Citations overview

A graph of different colored bars

Description automatically generated with medium confidence

**A graph of a bar graph

Description automatically generated with medium confidence**

A screenshot of a graph

Description automatically generated

Issues

* Low Leverage is driven primarily by “at\_be” which is attributed to Fama and French (1992)
* We only count the “seminal” paper. That is, we include Fama and French (1992), but not Fama and French (1993). Big issue if we want to calculate the total number of citations.
* ~~Rosenberg and Reid need many more cites, no?~~
* Fama-MacBeth is credited with first discovering the low beta effect, but not clear its true.

In-sample versus out-of-sample

(in-sample: anything before paper is published, in contrast to JKP)

A graph of a black and white dotted line

Description automatically generated with medium confidence

A graph with dots and a line

Description automatically generated

Citations and Performance

A graph showing the growth of the stock market

Description automatically generatedA screenshot of a graph

Description automatically generated

A graph of a graph showing a number of different colored lines

Description automatically generated with medium confidence

A graph of bar charts

Description automatically generated with medium confidence

**Initial mail from Toby**

Here is the paper pitch:

“Visiting the Factor Zoo” or “A Guide Through the Factor Zoo”

* + - Re-examining all of the out of sample performance of all discovered anomalies with a couple of twists:
      * Weight anomalies by their citations.  Does the citation-weighted average of factors/anomalies look as dismal as the EW average suggested in other studies (Harvey, Liu, and Zhou (2015), McLean and Pontiff (2015), Hou, Xue, and Zhang (2017))?  Or, does the field/academic market kind of get it right?
        + Think of citations as the currency academia uses to assess the importance/validity/robustness of research.
      * Correlate citation count with how robust the factor evidence is:
        + In-sample t-stat
        + Out-of-sample (in time) performance
        + External validity (other/international equity markets)
        + External validity (other asset classes)
        + Internal validity (across size groups, within vs. across industries, etc.)
        + Robustness to measure (e.g., different flavors of value give similar performance)
      * Can also assess whether anomalies are *economically* robust as well as statistically robust?  It’s not enough to say it has a positive t-stat > 2, but is it meaningful economically?  Could actually use the tangency portfolio of existing factors to see if a new factor added has value – does an anomaly significantly move the efficient frontier?  Is that movement correlated with citation count?
      * Basically, the question is do the factors the field focuses on (in terms of citations) have the most robust out of sample evidence, but also the best in-sample evidence, biggest impact on the model/efficient frontier, etc.  And, conversely, what are the factors that we don’t think matter, do they have weak evidence on all of these dimensions and do they have poor cite counts?

(\*I already have a lot of the cite counts, but we could do a better job. You already have the machinery to look at external, internal validity, cluster by theme to assess measurement robustness, etc.

**Abstract**

The proliferation of the factor zoo and questions about the robustness and replicability of factors has received much recent attention in our field.  While most studies approach the subject through a statistical lens, with a lively debate on the conclusions drawn, we take a different approach to assessing the factor zoo using basic economics.  Specifically, using citation counts as the currency for academic work, we infer the “price” of each factor based on the academic market for publication.  Mapping citations of factors to their efficacy in and out of sample, statistical reliability, internal and external validity, and their replicability, we find that the most cited factors have the strongest evidence backing them – being more reliable, more robust, and more significant.  A citation-weighted index of factors performs remarkably well out of sample and is reliably significant and stable. An inverse-citation-weighted index of factors is not. Moreover, factor cites are highly right-skewed, indicating that only a handful of factors matter for most published work, consistent with common asset pricing models containing only a few factors rather than the entire zoo. Looking at the cumulative time-series of citations, we find that cite counts predict future Sharpe ratios, suggesting that knowledge of the efficient frontier has grown from the search for factor premia.  The collective evidence is consistent with the market for academic work being fairly efficient, with citations capturing the implied robustness and importance of various asset pricing factors..

Tests

* Citation-weighted index of factors (cumulative returns)
  + Also, inverse citation weights
* Correlate citation count with how robust the factor evidence is:
  + In-sample t-stat
  + Out-of-sample (in time) performance
  + External validity (other/international equity markets)
  + External validity (other asset classes)
  + Internal validity (across size groups, within vs. across industries, etc.)
  + Robustness to measure (e.g., different flavors of value give similar performance)

**Qs:**

* What to do about the “one paper per factor” stuff?
* What dimensions do we want to test? E.g. does citations predict:
  + Out-of-sample returns in the US
  + OOS returns internationally
  + (see the stuff toby wrote in the first email)
* How do we create normalized citations?
  + Maybe from the data Mark collected