NBER WORKING PAPER SERIES

DISCONTINUED POSITIVE FEEDBACK TRADING AND THE DECLINE OF MOMENTUM PROFITABILITY

Itzhak Ben-David
Jiacui Li
Andrea Rossi
Yang Song

Working Paper 28624
http://www.nber.org/papers/w28624

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 2021, Revised June 2021

Previously circulated as "Discontinued Positive Feedback Trading and the Decline in Asset Pricing Factor Profitability." We thank Nicholas Barberis, Lauren Cohen, Sylvester Flood (Morningstar), Umit Gurun, Paul Kaplan (Morningstar), Dong Lou, Chris Malloy, Andrei Shleifer for helpful comments. We thank seminar participants at The Ohio State University, the University of Utah, the University of Washington, Hong Kong University of Science and Technology, and Arrowstreet Capital, as well as the National Bureau of Economic Research Behavior Finance Workshop for comments and George Aragon for sharing data. Ben-David is with The Ohio State University and the National Bureau of Economic Research, Li is with University of Utah, Rossi is with the University of Arizona, and Song is with the University of Washington. Ben-David is a co-founder and a partner in an investment advisor. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Discontinued Positive Feedback Trading and the Decline of Momentum Profitability
Itzhak Ben-David, Jiacui Li, Andrea Rossi, and Yang Song
NBER Working Paper No. 28624
March 2021, Revised June 2021
JEL No. G11,G24,G41

ABSTRACT

We argue that the June 2002 reform in Morningstar's mutual fund rating methodology explains a sizeable amount of the profitability decline in momentum-related factors and factor momentum strategies. Before the reform, fund ratings heavily depended on recent investment style performance, and ratings-chasing flows led to large style-level positive feedback trading. The reform disrupted this process, and factors that benefit from positive feedback trading experienced a precipitous return decline. The performance decline was specific to momentum-related strategies and was also limited to the U.S. market where the reform happened. Further validating the mechanism, factors that are negatively affected by the reform experienced a sharp return "kink" in mid 2002, while the unaffected factors did not. We estimate that the reform explains approximately a third and two thirds of the post-2002 profitability drop in momentum-related factors and factor momentum, respectively.

Itzhak Ben-David
The Ohio State University
Fisher College of Business
606A Fisher Hall
Columbus, OH 43210-1144
and NBER
ben-david.1@osu.edu

Jiacui Li
David Eccles School of Business
University of Utah
SFEBB 8123, 1655 Campus Center Dr
Salt Lake City, UT 84112
jiacui.li@eccles.utah.edu

Andrea Rossi
University of Arizona
Eller College of Management
Department of Finance
1130 E. Helen St.
Tucson, AZ 85721
rossi2@email.arizona.edu

Yang Song
Foster School of Business
University of Washington
4273 E Stevens Way NE
Seattle, WA 98195
songy18@uw.edu

1 Introduction

Over the last four decades, asset pricing researchers have identified hundreds of factors (anomalies) in the cross-section of stock returns. However, the profitability of these factors has declined noticeably over time. In particular, returns of momentum-related factors suddenly dwindled since mid-2002. The decline in profitability is evident in Panel (a) of Figure 1, which plots the returns of the momentum factor, other factors related to momentum (e.g. industry momentum), as well as the time-series factor momentum strategy of Ehsani and Linnainmaa (2021). Visibly, the profitability of these strategies declined sharply after mid-2002; for the momentum factor, returns dropped from 0.92% per month during the earlier period of January 1987 to June 2002 to merely 0.16% afterwards. In addition to the long-term profitability decline, there also appears to be a sharp "kink" around mid 2002. These findings of long-term profitability decline and short-term "kink" are robustness to alternative factor construction methodologies and have also been foreshadowed by earlier studies.[1]

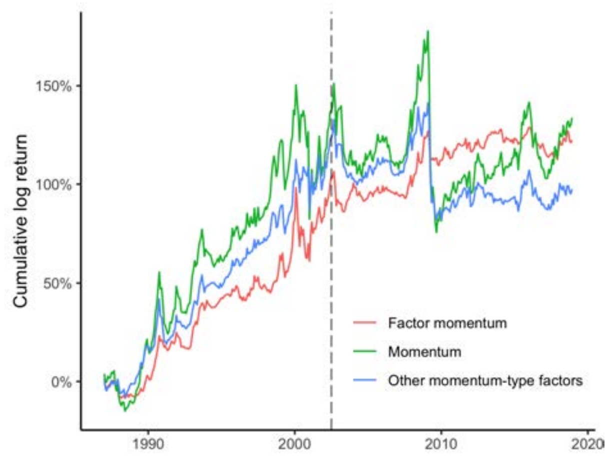
In this study, we argue that a seemingly innocuous institutional change—Morningstar's mutual fund rating methodology reform in June 2002—contributed to a non-negligible fraction of the post-2002 profitability decline of momentum-related strategies. Ben-David, Li, Rossi, and Song (2020a) show that Morningstar fund ratings drive correlated fund flows which create sizeable price pressures. In this paper, we further show that the pre-2002 methodology creates large positive feedback trading at the style-level, which contributes *additional* profits to strategies that benefit from positive feedback trading. Ceteris-paribus, the reform in 2002 caused a sudden halt to this process, negatively impacting the profits of both momentum-related factors and factor momentum strategies. Consistent with the fact that the methodology reform only happened in the U.S., this profitability decline is not seen in other countries. Using multiple methods, we estimate that the Morningstar reform can account for approximately one third and two thirds of the profitability decline in

[1] Appendix B.1 provides further details.

Figure 1. Morningstar Rating Methodology Change and Factor Return Decline

Panel (a) plots cumulative log returns of the $(t-12, t-1)$ momentum factor, other momentum-type factors, and the time-series factor momentum strategy. The other momentum-type factors include four classified to be in the momentum category by Hou et al. (2020): $(t-1,\ t-6)$ momentum (Jegadeesh and Titman, 1993), industry momentum (Moskowitz and Grinblatt, 1999), 52-week high (George and Hwang, 2004), and $(t-7,\ t-12)$ intermediate momentum (Novy-Marx, 2012). The factor momentum strategy is constructed following the methodology in Ehsani and Linnainmaa (2021) and based on our factor universe. Please see Section 2.2 and 5 for details. The vertical dashed line marks the Morningstar methodology reform in June 2002. Panel (b) examines the cross-section of post-reform profitability decline among 49 popular asset pricing factors. Factors are sorted into deciles based on their post-reform decline in exposure to Morningstar ratings, which is plotted on the horizontal axis by decile. The vertical axis plots the post-reform profitability decline. The green dashed line is the best linear fit. Rating exposure measures the degree to which each factor benefits from the positive feedback induced by Morningstar ratings and is defined in Section 2.3 (variable ExpSum(ΔRating)).

(a) Momentum-related factor returns (b) The cross-section of factor return decline



momentum-type factors and factor momentum, respectively.

How exactly does the Morningstar rating reform impact factor returns? Before the reform, ratings were based on each fund's past return ranking against all other equity funds. Because a significant fraction of fund returns were due to style exposures, funds in styles that recently performed well (poorly) receive high (low) ratings. Since ratings are a major driver of fund flows, flows to mutual funds with similar styles depended strongly on recent style performance. As funds scaled up or down their holdings in response to flows (e.g., Lou, 2012), their trading behavior caused substantial positive feedback effects at the style level. In June 2002, Morningstar reformed its methodology to rank funds *within* 3×3 size-value style categories. By doing so, ratings stopped depending on style returns, and the ratings-driven positive feedback trading at the style level came to a sudden halt.

Our empirical exercises, which investigate the impact of rating-induced feedback trading on the profitability of momentum-related strategies, proceed in four parts. As a preliminary step, we verify that style-level momentum indeed dwindled after June 2002. Before the reform, mutual funds in the style with top recent performance received flows that were higher *per month* by 1.7% of assets under management than funds in the bottom style. Consistent with the flows creating further price impact, stocks held by funds in the top style continue to outperform those in the bottom style by 84 bps per month. After the reform, the fund flow and return difference between the top and bottom styles became muted.

The first part of empirical exercises tests the prediction that the post-reform profitability decline should be most pronounced in momentum-type asset pricing factors. Existing work has identified many other mechanisms that cause factor returns to decline — changes in liquidity, arbitrage activity, or possible data-mining, and etc. Therefore, this part of analysis also includes a broad range of other asset pricing factors as controls: Compared to momentum-related factors, the average premium on the other factors should also be affected by mechanisms such as arbitrage activity, but not by the rating-induced positive feedback trading. We construct 49 popular long/short asset pricing factors which cover all the major categories.[2] To measure the exposure of factors to rating-induced effects, we follow Ben-David et al. (2020a) to measure rating exposure (a summary of recent rating changes) at the stock-level and aggregate it up to the factor level.

Consistent with our prediction, we find that momentum-related factors indeed suffered the largest drop in their rating exposure and also the largest profitability drop of approximately 0.80% per month after the reform date, while other factors only experienced a profitability drop of 0.29% on average. As an illustration, Panel (b) of Figure 1 shows that factors with larger decline in rating exposure (horizontal axis) experienced larger profitability declines after the reform (vertical axis).

[2]Following Hou et al. (2020), we use NYSE break points and value-weights to reduces the impact of microcap companies. Of these 49 factors, five are classified as of the momentum type by Hou et al. (2020).

To further test the rating-based mechanism, we conduct a region-based placebo test using the international versions of momentum factors constructed by other researchers. The 2002 Morningstar reform was specific to the US. Consistent with this fact, we do not observe similar declines in momentum profitability in other countries.

In the second part, we examine the impact on factor momentum strategies proposed by Ehsani and Linnainmaa (2021). Specifically, they show that strategies that go long or short in factors based on their recent twelve-month return generate excess returns. Because many factors have large and persistent size and value style loadings, factor momentum is highly related to style momentum, so we predict its profitability to also decline after mid 2002. Consistent with our conjecture, the rating exposure of factor momentum strategies declined dramatically after the reform, and time-series factor momentum returns based on our factor universe achieves 0.61% monthly returns before the reform and only 0.14% after. A similar result is obtained when using different factors constructed by other researchers. Finally, we also verify that this decline is unique to factor momentum based on U.S. factors.

In the third part, we use two methods to quantify how much of the post-2002 profitability decline can be attributed to the Morningstar reform. Following a standard procedure in asset pricing literature, we first create a long-short "rating factor" based on stock-level rating exposures. Spanning tests reveal that momentum-related factors and factor momentum indeed have high exposure to the rating factor, and according to this method, Morningstar can explain half and two thirds of the post-reform profitability decline of the momentum factor and factor momentum, respectively.

However, spanning tests are prone to over-estimate the explanatory power because they designate all returns *correlated* with the rating factor as "explained"—which is only valid in a statistical sense. We thus take a more direct approach to estimate the price impact of rating exposure using the 2002 shock, and then multiply that with the post-2002 decline of rating exposure of factors which can be directly measured. This method implies that the Morningstar reform can explain approximately one third of the decline of momentum-type factor returns and approximately two thirds of the decline in factor momentum.[3]

While the previous analyses focus on long-term return declines, in the last part, we zoom into a narrow window around the June 2002 reform to examine whether Morningstar ratings can explain the factor return "kink." We conduct an event study of factor-level ratings, fund flows, and returns. The reform caused exogenous style-level ratings changes in over 50% of funds, and the impact on factors is heterogeneous. Consistent with reform-induced rating changes driving factor price movements, only the factors that are affected by the reform experienced "kinks," i.e., sudden changes in flows and returns. The unaffected factors did not experience similar kinks. Using all other years other than 2002 as placebo tests, we confirm that the factor-level flow and return patterns documented are unique to 2002. Moreover, proxies for other possible influences on factor returns, such as arbitrage activity and liquidity, did not vary materially around the reform event. Overall, the findings indicate that the Morningstar reform can explain why some factors experienced a profitability kink in 2002, and also further shed light on the fact that rating-induced fund flows can have non-negligible impact on factor returns.

It is worth emphasizing that, as the quantification exercises reveal, our mechanism only explains *a subset* of the incremental decline of momentum-type strategies post 2002. Another major impact is the 2008–2009 momentum crash documented by Daniel and Moskowitz (2016). Thus, the mechanism we document should be thought of as a force that *exacerbated* momentum profitability during the relevant sample period; as the rating-induced feedback trading came to a stop, so did *a component* of momentum returns. More generally, momentum profitability is a complex phenomenon that likely defies a single explanation, and

[3]The fact that ratings have higher explanatory power on factor momentum is consistent with the mechanism: Morningstar only contributes to *style*-momentum but not *idiosyncratic*-momentum (Blitz, Huij, and Martens, 2011; Blitz, Hanauer, and Vidojevic, 2020). Because factors are diversified portfolios, factor momentum is not affected by idiosyncratic momentum, while stock momentum is.

many mechanisms have been proposed to date.[4] Our mechanism, however, stands out in its concreteness: we tie return predictability to directly measurable quantities of ratings and flows, and base our explanation on demand-based price effects which have been independently validated in several other studies (e.g., see Gabaix and Koijen, 2020; Ben-David et al., 2020a).

This paper contributes to the understanding of why factor strategies become less profitable over time. While existing studies emphasize the roles of liquidity (Khandani and Lo, 2011; Chordia, Subrahmanyam, and Tong, 2014; Lee and Ogden, 2015), arbitrage activity (Marquering, Nisser, and Valla, 2006; Green, Hand, and Soliman, 2011; Hanson and Sunderam, 2013; McLean and Pontiff, 2016; Calluzzo, Moneta, and Topaloglu, 2019; Cho, 2020), and possible data-mining (Harvey, Liu, and Zhu, 2016; Harvey, 2017; Hou et al., 2020; Huang, Song, and Xiang, 2020b; Falck, Rej, and Thesmar, 2021), we show that profitability decline can also arise from the removal of demand pressures that contributed to the profitability in the first place. The existing explanations are not mutually exclusive to ours. However, while the existing mechanisms are likely important contributors to the long-term factor profitability decline, they do not explain why profitability of momentum-related strategies dropped sharply after mid 2002—an event which also helps the identification of the Morningstar-based mechanism.

This paper is also related to recent attempts to understand the impact of demand on *systematic* components of asset prices. While the work on index composition changes clearly show that demand can impact the prices of individual stocks (Harris and Gurel, 1986; Shleifer, 1986; Wurgler and Zhuravskaya, 2002; Chang, Hong, and Liskovich, 2015), there is relatively less consensus on whether and how demand can shape factor-level price movements. This paper contributes to this line of inquiry by showing that ratings-induced demand can con-

[4] The other possible mechanisms include delayed information diffusion (Hong and Stein, 1999), behavioral biases (Daniel, Hirshleifer, and Subrahmanyam, 1998; Barberis, Shleifer, and Vishny, 1998; George and Hwang, 2004; Grinblatt and Han, 2005), as well as investor attention and media influences (Lee and Swaminathan, 2000; Hou, Xiong, and Peng, 2009; Hillert, Jacobs, and Müller, 2014). The literature on momentum is vast, and we cannot cover all of the explanations. Please see Jegadeesh and Titman (2011) for a survey of the literature.

tribute the expected return of asset pricing factors.[5] Our explanation is, thus, at sharp contrast to standard explanations of factor return profitability based on compensation for risk (Cochrane, 2011). Other studies on demand-based price effects use mutual fund flows (Teo and Woo, 2004; Coval and Stafford, 2007; Lou, 2012; Huang, Song, and Xiang, 2020a; Li, 2020), exchange-traded fund flows (Ben-David, Franzoni, and Moussawi, 2018; Brown, Davies, and Ringgenberg, 2021), and other sources of institutional investor demand (Parker, Schoar, and Sun, 2020; Ben-David, Franzoni, Moussawi, and Sedunov, 2021). More recently, Koijen and Yogo (2019) develop a structural methodology to estimate price impact, and Gabaix and Koijen (2020) show that the demand-induced price impact coefficient at the aggregate market level is large.

Our rating-induced positive feedback mechanism is consistent with the style-investing hypothesis in Barberis and Shleifer (2003). Positive feedback trading has also been identified in Teo and Woo (2004), Lou (2012), and Wahal and Yavuz (2013). Relative to these studies, we contribute by using an exogenous event to identify the price impact of rating-induced trading and by quantifying the explanatory power of our mechanism.

The rest of the paper is organized as follows. Section 2 details the data, factor universe, and variable construction. Section 3 explains how the Morningstar reform disrupts style-level positive feedback trading and makes testable predictions. Section 4 examines the impact of the reform on asset pricing factors, and Section 5 examines the impact on factor momentum. Section 6 conducts an event study around the reform date and Section 7 concludes. Robustness checks and additional tests are provided in the Appendix.

2 Data and Variable Construction

This section describes the data, our universe of asset pricing factors, and how we measure the impact of ratings on those factors.

[5] This paper focuses on the impact on *expected returns* of factors. Also using the 2002 Morningstar reform event, Ben-David et al. (2020a) show that correlated demand can exert large influence on *price fluctuations* of style portfolios.

2.1 Mutual Fund Data

We obtain monthly fund returns and total net assets (TNA) from the CRSP survivorship bias-free mutual fund data set. We use all U.S. domestic equity mutual funds. While funds are often marketed to different clients through different share classes, they invest in the same fund portfolio and typically only differ in the fee structure. Therefore, we aggregate all share classes at the fund level using Russ Wermers's MFLINKS (Wermers, 2000). We also obtain quarterly fund holdings from Thomson Reuters' S12, which is based on 13F filings.

We obtain Morningstar ratings and fund style categories from Morningstar Direct, and we merge them with the CRSP data using the matching table from Pastor, Stambaugh, and Taylor (2020). Because Morningstar assigns ratings at the share class level, we aggregate ratings at the fund level by TNA-weighting different share classes following Barber, Huang, and Odean (2016). We restrict our analysis to mutual funds with at least \$1 million TNA, and we winsorize fund flows at the 0.5% and 99.5% levels. We require the existence of 12 lags of monthly flows, returns, and ratings.

Most of our exercises start in January 1987 because Morningstar ratings are available from December of 1985, and the rating exposure variable ($\text{ExpSum}(\Delta\text{Rating})_{i,t-1}$, described in Section 2.3) uses 13 lagged months to construct. When requiring fund flow data, our sample starts in 1991 due to availability of monthly fund flow in CRSP.

2.2 Asset Pricing Factors

The main U.S. factor universe. We compute 49 popular stock-level characteristics that have been shown to predict returns; our choice of factors mostly follows Arnott, Clements, Kalesnik, and Linnainmaa (2019), and we restrict our attention to those that can be constructed using CRSP and Compustat data. Using the classification categories proposed in Hou et al. (2020), these 49 characteristics-based factors include 14 in the profitability category (e.g., return on assets), 13 in the investments category (e.g., share issuance), eight in the value/growth category (e.g., book-to-market ratio), six in the intangibles category (e.g., industry concentration), five in the momentum category (e.g., momentum of Jegadeesh and Titman, 1993), and three in the trading frictions category (e.g., Amihud illiquidity).

We follow the prescription in Hou et al. (2020) to limit the impact of microcaps in factor construction. Specifically, we use NYSE breakpoints to sort stocks into characteristics-based quintiles and then form value-weighted long-short factors. Appendix Table A.1 lists all asset pricing factors used in this paper.

International factors. As explained in Section 3, we expect the rating-induced post-reform profitability decline to be concentrated in U.S.-based factors, so non-U.S. factors can be used as a placebo test. For this purpose, we download the monthly global factor returns made available by Jensen, Kelly, and Pedersen (2021).[6] For constructing factor momentum strategies, we require lagged 12 months of returns and restrict attention to factors that have no data gaps since inception. The number of factors and countries covered gradually increase over the sample period, rising from 495 factors from 21 countries in 1987 to 3,615 factors from 26 countries by 2018; Appendix Figure A.1 and Table A.2 provide more details.

Partially due to differences between U.S. and non-U.S. markets, the factor construction methodology of Jensen et al. (2021) differs slightly from ours. In particular, they form long-short tercile portfolios. We use both the equal-weighted and the "capped value-weighted" returns they compute, where the latter is market value-weighted but also cap the market weight at the 80th NYSE percentile. The capping is intended to avoid one mega stock does not dominate a portfolio and is particularly relevant for less developed markets with fewer stocks. For brevity, we refer readers to the description in Jensen et al. (2021) for more details.

[6] Because this data source only contains factor returns, we cannot use their factors for our main U.S. factor-based exercise because we also need stock-level characteristics to compute the rating exposure measure. We downloaded this global factors data from Bryan Kelly's website on May 27th, 2021.

2.3 Measuring Rating Exposure and Flow-Induced Trading

We are interested in how Morningstar ratings and rating-induced fund flows lead to price pressure on asset pricing factors. To this end, we first measure ratings and flows at the stock level, and then we aggregate them up to the factor level.

Rating exposure. Ben-David et al. (2020a) show that Morningstar ratings induce fund flows which create stock price pressures. As a consequence, they find that a simple measure of "rating exposure"—an exponential sum of recent rating changes—strongly predict returns at the stock level. We follow their specification to measure rating exposure and briefly explain the methodology for the reader's convenience. We first define the average Morningstar rating of stock i in month t as the holding-weighted rating of all funds $\mathcal{J}$ that hold the stock:[7]

$$\text{Rating}_{i,t} = \frac{\sum_{\text{fund} j \in \mathcal{J}} \text{SharesHeld}_{i,j,t-1} \cdot \text{Rating}_{j,t}}{\sum_{\text{fund} j \in \mathcal{J}} \text{SharesHeld}^{\text{fund}}_{i,j,t-1}} \tag{1}$$

We then summarize the recent 12 months of stock-level rating changes with exponentially decaying weights:

$$\text{ExpSum}(\Delta\text{Rating})_{i,t-1} = \sum_{k=1}^{12} \tau_k \cdot (\text{Rating}_{i,t-k} - \text{Rating}_{i,t-k-1}), \tag{2}$$

where $\tau_k = \frac{12\cdot(1-\delta)}{1-\delta^{12}} \cdot \delta^{k-1}$ and $\sum_{k=1}^{12} \tau_k = 12$. The decay factor $\delta = 0.76$ implies a half-life of $\approx$ 2.58 months. Because the weights sum to 12 (months), in terms of units, ExpSum(ΔRating) should be interpreted as measuring a re-weighted version of rating changes over the previous year. Ben-David et al. (2020a) show that this measure strongly predicts stock returns, and the predictability is not sensitive to reasonable variations in the look-back horizon or weighting scheme.

Then, for each factor f, we measure its rating exposure by aggregating up the stock-level

[7] Note that Morningstar assigns ratings for mutual funds. The stock-level ratings are computed by us.

exposures:

$$\text{ExpSum}(\Delta\text{Rating})_{f,t-1} = \sum_{i \in \text{top quintile}} w^f_{i,t-1} \cdot \text{ExpSum}(\Delta\text{Rating})_{i,t-1} - \sum_{i \in \text{bottom quintile}} w^f_{i,t-1} \cdot \text{ExpSum}(\Delta\text{Rating})_{i,t-1} \tag{3}$$

where $w^f_{i,t-1}$ is the lagged market cap weight of stock i in the corresponding quintile portfolio.

Flow-induced trading. We also want to measure the amount of fund flow-induced trading in each factor. We first follow Lou (2012) to calculate flow-induced trading (FIT) for each stock i in each month t:

$$\text{FIT}_{i,t} = \frac{\sum_{\text{fund} j \in \mathcal{J}} \text{SharesHeld}_{i,j,t-1} \cdot \text{Flow}_{j,t}}{\sum_{\text{fund} j \in \mathcal{J}} \text{SharesHeld}_{i,j,t-1}}. \tag{4}$$

Here, flow of fund j in month t is defined as the net flow into the fund divided by the lagged TNA, following the literature (e.g., Coval and Stafford, 2007):

$$\text{Flow}_{j,t} = \frac{\text{TNA}_{j,t}}{\text{TNA}_{j,t-1}} - (1 + \text{Ret}_{j,t}). \tag{5}$$

In short, FIT is the total amount of non-discretionary mutual fund trading in stock i caused by fund flows. As explained in Lou (2012), whereas discretionary trading is likely to be related to fundamentals, FIT isolates the nondiscretionary trading that is only attributable to fund flows. We then aggregate FIT at the factor level following the same method as in Equation (3). Summary statistics for the stock- and factor-level data are in Panels A and B of Table 1.

Table 1. Summary Statistics

Panels A and B present summary statistics of the monthly data at the stock- and factor-levels, respectively. Per Equation (4), FIT is defined as the amount of mutual fund trading induced by fund flows as a fraction of shares held. Morningstar rating is measured in stars (1 to 5), and ExpSum(Δ Rating) is an exponentially weighted sum of the past 12 months of rating changes (defined in Equation (2)). Obs is the average number of observations per month. The last five columns report 1%, 25%, 50%, 75%, and 99% percentile distributions, respectively. Panel C reports the average number of global factors in different regions in Jensen et al. (2021) after applying the data filter described in Section 2.2.

Panel A: Stock-level summary statistics

	Obs	Mean	Std dev	1%	25%	50%	75%	99%
Market cap ($m)	4,405	3,443	16,172	7	110	396	1,587	56,645
Held by num funds	4,405	78.5	104.0	1.0	11.0	36.0	110.0	480.0
Return	4,405	1.10%	15.62%	−38.46%	−6.01%	0.44%	7.04%	52.24%
Rating	4,405	3.369	0.714	1.181	3.000	3.441	3.850	5.000
ExpSum(ΔRating)	4,405	−0.028	0.767	−2.358	−0.306	0.000	0.253	2.268
FIT	4,405	0.55%	2.44%	−4.49%	−0.46%	0.28%	1.22%	8.68%

Panel B: Factor-level summary statistics

	Obs	Mean	Std Dev	1%	25%	50%	75%	99%
Return	49	0.23%	3.69%	−9.94%	−1.60%	0.15%	2.00%	10.71%
Rating	49	0.018	0.207	−0.700	−0.062	0.016	0.103	0.554
ExpSum(ΔRating)	49	0.025	0.256	−0.745	−0.057	0.009	0.089	0.881
FIT	49	0.04%	0.43%	−1.23%	−0.14%	0.03%	0.20%	1.38%

Panel C: Number of global factors in Jensen et al. (2021)

Period	U.S.	Developed ex U.S.	Emerging	Europe	APAC ex Japan	Japan
1987–1990	152	844	38	595	176	111
1991–1994	152	1,980	154	1,412	592	129
1995–1998	152	2,627	244	1,841	898	132
1999–2002	152	2,864	319	1,981	1,067	136
2003–2006	152	2,967	367	2,047	1,151	137
2007–2010	152	3,045	389	2,103	1,193	139
2011–2014	152	3,058	397	2,110	1,205	139
2015–2018	152	3,063	399	2,113	1,211	139

3 Morningstar Rating Reform and The Disruption of Style-Level Positive-Feedback Trading

In this section, we describe the Morningstar rating methodology reform in June 2002 and explain why it led to a disruption of style-level positive feedback trading. Based on this mechanism, we then make testable predictions for the subsequent paper.

3.1 2002 Rating Methodology Reform

We now describe the Morningstar rating methodology reform in June 2002.

Methodology before the reform. After introducing its mutual fund rating system in 1985, Morningstar quickly became the industry leader in providing independent mutual fund ratings. To assign ratings, Morningstar first summarizes the past return performance of funds and conducts minor adjustments for total return volatility and expenses. Depending on the availability of data, the look-back horizon for past performance can be up to ten years, but more weight is applied to more recent periods.[8] Then, Morningstar ranks funds by their performance and assigns 1 to 5 star ratings with fixed proportions (10%, 22.5%, 35%, 22.5%, and 10%).[9]

The reform. While the rating methodology has been very stable over time, Morningstar implemented a major reform in June 2002.[10] After the reform, fund ratings were no longer based on how each fund ranked against *all* U.S. equity funds but only on fund rankings *within* style categories. For diverse U.S. equity funds (87% of all mutual funds in 2002), the style categories are the well-known 3×3 size-value matrix.[11] The change in methodology was announced in February 2002 and was first implemented in Morningstar's monthly ranking of funds at the end of June 2002.

This seemingly innocuous change had far-reaching consequences for the mutual fund industry. Before the change, fund ratings differed dramatically across styles based on recent style performance, as shown in Panel (a) of Figure 2 which plots the rating dispersion of

[8] For funds with over 10 years of history, Morningstar computes 3-year, 5-year, and 10-year past returns and combines them. The weights of the three horizons are set at 20%, 30%, and 50%, respectively. Because the three horizons are overlapping, however, the recent years are effectively given much more weight than more distant history.

[9] The Morningstar methodology is fully transparent. Appendix B of Ben-David et al. (2020a) provides further detail on the exact computation.

[10] The change was partially motivated by complaints from fund managers, arguing that they receive low ratings simply because their investment style performed poorly, but not because how they managed the funds. Please see Section 3 of Ben-David, Li, Rossi, and Song (2020b) for more details.

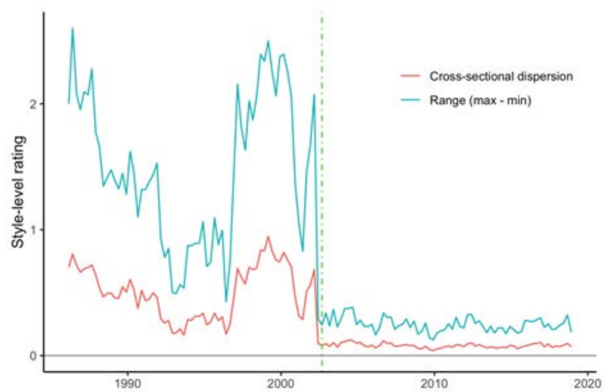
[11] Sector funds—the remaining 13%—were classified into 12 sectors (e.g., financials, utilities).

3 × 3 size-value fund styles. In the months before the methodology change, the top and bottom rated styles differed by up to 2 stars. After the reform, that difference dropped dramatically and ratings also become uncorrelated with past style performance.[12] Panel (b) plots the dispersion of style-level fund flows. Consistent with flows chasing ratings being a major driver of fund flows, style-level fund flows also became less dispersed after the reform.

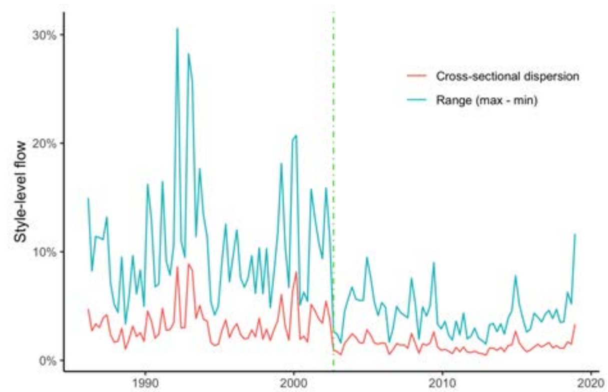
Figure 2. The Morningstar Methodology Reform and Style-Level Flows

Panel (a) and (b) plot the dispersion of quarterly fund ratings and TNA-weighted average fund flows by the 3 × 3 size-value Morningstar styles. Dispersion is measured either as cross-sectional standard deviation (red lines) or the difference between maximum and minimum values (blue lines). The vertical dashed line marks the June 2002 Morningstar methodology reform event.

(a) Style-level rating dispersion



(b) Style-level fund flow dispersion



Importantly for our identification purposes, investors continued to chase ratings in a similar manner before and after the reform. This has been shown by Ben-David, Li, Rossi, and Song (2019); Evans and Sun (2021); Ben-David et al. (2020a).[13] Therefore, the reform effectively re-directed fund flows to stop chasing style-level returns.

3.2 Disruption of Style-Level Positive-Feedback Trading

We now demonstrate that the 2002 reform disrupted style-level positive feedback trading. Based on this disruption, we conjecture that the reform should reduce the profitability of

[12]One may wonder why rating dispersion did not drop to exactly zero. A major reason is because Morningstar assigns ratings at share-class level, so taking an average over share classes would bring the dispersion to zero, but we compute average ratings at the fund-level.

[13]See, for example, Figure 1(b) and Figure 4(b) in Ben-David et al. (2020a).

momentum-type factors and factor momentum, and that the reduction should be specific to U.S.-based factors.

The pre-reform rating methodology generates a positive feedback loop at the style-level. This is illustrated in Panel (a) of Figure 3: funds in styles that performed well in the recent past get high ratings and attract inflows. Funds use the new flows to increase their investments in the same style of stocks, so the price of those stocks are pushed up even further. The mechanism also works in the other direction: funds in underperforming styles experience correlated outflows, resulting in downward price pressure on stocks associated with these styles. The post-June 2002 rating methodology, however, should cause a sudden disruption in this rating-induced positive feedback trading at the style level.

We confirm this style-level disruption in Panels (b) and (c) in Figure 3. Specifically, we sort the 3×3 Morningstar fund styles based on past-12-month returns—the typical look-back horizon used in studying momentum. Before the reform, funds in styles that recently performed well received higher average ratings and higher fund flows. The magnitudes are also large. Panel (b) shows that the average rating spread between funds in the top and bottom styles was about 0.8 stars before reform and shrank to almost zero after reform. Because rating attracts flows, Panel (c) shows that funds in the top style received about 1.7% higher flows per month than the bottom style before the reform, and that difference dropped to around 0.4% after the reform.[14] This disruption also has an impact on style returns. In Panel (d), we plot the TNA-weighted style-level fund returns. The top-ranked style exhibit approximately 0.8% higher monthly return than the bottom-ranked style before reform, and that difference disappeared after reform.[15]

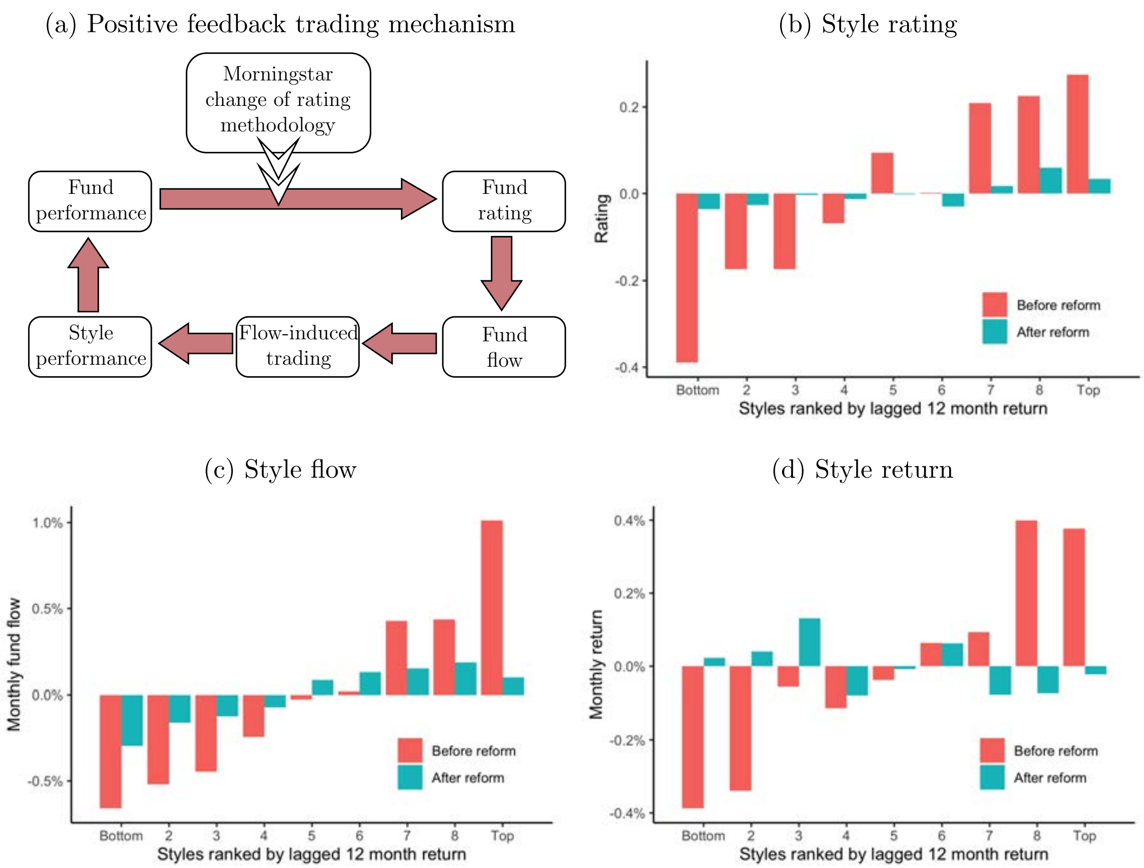
These findings lead us to conjecture that the 2002 Morningstar reform had a negative impact on strategies that load on style momentum, which includes several momentum-type

[14] The data in those graphs are demeaned within month to focus on cross-sectional patterns across styles.

[15] In unreported robustness checks, we find similar patterns when measuring returns using CAPM alpha, and the post-reform change in alpha spread is statistically significant at the 5% level. To alleviate the concern that fund returns may also be influenced by transaction costs and fees, we also repeated this exercise using the returns of the stocks held by the funds, rather than the fund returns. The results are unaffected.

Figure 3. Style-level Positive Feedback Trading Before and After Reform

This figure shows that the style-level positive feedback trading largely halted after the Morningstar methodology change in June 2002. The flow chart in Panel (a) illustrates how pre-2002 ratings generate positive style-level positive feedback trading. In panels (b) to (d), we sort the 3×3 Morningstar styles by their lagged 12 month returns. Panels (b) and (c) plots the TNA-weighted average rating and fund flows of the sorted styles. Panel (d) plot the return of funds in those styles. All variables are demeaned to focus on the cross-sectional difference across styles. This sample starts from 1991 due to monthly flow data availability in CRSP.



factors and the factor momentum strategies in (Ehsani and Linnainmaa, 2021). In the 49 factors we study, 5 are classified into the momentum category by Hou et al. (2020): $(t-1,\ t-12)$ and $(t-1,\ t-6)$ momentum (Jegadeesh and Titman, 1993), industry momentum (Moskowitz and Grinblatt, 1999), 52-week high (George and Hwang, 2004), and $(t-7,\ t-12)$ intermediate momentum (Novy-Marx, 2012). We predict that these factors should suffer large profitability declines after June 2002, and we will test this prediction in Section 4.

The reform does not impact idiosyncratic-level positive feedback trading. It is worth emphasizing that standard stock momentum strategies contain both a style-component and an idiosyncratic component. The former refers to the fact that styles with higher past returns continue to do so. The latter refers to the fact that, even after controlling for style-level effects, stocks with higher *idiosyncratic* past returns also have higher returns in the future (Blitz et al., 2011).

While the Morningstar reform disrupted style-level momentum, it did not disrupt idiosyncratic momentum. This is because the positive feedback mechanism we study works through the fund flows induced by ratings. The average stock is held by 78.5 funds, so for any given stock, there has to be a *correlated* change in the ratings of funds holding that stock in order to generate a sufficiently large rating-induced flow pressure. Therefore, while past style-level returns—which can induce correlated fund return changes—can have a large impact on a stock's rating, past idiosyncratic stock returns do not.

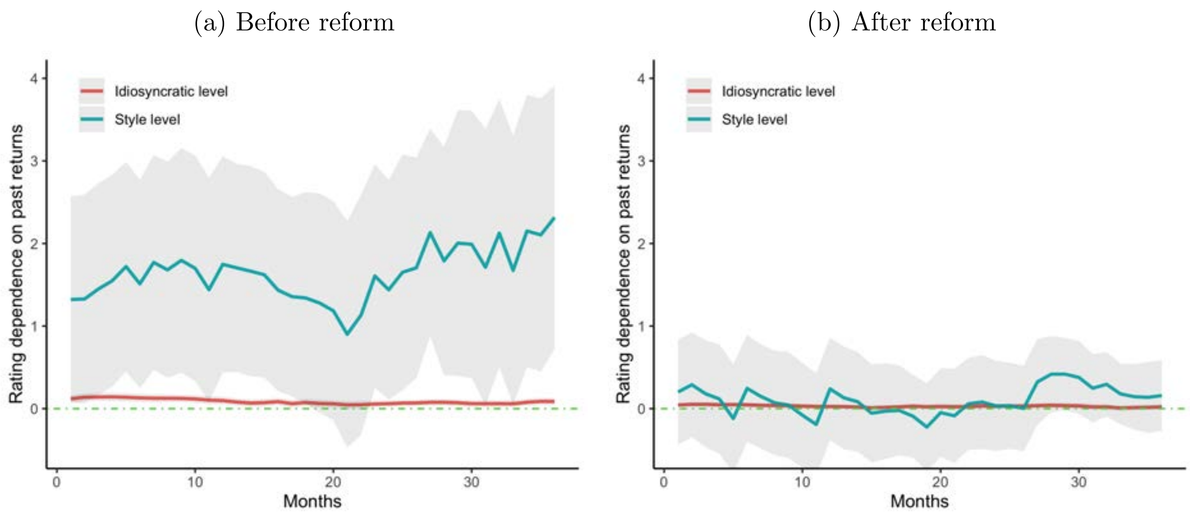
For a concrete example, consider a small cap growth stock that is held by many small cap growth funds. Suppose the stock's idiosyncratic return was high in the recent past. Because that stock is only a small part of each fund's portfolio, this shock is unlikely to have a sufficiently large effect on fund ratings. In contrast, suppose the style-level (small cap growth) return was high in the recent past. Under the pre-reform methodology, this means that all small cap funds would have performed well and thus receive higher ratings, leading to more positive feedback fund flows into all small cap growth stocks. After the methodology reform, this style-level positive feedback trading became muted by design.

Figure 4 illustrates these points using panel regressions of stock-level ratings on past 36 monthly lags of stock returns. To separately estimate the impact of different return components, we decompose each stock's return into:

$$\text{Ret}_{i,t} = \text{StyleRet}_{i,t} + \text{IdiosyncraticRet}_{i,t} \tag{6}$$

Figure 4. Morningstar Reform Only Impacted Style-Level Positive Feedback

This figure plots the panel regression coefficients of stock-level ratings (Equation (1)) on past 36 lags of monthly stock returns, which have been decomposed into style-level returns (3×3 Fama-French size-book/market styles) and idiosyncratic-level returns (the residual). Panel (a) and (b) plots the regression coefficients and the shaded areas represent 95% confidence intervals. The regressions control for month fixed effects and cluster standard errors by month.



where $\text{StyleRet}_{i,t}$ is defined as the market cap-weighted averaged return of the corresponding 3×3 Fama-French size-book/market style portfolio, and $\text{IdiosyncraticRet}_{i,t}$ is the residual. We regress stock ratings on 36 lags of each of these two components, controlling for month fixed effects, and plot the coefficients in Figure 4. Panel (a) shows that, before the reform, stock ratings heavily depended on past style-level returns but not idiosyncratic returns. This confirms that the Morningstar-induced positive feedback trading happens exclusively at the style-level. Panel (b) shows that, after the reform, the rating dependence on past style returns becomes muted.

This has important implications for the impact of this reform. The degree to which a trading strategy is impacted should *only* depend on its relationship with style momentum. As explained by Blitz et al. (2011), stock momentum includes both style-level momentum and idiosyncratic momentum. That is, in addition to style-level returns exhibiting positive autocorrelation, idiosyncratic components of returns also do. In contrast, because factors tend to be diversified portfolios in which idiosyncratic returns cancel out, factor momentum strategies do not load onto idiosyncratic momentum.

Testable Predictions. Based on the discussion in this section, we make three predictions.

1. **Disruption of momentum-related strategies.** Relative to other factors that do not depend on positive feedback trading, the rating exposure and profitability of momentum-related factors and factor momentum should decline more after the reform.

2. **Disruption is specific to the U.S.** Because the Morningstar reform is specific to the U.S.,[16] the post-reform profitability decline of momentum-related strategies should be concentrated in the U.S. market.

3. **Explanatory power.** The disruption of rating-induced positive feedback mechanism should explain a larger post-reform profitability decline for factor momentum, which primarily loads onto style momentum, than for the stock momentum factor, which also reflects idiosyncratic momentum.

It is worth emphasizing that our first two conjectures should be intended as ceteris-paribus predictions. As argued in the introduction, there are many other reasons that lead to time-varying changes in factor profitability. We are interested in the *incremental* impact of the disruption in feedback trading due to the Morningstar reform.

4 Effect of Reform on Factor Returns

In this section, we study the 49 asset pricing factors' rating exposure and profitability before and after the Morningstar reform. The non-momentum-related factors serve as controls, as they should be relatively unaffected by the reform.

[16] Appendix 2 of Morningstar (2016) lists all the historical major Morningstar rating methodology changes. The June 2002 change is unique to the U.S. market.

4.1 Which Factors Suffered Larger Profitability Declines?

As described in Section 2.3, we measure a factor's exposure to Morningstar ratings using $\text{ExpSum}(\Delta\text{Rating})_{f,t-1}$ defined in equation (3). In Panel (a) of Figure 5, we plot each factor's average post-reform rating exposure against the pre-reform values over the sample of 1987 to 2018. We mark factors from different categories using different colors. Clearly, before the reform, Morningstar served as an important tailwind for factors in the momentum categories (colored blue). After the reform, rating exposures of all factors shrunk. This plot is consistent with our conjecture that momentum-type factors were most positively affected by rating-induced positive-feedback trading before the reform, and they suffered the largest rating drop after the reform.

To visualize the cross-sectional differences across factor returns, in Panels (b) and (c), we plot the pre- and post-reform average factor return against the pre-reform rating exposure. Consistent with our prediction, factors that benefit from pre-reform rating exposure experienced high returns before the reform but not afterwards. For instance, the profits of the momentum factor was almost 1% each month but became negligible after June 2002. Other momentum-type factors, such as the 52-week-high factor, suffered similar declines in profitability.

Placebo Test: Momentum in Other Countries. We now test whether the post-reform momentum profitability drop is specific to the U.S. We use two sets of factor data constructed by other researchers. Because they both have different factor construction methodology, to be consistent, we use their versions of U.S. momentum factors for the comparison.

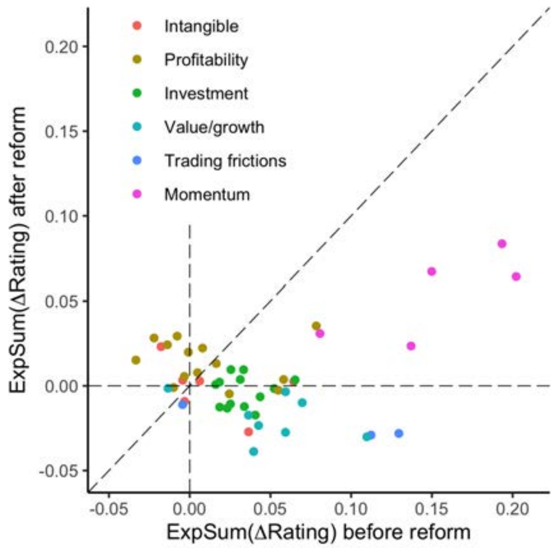
We first use the monthly momentum factors from Ken French's website—a standard data source for factor-based research.[17] In Panel A of Table 2, we compare his version of U.S. momentum factor against momentum in other developed markets, emerging markets, and

[17] His factor construction methodology slightly differs from ours, but the difference is not huge: the monthly correlation between his and our U.S. momentum factor is 96%. Specifically, the Fama-French construction forms 2×3 size-prior return independent sorts and defines the momentum factor as 1/2 (Small High + Big High) - 1/2 (Small Low + Big Low).

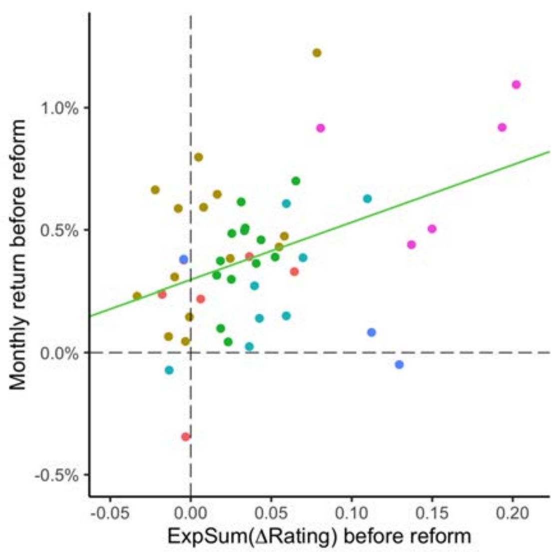
Figure 5. Factor Ratings and Return before versus after Reform

We compare factor statistics before and after the Morningstar methodology reform in June 2002. Panel (a) plots the post-reform ExpSum(ΔRating) (the exponentially-weighted sum of past-12-month rating changes) against the pre-reform values. Panels (b) and (c) plot average monthly factor returns before and after reform against pre-2002 ExpSum(ΔRating). The green lines in Panels (b) and (c) are best linear fits. The different colors for the data points represent the return factor classifications in Hou et al. (2020).

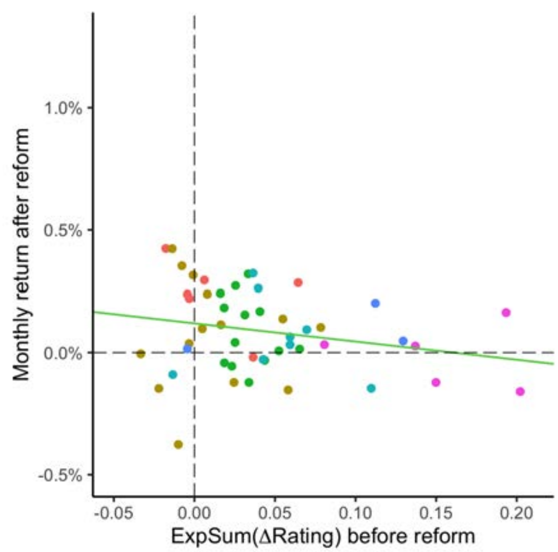
(a) Rating exposure before and after reform



(b) Factor returns before reform



(c) Factor returns after reform



also across other regions. Consistent with our prediction, only the U.S.-based momentum strategy experienced a large decline in profitability after 2002. In contrast, momentum profits were strong both before and after reform in all other markets except Japan.[18]

For robustness, we also produce the equivalence of Panel A using data from Jensen et al. (2021).[19] Panel B examines their equal-weighted returns while Panel C examines capped value-weighted returns. The conclusions are qualitatively unchanged. In fact, when judged using equal-weighted returns, momentum profits actually increased across the board, and U.S. is the only region where returns decreased, and the decline is statistically significant at the 5% level. Overall, these results are consistent with the Morningstar rating-based mechanism only disrupting U.S.-based momentum profitability.

4.2 Quantifying Explanatory Power of Ratings

How much of the decline in momentum-type factor profitability can be explained by the discontinuation of rating-induced feedback trading? In this section, we estimate the explanatory power using two methodologies, each of which has benefits and drawbacks:

1. **Spanning tests**: we form a "rating factor" and examine how much of the other factor returns can be spanned by it.

    - While this is commonly used in factor-based asset pricing, this approach may over-state explanatory power, as it effectively attributes all returns correlated with the rating factor as "explained," which is only valid in a statistical sense.

2. **Direct estimation:** we first estimate the price impact coefficient of ratings (λ) using the 2002 shock, which is well identified, and multiply it with the post-reform change in rating exposure of factors:

$$\underbrace{\lambda}_{\text{Price impact of ratings}} \times \left(\overline{\text{ExpSum}(\Delta\text{Rating})}_{f,\text{after 2002}} - \overline{\text{ExpSum}(\Delta\text{Rating})}_{f,\text{before 2002}} \right), \tag{7}$$

[18] That momentum strategy return is weak in Japan is a known result (Asness, Moskowitz, and Pedersen, 2013).

[19] Their methodology is also slightly different and explained in Section 2.2.

Table 2. Momentum Profitability Decline: The U.S. versus Other Markets

The table reports monthly momentum factor returns across markets before and after the Morningstar reform in June 2002. Panel A uses data from Ken French's website which starts in 1991. The other two panels use data from Jensen et al. (2021) which start from 1987 like our main exercises. Panels B and C report results based on equal-weighted and capped value-weighted returns, respectively. The standard errors are reported in the parenthesis, and coefficients statistically significant at the 10%, 5%, and 1% levels are denoted with ***, **, and * respectively.

Panel A: Ken French data

	U.S.	Market type		Regions		
		Developed ex U.S.	Emerging	Europe	APAC ex Japan	Japan
	(1)	(2)	(3)	(4)	(5)	(6)
Before reform	1.13***	0.82**	0.92***	1.11***	0.75	0.13
	(0.42)	(0.33)	(0.29)	(0.33)	(0.46)	(0.46)
After reform	0.03	0.59**	0.72***	0.75***	0.83***	0.11
	(0.32)	(0.23)	(0.19)	(0.28)	(0.24)	(0.25)
After − before	−1.10**	−0.23	−0.20	−0.36	0.08	−0.02
	(0.53)	(0.40)	(0.34)	(0.43)	(0.52)	(0.52)

Panel B: Jensen et al. (2021) data, equal-weighted

	U.S.	Market type		Regions		
		Developed ex U.S.	Emerging	Europe	APAC ex Japan	Japan
	(1)	(2)	(3)	(4)	(5)	(6)
Before reform	1.27***	0.74***	0.28	1.01***	0.22	−0.07
	(0.36)	(0.09)	(0.37)	(0.10)	(0.18)	(0.34)
After reform	0.32	0.94***	0.67***	1.06***	0.73***	0.17
	(0.29)	(0.07)	(0.16)	(0.09)	(0.09)	(0.23)
After − before	−0.94**	0.20*	0.39	0.05	0.51**	0.24
	(0.46)	(0.11)	(0.40)	(0.13)	(0.20)	(0.41)

Panel C: Jensen et al. (2021) data, capped value-weighted

	U.S.	Market type		Regions		
		Developed ex U.S.	Emerging	Europe	APAC ex Japan	Japan
	(1)	(2)	(3)	(4)	(5)	(6)
Before reform	0.92**	0.61***	0.78**	0.39**	0.79***	0.25
	(0.38)	(0.09)	(0.39)	(0.19)	(0.11)	(0.39)
After reform	0.05	0.58***	0.63***	0.60***	0.63***	0.02
	(0.29)	(0.08)	(0.20)	(0.10)	(0.11)	(0.25)
After − before	−0.87*	−0.03	−0.16	0.21	−0.16	−0.23
	(0.48)	(0.12)	(0.44)	(0.21)	(0.15)	(0.47)

$^{***}p < 1\%$, $^{**}p < 5\%$, $^{*}p < 10\%$

The key benefit of the direct approach is that rating exposure is *directly* measured. The drawback, which is shared with the spanning test method, is that it relies on strong functional form assumptions. Due to measurement errors introduced in matching fund ratings to stock holdings, we also expect this method to be prone to underestimation.

Spanning Tests. We first form a "rating factor": in each month, we sort stocks by $\text{ExpSum}(\Delta\text{Rating})_{i,t-1}$ into quintiles using NYSE break points, and then form the long-short value-weighted quintile factor portfolio.[20] We then use this factor to explain the other 49 factors in spanning regressions. Panel (a) of Figure 6 shows that each factor's loading on the rating factor, plotted on the vertical axis, is highly correlated to their bottom-up rating exposure measure ($\text{ExpSum}(\Delta\text{Rating})_{f,t-1}$), which is plotted on the horizontal axis. The factors with highest loadings are, as we expect, the momentum type ones. For instance, the standard $(t-1, t-12)$ momentum factor has a loading of 0.63 with a t-statistic of 9.94.

How much of factor returns can be explained? We run spanning regressions for each factor separately before and after the reform. To visualize the results, we sort factors into seven bins (so there are exactly 7 factors in each) by their average rating exposure $\text{ExpSum}(\Delta\text{Rating})_{f,t-1}$, and report the returns explained in Panels (b) and (c) in Figure 6. The top bin includes all five momentum-type factors plus the distress and size factors, all of which are labeled in Panel (a). The spanning test results are consistent with the prediction that explanatory power is mostly concentrated in momentum-type factors. Before the reform, the rating factor can explain 0.30% of 0.72% of the monthly return for the top bin of factors. After the reform, as ratings become less dispersed, the rating factor's return declines and so do the momentum-type factors.

When focusing on the momentum factor and other momentum-type factors, we find that the spanning method estimates that ratings can explain 0.41% and 0.28% of their post-reform return decline, respectively (second row in Table 3) This amounts to half of the overall profit decline for momentum and around one third of the other momentum-type factors. As

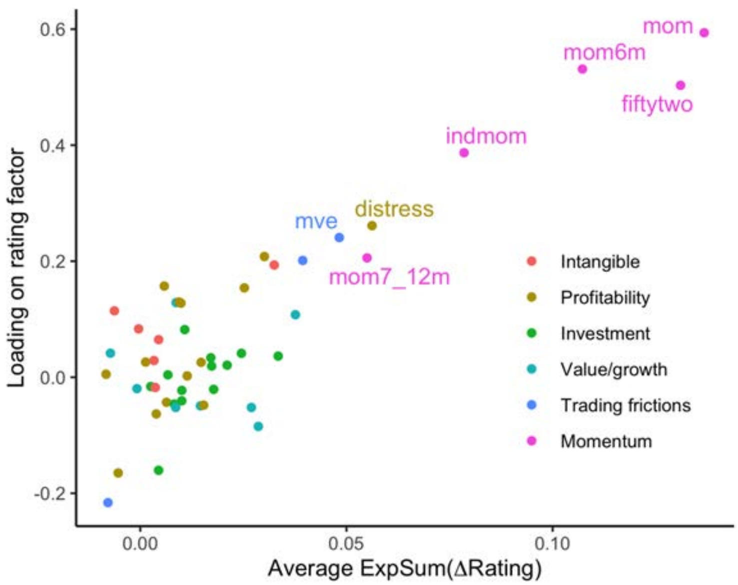
[20] Over our sample, this factor has an average monthly return of 0.41%.

discussed earlier, we consider the spanning tests as providing an *upper bound* estimate.

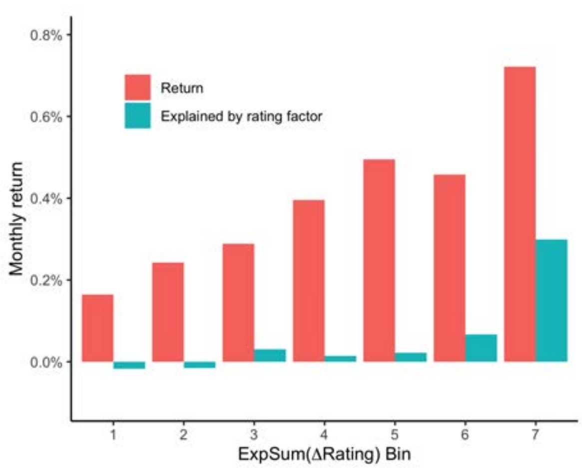
Figure 6. Spanning Test using the Rating Factor

We form a rating factor using long-short NYSE quintiles based on stock-level ExpSum(ΔRating), and examine whether it helps explain the other factors in this paper. The 49 factors are sorted into 7 bins (so that there are exactly 7 factors per bin) based on their average ExpSum(ΔRating). Panel (a) plots the factor loadings on the rating factor against their average ExpSum(ΔRating). Factors are colored by their classification categories in Hou et al. (2020). Factors in the top bin are labeled are they are, from left to right, $(t-7, t-12)$ intermediate momentum, distress risk, size, industry momentum, $(t-1, t-6)$ momentum, 52-week high, and standard $(t-1, t-12)$ momentum, respectively. Panels (b) and (c) plot the average factor returns (red bars) against returns explained by the rating factor (blue bars) in spanning tests before and after the June 2002 Morningstar methodology reform, respectively.

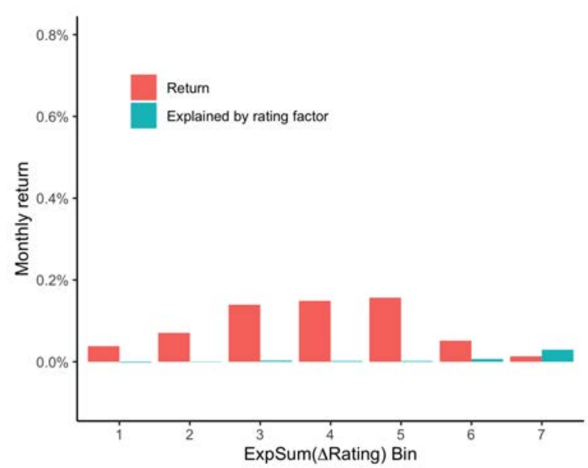
(a) Factor ExpSum(ΔRating) and rating factor loading



(b) Factor return explained, before reform



(c) Factor return explained, after reform



Direct estimation. We now use the second approach to estimate explanatory power. We use two approaches to estimate the price impact λ. In order to obtain a well-identified estimate, we run a factor-level return predicting regression using the 12-month window around the methodology reform event:

$$\mathrm{Ret}_{f,t} = \lambda \cdot \mathrm{ExpSum}(\Delta\mathrm{Rating})_{f,t-1} + X_{f,t-1} + \epsilon_{f,t}, \tag{8}$$

where the control $X_{f,t-1}$ includes factor returns over $t-1$, $t-2$ to $t-6$, and $t-7$ to $t-12$ months as well as factor- and time-fixed effects.[21] As discussed further in Section 6, using a short window means we primarily use of reform-induced rating variation, which reduces endogeneity concerns. To account for the cross-sectional factor return correlation, we adjust the standard errors using a feasible generalized least squares (FGLS) approach.[22] The estimation results are shown in Appendix Table B.2. For each star rating change in $\mathrm{Expsum}(\Delta\mathrm{Rating})_{f,t-1}$, we find that factor-level price impact is 2.27%, with a t-statistic of 4.28. The result is both statistically and economically significant. Appendix B.2 provides more details and robustness checks of this estimation.

We now apply the event-estimated λ coefficients to quantify the factor profitability decline

[21] These controls are motivated by the finding that factors exhibit momentum (Gupta and Kelly, 2019; Arnott et al., 2019; Ehsani and Linnainmaa, 2021).

[22] We use the full sample of factor returns to estimate the covariance matrix C of factor returns and incorporate C into the estimation. Specifically, let y be the vector of factor returns stacked together so that the first 49 entries are the first month, the next 49 entries are the second month, and so forth. Then, we estimate the covariance matrix of y to be

$$\hat{\Omega} = \begin{pmatrix} \hat{C} & 0 & \dots & 0 \\ 0 & \hat{C} & \dots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \dots & \hat{C} \end{pmatrix}$$

where $\hat{C}$ is the estimated contemporaneous return covariance matrix of the 49 factors. Let X denote the matrix of independent variables. Then, we estimate the regression coefficients and covariance using

$$\hat{b} = (X'\hat{\Omega}^{-1}X)^{-1}X'\hat{\Omega}^{-1}y,$$
$$\widehat{Var}(\hat{b}) = (X'\hat{\Omega}^{-1}X)^{-1}.$$

that can be explained by Morningstar. The results are shown in Columns (1) and (2) of Table 3. The first row shows that, after the reform, the monthly return of the momentum factor and other momentum-type factors dropped by 0.76% and 0.79%, respectively. Direct estimation suggests that rating-induced price pressures can explain 0.25% and 0.22% of the decline, which amounts to approximately one third of the profit declines. For comparison, we also report the estimates based on the spanning method which suggests that Morningstar ratings can explain around half of the profitability decline in the momentum factor and around one third of the other momentum factors. As discussed earlier, the spanning method tends to over-estimate, and the direct estimation methods tends to under-estimate. To be conservative, we take the latter method to inform our conclusion.

Table 3. Explanatory Power on Post-Reform Strategy Profitability Declines

We examine how much the Morningstar reform can explain the return decline of momentum-related factor strategies after June 2002. The first row reports the change in average monthly returns (in percent) from the pre-reform period (January 1987 to June 2002) to the post-reform period (July 2002 to December 2018). The next two rows present the estimated amount explained by Morningstar ratings. "Spanning method" refers to a regression-based approach in which we use a Morningstar rating-based factor to explain factor returns. "Direct estimate" is calculated based on multiplying the price impact parameter of ratings, estimated using the 2002 event, with the change in average $\text{ExpSum}(\Delta\text{Rating})_{f,t-1}$ after the reform. The last two rows present the *fraction* of return change explained by the Morningstar reform. Column (1) examines the momentum factor; Column (2) examines other factors in the momentum category; Column (3) and (4) examines the time-series and cross-sectional factor momentum strategies.

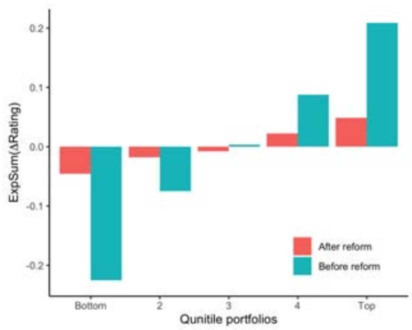
		Factors		Factor Momentum	
	Methodology	Momentum Factor	Other Mom-Type Factors	Time-series	Cross-section
		(1)	(2)	(3)	(4)
Return change (%)		-0.76	-0.79	-0.46	-0.37
Explained (%)	Spanning method	-0.41	-0.28	-0.30	-0.29
	Direct estimate	-0.25	-0.22	-0.31	-0.31
Fraction explained	Spanning method	0.545	0.355	0.647	0.791
	Direct estimate	0.330	0.274	0.677	0.829

The analysis so far focuses on long-short returns. To visualize the impact of our mechanism a more granular way, the middle column of Panels of Figure 7 plots the rating exposure,

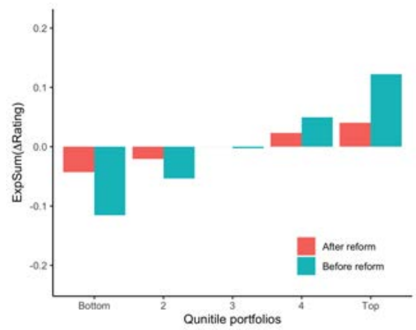
Figure 7. Effect of Morningstar Reform on Strategies, By Quintiles

This three rows of this figure plots the rating exposure (ExpSum(ΔRating)$_{t-1}$), flow-induced trading (FIT), and return of different strategies by quintile, respectively. The data is separated into the periods before reform (January 1987 to June 2002) and after reform (July 2002 to December 2018). The left column plot results for the cross-sectional factor momentum strategy, where each quintile represents a portfolio of factors. The middle column plot results for the stock momentum factor. The right column plot results for the factors that are not in the momentum category. All variables are demeaned to emphasize cross-sectional differences. The FIT results start in January 1991 due to availability of monthly fund flow data.

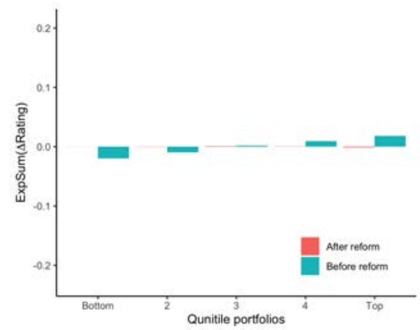
(a) ExpSum(ΔRating), Factor momentum



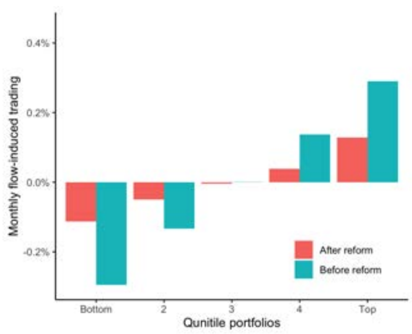
(b) ExpSum(ΔRating), Momentum



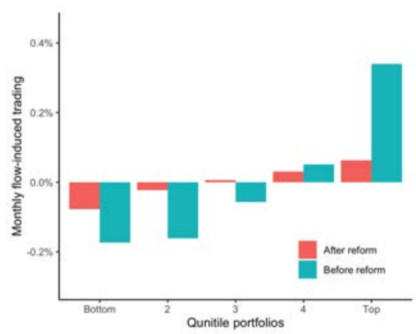
(c) ExpSum(ΔRating), non-momentum factors



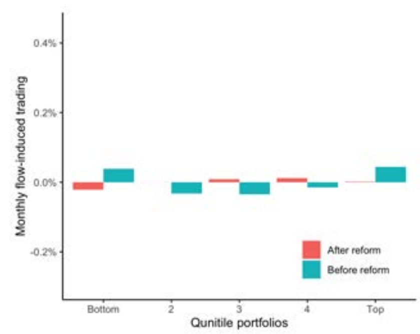
(d) FIT, Factor momentum



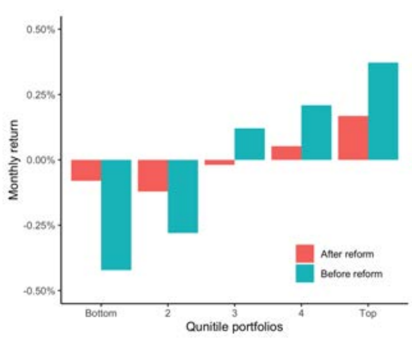
(e) FIT, Momentum



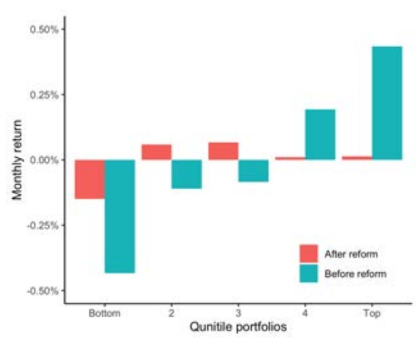
(f) FIT, non-momentum factors



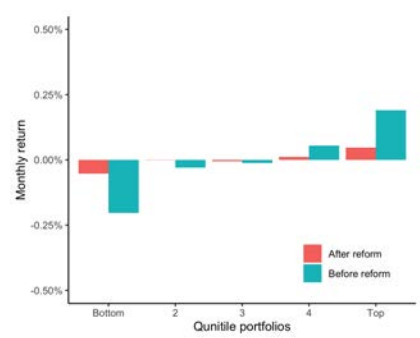
(g) Return, Factor momentum



(h) Return, Momentum



(i) Return, non-momentum factors



flow-induced trading, and returns of different quintile portfolios.[23] Panel (b) shows that, be-

[23] Because results for other momentum-type factors are very similar to that for momentum, we do not show them for brevity.

fore the rating reform, the long (short) legs of momentum experience significant upward (downward) rating changes. The right column of Panels shows that the same effect is barely present in the non-momentum-type factors. The second row shows broadly similar patterns for fund flow-induced trading. The third row shows that that post-2002 drop in long/short returns is more pronounced in momentum than other factors. Overall, these findings are consistent with our prediction that the rating-induced effect impacts momentum but not other asset pricing factors.

5 Effect of the Reform on Factor Momentum

Ehsani and Linnainmaa (2021) show that factors themselves exhibit momentum. Specifically, they propose two related strategies which can be implemented on any universe of factors, and they call the two strategies time-series factor momentum (TSFM) and cross-sectional factor momentum (CSFM). Both are long-short strategies with equal weights across factors in each leg. In TSFM, the long (short) leg consists of all factors with positive (negative) returns over the previous twelve months. In CSFM, the long (short) leg consists of factors with above (below) median past twelve month return. Therefore, while the CSFM portfolio has the same number of factors in each leg, the TSFM doesn't have to. We follow the factor momentum construction in Ehsani and Linnainmaa (2021) using our factor universe.

In this section, we show that factor momentum is highly exposed to style momentum, and as a consequence, is also affected by the Morningstar reform. For both TSFM and CSFM, their profitability drops after the reform and we estimate that Morningstar rating can explain approximately two thirds of the drop. This higher explanatory power, relative that of the stock momentum profitability decline, is expected: factor momentum does not load onto idiosyncratic momentum which the Morningstar mechanism is unrelated to.

It is important to note that our findings should be seen as supplements to, rather than contradictions to, Ehsani and Linnainmaa (2021). They show that factor momentum strategies have high returns, and argue that the persistence of profitability is related to difficulties of arbitrage in more systematic return components. They do not take a strong stance on what caused factor momentum in the first place. Our exercise focuses on providing a specific (and partial) economic explanation for why factor momentum arises in the first place.

5.1 Mechanism: Factor Momentum Loads on Style Momentum

We first show that factors have large and persistent style exposures. As a consequence, factor momentum is also affected by the most-reform halt of style-level rating-induced positive feedback trading.

We use a simple holdings-based approach to measure the style exposure of factors. For each stock i, let $w^{\pi}_{i,t}$ denote the fraction of its mutual fund holding in each of the 3×3 Morningstar size-value styles π in quarter t. Then, its size style exposure is given by $\sum_{\pi \in \text{three small cap styles}} w^{\pi}_{i,t} - \sum_{\pi \in \text{three large cap styles}} w^{\pi}_{i,t}$. Similarly, we define its value style exposure as $\sum_{\pi \in \text{three value styles}} w^{\pi}_{i,t} - \sum_{\pi \in \text{three growth styles}} w^{\pi}_{i,t}$. This holding-based exposure measure is easy to interpret: a stock that is only held by small (or large) cap style funds will have a size score of $+1$ (or -1). We then aggregate these stock-level style exposures up to the factor level. Because factors are long-short portfolios, we anticipate that their exposures will be bounded between -2 and $+2$.

To visualize the style exposure of factors, Panels (a) and (b) plot the annual average size and value exposure of factors. To illustrate time fluctuation in the exposures, in each Panel, we rank factors by their corresponding style exposures and plot the four factors at the 0, 1/3, 2/3, and 1 quantiles. Panel (a) shows that the size exposure of factors range from approximately 1.5 for the size factor to -1 for Ohlson's O-Score. This is a very large range, considering that the factor style exposure is designed to be bounded between -2 and 2. Panel (b) shows that value exposure ranges from approximately 0.5 for the cash flow to price factor to approximately -0.5 for the sales growth factor.

Not only do many factors have sizeable style exposures, more importantly, their factor exposures are very stable over time. This is shown in Panels (c) and (d): we plot factors' annual style exposures against their exposures in the previous year; each data point is a factor-year. The resulting points almost exactly fall on the 45% degree lines, implying that the style exposures of factors are very persistent. Therefore, we conclude that factor momentum loads onto style momentum and should be negatively impacted by the Morningstar reform.

5.2 The Decline of Factor Momentum Profits

To visualize how the factor momentum strategies are impacted by the Morningstar reform, Figure 9 plots the factor-level rating exposure ($\text{ExpSum}(\Delta\text{Rating})_{t-1}$) aggregate at the factor momentum strategy level. To further shed light on the source of the effect, we also decompose the ratings into the 3×3 style-level and the idiosyncratic components.[24] The graphs show that, before the reform, factor momentum had relatively high rating exposure. Style-level rating exposure dropped to effectively zero after the reform, and as a consequence, the overall rating exposure declined.

As predicted, factor momentum profits declined after the reform. This is reported in the first row of Columns (3) and (4) of Table 3. After the reform, TSFM and CSFM strategy monthly returns declined by 0.46% and 0.37%, respectively, from levels of 0.60% and 0.52% before the reform. That is, the profits of both strategies declined by three-fourths after the reform.

[24]Specifically, we first decompose *fund* ratings into a style-level component and an idiosyncratic residual:

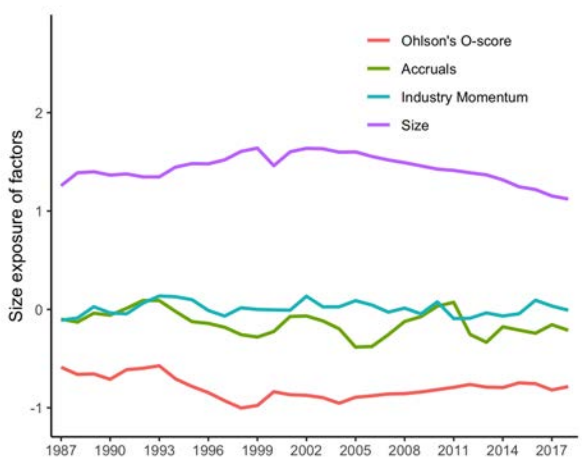
$$\text{Rating}^{\text{fund}}_{j,t} = \text{StyleRating}^{\text{fund}}_{j,t} + \text{IdiosyncraticRating}^{\text{fund}}_{j,t} \tag{9}$$

where $\text{StyleRating}^{\text{fund}}_{j,t}$ is the value-weighted average rating for the 3×3 style that fund j belongs to, and $\text{IdiosyncraticRating}^{\text{fund}}_{j,t}$ is defined as a residual. We then use this decomposition to derive separate style-level and idiosyncratic rating components at the stock-level and also the factor-level.

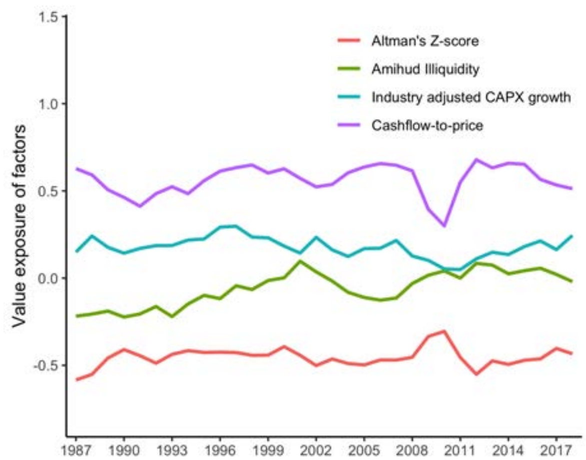
Figure 8. Factors Have Persistent Style Exposures

We measure size and value style exposures of factors using a portfolio-based approach. For instance, holding by funds in the small (large) cap Morningstar styles gets a size exposure score of +1 (-1), and we aggregate these exposure scores at the long-short factor level to measure the size exposure of the 49 factors considered in this paper. Value exposure is measured in a similar way. Panels (a) and (b) plot the annual average size and value exposures of factors over the sample. The factors chosen at those at the 0, 1/3, 2/3, and 1 quantiles of average exposures. Panels (c) and (d) examine persistence of style exposures by plotting factors' style exposures in the current year against exposures in the previous year. The green dashed lines represent 45% diagonal lines.

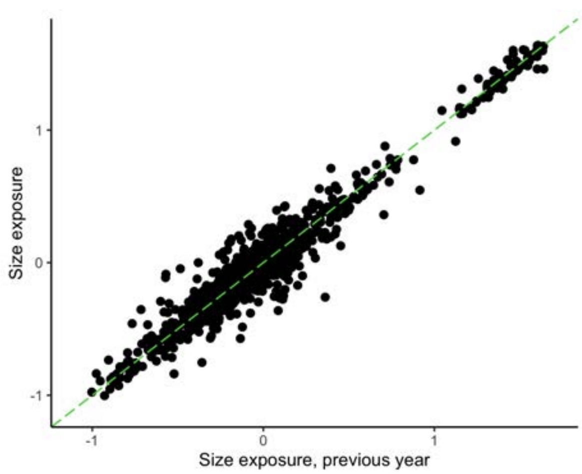
(a) Size exposure of factors



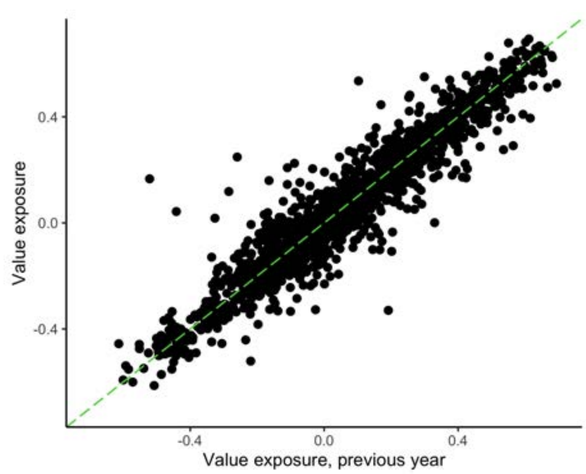
(b) Value exposure of factors



(c) Size exposure persistence



(d) Value exposure persistence



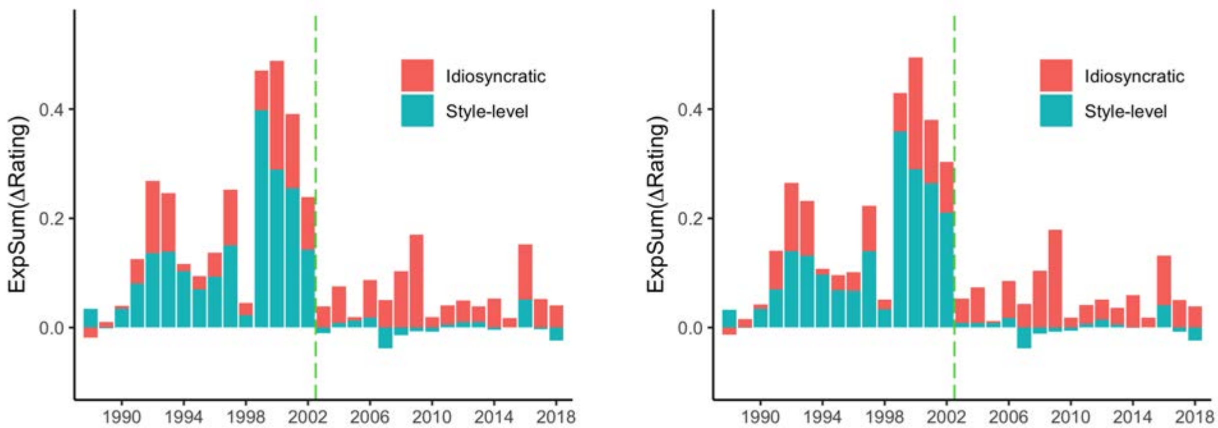
Placebo Test: Factor Momentum in U.S. and Other Countries. We now use the global factors in Jensen et al. (2021) to examine factor momentum strategies outside of U.S. The results are reported in Table 4. Because the results based on TSFM and CSFM are highly similar, we only report the former for brevity. Based on equal-weighted returns in Panel A,

Figure 9. Rating Exposure of Factor Momentum Strategies

Panels (a) and (b) plot the annual average ExpSum(ΔRating$_{t-1}$) for time-series and cross-sectional factor momentum strategies, respectively. Ratings are decomposed into style-level ratings (blue bars) and idiosyncratic ratings (red bars). To make the green vertical dashed line exactly delineating the reform date, each year y is defined as July of year $y-1$ to June of year y.

(a) Time-series factor momentum

(b) Cross-sectional factor momentum



the decline of factor momentum profits is specific to the U.S.; in fact, factor momentum profits increased in most of the regions except Europe. The capped value-weighted return-based results in Panel B are broadly similar, although with lower statistical significance (the U.S. post reform change is significant at the 10% level). Overall, these results are consistent with the prediction that the drop of factor momentum profits is concentrated in the U.S. market.

Quantifying Explanatory Power of Ratings. We then follow the same methodologies in Section 4.2 to quantify the explanatory power of Morningstar ratings. The results are reported in Columns (3) and (4) of Table 3. Both the spanning tests and the direct estimation method gives the same conclusion: Morningstar ratings can approximately account for two-thirds of the post-2002 decline of factor momentum profits. This higher explanatory power, relative to that of the stock momentum factor, is consistent with the mechanism: factor momentum does not load on idiosyncratic momentum and thus is more directly affected by the Morningstar reform.

Table 4. Factor Momentum Profitability Decline: U.S. versus Other Markets

The table reports monthly momentum factor returns across markets before and after the Morningstar reform in June 2002. We use the U.S. and global factors from Jensen et al. (2021). Panel A uses equal-weighted returns and Panel B uses the capped value-weighted returns. The standard errors are reported in the parenthesis.

Panel A: equal-weighted

	U.S.	Market type		Regions		
		Developed ex U.S.	Emerging	Europe	APAC ex Japan	Japan
	(1)	(2)	(3)	(4)	(5)	(6)
Before reform	1.12***	0.52***	0.41	0.58***	0.48***	0.18
	(0.30)	(0.11)	(0.28)	(0.13)	(0.12)	(0.18)
After reform	0.38**	0.54***	0.56***	0.46***	0.68***	0.40***
	(0.16)	(0.08)	(0.09)	(0.10)	(0.06)	(0.11)
After − before	−0.75**	0.02	0.16	−0.11	0.21	0.22
	(0.34)	(0.13)	(0.29)	(0.16)	(0.13)	(0.21)

Panel B: capped value-weighted

	U.S.	Market type		Regions		
		Developed ex U.S.	Emerging	Europe	APAC ex Japan	Japan
	(1)	(2)	(3)	(4)	(5)	(6)
Before reform	0.63*	0.34***	0.53	0.41***	0.35**	0.28
	(0.32)	(0.12)	(0.36)	(0.14)	(0.14)	(0.22)
After reform	0.06	0.23**	0.30**	0.21*	0.30***	0.04
	(0.13)	(0.10)	(0.13)	(0.12)	(0.08)	(0.11)
After − before	−0.58*	−0.11	−0.23	−0.20	−0.04	−0.24
	(0.35)	(0.15)	(0.39)	(0.19)	(0.16)	(0.25)

$^{***}p < 1\%$, $^{**}p < 5\%$, $^{*}p < 10\%$

To visualize the effect on different parts of the factor momentum strategy, the first column of Panels in of Figure 7 plots the rating exposure, flow-induced trading, and returns of the factor momentum strategy. To focus on the cross-sectional differences across different factors that compose the strategy, we examine the cross-sectional factor momentum strategy. The results are broadly consistent with our prediction that factor momentum is slightly more affected by the Morningstar-based mechanism than the stock momentum factor, which is shown in the middle column of Panels.

6 Which Factors Experienced "Kinks" in 2002?

So far, we have focused on explaining the long-term strategy profitability decline since mid 2002. As discussed in the introduction, there is an additional puzzle that many factors appear to have experienced a sharp kink in their returns in mid 2002. In this section, we use an event-study approach to examine whether this is explained by the Morningstar reform. In addition to explaining the kink, this exercise also sheds light on how rating-induced fund flows impact factor-level returns.

We zoom in on a one-year window (January to December 2002) around the rating reform. There are two benefits to using a short window. First, rating changes in this period are predominantly caused by the rating methodology change. Second, using a short window also reduces the chance that factor returns are impacted by other events such as the NYSE decimalization in early 2001 and the introduction of NYSE auto quoting in 2003 (Hendershott, Jones, and Menkveld, 2011).

6.1 Predicting How the Reform Would Impact Factors

We sort the 49 factors into quintiles based on how their rating is affected by the reform event. To alleviate endogeneity concerns, we sort the factors using the *predicted* reform-induced rating change computed using data in December 2001, which is *prior to* the event study window. Specifically, using data available up to December 2001, we estimate mutual fund ratings by following the pre-2002 and post-2002 Morningstar rating methodologies, and then aggregate these ratings up at the factor level. We then predict that each factor f will experience a rating change of

$$\text{PredictedChange}_f = \widehat{\text{Rating}}_{f,\text{Dec 2001}}^{\text{post-2002 methodology}} - \widehat{\text{Rating}}_{f,\text{Dec 2001}}^{\text{pre-2002 methodology}}, \tag{10}$$

where the two terms on the right hand side represent estimated factor-level rating under the two different rating methodologies, respectively.[25] Appendix B.3 explains the prediction process in more detail and verifies that the predictions can accurately forecasting actual factor rating changes at the reform event.

6.2 Event Study

Figure 10 plots what happened to the factors in 2002. Panel (a) plots average ratings of factors and shows a sharp methodology-induced change exactly at the event. Factors in quintile 1 suffer a drop of 0.43 stars, while those in quintile 5 experience a small increase of 0.19 stars. Panels (c) and (e) plot cumulatively monthly factor FIT and returns around the event, respectively. Quintile 1—the factors that benefited from ratings pre-event but suffered post-event—experienced a decline of 1% in monthly FIT and a sharp decline of -3.7% in monthly returns. At the same time, quintile 5 experienced an increase of 0.14% in monthly FIT and a slight increase of 0.75% in monthly returns.[26]

To alleviate the concern that the return and FIT changes could result from other reasons, we also conduct the same exercise in all years other than 2002. The results on rating, FIT, and return changes in other years are shown as the white bars in Panels (b), (d), and (f) with 95% confidence intervals. These panels show that the large change around June is unique to 2002.

6.2.1 Alternative Hypotheses to the Event Study Results

We now discuss the concern that the factor price fluctuations around June 2002 may be triggered by changes other than the Morningstar reform.

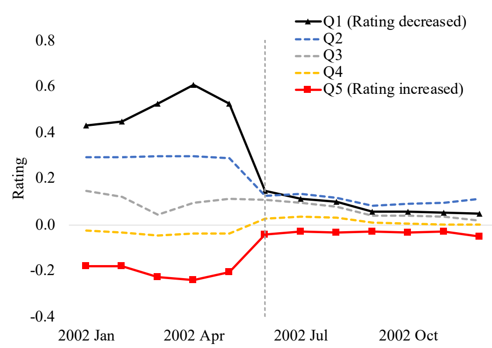
[25] We estimate the pre-2002 ratings under the old methodology, instead of using the actual pre-2002 ratings, to reduce estimation errors. Because we do not have exactly the same data set that Morningstar uses internally, our rating estimation contain errors. However, the same data-induced error is present in both terms in Equation (10), so we are able to difference it out.

[26] In a companion paper, we show that the implied style-level price impact coefficient (the reciprocal of demand elasticity) is approximately 5 (Ben-David et al., 2020a). That is, buying 1% of the market cap outstanding creates a price impact of approximately 5%. This magnitude is consistent with the existing literature that estimates the price impact of undiversifiable demand shocks (e.g., Gabaix and Koijen, 2020).

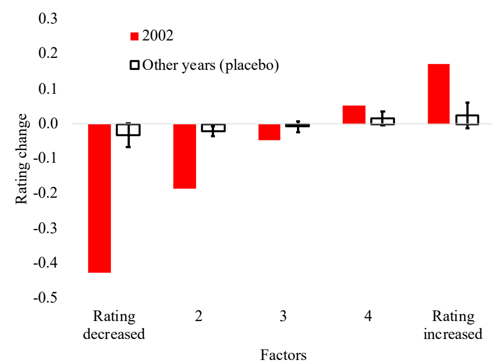
Figure 10. Stock Factors around the June 2002 Event

We perform event studies on the 49 factors using a 12-month window around the reform event (January to December 2002). In the left panels, we sort factors by their *predicted* reform-induced rating change into quintiles, and then plot the evolution of their ratings in Panel (a), cumulative fund flow–induced trading (FIT) in Panel (c), and cumulative returns in Panel (e). To alleviate endogeneity concerns, the rating change prediction only uses data up to December 2001 (prior to the event window). The dashed vertical line is the June 2002 reform event. The right panels conduct the same exercises in years other than 2002 as a placebo test. The red bars plot the average rating, flow-induced trading (FIT), and return changes after June (the average of July to December 2002 minus the average of January to June 2002), while the white bars plot the corresponding results for years other than 2002. The whiskers represent 95% confidence intervals. To focus on cross-sectional dispersion, all variables—ratings, returns, and flows—are demeaned by period.

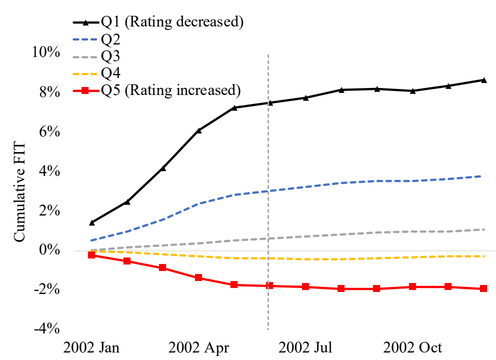
(a) Rating in 2002



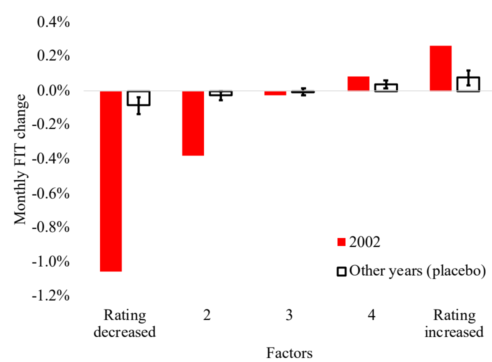
(b) Rating change, 2002 vs placebo



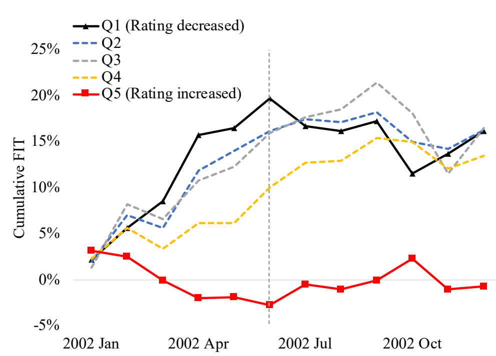
(c) Cumulative FIT in 2002



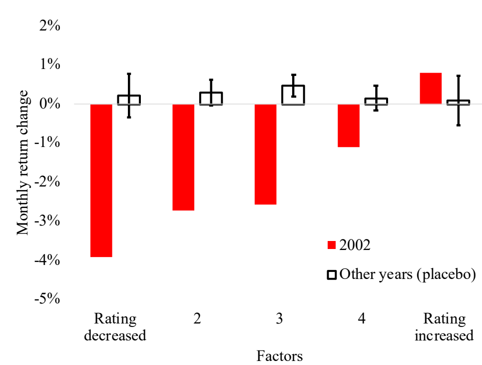
(d) FIT change, 2002 vs placebo



(e) Cumulative return in 2002



(f) Return change, 2002 vs placebo



Arbitrage activity. One natural worry is whether arbitrage forces in these factors have suddenly become stronger in mid 2002. A number of papers present evidence that factor profitability is related to arbitrage activity. For instance, Hanson and Sunderam (2013) argue that value and momentum strategy profits decrease when more capital is devoted to them. McLean and Pontiff (2016) show that factor profitability declines after the strategies were published in academic papers and link it to arbitrage actions. Relatedly, Lou and Polk (2018) show that a return-based measure of arbitrageur activity negatively predict momentum profits.

Did arbitrage activity change in June 2002? We use two measures proposed in the literature to proxy for arbitrage activity in factors. First, we follow Chen, Da, and Huang (2019) to construct a net arbitrage activity (NAT) measure. For each stock, the authors measure the long position of arbitrageurs using aggregate 13F holdings of hedge funds and the short position using aggregate short interest from Compustat.[27] The authors combine the long and short positions into a net position, and subtract the past four-quarter average to arrive at a measure of arbitrageur position changes, which they call NAT. We follow them to compute stock-level NAT and aggregate it at the factor level.

Second, we follow Lou and Polk (2018) to construct a correlation-based measure of arbitrage activity. These authors measure arbitrage activity in the momentum strategy by estimating excess return correlation within the long and short portfolios, which can be gen-

[27] We use the list of 13F institutions identified as hedge funds in Aragon, Li, and Lindsey (2018). We thank the authors for kindly sharing the data. It is worth noting that, while the short side of NAT is updated monthly, the long side relies on 13F holdings and is only updated quarterly.

erated by arbitrageurs trading in the factor.[28] We also compute this measure for all factors.[29]

We plot the evolution of these measures in the 12 month event window in Figure 11. As in Section 6, we sort factors into quintiles by their predicted rating change using data up to December 2001. Panel (a) plots the NAT measure, and Panel (b) plots the correlation-based measure. There is no noticeable change in either measure during the event window.

Changes in liquidity. One may also hypothesize that stock market liquidity increased dramatically in June 2002.[30] To examine this possibility, we aggregate the stock-level Corwin and Schultz (2012) bid-ask spread measure for the factors (averaging over the long and short legs) during this period. The results, plotted in Panel (c), show no evidence that liquidity changes account for our findings. Panel (d) shows that monthly trading turnover also had no clear change around the event.

In summary, we do not find any noticeable change in arbitrage trading activity or liquidity—two major forces that could impact factor returns—around June 2002. Thus, the event study supports the idea that Morningstar rating changes can exert tangible price impact on factor returns.

[28] Specifically, in any given month, they use the previous 52 weeks of data to compute a "comomentum" measure:

$$\begin{aligned}\text{CoMomentum}_t = \frac{1}{2}\cdot\Bigg[&\frac{1}{N^L(N^L-1)}\sum_i\sum_{j\neq i}\text{PartialCorr}(\text{Ret}_i,\text{Ret}_j)\\ &+\frac{1}{N^S(N^S-1)}\sum_i\sum_{j\neq i}\text{PartialCorr}(\text{Ret}_i,\text{Ret}_j)\Bigg],\end{aligned}$$

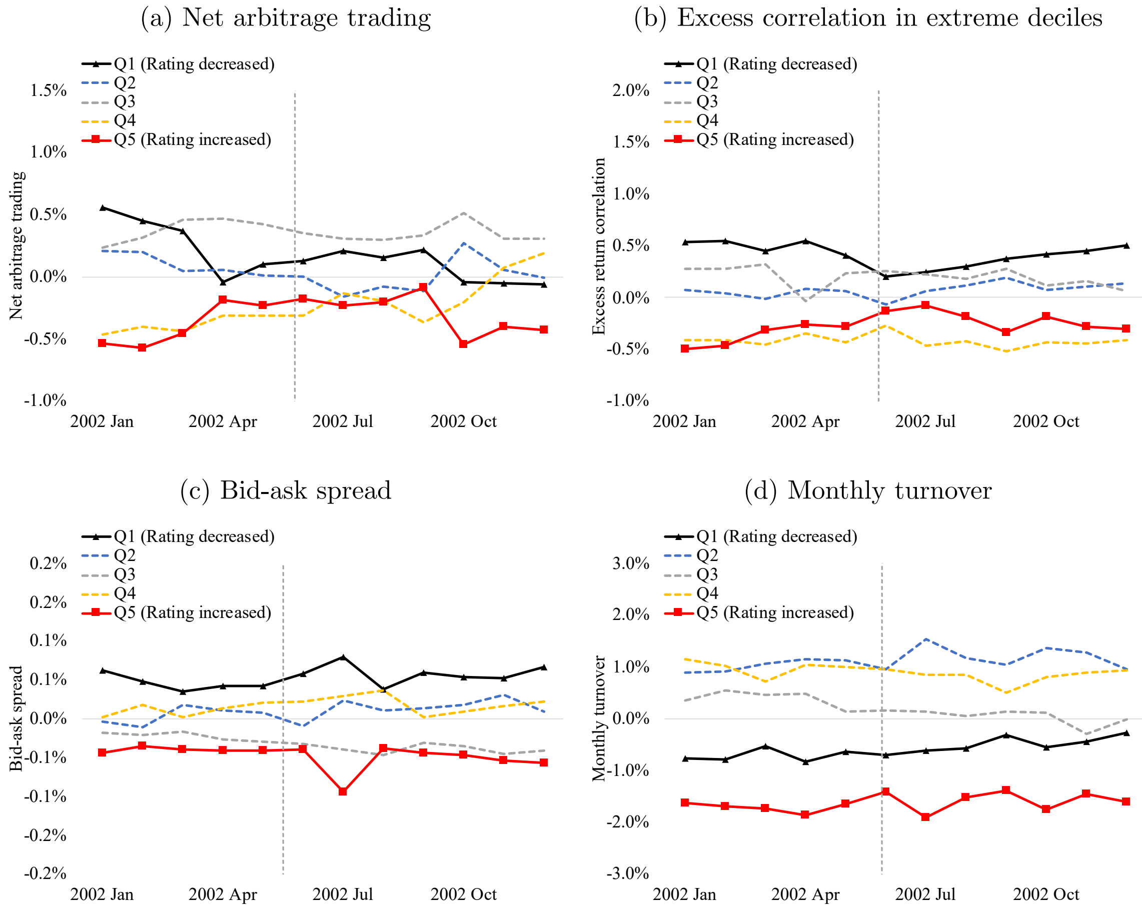
where N^L and N^S are the number of stocks in the long and short leg portfolios, respectively. To compute the partial return correlations, they first subtract Fama-French 30 industry returns from weekly stock returns and then regress the residuals on the Fama-French three factors to obtain alphas. Finally, they compute equal-weighted averages of the pairwise correlations of the alphas within the portfolios and take an average.

[29] As a sanity check on our replication of their methodology, consistent with Lou and Polk (2018), we find that this measure indeed negatively predicts returns of factors in the momentum category.

[30] Increasing liquidity may explain factor profitability declines through two possibility mechanisms. First, if a factor's profitability comes from demand price pressures, then increasing liquidity will reduce the price impact of such demand shocks. Second, if factor profitability is the result of arbitrageurs not being able to arbitrage away profits, then increasing liquidity may facilitate arbitrage effectiveness and thus reduce residual factor profitability. Of course, the asset pricing literature has also found evidence that illiquidity is a priced risk, so the changes may also come from changes in equilibrium-required rates of return (Amihud, 2002; Pástor and Stambaugh, 2003; Acharya and Pedersen, 2005).

Figure 11. Alternative Explanations: other Influences Around 2002

As in Figure 10, factors are sorted into quintiles by the predicted rating change using data in December 2001. Thus, quintile 1 (or 5) factors are those predicted to experience the largest rating decrease (increase) at the reform event. Panel (a) plots the net arbitrage trading measure in Chen et al. (2019). Panel (b) plots excess return correlation in extreme factor quintiles, a measure of arbitrage activity developed in Lou and Polk (2018). Panel (c) plots average bid-ask spread, measured following Corwin and Schultz (2012), of the long and short factor legs. Panel (d) plots the average monthly trading turnover of the long and short factor legs. To focus on cross-sectional dispersion, all variables are demeaned by month. In all panels, the vertical dashed line marks the methodology change event.



7 Conclusion

Since mid 2002, returns to momentum-type strategies — both the momentum factor and factor momentum — has dwindled substantially. In addition to the long-term profitability declines, many asset pricing factors have also experienced a "kink" in their returns right at mid-2002.

This paper finds that a significant part of the post-2002 factor profitability decline stems from a seemingly innocuous change in Morningstar's rating methodology. Before June 2002, Morningstar rated funds using their past performance ranking relative to U.S. equity funds. As a consequence, funds pursuing investment strategies associated with recently outperforming styles were rated higher than funds in recently underperforming styles. Investors chasing fund ratings led to significant style-level positive-feedback trading. After the reform, Morningstar rated funds using their past performance ranking against their 3×3 size-value style peers, causing an immediate halt to this positive-feedback trading. Because momentum-related factors and factor momentum benefits from the earlier positive-feedback trading, this halt caused a disruption to their profitability. We estimate that the Morningstar rating reform accounts for approximately a third and two thirds, respectively, of the post-2002 profitability decline of the momentum factor and factor momentum.

More broadly, our findings are in line with a number of recent studies indicating that demand effects can impact systematic price movements (Gabaix and Koijen, 2020; Li, 2020). For better identification, our paper focuses closely on the role of Morningstar ratings. However, it is possible that role of correlated demand, arising from other institutional features or frictions, may be even more consequential for asset pricing than is documented here. Therefore, unlike that assumed in classical "frictionless" asset pricing, demand effects may be a first-order driver of asset prices (Koijen and Yogo, 2019).

References

Acharya, Viral V, and Lasse Heje Pedersen, 2005, Asset pricing with liquidity risk, *Journal of Financial Economics* 77, 375–410.

Amihud, Yakov, 2002, Illiquidity and stock returns: cross-section and time-series effects, *Journal of Financial Markets* 5, 31–56.

Aragon, George O, Emma Li, and Laura Anne Lindsey, 2018, Exploration or exploitation? Hedge funds in venture capital, *Hedge Funds in Venture Capital (September 18, 2018)* .

Arnott, Robert D., Mark Clements, Vitali Kalesnik, and Juhani T. Linnainmaa, 2019, Factor momentum, Working paper, Dartmouth College.

Asness, Clifford S., Tobias J. Moskowitz, and Lasse Heje Pedersen, 2013, Value and momentum everywhere, *Journal of Finance* 68, 929–985.

Barber, Brad M., Xing Huang, and Terrance Odean, 2016, Which factors matter to investors? Evidence from mutual fund flows, *Review of Financial Studies* 29, 2600–2642.

Barberis, Nicholas, and Andrei Shleifer, 2003, Style investing, *Journal of Financial Economics* 68, 161–199.

Barberis, Nicholas, Andrei Shleifer, and Robert Vishny, 1998, A model of investor sentiment, *Journal of Financial Economics* 49, 307–343.

Ben-David, Itzhak, Francesco Franzoni, and Rabih Moussawi, 2018, Do ETFs increase volatility?, *Journal of Finance* 73, 2471–2535.

Ben-David, Itzhak, Francesco Franzoni, Rabih Moussawi, and John Sedunov, 2021, The granular nature of large institutional investors, *Management Science* forthcoming.

Ben-David, Itzhak, Jiacui Li, Andrea Rossi, and Yang Song, 2019, What do investors really care about?, Working paper, The Ohio State University.

Ben-David, Itzhak, Jiacui Li, Andrea Rossi, and Yang Song, 2020a, Advice-driven demand and systematic price fluctuations, Working paper, The Ohio State University.

Ben-David, Itzhak, Jiacui Li, Andrea Rossi, and Yang Song, 2020b, Why has mutual fund managers' skill disappeared?, Working paper, The Ohio State University.

Blitz, David, Matthias X Hanauer, and Milan Vidojevic, 2020, The idiosyncratic momentum anomaly, *International Review of Economics & Finance* 69, 932–957.

Blitz, David, Joop Huij, and Martin Martens, 2011, Residual momentum, *Journal of Empirical Finance* 18, 506–521.

Brown, David C., Shaun Davies, and Matthew Ringgenberg, 2021, ETF arbitrage and return predictability, *Review of Finance* forthcoming.

Calluzzo, Paul, Fabio Moneta, and Selim Topaloglu, 2019, When anomalies are publicized broadly, do institutions trade accordingly?, *Management Science* 65, 4555–4574.

Chang, Yen-Cheng, Harrison Hong, and Inessa Liskovich, 2015, Regression discontinuity and the price effects of stock market indexing, *Review of Financial Studies* 28, 212–246.

Chen, Andrew Y, and Tom Zimmermann, 2020, Open source cross-sectional asset pricing, *Available at SSRN* .

Chen, Yong, Zhi Da, and Dayong Huang, 2019, Arbitrage trading: The long and the short of it, *Review of Financial Studies* 32, 1608–1646.

Cho, Thummim, 2020, Turning alphas into betas: Arbitrage and endogenous risk, *Journal of Financial Economics* 137, 550–570.

Chordia, Tarun, Avanidhar Subrahmanyam, and Qing Tong, 2014, Have capital market anomalies attenuated in the recent era of high liquidity and trading activity?, *Journal of Accounting and Economics* 58, 41–58.

Cochrane, John H., 2011, Presidential address: Discount rates, *Journal of Finance* 66, 1047–1108.

Corwin, Shane A, and Paul Schultz, 2012, A simple way to estimate bid-ask spreads from daily high and low prices, *Journal of Finance* 67, 719–760.

Coval, Joshua, and Eric Stafford, 2007, Asset fire sales (and purchases) in equity markets, *Journal of Financial Economics* 86, 479–512.

Daniel, Kent D., David Hirshleifer, and Avanidhar Subrahmanyam, 1998, Investor psychology and security market under-and overreactions, *Journal of Finance* 53, 1839–1885.

Daniel, Kent D., and Tobias J. Moskowitz, 2016, Momentum crashes, *Journal of Financial Economics* 122, 221–247.

Ehsani, Sina, and Juhani T. Linnainmaa, 2021, Factor momentum and the momentum factor, *Journal of Finance* forthcoming.

Evans, Richard B., and Yang Sun, 2021, Models or stars: The role of asset pricing models and heuristics in investor risk adjustment, *Review of Financial Studies* 34, 67–107.

Falck, Antoine, Adam Rej, and David Thesmar, 2021, Why and how systematic strategies decay, Working paper, Massachusetts Institute of Technology.

Gabaix, Xavier, and Ralph S.J. Koijen, 2020, In search of the origins of financial fluctuations: The inelastic markets hypothesis, Working paper, Harvard University.

George, Thomas J., and Chuan-Yang Hwang, 2004, The 52-week high and momentum investing, *Journal of Finance* 59, 2145–2176.

Green, Jeremiah, John R.M. Hand, and Mark T. Soliman, 2011, Going, going, gone? The apparent demise of the accruals anomaly, *Management Science* 57, 797–816.

Green, Jeremiah, John R.M. Hand, and X. Frank Zhang, 2017, The characteristics that provide independent information about average US monthly stock returns, *Review of Financial Studies* 30, 4389–4436.

Grinblatt, Mark, and Bing Han, 2005, Prospect theory, mental accounting, and momentum, *Journal of Financial Economics* 78, 311–339.

Gupta, Tarun, and Bryan Kelly, 2019, Factor momentum everywhere, *Journal of Portfolio Management* 45, 13–36.

Hanson, Samuel G., and Adi Sunderam, 2013, The growth and limits of arbitrage: Evidence from short interest, *Review of Financial Studies* 27, 1238–1286.

Harris, Lawrence, and Eitan Gurel, 1986, Price and volume effects associated with changes in the S&P 500 list: New evidence for the existence of price pressures, *Journal of Finance* 41, 815–829.

Harvey, Campbell R., 2017, Presidential address: The scientific outlook in financial economics, *Journal of Finance* 72, 1399–1440.

Harvey, Campbell R., Yan Liu, and Heqing Zhu, 2016, ... and the cross-section of expected returns, *Review of Financial Studies* 29, 5–68.

Hendershott, Terrence, Charles M. Jones, and Albert J. Menkveld, 2011, Does algorithmic trading improve liquidity?, *Journal of Finance* 66, 1–33.

Hillert, Alexander, Heiko Jacobs, and Sebastian Müller, 2014, Media makes momentum, *The Review of Financial Studies* 27, 3467–3501.

Hong, Harrison, and Jeremy C. Stein, 1999, A unified theory of underreaction, momentum trading, and overreaction in asset markets, *Journal of Finance* 54, 2143–2184.

Hou, Kewei, Wei Xiong, and Lin Peng, 2009, A tale of two anomalies: The implications of investor attention for price and earnings momentum, *Available at SSRN 976394* .

Hou, Kewei, Chen Xue, and Lu Zhang, 2020, Replicating anomalies, *Review of Financial Studies* 33, 2019–2133.

Huang, Shiyang, Yang Song, and Hong Xiang, 2020a, Noise trading and asset pricing factors, Working paper, University of Washington.

Huang, Shiyang, Yang Song, and Hong Xiang, 2020b, The smart beta mirage, Working paper, University of Washington.

Jegadeesh, Narasimhan, and Sheridan Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, *Journal of Finance* 48, 65–91.

Jegadeesh, Narasimhan, and Sheridan Titman, 2011, Momentum, *Annu. Rev. Financ. Econ.* 3, 493–509.

Jensen, Theis Ingerslev, Bryan T Kelly, and Lasse Heje Pedersen, 2021, Is there a replication crisis in finance?, *Working paper* .

Khandani, Amir E., and Andrew W. Lo, 2011, Illiquidity premia in asset returns: An empirical analysis of hedge funds, mutual funds, and US equity portfolios, *Quarterly Journal of Finance* 1, 205–264.

Koijen, Ralph S.J., and Motohiro Yogo, 2019, A demand system approach to asset pricing, *Journal of Political Economy* 127, 1475–1515.

Lee, Charles M.C., and Bhaskaran Swaminathan, 2000, Price momentum and trading volume, *Journal of Finance* 55, 2017–2069.

Lee, Jieun, and Joseph P. Ogden, 2015, Did the profitability of momentum and reversal strategies decline with arbitrage costs after the turn of the millennium?, *Journal of Portfolio Management* 41, 70–83.

Li, Jiacui, 2020, What drives the size and value factors?, *Working paper* .

Lou, Dong, 2012, A flow-based explanation for return predictability, *Review of Financial Studies* 25, 3457–3489.

Lou, Dong, and Christopher Polk, 2018, Comomentum: Inferring arbitrage activity from return correlations, Working paper, London School of Economics.

Marquering, Wessel, Johan Nisser, and Toni Valla, 2006, Disappearing anomalies: A dynamic analysis of the persistence of anomalies, *Applied Financial Economics* 16, 291–302.

McLean, R. David, and Jeffrey Pontiff, 2016, Does academic research destroy stock return predictability?, *Journal of Finance* 71, 5–32.

Morningstar, 2016, The morningstar rating for funds, White paper, Morningstar.

Moskowitz, Tobias J., and Mark Grinblatt, 1999, Do industries explain momentum?, *Journal of Finance* 54, 1249–1290.

Novy-Marx, Robert, 2012, Is momentum really momentum?, *Journal of Financial Economics* 103, 429–453.

Novy-Marx, Robert, and Mihail Velikov, 2016, A taxonomy of anomalies and their trading costs, *The Review of Financial Studies* 29, 104–147.

Parker, Jonathan, Antoinette Schoar, and Yang Sun, 2020, Retail financial innovation and stock market dynamics: The case of target date funds, Working paper, Massachusetts Institute of Technology.

Pástor, Ľuboš, and Robert F. Stambaugh, 2003, Liquidity risk and expected stock returns, *Journal of Political Economy* 111, 642–685.

Pastor, Lubos, Robert F. Stambaugh, and Lucian A. Taylor, 2020, Fund tradeoffs, *Journal of Financial Economics* 614–634.

Shleifer, Andrei, 1986, Do demand curves for stocks slope down?, *Journal of Finance* 41, 579–590.

Teo, Melvyn, and Sung-Jun Woo, 2004, Style effects in the cross-section of stock returns, *Journal of Financial Economics* 74, 367–398.

Wahal, Sunil, and M. Deniz Yavuz, 2013, Style investing, comovement and return predictability, *Journal of Financial Economics* 107, 136–154.

Wermers, Russ, 2000, Mutual fund performance: An empirical decomposition into stock-picking talent, style, transaction costs, and expenses, *Journal of Finance* 55, 1655–1695.

Wurgler, Jeffrey, and Ekaterina Zhuravskaya, 2002, Does arbitrage flatten demand curves for stocks?, *Journal of Business* 75, 583–608.

Appendix A Data and Measures

A.1 Asset pricing factors

Table A.1 shows the list of 49 U.S. asset pricing factors used in this paper. Following Hou et al. (2020), we classify them into six categories: intangibles, investment, momentum, profitability, trading frictions, and value/growth. Figure A.1 and Table A.2 provide more details to the global factors we use from Jensen et al. (2021).

Figure A.1. Global Factors Data in Jensen et al. (2021)

This Figure tallies the global factors covered by Jensen et al. (2021). We divide data into different geographic areas and plot the number of countries and factors covered in Panels (a) and (b), respectively. As described in Section 2.2, we focus on factors that have full data since inception.

(a) Number of countries covered

(b) Number of factors

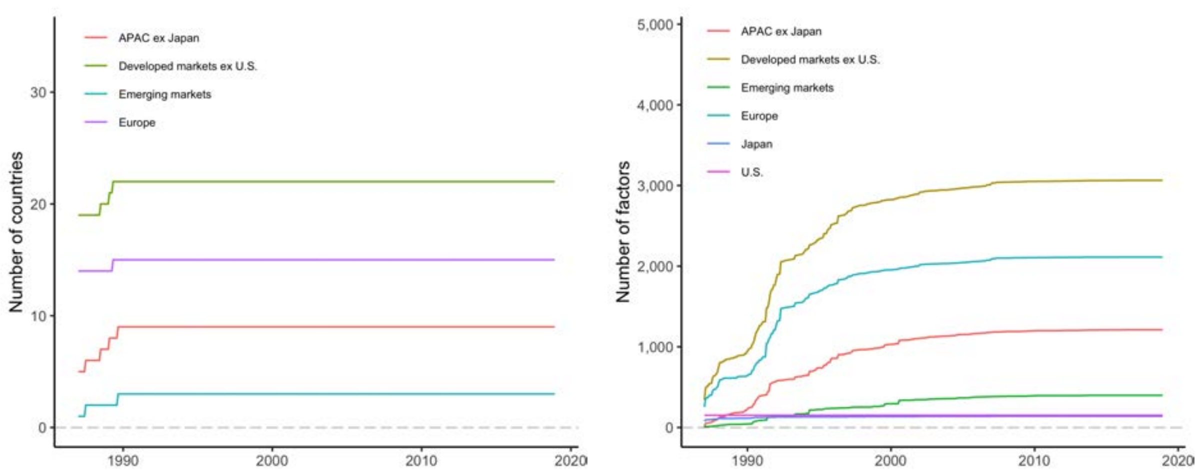


Table A.1. U.S. Asset Pricing Factors

The table lists the factors used in this study. The categorization is based on Hou et al. (2020).

Category	**Factor**	**Publication**
Intangibles (6)	Industry concentration	Hou and Robinson (JF 2006)
	Operating leverage	Novy-Marx (RF 2010)
	Firm age	Barry and Brown (JFE 1984)
	Advertising expense	Chan, Lakonishok, and Sougiannis (JF 2001)
	R&D expense	Chan, Lakonishok, and Sougiannis (JF 2001)
	Earnings persistence	Francis, LaFond, Olsson, and Schipper (AR 2004)
Investment (13)	Abnormal capital investment	Titman, Wei, and Xie (JFQA 2004)
	Accruals	Sloan (AR 1996)
	Asset growth	Cooper, Guylen, and Schill (JF 2008)
	Five-year share issuance	Daniel and Titman (JF 2006)
	Growth in inventory	Thomas and Zhang (RAS 2002)
	Industry-adjusted CAPEX growth	Abarbanell and Bushee (AR 1998)
	Investment growth	Xing (RFS 2008)
	Investment-to-assets	Hou, Xue, and Zhang (RFS 2015)
	Investment-to-capital	Xing (RFS 2008)
	Net operating assets	Hirshleifer, Hou, Teoh, and Zhang (JAE 2004)
	Net working capital changes	Soliman (AR 2008)
	One-year share issuance	Pontiff and Woodgate (JF 2008)
	Total external financing	Bradshaw, Richardson, and Sloan (JAE 2006)
Momentum (5)	52-week high	George and Hwang (JF 2004)
	Intermediate momentum ($t-7,\ t-12$)	Novy-Marx (JFE 2012)
	Industry momentum	Grinblatt and Moskwotiz (1999)
	Momentum ($t-2,\ t-6$)	Jegadeesh and Titman (JF 1993)
	Momentum ($t-1,\ t-12$)	Jegadeesh and Titman (JF 1993)
Profitability (14)	Cash-based profitability	Ball, Gerakos, Linnainmaa, and Nikolaev (JFE 2016)
	Change in asset turnover	Soliman (AR 2008)
	Distress risk	Campbell, Hilscher, and Szilagyi (JF 2008)
	Gross profitability	Novy-Marx (JFE 2013)
	Ohlson's O-score	Griffin and Lemmon (JF 2002)
	Operating profitability	Ball, Gerakos, Linnainmaa, and Nikolaev (JFE 2016)
	Piotroski's F-score	Piotroski (AR 2000)
	Profit margin	Soliman (AR 2008)
	QMJ profitability	Asness, Frazzini, Israel, Moskowitz, and Pederson (JFE 2018)
	Return on assets	Haugen and Baker (JFE 1996)
	Return on equity	Haugen and Baker (JFE 1996)
	Sales-minus-inventory growth	Abarbanell and Bushee (AR 1998)
	Sustainable growth	Lockwood and Prombutr (JFR 2010)
	Altman's Z-score	Dichev (JFE 1998)
Trading frictions (3)	Size	Banz (JFE 1981)
	Amihud illiquidity	Amihud (JFM 2002)
	Maximum daily return	Bali, Cakici, and Whitelaw (JFE 2011)
Value/Growth (8)	Book-to-market	Fama and French (JF 1992)
	Cash flow-to-price	Lakonishok, Shleifer, and Vishny (JF 1994)
	Earnings-to-price	Basu (JF 1977)
	Enterprise multiple	Loughran and Wellman (JFQA 2011)
	Sales growth	Lakonishok, Shleifer, and Vishny (JF 1994)
	Sales-to-price	Barbee, Mukherji, and Raines (FAJ 1996)
	Long-term reversals	Debondt and Thaler (JF 1985)
	Net payout yield	Boudoukh, Michaely, Richardson, and Roberts (JF 2007)

Journals: AR: Accounting Review, FAJ: Financial Analysts Journal, JAE: Journal of Accounting and Economics, JF: Journal of Finance, JFE: Journal of Financial Economics, JFQA: Journal of Financial and Quantitative Analysis, JFR: Journal of Financial Research, RAS: Review of Accounting Studies, RFS: Review of Financial Studies, RF: Review of Finance.

Table A.2. Countries Covered in Global Factors Data.

This table tallies the countries covered by the Jensen et al. (2021) global factors data. Panel A classifies countries or regions by market type and Panel B classifies by region.

Panel A: By market type

Market type	Country or region	Inception month	Number of factors: On average	At inception	By 2018
U.S.	United States	1987-01	152	152	152
Developed markets ex U.S.	UK	1987-01	143	90	150
	Netherlands	1987-01	133	77	143
	Japan	1987-01	133	80	139
	Germany	1987-01	125	2	152
	France	1987-01	125	2	150
	Sweden	1987-01	125	64	136
	Spain	1987-01	122	2	145
	Australia	1987-01	121	2	139
	Hong Kong	1987-01	120	2	144
	Switzerland	1987-01	119	2	146
	Denmark	1987-01	119	2	145
	Finland	1987-01	117	2	142
	Italy	1987-01	116	2	140
	Belgium	1987-01	116	2	143
	Singapore	1987-01	115	2	139
	Norway	1987-01	112	2	135
	Austria	1987-01	110	2	139
	New Zealand	1987-01	104	2	125
	Ireland	1987-01	102	2	126
	South Korea	1988-07	98	17	132
	Taiwan	1989-02	101	2	133
	Portugal	1989-05	97	2	121
Emerging markets	Malaysia	1987-01	110	2	133
	Thailand	1987-07	98	1	133
	India	1989-09	89	2	133

Panel B: By region					
Region	Country or region	Inception month	Number of factors		
			On average	At inception	By 2018
Europe	UK	1987-01	143	90	150
	Netherlands	1987-01	133	77	143
	Germany	1987-01	125	2	152
	France	1987-01	125	2	150
	Sweden	1987-01	125	64	136
	Spain	1987-01	122	2	145
	Switzerland	1987-01	119	2	146
	Denmark	1987-01	119	2	145
	Finland	1987-01	117	2	142
	Italy	1987-01	116	2	140
	Belgium	1987-01	116	2	143
	Norway	1987-01	112	2	135
	Austria	1987-01	110	2	139
	Ireland	1987-01	102	2	126
	Portugal	1989-05	97	2	121
APAC ex Japan	Australia	1987-01	121	2	139
	Hong Kong	1987-01	120	2	144
	Singapore	1987-01	115	2	139
	Malaysia	1987-01	110	2	133
	New Zealand	1987-01	104	2	125
	Taiwan	1989-02	101	2	133
	South Korea	1988-07	98	17	132
	Thailand	1987-07	98	1	133
	India	1989-09	89	2	133
Japan	Japan	1987-01	133	80	139

Appendix B Additional Empirical Results

B.1 The Momentum Profitability Drop In Mid 2002

This section examines the robustness of the finding that momentum-type factors became less profitable after mid 2002.

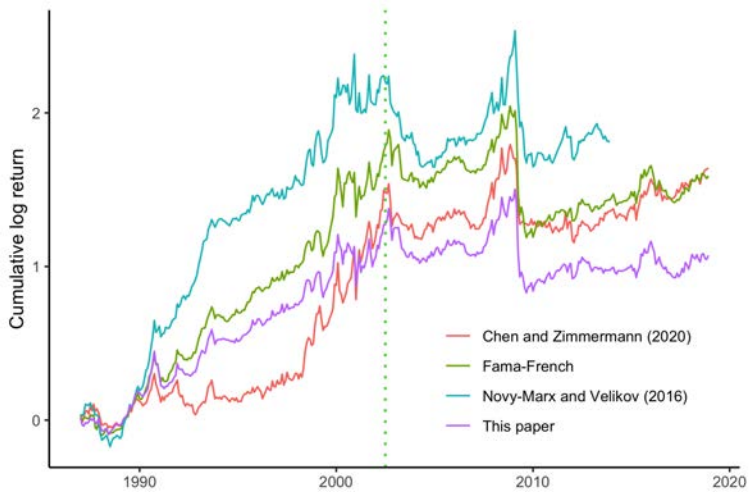
Robustness to alternative data sources. This finding is not specific to our factor universe or our factor construction methodology. Figure B.2 plots the cumulative return to momentum-related factors constructed from various sources. The purple line averages over the five momentum-type factors in this paper. The olive line uses the momentum factor downloaded from Ken French's website. The blue line, which ends in 2013 due to data availability, plots the average return of the two momentum-type factors in Novy-Marx and Velikov (2016). Finally, the red line plots the average return of the five momentum-type factors in Chen and Zimmermann (2020).[31] Despite differences in factor universe and factor construction, all four data sources show that momentum-related factors have suffered profitability declines since mid-2002.

Existing studies. We note that earlier studies have also shown evidence that suggests post-2002 return declines, even though detecting structural breaks is not their objective. For the reader's convenience, we present screenshots from those papers in Figure B.3. Panel (a) shows a chart from Green, Hand, and Zhang (2017), summarizing the average performance (equally-weighted as well as value-weighted) of 94 characteristics. Panel (b) shows a chart from Daniel and Moskowitz (2016) summarizing the performance to momentum strategy. In

[31] We obtain data in Novy-Marx and Velikov (2016) from Novy-Marx's website (`http://rnm.simon.rochester.edu/data_lib/ToAatTC/index.html`). We use their gross long-short factor returns in "returns to simple strategies". The two momentum-related factors are momentum and industry momentum. For data in Chen and Zimmermann (2020), we accessed the 0.1.2 version at `https://sites.google.com/site/chenandrewey/open-source-ap`, and used their "test asset portfolios" for NYSE-based value-weighted decile portfolios. We then constructed factors as long long the top decile and short the bottom decile. The five momentum-type factors include Junk Stock Momentum, 11 month residual momentum, 6 month residual momentum, 52 week high, and Industry Momentum.

Figure B.2. Momentum Profitability Decline After 2002: Other Data Sources

This figure plots the cumulative returns of momentum-related factors from various data sources. Please see the text for details about the data sources and factor construction process. The vertical dashed line marks the June 2002 Morningstar rating reform event.



both charts, we added a dashed line for June 2002.[32]

B.2 Estimating Price Impact Parameter for Explanatory Power Quantification

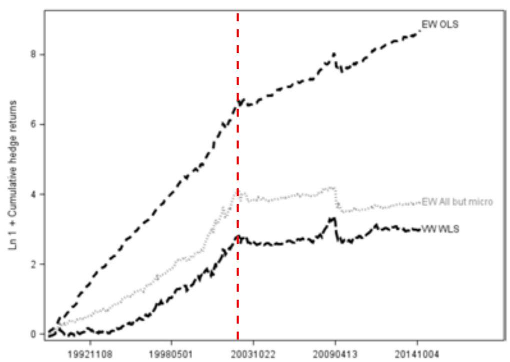
As described in Section 4.2, we use a twelve month window around the June 2002 methodology change to estimate the price impact of ratings. The regression results are shown in Figure B.2.

[32] Methodologically speaking, the finding of Green et al. (2017) is closer to our finding of factor momentum profit decline. Specifically, they investigate the profits to predicting stocks returns based on rolling multivariate Fama-MacBeth regressions with many stock characteristics. Therefore, their strategy ends up going long characteristics that recently performed well and short those that performed poorly – which is more similar to the factor momentum strategy in spirit. Even though they investigate characteristics and do not form factors, Cochrane (2011) notes that "portfolio sorts are really the same thing as nonparametric cross-sectional regressions," so the Green et al. (2017) also sheds light on factor-based results.

Figure B.3. Previous Evidence of Momentum-Type Strategy Profitability Decline

The figure presents charts in previous studies showing a kink in cumulative factor returns. In both panels, we added a red dashed line to mark the approximate location of June 2002 on the timeline. Panel (a) reproduces Figure 3 of Green et al. (2017). They study a strategy that uses 94 stock characteristics, and the different lines in the Figure represent different portfolio weighting methodologies. "EW OLS" refers to equal-weighting; "EW All but micro" refers to equal-weighting but excluding microcap stocks; "VW WLS" refers to value-weighted strategy. Panel (b) reproduces Figure 4b of Daniel and Moskowitz (2016) which plots the cumulative return to the momentum strategy. The Figures are taken from the latest SSRN versions of each paper: October 2016 version for Green et al. (2017), and July 2015 version of Daniel and Moskowitz (2016), with the authors' permissions.

(a) Green, Hand, and Zhang (2017, Fig 3)



(b) Daniel and Moskowitz (2016, Fig 4b)

Table B.2. Estimating Price Impact Coefficient (λ) Around the June 2002 Event

We use a panel regression to estimate the predictive relationship between monthly factor returns and the exponentially summed lagged ratings ($\text{ExpSum}(\Delta\text{Rating})_{f,t-1}$). The sample period is the 12 months around the reform (January to December 2002). We control for lagged factor returns in months $t-1$, $t-6$ to $t-2$, and $t-12$ to $t-7$. The four specifications differ in whether factor and month fixed effects are included. The standard errors in parentheses are adjusted for the cross-sectional correlation between factor returns using feasible generalized least squares.

Dependent variable:	Monthly factor return $\text{Ret}_{f,t}$(%)			
	(1)	(2)	(3)	(4)
$\text{ExpSum}(\Delta\text{Rating})_{f,t-1}$	2.270***	2.057***	2.388***	2.033***
	(0.534)	(0.493)	(0.516)	(0.478)
Lagged Returns	Yes	Yes	Yes	Yes
Factor FE	Yes	No	Yes	No
Month FE	Yes	Yes	No	No
Observations	588	588	588	588
Adjusted R^2	14.36%	7.21%	12.73%	5.86%

$^{***}p < 1\%$, $^{**}p < 5\%$, $^{*}p < 10\%$

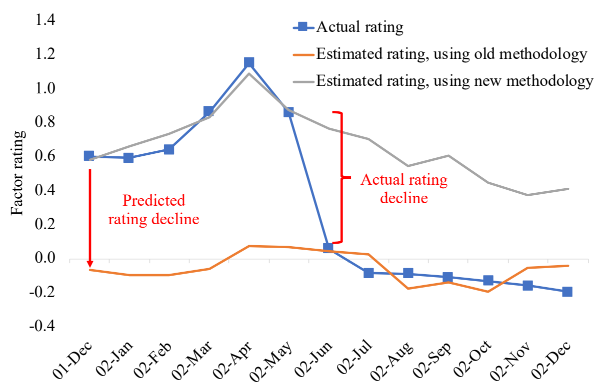
B.3 Event Study

Predicting Factor Rating Changes at the Reform Event In this section, we examine the accuracy of the factor-level rating change-prediction in Equation (10). We first illustrate the prediction method in Panels (a) and (b) of Figure B.4. Those two panels plot the two factors predicted to experience the largest rating decline (size) and increase (O-score). Our estimation matches actual ratings quite well. Before June 2002, the actual ratings closely match the estimated ratings under the old methodology (grey lines), and, after June 2002, the actual ratings closely match the estimated ratings under the new methodology (orange lines). Further, because the changes of factor-level ratings of factors over a few months is small, the predicted rating change using December 2001 data ends up being a reasonable predictor of the actual rating change in June 2002. This is further shown in Panel (c), where we plot the actual June 2002 rating changes of factors against the predicted changes. The latter explains the former with an R^2 of 84%.

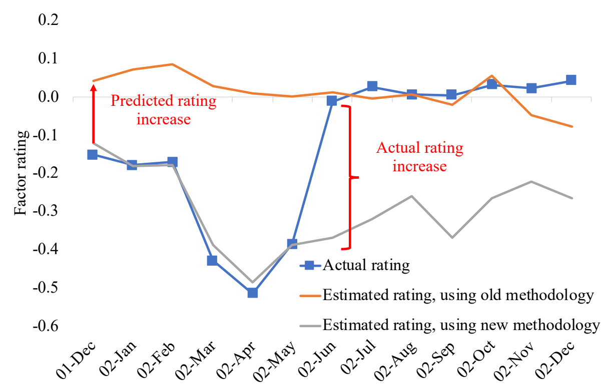
Figure B.4. Predicting Factor-level Rating Changes at the 2002 Reform Event

Panels (a) and (b) illustrate how we predict rating changes of factors at the June 2002 event using data in December 2001. Following Morningstar's rating construction process, we estimate ratings from ground up using fund returns. The grey lines plot the estimated rating under the old (pre-change) methodology, and the orange lines plot the estimated rating under the new (post-change) methodology. We use the difference between the two estimates in December 2001 (marked using red arrows) as the predicted rating change. The blue lines are the actual ratings. Panel (a) and (b) plot the factor with the largest predicted rating decline and increase, respectively (size and O-Score factors). Panel (c) compares the actual rating change in June 2002 against the predicted change using data in December 2001. The factors are sorted into quintiles based on the predicted rating change and colored differently.

(a) Most negatively affected factor



(b) Most positively affected factor



(c) Accuracy of rating change prediction