not quite everything.

Now we come to a central point of these regressions. The simple average annual excess return (no regressions) is 5.9%, but the intercept in this regression is somewhat lower at 5.5%.[60] Why is the regression intercept about 40 basis points lower? Well, because part of the historical return to the stock market comes from an increase in the Shiller P/E over the full period (that is, today's CAPE is quite high versus history). That is, over this entire period stocks got more expensive. The straight average return of 5.9% effectively gives the stock market full credit for this appreciation while the intercept in the regression gives it zero credit. Instead, the intercept gives an estimate of what the annual return would've been if we didn't see this (or any) valuation change. ABKW do a good job of motivating why this type of analysis may be important.

While the regression intercept and straight average return differ, and differ in the intuitive direction[61], the difference is not economically huge. We expect (and find) only small differences when "smeared" over very long periods (almost 50 years here) and when the change in valuation measure is not off-the-charts.

That's not to say this adjustment never matters. Over shorter periods, or periods that end with extreme valuations, the differences can become significant. For instance, doing the same exercise as above from

---

[57] In each case these regressions are done using rolling annual periods with the appropriate asymptotic adjustments to t-statistics. Results using non-overlapping calendar year returns are economically quite similar.
[58] We expect a negative coefficient as it's the change in earnings yield on the right-hand side meaning when it goes up the P/E goes down (and falling P/Es hurt returns).
[59] This is the portfolio that the Fama-French market closely approximates though it never gets to pure zero turnover buy-and-hold.
[60] You also get a t-statistic for the simple average of 2.41 versus 7.17 for the regression intercept. The 7.17 is much higher as knowing what happened to valuation contemporaneously tells us a ton (but not everything!) about return and makes our estimate of the intercept much more precise. Sadly it's not an implementable investment strategy as, of course, it involves knowledge of the future. I reserve the right to think a bit more about whether this intercept tells us anything extra about the reality of the unobservable true equity risk premium.
[61] In this case a straight average likely overestimates future stock returns as the part of past return from richening is likely unrepeatable and (perhaps) even mean reverting.

1968 until the peak of the technology bubble in March of 2000, you get a regression intercept of 4.5% but a straight average annual return of 6.5%. That is, instead of the 40 or so basis points difference we found over 1968-2016, we find a 200 basis points difference in the methods over the 1968-2000 period. A bit of this is the somewhat shorter period (about 32 vs. about 48 years) where valuation moves matter more as they are "smeared" across less time. But most of the far larger effect comes from the off-the-charts high valuations in March of 2000 (today's CAPE is high but not close to tech bubble peaks). When such conditions prevail the worries raised by ABKW can indeed be quite important.[62]

Now let's run similar regressions for the long-short factors (styles). Let's first regress the annual return of the book-to-price factor on the contemporaneous change in the B/P value spread[63]:

$$\text{Book-to-price factor } = 3.2\% - 0.84 * \text{Change in Factor B/P}, \quad R^2 = 82\%, \quad \text{simple avg.} = 3.0\%$$
$$(3.7) \quad (-14.5)$$

The average return of the book-to-price factor[64] over the whole period is 3.0%, only a trivial 20 basis or so points less than the 3.2% intercept. The simple realized return is a bit less than the regression intercept because, as ABKW observe, the book-to-price factor has ended the 1968-present period cheaper than it started. So realized returns are a bit lower (hurt by falling valuation) than the regression intercept as the regression intercept does not get similarly impacted. But since it's not tech bubble expensive, not even close, and because its effect is dispersed over nearly 50 years, this directionally correct observation of ABKW is also a very small one for value (they don't claim anything very different for value, as the regression and "top-line" methods are mostly similar here, the differences will come with the other factors).

For the book-to-price factor we, of course, find a significant negative relation between spread change and realized return. When valuations get cheaper (more expensive) the factor on average suffers (does better). But the relationship is a bit weaker, in coefficient and $R^2$, than we found for the stock market. This is because the "frictions" are more important even for relatively low turnover long-short factors like book-to-price. It turns out that even what I just called a low-turnover factor like value still, of course, entails more turnover than the market. Turnover is, again, one of the primary "frictions." The intuition behind this should be clear. If the value factor ends the period more expensive, but it's holding a different set of stocks (long, short, or both) than at the start of the period, then the owner of the value portfolio didn't actually get this seeming return from richening.

---

[62] ABKW highlight this same example early on though not using precisely the same analytics. As an example of the occasional importance of these effects the tech bubble is wonderful. As a comparison to today's factor pricing it's massively overdone.

[63] I do this analysis for only the book-to-price value spread. Repeating it using the sales-to-price value spread yields qualitatively similar results (stronger for the sales-to-price factor itself, a bit weaker for the other factors). Similarly we don't use the sales-to-price factor on the left-hand-side of the regression but results, again, are similar to those for book-to-price (in particular when using the sales-to-price value spread change for the sales-to-price factor itself). By-the-way, ABK find a similar strong (but weaker than for the full stock market) correlation here and utter "stop the presses!" Uh, no, this is what everyone would expect. It's the shocking low power of these regressions for the other factors that perhaps deserve that exclamation.

[64] Note that this is not from a regression it's just the straight average of the left hand side returns.

Now let's look at the other factors for some real fireworks over 1968-2016 (or really a lack of fireworks!)[65]:

Momentum factor[66] = 4.7% - 0.13 * Change in Factor B/P,   $R^2$ = 25%,   simple avg. = 4.8%
                  (2.8)   (-5.1)

Profitability factor = 1.4% - 0.64 * Change in Factor B/P,   $R^2$ = 64%,   simple avg. = 1.6%
                  (1.6)   (-10.1)

BAB factor = 4.0% - 0.13 * Change in Factor B/P,   $R^2$ = 16%,   simple avg. = 4.3%
                  (3.1)    (-3.2)

Profitability yields a weaker contemporaneous relationship (slope and R-squared) than book-to-price and momentum and BAB much weaker.[67,68,69,70] What's going on? Well, the frictions are just bigger. The

---

[65] The size of these factor premias (the intercepts in these regressions and the straight factor returns) are lower than what many are used to as, recall, I'm working only within large capitalization stocks not averaging results across both large and small stocks as is common in the Fama-French methodology.

[66] Here's a really odd one from ABKW. Notice the momentum factor intercept here is statistically significant. Yet, in ABKW Table 1 full sample they find a momentum premium of 3.93% (labelled "long-term return") that they mark as statistically significant yet the near identical "performance net of valuation change" of 3.83% that is marked insignificant. How'd they make the significance disappear? I believe what they're doing, mistakenly, is subtracting the return from valuation change each month and taking the variance of that difference. Thus, the variance of the return from valuation change raises the volatility of the momentum premium when in fact if you adjust for it, it should lower this volatility (as it does in the regression intercepts here). They have slipped in a result that the momentum premium is still large but now statistically insignificant after adjusting for valuation change – and I believe it's simply an error.

[67] ABKW find over their full period that the returns to their version of low beta drop by about 1/3 using their regression method. In contrast I find almost no change (approximately 30 basis points between the 4.0% intercept and the 4.3% average return). However, if I remove the hedge ratio needed to keep BAB (what they call low beta vs. high beta) from having a permanent short market position I also find about a 30% reduction in BAB returns using the regression method (i.e., the intercept vs. simple average difference is more like ABKW's). This is likely because a permanently short (as I suspect ABKW are using) portfolio wrongly mixes in some of the more powerful predictability found in the regressions of the market on changing valuations with the far weaker long-short factor on valuation-change regressions. Also, the 4.0% intercept in the BAB regression above only drops to 3.4%, still quite strong, if one stops the regression a decade ago. The BAB results are not just the result of a strong decade. This would not surprise Fischer Black (Black 1972) who discussed this effect in the 1970s!

[68] ABKW say "Low beta's end-point in relative valuation is near an all-time peak, meaning the historical link between relative valuation levels and returns will seem weak, even if it's not (since these recent high valuations have not yet had the chance to mean revert)." They are saying that their and my finding, that valuations don't drive BAB returns much at all, are perhaps the result of BAB ending so expensive. You have to admire their tenacity. When faced with an empirical rejection (which I also find a bit surprising as remember, we expect value to work) they retreat to conjecture but can't quite bring themselves to test their conjecture! Well if you think the regressions are broken by the "flood of capital" at the end point we can fix that. Re-running my regression over different end points, or dividing the sample in half, yields little change to the "friction" making their story quite unlikely. The new ABK repeats the conjecture from ABKW that perhaps valuation fails (predictively and contemporaneously) for BAB because of the end point. Curiously, again, this is left to conjecture in both papers. Apparently I alone have mastered the technology of running a regression over sub-periods, or alternatively, a supportive "conjecture" is more useful than a conjecture rejecting test.

turnover is higher (particularly of course for momentum) and in the case of BAB there is an additional friction that comes from unequal dollars long and short. Essentially, ABKW get the sign on this effect right but gigantically overstress the magnitude. In plain English, choosing momentum as an example, the regression estimate says the top-line method driving most of ABKW's strongest statements exaggerates the effect by 87% over the last decade.[71]

Earlier I compared the intercept in these contemporaneous regressions[72] with the stand-alone average return on the factor. For the stock market I found little long-term difference though in the expected direction given the market today is expensive versus history (and again if one was at the end of the tech bubble then accounting for the spread change matters much more). Well, for none of the long-short factors is the intercept significantly different from the straight average of the left-hand-side returns (for momentum it rounds to equal at one decimal place all the way up to BAB where it's a measly approximately 30 basis points). Meaning, unlike the claims of ABKW, factoring in changes in valuation does not change our conclusions about <u>any</u> of these factors over the long-term. There is just zero evidence that the historical realized return of these factors is driven by long-term richening in their valuations.

Any factor is subject to the possibility that the evidence behind it may be the result of data mining. But, again, ABKW don't make this very general point. They say: "Many of the most popular new factors and strategies have succeeded solely because they have become more and more expensive." And they say that about both the short- and (repeatedly) the long-term. And they say it again, and again, and again, across each of their recent pieces. In other words, they allege a very specific point – that the non-value factors have all richened and researchers have mistaken this richening for inherent alpha (or in this discussion – regression intercept).[73] ABKW come to their conclusions by stressing (and often only citing)

---

[69] By-the-way, examining ABKW's Figure 2 (second panel middle graph on left; the red line) the valuations for low vs. high beta were very high but have recently come down to much more normal (without a concomitant factor return crash, remember the rather giant frictions here!). Their text and their figure are at some odds as the text discusses the super-high endpoint that isn't there (though please do not worry as in the new ABK it is back to just very high!).

[70] The BAB strategy I study here is somewhat stronger than the corresponding low vs. high beta results in ABKW. I'm not sure why. Some of it may be that they might not use a hedge ratio, but that would not explain all the difference. My results are consistent with the large capitalization only part of the returns found here and in general with the results in Frazzini and Pedersen (2014).

[71] While I disagree strongly with their focus on the simply wrong top line results I do credit ABKW with promoting these regressions. They are very useful. I have some doubts about ABKW's implementation (again, logs or not?) but formulated correctly they are a very neat way to demonstrate the effect (or lack thereof for many factors) of changing valuations on factor or market returns.

[72] Recall that these intercepts are an estimate of the annual return on the factor in the case of an unchanged right-hand-side value spread.

[73] In discussing their observation that "price matters" the new ABK paper says "We marvel that these observations are controversial in some circles." They are saying it is all just common sense. I can't help feeling I'm being discussed here without formal reference :) They should stop "marveling" and read the criticisms. This is not simple prairie home companion ah shuck simple stuff. It is indeed common sense that price matters – common sense I've promoted for 25 years and fought for when necessary. In fact it's why factor investors almost invariably include substantial exposure to the value factor. But it's not common sense that timing the non-value factors with valuation itself is helpful or harmful when the regular value factor is already a big part of the portfolio. It's certainly

the top-line change in the value spread and not the regression method (or some other similar approach) to account for "frictions." This is not a matter of personal choice among two plausible methodologies where perhaps you want to give weight to both methodologies.[74] ABKW are wrong when they refer to the regression method as "more conservative" versus the top-line. It is only "more conservative" in that the top-line method they focus on is simply wrong and always wrong in the direction of exaggeration. The regression results are at least a best guess at unbiased results, unlike the top-line method which are very biased to exaggeration, and this best guess does not support ABKW's story at all over the long-term.[75]

## VI. Conclusion

Where does all this analysis and discussion leave me? I believe in value investing and its close cousin contrarian investing. In particular, I believe in it when implemented with diversification (e.g., the long-short value factor) and combined with other diversifying factors. But, used alone, for narrow decisions like market or factor timing, when not at screaming historical extremes, pure valuation is just not a very strong strategy. Even at extremes it's quite the wild ride. This is clear when tested using real-world implementable strategies.[76] Furthermore, regardless of the power of valuation-based timing, I find current measurements are less extreme than do ABKW (using just book-to-price).[77] I also find that whatever modest power value-based timing has, standalone it provides little diversification to an investor already exposed to the regular diversified value factors. I find this unsurprising as value-based factor timing is (mostly) just a narrower time-varying form of value itself. This last part is preliminary and does call for more research over alternative timing methods and different metrics to evaluate the portfolio.

not common sense (it's not even right) that the long-term returns of the non-value factors are driven by richening and therefore not even valid to begin with, an assertion each of their works repeats again and again. So, invoking "common sense" when your tests, recommendations, and accusations (the industry has data mined it all!) are somewhere between dodgy and invalid, is not any kind of sense common or otherwise. I marvel at their marveling.
[74] Even handedness is only a simulacrum of "fair" if one option is clearly and obviously wrong. Yet they can't even muster such even handedness! By the time the webinar rolls around it's not a choice of two methods it's only the silly top-line results that are mentioned (check it out around 18:30). That is simply not ok. You just don't get to say "well, some say 2+2=4 some say 2+2=5, there is a debate" and then move on to tacitly and consistently use 2+2=5. It is really quite outrageous and I call on them to fully take back their statements throughout both written works and the webinar that richening valuations have driven the near 50 year results for these non-value factors.
[75] They do make some points I agree with even about the long-term. For instance, they note that the profitability factor is a conundrum as you seem to be getting paid for owning better more comfortable assets. While not dispositive it's an interesting and reasonable point (we discuss that ourselves here including discussing the rather subtle difference between explaining profitability adding value in the presence of the value factor versus it working on its own). This is an absolute fair point to raise when discussing whether profitability is the result of data mining or has "structural alpha" but has nothing to do with the valuation changes ABKW discuss. They also correctly note that the industry is in quite a confused state defining "quality."
[76] Unlike the long-horizon regressions used, and sometimes abused, by ABKW, or even the shorter ones added, welcomingly, by ABK.
[77] Again, the new ABK finds similar results using four measures, even if the new ABK's tone resembles ABKW while their numbers resemble mine.

Part of my debate with ABKW is about magnitude not sign and that is where our differences are most understandable and mild. I agree with ABKW that much of the world engages in performance chasing or what I often call "momentum investing at a value time horizon" (see [peeve #3](#)). While momentum investing is, on average, effective, at shorter (say 3-12 months) horizons, sadly many investors seem to act like trend followers at 3-5 year horizons when they should be acting as (usually very mild) contrarians.[78] I like to say that performance chasing at these horizons is not merely "wrong" but "backwards" – an extra bad kind of wrong. Still, avoiding performance chasing is not the same as doing a significant amount of the opposite. I infer from their paper that they'd be significantly contrarian factor timers.[79] In that respect I think the evidence says "sin a little" (if timing is a sin) or perhaps don't sin at all if you already have significant exposure to the naturally highly contrarian regular old value factor itself.[80] My advice is to do a little (none is still fine barring far larger extremes) pure contrarian timing but rather focus on finding good factors you believe in[81] and stick with them long-term. In fact, stick with them like grim death as even good factors have bad periods (bad periods that become far harder to stick with when you add in significant contrarian timing by the way). If one must "sin" then, please, sin only a bit and only an amount you can stick to through the tough times.

A place I strongly disagree with ABKW is their repeated comparisons to the technology bubble of 1999-2000 and other very strong statements made about today's rather mundane factor valuations. These are histrionic and unwarranted by ABKW's own modest findings let alone my even more circumspect ones regarding current valuations (again more circumspect results shared by the newer ABK). Some factors are modestly cheap today, some are modestly expensive, a few actually on the high side (but not above past highs) versus history, but none are vaguely comparable to the extremes of 1999-2000 which <u>blew away</u> prior experience. You can't lead with a comparison to 1999-2000, put "Horribly Wrong" in your title, predict a "factor crash," encourage investors to "double down," and then defend yourself from being called histrionic with something like "well we just mean, you know, somewhat worse than usual

---

[78] Of course I mean active investors (even if "active" here is smart beta) as taken all together all investors can't lean the same way.

[79] The conclusion of the newer ABK specifically says otherwise cautioning against too much timing. I think this is just CYA stuff. If they aren't urging significant factor timing now I have to ask what's the point of this whole back and forth? What's the point of "horribly wrong"? Their conclusion employs a Charlie Munger quote implying that ignoring them is "stupid" and ends with the amazing straw man "Before leaping into the next great strategy, do we really want to follow the counsel of many who argue that we should turn a blind eye to current valuation levels?" Again, I have a sinking feeling they mean me. If by pointing out that significant exposure to the value factor already means significant consideration of valuation and adding factor timing risks doubling up on this one risk/return factor, and that many of the factors they labeled in ABKW as very expensive are much more moderately so using more robust multi-factor measure in my Asness (2016) and indeed their follow-up ABK, and that their assertions that richening have driven a near fifty year "alpha mirage" in many of the non-value factors is simply flat out wrong, adds up to "turning a blind eye to current valuation levels" then I plead guilty.

[80] Ironically one of the behavioral biases that may help give rise to the value effect is "overconfidence" and I'm essentially saying that to add a lot of valuation based factor timing on top of the regular value factor is showing a lot of "overconfidence about benefiting from others' overconfidence."

[81] In other words, factors where you believe their historical results are not the result of data mining (I defend the factors from the specifics charges of data mining brought by ABKW and ABK but not from all such charges) and have not been arbitraged away (likely value spreads well worse than old historical extremes).

and you really must act carefully." As I've said before, shouting fire in a (surprisingly not so) crowded factor theater is wrong. Pretending you didn't so shout is weak.

The part of ABKW I find most unreasonable is, unfortunately, the major new takeaway of their paper[82] – the idea that factor "richening" (again, for only the non-value factors) has driven data mining crazed researchers to overstate long-term results. This is just wrong. Over the typical long horizons studied these richenings have had trivial impact. I find this result very intuitive given long (near 50 year) time horizons smear the effect of valuation changes and because of the "frictions" that exist between valuation changes and factor returns.

That's not to say ABKW is all bad. ABKW do a service when they remind us that any richening or cheapening is generally not repeatable and shouldn't be considered "inherent" factor return. That is something too easy to forget. But this only has practical significance over relatively short periods or sometimes near the peak (or valley) of great epochal financial events.[83,84,85] Unfortunately, ABKW do a

---

[82] They summarize their paper with four key points and this is number one and it's the result quoted under Rob Arnott's portrait in big font on the front page!

[83] ABKW assert at one point that "many of these alpha claims are based on a 10- to 15-year backtest." That would indeed be a shorter period more likely affected by factor richening or cheapening. While some may run such short-term back tests this simply isn't the norm for the main contenders I discuss here that underlie most research published in credible journals. Momentum, BAB, and profitability all are studied for way longer than 10-15 years (in fact, only profitability is a relatively recent finding while momentum and BAB are truly "out of sample" for the last 10-15 years, the period ABKW believe researchers are data mining over!). However, I readily concede that it's dangerous if a researcher publishes results over only 10 to 15 years and much more so if that's a period where valuations changed a lot and the researcher ignores it. In general ABKW's discussion of recent results (last 10-15 years) versus full period (1967-2016) results can be difficult to untangle (when do they mean the last 10-15 years vs. the full sample?). Though statements like this from ABKW make it hard to believe they aren't erroneously talking about the very long horizon: "Our parable holds a relevant lesson for smart beta investors: a lengthy return history, even 50 years, does not guarantee a correct conclusion. Investors need to look under the hood to understand how a strategy or factor produced its alpha. We compare several popular strategies' current valuations relative to history, and find that for many, much of the historical value-add—in some cases, all!—has come primarily from the "alpha mirage" of rising valuations."

[84] In yet another oddness, ABKW repeatedly imply that researchers have data mined over the last 10-15 years to find these non-value factors, yet only profitability has had 10 or 15 year returns above its average 1968-2016 results (and still not statistically significantly in a regression on a last 10 or last 15 year dummy variable). Momentum and BAB are (again not statistically significantly) below theirs. Even over the last five years momentum is average, and BAB a tad better than average, and again only profitability non-trivially better than average (though still not statistically significantly so regressing on a last five years dummy). The last five years indeed might currently contribute to these strategies' recent popularity, I certainly don't disagree that investors chase performance. But these five years are certainly not the period researchers may or may not have data mined over (the real point of ABKW's story). The core tacit assumption, that these factors have had way better-than-normal 10-15 year returns does not seem compelling.

[85] Similarly, they assert that some of the original published work on factors was the result of data mining over recent returns. For instance, in footnote 3 of ABK they say "It bears mention that Sanjoy Basu's seminal paper on the value effect appeared in 1977 after an extraordinary five-year run in which the value effect both delivered exceptional returns, following the collapse of the Nifty Fifty… Was Basu engaged in inadvertent data mining, finding the value effect after, and because of, a surge in both relative valuation and relative performance? Probably." The only problem with this conclusion is they quite wrongly look at the 1977 date of publication, not the fact that the sample period that Basu studied which ended in 1971. Those familiar with the process of writing

large disservice when they accuse broad swaths of the industry of blissfully ignoring valuation changes and thus mistaking factor richening for true inherent return in a wild bacchanalia of marketing and tenure-seeking data mining. They have effectively said "you all have had it wrong for years, let us educate you as to your mistake." They have repeated this again and again (the 50 year alpha "mirage"!). Yet it is they who are flat out wrong on this topic. An apology and retraction, at the least, is warranted, rather than repeating the falsehood again and again.

Further, I admit to being vexed at ABKW's self-serving argument (sell non-value factors buy value!), all the while occasionally protesting "we don't mean anything drastic" complete with constructing a facile moving target in terms of advice (sometimes it's only do a little, sometimes it sounds like bet the farm) and reasoning for why you should favor their form of "smart beta" not others.[86,87] Their argument comes not just from regular run-of-the mill contrarian factor timing, though they do discuss this at great length, but most directly from their delegitimizing of the non-value factors' long-term tests. In fact, they

---

and reading academic papers in top journals know this delay is a very common occurrence as the review, acceptance and publication process is a lengthy one, often ending long after the research that went into the paper has been completed. Basu's sample period actually ended following a rather middling period for value which was only then followed by very strong value performance, nearly the opposite of the ABK story. Basu did not data mine but was rather rewarded with good out-of-sample performance! Again, we see that really neat stories, sometimes near-morality parables, can wither in the face of correctly examined data. The Basu story fit the ABKW/ABK framework to a tee – they just got it backwards. As an aside I've done this analysis in terms of book-to-price, following ABK's footnote 3, but Basu actually studied the E/P strategy (the results are more in the middle but absolutely don't support the ABK parable). If you're curious, the analogous accusation (Arnott mentioned Banz (1981) for small cap on our recent joint panel) is backwards for small cap (the size premium is actually negative for the last five years of Banz's paper's sample data) and wrong for low beta (using Black, Jensen, and Scholes (1972)) and momentum (Jegadeesh and Titman (1993)) where the last five years of their data was slightly below average. In each case the first five years of out-of-sample results (after the data in the paper generally about the length of the pre-publication period) were strong. Finally, if one wants an example of finding a new result (that wasn't really new but that's a different battle!) after a celebrated period of returns, you can actually look at the "discovery" of fundamental indexing after the gigantic value comeback following the tech bubble. Ironically they name it smart beta. Aside from the above how was the play Mrs. Arnott?

[86] If all ABKW (and ABK and the webinar) are saying is "don't abandon value now!" then, sure, I'm right there with them. I always have been. But somehow "horribly wrong!" "factor crash" and "double down" make me think that's not the overall recommendation. They move back and forth between seeming strong prescriptions for an implied heavy amount of timing, and protestations of reasonableness and mildness, with disquieting ease (i.e., "We would offer a word of caution to those who wish to use these findings to aggressively time strategies: timing strategies leads to greater concentration of risk, so lacks diversification. A noisy signal and weak ability to diversify is a recipe for disappointment").

[87] In debate (live and in writing) ABKW/ABK seem to want to reduce their position to "prices matter and isn't that just common sense?" Of course prices matter but they are saying far more than that, and that's where the trouble starts! They're claiming some strategies are mirages created by data miners ignorant of the effect of valuation changes over the long-term. They're claiming it's surprising and new that value spreads predict factor returns despite us already knowing value works. Their implying that this in-sample predictability matters a lot for a portfolio without studying real world implementable versions or considering that an investor with significant exposure to the value factor already has made a bold statement that price matters. There are some hard issues here and I may yet be proven wrong. But it's not just "common sense" at all. It may indeed be the case that some factors have gotten more expensive partly due to flows, themselves due to recent popularity. But, again, comparisons to "bubbles" are way too extreme and what and how much to do with this information is far from clear.

move back and forth quite often between these two very different arguments with confusing ease ("hey, be a 5 year contrarian" vs. "these other factors aren't even real!"). If the factors are not real, we do not even need to engage in the rest of the discussion about their current valuation.

What do I think investors should do? Going forward I do not know whether value or the non-value factors will do better. I wasn't smart enough ten years ago to time my way out of value and into the other factors (though thankfully did enough of them!). I hope, and would expect based on history, that should value win over the others going forward that the combination of all would still do ok. I do not root against value investing going forward. As always I come to praise not bury value investing. But, barring pricing that is historically off-the-charts, something not present as of this writing, I continue to favor factor diversification over the (venal) sin of a little too much factor timing and the (cardinal) sin of extreme factor timing and the super-duper-cardinal sin of dismissing near 50-year backtests on false charges. Overall, as advice to investors (which I don't purport to be giving of course), I stress finding what factors you believe in and staying diversified across them unless we one day see far more extreme pricing than we do today.

**References**

Arnott, R., Beck, N., Kalesnik, V., and West, J. "How Can "Smart Beta" Go Horribly Wrong?" Fundamentals, Research Affiliates, 2016.

Arnott, R., Beck, N., and Kalesnik, V., "To Win with 'Smart Beta' Ask If the Price is Right" Fundamentals, Research Affiliates, 2016b.

Arnott, R., Webinar "How Can "Smart Beta" Go Horribly Wrong?" Research Affiliates, 2016, https://www.youtube.com/watch?v=npXPFnrM17Y&list=PLTcDPNRSpDnydPFSxH6Y8EDs9tMy4KWAW&index=3.

Asness, C. "The Value of Fundamental Indexing," Institutional Investor, October 16, 2006

Asness, C. , Friedman J., Krail R., Liew J. "Style Timing: Value vs. Growth." The Journal of Portfolio Management, Vol. 26, No. 3, (Spring 2000), pp 50-60.

Asness C., and Frazzini, A. "The Devil in HML's Details," The Journal of Portfolio Management, Vol. 39, No. 4, (2013), pp. 49-68.

Asness, C., Frazzini, A., Israel, R., and Moskowitz, T. "Fact, Fiction and Momentum Investing," AQR Capital Management, 2014.

Asness, C., Frazzini, A. Israel, R., and Moskowitz, T. "Fact, Fiction and Value Investing," AQR Capital Management, 2015.

Asness, C., Ilmanen, A., and Maloney, T. "Market Timing Is Back in the Hunt for Investors," Institutional Investor, November 11, 2015.

Asness, C., "How Can a Strategy Still Work If Everyone Knows About It?" Forum post, AQR Capital Management, 2015.

Asness, C., "The Siren Song of Factor Timing," The Journal of Portfolio Management, Special QES Issue, Vol. 42, No. 5, pp. 1-6, 2016.

Banz, Rolf W. "The Relationship Between Return and Market Value of Common Stocks." *Journal of Financial Economics*. Vol 9, Issue 1, (March 1981). Pp 3-18.

Basu, Sanjoy. "Investment Performance of Common Stocks in Relation to Their Price-Earnings Ratios: A Test of the Efficient Market Hypothesis." *Journal of Finance.* Vol 32. No 3. (June 1977). Pp-663-682.

Black, F., "Capital market equilibrium with restricted borrowing," Journal of Business, 45, 3, pp. 444-455, 1972.

Black, F., Jensen, M.C., and Scholes, M. "The Capital Asset Pricing Model: Some Empirical Tests," in Studies in the Theory of Capital Markets. Michael C. Jensen, ed. New York: Praeger, (1972), pp. 79-121.

Cohen, R., Polk, C., and Vuolteenaho, T. "The Value Spread," The Journal of Finance, Vol. 58, No. 2, (2003), pp. 609-642.

Fama, E., and French, K. "Permanent and Temporary Components of Stock Prices," Journal of Political Economy, Vol 96, No. 2 (1988), pp. 246-273.

Fama, E., and French, K. "The Cross-Section of Expected Stock Returns," Journal of Finance, Vol. 47, No. 2, (1992), pp. 427-465.

Fama E., and French, K. "Common Risk Factors in the Returns on Stocks and Bonds," Journal of Financial Economics, Vol. 33, (1993), pp. 3-56.

Frazzini, A., and Pedersen, L. "Betting Against Beta," Journal of Financial Economics, Vol. 111, No. 1, (2014) pp. 1-25.

Geczy, C., and Samonov, M. "Two Centuries of Price Momentum," Forthcoming in Financial Analysts Journal. 2016.

Ilmanen, A., Nielsen, L., and Chandra, S. "Are Defensive Stocks Expensive? A Closer Look at Value Spreads," 2015.

Jegadeesh, N., and Titman, S. "Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency," Journal of Finance, Vol. 48, No. 1, (1993), pp. 65-91.

Novy-Marx, R. "The Other Side of Value: The Gross Profitability Premium," Journal of Financial Economics, Vol. 108, No. 1, (2013), pp. 1-29.

## Appendix A

The valuation factors follow the method of Fama and French (1993) except we use up-to-date price as studied by Asness and Frazzini (2013). The book-to-price (B/P) data comes from the AQR data library and sales-to-price (S/P) is calculated using the same data sources (CRSP, Compustat, XpressFeed Global).  Data for the momentum factor and Betting-against-beta (BAB) factor come from the AQR data library.  Data for all factors is for large cap U.S. stocks and are capitalization weighted long the 1/3 best stocks on the respective measure and short the 1/3 worst.  Data are from January 1968 through January 2016. Additional information on sources and calculation methodologies are described on the website along with the data sets.