

FACTOR / STYLE INVESTING

Lies, Damned Lies, and Data Mining

April 12, 2017

We are the whipping boy for [a recent article](#) on the dangers of data mining in our field. And the whipping is delivered largely based on an unsupported shot taken by my frequent foil and sparring partner, Rob Arnott. Before I take on this attack¹, we need to back up a bit.

Data mining, that is searching the data to find in-sample patterns in returns that are not real but random, and then believing you've found truth, is a real problem in our field. Random doesn't tend to repeat so data mining often fails to produce attractive real life returns going forward. And given the rewards to gathering assets, often made easier with a good "backtest," the incentive to data mine is great. We've talked about it endlessly for years and written on it many times. But we're not nihilists who believe everything is data mining.^{2,3} We are more likely to believe in-sample evidence when it's also accompanied by strong out-of-sample evidence (across time, geography, and asset class⁴, and an economic story that makes sense.⁵ In that case, and barring exceptionally convincing evidence something has changed, we not only believe in it but will stick to it like grim death through its inevitable ups and downs. After many years of research and managing portfolios, we believe there are at least [four widely known types of factors](#) that are real (that is, they don't just look good because of data mining).^{6,7,8} People are often shocked that we believe in only a few core investment concepts – somehow they think there are many more. Nope. For instance: No small firm effect. No January effect. No Super Bowl effect – though if you do believe that indicator, you should be shorting stocks this year because of Tom Brady; sorry if that's deflating.

But all of this didn't stop us from being the cannon fodder in this [new article](#). Are we data miners? Heck no. We've always explicitly stood for the opposite. The list of things with great backtests we don't believe in is legion, and the ones we do believe in have, again, tended to work through time, in a multitude of asset classes and across geographies, with many of these being out-of-sample tests of the original findings. But, when you are trying your best to come up with a story, you find people who will say what you need them to (a journalistic version of data mining!) about someone vaguely interesting (I guess we're vaguely so). So the reporter asked a non-objective guy with whom I've feuded to opine on me and by extension AQR.^{9,10} It's no secret to readers of this perspectives column, and our work in general, that we have had an intellectual dispute with Rob Arnott on the subject of whether the main factors commonly discussed are something one should time (get in/out based on how expensive they look versus history).¹¹ A secondary debate has indeed been whether some of the main factors in finance¹² are the result of long-term data mining. Well, perhaps because he lost the first point, Arnott upped the secondary topic to primary and unleashed on me in the Businessweek piece, leading with a tiny bit of honey about my "[outstanding](#)" prior work but then bringing on a big bowl of vinegar flavored whoop-ass.¹³ Rob says,

"I think Cliff has done some outstanding work over the years," but adds that he's "insufficiently skeptical about the pervasiveness of data-mining and its impact even in the factors he uses."

That is, he says I'm a data-miner. That may seem like an innocuous little comment actually prefaced with a kind of compliment. It's not. It's a damning accusation that's provably false, backwards in fact. Worse, it's a falsehood meant to deflect and confuse as it kind of rhymes with a separate dispute we've been having, the "secondary debate" mentioned above – a dispute Rob's been ducking. So, if you just read Rob's comment on its own by most peoples' standards I'm overreacting here. Admittedly that's kind of my go-to move. But, in the broader context, the ongoing debate and what a serious and backwards "shot" he really took, I think I'm reacting appropriately. Of course, I usually think that...

After Rob's quote the article provides a response from me. They actually ran a somewhat truncated version of what I said. Here is the verbatim response I sent the reporter to Rob's above comment,

"Rob and AQR largely believe in a very similar set of factors like value, low risk, and momentum, to which we think we've both applied a lot of a priori skepticism. Protestations otherwise are marketing tactics and reflect an ongoing confusion between factor timing, which he believes in more than we do, and long term factor efficacy."

That kind of says it all but way too briefly and calmly for my taste; hence, this longer version you're reading now.

In the first part of the above quote I was making a very simple point. Rob and Research Affiliates publicly claim to believe in, and run investment products based on, factors that largely overlap with AQR. Now, for competitive reasons I wish they'd stop, but it's a free country. Value, low risk, and momentum are all things to various degrees we both believe in, and when it comes to investing in equities, it covers a large part of what each of us do.¹⁴ Check out how [his firm describes](#) one of its products. What exactly does he claim we believe in because

of data mining that isn't in his list here? I mean, if I and AQR are data miners, then [double data-mining on you](#) Rob!¹⁵ That he'd accuse us of being "insufficiently skeptical" about the dangers of data mining isn't just at odds with our long history of the exact opposite, but bat**** crazy when it's mostly the stuff he believes in too. I guess he's hoping nobody noticed. I noticed. I notice things, particularly when they are about me and they are so very noticeable.¹⁶

Let me be clear. Rob doesn't actually think the factors we at AQR believe in are data mined any more than he believes his own are data mined. That's a smokescreen. What Rob is doing here is the time honored strategy of the best defense is a good offense combined with the old adage about pounding the table when you've got nothing. Separate from this kerfuffle Rob has actually accused most of the field of applied finance of data mining in a very specific way and we've shown he's wrong (at least massively exaggerating). Apparently he doesn't like that so we have this deflection. Please note, he's not wrong that data mining is a big problem, everybody reasonable thinks that certainly including me. Whoever shouts it louder at others doesn't necessarily believe it more. But his very specific accusation against the field about a very specific type of data mining has no teeth. Unfortunately to understand this we have to get much more into the geeky weeds, sorry...

Rob has made claims in various papers that some of the major factors that much of the academic/practitioner world (not just AQR, and oddly again including him) has found to work historically are not just due to generic data mining but to a highly specific form of data mining he has uncovered. This is important. He's not just crying that people data mine, which is always a dangerous possibility. He thinks he's uncovered precisely the error they're making. The highly specific type of data mining that he alleges is that some of these factors have richened over the long haul leading to a one-time long term windfall to an investor in that factor (or, more likely, to the backtest of that factor). He claims researchers have mistaken this windfall for repeatable return. It's a good story that can indeed apply at relatively short horizons and at the peaks or troughs of major factor bubbles/depressions. But as I [exhaustively show](#), based on a technique from Rob and his colleagues' own papers, his very strong assertions are just wrong. The effect he discusses just doesn't affect the long term results enough to matter for the factors in question. In a nutshell, if a factor richens by 100% over 50 years (let's ignore compounding), a simplistic, and ultimately wrong, approach says the investor gets an extra 2% a year from this richening (this is indeed the approach Rob and colleagues focus on assuming the investor gets the full 2%). If true, this could indeed bias researchers to find and love factors that have richened. But, as Rob [shows himself](#) ¹⁷, (but then promptly ignores) if a factor richens like this you don't get nearly 2% a year in reality, and that difference is most severe for higher turnover factors (though the difference exists for them all). The factor at the end that is 100% more expensive than the factor at the beginning is composed of very different positions. It's not a single asset that's richened but rather turnover has led the factor investor to own a very different portfolio over time. As such, the investor in the richening factor benefits in a much smaller way as some, often a lot, of the richening didn't happen to the stocks they actually owned but to the stocks the strategy ended up owning at the end.¹⁸ When you adjust for this, over the long haul (like the 50+ years used by most credible researchers) none of the factors Arnott and crew examine are seriously harmed by adjusting for long-term richening.

Essentially, despite them (very briefly) flagging the turnover issue themselves and presenting the outline of a way to deal with it (fleshed out in great detail and utilized in my [previously referenced](#) work), they then completely ignore all this and still scream "bad researchers have data mined over factor richening and we're here to save the day!" They are saying that the entire rest of the field has missed something huge and important and has thus misled people. They say this despite their own evidence to the contrary that I expanded upon. I called him out [here](#) and in the last bullet [here](#) suggesting that they either retract their assertion or prove me wrong.¹⁹ He has ignored this as, in general, he references little relevant research by others and none of our recent work as he's writing repeated breathless white papers (i.e., watch out for the crash - we're at the 85th percentile!) rather than participating in give-and-take debate (I promise I'm really ok with both the give and the take as long as we are actually addressing each other's points not just taking shots in the media).

Let's try to be super clear with a flowchart:

Arnott accuses most of the industry (academics and practitioners) of a specific form of data mining based on their mistaking factor richening for true factor return.



Asness shows that this specific accusation is simply bad math (though, of course, acknowledging that other forms of data mining are always a concern). Asness repeatedly calls on Arnott to defend this specific broadside he undertook against academics and practitioners everywhere. Arnott, to date, declines.



Arnott calls Cliff and AQR data miners in recent Bloomberg/ Businessweek article, presumably as a deflection from the real debate on the very specific accusation he's made regarding data mining. Or perhaps it was just a revenge shot for past transgressions (i.e., Asness pointing out that both fundamental indexing and value-based factor timing are just systematic value investing and not new findings). Only the shadow knows.



Asness, befuddled, points out that Arnott believes in largely the same very limited robust set of factors as AQR/Asness, and thus wonders, in some awe, how Arnott could thus make that particular accusation with a straight face?



Arnott keeps face straight.



Asness writes this screed.

Rob wants to debate “Resolved: Data mining is a real problem” with himself in the affirmative and me cast as the doubter. That’s not the debate. We both agree on that, and in fact mostly on what factors pass the “it’s not data mining” test. The actual debate is “Resolved: Researchers massively mistake factor richening for true factor return” with Rob arguing for and me against the proposition. That’s a debate I’ve done my part in but where he’s yet to engage. If he wants to, and I’m wrong, so be it. But this ain’t it, and his comments ain’t right.

At AQR we pride ourselves on minimizing data mining. Nobody in our field is perfect on this front but we’ve had the discipline to walk away from good-looking factors we don’t trust. We have no desire to find things that have worked in the past that won’t work going forward. In many ways we’ve led this fight against data mining for many years. We believe in a small subset of things²⁰ that have worked out-of-sample through time, out-of-sample across geography, out-of-sample across asset class, and importantly, are explained by an economic story that’s not just “the data says so.” So, finding myself the victim of a drive-by shooting accusation of data mining by someone who believes in largely the same things we do, even if he occasionally renames them and claims them in the name of the Kingdom of Arnett, was pretty jarring. I would put my “respect for data” (and wood) and “sufficient cynicism about data mining” up against Rob’s any day (by the way, I’m not saying he doesn’t have it too).

Bottom line, his accusation that the industry has data mined over richening is false, his accusation that I’m a data miner is false (and particularly hypocritical), and I think he knows both these things and says them anyway.

Aside from that, Mrs. Asness and I very much enjoyed the play.

[1] I’m repeatedly told that the smart thing to do when you get press you don’t like is to ignore it. Consultants and other “professionals” say things like “you’ll only draw more attention to it.” Well that’s probably the right strategy when that bad press has a good point (not the case here!). It might even be the right strategy when that bad press is utterly false (the case here!). But, for better or worse, silence in the face of an unanswered falsehood, from people who should know better, is just never going to be how I roll... Did that sound cool? The “roll” thing? I was trying for cool.

[2] For instance, we defend the main Fama-French factors and momentum from the data mining charge.

[3] The closest person I ever met to believing nearly everything was data mining was Fischer Black. I remember a colleague presenting a model and Fischer kept, puckishly, saying “it’s an interesting DM model.” None of us could figure out what he meant. You see these were the pre-Euro days and we all kept thinking “but this model isn’t about the Deutsche Mark?” But, it was Fischer so we all kept assuming he knew something we didn’t. Eventually he told us what he meant. It was a bit of a let-down.

[4] Here’s an example of all three.

[5] And hopefully even leads to a theoretical model with other testable implications that can back-up or refute the story.

[6] Note, AQR does more than just straight factor investing, and I am not implying it’s this simple or limited everywhere. In fact, even in factor investing there are some things we think are still “alpha”, but we don’t write about them, and Rob wouldn’t know about them to critique them! We also think what you believe in is only part of the battle. There is a lot of “craftsmanship,” for want of a better word, in how one implements these beliefs; a discussion for another day.

[7] For stocks it’s actually only three: value, momentum, and quality. We lose one for stocks as the fourth, carry, is essentially the same as value for equities (it differs in important ways for bonds, FX, and commodities).

[8] Now, the number of ways we measure these factors is indeed larger than just three. That is, we use many separate measures for each of the three themes of value, momentum, and quality. That is not data mining. It’s about measuring the concept robustly – almost exactly the opposite action in spirit. Data mining is about overfitting randomness or errors. Since every measure of a factor contains error, averaging across a host of similar related measures can reduce those errors. It’s a separate more minor gripe but Robert Novy-Marx’s comments on our work (only touched on in this article which confusingly excerpts a longer back-and-forth) misses this point entirely.

[9] And one other guy who used to work with us, it didn’t work out, and now works with a rival – a more minor issue but definitely a pattern (two for two) for this article. Isn’t there anyone on Earth he could’ve found who thinks I’m ok? Was my mom not taking his calls? There’s a dude I had a playground fight with in 3rd grade. I called him names and he beat the crap out of me. I assume he was next on this reporter’s list.

[10] By the way, it’s ironic that Gross Profitability is one of the major factors Rob thinks is data mining and is also the major academic contribution of Robert Novy-Marx, the other guy cited in the article. Can’t they just argue with each other, cut out the middle-man, and leave me out of it?

[11] This is our second big public dispute. The [first came](#) when Rob renamed other people's work on value investing as "fundamental indexing" and declared it was barely related to value (he's since migrated that story somewhat back towards reality without ever acknowledging what he asserted for the first few years). Ironically part of our current dispute over factor timing is him finding value works (weakly) for this timing task and me pointing out, yet again, that yes, that's just systematic value investing itself, not anything new. Plus ça change...

[12] Sadly this is far from unique to AQR – I wish they were just ours!

[13] That was a mixed metaphor including honey, vinegar flavored whoop-ass whatever that is... Ok, it really doesn't hold together, sorry.

[14] It's likely that we do it in different proportions with different specific implementations, but this is still about basically shared beliefs.

[15] In fact, despite our several major disputes and his off-hand accusation here, "data miner" is not something I'd call Rob (nor would I say what Captain Kirk says in the link!)

[16] Now, if you want to accuse me of something that has more bite, may I suggest "Cliff gets way too emotional and, even if he's usually right, his temper tantrums really sometimes hurt him"? Yeah, you got me. I snuck in the "even if he's usually right" part. That's optional, but I recommend it.

[17] If you're curious look at the "regression coefficient-adjusted" lines [here](#) in the tables that he, for some reason, presents but then promptly ignores as he then repeatedly, and only, refers to the greatly exaggerated results instead.

[18] Turnover is perhaps the primary but not the only reason that factor richening does not add or subtract one for one from factor return. [Ilmanen, Nielsen, and Chandra \(2015\)](#) explore these other effects they call "wedges".

[19] Hey, if that can be done, I promise to acknowledge it and apologize. It's certainly possible that the answer is somewhere in between his wild claims and the math I show (I'm certainly capable of missing something!).

[20] Each, again, implemented with many robust measures if possible.

This document is not intended to, and does not relate specifically to any investment strategy or product that AQR offers. It is being provided merely to provide a framework to assist in the implementation of an investor's own analysis and an investor's own view on the topic discussed herein.

This document has been provided to you solely for information purposes and does not constitute an offer or solicitation of an offer or any advice or recommendation to purchase any securities or other financial instruments and may not be construed as such. The factual information set forth herein has been obtained or derived from sources believed by the author and AQR Capital Management, LLC ("AQR") to be reliable but it is not necessarily all-inclusive and is not guaranteed as to its accuracy and is not to be regarded as a representation or warranty, express or implied, as to the information's accuracy or completeness, nor should the attached information serve as the basis of any investment decision. This document is not to be reproduced or redistributed to any other person. The information set forth herein has been provided to you as secondary information and should not be the primary source for any investment or allocation decision. Past performance is not a guarantee of future performance. Diversification does not eliminate the risk of experiencing investment losses.

This material is not research and should not be treated as research. This paper does not represent valuation judgments with respect to any financial instrument, issuer, security or sector that may be described or referenced herein and does not represent a formal or official view of AQR. The views expressed reflect the current views as of the date hereof and neither the author nor AQR undertakes to advise you of any changes in the views expressed herein.

The information contained herein is only as current as of the date indicated, and may be superseded by subsequent market events or for other reasons. Charts and graphs provided herein are for illustrative purposes only. The information in this presentation has been developed internally and/or obtained from sources believed to be reliable; however, neither AQR nor the author guarantees the accuracy, adequacy or completeness of such information. Nothing contained herein constitutes investment, legal, tax or other advice nor is it to be relied on in making an investment or other decision. There can be no assurance that an investment strategy will be successful. Historic market trends are not reliable indicators of actual future market behavior or future performance of any particular investment which may differ materially, and should not be relied upon as such. Diversification does not eliminate the risk of experiencing investment losses.

The information in this paper may contain projections or other forward-looking statements regarding future events, targets, forecasts or expectations regarding the strategies described herein, and is only current as of the date indicated. There is no assurance that such events or targets will be achieved, and may be significantly different from that shown

here. The information in this document, including statements concerning financial market trends, is based on current market conditions, which will fluctuate and may be superseded by subsequent market events or for other reasons.