Factor Momentum and the Momentum Factor*

Sina Ehsani† Juhani Linnainmaa‡

First draft: March 2017

This draft: December 2020

Abstract

Momentum in individual stock returns emanates from momentum in factor returns. Most factors are positively autocorrelated: the average factor earns a monthly return of 6 basis points following a year of losses and 51 basis points following a positive year. We find that factor momentum concentrates in factors that explain more of the cross section of returns and that it is not incidental to individual stock momentum: momentum-neutral factors display more momentum. Momentum found in high-eigenvalue PC factors subsumes all forms of individual stock momentum. Our results suggest that momentum is not a distinct risk factor; it times other factors.

JEL classification: G11, G12, G40

Keywords: Factors; Anomalies; Momentum

*We thank Huaizhi Chen (discussant), Ing-Haw Cheng, John Cochrane, Dong Lou (discussant), Amit Goyal, Christian Goulding, Mark Grinblatt, Serhiy Kozak, Jon Lewellen, Michele Mazzoleni, Paulo Maio (discussant), Stefan Nagel (editor), Sheridan Titman, Ivo Welch, and Guofu Zhou, the associate editor, and two anonymous referees for valuable comments. We also thank Louis Yang for outstanding research assistance. We are grateful for the feedback by the conference and seminar participants at Washington University in St. Louis, DePaul University, Ohio State University, Hong Kong University of Science and Technology, Hong Kong Polytechnic University, University of New South Wales, Baruch College, University of Texas–Rio Grande Valley, Clemson University, Northern Illinois University, Monash University, La Trobe University, Australian National University, University of Technology Sydney, UBS Sydney, University of California–Riverside, University of Oxford, Northern Trust Asset Management, Kepos Capital, 2018 Mark Grinblatt Retirement Conference at UCLA, 2018 Midwest Finance Association meetings, 2019 Western Finance Association meetings, 2019 UBS Quant Conference, 2020 American Finance Association meetings, the 3rd Wolfe Global Quantitative and Macro Investing Conference, and Q Group Fall 2020 Conference.

†School of Business, Northern Illinois University. Email: sehsani@niu.edu.

‡Dartmouth College and NBER. Linnainmaa is also affiliated with Citadel and Research Affiliates. Neither Citadel nor Research Affiliates provided any funding for this research. Email: juhani.t.linnainmaa@tuck.dartmouth.edu. Tel: (603) 646-3160.

1 Introduction

Momentum appears to violate the efficient market hypothesis in its weakest form. Past returns should not predict future returns if asset prices respond to new information immediately and to the right extent—unless past returns correlate with changes in systematic risk. Researchers have sought to explain momentum with time-varying risk, behavioral biases, and trading frictions.[1] At the same time, the pervasiveness of momentum over time and across asset classes has given momentum the status of an independent factor: models without momentum cannot explain it and those with momentum cannot explain anything more than just momentum (Fama and French, 2016).[2] In this paper we show that momentum is a dynamic portfolio that times other factors. Rather than being unrelated to the other factors, momentum relates to *all* of them.

We first show that factors' prior returns are informative about their future returns. Small stocks, for example, are likely to outperform big stocks when they have done so over the prior year. This effect is economically and statistically large among the 20 factors we initially study: The average factor earns 51 basis points per month following a year of gains but just 6 basis points following a year of losses. This difference is significant with a t-value of 4.22. This result is not specific to the use of obscure asset pricing factors: we use off-the-shelf factors that are regularly updated and published by academics and a hedge fund.

Factor momentum bets on the continuation in factor returns. A *time-series* momentum strategy, which is long factors with positive returns and short those with negative returns, earns an annualized return of 3.9% (t-value = 7.01). This strategy dominates the *cross-sectional* strategy, which is long

[1] See, for example, Conrad and Kaul (1998), Berk et al. (1999), Johnson (2002), and Sagi and Seasholes (2007) for risk-based explanations; Daniel et al. (1998), Hong and Stein (1999), Frazzini et al. (2012), Cooper et al. (2004), Griffin et al. (2003), and Asness et al. (2013) for behavioral explanations; and Korajczyk and Sadka (2004), Lesmond et al. (2004), and Avramov et al. (2013) for trading friction-based explanations.

[2] Jegadeesh (1990) and Jegadeesh and Titman (1993) document momentum in the cross section of stocks, Jostova et al. (2013) in corporate bonds, Beyhaghi and Ehsani (2017) in corporate loans, Hendricks et al. (1993), Brown and Goetzmann (1995), Grinblatt et al. (1995), and Carhart (1997) in mutual funds, Baquero et al. (2005), Boyson (2008), and Jagannathan et al. (2010) in hedge funds, Bhojraj and Swaminathan (2006), Asness et al. (2013), and Moskowitz et al. (2012) in major futures contracts, Miffre and Rallis (2007) and Szakmary et al. (2010) in commodity futures, Menkhoff et al. (2012) in currencies, and Lee et al. (2014) in credit default swaps.

factors with above-median returns and short those with below-median returns, because it is a pure bet on the autocorrelations in factor returns. A cross-sectional strategy, by contrast, also bets that a high return on a factor predicts low returns on the other factors (Lo and MacKinlay, 1990); in the data, however, high return on a factor typically predicts high returns also on other factors.

Why are factors autocorrelated? We show that Kozak et al.'s (2018) model of sentiment investors leads to factor reversal or momentum depending on the persistence of sentiment. If sentiment is sufficiently persistent, this persistence carries over to factor returns. Although arbitrageurs know that factor premiums are predictable, they do not trade sufficiently aggressively to neutralize this effect because, by doing so, they would expose themselves to factor risk. This model predicts that momentum should concentrate in more systematic factors, much as in Kozak et al. (2018) it is the sentiment-driven demand component that aligns with covariances that distorts asset prices.

We extract principal components from 47 factors from Kozak et al. (2020). We find that factor momentum concentrates in the high-eigenvalue PCs, that is, in factors that explain more of the cross section of returns. A strategy that trades the first ten high-eigenvalue PCs has a five-factor model alpha that is significant with a t-value of 6.51. Momentum in this set of PCs either greatly reduces the momentum in the other subsets of PCs (the first half of the sample) or fully subsumes it (the second half). The finding that momentum concentrates into high-eigenvalue factors is consistent with the absence of near-arbitrage opportunities; if low-eigenvalue factors exhibited momentum, arbitrageurs could profit from this effect without assuming much factor risk. Haddad et al. (2020) find that predictability based on factors' valuation ratios also concentrates this same way.

Momentum in factor returns transmits into the cross section of security returns, and the amount that transmits depends on the dispersion in factor loadings. The more these loadings differ across assets, the more of the factor momentum shows up as *cross-sectional* momentum in individual security returns. This transmission mechanism sets up the main hypothesis that we test in this paper: Do individual stock returns display momentum beyond that which emanates from factor

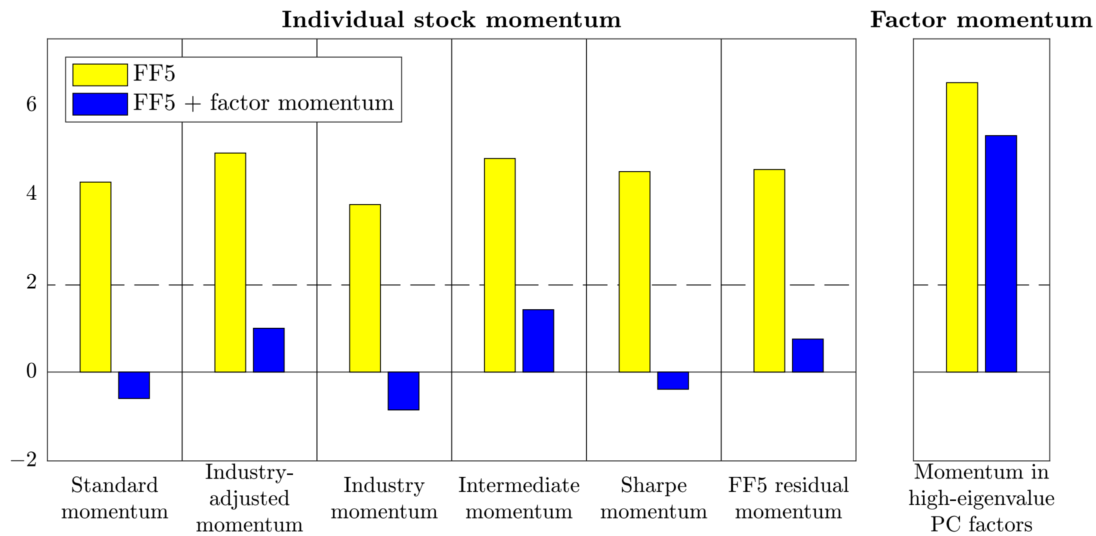

Figure 1: **Individual stock momentum versus factor momentum.** This figure shows t-values associated with alphas for six momentum strategies that trade individual stocks and a factor momentum strategy that trades the first ten high-eigenvalue principal components extracted from 47 factors. For individual stock momentum strategies, we report t-values from the five-factor model (yellow bars) and this model augmented with the factor momentum strategy (blue bars). The blue residual momentum regression also includes betting-against-beta factors. For factor momentum we report t-values from the five-factor model (yellow bar) and this model augmented with the first five individual stock momentum strategies (blue bar). The dashed line denotes a t-value of 1.96.

returns? Our empirical strategy in testing this hypothesis is to confront various strategies that trade individual stock momentum with factor momentum.

We begin by pricing portfolios sorted by prior one-year returns. We find that factor momentum, if anything, prices portfolios sorted by prior one-year returns better than Carhart's (1997) UMD factor, a factor that *directly* targets momentum in stock returns. When we augment the five-factor model with a factor that trades momentum in the high-eigenvalue PCs, mean absolute alphas are negligible and the GRS test does not reject the null that these alphas are jointly zero.

Factor momentum also explains other forms of stock momentum: industry momentum, industry-adjusted momentum, intermediate momentum, and Sharpe ratio momentum. The left-hand side of Figure 1 shows two t-values for each version of individual stock momentum. The first is that associated with the strategy's five-factor model alpha; the second is from a model that also captures momentum found in the first ten high-eigenvalue PCs. Factor momentum renders these forms of individual stock momentum strategies statistically insignificant. The right-hand side of the same figure shows that a five-factor model augmented with the aforementioned five forms of individual stock momentum leaves factor momentum with an alpha that is significant with a t-value of 5.32.

Residual momentum strategies are of independent interest. A strategy that selects stocks based on their CAPM residuals is more profitable than a strategy that selects stocks based on their total past returns. However, as we remove additional factors from stock returns—such as value and size—residual momentum strategies weaken. This pattern, too, appears to relate to factor momentum. If an investor works with a misspecified asset pricing model, residual momentum strategies profit from "omitted-factor momentum" even when firm-specific innovations are IID. If the factors in the investor's model are less autocorrelated than those it omits, residuals display *more* momentum than total stock returns. The finding that residuals exhibit more momentum than raw stock returns should therefore not be construed as evidence that *firm-specific* returns display momentum. Consistent with this omitted-factor argument, no residual momentum strategy is significant net of factor momentum—Figure 1 shows $t(\hat{\alpha})$s for one such strategy. In short, our empirical tests indicate that momentum found in the first ten PCs subsumes all versions of individual stock momentum. We uncover no evidence of momentum in stock returns beyond that emanating from factor momentum.

If factors are linear combinations of individual stocks, is not factor momentum ultimately a reflection of individual stock momentum? Our result that the nature of factors matters—more systematic factors display more momentum—is one step towards illustrating that factors are distinct from individual stocks. We also take a step further by constructing *momentum-neutral factors*; these are factors whose weights are as close as possible to the original factors but orthogonal to past stock returns. An investor investing in a momentum-neutral size factor, for example, would buy and sell small and large stocks that are identical in terms of their past returns. We show that momentum-neutral factors exhibit *more* momentum than standard factors and that factor momentum in momentum-neutral factors subsumes standard factor momentum. Factor momentum is therefore not merely incidental to individual stock momentum. Of all the factor momentum strategies we consider, the one with the highest Sharpe ratio is the one that trades momentum in the high-eigenvalue PC factors extracted from momentum-neutral factors. This strategy's five-factor model alpha is significant with a t-value of 7.53.

Our results suggest that momentum is, in large part, about timing other factors. This characterization of momentum resolves the perennial question about covariances and momentum (Cochrane, 2011, p. 1075): "...why should all the momentum stocks then rise and fall together the next month, just as if they are exposed to a pervasive, systematic risk?" Momentum stocks comove because they are exposed to the same systematic risks; winners, for example, load positively on factors that have done well and negatively on those that have done poorly. Because momentum's loadings change over time, we are easily left with the impression that momentum is distinct from other risk factors.

A clarifying note about momentum's status as a distinct risk factor is in order. Momentum is distinct from, for example, the five factors of the Fama-French model in the sense that a static combination of these factors does not span momentum. Our result is that we can capture all of momentum profits by timing other factors. We could also, alternatively, redefine the factors we already have and push momentum back into them. Ehsani and Linnainmaa (2020a) show that UMD is spanned in an unconditional regression against what they call a "time-series efficient" Fama-French five-factor model. There is no need to construct a separate momentum factor from security-level data; the factors we already have will do. However, even if one accepts our conclusion that momentum is not a distinct factor, it does not mean that investors can ignore momentum; to capture it, investors still need to time the other factors, redefine these other factors or, if they so insist, trade individual stock momentum as if it was distinct from the other factors.

Our results relate to McLean and Pontiff (2016), Avramov et al. (2017), and Zaremba and Shemer (2017) who show that anomaly returns predict the cross section of anomaly returns at the one-month and one-year lags. Arnott et al. (2019) show that short-term cross-sectional factor momentum explains short-term industry momentum. That alternative form of factor momentum, however, explains none of individual stock momentum, consistent with the finding of Grundy and Martin (2001) that industry momentum is largely unrelated to stock momentum.

2 Factor Momentum in Off-the-Shelf Factors

2.1 Data

We take monthly factor data from three sources: Kenneth French's, AQR's, and Robert Stambaugh's data libraries.[3] Table 1 lists the factors, start dates, average annualized returns, standard deviations of returns, and t-values associated with the average returns. If the factor return data are not provided, we compute factor return as the average return on the three top deciles minus that on the three bottom deciles, where the top and bottom deciles are defined in the same way as in the original study.

The 15 anomalies that use U.S. data are size, value, profitability, investment, momentum, accruals, betting against beta, cash-flow to price, earnings to price, liquidity, long-term reversals, net share issues, quality minus junk, residual variance, and short-term reversals. Except for the liquidity factor of Pástor and Stambaugh (2003), the return data for these factors begin in July 1963; those for the liquidity factor begin in January 1968. The seven global factors are size, value, profitability, investment, momentum, betting against beta, and quality minus junk. Except for the momentum factor, the return data for these factors begin in July 1990; those for the momentum factor begin in November 1990. We call this set of 22 factors the "off-the-shelf" factors. We later study a broader set of 47 U.S. factors.

Table 1 shows significant variation in average annualized returns. The global size factor, for example, earns 1.1%, while both the U.S. and global betting against beta factors earn almost 10%.

[3]These data sets are available at http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html, https://www.aqr.com/insights/datasets, and http://finance.wharton.upenn.edu/~stambaug/.

Table 1: Descriptive statistics

This table reports the start date, the original study, and the average annualized returns, standard deviations, and t-values for 15 U.S. and seven global factors. The universe of stocks for the global factors is the developed markets excluding the U.S. The end date for all factors is December 2019.

Factor	Original study	Start date	Annual return Mean	SD	t-value
U.S. factors					
Size	Banz (1981)	Jul 1963	2.7%	10.4%	1.97
Value	Rosenberg et al. (1985)	Jul 1963	3.7%	9.7%	2.82
Profitability	Novy-Marx (2013)	Jul 1963	3.1%	7.5%	3.13
Investment	Titman et al. (2004)	Jul 1963	3.3%	6.9%	3.59
Momentum	Jegadeesh and Titman (1993)	Jul 1963	7.8%	14.5%	4.02
Accruals	Sloan (1996)	Jul 1963	2.8%	6.6%	3.19
Betting against beta	Frazzini and Pedersen (2014)	Jul 1963	9.8%	11.2%	6.55
Cash-flow to price	Rosenberg et al. (1985)	Jul 1963	3.4%	8.6%	2.94
Earnings to price	Basu (1983)	Jul 1963	3.5%	8.9%	2.95
Liquidity	Pástor and Stambaugh (2003)	Jan 1968	4.4%	11.6%	2.77
Long-term reversals	Bondt and Thaler (1985)	Jul 1963	2.5%	8.7%	2.16
Net share issues	Loughran and Ritter (1995)	Jul 1963	2.8%	8.2%	2.52
Quality minus junk	Asness et al. (2019)	Jul 1963	4.6%	7.7%	4.47
Residual variance	Ang et al. (2006)	Jul 1963	1.6%	17.3%	0.68
Short-term reversals	Jegadeesh (1990)	Jul 1963	6.0%	10.6%	4.21
Global factors					
Size	Banz (1981)	Jul 1990	1.1%	7.1%	0.83
Value	Rosenberg et al. (1985)	Jul 1990	4.0%	7.4%	2.92
Profitability	Novy-Marx (2013)	Jul 1990	4.3%	4.7%	4.91
Investment	Titman et al. (2004)	Jul 1990	1.9%	6.1%	1.74
Momentum	Jegadeesh and Titman (1993)	Nov 1990	7.9%	12.1%	3.54
Betting against beta	Frazzini and Pedersen (2014)	Jul 1990	9.6%	9.7%	5.70
Quality minus junk	Asness et al. (2019)	Jul 1990	6.3%	6.8%	5.06

Factors' volatilities also vary significantly. The global profitability factor, for example, has an annualized standard deviation of returns of just 4.7%; at the other extreme, the volatility of the residual variance factor is 17.3%.

2.2 Factor returns conditional on past returns

Table 2 shows that factors' prior returns significantly predict their future returns. We estimate time-series regressions in which the dependent variable is a factor's month t return, and the explanatory variable is an indicator variable for the factor's performance over the prior year from month $t-12$ to $t-1$. This indicator variable takes the value of one if the factor's return is positive, and zero otherwise.[4]

The intercepts in Table 2 measure the average factor returns earned following a year of underperformance. The slope coefficient represents the average return difference between the up- and down-years. In these regressions all slope coefficients, except that for the U.S. momentum factor, are positive. Six of the estimates are significant at the 5% level and an additional four at the 10% level. Although all factors' unconditional means are positive (Table 1), the intercepts show that six anomalies earn a negative average return following a year of underperformance. The first row shows that the amount of predictability in factor premiums is economically and statistically large. We estimate this pooled regression using data on the 20 non-momentum factors. The average anomaly earns a monthly return of just 6 basis points (t-value $= 0.72$) following a year of underperformance. When the anomaly's return over the prior year is positive, this return increases by 45 basis points (t-value $= 4.22$) to 51 basis points.

[4] Table A1 shows estimates from regressions of factor returns on prior one-year factor returns. We present the indicator-variable specification of Table 2 as the main specification because it is analogous to a strategy that signs the positions in factors based on their prior returns. Christoffersen and Diebold (2006) show that the *signs* of returns may display serial dependence even if means are unpredictable. Sign autocorrelation and the lack of autocorrelation in means can coexist if means are positive and volatility is serially dependent. The regressions in Table 2 are of the "return-on-sign" rather than "sign-on-sign" variety and therefore not subject to this mechanism; they show that signs predict differences in conditional means. The pooled estimate of 0.25 (t-value $= 2.59$) in Table A1's "return-on-return" regression also indicates that mean returns are autocorrelated.

Table 2: Average factor returns conditional on their own past returns

This table reports estimates from regressions in which the dependent variable is a factor's monthly return and the independent variable takes the value of one if the factor's average return over the prior year is positive and zero otherwise. We estimate these regressions using pooled data (first row) and separately for each anomaly (remaining rows). The pooled data exclude the two momentum factors. In the pooled regression we cluster standard errors by month. Table 1 reports the factor start dates. The sample ends in December 2019.

	Intercept		Slope	
Anomaly	$\hat{\alpha}$	$t(\hat{\alpha})$	$\hat{\beta}$	$t(\hat{\beta})$
Pooled	0.06	0.72	0.45	4.22
		U.S. factors		
Size	−0.10	−0.62	0.58	2.51
Value	0.04	0.20	0.41	1.78
Profitability	0.04	0.22	0.34	1.67
Investment	0.12	0.97	0.24	1.55
Momentum	0.72	2.70	−0.09	−0.29
Accruals	0.15	1.18	0.10	0.65
Betting against beta	−0.22	−0.63	1.32	3.53
Cash-flow to price	0.13	0.78	0.24	1.16
Earnings to price	0.10	0.62	0.30	1.46
Liquidity	0.16	0.74	0.36	1.29
Long-term reversals	−0.25	−1.66	0.76	3.85
Net share issues	0.17	1.32	0.09	0.49
Quality minus junk	0.09	0.65	0.43	2.51
Residual variance	−0.46	−1.64	1.06	2.74
Short-term reversals	0.49	1.43	0.01	0.04
		Global factors		
Size	−0.06	−0.39	0.28	1.33
Value	0.04	0.15	0.47	1.77
Profitability	0.14	1.03	0.26	1.62
Investment	−0.06	−0.41	0.38	1.94
Momentum	0.67	1.77	0.02	0.04
Betting against beta	0.19	0.58	0.84	2.30
Quality minus junk	0.39	1.76	0.12	0.49

2.3 Average returns of time-series and cross-sectional factor momentum strategies

We measure the profitability of strategies that take long and short positions in factors based on their prior returns. A time-series momentum strategy is long factors with positive returns over the prior year (winners) and short those with negative returns (losers). A cross-sectional momentum strategy is long factors that earned above-median returns relative to the other factors over the prior one-year period (winners) and short those with below-median returns (losers). We rebalance both strategies monthly.[5] We exclude the U.S. and global stock momentum factors to avoid inducing a mechanical correlation between factor momentum and individual stock momentum. The two factor momentum strategies therefore trade a maximum of 20 factors. The number of factors starts at 13 in July 1964 and increases to 20 by July 1991 because of the variation in the factors' start dates.

Table 3 shows the average returns for the time-series and cross-sectional factor momentum strategies as well as for an equal-weighted portfolio of all 20 factors. The annualized return on the average factor is 4.1% with a t-value of 7.77. In the cross-sectional strategy, both the winner and loser portfolios have the same number of factors. In the time-series strategy, the number of factors in these portfolios varies. For example, if five factors have above-zero and 15 factors below-zero returns over the prior year, the winner strategy is long five factors and the loser strategy is long the remaining 15 factors. The time-series momentum strategy takes positions in all 20 factors with the sign of the position in each factor determined by the factor's prior return. We report returns both for the factor momentum strategies as well as for the loser and winner portfolios.

Consistent with the results on the persistence in factor returns in Table 2, both winner strategies outperform the equal-weighted benchmark, and the loser strategies underperform it. The portfolio of time-series winners earns an average return of 5.9% (t-value = 10.03), and cross-sectional winners earn an average return of 6.5% (t-value = 8.98). The two loser portfolios earn average returns of

[5] In Appendix A we construct alternative strategies in which the formation and holding periods range from one month to two years.

Table 3: Average returns of time-series and cross-sectional factor momentum strategies

This table reports annualized average returns, standard deviations, and Sharpe ratios for different combinations of up to 20 factors. The number of factors increases from 13 in July 1964 to 20 by July 1991 (see Table 1). The equal-weighted portfolio invests in all factors with equal weights. The time-series factor momentum strategy is long factors with positive returns over the prior one-year period (winners) and short factors with negative returns (losers). The cross-sectional momentum strategy is long factors with above-median returns relative to other factors over the prior year (winners) and short factors with below-median returns (losers). The time-series strategy is on average long 11 factors and short 6 factors. The cross-sectional strategy is balanced because it selects factors based on their relative performance. We rebalance all strategies monthly. The sample begins in July 1964 and ends in December 2019.

	Annualized return			
Strategy	Mean	SD	t-value	Sharpe ratio
Equal-weighted portfolio	4.10	3.93	7.77	1.04
Time-series factor momentum	3.92	4.16	7.01	0.94
Winners	5.93	4.41	10.03	1.35
Losers	0.76	5.26	1.08	0.14
Cross-sectional factor momentum	2.40	3.55	5.04	0.68
Winners	6.45	5.35	8.98	1.21
Losers	1.69	5.26	2.39	0.32

0.8% and 1.7%, and the t-values associated with these averages are 1.08 and 2.39.

The momentum strategies are about the spreads between the winner and loser portfolios.[6] The time-series factor momentum strategy earns an annualized return of 3.9% (t-value = 7.01); the cross-sectional strategy earns a return of 2.4% (t-value = 5.04). Because time-series losers earn premiums that are close to zero, the choice of being long or short a factor following periods of negative returns is muted from the viewpoint of average returns. However, the time-series momentum strategy has a lower standard deviation than the winner portfolio alone (4.2% versus 4.4%) because it diversifies across all factors.

The difference between the time-series and cross-sectional factor momentum strategies is statistically significant. In a regression of the time-series strategy on the cross-sectional strategy, the estimated slope is 1.0 and the alpha of 1.5% is significant with a t-value of 5.14. In the reverse

[6]The mean return of the cross-sectional strategy is half of the difference between its winner and loser legs. The mean for the time-series strategy is closer to the mean of its winner leg because the strategy, on average, includes more long than short positions.

regression of the cross-sectional strategy on the time-series strategy, the estimated slope is 0.7 and the alpha of -0.5% has a t-value of -1.83. The time-series factor momentum therefore subsumes the cross-sectional strategy, but not vice versa.

Investors can capture the momentum premium without prespecifying which leg of the factor *on average* earns a higher return. Consider, for example, the discovery of a new A-minus-B factor, AMB. An investor who believes that this factor associates with an unconditional return premium needs to determine, ex ante, which leg outperforms the other. She would rely on historical data, an economic model, or on both together to make this determination. An investor who seeks to profit from the autocorrelation in factor returns, by contrast, does not need an estimate of the factor's unconditional mean. An investor who believes that AMB displays momentum would invest in AMB after a year of gains and in its reverse, "BMA," after a year of losses. Table 2 suggests that an investor with perfect foresight about the signs of the factors' unconditional premiums would have earned approximately the same return as the momentum investor.

Figure 2 plots the cumulative returns associated the equal-weighted portfolio and the winner and loser portfolios of Table 3. We leverage the strategies in this figure so that their volatilities are equal to that of the equal-weighted portfolio. Consistent with its near zero monthly premium, the total return on the time-series loser strategy remains close to zero even at the end of the 56-year sample period. The time-series winner strategy, by contrast, has earned twice as much as the passive strategy by the end of the sample period. Although the cross-sectional winner strategy in Panel A of Table 3 earns the highest average return, it is more volatile, and therefore underperforms the time-series winner strategy on a volatility-adjusted basis. The cross-sectional loser strategy earns a higher return than the time-series loser strategy: factors that underperformed other factors but that still earned *positive* returns tend to earn positive returns going forward. The winner-minus-loser gap is therefore wider for the time-series strategy than what it is for the cross-sectional strategy.

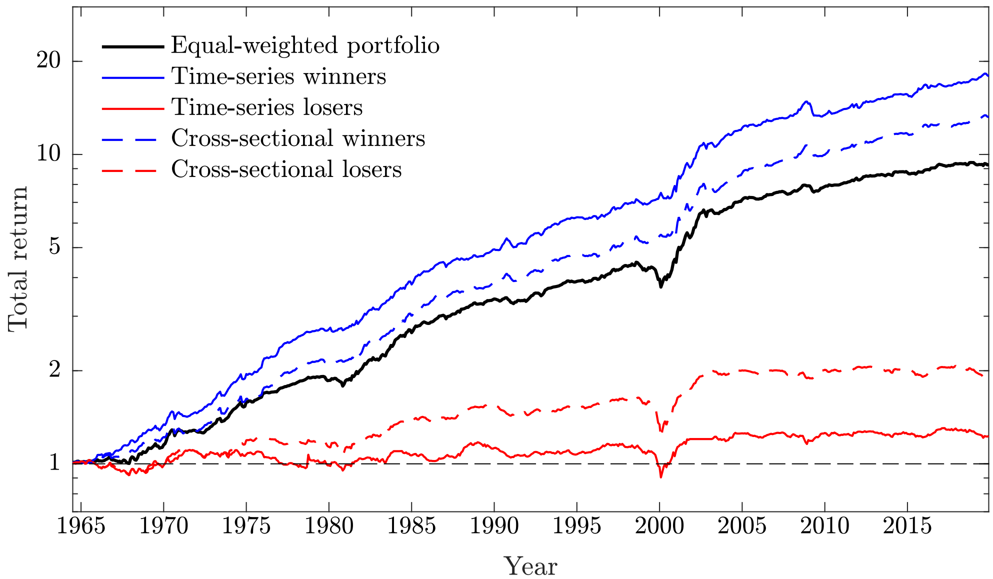

Figure 2: **Profitability of time-series and cross-sectional factor momentum strategies, July 1964–December 2019.** This figure displays total return on an equal-weighted portfolio of all factors and the returns on factors partitioned into winners and losers by their past performance. Time-series winners and losers are factors with above- or below-zero return over the prior year. Cross-sectional winners and losers are factors that have out- or underperformed the median factor over this formation period. Each portfolio is rebalanced monthly and each portfolio's standard deviation is standardized to equal that of the equal-weighted portfolio.

2.4 Decomposing factor momentum profits: Why does the cross-sectional strategy underperform the time-series strategy?

The difference between the cross-sectional and time-series factor momentum strategies is significant. In this section we use Lo and MacKinlay (1990) and Lewellen (2002) decompositions to quantify the sources of profits to each strategy, and to identify the primary cause of their difference. The cross-sectional decomposition chooses portfolio weights that are proportional to demeaned past returns. We study the performance of this "linear-weight" factor momentum strategy only in this section (and in the related Appendix B in the Appendix) because it lends itself to the analytical Lo-MacKinlay decomposition. The weight on factor f in month t is positive if the factor's past return is above average and negative if it is below average:[7]

$$w_t^f = r_{-t}^f - \bar{r}_{-t}, \tag{1}$$

where r_{-t}^f is factor f's past return over some formation period such as from month $t-12$ to month $t-1$ and $\bar{r}_{-t}$ is the cross-sectional average of all factors' returns over the same formation period. The month-t return that results from the position in factor f is therefore

$$\pi_t^f = (r_{-t}^f - \bar{r}_{-t})\, r_t^f, \tag{2}$$

where r_t^f is factor f's return in month t. Averaging the profits in equation (2) across the F factors and taking expectations, we get

$$\mathrm{E}[\pi_t^{\mathrm{XS}}] = \mathrm{E}\big[\sum_{f=1}^{F} \frac{1}{F}(r_{-t}^f - \bar{r}_{-t})r_t^f\big] = \frac{1}{F}\sum_{f=1}^{F} \mathrm{cov}(r_{-t}^f, r_t^f) - \mathrm{cov}(\bar{r}_{-t}, \bar{r}_t) + \frac{1}{F}\sum_{f=1}^{F}(\mu^f - \bar{\mu})^2, \tag{3}$$

where μ^f is factor f's unconditional expected return. The three potential sources of profits can be isolated by writing equation (3) in matrix notation,

$$\begin{aligned} \mathrm{E}[\pi_t^{\mathrm{XS}}] &= \frac{1}{F}\mathrm{Tr}(\Omega) - \frac{1}{F^2}1'\Omega 1 + \sigma_\mu^2 \\ &= \frac{F-1}{F^2}\mathrm{Tr}(\Omega) - \frac{1}{F^2}(1'\Omega 1 - \mathrm{Tr}(\Omega)) + \sigma_\mu^2, \end{aligned} \tag{4}$$

where $\Omega = \mathrm{E}\big[(r_{-t}^f - \mu)(r_t^f - \mu)'\big]$ is the autocovariance matrix of factor returns, $\mathrm{Tr}(\Omega)$ is the trace of this matrix, and σ_μ^2 is the cross-sectional variance of mean factor returns.

Equation (4) separates cross-sectional momentum profits to three sources:

1. Positive autocovariances in factor returns: a past high factor return signals future high return.

[7] The key idea of the Lo and MacKinlay (1990) decomposition is the observation that, by creating a strategy with weights proportional to past returns, the strategy's expected return is the expected product of lagged and future returns. This expected product can then be expressed as the product of expectations plus the covariance of returns.

Table 4: Decomposition of factor momentum profits

This table decomposes the profits of the cross-sectional and time-series factor momentum profits using equations (4) and (5). We report the premiums in percentages per year. We multiply the cross-serial covariance term by -1 so that these terms represent their net contributions to the returns of the cross-sectional and time-series strategies. We compute the standard errors by block bootstrapping the factor return data by month. When month t is sampled, we associate month t with the factors' average returns from month $t-12$ to $t-1$ to compute the terms in the decomposition. The sample begins in July 1964 and ends in December 2019.

Strategy	Decomposition	Annualized premium (%)	Standard error
Cross-sectional	Autocovariances	2.54	0.97
factor momentum	− Cross-serial covariances	−1.00	0.50
	+ Variance of mean returns	0.48	0.15
	= Cross-sectional factor momentum	2.16	0.66
Time-series	Autocovariances	2.67	1.02
factor momentum	+ Mean squared returns	1.76	0.42
	= Time-series factor momentum	4.51	1.00

2. Negative cross-serial covariances: a past high factor return signals low returns on other factors.

3. Cross-sectional variance of mean returns: some factors earn persistently high or low returns.

The last term is independent of the autocovariance matrix; that is, factor "momentum" can emerge even in the absence of any time-series predictability. A cross-sectional strategy is long the factors with the highest past returns and short the factors with the lowest past returns; therefore, if past returns are good estimates of factors' unconditional means, a cross-sectional momentum strategy earns positive returns even in the absence of auto- and cross-serial covariance patterns (Conrad and Kaul, 1998).

Table 4 shows that the cross-sectional momentum strategy in equation (4) earns an average annualized return of 2.16% with a t-value of 3.26. The autocovariance term contributes an average of 2.54%, more than all of the cross-sectional strategy's profits. The cross-serial covariance term is positive and, therefore, makes a *negative* contribution (−1.00% per year) to the cross-sectional strategy's profits. A positive return on a factor predicts positive returns also on the other factors, and the cross-sectional strategy loses by trading against this cross-predictability.[8] This negative term more than offsets the positive contribution of the cross-sectional variation in means (0.48% per year).

Whereas the cross-sectional strategy's weights are based on the factors' *relative* performance, those of the time-series strategy are based on their *absolute* performance. The weight on factor f in month t is its return over the formation period, $w_t^f = r_{-t}^f$. Following Moskowitz et al. (2012), the time-series momentum strategy's expected return decomposes as:

$$\mathrm{E}[\pi_t^{\mathrm{TS}}] = \frac{1}{F}\mathrm{E}\big[\sum_{f=1}^{F} r_{-t}^f r_t^f\big] = \frac{1}{F}\sum_{f=1}^{F}\big[\mathrm{cov}(r_{-t}^f, r_t^f) + (\mu^f)^2\big] = \frac{1}{F}\mathrm{Tr}(\Omega) + \frac{1}{F}\sum_{f=1}^{F}(\mu^f)^2. \qquad (5)$$

Equation (5) shows that the time-series momentum profits stem either from autocorrelation in factor returns or from mean returns that are either very positive or negative.[9]

Table 4 shows that the time-series strategy earns an annualized return of 4.51% (t-value = 4.49). The decomposition of these profits into the autocorrelation and mean-squared components shows that this premium largely derives from the autocorrelation in factor returns; the annualized premiums associated with these two components are 2.67% (t-value of 2.62) and 1.76% (t-value = 4.14). The time-series strategy outperforms the cross-sectional strategy because it does not bet on factors displaying negative cross-serial covariances; it is a pure bet on factor autocorrelations.

[8]The cross-sectional strategy is always long and short the same number of factors. If the cross-serial covariance term is non-zero, this balance is not optimal. Suppose, for example, that all factors have earned positive profits over the prior year. The positive autocovariance and cross-serial covariance terms then predict positive returns on *all* factors. By the virtue of being long and short the same number of factors, the cross-sectional strategy loses by shorting factors with poor performance *relative* to other factors.

[9]Autocovariances appear in the decompositions of both the cross-sectional and time-series strategies. The scaling factor of the autocovariance term, however, is different. In the cross-sectional decomposition we isolate the diagonal elements of the covariance matrix to attribute this strategy's profits to the auto- and cross-covariance components. The last terms in the two decomposition also differ. In the cross-sectional decomposition in equation (4), it is the variance of mean returns; in the time-series decomposition in equation (5), it is the sum of squared mean returns.

3 Factor Momentum and the Covariance Structure of Returns

3.1 Factor momentum in economies with sentiment investors

Why are factors autocorrelated? In this section we build on the Kozak et al. (2018) model to derive the conditions under which factors exhibit momentum and characterize the properties of the factors that exhibit the most momentum. We first describe the key elements of the Kozak et al. (2018) model. The economy has two types of risk-averse investors: fully rational arbitrageurs and sentiment investors with distorted beliefs about asset returns' true distributions. Asset cash flows are IID and the covariance matrix of these cash flows has a few dominant factors. Sentiment investors' demand has an additional sentiment-driven demand component. Sentiment investors cannot take substantial leverage or short extensively. By market clearing, rational arbitrageurs trade against sentiment investors. Kozak et al. (2018) study the extent to which, and under what conditions, sentiment distorts asset prices.

The key finding of Kozak et al. (2018) is that arbitrageurs almost fully subsume any sentiment-driven demand not aligned with common factor covariances. The intuition is that arbitrageurs can make these profitable trades without assuming any factor risk, therefore neutralizing these components of sentiment investors' demand. Conversely, arbitrageurs are reluctant to take the other side of those sentiment-driven trades that align with common factor covariances; such trades would expose them to factor risk. This dichotomy implies that even if sentiment-driven demand has nothing to do with the covariances of cash flows, those mispricings that align with covariances remain. Kozak et al.'s (2018) conclusion is that the absence of near-arbitrage opportunities together with the substantial commonality in asset returns ensures that the stochastic discount factor can be represented as a function of a few dominant factors. The ability to do so provides no clues as to whether pricing is rational or subject to behavioral distortions.

We now derive the condition under which asset returns and the factors in this model are autocorrelated. In what follows, we assume that the reader is familiar with Sections III and IV and

Appendix C of the original paper. Kozak et al. (2018, equation (C5)) gives the realized returns as

$$R_{t+1} = D_{t+1} + a_1(\xi_{t+1} - \xi_t) - R_f\,(a_0 + a_1\xi_t), \tag{6}$$

where R_{t+1} is an $N \times 1$ vector of asset returns, D_{t+1} are the dividends, R_f is the risk-free rate, a_0 and a_1 are vectors of constants, and ξ_t is the sentiment-investor demand. This demand follows an AR(1) process, $\xi_{t+1} = \mu + \phi\xi_t + \nu_{t+1}$, with $\text{var}(\nu_{t+1}) = \omega^2$. Sentiment investors' demand is distorted in direction δ by the amount ξ_t. From equation (6), the return autocovariance matrix is

$$\begin{aligned} \text{cov}(R_t, R_{t+1}) &= a_1 a_1' \text{cov}(\xi_t - R_f\,\xi_{t-1}, \xi_{t+1} - R_f\,\xi_t) \\ &= a_1 a_1' \sigma^2 \left[(1 + R_f^2)\phi - R_f - R_f\,\phi^2\right], \end{aligned} \tag{7}$$

where the second row uses the properties of the AR(1) process, $\sigma^2 \equiv \text{var}(\xi_t) = \frac{\omega^2}{1-\phi^2}$ and $\text{cov}(\xi_t, \xi_{t+h}) = \phi^{|h|}\sigma^2$.

Kozak et al. (2018) note that a_1 can be solved from the arbitrageurs' first-order condition (equation (C10) in KNS) combined with the market-clearing condition (equation (31) in KNS) using the method of undetermined coefficients. Specifically, b_2 appears in the term multiplying ξ_t in the first-order condition and, because market clearing has to hold for any value of ξ_t, this slope must be zero. Collecting the terms, a_1 can be written as

$$a_1 = \frac{\gamma\theta\Gamma\delta}{R_f + \frac{1}{1+2b_2\omega^2}\left(\frac{\gamma\theta\delta' a_1}{2b_2} - \phi\right) - \frac{\gamma\theta\delta' a_1}{2b_2}}, \tag{8}$$

and therefore[10]

$$a_1 a_1' = \frac{\gamma^2\theta^2\Gamma\delta\delta'\Gamma}{\left[R_f + \frac{1}{1+2b_2\omega^2}\left(\frac{\gamma\theta\delta' a_1}{2b_2} - \phi\right) - \frac{\gamma\theta\delta' a_1}{2b_2}\right]^2} = \Gamma\delta\delta'\Gamma c_0. \tag{9}$$

[10]Constant $c_0 > 0$ has (scalar) $\delta' a_1$ in the denominator; it could be eliminated by premultiplying both sides of equation (8) by δ', solving for $\delta' a_1$, and plugging it back into this expression. However, for our purposes, the value of the denominator does not matter, and it has to be positive for the solution for a_1 to exist.

The factors in Kozak et al. (2018) are the eigenvectors of the covariance matrix of asset cash flows, $\Gamma = Q\Lambda Q$, where Q is the matrix of eigenvectors and Λ is a diagonal matrix with the eigenvalues. Following Kozak et al. (2018), we consider factor q_k, which is the kth principal component. The autocovariance of this factor is

$$\begin{aligned} \mathrm{cov}(PC_t^k, PC_{t+1}^k) &= \mathrm{cov}(q_k' R_t, q_k' R_{t+1}) = q_k' \mathrm{cov}(R_t, R_{t+1}) q_k \\ &= q_k' a_1 a_1' q_k \sigma^2 \left[(1 + R_f^2)\phi - R_f - R_f \, \phi^2 \right] \\ &= q_k' \Gamma \delta \delta' \Gamma q_k c_0 \sigma^2 \left[(1 + R_f^2)\phi - R_f - R_f \, \phi^2 \right]. \end{aligned} \tag{10}$$

Kozak et al. (2018, equation (16)) characterize the association between the principal components and δ by expressing δ as a linear combination of the principal components, $\delta = Q\beta$. With this mapping together with the eigenvalue decomposition of the covariance matrix, the term $q_k' \Gamma \delta \delta' \Gamma q_k$ in equation (10) becomes

$$q_k' \Gamma \delta \delta' \Gamma q_k = q_k' Q \Lambda \beta \beta' \Lambda Q' q_k = \iota_k' \Lambda \beta \beta' \Lambda \iota_k = \lambda_k^2 \beta_k^2, \tag{11}$$

where ι_k is a vector of zeros with one as the kth element. The autocovariance of the kth principal component is therefore

$$\mathrm{cov}(PC_t^k, PC_{t+1}^k) = \lambda_k^2 \beta_k^2 c_0 \left[(1 + R_f^2)\phi - R_f - R_f \, \phi^2 \right]. \tag{12}$$

When are factors serially correlated? The bracketed expression in equation (12) determines the sign of the autocovariance. This expression is quadratic and concave in ϕ with two roots: $\phi = \frac{1}{R_f}$ and $\phi = R_f$; factors therefore positively correlate when sentiment is sufficiently persistent, $\phi \in (\frac{1}{R_f}, 1]$. The persistence in sentiment drives the momentum in factors for the same reason as factor premiums align with covariances in Kozak et al. (2018): although arbitrageurs are aware that factors exhibit either reversals (when $\phi < \frac{1}{R_f}$) or momentum (when $\phi > \frac{1}{R_f}$), they are reluctant to trade so aggressively that they would neutralize this pattern because, by doing so, they would assume factor risk. Autocorrelation in factor returns emerges from the connection between sentiment and prices. If sentiment is high today, so are prices. But mean reversion in sentiment would mean that both sentiment and prices are lower tomorrow. The extent to which sentiment autocorrelates therefore pins down the dynamics of factor returns.

In this model sentiment has to be highly correlated to generate factor *momentum*. With an average monthly risk-free rate of 0.39% between July 1965 through and December 2018, the momentum threshold is $\phi > 0.996$. Is this, then, a reasonable mechanism for driving factor momentum? Perhaps. First, the first-order autocorrelation in Baker and Wurgler (2006) sentiment index over the same 1965–2018 period is 0.986, and the Dickey and Fuller (1979) test does not reject the null hypothesis of an unit root at the 10% level.[11] By extension, we also cannot reject the null hypothesis that ϕ is above the critical threshold for factor momentum. Moreover, if Baker and Wurgler (2006) index measures sentiment, it does so with noise; the latent sentiment index could be highly persistent. Second, the Kozak et al. (2018) model is a stylized model for tractability; the risk-free rate, the sentiment index, and the effect of the sentiment on stock returns, for example, are all exogenous, and cash flows are IID with a fixed covariance matrix. The model's qualitative prediction—that persistence in sentiment can generate factor momentum—can be true even if it were to miss the mark on quantities. Factors *are* positively autocorrelated in the data, which implies that if a model in the spirit of Kozak et al. (2018) generates those data, sentiment must be sufficiently autocorrelated to clear the hurdle in such a generalized model.

What factors have more momentum in the Kozak et al. (2018) model? Equation (12) shows that those high-eigenvalue factors that line up with δ have more momentum. This result again parallels the distortion result in Kozak et al. (2018): sentiment-driven demand component δ has a

[11]The Dickey-Fuller test statistic with 641 months of data is -2.36. The 10% critical z-value to reject the null hypothesis of a unit root is -2.57.

large impact on SDF variance only when δ lines up "primarily with the high-eigenvalue (volatile) PCs of asset returns" (p. 1203). Our analysis suggests that the high-eigenvalue factors are also those that should display more factor momentum.

3.2 High-variance principal components and factor momentum

The prediction that momentum should concentrate into high-eigenvalue factors transcends the specifics of the sentiment model. Both Kozak et al. (2018) and Haddad et al.'s (2020), for example, assume the absence of near-arbitrage opportunities to motivate their study of the extent to which low-order PCs explain unconditional differences in expected returns and generate time-series predictability.

We use data on 54 factors from Kozak et al. (2020) to measure factor momentum's concentration into high-eigenvalue principal components.[12] We exclude the seven predictors that relate to momentum or that combine momentum with other characteristics.[13] Similar to Kozak et al. (2020), we exclude all-but-microcaps from analysis to ensure that the very small and illiquid stocks do not unduly influence the results.[14] The characteristics are expressed as weights on zero-investment long-short factors. Each firm characteristic $c_{i,t}$, where i indexes firms, is first transformed into a cross-sectional rank, $rc_{i,t} = \frac{\text{rank}(c_{i,t})}{n_t+1}$, where n_t is the number of stocks in month t. These ranks are then centered around zero and normalized by the sum of absolute deviations from the mean,

$$w_{i,t} = \frac{rc_{i,t} - \overline{rc}_{i,t}}{\sum_{i=1}^{n_t} |rc_{i,t} - \overline{rc}_{i,t}|}. \tag{13}$$

If a firm's characteristic $c_{i,t}$ is missing, we set the weight corresponding to this characteristic to zero (Kozak et al., 2020). Month t return on a factor based on characteristic j is then $f_t =$

[12]We thank Serhiy Kozak for making these data available at https://www.serhiykozak.com/data.

[13]The characteristics we exclude from the original list are (1) momentum (6m), (2) industry momentum, (3) value-momentum, (4) value-momentum-profitability, (5) momentum (1 year), (6) momentum-reversal, and (7) industry momentum-reversal.

[14]Following Kozak et al. (2020) we compute the total market value of all common stocks traded on NYSE, Amex, and Nasdaq in month t and exclude stocks with market values less than 0.01% of the total market value.

$\sum_{i=1}^{n_{t-1}} w_{i,t-1} r_{i,t}$. Table A2 lists the 47 characteristics and the annualized CAPM alphas for long-short factors based on these characteristics. The factors are not re-signed based on the direction into which each characteristic predicts returns; the size factor, for example, is long large stocks and short small stocks and therefore earns a negative average return.

Table 5 reports on the profitability of factor momentum strategies that trade PC factors extracted from these 47 factors. To avoid a lookahead bias, we compute month $t+1$ returns on PC factors using only information that is available as of the end of month t. Our out-of-sample procedure consists of five steps:

1. Compute eigenvectors using daily returns on the 47 factors from July 1973 through the end of month t from the correlation matrix of factor returns.

2. Compute monthly returns for the PC factors up to month $t+1$ using these eigenvectors. PC factor f's return is $r_{f,t}^{pc} = \sum_{j=1}^{47} v_j^f r_{j,t}$, where v_j^f is the jth element of the fth eigenvector and $r_{j,t}$ is the return on individual factor j.

3. Compute individual factors' variances using data up to month t. Demean and lever the PC factors so that their variances up to month t are equal to the variance of the average individual factor and that their average returns up to month t are zero.

4. Construct a factor momentum strategy that is long factors with positive average returns from month $t-11$ to t and short factors with negative average returns.

5. Compute the return on the resulting factor momentum strategy in month $t+1$.

This strategy's return in month $t+1$ is out-of-sample relative to the computation of the eigenvectors in the first step, which uses data only up to the end of month t. Similarly, the demeaning and leveraging in the third step only use information up to the end of month t.[15] When we construct

[15]Goyal and Jegadeesh (2017) and Huang et al. (2020) note that time-series momentum strategies that trade individual assets (or futures contracts) are not as profitable as they might seem because they are net long assets with positive risk premiums. The argument is that, because average returns are positive, a time-series strategy buys more

Table 5: Factor momentum in high- and low-eigenvalue factors

This table reports estimates from time-series regressions in which the dependent variable is the return on factor momentum. We construct factor momentum strategies from the 47 factors listed in Table A2 using either the individual factors or principal component extracted from these factors. We compute the factor PC momentum strategy's month $t+1$ return in five steps: (1) compute eigenvectors from the correlation matrix of daily factor returns from July 1963 up to the end of month t; (2) compute monthly returns for PC factors up to month $t+1$ using these eigenvectors; (3) demean and lever up or down all PC factors so that their average returns up to month t are zero and their time-series variances match that of the average original factor up to month t; (4) take long positions in the PC factors with positive average returns from month $t-11$ to t and short positions in factors with negative average returns; (5) compute the return on the resulting strategy in month $t+1$. This strategy's returns are out-of-sample relative to the computation of the eigenvectors in step (1). We similarly lever individual factor returns so that when we compute month $t+1$ return on the strategy that trades these factors, these factors' variances up to month t are all equal to the average factor's variance up to month t. Panel A reports monthly average returns and t-values for momentum strategies that trade subsets of PC factors ordered by eigenvalues. Panels B and C reports estimates from regressions that explain the returns of momentum strategies with each other. The two intercepts correspond to the first and second halves of the sample. The sample begins in July 1973 and ends in December 2019. The first half runs from July 1973 through September 1996 and the second half from October 1996 through December 2019.

Panel A: Factor momentum in subsets of PC factors ordered by eigenvalues

Set of	Full sample		First half		Second half	
PCs	$\bar{r}$	$t(\bar{r})$	$\bar{r}$	$t(\bar{r})$	$\bar{r}$	$t(\bar{r})$
1–10	0.19	7.07	0.27	8.49	0.11	2.60
11–20	0.13	5.23	0.20	6.13	0.05	1.50
21–30	0.10	5.02	0.18	7.93	0.02	0.63
31–40	0.10	4.05	0.16	5.07	0.04	1.08
41–47	0.07	2.51	0.09	2.71	0.06	1.17

Panel B: Explaining factor momentum in low-eigenvalue PC factors

Explanatory	Set of PCs			
variable	11–20	21–30	31–40	41–47
$\alpha_{\text{first half}}$	0.12	0.12	0.06	−0.01
	(3.50)	(4.27)	(1.86)	(−0.31)
$\alpha_{\text{second half}}$	0.02	0.00	0.00	0.03
	(0.56)	(−0.16)	(−0.10)	(0.78)
$\text{FMOM}_{\text{PC1–10}}$	0.34	0.28	0.34	0.43
	(9.78)	(9.50)	(9.50)	(10.64)
FF5	Y	Y	Y	Y
N	558	558	558	558
Adj. R^2	20.8%	21.7%	20.3%	22.4%

Panel C: Explaining factor momentum in high-eigenvalue PC factors

Explanatory variable	Regression (1)	(2)	(3)	(4)	(5)
$\alpha_{\text{first half}}$	0.17	0.16	0.20	0.22	0.10
	(4.63)	(4.36)	(5.38)	(6.19)	(2.92)
$\alpha_{\text{second half}}$	0.08	0.09	0.09	0.08	0.06
	(2.30)	(2.59)	(2.57)	(2.16)	(1.94)
$\text{FMOM}_{\text{PC11–20}}$	0.43				0.26
	(9.78)				(6.26)
$\text{FMOM}_{\text{PC21–30}}$		0.51			0.29
		(9.50)			(5.83)
$\text{FMOM}_{\text{PC31–40}}$			0.41		0.20
			(9.50)		(4.65)
$\text{FMOM}_{\text{PC41–47}}$				0.39	0.21
				(10.64)	(5.66)
FF5	Y	Y	Y	Y	Y
N	558	558	558	558	558
Adj. R^2	24.6%	24.0%	24.0%	26.6%	40.5%

time-series factor momentum strategies using the original factors, we similarly scale all factors to have the same volatility up to the end of month t so that they are comparable with the PC factors. Any instability in factor rotations does not affect the momentum signal. In the procedure above we compute month $t+1$ return and the average return from month $t-11$ to t using the same time-t eigenvectors. That is, even if the rotation of the factors changes over, say, a six-month period, this instability does not matter because we fix the rotation each month before we look one year backward and one month forward in time. We use daily factor returns starting in July 1963 to compute the eigenvectors; we require at least ten years of data to extract the principal components. The returns on the factor momentum strategies therefore begin in July 1973.[16]

often than it sells. We compute the PC eigenvectors from the correlation matrix which is equivalent to computing PCs from the covariance matrix of demeaned factors. All our PC factors earn zero mean returns *in-sample*. The momentum strategy that trades these PC factors is therefore a pure bet on autocorrelations and not subject to the Goyal and Jegadeesh (2017) bias.

[16]The PC factors are quite stable. A comparison between three alternative sets of PC factors illustrates this

Panel A of Table 5 shows average returns and t-values for momentum strategies that trade subsets of PC factors ordered by eigenvalues. Over the full sample, the strategy that trades the first ten PC factors earns 20 basis points per month (t-value = 7.33). Because the Kozak et al. (2020) factors have low volatilities, so do the PC factors and, by extension, these momentum strategies. The average returns and t-values are lower among the low-eigenvalue PC factors. A momentum strategy trading the last set of PC factors, for example, earns 7 basis points per month (t-value = 2.61). Factor momentum strategies are less profitable in the second half of the sample; although the strategy that trades the first ten PC factors is statistically significant at the 1% level in the second half, the others are not significant even at the 10% level.[17]

In Panels B and C we measure the extent to which the momentums found in the five subsets of factors correlate with and subsume each other. In Panel B we report estimates from regressions such as

$$\mathrm{FMOM}_t^{\mathrm{PC11-20}} = \alpha_{\text{first half}} \mathbb{1}_{t \in \text{first half}} + \alpha_{\text{second half}} \mathbb{1}_{t \in \text{second half}} + b\, \mathrm{FMOM}_t^{\mathrm{PC1-10}} + \mathrm{FF5} + \varepsilon_t^{\mathrm{PC11-20}}, \quad (14)$$

where the two alphas represent the $\mathrm{FMOM}^{\mathrm{PC11-20}}$ strategy's incremental returns over the $\mathrm{FMOM}^{\mathrm{PC1-10}}$ strategy in the first and second halves of the sample and FF5 denote the five factors of the Fama-French model. In Panel C we reverse this regression to explain the momentum present in the first ten PC factors with that found in the other (or all) subsets of PC factors.

The momentum strategies significantly correlate with each other; the t-values for the slope estimates $\hat{b}$ are close to ten. This correlation is noteworthy because the individual PC factors are,

stability. We first compute returns on PC factors in month t from a covariance matrix estimated using daily data (1) up to the end of month $t-1$, (2) up to the end of month $t-2$, or (3) over the full sample period. We then compute average returns for the first ten PC factors for each of these alternative definitions. The correlation between the first two definitions (timely and not-so-timely PC factors) is 0.977 because the covariance matrix rarely changes that much from one month to the next. The correlation between the first and third definitions (timely and full-sample PC factors) is 0.810. That is, factor PC returns computed using the full-sample covariance matrix are in considerable agreement with those extracted using information available in real time.

[17]This deterioration in the momentum profits parallels that experience by the individual stock momentum. UMD's average return in the first half is 81 basis points per month (t-value = 4.00); in the second half, it is 38 basis points (t-value = 1.21).

by definition, orthogonal to each other in-sample (and approximately orthogonal out-of-sample). These positive correlations indicate that all sets of PC factors display *momentum* in a synchronized way: they all tend to be profitable or unprofitable at the same time. The $\alpha_{\text{first half}}$ intercepts in Panel B show that, during the first half of the sample, the strategy that trades the first ten PC factors spans the last two strategies but not the strategies that trade PC factors 11–20 and 21–30. This slow decay of alphas stands in contrast to Kozak et al.'s (2018) who find that a model with a small number of low-order PC factors does well in explaining the expected returns on anomaly portfolios. The significant alphas in Panel B suggest that, during the first half of the sample, a large number of PCs is required to capture all momentum profits.

One possible explanation for this slow alpha decay relates to arbitrage activity. Arbitrageurs might have learned more about momentum (and how to harvest it) over time. The result that a small number of low-order PCs suffice to characterize the stochastic discount factor relies on the assumption of the absence of near-arbitrage opportunities. If such opportunities were more plentiful in the past (because arbitrageurs did not know about them), but have grown scarce henceforth, we would expect alphas to decay faster later in the sample. Consistent with this argument, the momentum in the first ten PC factors subsumes momentums found in all other sets of PC factors during the second half of the sample. Panel C shows that these spanning results do not work both ways. Momentum in the first ten PC factors is informative about the cross section of returns both in the first and second halves when we control for any or all other momentum strategies.

The result that PC factors—and, in particular, the high-eigenvalue factors—exhibit more momentum suggests that momentum is intertwined with the covariance structure of returns. In addition to being consistent with Kozak et al.'s (2018) model of sentiment investors and, more generally, the absence of near-arbitrage opportunities, it is also consistent with Haddad et al.'s (2020) empirical finding that it is the high-eigenvalue PC factors that predictable using the value spreads of Cohen et al. (2003).

The finding that more systematic factors are more autocorrelated is specific neither to the Kozak et al. (2020) factors or the use of the principal-components methodology. Appendix C shows that the same result is true also for the 14 U.S. factors from Table 1. Factors based on characteristics that explain more of the cross-sectional variation are also the ones more predictable by their own past returns. Size, market beta, idiosyncratic volatility, and quality-minus-junk, for example, all are among the most predictable; at the same time, the characteristics underneath these factors explain more of the cross-sectional variation in returns.

4 Factor Momentum and Individual Stock Momentum

4.1 Transmission of factor momentum into the cross section of stock returns

If stock returns obey a factor structure, factor momentum transmits into the cross section of stock returns in the form of *cross-sectional* stock momentum of Jegadeesh and Titman (1993). In multifactor models of asset returns, such as the Intertemporal CAPM of Merton (1973) and the Arbitrage Pricing Theory of Ross (1976), multiple sources of risk determine expected returns. Consider a factor model in which asset excess returns obey an F-factor structure,

$$R_{i,t} = \sum_{f=1}^{F} \beta_i^f r_t^f + \varepsilon_{i,t}, \tag{15}$$

where $R_{i,t}$ is stock i's excess return, r_t^f is the return on factor f, β_i^f is stock i's beta on factor f, and $\varepsilon_{i,t}$ is the stock-specific return component. We assume that the factors do not exhibit any lead-lag relationships with the stock-specific return components, that is, $\mathrm{E}[r_{t'}^f \varepsilon_{i,t}] = 0$.

We now assume that asset prices evolve according to equation (15) and examine the payoffs to a cross-sectional momentum strategy; this strategy, as before, chooses weights that are proportional to stocks' performance relative to the cross-sectional average. The expected payoff to the position in stock i is

$$\mathrm{E}[\pi_{i,t}^{\mathrm{mom}}] = \mathrm{E}\big[(R_{i,-t} - \bar{R}_{-t})(R_{i,t} - \bar{R}_t)\big], \tag{16}$$

where $\bar{R}$ is the return on an equal-weighted index. Under the return process of equation (15), this expected profit becomes

$$\begin{aligned}\mathrm{E}[\pi_{i,t}^{\mathrm{mom}}] \;=\;& \sum_{f=1}^{F}\big[\mathrm{cov}(r_{-t}^{f}, r_t^{f})\,(\beta_i^f - \bar{\beta}^f)^2\big] + \sum_{f=1}^{F}\sum_{g\neq f}^{F}\big[\mathrm{cov}(r_{-t}^{f}, r_t^{g})\,(\beta_i^g - \bar{\beta}^g)\,(\beta_i^f - \bar{\beta}^f)\big] \\ & + \mathrm{cov}(\varepsilon_{i,-t}, \varepsilon_{i,t}) + (\eta_i - \bar{\eta})^2,\end{aligned} \tag{17}$$

where η_i is stock i's unconditional expected return. The expectation of equation (17) over the cross section of N stocks gives the expected return on the cross-sectional momentum strategy,

$$\begin{aligned}\mathrm{E}[\pi_t^{\mathrm{mom}}] \;=\;& \underbrace{\sum_{f=1}^{F}\big[\mathrm{cov}(r_{-t}^{f}, r_t^{f})\,\sigma_{\beta^f}^2\big]}_{\text{factor autocovariances}} + \underbrace{\sum_{f=1}^{F}\sum_{g\neq f}^{F}\big[\mathrm{cov}(r_{-t}^{f}, r_t^{g})\,\mathrm{cov}(\beta^f, \beta^g)\big]}_{\text{factor cross-serial covariances}} \\ & + \underbrace{\frac{1}{N}\sum_{i=1}^{N}\big[\mathrm{cov}(\varepsilon_{i,-t}, \varepsilon_{i,t})\big]}_{\substack{\text{autocovariances} \\ \text{in residuals}}} + \underbrace{\sigma_\eta^2}_{\substack{\text{variation in} \\ \text{mean returns}}},\end{aligned} \tag{18}$$

where N is the number of stocks and $\sigma_{\beta^f}^2$ and σ_η^2 are the cross-sectional variances of the portfolio loadings and stocks' unconditional expected returns.[18]

Equation (18) shows that the profits of the cross-sectional stock momentum strategy can emanate from four sources:

1. Positive autocorrelation in factor returns induces momentum profits through the first term. Cross-sectional variation in betas amplifies this effect. The amount of momentum in the cross section of stocks depends on the number of factors and how autocorrelated the typical factor

[18] Equation (18) does not assume that there are no arbitrage opportunities. If there are no arbitrage opportunities, then the firm-specific component $\varepsilon_{i,t}$ is mean zero and the last term in the decomposition, σ_η^2, represents variation in stocks' *risk premiums*. If there are arbitrage opportunities, this term also picks up cross-sectional variation in mispricings.

is. Even if the typical factor is only weakly autocorrelated, the cross section of stocks will display a lot of momentum if the number of factors is large.[19]

2. The lead-lag return relationships between factors could also contribute to stock momentum profits. The strength of this effect depends both on the cross-serial covariance in factor returns and the covariances between factor loadings. This condition is restrictive: the cross-serial covariances of returns and the covariances of betas have to have the same signs. It would need to be, for example, that, first, SMB's return today positively predicts HML's return tomorrow and, second, that stocks' SMB and HML loadings positively correlate.

3. Autocorrelation in firm-specific returns can also add to the profits of the cross-sectional momentum strategy.

4. The cross-sectional variation in mean returns of individual securities contributes to momentum profits through the Conrad and Kaul (1998) mechanism.

4.2 Pricing momentum-sorted portfolios with momentum and factor momentum

Does factor momentum contribute to the returns of cross-sectional momentum strategies? In Table 6 we examine the connection between individual stock and factor momentum. In Panel A we compare the performance of four asset pricing models in pricing portfolios sorted by prior one-year returns skipping a month. This sorting variable is the same as that used to construct Carhart's (1997) UMD factor. The first model is the Fama-French five-factor model; the second model is this model augmented with the UMD factor; and the third and fourth models augment the five-factor model with factor momentum (FMOM) constructed either from the 20 factors listed in Table 1 or from the ten high-eigenvalue PC factors from Table 5. The factor momentum strategies are long

[19] In Section 4.5.1 we set up simulations in which all momentum emanates from factor autocorrelations. We build on these simulations in Appendix D.3 to demonstrate how the cross section of stocks aggregates the autocorrelations found across multiple factors.

factors with positive returns over the prior year and short those with negative returns.[20] We report alphas for the deciles and the factor loadings against UMD and FMOM.

Stock momentum is evident in the alphas of the Fama-French five-factor model. The alphas for the loser and winner portfolios are -0.75% and 0.57% per month (t-values = -4.05 and 4.82). The average absolute alpha across the deciles is 26 basis points. We significantly improve the model's ability to price these portfolios by adding UMD. The average absolute monthly alpha falls to 12 basis points, and the return on the long-short portfolio falls from 1.33% to 0.27%. Yet, the alpha associated with the long-short portfolio is statistically significant with a t-value of 2.43.

The model augmented with factor momentum constructed from the 20 individual factors performs just as well as—or even better than—the Carhart (1997) six-factor model. The average absolute alpha falls to 11 basis points per month; the Gibbons et al. (1989) test statistic falls from 3.10 to 2.33; and the alpha of the high-minus low portfolio falls from 0.27% to 0.20% (t-value = 0.99). Similar to the Carhart (1997) model, the estimated slopes against factor momentum increase monotonically from bottom decile's -2.46 to top decile's 1.42.

A momentum strategy that trades the first ten PC factors performs just as well in pricing the momentum-sorted portfolios. The absolute pricing error is 9 basis points per month and GRS-test does not reject the null that the alphas across the ten test portfolios are zero.[21] The fact that the five-factor model augmented with factor momentum performs as well as (or better than) the Carhart six-factor model is surprising. The Carhart model sets a high bar because both the factor and the test assets sort on the same variable; that is, UMD targets momentum as directly as, say,

[20] The first term in equation (18), which links cross-sectional momentum to factor momentum, multiplies factor autocovariances with cross-sectional dispersion in betas. If there is no dispersion in betas, factor autocorrelation cannot transmit into the cross section. In the data the differences in beta dispersions are not large enough for this effect to matter, perhaps because each factor is defined using cross-sectional spread in characteristics or, in the case of the liquidity factor, cross-sectional variation in estimated betas. A factor momentum strategy that gives factors weights proportional to the cross-sectional variances of their betas earns an average return of 0.31% (t-value = 7.03) from July 1965 through December 2019; the unweighted strategy earns an average return of 0.34% (t-value = 6.86) over this period. In this computation we estimate betas for individual stocks from univariate regressions using five years of monthly data up to month t, requiring a minimum of two years of data, and compute month $t+1$ returns using this information. The correlation between the weighted and unweighted strategies is 0.95.

[21] The sample period in the last column starts in July 1973 because it uses the PC factors. Over this sample period the winners-minus-losers portfolio's five-factor model alpha is 1.14% (t-value = 3.72).

Table 6: Pricing momentum-sorted portfolios with momentum and factor momentum

Panel A compares the performance of four asset pricing models in explaining the monthly excess returns on ten portfolios sorted by prior one-year returns skipping a month, $r_{t-12,t-2}$: (1) the Fama-French five-factor model; (2) the five-factor model augmented with Carhart's (1997) UMD factor; and (3) and (4) the five-factor model augmented with factor momentum. Factor momentum is long the factors with positive prior one-year returns and short those with negative returns. $\text{FMOM}_{\text{ind.}}$ uses the 20 factors listed in Table 1; $\text{FMOM}_{\text{PC1–10}}$ uses the first 10 PC factors from Table 5. We report alphas for these models and the loadings against the UMD and FMOM factors. Panel B reports estimates from regressions of Carhart's (1997) UMD factor against the five-factor model (first row) and this model augmented with a momentum strategy that trades different subsets of PC factors (other rows). The sample in Panel A, except for the last column, begins in July 1964 and ends in December 2019. The sample in Panel A's last column and Panel B begins in July 1973.

	Asset pricing model						
	FF5	FF5 + UMD		FF5 + $\text{FMOM}_{\text{ind.}}$		FF5 + $\text{FMOM}_{\text{PC1–10}}$	
Decile	$\hat{\alpha}$	$\hat{\alpha}$	$\hat{b}_{\text{umd}}$	$\hat{\alpha}$	$\hat{b}_{\text{fmom}}$	$\hat{\alpha}$	$\hat{b}_{\text{fmom}}$
Losers	−0.75	−0.10	−0.93	−0.04	−2.46	−0.02	−3.65
	(−4.05)	(−0.94)	(−36.59)	(−0.28)	(−20.06)	(−0.09)	(−12.95)
2	−0.35	0.13	−0.70	0.16	−1.78	0.18	−2.66
	(−2.74)	(2.08)	(−46.76)	(1.54)	(−21.26)	(1.36)	(−14.05)
3	−0.20	0.18	−0.54	0.17	−1.30	0.20	−2.06
	(−1.90)	(2.92)	(−38.35)	(1.93)	(−17.78)	(1.83)	(−13.16)
4	−0.16	0.07	−0.33	0.12	−0.95	0.16	−1.43
	(−1.93)	(1.20)	(−22.77)	(1.69)	(−16.70)	(1.93)	(−11.67)
5	−0.16	−0.04	−0.17	−0.02	−0.47	0.02	−0.78
	(−2.45)	(−0.65)	(−12.30)	(−0.39)	(−9.07)	(0.22)	(−7.49)
6	−0.13	−0.09	−0.05	−0.07	−0.22	−0.05	−0.41
	(−2.05)	(−1.46)	(−3.52)	(−1.02)	(−4.26)	(−0.65)	(−3.90)
7	−0.12	−0.16	0.07	−0.14	0.09	−0.12	0.06
	(−1.94)	(−2.72)	(4.73)	(−2.32)	(1.83)	(−1.70)	(0.58)
8	0.04	−0.11	0.22	−0.09	0.44	−0.10	0.69
	(0.62)	(−2.05)	(16.96)	(−1.34)	(8.42)	(−1.44)	(6.68)
9	0.08	−0.14	0.33	−0.11	0.66	−0.14	0.94
	(1.08)	(−2.46)	(23.85)	(−1.45)	(11.04)	(−1.61)	(7.56)
Winners	0.57	0.17	0.57	0.16	1.42	0.01	2.38
	(4.82)	(2.32)	(32.93)	(1.60)	(17.21)	(0.05)	(14.26)
Winners	1.33	0.27	1.51	0.20	3.88	0.02	6.03
− Losers	(4.91)	(2.43)	(56.81)	(0.99)	(23.13)	(0.09)	(15.84)
N	666	666		666		558	
Avg. $\lvert\hat{\alpha}\rvert$	0.26	0.12		0.11		0.10	
GRS F-value	4.24	3.10		2.33		1.30	
GRS p-value	0.00%	0.04%		1.06%		20.29%	

Panel B: Pricing UMD with momentum in subsets of PC factors

Subset of PCs	Alpha $\hat{\alpha}$	Alpha $t(\hat{\alpha})$	Factor momentum $\hat{b}_{\text{fmom}}$	Factor momentum $t(\hat{b}_{\text{fmom}})$	FF5	R^2
None	0.62	3.36	.	.	Y	10.7%
1–10	−0.09	−0.59	3.90	17.50	Y	42.5%
11-20	0.35	1.95	2.05	6.89	Y	17.6%
21–30	0.28	1.60	3.14	9.07	Y	22.1%
31–40	0.34	2.01	3.03	10.91	Y	26.4%
41–47	0.40	2.33	2.56	10.54	Y	25.5%

HML targets portfolios sorted by book-to-market.

In Panel B we report estimates from time-series regressions in which Carhart's (1997) momentum factor, UMD, is the dependent variable. The model on the first row is the Fama-French five-factor model and the other rows augment this model with strategies that trade momentum in different subsets of the PC factors. Momentum found in the high-eigenvalue factors explains all of UMD's profits; the alpha is −6 basis points. Strategies based on the other subsets perform worse, leaving UMD with substantial alphas that are, except for the 21–30 set, statistically significant at the 5% level. Across the five models, momentum in the high-eigenvalue factors also explains the most of the time-series variation in UMD's return at $R^2 = 43\%$.

4.3 Alternative momentum factors: Spanning tests

In Table 7 we show that, in addition to the "standard" individual stock momentum of Jegadeesh and Titman (1993), factor momentum also subsumes other cross-sectional momentum strategies. We construct three other momentum factors using the UMD methodology: Industry-adjusted momentum of Cohen and Polk (1998) sorts stocks' by their industry-adjusted returns; intermediate momentum of Novy-Marx (2012) sorts stocks by their returns from month $t-12$ to $t-7$; and Sharpe ratio momentum of Rachev et al. (2007) sorts stocks by the returns scaled by the volatility of returns. We also construct the industry momentum strategy of Moskowitz and Grinblatt (1999). This strategy sorts 20 industries based on their prior six-month returns and takes long and short positions in the top and bottom three industries.

Panel A of Table 7 introduces these alternative momentum factors alongside the two factor momentum strategies: momentum in the 20 individual factors and that in the first ten PC factors. All momentum factors earn statistically significant average returns and Fama-French five-factor model alphas. Although the average returns associated with the two factor momentum strategies are the lowest, they are also the least volatile. Their Sharpe and information ratios, which are proportional to the t-values associated with the average returns and five-factor model alphas, are the highest among all the momentum strategies.

Panel B shows estimates from spanning regressions in which the dependent variable is one of the individual stock momentum factors. The model is the Fama-French five factor model augmented with one of the factor momentum strategies. These regressions have two interpretations. From an investment perspective, a statistically significant alpha implies that an investor would have earned a higher Sharpe ratio by having traded the left-hand side factor in addition to the right-hand side factors (Huberman and Kandel, 1987). From an asset pricing perspective, a statistically significant alpha implies that the asset pricing model that only contains the right-hand side factors is dominated by a model that also contains the left-hand side factor (Barillas and Shanken, 2017). Although all definitions of momentum earn statistically significant average returns and five-factor model alphas, both factor momentums span all of them. The maximum t-value across the ten specifications is industry-adjusted momentum's 1.67 when explained with the momentum in individual factors.[22]

Panel C shows that none of the alternative definitions of individual stock momentum span either version of factor momentum. Across the 12 specifications, the lowest t-value for factor momentum's alpha is 4.07. The last row augments the Fama-French five-factor model with all five individual

[22] We exclude the market factor from the set of individual factors when constructing the factor momentum strategy that trades individual factors. The results here and elsewhere are not sensitive to the decision to exclude or include the market factor. Panel B's alphas, for example, range from −1 basis point (t-value = −0.10) to 13 basis points (t-value = 1.60) when the momentum strategy $\text{FMOM}_{\text{ind.}}$ trades the market factor as well.

Table 7: Alternative definitions of momentum: Spanning tests

Panel A reports monthly average returns and Fama-French five-factor model alphas for alternative momentum factors. Every factor, except for industry momentum, is similar to the UMD factor of Jegadeesh and Titman (1993) ("standard momentum"). We sort stocks into six portfolios by market values of equity and prior performance. A momentum factor's return is the average return on the two high portfolios minus that on the two low portfolios. Industry momentum uses the (Moskowitz and Grinblatt, 1999, Table I) methodology; it is long the top three industries based on their prior six-month returns and short the bottom three industries, with each stock classified into one of 20 industries. Panel A also reports references for the original studies that use these alternative definitions. Panel B reports estimates from regressions in which the dependent variable is the monthly return on one of the individual stock momentum strategies. These regressions augment the five-factor model with momentum found in either the 20 individual factors or the first ten PC factors. Panel C reports estimates from regressions in which the dependent variable is one of the factor momentum strategies. These regressions augment the five-factor model with one of the individual stock momentum factors (UMD*) or, on the last row, with all five individual stock momentum factors. The sample begins in July 1964 and ends in December 2019 except for the regressions that use the PC-based factor momentum, in which the sample begins in July 1973.

Panel A: Factor means and Fama-French five-factor model alphas

		Monthly returns			FF5 model	
Momentum definition	Reference	$\bar{r}$	SD	$t(\bar{r})$	$\hat{\alpha}$	$t(\hat{\alpha})$
	Individual stock momentum					
Standard momentum	Jegadeesh and Titman (1993)	0.64	4.22	3.93	0.70	4.28
Ind.-adjusted momentum	Cohen and Polk (1998)	0.41	2.64	3.96	0.50	4.93
Industry momentum	Moskowitz and Grinblatt (1999)	0.63	4.60	3.54	0.69	3.77
Intermediate momentum	Novy-Marx (2012)	0.48	3.02	4.12	0.56	4.81
Sharpe ratio momentum	Rachev et al. (2007)	0.55	3.59	3.94	0.63	4.51
	Factor momentum					
Momentum in individual factors		0.33	1.20	7.01	0.29	6.21
Momentum in PC factors 1–10		0.19	0.64	7.07	0.18	6.51

stock momentum factors. In these specifications factor momentums' alphas are significant with t-values of 4.30 (individual factors) and 5.45 (PC factors 1–10).

Table 7 indicates that factor momentum contains information not present in any other forms of momentum and yet, at the same time, no other form of momentum is at all informative about the cross section of stock returns when controlling for factor momentum. These spanning results suggest that individual stock momentum is, at least in large part, a manifestation of factor momentum.

Panel B: Regressions of individual stock momentum strategies on factor momentum

Individual stock momentum, UMD*	Factor momentum: Individual factors $\hat{\alpha}$	Individual factors $\text{FMOM}_{\text{ind.}}$	PC factors 1–10 $\hat{\alpha}$	PC factors 1–10 $\text{FMOM}_{\text{PC1–10}}$	FF5
Standard momentum	0.00 (−0.04)	2.43 (24.72)	−0.09 (−0.60)	3.90 (17.52)	Y
Industry-adjusted momentum	0.14 (1.67)	1.23 (17.63)	0.10 (0.99)	1.90 (12.68)	Y
Industry momentum	0.02 (0.12)	2.32 (18.83)	−0.16 (−0.85)	4.10 (15.51)	Y
Intermediate momentum	0.15 (1.51)	1.41 (17.72)	0.17 (1.40)	2.20 (12.64)	Y
Sharpe ratio momentum	0.02 (0.19)	2.12 (25.45)	−0.05 (−0.39)	3.63 (19.74)	Y

Panel C: Regressions of factor momentum on individual stock momentum strategies

Individual stock momentum, UMD*	Dependent variable: Momentum in individual factors $\hat{\alpha}$	Momentum in individual factors UMD*	Momentum in PC factors 1–10 $\hat{\alpha}$	Momentum in PC factors 1–10 UMD*	FF5
Standard momentum	0.15 (4.44)	0.20 (24.72)	0.13 (5.53)	0.09 (17.52)	Y
Industry-adjusted momentum	0.16 (4.07)	0.26 (17.63)	0.13 (5.17)	0.12 (12.68)	Y
Industry momentum	0.19 (4.88)	0.15 (18.83)	0.14 (5.88)	0.07 (15.51)	Y
Intermediate momentum	0.16 (4.15)	0.23 (17.72)	0.13 (4.95)	0.10 (12.64)	Y
Sharpe ratio momentum	0.14 (4.20)	0.23 (25.45)	0.11 (5.17)	0.11 (19.74)	Y
All of above	0.14 (4.30)	.[†]	0.12 (5.32)	.[†]	Y

[†]Note: These regressions include all five individual stock momentum factors on the RHS at the same time: standard momentum, industry-adjusted momentum, industry momentum, intermediate momentum, and Sharpe ratio momentum.

An investor who trades individual stock momentum indirectly times factors; she would do better by timing the factors directly.

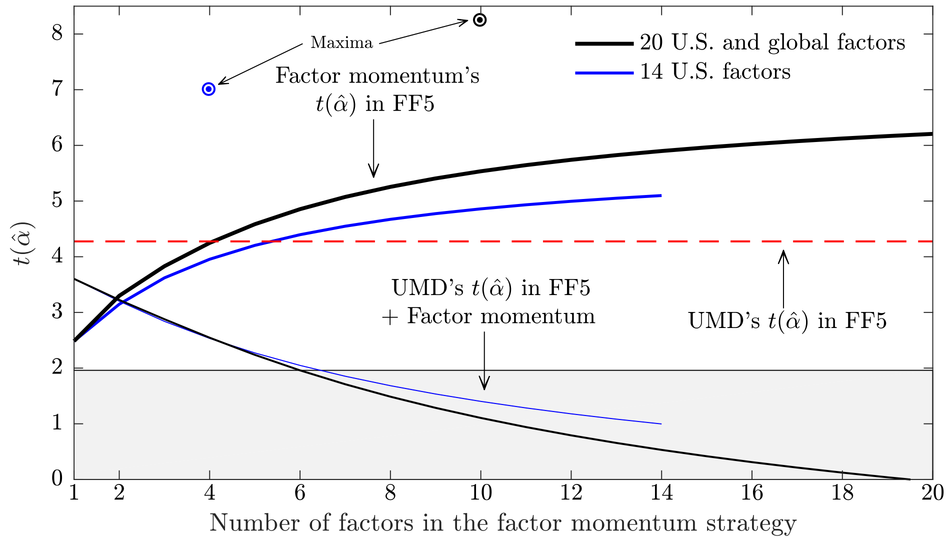

Figure 3: **Individual stock momentum versus factor momentum as a function of the number of factors.** We form all subsets of the 14 U.S. factors (blue lines) or 20 U.S. and global factors (black lines) listed in Table 1 and form time-series factor momentum strategies that trade these factors. A time-series factor momentum strategy is long factors with positive returns over the prior year and short those with negative returns. The thick line represents the factor momentum strategy's average $t(\hat{\alpha})$ from the Fama-French five-factor model regression; the thin line represents UMD's average $t(\hat{\alpha})$ from a regression that augments the five-factor model with the factor momentum strategy; and the dashed line denotes UMD's $t(\hat{\alpha})$ from the Fama-French five-factor model regression. The circles denote the combinations with the highest t-values in the two universes of factors. The shaded region indicates t-values below 1.96.

4.4 Individual stock momentum versus factor momentum with alternative sets of factors

The factor momentum strategy that trades individual factors takes positions in up to 20 factors. Tables 6 and 7 show that this factor momentum explains individual stock momentum. In Figure 3 we measure the extent to which this result is sensitive to the number and identity of the factors included in the version of factor momentum that trades the 20 individual factors.

We first construct all possible combinations of factors, ranging from one factor to the full set of 20 factors. We then construct a factor momentum strategy from each set of factors and estimate two regressions. The first regression is the Fama-French five-factor model with factor momentum as the dependent variable. The dependent variable in the second regression is UMD and the model is the Fama-French five-factor model augmented with factor momentum. We record the t-values associated with the alphas from all possible models, and plot averages of these t-values as a function of the number of factors.[23] The black lines in Figure 3 denote these combinations drawn from the full set of 20 factors. We also construct all possible factor momentum strategies that trade only the 14 U.S. factors. The blue lines in Figure 3 denote these combinations. We also plot, for reference, the t-value associated with UMD's alpha in the five-factor model.

Figure 3 shows that the t-value associated with factor momentum's five-factor model alpha monotonically increases in the number of factors. Consider first strategies drawn from the full set of 20 factors. When factor momentum alternates between long and short positions in just one factor, the average t-value is 2.49; when it trades 10 factors, it is 5.54; and when we reach 20 factors, it is 6.21. At the same time, factor momentum's ability to span UMD improves. The typical one-factor factor momentum strategy leaves UMD with an alpha that is statistically significant with a t-value of 3.60. However, when the number of factors increases to 10, this average t-value has decreased to 1.10; and with all 20 factors, this t-value is -0.04. The patterns are the same when we limit the analysis to the 14 U.S. factors. For example, the average t-value associated with UMD's alpha is 1.40 when we construct factor momentum from 10 U.S. factors.[24]

These estimates suggest that factor momentum's ability to span UMD is not specific to the set of factors used; as the number of factors increases, the autocorrelations found within most sets of factors aggregate to explain individual stock momentum. Figure 3 supports our thesis that individual stock momentum is an aggregation of the autocorrelations found in factor returns; the more factors we identify, the better we capture UMD's return.

[23]The sample begins in July 1964 and ends in December 2019. Because some factors have later start dates, we exclude those factor combinations that would result in a sample that does not span the full 1964–2019 period. There are, for example, $\frac{20!}{(20-6)!6!} = 38{,}760$ six-factor combinations. We exclude those seven combinations that would result in start dates later than July 1964. The total number of one- to twenty-factor combinations is 1,048,575; 1,048,448 of these span the full sample period.

[24]The t-values we report in Figure 3 are averages of various combinations. We could indulge in some data dredging and ask which combinations of factors display the most momentum. Among the 14 U.S. factors, a combination of four factors produces a strategy with a $t(\hat{\alpha})$ of 6.99; in the set of all 20 factors, the highest t-value of 8.24 belongs to a ten-factor strategy. The blue and black circles in Figure 3 denote these maxima. More "powerful" factor momentum strategies than the 20-factor version therefore lurk within this set of factors. We use the all-factor strategy to err on the side of caution; any strategy that uses a subset of all available factors would need to be justified on an ex-ante basis or subjected to tests that address the multiple hypothesis testing problem.

4.5 Do firm-specific returns display momentum?

4.5.1 Simulation evidence

If factor momentum drives *all* momentum in the cross section of stock returns, firm-specific returns should not display any continuation. A natural test would therefore be to measure momentum in firm-specific returns. However, when these returns have to be estimated as residuals from factor models, we encounter three problems: (1) we do not know the identities of all factors, (2) we do not observe true factor returns, and (3) we can only estimate stocks' factor loadings with noise. It is therefore not possible—absent a natural experiment that would allow us to identify true firm-specific returns—to attribute conclusively cross-sectional momentum into effects emanating from "factor momentum" and "residual momentum."

To illustrate the issue arising from omitted factors, suppose that two systematic factors drive excess stock returns:

$$R_{i,t} = \beta_{i,1} F_{1,t} + \beta_{i,2} F_{2,t} + \varepsilon_{i,t}. \tag{19}$$

A researcher who knows only about the first factor then estimates the residual as

$$\hat{\varepsilon}_{i,t} = [r_{i,t} - \beta_{i,1} F_{1,t}] + \beta_{i,2} F_{2,t}, \tag{20}$$

where we assume that the researcher observes the true factor F_1 and stock i's beta against it. If the researcher does not have the full set of factors, *estimated* residuals can display momentum even if firm-specific returns are IID. In terms of equation (20), $\hat{\varepsilon}_{i,t}$ would display momentum if the omitted factor, F_2, displays momentum. What the researcher views as the firm-specific residual is only the residual net of known factors. This problem of *omitted-factor momentum* grows worse if we do not observe true factors or betas.

Table 8 demonstrates the difficulty of disentangling factor momentum from residual momentum. We simulate returns from an economy in which only factor returns are positively autocorrelated.

Table 8: Residual momentum versus factor momentum: Simulations

This table reports average t-values from simulations that assess the strengths of individual stock momentum, residual momentum, and factor momentum in an economy in which only factor returns are positively autocorrelated. We simulate 672 months of returns from a market with 2,000 stocks. Ten systematic factors and IID firm-specific innovations drive stock returns. In "Symmetric factors" specification the ten factors have the same variance, their risk premiums are equally persistent, and stocks' betas against these factors are mean zero. In "Uncorrelated market factor" specification the first factor explains five times as much of the cross section of the stock returns as each of the other nine factors, the first factor's risk premium is uncorrelated, and stocks' betas against the first factor have a mean of one; the betas against the other nine factors are mean zero. Appendix D details these simulations. The individual stock momentum strategy is long the top decile of stocks with the highest average returns over the prior year and short the bottom decile. Residual momentum strategy is long the top decile and short the bottom decile formed by firm-specific residuals over the prior year. We estimate firm-specific residuals from a factor model with $1, 2, \ldots$ or 10 factors from month $t - 72$ to $t - 13$ and use the resulting betas to compute residuals for months $t - 12$ to $t - 1$. The factor momentum strategy is long factors with positive returns over the prior year and short factors with negative returns. Column "Number of known factors" indicates the number of factors used to compute firm-specific residuals and for trading factor momentum. We report average t-values from 10,000 simulations.

	Symmetric factors		Uncorrelated market factor	
Number of known factors	Residual momentum	Factor momentum	Residual momentum	Factor momentum
1	5.65	1.55	5.62	−0.02
2	5.28	2.23	5.21	0.84
3	4.86	2.72	4.82	1.46
4	4.45	3.14	4.38	2.00
5	4.01	3.49	3.94	2.46
6	3.52	3.85	3.42	2.88
7	2.97	4.16	2.91	3.23
8	2.33	4.44	2.26	3.56
9	1.51	4.69	1.42	3.88
10	0.00	4.95	−0.01	4.16
Individual stock momentum	6.01		4.68	

Ten systematic factors and IID firm-specific innovations drive stock returns. Factors' risk premiums follow AR(1) processes, and a factor's return is the sum of its risk premium and an IID innovation. We simulate data under two assumptions about factors. In the "Symmetric factors" specification all ten factors have the same variance and all factors' risk premiums are equally persistent. In the "Uncorrelated market factor" specification the first factor explains five times as much of the cross section of returns as the other nine factors and its risk premium is uncorrelated. Appendix D details these simulations. In the simulations the average non-market factor's autocorrelation is similar to that in the data: the correlation between month t returns and the average returns from month $t-12$ to $t-1$ is 0.25 in the data (see Table A1) and 0.20 in the simulations (see Table A8).

We construct three momentum strategies using the simulated returns. Individual stock momentum is long the top decile of stocks with the highest returns over the prior year and short the bottom decile. Residual momentum strategy is long and short the top and bottom decile of stocks based on their firm-specific residuals over the prior year. We estimate firm-specific residuals from a factor model with $1, 2, \ldots$ or 10 factors from month $t-72$ to $t-13$ and use these beta estimates to compute residuals from month $t-12$ to $t-1$. The factor momentum strategy is long factors with positive returns over the prior year and short factors with negative returns. Column "Number of known factors" in Table 8 indicates the number of factors used to compute firm-specific residuals and for trading factor momentum. We report average t-values from 10,000 simulations.

Stock returns display momentum in these simulations. In the "Symmetric factors" specification, the individual stock momentum strategy is significant with an average t-value of 6.01. Although there is no momentum in stock returns per se, this strategy is profitable because it indirectly bets on the persistence in factor risk premiums: stocks with the best performance over the prior year, on average, load positively on factors with high past returns and negatively on factors with low past returns; stocks with the worst performance have, on average, the opposite loadings. Because there is no momentum in firm-specific residuals, residual momentum strategies are, on average, less profitable than the individual stock momentum strategy. Residual momentum strategies, however, remain statistically significantly profitable as long as the investor's asset pricing model omits two factors; at this point, the residual momentum's strategy average t-value is 2.33. When the investor has the full factor model, the estimated residuals, on average, equal firm-specific innovations and the residual momentum strategy earns no profits.

Factor momentum strategies are profitable because factors positively autocorrelate. The greater the number of factors the investor knows, the more profitable these strategies are. If the investor knows just two factors, the factor momentum strategy's average t-value is 2.23; if the investor knows all ten factors, its t-value is 4.95.[25]

Residual momentum is stronger than individual stock momentum if the total amount of momentum in the known factors is less than that in the omitted factors. The "Uncorrelated market factor" specification illustrates this possibility by assuming that the first factor is more systematic than the other factors and serially uncorrelated. If an investor computes firm-specific residuals from a one-factor model with only the uncorrelated factor, residuals become better measures of the remaining nine autocorrelated factors. Although individual stock momentum strategy's average t-value in these simulations is 4.68, the residual momentum strategy under a one-factor model has a t-value of 5.62.

The residual momentum strategy becomes less profitable as we begin removing the autocorrelated factors. It remains more profitable than the individual stock momentum strategy up to three factors. Although the factor momentum strategy becomes more profitable as the investor adds more factors, this strategy continues to be hurt by including also the market factor. If the investor traded only the nine autocorrelated factors, then the factor momentum strategy would be profitable with an average t-value of 4.69—this case corresponds to the "Symmetric factors" specification with nine known factors.

Table 8 shows that attempts to disentangle factor momentum from residual momentum run into an omitted-variables problem. If the asset pricing model is incomplete, a residual momentum strategy may be profitable by the virtue of trading omitted-factor momentum.

[25] Although all momentum resides in factors, factor momentum is not as profitable as individual stock momentum because it takes long and short positions in all ten factors. This strategy therefore also invests in factors whose past returns (and therefore estimated risk premiums) are close to zero. The individual stock momentum and residual momentum strategies, by the virtue of taking positions only in the top and bottom deciles of stocks, implicitly bet more on the factors with large positive or negative past returns and, therefore, large positive or negative risk premiums.

Table 9: Residual momentum versus factor momentum: Actual data

This table reports average returns and alphas for UMD-style individual stock momentum strategies. The strategy on the first row sorts stocks by their raw returns from month $t-12$ to $t-2$. The strategies on the other rows sort stocks by their estimated residuals based on the CAPM or the Fama-French three- or five-factor model. The independent variable is a factor momentum strategy that trades either the 20 off-the-shelf factors or the first then high-eigenvalue PC factors. The strategy on the first row, which is the same as the standard UMD, trades the same stocks that are included in the residual momentum strategies. t-values are in parentheses. The sample, except for the last column, begins in July 1967 and ends in December 2019. In the last column the sample begins in July 1973.

		Control for factor momentum			
Sorting	Average	Individual factors		PC factors 1–10	
variable	return	$\hat{\alpha}$	$\hat{b}_{\text{fmom}}$	$\hat{\alpha}$	$\hat{b}_{\text{fmom}}$
Raw returns	0.45	−0.19	1.96	−0.29	3.69
	(2.88)	(−1.45)	(19.16)	(−2.04)	(17.18)
CAPM residuals	0.58	0.08	1.53	−0.05	3.08
	(4.29)	(0.67)	(16.68)	(−0.38)	(16.58)
FF3 residuals	0.44	0.15	0.90	0.00	2.09
	(3.83)	(1.35)	(10.27)	(0.04)	(11.92)
FF5 residuals	0.37	0.17	0.63	0.00	1.72
	(3.39)	(1.52)	(7.32)	(−0.03)	(10.13)

4.5.2 Actual data

In Table 9 we examine the profitability of three residual momentum strategies. We compute residuals from the CAPM and the Fama-French three- and five-factor models. We estimate stocks' factor loadings using data from month $t-72$ to $t-13$, requiring a minimum of three years of data, and compute average residuals from month $t-12$ to $t-2$. We then construct UMD-like residual momentum strategies by sorting stocks into six portfolios by size and past return and taking long and short positions in the winner and loser portfolios. For the sake of comparability, we recompute UMD in this table so that it is based on the same universe of stocks as that used for the residual momentum strategies.

The individual stock momentum strategy earns an average return of 45 basis points per month (t-value = 2.88) in this sample. This strategy's alpha turns negative when we control for factor momentum strategy that trades either the 20 individual factors or the first ten high-eigenvalue PC factors.

A momentum strategy based on the CAPM residuals is more profitable than the strategy based on stocks' raw returns. This strategy earns an average return of 58 basis points (t-value = 4.29). This increase is consistent with the notion that the market factor is weakly serially correlated relative to its importance in explaining cross-sectional variation in realized returns; removing this factor renders the residuals *more* informative about the other factors. This residual momentum strategy also correlates significantly with the two factor momentum strategies. Its alpha net of factor momentum in the individual factors is 8 basis points (t-value = 0.67) and −5 basis points net of the momentum in the PC factors.

The momentum strategy based on the three-factor model residuals is less profitable (44 basis points, t-value = 3.83) than that based on CAPM residuals, and the strategy based on the five-factor model residuals is less profitable still (37 basis points, t-value = 3.39). These profits decrease because the additional factors we now expunge—size, value, and so forth—contribute meaningfully to momentum profits. These strategies' alphas are also close to zero when we control for factor momentum. For example, the alpha of the five-factor model residual strategy is zero (t-value = −0.03) when we control for the momentum found in the first ten PC factors.

The estimates in Table 9 are consistent with the momentum in individual stock returns emanating from factors. When we remove a factor that is very systematic relative to its autocorrelation, the momentum strategy becomes more profitable. But once we remove factors that contribute to momentum profits, the strategy becomes increasingly less profitable. The extent to which residuals display momentum net of factor momentum depends on the properties of the factors that remain outside the factor model used to estimate the residuals. The insignificant alphas in Table 9 suggest that, whatever these additional factors beyond the five-factor model are, they are still largely the same as those found within the first ten PC factors.[26]

[26]The absence of residual momentum is also consistent with the predictions of Kozak et al.'s (2018) model. Only

In Appendix B we return to Section 4.1's linear-weight decomposition and reach the same conclusion about the tension between factor momentum, residual momentum, and omitted factor momentum. Under the CAPM the decomposition attributes almost all of the profits to the autocorrelation in firm-specific returns; but in a seven-factor model, it attributes most of the returns to the autocorrelations in factor returns.

An additional observation about the nature of residual momentum strategies is in order. These strategies can appear profitable not only because there is momentum in firm-specific returns, but because these strategies also implicitly bets against betas. To see why, note that a firm's estimated residual return is its return minus the product of the estimated factor loadings and factor returns. Firms with high residual returns have either earned high returns or have low estimated betas; and those with low residuals have either earned low returns or have have high estimated betas. A residual momentum strategy is therefore also, in part, long low-beta stocks and short high-beta stocks. We show in Appendix E that residual momentum strategies indeed make significant bets against betas; that their five-factor model alphas exceed their average returns because of these bets; and that, by controlling for betting-against-beta factors, factor momentum strategies span the three residual momentum strategies.

5 Momentum vis-à-vis other factors

5.1 Unconditional and conditional correlations with the momentum factor

The puzzling feature of individual stock momentum is its low correlations with other factors. Over the July 1963 through December 2019 period, the adjusted R^2 from regressing UMD on the Fama-French five-factor model is just 9%. These estimates might imply that factors *unrelated* to the market, size, value, profitability, and investment factor must explain the remaining 91% of the

those sentiment-driven demand components that align with covariances distort prices; there are no firm-specific distortions and, by extension, no firm-specific momentum.

variation or, alternatively, that momentum is a distinct risk factor.

The *unconditional* correlations between UMD and the other factors, however, significantly understate their associations. Consider, for example, the size factor. If size has performed well, UMD will, by construction, be long small-cap stocks and short large-cap stocks. Because both UMD and SMB are now long small-cap stocks and short large-cap stocks, we expect them to correlate positively the next month. If, on the other hand, size has performed poorly, UMD will be short small-cap stocks and long large-cap stocks. We would therefore expect UMD and SMB to correlate negatively. The same mechanism should hold for all factors: if a factor has performed well, UMD will be long that factor, and UMD and the factor will positively correlate; but if the factor has performed poorly, UMD will be short that factor and the correlation will be negative.

In Table 10 we report factors' correlations with UMD. We report three correlations: unconditional correlation, correlation conditional on the factor's return over the prior year being positive, and correlation conditional on this return being negative. The unconditional correlations between UMD and the factors are low; 11 out of the 20 correlations with the individual factors are positive, and the correlation between UMD and the portfolio of all 20 factors is 0.04. The correlations conditional on past returns, however, are remarkably different. Except for the short-term reversals factor, all factors correlate more with UMD when their past returns are positive.[27] For 17 of these 19 factors, the difference is statistically significant at the 5% level. The first row assigns all factors into two groups based on their past returns. The basket of factors with positive past returns has a correlation of 0.45 with UMD; the basket of factors with negative returns has a correlation of −0.51. Table A3 in the Appendix shows that both the long and short legs contribute to this pattern: the short leg drives the positive correlation following a year of positive returns and the long leg drives the negative correlation following a year of negative returns.

Because the unconditional correlations between momentum and the other factors are close to

[27] The short-term reversals factor has almost 100% turnover per month (Novy-Marx and Velikov, 2016). Any association between past factor returns and current holdings therefore breaks down.

Table 10: Unconditional and conditional correlations with the momentum factor

This table reports correlations between UMD and factor returns: ρ is UMD's unconditional correlation with the factor, ρ^+ is the correlation conditional on the factor's return over the prior year being positive, and ρ^- is the correlation conditional on the prior-year return being negative. The first row takes the average of all 20 factors or averages of factors with positive or negative returns over the prior year. The z-value in the last column is from a test that the conditional correlations are equal. This test uses Fisher's (1915) z-transformation, $1 \Big/ \sqrt{\frac{1}{N^+-3} + \frac{1}{N^--3}} \left(\tanh^{-1}(\hat{\rho}^+) - \tanh^{-1}(\hat{\rho}^-)\right) \sim N(0,1)$, where $\tanh^{-1}(x) = \frac{1}{2}\frac{\ln(1+x)}{\ln(1-x)}$ and N^+ and N^- are the number of observations used to estimate ρ^+ and ρ^-.

	Unconditional correlation	Conditional correlations		H_0: $\hat{\rho}^+ = \hat{\rho}^-$
Factor	$\hat{\rho}$	$\hat{\rho}^+$	$\hat{\rho}^-$	z-value
Pooled	0.04	0.45	−0.51	18.37
			U.S. factors	
Size	−0.04	0.16	−0.39	7.20
Value	−0.20	0.17	−0.58	10.45
Profitability	0.11	0.46	−0.41	11.22
Investment	−0.03	0.19	−0.37	7.13
Accruals	0.13	0.30	−0.15	5.46
Betting against beta	0.18	0.41	−0.22	6.70
Cash-flow to price	−0.13	0.23	−0.59	11.38
Earnings to price	−0.17	0.20	−0.61	11.50
Liquidity	−0.03	0.03	−0.14	2.15
Long-term reversals	−0.09	0.10	−0.43	7.02
Net share issues	0.11	0.36	−0.42	10.44
Quality minus junk	0.28	0.46	−0.41	11.00
Residual variance	0.21	0.67	−0.56	18.44
Short-term reversals	−0.30	−0.39	−0.19	−2.28
			Global factors	
Size	0.07	0.09	0.05	0.35
Value	−0.16	0.15	−0.48	5.81
Profitability	0.27	0.33	−0.02	2.60
Investment	0.06	0.40	−0.43	7.99
Betting against beta	0.22	0.24	0.15	0.73
Quality minus junk	0.42	0.48	−0.17	4.87

zero, most factor models, such as the five-factor model, explain none of momentum profits. This result, however, does not imply that momentum is "unrelated" to the other factors. Table 10 shows that the unconditional correlations are close to zero only because these correlations are significantly time-varying. Momentum, in fact, appears to relate to *all* factors; it is just that momentum switches between being long and short other factors, thereby producing unconditional correlations close to zero. This argument of time-varying loadings also suggests a solution to the puzzle that Cochrane (2011, p. 1075) poses when discussing a behavioral explanation for momentum:

> "For example, "extrapolation" generates the slight autocorrelation in returns that lies behind momentum. But why should all the momentum stocks then rise and fall together the next month, just as if they are exposed to a pervasive, systematic risk?"

Momentum stocks indeed comove because of pervasive, systematic risks. Winners, for example, are stocks that positively load on factors that have performed well and negatively on those that have done poorly.[28]

5.2 Momentum in momentum-neutral factors

Does momentum reside in factors or individual stocks? Because factors are portfolios of stocks, factor returns ultimately emanate from individual stock returns. The argument that factor momentum drives individual stock momentum is that the individual stocks that make up the factor are inconsequential: momentum is about the factor loadings (or characteristics) associated with the factors, not about specific companies.

A source of some ambiguity in showing any resemblance of causality is that individual stock momentum may induce incidental momentum into factor returns. If the size factor, for example,

[28]The five-factor model and the 9% adjusted R^2 that it gives to UMD illustrates this issue. If, instead of regressing UMD on the five factors, suppose that we split each factor into two parts: HML_t^{up}, $\text{HML}_t^{\text{down}}$, SMB_t^{up}, $\text{SMB}_t^{\text{down}}$, and so forth, where

$$\text{HML}_t^{\text{up}} = \begin{cases} \text{HML}_t & \text{if HML's prior-year return is positive,} \\ 0 & \text{otherwise,} \end{cases}$$

and similarly for the other factors. This *conditional* five-factor model explains 49% of the variation in UMD's returns.

has performed well over the prior year, then the stocks in this factor's long leg have, by definition, higher past returns than those in its short leg. The existence of individual stock momentum alone would then predict that the size factor should continue to perform well. This incidental momentum effect could give rise to what looks like factor momentum, even if momentum does not reside in factors. In this section we quantify the extent to which this mechanism contributes to factor momentum profits. Whereas the tests in Section 4 measure whether individual stock momentum survives when we control for factor momentum—it does not—we now approach this question about the connection between individual stock and factor momentum from the opposite direction: how much of factor momentum survives when we control for individual stock momentum?

We examine the origins of momentum by measuring factor momentum in *momentum-neutral* factors. A factor has incidental momentum if $\sum_{i=1}^{N} w_{i,t} r_{i,t-12,t-2} \neq 0$, that is, if the factor's past return, computed using its *current* weights $w_{i,t}$, is nonzero. An investor who invests in such a factor may indirectly benefit from the momentum in stock returns. We construct momentum-neutral factors by taking the Kozak et al. (2020) factors and twisting the factor weights as little as possible to render them orthogonal with respect to past returns. The objective is to find new weights x_i such that

$$\min_{x_i} \sum_i (w_i - x_i)^2 \quad \text{s.t.} \ \sum_{i=1}^{N} x_i = 0 \ \text{ and } \ \sum_{i=1}^{N} x_i r_{i,t-12,t-2} = 0. \tag{21}$$

We show in Appendix F that the weights x_i are equivalent to the residuals from a cross-sectional regression of the original factor weights on past returns:

$$w_{i,t} = a + b\, r_{i,t-12,t-2} + x_{i,t}. \tag{22}$$

We call the factors with weights $x_{i,t}$ *momentum-neutral* factors. The idea of momentum-neutral factors applies to all factors. An investor investing in value might, for example, notice that the returns on the stocks in the long and short legs over the prior year differ. To avoid an incidental bet on stock-level momentum, the investor could alter the weights to render the value and growth portfolios perfectly indistinguishable from each other based on past returns: the value portfolio that the investor buys has performed exactly as well (or poorly) as the growth portfolio he sells. Momentum-neutrality means that at, any point in time, $\sum_{i=1}^{N} x_i r_{i,t-12,t-2} = 0$. This condition does *not* mean that the factor's past return is zero: a factor's return is based on time t weights and future returns; momentum neutrality is about time t weights and *past* returns.

The weights of the standard and momentum-neutral factors are close to each other. Consider, for example, the value factor which invests in stocks based on their book-to-market ratios. Most of the cross-sectional return variation in this factor's weights are unrelated to past returns. The average and median R^2s from the regression in equation (22) are 2.9% and 1.5%. These numbers imply that the original factor weights and the momentum-neutral weights are very close: the average correlation between them is $\sqrt{1 - 0.029} = 0.99$. The average R^2s of the 47 factors range from 0.4% (growth in long-term net operating assets) to 10.2% (Asness and Frazzini's (2013) monthly version of value); the average across all factors is 2.4%.

Table A2 in the Appendix lists shows the annualized CAPM alphas for momentum-neutral versions of the 47 factors. Momentum-neutral factors typically earn similar premiums as the original factors but with lower volatility; they therefore often earn higher information ratios. Consider, for example, those 37 original factors whose premiums are statistically significant at the 5% level. Momentum-neutral versions of 30 of these factors earn higher information ratios than the original factors.

Table 11 shows that the factor momentum strategy that trades momentum-neutral factors is *more profitable* than the one that trades the original factors. The strategy that trades the ten high-eigenvalue principal components extracted from the standard factors earns a CAPM alpha that is significant with a t-value of 6.51; this estimate correspond to an annualized information ratio

Table 11: Factor momentum in momentum-neutral factors

This table reports estimates from time-series regressions in which the dependent variable is the return on a factor momentum strategy. We construct time-series factor momentum strategies from the 47 factors listed in Table A2, using either the original or momentum-neutral versions of these factors. Momentum-neutral factors adjust factor weights so that they are orthogonal to individual stock returns from month $t-12$ to $t-2$. We extract the PC factors from either the original or momentum-neutral factors and trade momentum in the first ten high-eigenvalue PC factors. The independent variables are the five factors of the Fama-French model and the other factor momentum strategy. The sample begins in July 1973 and ends in December 2019.

	Dependent variable			
	Momentum in original factors		Momentum in momentum-neutral factors	
Independent variable	(1)	(2)	(3)	(4)
Alpha	0.18	0.03	0.15	0.06
	(6.51)	(1.45)	(7.53)	(3.91)
Momentum in original factors				0.52
				(24.83)
Momentum in momentum-neutral factors		1.01		
		(24.83)		
FF5 factors	Y	Y	Y	Y
N	558	558	558	558
R^2	2.4%	53.9%	5.9%	55.5%

of 0.96. This t-value increases to 7.53 when we construct the strategy using momentum-neutral PC factors; this estimate corresponds to an information ratio of 1.10. Controlling for momentum-neutral factors, the strategy that trades the PCs based on the original factors has an alpha of 3 basis points (t-value = 1.45). The strategy that trades momentum-neutral factors, however, remains significant with a t-value of 3.91 when we reverse this regression.

The finding that factor momentum in momentum-neutral factors subsumes that in the original factors rejects the possibility that factor momentum is merely incidental momentum. In fact, the results indicate that incidental momentum explains *none* of the factor momentum profits.

6 Conclusion

Positive autocorrelation is a pervasive feature of factor returns. Factors with positive returns over the prior year earn significant premiums; those with negative returns earn premiums that are indistinguishable from zero. Factor momentum is a strategy that bets on these autocorrelations in factor returns.

Factor momentum transmits into the cross section of stock returns through variation in stocks' factor loadings. Consistent with this mechanism, we show that factor momentum explains both the "standard" momentum of Jegadeesh and Titman (1993) and other forms of it: industry-adjusted momentum, industry momentum, intermediate momentum, Sharpe momentum, and three versions of residual momentum. By contrast, these other momentum factors do not explain factor momentum. An empirical model that controls for momentum found in high-eigenvalue PC factors describes the data well. We find no evidence of any momentum in individual stock returns that does not originate from the factors.

Our results imply that momentum is not a distinct factor; rather, it is the sum of the autocorrelations found in the other factors. An investor trading momentum bears systematic risk because all winners and all losers have similar factor exposures. We are left with the impression that momentum is unrelated to the other factors only because these loadings change over time.

Factor momentum may stem from mispricing. We show that Kozak et al.'s (2018) model with sentiment investors produces factor momentum when sentiment is sufficiently persistent. This model predicts that, if so, momentum should concentrate in factors that explain more of the cross section of stock returns. The data support this prediction; factor momentum, in particular during the second half of the sample, concentrates in high-eigenvalue factors.

We leave two questions for future research. First, although factor momentum is consistent with Kozak et al.'s (2018) model of sentiment investors, this consistency does not imply that factor momentum must stem from mispricing. The point of Kozak et al. (2018), after all, is that the extent to which covariances align with premiums provides no clues as to whether factor premiums compensate for risks or reflect mispricings. The finding that momentum resides in the high-eigenvalue factors is a more general implication of the absence of near-arbitrage opportunities. If we were to write down a rational model with time-varying risk premiums, would such a model provide additional—and distinguishing—predictions about factor momentum? Second, although we find no residual momentum net of factor momentum, this result does not conclusively prove that firm-specific returns are serially uncorrelated. Because of complications stemming from omitted factors and estimated betas, we cannot settle this issue if we do not observe the true asset pricing model. To pass the final judgment on the divide between factor and firm-specific momentum, one would need to device a method for extracting—or find a natural experiment for identifying—firm-specific returns.

References

Ang, Andrew, Robert J. Hodrick, Yuhang Xing, and Xiaoyan Zhang, 2006, The cross-section of volatility and expected returns, *Journal of Finance* 61, 259–299.

Arnott, Rob, Mark Clements, Vitali Kalesnik, and Juhani T. Linnainmaa, 2019, Factor momentum, Working paper.

Asness, Cliff, and Andrea Frazzini, 2013, The devil in HML's details, *Journal of Portfolio Management* 39, 49–68.

Asness, Clifford S., Andrea Frazzini, and Lasse Heje Pedersen, 2019, Quality minus junk, *Review of Accounting Studies* 24, 34–112.

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

Avramov, Doron, Si Cheng, Amnon Schreiber, and Koby Shemer, 2017, Scaling up market anomalies, *Journal of Investing* 26, 89–105.

Avramov, Doron, Tarun Chordia, Gergana Jostova, and Alexander Philipov, 2013, Anomalies and financial distress, *Journal of Financial Economics* 108, 139–159.

Baker, Malcolm, and Jeffrey Wurgler, 2006, Investor sentiment and the cross-section of stock returns, *Journal of Finance* 61, 1645–1680.

Banz, Rolf W., 1981, The relationship between return and market value of common stocks, *Journal of Financial Economics* 9, 3–18.

Baquero, Guillermo, Jenke Ter Horst, and Marno Verbeek, 2005, Survival, look-ahead bias, and persistence in hedge fund performance, *Journal of Financial and Quantitative Analysis* 40, 493–517.

Barillas, Francisco, and Jay Shanken, 2017, Which alpha? *Review of Financial Studies* 30, 1316–1338.

Basu, Sanjoy, 1983, The relationship between earnings' yield, market value and return for nyse common stocks: Further evidence, *Journal of Financial Economics* 12, 129–156.

Berk, Jonathan B., Richard C. Green, and Vasant Naik, 1999, Optimal investment, growth options, and security returns, *Journal of Finance* 54, 1553–1607.

Beyhaghi, Mehdi, and Sina Ehsani, 2017, The cross-section of expected returns in the secondary corporate loan market, *Review of Asset Pricing Studies* 7, 243–277.

Bhojraj, Sanjeev, and Bhaskaran Swaminathan, 2006, Macromomentum: Returns predictability in international equity indices, *Journal of Business* 79, 429–451.

Black, Fischer, 1993, Beta and return, *Journal of Portfolio Management* 20, 8–18.

Bondt, Werner F. M., and Richard Thaler, 1985, Does the stock market overreact? *Journal of Finance* 40, 793–805.

Boyson, Nicole M., 2008, Hedge fund performance persistence: a new approach, *Financial Analysts Journal* 64, 27–44.

Brown, Stephen J., and William N. Goetzmann, 1995, Performance persistence, *Journal of Finance* 50, 679–698.

Carhart, Mark M., 1997, On persistence in mutual fund performance, *Journal of Finance* 52, 57–82.

Christoffersen, Peter F., and Francis X. Diebold, 2006, Financial asset returns, direction-of-change forecasting, and volatility dynamics, *Management Science* 52, 1273–1287.

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

Cohen, Randolph B, Christopher Polk, and Tuomo Vuolteenaho, 2003, The value spread, *Journal of Finance* 58, 609–641.

Cohen, Randolph B., and Christopher K. Polk, 1998, The impact of industry factors in asset-pricing tests, Kellogg Graduate School of Management working paper.

Conrad, Jennifer, and Gautam Kaul, 1998, An anatomy of trading strategies, *Review of Financial*

Studies 11, 489–519.

Cooper, Michael J., Roberto C. Gutierrez, and Allaudeen Hameed, 2004, Market states and momentum, *Journal of Finance* 59, 1345–1365.

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

Daniel, Kent, Lira Mota, Simon Rottke, and Tano Santos, 2020, The cross-section of risk and returns, *Review of Financial Studies* 33, 1927–1979.

Dickey, David A., and Wayne A. Fuller, 1979, Distribution of the estimators for autoregressive time series with a unit root, *Journal of the American Statistical Association* 74, 427–431.

Ehsani, Sina, and Juhani T. Linnainmaa, 2020a, Time-series efficient factors, Working Paper.

Ehsani, Sina, and Juhani T. Linnainmaa, 2020b, Do firm-specific returns really display momentum? Working Paper.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.

Fama, Eugene F., and Kenneth R. French, 2015, A five-factor asset pricing model, *Journal of Financial Economics* 116, 1–22.

Fama, Eugene F., and Kenneth R. French, 2016, Dissecting anomalies with a five-factor model, *Review of Financial Studies* 29, 69–103.

Fisher, Ronald A., 1915, Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population, *Biometrika* 10, 507–521.

Frazzini, Andrea, Ronen Israel, and Tobias J. Moskowitz, 2012, Trading costs of asset pricing anomalies, Fama-Miller Working Paper 14-05.

Frazzini, Andrea, David Kabiller, and Lasse Heje Pedersen, 2018, Buffett's alpha, *Financial Analysts Journal* 74, 35–55.

Frazzini, Andrea, and Lasse Heje Pedersen, 2014, Betting against beta, *Journal of Financial Economics* 111, 1–25.

Gibbons, Michael R., Stephen A. Ross, and Jay Shanken, 1989, A test of the efficiency of a given portfolio, *Econometrica* 57, 1121–1152.

Goyal, Amit, and Narasimhan Jegadeesh, 2017, Cross-sectional and time-series tests of return predictability: What is the difference? *Review of Financial Studies* 31, 1784–1824.

Griffin, John M., Xiuqing Ji, and J. Spencer Martin, 2003, Momentum investing and business cycle risk: Evidence from pole to pole, *Journal of Finance* 58, 2515–2547.

Grinblatt, Mark, Sheridan Titman, and Russ Wermers, 1995, Momentum investment strategies, portfolio performance, and herding: A study of mutual fund behavior, *American Economic Review* 85, 1088–1105.

Grundy, Bruce D, and J. Spencer Martin Martin, 2001, Understanding the nature of the risks and the source of the rewards to momentum investing, *Review of Financial studies* 14, 29–78.

Haddad, Valentin, Serhiy Kozak, and Shrihari Santosh, 2020, Factor timing, *Review of Financial Studies* 33, 1980–2018.

Hendricks, Darryll, Jayendu Patel, and Richard Zeckhauser, 1993, Hot hands in mutual funds: Short-run persistence of relative performance, 1974–1988, *Journal of Finance* 48, 93–130.

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.

Huang, Dashan, Jiangyuan Li, Liyao Wang, and Guofu Zhou, 2020, Time-series momentum: Is it there? *Journal of Financial Economics* 135, 774–794.

Huberman, Gur, and Shmuel Kandel, 1987, Mean-variance spanning, *Journal of Finance* 42, 873–888.

Jagannathan, Ravi, Alexey Malakhov, and Dmitry Novikov, 2010, Do hot hands exist among hedge fund managers? An empirical evaluation, *Journal of Finance* 65, 217–255.

Jegadeesh, Narasimhan, 1990, Evidence of predictable behavior of security returns, *Journal of Finance* 45, 881–898.

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

Johnson, Timothy C., 2002, Rational momentum effects, *Journal of Finance* 57, 585–608.

Jostova, Gergana, Stanislava Nikolova, Alexander Philipov, and Christof W. Stahel, 2013, Momentum in corporate bond returns, *Review of Financial Studies* 26, 1649–1693.

Kenney, John Francis, and Ernest Sydney Keeping, 1951, *Mathematics of Statistics: Part Two* (D. Van Nostrand Company, New York, NY).

Korajczyk, Robert A., and Ronnie Sadka, 2004, Are momentum profits robust to trading costs? *Journal of Finance* 59, 1039–1082.

Kozak, Serhiy, Stefan Nagel, and Shrihari Santosh, 2018, Interpreting factor models, *Journal of Finance* 73, 1183–1223.

Kozak, Serhiy, Stefan Nagel, and Shrihari Santosh, 2020, Shrinking the cross-section, *Journal of Financial Economics* 135, 271–292.

Lee, Jongsub, Andy Naranjo, and Stace Sirmans, 2014, CDS momentum: Slow moving credit ratings and cross-market spillovers, Working paper.

Lesmond, David A., Michael J. Schill, and Chunsheng Zhou, 2004, The illusory nature of momentum profits, *Journal of Financial Economics* 71, 349–380.

Lewellen, Jonathan, 2002, Momentum and autocorrelation in stock returns, *Review of Financial Studies* 15, 533–564.

Lewellen, Jonathan, 2015, The cross-section of expected stock returns, *Critical Finance Review* 4, 1–44.

Lo, Andrew W., and A. Craig MacKinlay, 1990, When are contrarian profits due to stock market overreaction? *Review of Financial Studies* 3, 175–205.

Loughran, Tim, and Jay R. Ritter, 1995, The new issues puzzle, *Journal of Finance* 50, 23–51.

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

Menkhoff, Lukas, Lucio Sarno, Maik Schmeling, and Andreas Schrimpf, 2012, Carry trades and global foreign exchange volatility, *Journal of Finance* 67, 681–718.

Merton, Robert C., 1973, An intertemporal capital asset pricing model, *Econometrica* 41, 867–887.

Miffre, Joëlle, and Georgios Rallis, 2007, Momentum strategies in commodity futures markets, *Journal of Banking and Finance* 31, 1863–1886.

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

Moskowitz, Tobias J., Yao Hua Ooi, and Lasse Heje Pedersen, 2012, Time series momentum, *Journal of Financial Economics* 104, 228–250.

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

Novy-Marx, Robert, 2013, The other side of value: The gross profitability premium, *Journal of Financial Economics* 108, 1–28.

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

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

Rachev, Svetlozar, Teo Jašić, Stoyan Stoyanov, and Frank J. Fabozzi, 2007, Momentum strategies based on reward–risk stock selection criteria, *Journal of Banking and Finance* 31, 2325–2346.

Rosenberg, Barr, Kenneth Reid, and Ronald Lanstein, 1985, Persuasive evidence of market inefficiency, *Journal of Portfolio Management* 11, 9–16.

Ross, Stephen A., 1976, The arbitrage theory of capital asset pricing, *Journal of Economic Theory* 13, 341–360.

Sagi, Jacob S., and Mark S. Seasholes, 2007, Firm-specific attributes and the cross-section of momentum, *Journal of Financial Economics* 84, 389–434.

Sloan, Richard, 1996, Do stock prices fully reflect information in accruals and cash flows about future earnings? *The Accounting Review* 71, 289–315.

Szakmary, Andrew C., Qian Shen, and Subhash C. Sharma, 2010, Trend-following trading strategies in commodity futures: A re-examination, *Journal of Banking and Finance* 34, 409–426.

Titman, Sheridan, KC John Wei, and Feixue Xie, 2004, Capital investments and stock returns, *Journal of Financial and Quantitative Analysis* 39, 677–700.

Zaremba, Adam, and Jacob Shemer, 2017, Is there momentum in factor premia? Evidence from international equity markets, *Research in International Business and Finance* 1, 10–20.

Table A1: Autocorrelations in factor returns

Table 2 in the main text reports estimates from regressions in which the dependent variable is a factor's return in month t and the explanatory variable is an indicator variable that takes the value of one if the factor's return over the prior year is positive and zero other. This table reports estimates from regressions in which the dependent variable is a factor's monthly return and the independent variable is the factor's average return over the prior year. We estimate these regressions using pooled data (first row) and separately for each anomaly (remaining rows). We cluster the standard errors by month in the pooled regression.

	Intercept		Slope	
Anomaly	$\hat{\alpha}$	$t(\hat{\alpha})$	$\hat{\beta}$	$t(\hat{\beta})$
Pooled	0.27	5.27	0.25	2.59
		U.S. factors		
Accruals	0.22	2.82	−0.03	−0.23
Betting against beta	0.41	2.22	0.50	3.38
Cash-flow to price	0.23	2.14	0.15	0.92
Investment	0.21	2.59	0.24	1.43
Earnings to price	0.24	2.12	0.18	1.12
Book-to-market	0.21	1.75	0.26	1.60
Liquidity	0.35	2.44	0.09	0.54
Long-term reversals	0.12	1.18	0.41	3.26
Net share issues	0.15	1.64	0.34	1.92
Quality minus junk	0.28	3.27	0.27	1.76
Profitability	0.20	1.71	0.26	1.08
Residual variance	0.10	0.51	0.20	1.10
Market value of equity	0.17	1.45	0.29	1.80
Short-term reversals	0.50	2.76	−0.01	−0.04
Momentum	0.64	3.91	0.00	−0.02
		Global factors		
Betting against beta	0.59	3.05	0.31	1.82
Investment	0.11	1.00	0.31	1.26
Book-to-market	0.19	1.34	0.44	2.26
Quality minus junk	0.44	3.82	0.12	0.55
Profitability	0.29	3.46	0.16	0.82
Market value of equity	0.09	0.81	0.20	1.02
Momentum	0.75	3.86	−0.09	−0.45

Table A2: Standard and momentum-neutral factors based on the Kozak et al. (2020) characteristics

This table reports annualized CAPM alphas and t-values associated with these alphas for 47 factors. These factors are based on the Kozak et al. (2020) characteristics except for the seven characteristics related to momentum. Each characteristic is converted into a centered cross-sectional rank normalized by the average absolute deviation from the mean. A factor's return in month t is the product of stock returns in month t and these characteristics ("weights") in month $t-1$. The factors are not re-signed; the size factor, for example, is long large stocks and short small stocks. The sample excludes each month stocks with market values less than 0.01% of the total market value of all common stocks traded on NYSE, Amex, and Nasdaq. Original factors use the raw firm characteristics. Momentum-neutral factors use weights orthogonal to stocks' returns over the prior year (see Appendix F). Freq. is the frequency at which the characteristics are recomputed: A = annual, Q = quarterly, and M = monthly.

			Original		Momentum-neutral	
#	Factor	Freq.	$\hat{\alpha}$	$t(\hat{\alpha})$	$\hat{\alpha}$	$t(\hat{\alpha})$
1	Size	A	−0.58%	−0.99	−0.13%	−0.28
2	Value	A	2.45%	3.48	2.44%	3.95
3	Gross profitability	A	0.77%	1.39	0.62%	1.20
4	Cash flow duration	M	−2.55%	−3.65	−2.58%	−4.11
5	Value-Profitability	M	3.11%	5.84	2.79%	5.90
6	Piotroski's F-score	A	1.62%	3.98	1.34%	4.97
7	Debt issuance	A	0.13%	0.28	0.18%	0.56
8	Share repurchases	A	1.59%	3.47	1.45%	4.80
9	Share issuance, annual	A	−2.57%	−5.52	−2.47%	−6.30
10	Accruals	A	−0.98%	−3.13	−0.80%	−2.68
11	Asset growth	A	−2.29%	−5.03	−2.33%	−5.69
12	Asset turnover	A	0.97%	1.90	0.75%	1.55
13	Gross margins	A	−0.21%	−0.50	−0.12%	−0.32
14	Dividend yield	A	1.75%	3.28	1.87%	3.78
15	Earnings/Price	A	3.19%	4.51	2.97%	5.09
16	Cash flow/Market value of equity	A	2.77%	4.09	2.62%	4.40
17	Net operating assets	A	−2.02%	−5.15	−1.81%	−5.03
18	Investment	A	−2.10%	−5.31	−2.01%	−5.57
19	Investment-to-capital	A	−2.28%	−3.26	−2.32%	−3.99
20	Investment growth	A	−1.76%	−5.25	−1.63%	−5.39
21	Sales growth	A	−1.66%	−3.30	−1.84%	−4.19
22	Leverage	A	1.84%	2.34	1.77%	2.62
23	Return on assets	A	0.73%	1.64	0.65%	1.57
24	Return on equity	A	0.82%	1.79	0.68%	1.70
25	Sales-to-Price	A	2.73%	3.72	2.46%	3.97
26	Growth in LTNOA	A	−0.26%	−0.86	−0.20%	−0.72
27	Dividend growth	A	−0.70%	−2.31	−0.76%	−2.62
28	Abnormal investment	A	−0.19%	−0.32	−0.40%	−0.68
29	Short interest	M	−0.95%	−2.05	−0.96%	−2.25
30	Long-term reversals	M	−1.73%	−3.02	−1.76%	−3.34

(continued)

			Original		Momentum-neutral	
#	Factor	Freq.	$\hat{\alpha}$	$t(\hat{\alpha})$	$\hat{\alpha}$	$t(\hat{\alpha})$
31	Value	M	1.94%	2.23	2.80%	4.91
32	Share issuance	M	−2.62%	−5.01	−2.76%	−6.70
33	PEAD (SUE)	M	2.62%	4.42	1.57%	3.61
34	Return on book equity	M	2.43%	4.08	1.53%	3.45
35	Return on market equity	M	4.58%	5.30	4.03%	6.75
36	Return on assets	Q	2.21%	3.48	1.29%	2.50
37	Short-term reversals	M	−2.19%	−3.09	−2.42%	−4.53
38	Idiosyncratic volatility	M	−3.40%	−3.75	−3.27%	−5.16
39	Beta arbitrage	M	−2.92%	−3.73	−3.08%	−4.74
40	Seasonality	M	1.74%	3.90	1.49%	3.72
41	Industry relative reversals	M	−3.81%	−7.79	−3.85%	−9.75
42	Industry relative reversals (low vol)	M	−5.00%	−15.18	−5.04%	−16.45
43	Composite issuance	M	−2.73%	−6.45	−2.70%	−7.30
44	Price	M	0.55%	0.75	−0.51%	−1.13
45	Firm age	M	1.69%	2.46	1.79%	4.06
46	Share volume	M	−2.29%	−2.78	−2.66%	−4.35
47	Initial public offering	M	−6.87%	−2.99	−5.59%	−4.92

Table A3: Conditional covariances with the momentum factor: Decomposition

This table reports covariances between UMD and factor returns. It is similar to Table 10 except that (1) we report covariances instead of correlations and (2) we decompose covariances into two components: the covariance between UMD and each factor's long leg (L) and that between UMD and the *negative* of each factor's short leg ($-S$). These covariances add up to the total covariance between UMD and the factor. We compute the covariance between UMD's and factors' month-t returns, conditioning on the sign of the factor's average return from month $t-12$ to $t-1$. The data are for the 12 U.S. factors with portfolio-level data. The data begin in July 1963 and end in December 2019.

	Conditional on a year of positive factor returns			Conditional on a year of negative factor returns		
	Cov. between UMD and:			Cov. between UMD and:		
Factor	L	$-S$	$L-S$	L	$-S$	$L-S$
Pooled	−3.51	6.16	2.64	−4.28	−1.18	−5.46
Size	−5.54	6.94	1.39	1.36	−6.43	−5.08
Value	−3.62	5.42	1.80	−6.37	−2.29	−8.67
Profitability	−3.91	7.50	3.59	−0.64	−3.90	−4.54
Investment	−3.80	5.31	1.51	−2.39	−0.78	−3.17
Accruals	−0.49	2.69	2.20	−5.34	4.00	−1.34
Cash-flow to price	−2.98	5.21	2.23	−5.32	−1.93	−7.25
Earnings to price	−3.76	5.68	1.91	−4.70	−3.20	−7.90
Long-term reversals	−4.81	5.94	1.14	−4.81	0.96	−3.84
Net share issues	−4.71	8.97	4.26	−0.09	−3.04	−3.12
Quality minus junk	7.31	−4.75	2.56	−11.39	6.40	−4.99
Residual variance	−5.14	21.22	16.05	2.07	−11.74	−9.67
Short-term reversals	−5.12	1.47	−3.64	−10.33	5.56	−4.77

Internet Appendix

A Cross-sectional and time-series factor momentum strategies with different formation and holding periods

In the main text we form cross-sectional and time-series factor momentum strategies using twelve-month formation and one-month holding periods. Table A4 examines the performance of these strategies with formation and holding periods ranging from one month to two years. When the holding period is longer than a month, we correct standard errors using Jegadeesh and Titman's (1993) overlapping-portfolio approach. In month t, the return on a strategy with an h-month holding period is computed as the average return across h strategies. These h strategies correspond to portfolios formed every month between months $t-h$ and $t-1$. This approach produces a single time series for each formation period-holding period combination.

Panel A of Table A4 examines the performance of time-series factor momentum strategies. The time-series strategy with the one-month formation and holding periods earns an average return of 33 basis points (t-value = 6.97). These strategies typically remain profitable also with longer formation and holding periods. All time-series momentum strategies are the most profitable when held for one month; at longer holding periods, this strategy's performance deteriorates because it cannot immediately rebalance away from factors whose average returns turn negative.

Panel B of Table A4 shows that a cross-sectional momentum strategy with one-month formation and holding periods earns an average return of 28 basis points per month (t-value of 6.07). This strategy is the most profitable among all cross-sectional strategies. The profits on this short-term strategy decay quickly: the return on a strategy with one-month formation and three-month holding period is small and insignificant. Some of the strategies with longer formation periods, although less profitable initially, earn statistically significant profits at the three- and six-month holding periods.

Table A4: Average returns of time-series and cross-sectional factor momentum strategies: Alternative formation and holding periods

This table reports annualized average returns and t-values for time-series and cross-sectional factor momentum strategies that trade the 20 non-momentum factors listed in Table 1. The time-series factor momentum strategy is long factors with positive returns over a formation period that ranges from one month to two years and short factors with negative returns. The cross-sectional momentum strategy is long factors that earned above-median returns relative to other factors over the same formation period and short factors with below-median returns. We vary the rebalancing frequency from one month to two years. We use the Jegadeesh and Titman (1993) approach to correct standard errors for overlapping returns when the holding period is longer than a month. The sample begins in July 1964 and ends in December 2019.

Panel A: Time-series factor momentum

Holding	Formation period						Formation period					
period	1	3	6	12	18	24	1	3	6	12	18	24
	Average returns						t**-values**					
1	0.33	0.28	0.31	0.33	0.27	0.24	6.97	5.57	6.54	7.02	6.13	5.61
3	0.06	0.12	0.19	0.24	0.23	0.19	1.27	2.56	4.32	5.22	5.34	4.39
6	0.12	0.09	0.21	0.19	0.19	0.17	2.62	2.02	4.99	4.57	4.61	4.08
12	0.14	0.10	0.10	0.09	0.08	0.09	3.17	2.36	2.33	2.09	1.97	2.16
18	0.03	0.00	0.02	0.03	0.07	0.08	0.82	0.00	0.44	0.79	1.62	1.97
24	0.06	0.03	0.05	0.10	0.08	0.14	1.48	0.79	1.30	2.35	1.84	3.22

Panel B: Cross-sectional factor momentum

Holding	Formation period						Formation period					
period	1	3	6	12	18	24	1	3	6	12	18	24
	Average returns						t**-values**					
1	0.27	0.20	0.18	0.20	0.13	0.12	6.96	4.76	4.62	5.04	3.62	3.15
3	0.04	0.04	0.09	0.11	0.08	0.07	0.92	1.06	2.50	2.84	2.32	1.93
6	0.06	0.04	0.11	0.09	0.07	0.03	1.76	1.22	2.99	2.47	1.94	0.92
12	0.09	0.04	0.00	−0.03	−0.01	−0.04	2.57	1.25	−0.13	−0.72	−0.41	−1.03
18	−0.03	−0.06	−0.07	−0.03	−0.02	0.03	−0.95	−1.82	−1.83	−0.97	−0.59	0.84
24	0.02	−0.03	0.00	0.01	0.04	0.03	0.53	−0.89	0.12	0.30	1.14	0.90

B Decomposing equity momentum profits

Equation (18) attributes momentum in the cross section of stock returns to four sources: (1) factor autocovariances, (2) factor cross-serial covariances, (3) differences in unconditional mean returns, and (4) autocovariances in residuals. This decomposition assumes that strategy weights are proportional to assets' past returns. In this appendix we compute the terms of this decomposition using four different asset pricing models. Except for the last component—the dispersion in mean returns—the decomposition estimates are sensitive to the choice of the asset pricing model.

Our test assets are 100 value-weighted portfolios sorted by stocks' returns from month $t-12$ to $t-2$ using NYSE breakpoints. The sample includes all stocks listed on NYSE, AMEX, and Nasdaq, identified on CRSP as ordinary common shares (share codes 10 and 11). We require a stock to have non-missing returns from month $t-12$ to $t-2$, a non-missing return in month t, and a non-missing market value of equity at the end of month $t-1$. We use momentum-sorted portfolios to maximize the dispersion in past returns and to reduce noise.[29] Portfolio p's contribution to the linear-weight strategy's profits at time t equals

$$\pi_{p,t}^{\text{mom}} = (R_{-t}^{p} - \bar{R}_{-t})(R_{t}^{p} - \bar{R}_{t}), \tag{A-1}$$

where R_{-t}^{p} is portfolio p's return from month $t-12$ to $t-2$ (denoted by $-t$) and $\bar{R}_t$ is the equal-weighted average return across all portfolios. This linear-weight momentum strategy earns an average monthly return of 0.69% (t-value = 2.66).

[29] A momentum strategy with weights proportional to past returns (equation A-1) is not profitable when applied to the cross section of individual stock returns. It earns an average return of 0.4% per month (t-value = 1.04).

Section 4.1 shows that linear strategy's profit decomposes into four terms,

$$\mathrm{E}[\pi_t^{\mathrm{mom}}] = \underbrace{\sum_{f=1}^{F}\left[\mathrm{cov}(r_{-t}^{f}, r_t^{f})\,\sigma_{\beta^f}^2\right]}_{\text{factor autocovariances}} + \underbrace{\sum_{f=1}^{F}\sum_{g\neq f}^{F}\left[\mathrm{cov}(r_{-t}^{f}, r_t^{g})\,\mathrm{cov}(\beta^f, \beta^g)\right]}_{\text{factor serial cross-covariances}} + \underbrace{\sigma_\eta^2}_{\substack{\text{variation in}\\ \text{mean returns}}} + \underbrace{\frac{1}{N}\sum_{i=1}^{N}\left[\mathrm{cov}(\varepsilon_{i,-t}, \varepsilon_{i,t},)\right]}_{\substack{\text{autocovariances}\\ \text{in residuals}}}. \tag{A-2}$$

This decomposition assumes that the stock-specific return component, $\varepsilon_{i,t}$, does not correlate across firms or with factors either concurrently or at leads or lags.

In Table A5 we show factor-level estimates of this decomposition under the Fama and French (2015) model to illustrate how betas and covariances interact and aggregate. The first two terms depend on the autocovariance matrix of factor returns and the covariance matrix of factor betas. We compute the autocovariance matrix between factors' month $t-12$ to $t-2$ and month t returns. The element (f, g) in this matrix is the covariance $\mathrm{cov}(r_{-t}^{f}, r_t^{g})$. The diagonal elements are factor autocovariances and the off-diagonal elements are factor cross-serial covariances. All diagonal elements and most off-diagonal elements are positive: a high past return on a factor predicts high returns not only on the factor itself but on *other factors* as well. The market factor is an exception: high past returns on the market predict lower returns on the other factors. Conversely, high returns on the size (SMB), value (HML), profitability (RMW), and investment (CMA) factors predict low future market returns.

Factor auto- and cross- covariances transmit into the cross-section through the covariance matrix of betas. We estimate each portfolio's beta every month using three months of daily data from month $t-2$ to t. Panel B of Table A5 shows that the pairwise covariances between factor betas are typically positive. Assets with high market betas, for example, also have higher betas against the other factors. The exceptions are the investment and value factors, whose betas are negatively correlated.

Table A5: Decomposing equity momentum profits under the Fama and French (2015) five-factor model

This table reports estimates of the terms in equation (18) under the assumption that the Fama and French (2015) five-factor model governs asset returns. We use data on 100 momentum-sorted portfolios; each stock is assigned to a portfolio based on its return from month $t-12$ to $t-2$. The first panel shows the estimates of the auto- and cross-serial covariances between factor returns, $\text{cov}(r^f_{-t}, r^f_t)$. The second panel shows the covariance matrix of factor betas, $\text{cov}(\beta^g, \beta^f)$. We estimate portfolio betas using daily data from month $t-2$ to t. Panel C shows the four terms of the decomposition in equation (18).

Panel A: Autocovariance matrix of factor returns

Average factor return	Month t return				
from month $t-12$ to $t-2$	MKTRF	SMB	HML	RMW	CMA
MKTRF	0.00	−0.38	−0.14	−0.04	−0.01
SMB	−0.38	0.24	0.14	0.01	0.14
HML	0.03	0.31	0.16	0.00	0.04
RMW	−0.04	0.16	0.10	0.09	0.13
CMA	−0.10	0.17	0.09	−0.02	0.08

Panel B: Covariance matrix of factor betas

	MKTRF	SMB	HML	RMW	CMA
MKTRF	0.08				
SMB	0.04	0.19			
HML	0.04	0.05	0.33		
RMW	0.01	0.03	0.09	0.36	
CMA	0.03	0.00	−0.12	0.04	0.42

Panel C: Decomposition estimates

Factor autocovariances	$\sum_{f=1}^{F} \left[\text{cov}(r^f_{-t}, r^f_t)\, \sigma^2_{\beta_f}\right]$	0.16%
+ Factor cross-serial covariances	$\sum_{f=1}^{F} \sum_{g \neq f}^{F} \left[\text{cov}(r^f_{-t}, r^g_t)\, \text{cov}(\beta^f, \beta^g)\right]$	−0.02%
+ Variance of mean returns	σ^2_η	0.11%
+ Residual autocovariances	$\frac{1}{N} \sum_{i=1}^{N} [\text{cov}(\varepsilon_{i,-t}, \varepsilon_{i,t})]$	0.43%
= Total		0.69%

We use the autocovariance matrix and covariance matrix of factor betas to compute the contributions of the auto-covariances and cross-serial covariances to the profits of the stock momentum strategy. The total contribution of the autocovariance term, 0.16%, is the sum of the pairwise products of the diagonal elements of the autocovariance and beta matrices. The contribution of the cross-covariance term, -0.02%, is the sum of the pairwise products of the off-diagonal elements of the two matrices.

The third term in equation (A-2), variation in mean returns, is the only component that does not depend on the choice of the asset pricing model. The estimate for this term, computed as the cross-sectional variance of the full-sample average returns, is 0.11%.[30] This Conrad and Kaul (1998) mechanism therefore accounts for one-sixth of the total profits.

The last term is the auto-covariance between past and future residuals. We estimate this term as the strategy's total return minus the sum of the first three components. This estimate, by the virtue of being the remainder, also includes covariances between firm-specific returns and between firm-specific returns and factors, if any; these terms will be nonzero if the asset pricing model is misspecified. Because the strategy's average monthly return in 0.69%, this five-factor model decomposition attributes $0.69\% - [0.16\% + (-0.02\%) + 0.11\%] = 0.43\%$ to the autocovariance in residuals.

The estimates of the decomposition are sensitive to the choice of the factor model. The estimates of firm-specific returns are subject to the omitted-factor bias: if any factors omitted from the model display momentum, then this momentum will be attributed to the residuals. We illustrate this sensitivity in Table A6 by comparing four asset pricing models: CAPM, Fama and French (1993) three-factor model, Fama and French (2015) five-factor model, and a seven-factor model that augments the Fama and French (2015) model with the betting-against-beta and quality-

[30] The number of observations affects the estimate of the cross-sectional variance. Kenney and Keeping (1951) show that if $\hat{\sigma}$ is the sample's standard deviation, then $\frac{\hat{\sigma}}{b(N)}$ is an unbiased estimator of the actual standard deviation. The adjustment factor $b(N) = \sqrt{\frac{2}{N-1}}\,\Gamma\left(\frac{N}{2}\right)/\,\Gamma\left(\frac{N-1}{2}\right)$, in which N is the number of observations and $\Gamma(\cdot)$ is the gamma function. In our sample, N is 100 and the adjustment factor is therefore $b(N) = 0.997$.

Table A6: Decomposition of equity momentum profits: Comparing four asset pricing models

This table reports estimates for each component of the momentum-profit decomposition in equation (18). We use data on 100 momentum-sorted portfolios and compute the decompositions using four asset pricing models: CAPM, Fama and French (1993) three-factor model, Fama and French (2015) five-factor model, and a seven-factor model that augments the Frazzini et al. (2018) model with the profitability and investment factors and removes the momentum factor. We compute standard errors by block bootstrapping the data by month 1,000 times. When month t is sampled, we link month t returns with factors' average returns from month $t-12$ to $t-2$ and month t estimates of portfolio betas. For each simulated sample we recompute the factor autocovariance and beta covariance matrices, the variance of average portfolio returns, and the total return on the momentum strategy. We compute the first three terms of the decomposition using this information. The last term, residual autocovariances, is the difference between the total strategy return and the sum of the first three components.

	Asset pricing model			
Component	CAPM	FF3	FF5	FF5 + BAB + QMJ
Factor autocovariances	0.00	0.09	0.16	0.40
	[0.03]	[0.07]	[0.13]	[0.22]
Factor cross-covariances		−0.03	−0.02	−0.09
		[0.02]	[0.03]	[0.27]
Variance of mean returns	0.11	0.11	0.11	0.11
	[0.08]	[0.08]	[0.08]	[0.08]
Residual autocovariances	0.58	0.52	0.43	0.28
	[0.21]	[0.18]	[0.17]	[0.18]
Total	0.69	0.69	0.69	0.69
	[0.26]	[0.26]	[0.26]	[0.26]

minus-factors. This last model is Frazzini et al.'s (2018) model without the momentum factor but augmented with the profitability and investment factors. We decompose momentum profits using each model and block bootstrap the data 1,000 by month to compute standard errors.

Table A6 shows a "tradeoff" between factor and residual autocovariances. As we increase the number of factors, the estimate of factor autocovariances increases from 0.00% in the CAPM to 0.40% in the seven-factor model. At the same time, the amount of profits attributed to residual autocovariances decreases from 0.58% to 0.28%. This pattern shows the limitations of using a decomposition such that in equation (A-2) to draw inferences about the relative contributions of "factor" and "residual" momentum. Any factors omitted from the asset pricing model will be pushed into the *estimates* of firm-specific returns. The mechanics of this decomposition further complicate this issue: if the factor model is incomplete, the assumption that firm-specific innovations do not correlate with each other or factors is violated. In Section 4.5 of the paper we construct alternative residual momentum strategies and find that a large share of "residual" momentum may stem from omitted-factor momentum.

C Are more systematic factors more autocorrelated?

In Table A7 we measure how systematic each factor is by measuring their ability to explain variation in the cross section of stock returns. Because each factor is based on firm characteristics, we can estimate cross sectional regressions of monthly stock returns against these characteristics, one at a time, and record the adjusted R^2s.[31] We then compute the average and median R^2s across these cross-sectional regressions. A characteristic with a high average R^2 explains more of the comovement in stock returns. Because we estimate these regressions month by month, a characteristic associated with a high average R^2 does not necessarily explain more of the differences in *average* returns.[32]

Some characteristics explain more of the cross-sectional variation in stock returns than others. Table A7 shows that size, market beta, idiosyncratic volatility, and quality-minus-junk all explain, on average, more than 1% of the cross-sectional variation monthly returns. Other characteristics, such as accruals and Pástor and Stambaugh's (2003) liquidity betas, explain at most 0.2% of this variation. We report both average and median R^2s because the distributions are right-skewed.

Table A7 shows that cross-sectional R^2s align with factor autocorrelations. As in Table 2 we measure time-series variation in factor premiums by regressing month-t return on an indicator variable based on the average return over the prior 12 months. Factors based on characteristics that explain more of the cross-sectional variation are also more autocorrelated. Size, market beta, idiosyncratic volatility, and quality-minus-junk, for example, all show significant variation in factor premiums based on their past returns; factors such as accruals and liquidity, by contrast, do not. The correlations at the bottom of the table show that the association between the R^2s and autocorrelations is economically significant. Depending on whether we gauge autocorrelation with

[31] The quality-minus-junk factor of Asness et al. (2019) is based on three components: profitability, quality, and safety. In the quality-minus-junk regressions we include the three z-scores associated with these components.

[32] Lewellen (2015, p. 10, emphasis in original), for example, notes that "The [Fama-MacBeth] R^2 provides information mostly about the fraction of *contemporaneous* volatility explained by characteristic-based portfolios, not about the *predictive ability* of the characteristics."

Table A7: Are more systematic factors more autocorrelated?

We measure the extent to which time-series variation in factor premiums correlates with the degree to which the factor relates to systematic variation in stock returns. "Time-varying premium" is the slope from a univariate regressions in which the dependent variable is a factor's monthly return and the independent variable takes the value of one if the factor's average return over the prior year is positive and zero otherwise. "Characteristic R^2" is the average or median R^2 from cross-sectional regressions of monthly stocks returns against the characteristic on which the factor is based. The sample period starts in July 1963 and ends in December 2019.

	Time-varying premium		Characteristic R^2	
Factor	$\hat{\beta}$	$t(\hat{\beta})$	Mean	Median
Size	0.58	2.51	1.22%	0.37%
Value	0.41	1.78	0.71%	0.25%
Profitability	0.34	1.67	0.22%	0.01%
Investment	0.24	1.55	0.32%	0.11%
Accruals	0.10	0.65	0.17%	0.04%
Betting against beta	1.32	3.53	1.43%	0.52%
Cash-flow to price	0.24	1.16	0.38%	0.13%
Earnings to price	0.30	1.46	0.42%	0.14%
Liquidity	0.36	1.29	0.12%	0.03%
Long-term reversals	0.76	3.85	0.48%	0.12%
Net share issues	0.09	0.49	0.23%	0.05%
Quality minus junk	0.43	2.51	1.12%	0.53%
Residual variance	1.06	2.74	1.88%	0.58%
Short-term reversals	0.01	0.04	0.82%	0.21%
Correlations	$\hat{\rho}$	$t(\hat{\rho})$		
$\rho(\hat{\beta}, R^2_{\text{mean}})$	0.74	3.78		
$\rho(\hat{\beta}, R^2_{\text{median}})$	0.71	3.48		
$\rho(t(\hat{\beta}), R^2_{\text{mean}})$	0.56	2.33		
$\rho(t(\hat{\beta}), R^2_{\text{median}})$	0.58	2.44		

the point estimates or their t-values, or R^2s as the time-series averages or medians, the correlations range from 0.56 to 0.74. Because these correlations are so high, the alignment between R^2s and time-series variations in the premiums is statistically significant despite the small sample of just 14 factors. These results are consistent with the finding in Table 5 that high-eigenvalue factors—more systematic factors—exhibit more momentum.

D Individual stock momentum, residual momentum, and factor momentum in simulated economies

D.1 Processes

We simulate T months of returns for N stocks. Stock i's excess return in month t is

$$R_{i,t} = \sum_{f=1}^{F} \beta_i^f r_t^f + \varepsilon_{i,t}, \tag{A-3}$$

where F is the number of factors, β_i^f is stock i's loading against factor f, r_t^f is factor f's return and $\varepsilon_{i,t}$ is a firm-specific return. We draw firm-specific returns from a normal distribution $N(0, \sigma_\varepsilon^2)$. We draw stocks' betas against each factor from a normal distribution, $\beta_i^f \sim N(\beta_0^f, \sigma_{\beta f}^2)$. All draws are uncorrelated across stocks and factors.

Factor f's return in month t is

$$r_t^f = \lambda_t^f + \eta_{f,t}, \tag{A-4}$$

in which λ_t^f is the factor's risk premium in month t and $\eta_{f,t}$ is a factor-specific innovation drawn from a normal distribution $N(0, \sigma_{\eta,f}^2)$.

A factor's risk premium follows an AR(1) process with a zero mean,

$$\lambda_t^f = \kappa_f \lambda_{t-1}^f + \omega_{f,t}, \tag{A-5}$$

in which κ_f controls the persistence of the risk premium process and $\omega_{f,t}$ is the shock to the process drawn from a normal distribution $N(0, \sigma_{\omega,f}^2)$. We draw the initial value of λ_0^f from the stationary distribution of the risk-premium process, $N\Big(0, \sigma_{\omega,f}^2/(1-\kappa_f^2)\Big)$.

D.2 Parameter values

We use the following parameters to simulate factor and stock returns in the "Symmetric factors" specification of Table 8:

- Number of months, stocks, and factors: $T = 672$ months, $N = 2{,}000$ stocks, and $F = 10$ factors.
- Distributions of betas: $\beta_0^f = 0$ and $\sigma^2_{\beta^f} = 1$ for all factors.
- Persistence of the risk premium process: $\kappa_f = 0.9$ for all factors.
- We normalize the total variance of stock returns to $\text{var}(r_{i,t}) = 1$ and assume that:
  - 40% of the variation is due to factors and the remaining 60% to firm-specific returns
  - Each factor contributes equally to the total variation in stock returns
  - 5% of the variation in factor returns is due to variation in the risk premium and the remaining 95% to factor-specific innovations, $\eta_{f,t}$.
- With these proportionality assumptions, the variances of the shocks are: $\sigma^2_\varepsilon = 0.6$, $\sigma^2_{\omega,f} = 0.00038$, and $\sigma^2_{\eta,f} = 0.038$.

Because we compute t-values associated with the profits of the momentum strategies, the normalization assumption $\text{var}(r_{i,t}) = 1$ does not matter: both the average returns and variances scale with this assumption, and the scale cancels out. The values of $\sigma^2_{\omega,f}$ and $\sigma^2_{\eta,f}$ follow from the proportionality assumptions.

We modify the following assumptions to simulate factor and stock returns in the "Uncorrelated market factor" specification of Table 8:

- Distributions of betas: $\beta_0^{\text{mkt}} = 1$ for the market factor, $\beta_0^f = 0$ for factors $2, \ldots, F$, and $\sigma^2_{\beta^f} = 1$ for all factors.

- Persistence of the risk premium process: $\kappa_{\text{mkt}} = 0$ for the market factor and $\kappa_f = 0.9$ for factors $2, \ldots, F$.

- Market factor's contribution to the total variation in stock returns is five times as large as that of factors $2, \ldots, F$.

- The variances of the shocks are now: $\sigma_\varepsilon^2 = 0.6$, $\sigma_{\omega,\text{mkt}}^2 = 0.00357$, $\sigma_{\eta,\text{mkt}}^2 = 0.06786$, $\sigma_{\omega,f}^2 = 0.00027$ for $f = 2, \ldots, 10$, and $\sigma_{\eta,f}^2 = 0.02714$ for $f = 2, \ldots, 10$.

The other parameter values are the same as in the "Symmetric factors" specification.

D.3 Aggregation mechanism in simulations

In the simulated economy described above, firm-specific returns are uncorrelated but factor returns are positively autocorrelated because factor risk premiums follow AR(1) processes. The autocorrelation in factor returns transmits into the cross section of stock returns because stocks have different loadings against the factors. A portfolio of winner stocks, on average, loads positively on factors with high past returns and negatively on those with low past returns. For a portfolio of loser stocks the loadings are the opposite. A strategy that buys winner stocks and sells loser stocks therefore indirectly trades the momentum in factors.

Equation (18), which we derive under the assumption of a "linear" portfolio rule, suggests that the amount of momentum in individual stock returns depends on the *number* of factors: even if each factor is only mildly positively autocorrelated, the total amount of momentum in individual stock returns is large if there are many factors. We illustrate this aggregation mechanism in Table A8 by varying the number of factors in the economy (F) between one and fifty and the persistence of the risk premium process (κ) between 0.5 and 0.95. We keep the other parameters of the simulation at the values described in Section D.2. We consider an individual stock momentum strategy that buys the top decile of stocks and sells the bottom decile of stocks based on their prior-year returns. We

Table A8: Transmission of factor momentum into the cross section of stock returns in simulated economies

This table reports average t-values from simulations that measure the strength of individual stock momentum in an economy in which only factor returns are positively autocorrelated. We simulate 672 months of returns from a market with 2,000 stocks. Systematic factors and IID firm-specific innovations drive stock returns. All factors have the same variance, their risk premiums are equally persistent, and stocks' betas against these factors are mean zero. In this table the number of systematic factors ranges from one to fifty and the persistence of the factor risk premium process from 0.5 to 0.95. Appendix D.2 gives the values of the other parameters used in the simulations. The individual stock momentum strategy is long the top decile of stocks with the highest average returns over the prior 12 months and short the bottom decile. *Factor autocorrelation* is the average slope from a regression of factor's month t return against its average return from month $t-12$ to $t-1$. We report average t-values from 10,000 simulations.

Number of	Risk premium persistence, κ					
factors	0.50	0.60	0.70	0.80	0.90	0.95
1	0.28	0.44	0.69	1.12	1.85	2.41
2	0.47	0.69	1.00	1.60	2.66	3.50
5	0.70	1.08	1.64	2.60	4.26	5.64
10	1.03	1.50	2.30	3.63	6.01	7.92
20	1.45	2.14	3.21	5.03	8.36	11.00
50	2.12	3.16	4.79	7.53	12.53	16.53
Factor autocorrelation	0.01	0.03	0.06	0.12	0.20	0.26

compute t-value associated with the average return of this long-short strategy, repeat the simulation 10,000 times, and report the average t-values.

Table A8 shows that the amount of momentum in the cross section of individual stocks increases both in the number of factors and the persistence in factor returns. When the risk premiums are the least persistent, $\kappa = 0.5$, "UMD's" t-value in the average simulation is above 2.0 only when there are $F = 50$ factors. In this case factor returns are only weakly positively autocorrelated; the last row shows that the correlation between factors' month t returns and their average returns from month $t-12$ to $t-1$ is just 0.01. As we increase the persistence of the risk premiums, momentum grows stronger. For example, when $\kappa = 0.95$, individual stock momentum strategy is statistically significant at the 5% level in the average simulation even if there is only one systematic factor in the economy.

E Residual momentum and bets against betas

Table 9 reports estimates from univariate regressions that show that factor momentum subsumes residual momentum strategies. The alphas are economically small and statistically insignificant. Table A9 shows, first, that this result holds when controlling for additional factors and, second, that residual momentum strategies are profitable for reasons *unrelated* to momentum.

In Table A9 we report these strategies average returns (first row) and alphas from three asset pricing models:

M1: Fama-French five-factor model.

M2: Model 1 augmented with a factor momentum strategy that trades the first ten PC factors.

M3: Model 2 augmented with five betting-against-beta factors.

We construct the betting-against-beta factors in Model 3 by sorting stocks into six portfolios based on size and one of the beta estimates. We use the same five-factor model beta estimates we use to compute residual returns. A betting-against-beta factor is long the two low-beta portfolios and short the two high-beta portfolios. The first of these factors is therefore similar to the BAB factor of Frazzini and Pedersen (2014).

Panel A of Table A9 shows that residual momentum strategies' five-factor model alphas exceed their average returns. The five-factor model residual momentum strategy, for example, earns a five-factor model alpha of 51 basis points per month (t-value $= 4.56$) whereas this strategy's average return is only 33 basis points (t-value $= 2.76$). This difference implies that the residual momentum strategy, on average, has significant negative loadings against the five factors of the five-factor model. These negative loadings emerge naturally because residuals must be estimated from the data. A stock's five-factor model residual, for example, is computed as

$$\hat{e}_{i,t} = r_{i,t} - r_{f,t} - \left[\hat{\beta}_{i,t}^{\text{mktrf}}\text{MKTRF}_t + \hat{\beta}_{i,t}^{\text{smb}}\text{SMB}_t + \cdots + \hat{\beta}_{i,t}^{\text{cma}}\text{CMA}_t\right], \tag{A-6}$$

Table A9: Residual momentum versus factor momentum: Alternative factor models

Panel A reports average returns (first row) and alphas (other rows) for UMD-style individual stock momentum strategies. We sort stocks into portfolios either by their raw returns from month $t-12$ to $t-2$ or by their CAPM, FF3, or FF5 residuals. The independent variable is a factor momentum strategy based on the first ten high-eigenvalue PC factors. The regressions differ in the set of additional factors included in the asset pricing model. These additional factors are the five factors of the Fama-French five-factor model and the betting-against-beta versions of these factors. A betting-against-beta factor is long low-beta stocks and short high-beta stocks. We construct these factors using the same methodology as for HML. The estimated betas are the same as those used to compute residual returns. t-values are in parentheses. Panel B reports average value-weighted betas for the winner and loser portfolios and for their difference. The winner portfolio, for example, is the equal-weighted average of the small-winner and big-winner portfolios that underlie the FF5 residual momentum strategy. The sample begins in July 1973 and ends in December 2019.

Panel A: Average returns and alphas of individual stock and residual momentum strategies

Sorting variable	$\hat{\alpha}$	$t(\hat{\alpha})$	Factor momentum $\hat{b}$	Factor momentum $t(\hat{b})$	Additional factors FF5	Additional factors FF5 BAB	R^2
Raw	0.42	2.44			N	N	.
returns	0.48	2.82			Y	N	10.1%
	−0.18	−1.26	3.61	17.59	Y	N	42.3%
	−0.23	−1.65	3.42	16.82	Y	Y	46.1%
CAPM	0.55	3.74			N	N	.
residuals	0.67	4.57			Y	N	10.8%
	0.11	0.90	3.04	17.25	Y	N	42.0%
	0.04	0.30	2.82	16.87	Y	Y	50.0%
FF3	0.41	3.22			N	N	.
residuals	0.56	4.64			Y	N	17.0%
	0.18	1.65	2.09	13.13	Y	N	36.6%
	0.09	0.82	1.90	12.48	Y	Y	44.8%
FF5	0.33	2.76			N	N	.
residuals	0.51	4.56			Y	N	19.0%
	0.20	1.87	1.74	11.39	Y	N	34.3%
	0.07	0.74	1.52	10.60	Y	Y	44.4%

in which the betas are estimated using historical data as described in Section 4.5.2. A residual momentum strategy is therefore long stocks with high returns *and* low betas and short stocks with low returns and high betas.

Panel B of Table A9 shows that these unintended bets against betas are economically important and statistically significant. We report value-weighted betas for the winner and loser portfolios that

Panel B: Estimated betas when sorting stocks into portfolios by FF5 residuals

Estimated portfolio betas	Residual momentum portfolio Winners	Losers	Winners − Losers	t-value
Market, $\hat{\beta}^{\text{mktrf}}$	0.97	1.17	−0.19	−19.90
Size, $\hat{\beta}^{\text{smb}}$	0.46	0.48	−0.02	−1.59
Value, $\hat{\beta}^{\text{hml}}$	0.08	0.15	−0.07	−2.91
Profitability, $\hat{\beta}^{\text{rmw}}$	−0.28	0.22	−0.50	−20.40
Investment, $\hat{\beta}^{\text{cma}}$	−0.22	0.10	−0.32	−11.58

underlie the FF5 residual momentum strategy, and the winner-minus-loser difference. The winner portfolio, for example, on average, is long stocks with an estimated market beta of 0.97; the stocks in the loser portfolio, by contrast, have an average market beta of 1.17. The time-series difference between these betas is significant with a t-value of −19.90; that is, the residual momentum strategy embeds a significant bet against market betas. The market factor is not an exception. The bets against the profitability and investment factors are particularly pronounced. We know from Black (1993) that the bets against market betas are profitable on a market-adjusted basis, and from, for example, Daniel et al. (2020) that the bets against the other four betas also generate significant abnormal returns when adjusted for the five-factor model.[33]

The other factor models in Panel A of Table A9 show that the betting-against-beta effect is indeed important. Model 2, when applied to the strategy that selects stocks based on the residuals from the five-factor model, leaves this strategy with an alpha of 20 basis points (t-value = 1.87). This 20-basis point alpha is almost the same as the 19-basis point difference between the strategy's five-factor model alpha (51 basis points) and its average return (33 basis points). When we augment the five-factor model with the betting-against-beta factors, much of the remaining alpha vanishes; it is now just 7 basis points (t-value = 0.74). Although residual momentum strategies

[33] A betting-against-beta strategy, constructed as an HML-style factor that is long low market-beta stocks and short high market-beta stocks, earns an average return of 13 basis points per month (t-value = 0.82) between July 1973 and December 2019. This strategy's CAPM alpha, however, is 47 basis points per month (t-value = 4.22) because its loading against the market factor is −0.64. (We use the same methods as in Table A9 to estimate stocks' market betas.) That is, the betting-against-beta strategy is profitable on a *market-adjusted* basis. Frazzini and Pedersen (2014) render their betting-against-beta factor approximately market neutral by levering the long and short legs up and down to have betas of one.

could be profitable because *firm-specific* returns display momentum, they significantly benefit from the betting-against-beta effect.[34] The estimates in Table A9 show that once we control for the betting against beta effect—to absolve the residual momentum strategy from these incidental beta bets—residual momentum is not profitable net of factor momentum.

[34] Ehsani and Linnainmaa (2020b) expand on this analysis in a companion paper to this work. We analytically decompose the profits that residual momentum strategies earn into three components: (1) a positive bet on momentum, (2) a positive bet on the betting-against-beta strategy, and (3) a negative bet on momentum in the factors included in the model. When the residual momentum strategies are rendered beta-neutral already in the portfolio formation stage, the pattern in the five-factor model alphas disappears.

F Constructing momentum-neutral factors

In this appendix we define and compute momentum-neutral factor weights. We also show that these weights are equivalent to the residuals from a regression of the original factor weights on prior stock returns. We start from the original factor weights w_i from equation (13). These weights are proportional to stock characteristic c and sum to zero: $\sum_{i=1}^{N} w_i = 0$. We define momentum-neutral factors as factors computed using weights x_i that are: (1) as close as possible in terms of mean-squared errors to the original weights and (2) orthogonal to individual stock returns over the prior year skipping a month, $r_{i,t-12,t-2}$. The objective is to find new weights x_i such that

$$\min_{x_i} \frac{1}{2} \sum_i (w_i - x_i)^2 \quad \text{s.t.} \sum_i x_i = 0 \text{ and } \sum_i x_i r_{i,t-12,t-2} = 0. \tag{A-7}$$

The Lagrangian to this problem is,

$$\mathcal{L} = \frac{1}{2} w'w - x'w + \frac{1}{2} x'x + \lambda_0 x'\mathbf{1} + \lambda_1 x'r, \tag{A-8}$$

where w is an $N \times 1$ vector of "old" weights, x is an $N \times 1$ vector of "new" weights, $\mathbf{1}$ is an $N \times 1$ vector of ones, r is an $N \times 1$ vector of past returns $r_{i,t-12,t-2}$, and λ_0 and λ_1 are the Lagrange multipliers. The first-order conditions are

$$\frac{\partial \mathcal{L}}{\partial x} = x - w + \lambda_0 \mathbf{1} + \lambda_1 r = 0, \tag{A-9}$$
$$\frac{\partial \mathcal{L}}{\partial \lambda_0} = x'\mathbf{1} = 0, \text{ and} \tag{A-10}$$
$$\frac{\partial \mathcal{L}}{\partial \lambda_1} = x'r = 0. \tag{A-11}$$

The solutions for λ_0 and λ_1 are:

$$\lambda_0 = -\frac{1}{N} r'\mathbf{1} \frac{w'r}{r'r - \frac{1}{N}(r'\mathbf{1})^2} \quad \text{and} \quad \lambda_1 = \frac{w'r}{r'r - \frac{1}{N}(r'\mathbf{1})^2}. \tag{A-12}$$

The weights of the momentum-neutral portfolio are then

$$x = w - \lambda_0 \mathbf{1} - \lambda_1 r. \tag{A-13}$$

The definitions of λ_0 and λ_1 are the same as the estimates of the intercept and the slope from a univariate regression of portfolio weights w_i on past stock returns:

$$w_i = \beta_0 + \beta_1 r_i + x_i. \tag{A-14}$$

The estimated slope from this regression, for example, is

$$\hat{\beta}_1 = \frac{\text{cov}(w_i, r_i)}{\text{var}(r_i)} = \frac{\frac{1}{N}\sum_{i=1}^N w_i r_i - \frac{1}{N^2}\sum_{i=1}^N w_i \sum_{i=1}^N r_i}{\frac{1}{N}\sum_{i=1}^N r_i^2 - \frac{1}{N^2}(\sum_{i=1}^N r_i)^2} = \frac{w'r}{r'r - \frac{1}{N}(r'\mathbf{1})^2} = \lambda_1, \tag{A-15}$$

where $\sum_{i=1}^N w_i = 0$ because the original weights are long-short weights. The weights in equation (A-13) are therefore the same as those as the residuals from the regression in equation (A-14).