Five-factor residual momentum: a comparative analysis

José Gonçalo Maia Barbosa Valente Teixeira

Dissertation written under the supervision of professor Pedro Barroso, Ph.D.

Dissertation submitted in partial fulfilment of requirements for the MSc in Finance, at the Universidade Católica Portuguesa, May 2023.

Five-factor residual momentum: a comparative analysis

José Gonçalo Maia Barbosa Valente Teixeira

Abstract

Momentum strategies entail taking on time-varying exposures to the five factors of Fama and French. We show that by ranking stocks on the portion of their return not explained by such factors (residual return) improves momentum's Sharpe ratio and results in less volatile performance. By comparing residual momentum with standard momentum we find that neither strategy can be explained by priced risk factors or by industry effects, but the residual momentum strategy subsumes traditional momentum. Not all of the Fama and French factors contribute equally to residual momentum and the simpler version of the strategy delivers the best performance. Our results also show that the profitability of standard momentum is not explained by the risk of factor timing.

Keywords Anomalies, Momentum, Residual returns, Time-varying risk

Momentum residual com cinco fatores: uma análise comparativa

José Gonçalo Maia Barbosa Valente Teixeira

Resumo

As estratégias baseadas em *momentum* implicam a assunção de uma exposição aos cinco fatores de Fama e French que varia no tempo. Nós mostramos que ao ordenar ações com base na componente do seu retorno que não é explicada por estes fatores (retornos residuais) é possível melhorar o rácio de Sharpe do *momentum* e resulta num desempenho menos volátil. Ao comparar o *momentum* residual com a versão tradicional de *momentum* verificamos que nenhuma das estratégias pode ser explicada por fatores de risco conhecidos nem pelo desempenho de indústrias como um todo, mas que o *momentum* residual explica o *momentum* tradicional. Nem todos os fatores de Fama e French contribuem igualmente para o desempenho do *momentum* residual e a versão mais simples da estratégia é a que tem o melhor desempenho. Os nossos resultados também indicam que o desempenho do *momentum* tradicional não é explicado pelo risco de tentar implementar o *timing* de fatores.

Palavras-chave Anomalias, *Momentum*, Retornos residuais, Risco variável no tempo

Acknowledgements

Firstly, I would like to thank my supervisor, Pedro Barroso for his guidance and for the thoughtfulness with which he reviewed my work. To my friends, I appreciate all the fun moments and conversations that helped me take my mind off things during this rather stressful past few months. Finally, to my family and especially my parents for all the love, patience, support and for always giving me the freedom to pursue my own choices.

Contents

1 Introduction

Some recurring patterns in average stock returns are traditionally deemed to be anomalies if they are not explained by the cornerstone model of classical finance theory, the Capital Asset Pricing Model (CAPM) developed by Treynor (1962), Sharpe (1964) and Lintner (1965), which describes the behavior of capital markets in a one-period equilibrium setting where investors follow the mean-variance optimization framework of Markowitz (1952). One such anomaly is momentum, which consists of the persistence of asset returns in the cross-section[1]. This anomaly has garnered the attention of researchers because of the challenge it poses to the weak form of market efficiency described by Fama (1970).

The seminal work of Jegadeesh and Titman (1993, 2001) documented the profitability of momentum in U.S. equity markets by analyzing the performance of long-short portfolios that buy the best performing stocks and sell the laggards. Despite the positive returns that this strategy delivers its implementation exposes the investor to a particular type of risk that makes momentum unique among asset pricing anomalies. As Grundy and Martin (2001) have shown, if the performance divergence between winners and losers is driven by the relative performance of cross-sectional risk factors, then momentum can be interpreted as a strategy that times the performance of other systematic equity strategies like size or value, in which case the high returns earned by momentum might be understood as compensation for bearing this factor timing risk.

To analyze this issue we develop a comparative analysis between the standard momentum strategy and a residual momentum strategy that sorts the winners from the losers based on residual (or idiosyncratic) returns relative to the five-factor model of Fama and French (2015). This builds on previous work by Blitz, Huij, and Martens (2011) and Blitz, Hanauer, and Vidojevic (2020) who conduct similar research using the three-factor model of Fama and French (1993).

We find that the five-factor residual momentum strategy delivers better risk-adjusted perfor-

[1] Moskowitz, Ooi, and Pedersen (2012) investigate the time-series momentum phenomenon, which focuses on the persistence of an asset's own past return.

mance compared to traditional momentum when considering volatility as the single measure of risk, but it fails to meaningfully reduce the well known crash risk of momentum. Residual momentum's performance is inferior when controlling for the performance of factor-mimicking portfolios, but after introducing interaction terms that capture the recent performance of each factor into the model, we find that the performance of both strategies is very similar. Interestingly, raw return momentum still exhibits positive and statistically significant returns in this later specification, suggesting that factor timing risk does not explain its profitability. We also show that not all factors are equally relevant to the residual momentum strategy. Changing our original procedure by building two simpler variants (one that controls only for the market factor, and another that controls for all three of the Fama and French (1993) factors) we conclude that less is more: the best performing of all momentum strategies ranks stocks on their idiosyncratic returns relative to the CAPM, and performance deteriorates as we add more risk factors. In addition, we show that the profitability of both raw return and five-factor residual return momentum are not explained by momentum across industries. Finally, analyzing the performance of both strategies by calendar month we show that residual momentum is less prone to the January effect and that its performance is much more stable across the year.

The remainder of the paper is structured as follows: Section 2 expands on the relevant literature; Section 3 describes the data sources and methodology for the construction of our portfolios; Section 4 presents our empirical methodology and the conclusions that can be drawn from our statistical work; Section 5 presents three modifications for robustness that stress test some of our key assumptions; Section 6 concludes.

2 Literature Review

While the first empirical analysis of momentum in U.S. equity markets dates as far back as the work of Levy (1967) on the performance of stock picking using relative strength signals, the paper that kickstarted much of the academic discussion around this phenomenon and developed the portfolio construction methodology that became standard practice in the literature was Jegadeesh and Titman (1993). This study comprises the period between 1965 and 1989, and in it the authors find that stocks that have performed better in the recent past tend to continue their outperformance, and vice-versa. Experimenting with different formation and holding periods for the construction of portfolios, they show that a zero-cost strategy that buys the winners and sells the losers earns returns that are, for the most part, statistically significant, with average monthly returns ranging from 0.32% to 1.49% depending on the periods used. Extending their analysis further into the late 1990s, Jegadeesh and Titman (2001) find that momentum continued to exhibit statistically significant profits.

The momentum phenomenon, however, is not limited to U.S. equity markets, or even to equity markets in general, for that matter. Rouwenhorst (1998, 1999) and Fama and French (2012) show the existence and behavior of momentum across a vast array of both developed and emerging markets[1]. Additionally, the profitability of momentum is shown to hold for currencies (Okunev and White (2003)), commodities (Erb and Harvey (2006)) and even across countries (Asness, Liew, and Stevens (1997)). Finally, Asness, Moskowitz, and Pedersen (2013) bring all of these strands of literature together in a very comprehensive analysis where they document the existence and examine the dynamics of value and momentum premia within and across various asset classes and geographies.

With all this in mind, we should caution that the pervasiveness of momentum does not make it a free lunch, and a simple empirical example will clarify this. If we download data on portfolios sorted on their past returns from Kenneth French's data library[2], we will see that

[1] With the notable exception of the Japanese market, for a deeper study of this particular case, see Asness (2011) and Hanauer (2014)

[2] This dataset can be found at https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html, under "10 Port-

the winners-minus-losers (WML) strategy delivered an average monthly return of 1.15% and a Sharpe ratio of 0.51 from January 1927 to September 2022. The problem with this simple analysis is that the Sharpe ratio is a performance measure designed for the Gaussian world of Markowitz (1952); therefore, it obscures risks that come from higher-order moments of the distribution of returns. Digging a little depper, we will find that the WML portfolio sports a kurtosis of 19.21, which is not surprising when we consider that in its worst month it suffered a collapse of 77%.

Given the latent danger associated with investing in momentum strategies, it should be come as no surprise that much of the literature on this topic has focused on discovering the source of this crash risk so it can be mitigated. Grundy and Martin (2001) laid much of the initial groundwork by showing that momentum strategies entail time-varying exposures to other risk factors. That is, momentum is unique amongst anomalies in the sense that it can be thought of as a factor-timing strategy to the extent that common factors are driving the performance of past winners and losers. To illustrate, consider, for example, the behavior of momentum following a bear market: the past winners will be, by definition, stocks with a low beta over the formation period and vice-versa for the losers; therefore, as stocks recover, a WML portfolio has a negative exposure to the market which causes it to crash. Daniel and Moskowitz (2016) adopt the framework of Merton (1974) in viewing equity as a call option on the firm's assets to explain this behavior. The high-beta stocks that the WML portfolio is short on following market drawdowns tend to be distressed firms, so the short leg of the portfolio is akin to writing out-of-the-money options on the assets of the underlying firms. As the market recovers, these options go into the money, and the convexity associated with this process lies behind the crash risk of momentum.

To manage this risk, Grundy and Martin (2001) propose hedging out the factor exposure of the momentum portfolio, thereby removing the factor timing dimension. The issue is that the strategy that they follow uses forward-looking betas, so it obviously could not be implemented by an investor in real time. Daniel and Moskowitz (2016) show that using ex-ante betas fails to reduce crashes meaningfully. Barroso and Santa-Clara (2015) explain this failure by decomposing the risk of momentum into market and idiosyncratic components. Using AR(1) regressions, they find that most of the risk of WML is not only specific to the strategy but also that this dimension of risk is highly predictable, unlike the systematic portion. Building

folios Formed on Momentum"

on this insight, they devise a highly effective strategy to reduce the crash risk by scaling the investment in the WML portfolio by its 6-month realized volatility relative to a constant volatility target. This decreases the kurtosis in their sample from 18.24 to just 2.68. Other variations of volatility-scaling have been applied to momentum: Daniel and Moskowitz (2016) developed a dynamically scaled version that uses a GARCH model to predict the volatility of momentum instead of using a constant target, and Wang and Yan (2021) consider only downside volatility when leveraging or deleveraging the WML portfolio. For a comparison of these different risk management strategies, see Hanauer and Windmüller (2023).

Gutierrez and Prinsky (2007) propose a different approach to reduce the dynamic factor tilts of momentum by ranking stocks on their residual returns relative to the three-factor model of Fama and French (1993) and observe that this eliminates the performance reversal that plain momentum suffers beyond the first year. Blitz, Huij, and Martens (2011) use this same strategy for risk reduction purposes and notice that a WML portfolio built on residual returns is less prone to crashes and exhibits half the volatility of raw return momentum, while Chaves (2016) shows that most of the reduction in risk can be obtained by curbing exposure to the market portfolio alone, which supports the framework of Daniel and Moskowitz (2016). Blitz, Hanauer, and Vidojevic (2020) conduct time-series GRS, Fama and MacBeth (1973) regressions and factor-spanning tests and conclude that other commonly used asset pricing factors do not explain this phenomenon. While the profitability associated with traditional momentum can be squared with the efficient markets hypothesis by interpreting it as adequate compensation for bearing the risk of dramatic crashes (analogous to the premium one receives for writing an option), the volatility scaled and residual versions of momentum present larger challenges.

3 Data and methodology

3.1 Data sources

The primary data source used in this study is the Center for Research in Security Prices (CRSP) monthly stock file. Consistent with most of the literature, we exclude closed-end funds, certificates, American Depository Receipts (ADRs), Shares of Beneficial Interest (SBIs), unit trusts, Real Estate Investment Trusts (REITs) and companies incorporated outside the U.S., therefore our investment universe comprises all common stocks (share codes 10 or 11) that traded on the NYSE, AMEX or NASDAQ. In addition, we exclude stocks when their price is below $1 to minimize microstucture issues. Data on the five factors of Fama and French (2015) and the one-month Treasury bill rate (which serves as a proxy for the risk-free rate) were downloaded from Kenneth French's online data library. Furthermore, given that we require 36 months of return data for the residual momentum calculation, we exclude stocks that do not have a complete return history over the 36-month rolling window when constructing both strategies (even though this is not a necessary step to build the total return momentum signal) to ensure that the performance differences between the total return and residual return strategies are not due to sample selection bias. The period used in our analysis starts in January 1969 and ends in December 2021. As we will explain below, we use a 36 month rolling window to estimate residual momentum, which means that the earliest possible start date would be July 1966 given that five factors return data are only available from July 1963 onward. We decided to extend the start data further in time to preserve the possibility of using longer rolling windows without changing the start date of our analysis[1]

[1] To reliably estimate factor betas we need a somewhat large sample, so it is not very reliable to use windows much shorter than 36 months.

3.2 Momentum calculation

3.2.1 Total return momentum

We construct total return momentum portfolios in a manner consistent with the most relevant literature (Jegadeesh and Titman (1993); Fama and French (1996)): to form a portfolio for month t, we rank stocks on their cumulative return from month t - 12 to month t - 2. We obviously cannot use a stock's return on month t because that would induce look-ahead bias, and we exclude the return in month t - 1 to avoid the short-term reversal effect documented by Jegadeesh (1990) and Lehmann (1990). Stocks are then grouped into deciles to form portfolios, with the top decile (portfolio 10) being the winners' portfolio and the bottom decile (portfolio 1) the losers' and held for one month until the portfolio is rebalanced again following the same methodology.

3.2.2 Five-factor residual momentum

The methodology for the calculation of portfolios ranked on residual momentum is conceptually the same as for total return momentum, with a few added twists. First, we regress each stock's excess return on the five-factor model of Fama and French (2015).:

$$r_{i,t} = \alpha_i + \beta_{1,i} RMRF_t + \beta_{2,i} SMB_t + \beta_{3,i} HML_t + \beta_{4,i} RMW_t + \beta_{5,i} CMA_t + \varepsilon_{i,t} \tag{1}$$

where $r_{i,\,t}$ is the return of stock i in month t in excess of the risk-free rate, $RMRF_t$, SMB_t, HML_t, RMW_t and CMA_t are the excess returns of the factor-mimicking portfolios in month t. Given that we need a relatively large sample to minimize the standard errors associated with the parameters' estimates, we use a 36-month rolling window (i.e., the model that is used to calculate the residual in month t considers data between months t - 36 and t - 1) as in Blitz, Huij, and Martens (2011) to fit the regression model.

The five-factor pricing model includes the market (RMRF), size (SMB) and value (HML) factors of Fama and French (1993) and adds two novel factors: robust-minus-weak (RMW) and conservative-minus-aggressive (CMA). They are both constructed as long-short portfolios that seek to capture the return spread between high-profitability and low-profitability stocks (RMW) and high-investment and low-investment stocks (CMA)[2].

[2] Profitability is defined as operating profit divided by book value of equity, while the proxy for investment is the percentage change in total assets. Both factors are constructed in the same way as HML, that is, the sorts are done within groups of stocks of similar size in order to disentangle the size factor from the factor that is being measured. For more details on factor construction, as well as for an in-depth analysis of the performance

The residual return is then calculated as the difference between each stock's realized excess return in month t and the excess return predicted by the regression model:

$$e_{i,t} = r_{i,t} - \widehat{\alpha_i} - \widehat{\beta_{1,i}} RMRF_t - \widehat{\beta_{2,i}} SMB_t - \widehat{\beta_{3,i}} HML_t - \widehat{\beta_{4,i}} RMW_t - \widehat{\beta_{5,i}} CMA_t \quad (2)$$

We noticed some differences in the literature regarding the calculation of residual momentum. For example, Blitz, Huij, and Martens (2011) exclude the alpha when estimating the residual return, arguing that, because the alpha is estimated using a 36-month window, which extends further back than the 11-month formation period, it is capturing not only the intermediate-term momentum phenomenon but also the long-term reversal effect of DeBondt and Thaler (1985). On the other hand, Gutierrez and Prinsky (2007) and Blitz, Hanauer, and Vidojevic (2020) include the alpha. We decided to follow this latter approach to remain consistent with the most recent literature.

Finally, to obtain the momentum score, we standardize the cumulative 12 - 1M residual return by the cumulative volatility of monthly residuals over the same period to remain consistent with Gutierrez and Prinsky (2007), Blitz, Huij, and Martens (2011) and Blitz, Hanauer, and Vidojevic (2020). This adjustment is made to better isolate the portion of the firm-specific return that results from news instead of noise. The residual momentum score for stock i in month t is therefore given by:

$$ResidualMomentumScore_{i,t} = \frac{\sum_{t-12}^{t-2} e_{i,t}}{\sqrt{\sum_{t-12}^{t-2} (e_{i,t} - \overline{e_i})^2}} \quad (3)$$

After obtaining the residual momentum scores, stocks are again sorted into deciles and held for one month, just like in the total return momentum strategy.

3.3 Portfolio construction

After sorting stocks into deciles according to their momentum scores, we must decide how to weigh each of the stocks. The common practice in the residual momentum literature (Gutierrez and Prinsky (2007) and Blitz, Huij, and Martens (2011)) is to equal-weight the stocks in each portfolio after sorting them using NYSE-AMEX-NASDAQ return breakpoints. This is not, in our view, a proper approach because, for the conclusions of empirical asset pricing studies to

of the five-factor model in describing the cross-section of stock returns, see Fama and French (2015).

be valid, the results of the strategy under investigation should be as similar as possible to the results that an investor would have obtained from a real-time implementation. The goal of equal weighting is to reduce the dominance of a few very large companies that might result from a value-weighted implementation. Even though this might be a valid concern, by equally weighting every stock, one incurs the risk of having results that are reflective of the performance of tiny companies, which are 3% of total market capitalization but account for 60% of the total number of companies (Fama and French (2008)). As such, a strategy that equally weights stocks from the NYSE-AMEX-NASDAQ universe is, at best, severely capacity-constrained or, at worst, not implementable at all due to the very low liquidity in microcaps.

The question of whether empirical studies produce results divorced from those that could be obtained in the real world sparked discussions of a "replication crisis" in finance (Harvey (2017)), where concerns arose about whether researchers are incentivized to produce statistically significant results, even if it comes at the cost of producing findings that cannot be replicated or implemented in actual markets. One way through which this form of "p-hacking" can be achieved in cross-sectional equity studies is precisely by having results that depend heavily on the performance of small, illiquid stocks. Hou, Xue, and Zhang (2020) conducted a comprehensive study attempting to replicate 452 anomalies after controlling for microcaps via value-weighted returns, portfolio sorts with NYSE breakpoints, and excluding stocks with a market capitalization below the 20th percentile NYSE market capitalization. They find that 65% of anomalies fail to show any profitability at a 5% significance level. Even though the residual momentum of Blitz, Huij, and Martens (2011) is amongst the anomalies that successfully replicate, we decided to still exclude stocks with a capitalization below the NYSE's 20th percentile[3]. This filter, in addition to the ones mentioned in subsection 3.1, results in an investment universe of 926 stocks in January 1969 and 1,544 in December 2021. We will still equal-weight our portfolios to stay consistent with the literature on residual momentum, but we will stress test this decision further in our section on robustness tests, where we will impose an additional restriction on our investment universe to include only the larger and most liquid stocks.

[3] Data on the NYSE's market capitalization percentiles are available on Kenneth French's data library.

4 Empirical results

In this section we will present the main results of our empirical analysis. We will begin by comparing the risk and return profiles of the total return and residual return momentum strategies by analyzing their descriptive statistics and performance when controlling for the returns of factor-mimicking portfolios. We will then further our assessment by accounting for other risk factors not usually included in traditional asset pricing models, that is, the time-varying nature of the risk of those factors; next, we shall investigate which factors contribute the most to the performance differential between total return and residual return by separately residualizing returns relative to the CAPM and the three-factor model of Fama and French (1993) and comparing these results to those delivered by our baseline residual momentum strategy, which residualizes returns in relation to the five-factor model of Fama and French (2015).

4.1 Descriptive statistics and cumulative returns

To start the analysis, we report in Table 1 descriptive statistics of the decile portfolios formed on total return and five-factor residual return momentum over the 1969-2021 period.

We can observe that, for both strategies, there is an essentially monotonic increase in the average annualized return as we move up from the losers' portfolio (P1) to that of the winners (P10). Even though the average annualized return for the zero-cost P10-P1 residual momentum portfolio (7.50%) is slightly lower than that of the total momentum, the decrease in volatility from residualization more than offsets that, thus resulting in a considerably higher Sharpe ratio. Furthermore, residualizing returns virtually eliminates the negative skew associated with plain momentum. Finally, there is a decrease in kurtosis, which indicates that residual momentum has a moderately more negligible crash risk than plain momentum. However, from a risk management standpoint, it does not appear to be as effective in this regard as the volatility scaling approaches of Barroso and Santa-Clara (2015), Daniel and Moskowitz (2016) and Wang and Yan (2021). This is most likely explained by the findings of Barroso and Santa-Clara (2015), which show that the systematic component of momentum risk is very unpredictable. This means that residualizing returns with real-time betas is not a very effective strategy because there is a very low correlation between *ex-ante* and *ex-post* betas[1] as Barroso (2016) has shown.

Table 1. Descriptive statistics of decile portfolios

This table reports descriptive statistics of decile portfolios constructed as univariate sorts on total (WML) and residual return momentum (rWML). We include all common stocks (CRSP share code 10 or 11) that traded on the NYSE, AMEX or NASDAQ (CRSP exchange codes 1, 2 or 3) from January 1969 to December 2021 with share prices above $1 and market capitalization above the 20th percentile of NYSE listed companies. Total return momentum is calculated as the 12M-2M cumulative return. Residual momentum is the volatility scaled 12M-2M residual return over the past 36 months using the five factor model of Fama and French (2015). For each portfolio we report the maximum and minimum monthly returns, the average annualized return, annualized volatility, skewness, kurtosis and Sharpe ratio. Returns are presented in percentage and portfolios are equal weighted with rebalancing done on a monthly basis.

		Maximum	Minimum	Average Return	Volatility	Skewness	Kurtosis	Sharpe Ratio
WML	P1	55.63	-30.33	7.14	27.04	0.66	8.56	0.26
	P2	38.11	-26.74	12.15	21.36	0.13	7.20	0.57
	P3	29.91	-26.03	12.61	18.96	-0.12	7.15	0.67
	P4	22.53	-25.00	12.89	17.58	-0.32	6.06	0.73
	P5	25.15	-26.40	13.89	16.95	-0.45	6.96	0.82
	P6	22.21	-26.41	13.91	16.61	-0.56	6.38	0.84
	P7	20.95	-27.97	14.13	16.58	-0.70	6.66	0.85
	P8	18.95	-29.46	14.08	17.33	-0.80	6.99	0.81
	P9	20.11	-28.60	15.29	18.17	-0.75	6.18	0.84
	P10	26.36	-33.09	16.10	22.19	-0.64	6.02	0.73
	P10 - P1	24.09	-50.65	8.95	20.28	-1.54	13.88	0.44
rWML	P1	30.12	-24.54	8.98	21.16	-0.03	4.95	0.42
	P2	26.58	-23.99	10.89	19.80	-0.28	5.25	0.55
	P3	25.60	-25.28	12.15	19.07	-0.29	5.67	0.64
	P4	27.89	-25.75	12.60	18.71	-0.32	5.64	0.67
	P5	24.86	-26.81	13.25	18.50	-0.45	6.09	0.72
	P6	18.10	-27.50	13.66	18.21	-0.65	5.99	0.75
	P7	24.68	-28.64	15.13	18.08	-0.54	6.07	0.84
	P8	21.94	-27.78	14.88	18.46	-0.62	5.79	0.81
	P9	19.59	-28.34	14.86	18.58	-0.61	5.60	0.80
	P10	21.54	-29.16	16.48	19.73	-0.42	5.20	0.84
	P10 - P1	21.57	-21.21	7.50	12.11	-0.00	9.26	0.62

In Figure 1, we plot the cumulative returns over our sample for the total return and residual return strategies, as well as for the market risk premium[2]. We use the market premium instead of the raw market return because it is a zero-cost strategy, like both momentum strategies. In this case, the long leg of the portfolio is the CRSP value-weighted portfolio, and the short leg is equivalent to borrowing at the 1-month Treasury bill rate. An alternative approach for

[1] The analysis of Barroso and Santa-Clara (2015) focuses only on market risk, but given that the five-factor residual momentum does not mitigate crash risk it is probable that this explanation holds for other factors as well.

[2] Data for the market risk premium is available on Kenneth French's data library. The market risk premium is calculated as the return of the CRSP value-weighted portfolio minus the one-month Treasury bill rate.

comparison purposes would be to calculate the return of the momentum portfolios as fully-funded portfolios and compare those against the market return.

We can see that momentum portfolios have beaten the market since the beginning of our sample, although this performance differential has narrowed substantially since the Global Financial Crisis, with momentum languishing as the market entered the last decade's bull market.

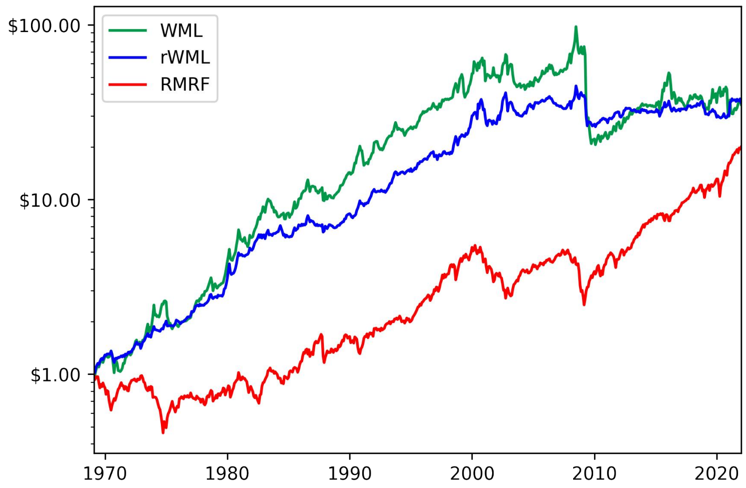

Figure 1. Cumulative returns
This figure shows the cumulative return of long-short portfolios that invest in total return momentum, residual momentum and the market (funded at the risk-free rate). Residual momentum is calculated using residual returns relative to the five-factor model of Fama and French (2015) using a 36-month rolling window. Both momentum strategies use the traditional 12M - 2M formation period. We consider all stocks that traded on the NYSE, AMEX or NASDAQ between January 1969 and December 2021 with a share price above $1 and a market capitalization above the 20th percentile market capitalization of companies listed on the NYSE. Momentum portfolios are equal weighted and rebalancing is done on a monthly basis.

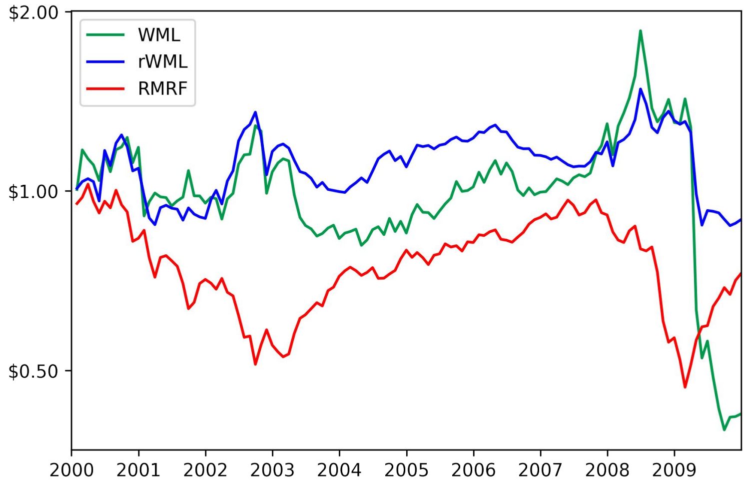

Figure 2. Cumulative returns between 2000-2009
This figure shows the cumulative return of long-short portfolios that invest in total return momentum, residual momentum and the market (funded at the risk-free rate) over the 2000-2009 period. Residual momentum is calculated using residual returns relative to the five-factor model of Fama and French (2015) using a 36-month rolling window. Both momentum strategies use the traditional 12M - 2M formation period. We consider all stocks that traded on the NYSE, AMEX or NASDAQ between January 1969 and December 2021 with a share price above $1 and a market capitalization above the 20th percentile market capitalization of companies listed on the NYSE. Momentum portfolios are equal weighted and rebalancing is done on a monthly basis.

Moving on to Figure 2, we zoom in on the decade leading up to the crisis and the associated bear market, and the behavior we observe matches the findings of Daniel and Moskowitz (2016): momentum's crash coincides with recovery from the bear market's nadir and even though the residual momentum portfolio is not at all immune to crash risk, it appears that ranking stocks on their residual returns somewhat mitigates the extent of the downturn, again supporting Daniel and Moskowitz (2016)'s thesis, which ties the crash risk to factor reversal. Another interesting observation we can draw from this figure is that during the bear market of 2000-2003 that followed the dot-com bust, the market's drawdown was almost as severe as the one it suffered during the Global Financial Crisis. The difference being that both the crash and the recovery in 2008-2009 occurred in a much more compressed timeframe compared to the events of 2000-2003. It therefore appears that if big market moves take place over a longer period of time, momentum is not as prone to crashes, perhaps because its factor positioning is not as out of step with subsequent factor performance if crashes and recoveries take longer to play out.

4.2 Asset spanning regressions

In this subsection, we run asset-spanning tests to assess whether total return and residual momentum can be considered separate factors that enhance the efficient frontier or whether other known factors subsume their behavior.

We regress the returns of both momentum strategies on three separate asset pricing models: (i) the Capital Asset Pricing Model (CAPM) of Sharpe (1964) and Lintner (1965); (ii) the three-factor model of Fama and French (1993); (iii) the five-factor model of Fama and French (2015).

Table 2. Asset spanning regressions of decile portfolios
This table reports the performance of decile portfolios constructed as univariate sorts on total (WML) and residual return momentum (rWML). We include all common stocks (CRSP share code 10 or 11) that traded on the NYSE, AMEX or NASDAQ (CRSP exchange codes 1, 2 or 3) from January 1969 to December 2021 with share prices above \$1 and market capitalization above the 20th percentile of NYSE listed companies. Total return momentum is calculated as the 12M-2M cumulative return. Residual momentum is the volatility scaled 12M-2M residual return over the past 36 months using the five factor model of Fama and French (2015). For each portfolio we report monthly alphas (in percentage) relative to the CAPM, the three factor model of Fama and French (1993), the five factor model of Fama and French (2015) and associated t-statistics. We also show the coefficients associated with each of the five factors. To account for heteroscedasticity, t-stats are computed using White (1982) standard errors. Portfolios are equal weighted and rebalancing is done on a monthly basis. **Note:** *, ** and *** denote statistical significant results at a 5%, 1% and 0.1% level, respectively.

		CAPM α	T-stat	3FF α	T-stat	5FF α	T-stat	β_{RMRF}	β_{SMB}	β_{HML}	β_{RMW}	β_{CMA}
WML	P1	-0.20	-1.12	-0.42	-2.86	-0.41	-2.43	1.32^{***}	0.70^{***}	0.65^{***}	0.07	-0.15
	P2	0.35	2.73	0.16	1.59	0.09	0.77	1.15^{***}	0.51^{***}	0.51^{***}	0.27^{***}	-0.07
	P3	0.44	4.28	0.27	3.41	0.16	1.94	1.07^{***}	0.42^{***}	0.46^{***}	0.31^{***}	-0.07
	P4	0.50	5.53	0.35	5.09	0.23	3.31	1.02^{***}	0.37^{***}	0.36^{***}	0.32^{***}	0.04
	P5	0.58	6.91	0.45	7.22	0.30	5.11	1.00^{***}	0.34^{***}	0.34^{***}	0.33^{***}	0.12^{*}
	P6	0.61	7.63	0.48	7.92	0.32	5.93	1.00^{***}	0.31^{***}	0.27^{***}	0.34^{***}	0.14^{**}
	P7	0.62	7.73	0.51	8.07	0.34	5.70	0.99^{***}	0.31^{***}	0.21^{***}	0.34^{***}	0.19^{***}
	P8	0.65	6.92	0.49	6.95	0.34	4.89	1.02^{***}	0.35^{***}	0.19^{***}	0.33^{***}	0.17^{**}
	P9	0.68	7.07	0.60	7.37	0.43	5.49	1.04^{***}	0.38^{***}	0.12^{***}	0.32^{***}	0.20^{**}
	P10	0.73	7.21	0.62	5.27	0.49	4.17	1.13^{***}	0.59^{***}	-0.02	0.22^{**}	0.18
	P10 - P1	0.85	4.69	1.04	4.77	0.90	3.76	-0.19^{**}	-0.11	-0.67^{***}	0.15	0.33
rWML	P1	0.06	0.55	-0.03	-0.29	-0.001	-0.102	1.12^{***}	0.51^{***}	0.22^{***}	-0.04	-0.01
	P2	0.24	2.69	0.15	2.16	0.12	1.66	1.1^{***}	0.49^{***}	0.25^{***}	0.08	-0.02
	P3	0.37	4.32	0.28	4.73	0.24	3.86	1.06^{***}	0.50^{***}	0.22^{***}	0.07	0.03
	P4	0.42	5.32	0.32	6.58	0.29	5.57	1.05^{***}	0.49^{***}	0.23^{***}	0.09^{**}	-0.00
	P5	0.48	6.04	0.38	8.07	0.35	6.89	1.04^{***}	0.49^{***}	0.24^{***}	0.08^{*}	0.01
	P6	0.52	6.79	0.43	9.26	0.38	8.15	1.03^{***}	0.47^{***}	0.19^{***}	0.15^{***}	0.01
	P7	0.65	8.63	0.56	13.08	0.52	11.59	1.02^{***}	0.48^{***}	0.20^{***}	0.08^{***}	0.03
	P8	0.61	7.63	0.54	10.93	0.52	10.44	1.01^{***}	0.51^{***}	0.17^{***}	0.06	-0.00
	P9	0.63	7.71	0.57	10.74	0.56	9.78	1.00^{***}	0.53^{***}	0.06	0.00	0.04
	P10	0.72	7.40	0.71	9.50	0.71	9.10	1.03^{***}	0.53^{***}	-0.06	-0.03	0.09
	P10 - P1	0.66	4.76	0.75	5.35	0.72	4.88	-0.09^{*}	0.01	-0.28^{**}	0.01	0.10

Starting with the CAPM, we note that both momentum portfolios exhibit negative exposure to the market, albeit the coefficient is not statistically significant; however when we extend the model to include the remaining factors, the exposure to the market portfolio becomes statistically significant at the 5% level, but rWML is much closer to being a market neutral strategy. Additionally, both strategies exhibit a negative and statistically significant loading on the value factor, resulting in both having bigger alphas relative to the three-factor model than relative to the CAPM, given that Asness, Moskowitz, and Pedersen (2013) show that there is a pervasive negative correlation between value and momentum within and across asset classes. However, the fact that residual momentum shares this characteristic indicates that it can bring diversification benefits to investors who have exposure to the market and value premiums, in addition to being a profitable strategy on its own. It is also worth noting that exposure to factors is much more stable across deciles for rWML than for WML: total return momentum's losers portfolio is more sensitive to the market factor, and it also tilts towards value and size relative to the winners portfolio, while for residual momentum there is no big change in factor exposure, except for a relatively small tilt away from value in portfolios P9 and P10. Finally, the alphas of the momentum strategies are positive and statistically significant with respect to all models, indicating that these factors do not explain the performance of momentum. We observe an essentially monotonic increase in alphas as we go from the bottom decile to the top decile. The alpha of total return momentum is larger than that of residual momentum regardless of the model used, but saying that one strategy is better than another by simply looking at their alphas can be misleading if the strategies in question have different levels of residual volatility.

To get around this issue, we will study how each strategy contributes to expanding the efficient frontier by enhancing our models through the inclusion of total return momentum and residual momentum as factors in right-hand side of the regression equation. Specifically, we analyze the performance of total return momentum by controlling for residual momentum and vice-versa. The results of this test are presented on Table 3 below.

Table 3. Asset spanning regressions of decile portfolios
This table reports the performance of decile portfolios constructed as univariate sorts on total (WML) and residual return momentum (rWML). We include all common stocks (CRSP share code 10 or 11) that traded on the NYSE, AMEX or NASDAQ (CRSP exchange codes 1, 2 or 3) from January 1969 to December 2021 with share prices above $1 and market capitalization above the 20th percentile of NYSE listed companies. Total return momentum is calculated as the 12M-2M cumulative return. Residual momentum is the volatility scaled 12M-2M residual return over the past 36 months using the five factor model of Fama and French (2015). For each strategy we report monthly alphas (in percentage) relative to the five factors of Fama and French (2015). For total return momentum we control for the performance of residual return momentum and vice-versa. To account for heteroscedasticity, t-stats are computed using White (1982) standard errors. Portfolios are equal weighted and rebalancing is done on a monthly basis. **Note:** *, ** and *** denote statistical significant results at a 5%, 1% and 0.1% level, respectively.

	α	T-stat	β_{RMRF}	β_{SMB}	β_{HML}	β_{RMW}	β_{CMA}	β_{WML}	β_{rWML}
WML	0.00	0.43	-0.08	-0.12	-0.35^{***}	0.14	0.21		1.15^{***}
rWML	0.32	3.15	-0.01	0.06	0.01	-0.05	-0.04	0.44^{***}	

This analysis shows that residual return momentum subsumes total return momentum because when controlling for the former's performance the latter's alpha becomes insignificant (t-stat of just 0.43), while the reverse is not true; even after including WML in the model rWML exhibits a very robust and significant monthly alpha of 0.32%. We can therefore say that the five-factor model enhanced with residual momentum fully captures the performance of standard momentum, while the opposite is not true. When it comes to factor coefficients, both momentum based strategies display positive and significant loadings relative to each other, but this is hardly surprising as we should expect to see some overlap between winners on a total return basis and on a residual return basis.

4.3 Time-varying exposure to risk factors

We will now analyze the time-varying nature of factor exposures associated with the momentum strategy to investigate whether the profitability of total return momentum can be understood as a reward for bearing the risk of factor timing that can arise as a result of buying winners and selling losers. As Grundy and Martin (2001) have argued, the dynamics surrounding the changing factor tilts of momentum can be easily understood through the lens of the traditional one-factor model, the CAPM. Suppose that we are in a period following a bull market in equities; in this setting, the winners are expected to have substantially larger betas than the losers[3], as such, a top decile minus bottom decile momentum strategy is going to be long high beta stocks and short low beta stocks following an upturn in the stock market, and vice-versa following bear markets. Daniel and Moskowitz (2016) point to this characteristic of momentum as the source of its crash risk by viewing the high beta stocks that comprise the short leg of the winners minus losers portfolio coming out of a bear market through the prism of Merton (1974)'s model (i.e.: as out of the money call options on the assets of distressed firms, which suddenly go in the money as the market recovers). The issue with this time-varying exposure to factors is that it makes it harder to analyze momentum as a standalone phenomenon because, through this lens, it can be argued that momentum is not really a strategy of its own but rather a strategy whose profitability hinges on timing other factors (as Ehsani and Linnainmaa (2022) have shown) [4], therefore, attempting to minimize the dynamic factor tilts by selecting stocks on their firm-specific returns can not only help us to get closer to a more "pure" implementation of momentum but also evaluate the extent to which the profitability of plain momentum is a reward for timing other systematic risks.

As a first step, we wanted to grasp the variation over time of the betas of both winners-minus-losers portfolios. To this end, we employ the same methodology to estimate betas that we used to calculate the residual return for each stock in equation (1): we use a 36-month rolling window regression model with a minimum number of 36 observations to estimate the loadings that the individual stocks in our sample place on each factor of the Fama and French (2015) five-factor model. After collecting these estimates, we compute the average beta for all

[3] This does not necessarily mean that performance increases monotonically with a stock's beta. This would be true if we lived in a world where exposure to market risk was the only source of return, which is in fact not the case, but still, even though the CAPM does not perfectly explain the cross-section of returns, high beta stocks earn higher average returns than low beta stocks, even though the higher returns of high beta stocks are not as high as predicted by the CAPM and the lower returns of low beta stocks are not as low as predicted by the CAPM (see Frazzini and Pedersen (2015))

[4] This interpretation of momentum is not uncontested. For differing viewpoints, see Fan, Li, Liao, and Liu (2020).

decile portfolios in both momentum strategies; given that we are dealing with equal weighted portfolios, this is simply the average of the betas of the stocks in each portfolio. Finally, the beta of the winners-minus-losers portfolios is just the difference between the beta of the top decile portfolio and the beta of the bottom decile. These time-series data are plotted in Figure 3.

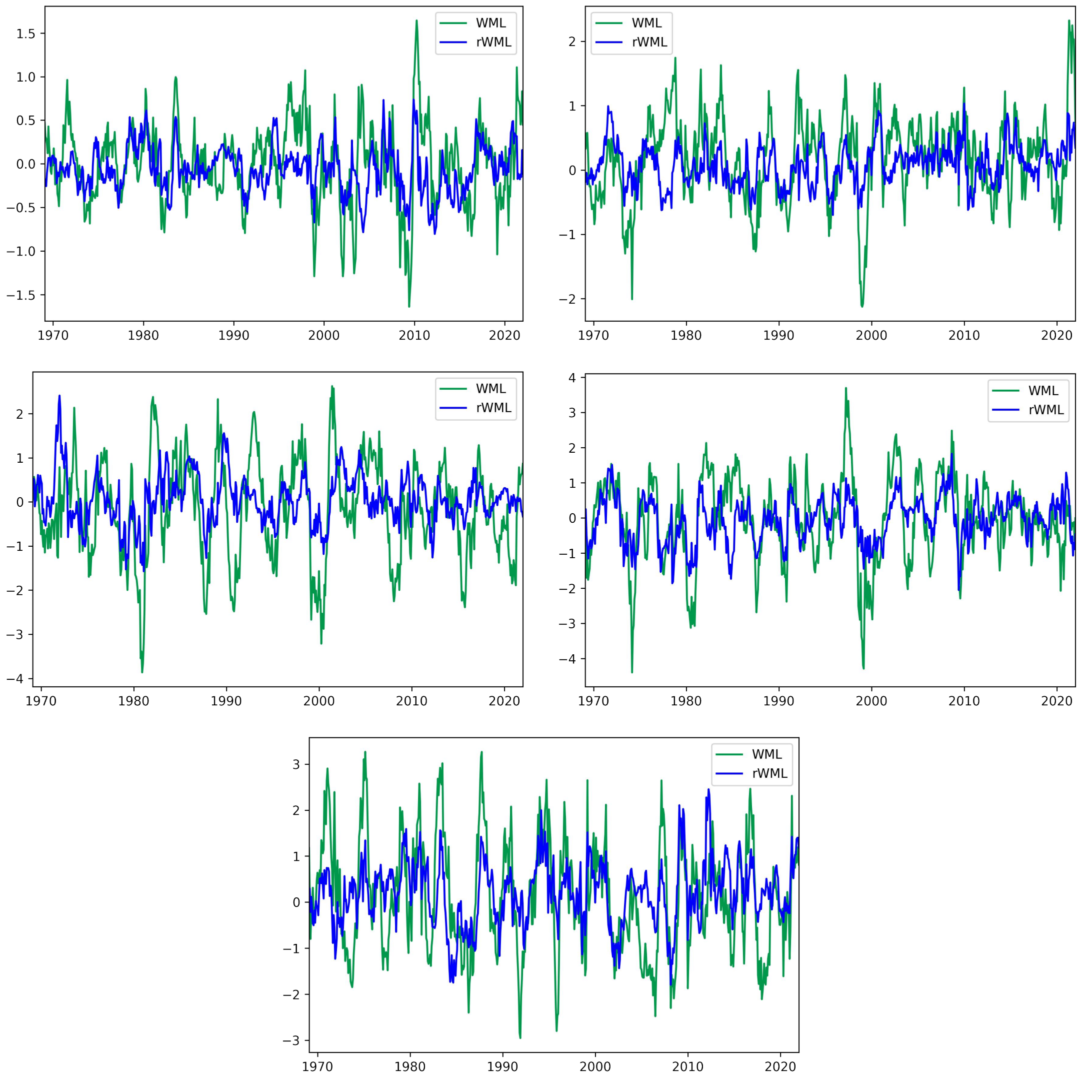

Figure 3. Time-varying betas
This figure shows the time-varying betas of long-short portfolios that invest in total return momentum and residual momentum and the market. Residual momentum is calculated using residual returns relative to the five-factor model of Fama and French (2015) using a 36-month rolling window. Both momentum strategies use the traditional 12M - 2M formation period. We consider all stocks that traded on the NYSE, AMEX or NASDAQ between January 1969 and December 2021 with a share price above $1 and a market capitalization above the 20th percentile market capitalization of companies listed on the NYSE. Momentum portfolios are equal weighted and rebalancing is done on a monthly basis. Betas are calculated using a 36-month rolling window, and each portfolio's beta is simply the average of the individual stocks' betas. The beta of the long-short portfolio is the beta of the top decile portfolio minus the beta of the bottom decile portfolio.

In each of these charts, we can see that the betas of the top decile minus bottom decile residual momentum approach are considerably more stable over time than those of its total return momentum counterpart. This means that residual momentum does a better job of capturing the firm-specific component of momentum, and it is not as susceptible to take on extreme loadings on factors that have performed well over the formation period. Focusing on the market beta chart in particular, we find an explanation for why residual momentum suffered a much less severe crash during the Global Financial Crisis (2). Zooming in on the 2000-2009 period on 4, we can see that the betas of both strategies around early 2009, when the stock market bottomed and momentum crashed. Total return momentum had a considerable negative beta exposure (as it loaded up on the kind of low beta stocks that held up better during the bear market and shorted the high beta stocks that crashed during the crisis); residual momentum meanwhile also had a negative beta (which is why it still suffered when the market recovered), but it was much closer to being market-neutral so its crash was much shallower.

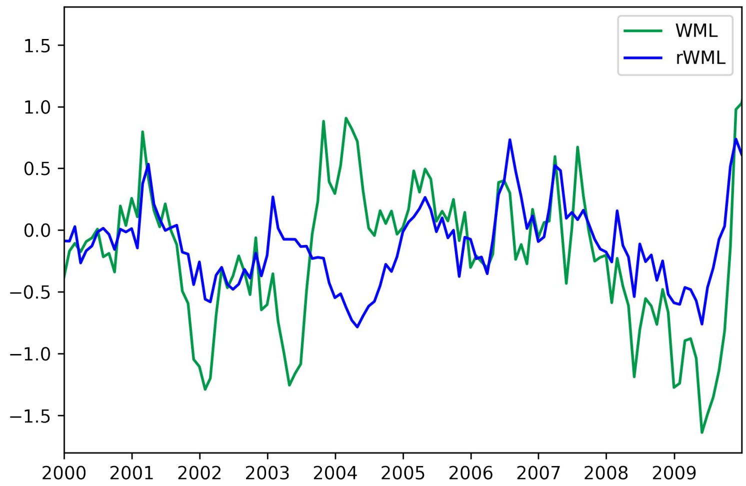

Figure 4. Time-varying betas
This figure shows the time-varying market beta of long-short portfolios that invest in total return momentum and residual momentum and the market. Residual momentum is calculated using residual returns relative to the five-factor model of Fama and French (2015) using a 36-month rolling window. Both momentum strategies use the traditional 12M - 2M formation period. We consider all stocks that traded on the NYSE, AMEX or NASDAQ between January 1969 and December 2021 with a share price above $1 and a market capitalization above the 20th percentile market capitalization of companies listed on the NYSE. Momentum portfolios are equal weighted and rebalancing is done on a monthly basis. The market beta is calculated using a 36-month rolling window, and each portfolio's beta is simply the average of the individual stocks' betas. The beta of the long-short portfolio is the beta of the top decile portfolio minus the beta of the bottom decile portfolio.

To develop a better understanding of the performance of these momentum strategies when controlling for factor timing, we need a proper econometric structure for our analysis. Given that we are attempting to estimate the relationship between the returns to the long-short momentum portfolios and the performance of factor-mimicking portfolios during the formation period used to select stocks on their momentum (in our case, prior 12 months excluding the most recent month), we will adopt a conditional framework like Grundy and Martin (2001) and Blitz, Huij, and Martens (2011). Grundy and Martin (2001) do so by adding interaction terms to the factor models that capture the performance of each factor during the formation period, defining three possible states: *up*, when the factor's average return during the formation period exceeded its historical average return by more than one standard deviation; *down*, when it was below the historical average by more than one standard deviation and *flat* when it was between these two scenarios. Blitz, Huij, and Martens (2011) on the other hand, create a dummy variable for the *up* state, which takes the value of 1 if the factor has a positive return over the formation period and 0 otherwise. We do not adopt the same model as Grundy and Martin (2001) for the sake of parsimony; because we are working with five different factors, we would have to add 15 interaction terms (3 for each factor). In addition, Grundy and Martin (2001) define each factor's *up* state relative to the historical performance of the factor, which we believe is not an entirely adequate way to capture factor momentum because it is possible for a given factor's performance to be above its historical average and yet still be underperforming relative to other factors. The model of Blitz, Huij, and Martens (2011) circumvents this issue by considering factors to be in an *up* state if the cumulative return over formation period is positive, implicitly assuming that zero is a good estimate for the average performance of a factor over 11 months. We adopt this definition, but tweak this later assumption, by considering each factor's average monthly return over the formation period, instead of the cumulative return. As such, our dummy variables are defined as follows: a factor-mimicking portfolio is in an *up* state if its average monthly return during the formation period is positive. To avoid look-ahead bias we do not use the historical average for the whole period. Instead, for each month t, we compute the historical return as the average from the factor's inception date up until month t - 1 using an expanding window[5]. With all this being said, we estimate the following model:

$$
\begin{aligned}
r_{i,t} = \alpha_i &+ \beta_{1,i} RMRF_t + \beta_{2,i} SMB_t + \beta_{3,i} HML_t + \beta_{4,i} RMW_t + \beta_{5,i} CMA_t + \\
&+ \beta_{6,i} RMRF\_UP_t RMRF_t + \beta_{7,i} SMB\_UP_t SMB_t + \\
&+ \beta_{8,i} HML\_UP_t HML_t + \beta_{9,i} RMW\_UP_t RMW_t + \\
&+ \beta_{10,i} CMA\_UP_t CMA_t + \varepsilon_{i,t}
\end{aligned} \tag{4}
$$

This simply adds five dummy variables (in the interaction terms) to the five-factor model of equation 1, where each dummy $RMRF\_UP_t$, SMB_UP_t, HML_UP_t, RMW_UP_t and CMA_UP_t takes the value of 1 if its respective factor's average return over months t - 12 and t - 2 (the formation period) is positive, and of 0 when otherwise.

[5] For the market, size and value factors our expanding window starts in July 1926, while for profitability and investment it starts in July 1963.

Table 4. Regression with interaction terms
This table reports the performance of long-short portfolios on total (WML) and residual return momentum (rWML). We include all common stocks (CRSP share code 10 or 11) that traded on the NYSE, AMEX or NASDAQ (CRSP exchange codes 1, 2 or 3) from January 1969 to December 2021 with share prices above $1 and market capitalization above the 20th percentile of NYSE listed companies. Total return momentum is calculated as the 12M-2M cumulative return. Residual momentum is the volatility scaled 12M-2M residual return over the past 36 months using the five factor model of Fama and French (2015). For each portfolio we report alphas relative to a model which has as its independent variables the five factors of Fama and French (2015) plus additional interaction terms that capture factor performance over the formation period of the portfolios (equation 4). To account for heteroscedasticity, t-stats are computed using White (1982) standard errors. Portfolios are equal weighted and rebalancing is done on a monthly basis. **Note:** *, ** and *** denote statistical significant results at a 5%, 1% and 0.1% level, respectively and t-stats are in parentheses.

	α	β_{RMRF}	β_{SMB}	β_{HML}	β_{RMW}	β_{CMA}	$\beta_{RMRF_{UP}}$	$\beta_{SMB_{UP}}$	$\beta_{HML_{UP}}$	$\beta_{RMW_{UP}}$	$\beta_{CMA_{UP}}$
WML	0.58^{**}	-0.57^{***}	-0.62^{***}	-1.08^{***}	0.20	0.04	0.66^{***}	0.82^{***}	1.00^{***}	0.17	0.04
	(2.87)	(-4.63)	(-4.73)	(-4.88)	(1.20)	(0.16)	(4.45)	(4.36)	(4.70)	(0.62)	(0.15)
rWML	0.58^{***}	-0.26^{***}	-0.13	0.37^{*}	-0.10	0.04	0.29^{***}	0.23	0.32	0.30^{*}	-0.08
	(4.26)	(-3.76)	(-1.65)	(-3.07)	(-0.82)	(0.21)	(3.51)	(2.00)	(2.47)	(1.62)	(-0.43)

The results for total return momentum indicate that its exposure to the market factor has strong time-variation and it is statistically significant; if the market's average monthly return over the formation period is positive, total return momentum's return is expected to go up by 0.66%, and vice-versa, supporting the thesis of Daniel and Moskowitz (2016). The same is clearly true for the size and value factors, but not for the profitability and investment factors introduced by Fama and French (2015). The performance of residual momentum shows that residualizing returns is effective in somewhat neutralizing the tendency of plain momentum to load up on the factors with the best recent performance. Its coefficient on the market interaction term is less than half of that of standard momentum and this might explain the smaller kurtosis of residual momentum (subsection 4.1) if we accept Daniel and Moskowitz (2016)'s explanation for the drivers of the crash risk. Residual momentum's exposure to the recent performance of the size factor is not statistically significant, unlike that of total return momentum, suggesting that the former strategy is better insulated against periods of poor performance in the small-cap segment of the market, however residual momentum tends to tilt towards profitable firms when these have done well, as shown by the positive and significant coefficient on RMW_{UP}. We can also see that both momentum strategies are negatively correlated with the value factor, so like its total return counterpart residual momentum can also provide diversification benefits to an investor who is long cheap stocks (Asness, Moskowitz, and Pedersen (2013)). Interestingly, even though total return momentum is characterized by a quite strong negative relationship with value, it tilts towards value stocks when these have been performing well, suggesting that the negative correlation between the two factors occurs mostly when value has underperformed relative to history. Finally, the alphas of both strategies are positive and statistically significant, even when incorporating the dynamic exposure to factors.

4.4 Importance of different factors for the residual momentum strategy

The goal of the residual momentum strategy is to sort stocks on their abnormal returns relative to an asset pricing model in order to better capture the stock-specific component of momentum. We use the five factor model of Fama and French (2015) in our analysis because it is one of the most recently developed models in empirical asset pricing and also because there is not, as far as we are aware, any work published in the literature that studies the residual (or idiosyncratic) momentum phenomenon using this particular model, but in fact this sort of strategy can be constructed by orthogonalizing returns against any model (preferably one grounded in sound economic analysis to avoid data snooping) that does a reasonably good job of describing the cross-section of average stock returns. Nevertheless, there is no reason to expect an equal impact from each factor to the performance of the strategy; actually, as we have shown on Table 4, the coefficients on the interaction terms that capture the dynamic factor tilts are not all equal and not all statistically significant. The work of Chaves (2016) showed that the majority of the improvements that accrue to an investor who follows a stock-specific momentum strategy are obtained by lowering exposure to the market factor and that there are not many incremental gains to be had from adding the size and value factors, like Blitz, Huij, and Martens (2011) and Blitz, Hanauer, and Vidojevic (2020) do.

To quantify this impact ourselves we will compare our baseline strategy to two other very similar ones: the CAPM residual momentum strategy and the Fama and French (1993) three-factor model residual momentum. The methodology to calculate residual returns is essentially unchanged; we calculate them per equation 2, the only difference is that we use only the market factor to calculate the predicted excess return in the CAPM specification, and the market, size and value factors in the three-factor specification. Once again, we build equal weighted portfolios of U.S. common stocks (excluding microcaps). We report the performance of top decile, bottom decile and top minus bottom decile portfolios.

Table 5. Importance of factors for the residual momentum strategy
This table reports the performance of bottom decile, top decile and top minus bottom decile portfolios constructed as univariate sorts on 12M - 2M residual momentum calculated using (i) the CAPM, (ii) the three-factor model of Fama and French (1993) and the five-factor model of Fama and French (2015). We include all common stocks (CRSP share code 10 or 11) that traded on the NYSE, AMEX or NASDAQ (CRSP exchange codes 1, 2 or 3) from January 1969 to December 2021 with share prices above $1 and market capitalization above the 20th percentile of NYSE listed companies. For each portfolio we calculate average monthly returns, monthly volatilies, skewness, kurtosis and Sharpe ratios. To account for heteroscedasticity, t-stats are computed using White (1982) standard errors. Portfolios are equal weighted and rebalancing is done on a monthly basis.

	Mkt-Rf			Mkt-Rf, SMB, HML			Mkt-Rf, SMB, HML, RMW, CMA		
	P1	P10	P10 - P1	P1	P10	P10 - P1	P1	P10	P10 - P1
Average Return	0.62	1.46	0.84	0.73	1.40	0.67	0.75	1.37	0.62
Volatility	6.49	5.97	4.67	6.02	6.04	3.90	6.11	5.70	3.50
Skewness	0.10	−0.26	0.48	0.06	−0.39	0.21	−0.03	−0.42	−0.00
Kurtosis	5.02	6.38	12.27	5.20	5.58	10.49	4.95	5.20	9.26
Sharpe Ratio	0.33	0.85	0.62	0.42	0.80	0.60	0.42	0.84	0.62
CAPM α	−0.10	0.80	0.90	0.05	0.72	0.67	0.06	0.72	0.66
t-stat	−0.79	6.70	4.83	0.47	6.47	4.24	0.55	7.40	4.76
3FF α	−0.18	0.81	0.98	−0.04	0.73	0.77	−0.03	0.72	0.75
t-stat	−1.48	8.48	5.14	−0.40	8.24	4.93	−0.29	9.50	5.35
5FF α	−0.12	0.80	0.92	−0.00	0.74	0.74	−0.01	0.71	0.72
t-stat	−0.83	7.88	4.16	−0.09	8.05	4.37	−0.10	9.10	4.88

The descriptive statistics indicate that each of the three approaches to residual momentum are roughly comparable. The Sharpe ratio indicates that all of them do an equally good job at reducing risk from the second momentum of the distribution compared to the total return momentum strategy (see Table 1). As we add factors to the return orthogonalization process, we can see that the distribution goes from slightly positively skewed to basically symmmetric, but on the other hand the kurtosis is considerably reduced in the five-factor specification of residual momentum compared to the simple market-adjusted implementation. In fact, residualizing returns against only the CAPM does very little to mitigate crash risk, given that the kurtosis of this strategy is 12.27 compared to 13.88 for total return momentum. This is in line with previous findings by Barroso and Santa-Clara (2015), who demonstrate that most of the risk inherent to the momentum strategy is idiosyncratic, and that the systematic component is very unpredictable, so calculating residual returns using real-time betas is of no use when it comes to trimming the tails of the distribution, as Daniel and Moskowitz (2016) also showed.

When we look at the alphas of each strategy relative to the CAPM, the three-factor model of Fama and French (1993) and the five-factor model of Fama and French (2015), we arrive at even stronger conclusion than that of Chaves (2016). It seems that there are losses in performance by adding factors to factor orthogonalization. Although all strategies present us with positive and statistically significant alphas relative to all pricing models, the CAPM adjusted strategy clearly dominates, with substantially higher alphas no matter the factors we control for. It is not surprising to see that reducing exposure to the profitability and investment factors is not very important if we keep the results of 4 in mind. While total return momentum has statistically significant exposure to the $RMRF_{UP}$, SMB_{UP} and HML_{UP}, its exposure to RMW_{UP} and CMA_{UP} is not at all significant. This means that it tends to increase its exposure to the market, size and value factors when these have done well, but not to the profitability and investment factors. Therefore, given that momentum does not exhibit dynamic factor tilts towards RMW and CMA, it is only natural that adding these two factors does not yield measurable improvements in performance.

4.5 Industry adjustment

The impact of industry effects on the performance of momentum strategies was first studied by Moskowitz and Grinblatt (1999). Their analysis builds the strategy using a 6-month formation window and a 6-month holding period. To assess whether industry performance accounts for the profitability of the momentum anomaly, the authors form an industry momentum strategy, which sorts portfolios of stocks in the same industry instead of individual stocks and buys the top three performing industries while selling the bottom three performers. They show that the performance of such a strategy is identical to that of momentum in individual equities, and to further investigate the issue, they build industry-adjusted portfolios by sorting stocks on their individual returns in excess of the respective industry's returns, finding the profitability of this latter strategy to be negligible, thereby concluding that the profitability of momentum is due to momentum in industries, not in stocks *per se*. Grundy and Martin (2001) find that the results of Moskowitz and Grinblatt (1999) occur because the formation period is contiguous with the investment period; in other words, using the standard definition of momentum, which excludes the most recent month, there is no evidence that industry effects explain momentum. More recently, Blitz, Huij, and Martens (2011) conducted a Principal Component Analysis (PCA) to capture the impact of industry effects on the performance of their residual momentum portfolios (which are built using Fama and French (1993) three-factor residual returns) and conclude that industry effects do not subsume the profitability of their strategy; Blitz, Hanauer, and Vidojevic (2020) employ a similar methodology to that of Moskowitz and Grinblatt (1999) and find that not only is residual momentum profitable after adjusting for industry effects, but that total return momentum is profitable as well, a result which is in accordance with the conclusions of Grundy and Martin (2001).

To investigate this issue the first step is to to classify the stocks in our samples by their industry. To do so, we retrieve their Standard Industrial Classification (SIC) codes and group stocks into industries using the first two digits of the SIC code and using the same classification criteria employed by Moskowitz and Grinblatt (1999), which are presented in Table 6 below:

Table 6. Description and statistics of industries
In this table we report summary statistics and of our 20 industries, including the two-digit SIC codes we use to group stocks. The average number of companies in each industry is reported, as well as the minimum number of companies that appears in each industry (reported in parentheses). In addition, we report the average percentage of total market capitalization for each industry, as well as the industry average return in excess of the 1-month Treasury bill rate.

Industry	SIC Codes	Avg. No. of Stocks	Avg. % of Market Cap.	Avg. Excess Return
1. Mining	10-14	176.41 (94)	3.45%	0.65%
2. Food	20	92.98 (56)	4.34%	0.89%
3. Apparel	22-23	70.46 (17)	0.56%	0.52%
4. Paper	26	37.13 (15)	1.67%	0.82%
5. Chemical	28	227.25 (107)	11.02%	1.06%
6. Petroleum	29	24.87 (11)	7.29%	0.95%
7. Construction	32	33.17 (9)	0.79%	0.90%
8. Prim. Metals	33	59.02 (25)	1.38%	0.73%
9. Fab. Metals	34	80.54 (35)	1.16%	0.92%
10. Machinery	35	221.13 (110)	7.47%	0.87%
11. Electrical Equip.	36	277.80 (145)	6.19%	1.03%
12. Transport Eq.	37	82.15 (55)	3.24%	0.79%
13. Manufacturing	38-39	226.92 (69)	3.28%	1.01%
14. Railroads	40	10.50 (3)	0.62%	1.19%
15. Other Transportation	41-47	45.17 (23)	0.42%	0.70%
16. Utilities	49	152.41 (78)	6.44%	0.68%
17. Dept. Stores	53	35.56 (11)	3.08%	0.70%
18. Retail	50-52, 54-59	344.39 (127)	5.01%	0.87%
19. Financial	60-69	660.78 (92)	14.91%	0.73%
20. Other	other	901.32 (220)	17.67%	0.95%

To investigate the effects of industries on our strategies, we adopt a methodology similar to the one Moskowitz and Grinblatt (1999) and Blitz, Hanauer, and Vidojevic (2020) use to separate industry momentum from individual stock momentum. To this end, we adjust each stock's momentum signal by subtracting from it the average of the respective industry signal. For total return momentum, this means subtracting from each stock's 12M - 2M cumulative return its industry's average cumulative return over the same period. In the case of residual momentum, the adjustment is made by subtracting the industry average residual momentum score (computed as in equation 3) from each stock's residual momentum score.

Table 7. Performance of industry-adjusted momentum strategies
This table presents a comparison between the results of total return momentum (WML), residual momentum (rWML) and their industry-adjusted counterparts. We include all common stocks (CRSP share code 10 or 11) that traded on the NYSE, AMEX or NASDAQ (CRSP exchange codes 1, 2 or 3) from January 1969 to December 2021 with share prices above $1 and market capitalization above the 20th percentile of NYSE listed companies. Total return momentum is calculated as the 12M-2M cumulative return. Residual momentum is the volatility scaled 12M-2M residual return over the past 36 months using the five factor model of Fama and French (2015). The industry-adjusted versions of each strategy subtract from each stock's trading signal the average trading signal for its industry (using the groupings presented in Table 6. For each portfolio we report average monthly returns, monthly volatility, skewness, kurtosis and Sharpe ratios. Portfolios are equal weighted and rebalancing is done on a monthly basis.

			Average Return	Volatility	Skewness	Kurtosis	Sharpe Ratio
WML	Total	P1	0.60	7.81	0.66	8.56	0.26
		P10	1.34	6.41	-0.64	6.02	0.73
		P10 - P1	0.75	5.85	-1.54	13.88	0.44
	Industry Adjusted	P1	0.74	7.43	0.21	6.29	0.35
		P10	1.41	6.95	-0.19	6.21	0.70
		P10 - P1	0.67	4.94	-0.79	13.48	0.47
rWML	Total	P1	0.75	6.11	-0.03	4.95	0.42
		P10	1.37	5.70	-0.42	5.20	0.84
		P10 - P1	0.62	3.50	-0.00	9.26	0.62
	Industry Adjusted	P1	0.87	5.91	-0.00	6.26	0.51
		P10	1.39	5.81	-0.44	4.78	0.83
		P10 - P1	0.52	3.48	-0.47	8.41	0.52

Our results show that industries play some role in the performance of rWML. After sorting stocks on their residual momentum score in excess of their industry's average, the Sharpe ratio of the strategy drops from 0.62 to 0.52, while the performance of total return momentum actually improves slightly after subtracting the industry average cumulative return over the formation period, so residual momentum is partly driven by momentum in industries, but given that the decrease in profitability for the industry-adjusted version is very small we can confidently say that industry effects are not driving the performance of either strategy.

4.6 Calendar month effects

The last of our empirical studies concerns the performance of our momentum strategies throughout the calendar year. This analysis is motivated by the work of Jegadeesh and Titman (1993, 2001) and Grinblatt and Moskowitz (2004) who find that total return momentum experiences seasonal patterns that are quite strong. The reason for this is related to the tax-loss selling effect, whereby fund managers sell small-cap loser stocks in December (to take advantage of a tax deduction), which results in falling prices for these kinds of stocks in December before they recover in January[6]. Given that the short leg of a total return momentum strategy is typically tilted towards stocks with these characteristics, momentum tends to perform well in December and poorly in January.

Given the econometric results obtained in subsection 4.3, we know residual momentum's exposure to the size factor is statistically insignificant. In contrast, total return momentum exhibits a negative and very significant exposure to size, so we expect residual momentum to be less exposed to seasonality (at least in this regard). The performance of both strategies by calendar month is presented in Table 8:

Table 8. Performance of total return and residual return momentum over the calendar year
This table reports the average performance of decile portfolios constructed as univariate sorts on total (WML) and residual return momentum (rWML) over the calendar year. We include all common stocks (CRSP share code 10 or 11) that traded on the NYSE, AMEX or NASDAQ (CRSP exchange codes 1, 2 or 3) from January 1969 to December 2021 with share prices above $1 and market capitalization above the 20th percentile of NYSE listed companies. Total return momentum is calculated as the 12M-2M cumulative return. Residual momentum is the volatility scaled 12M-2M residual return over the past 36 months using the five factor model of Fama and French (2015). We report the average return in each calendar month for bottom decile, top decile and top minus bottom decile portfolios and associated t-statistics.

		Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
WML	P1	2.95	0.36	1.50	2.88	0.40	-1.10	-0.75	0.74	-1.68	-1.56	1.89	1.50
		(2.43)	(0.38)	(1.46)	(2.00)	(0.42)	(-1.40)	(-0.76)	(0.72)	(-1.60)	(-1.37)	(1.60)	(1.89)
	P10	2.14	1.71	1.76	1.68	1.12	1.28	0.58	0.61	0.07	0.13	2.38	2.64
		(2.14)	(2.28)	(1.95)	(2.22)	(1.48)	(1.93)	(0.68)	(0.69)	(0.08)	(0.10)	(2.71)	(3.70)
	P10 - P1	-0.82	1.35	0.27	-1.20	0.72	2.37	1.33	-0.14	1.75	1.69	0.49	1.14
		(-0.81)	(2.03)	(0.44)	(-1.01)	(1.05)	(3.51)	(1.87)	(-0.19)	(2.53)	(2.19)	(0.55)	(1.69)
rWML	P1	1.93	0.75	0.92	1.65	0.84	-0.56	-0.15	0.44	-0.78	-0.17	2.33	1.78
		(1.59)	(0.79)	(0.90)	(1.14)	(0.90)	(-0.71)	(-0.15)	(0.42)	(-0.74)	(-0.15)	(1.96)	(2.24)
	P10	2.29	1.18	1.66	1.79	1.08	1.63	0.70	1.09	0.06	0.22	2.19	2.59
		(2.30)	(1.57)	(1.85)	(2.37)	(1.43)	(2.45)	(0.83)	(1.23)	(0.07)	(0.17)	(2.50)	(3.62)
	P10 - P1	0.37	0.43	0.74	0.15	0.23	2.18	0.85	0.65	0.84	0.39	-0.13	0.81
		(0.36)	(0.64)	(1.23)	(0.12)	(0.34)	(3.23)	(1.20)	(0.88)	(1.21)	(0.51)	(-0.15)	(1.06)

We can tell straight away that residual momentum delivers much more stable performance over the calendar year. Concerning the tax-loss selling effect, we note that total return momentum earns statistically significant profits in December and negative, albeit insignificant, returns in January. we should emphasize that even though the long-short portfolio does not earn a statistically significant profit in January, the bottom decile of total return momentum performs very strongly in the first month of the year, which is consistent with the January effect because small stocks tend to outperform in January, and total return momentum has a negative loading on the size factor. The short leg of residual momentum also performs better in January relative to the rest of the year, but this positive return is not significant. In conclusion, losers on a total return basis perform better in January, even though this does not seep through in a statistically

[6] For research on this phenomenon, see the work of Dyl (1977), Roll (1983), Reinganum (1983) and Jones, Pearce, and Wilson (1987), for example.

significant manner to the long-short portfolio. Residual momentum is not as susceptible to the January effect and is much less volatile throughout the year.

5 Robustness tests

In this section we implement some robustness tests to analyze whether the results we have obtained are valid under different settings. We start by restricting our sample to large-capitalization stocks to approach the results that are closer to the ones a large institutional investor would have obtained had these strategies been implemented in real time. Additionally, we also assess the performance of total and residual momentum under different formation and holding periods. Finally, none of the work we have done so far included transaction costs, which we know can be fairly substantial, especially for strategies with high turnover. In fact, an anomaly is only really an anomaly if its returns are significant after accounting for transaction costs, so we investigate this issue to conclude our empirical analysis.

5.1 Large-cap universe

As previously mentioned in 3.3, we now restrict our sample further to include only stocks with a large market capitalization. As mentioned, this is an important robustness test to ensure that our results could be implemented at scale, especially because our studies are conducted using equal weighted portfolios, which increases the risk of the results being skewed by small companies. To do this end, we define large companies as those with a market capitalization greater than the 70th percentile market capitalization of stocks traded on the NYSE. This narrows down our investment universe to just 343 stocks in January 1969 and 513 in December 2021, so we are left with a subset of fairly large and liquid securities, minimizing the drawbacks of equal weighting. We report descriptive statistics for the decile portfolios of total return and residual momentum in Table (tabela 6)

Table 9. Descriptive statistics of decile portfolios in the large-cap sample
This table reports descriptive statistics of decile portfolios constructed as univariate sorts on total (WML) and residual return momentum (rWML). We include all common stocks (CRSP share code 10 or 11) that traded on the NYSE, AMEX or NASDAQ (CRSP exchange codes 1, 2 or 3) from January 1969 to December 2021 with share prices above $1 and market capitalization above the 70th percentile of NYSE listed companies. Total return momentum is calculated as the 12M-2M cumulative return. Residual momentum is the volatility scaled 12M-2M residual return over the past 36 months using the five factor model of Fama and French (2015). For each portfolio we report the maximum and minimum monthly returns, the average return (returns are in percentage), the volatility, skewness, kurtosis and Sharpe ratio. Portfolios are equal weighted and rebalancing is done on a monthly basis.

		Maximum	Minimum	Average Return	Volatility	Skewness	Kurtosis	Sharpe Ratio
WML	P1	31.23	-24.15	7.58	23.31	0.03	4.79	0.33
	P10	35.54	-31.01	15.33	22.09	-0.26	9.18	0.36
	P10 - P1	43.01	-32.03	7.75	21.65	-0.02	9.18	0.36
rWML	P1	19.20	-19.96	7.87	18.82	-0.20	3.73	0.42
	P10	20.39	-25.71	14.17	17.55	-0.52	5.29	0.81
	P10 - P1	16.66	-13.63	6.30	12.72	0.03	5.21	0.50

The behavior of both strategies in the large-cap sample is very similar to their original versions. We still observe an almost doubling in average returns when we go from the bottom decile to the top decile, but excluding smaller stocks results in a decrease in average returns and Sharpe ratios, however, the difference in Sharpe ratio between residual and total return momentum still remains, which indicates that orthogonalizing returns is an effective way to reduce volatility, even in the large-cap sample. The negative skew of total return momentum goes away in the large-cap sample, and the distribution becomes essentially symmetric, while there is a substantial decrease in kurtosis for both, suggesting that the risk of large crashes is more prevalent in small-caps. Again, we invoke the explanation of Daniel and Moskowitz (2016): if the source of the crash risk lies in the option-like behavior of distressed firms that momentum is short on following bear markets, it is perfectly reasonable to expect this characteristic to be less expressive among large-caps, given that financially distressed firms tend to be smaller (Campbell, Hilscher, and Szilagyi (2008)).

With regards to the performance after controlling for factor performance (Table 10), our main conclusions from subsection 4.2 are unchanged. Both strategies deliver positive, and highly significant alphas relative to all pricing models, with total return momentum being a superior strategy from this standpoint. Residual momentum is still very much a market-neutral strategy, while total return momentum has a slight negative tilt on the market factor. The negative correlation between both strategies and the value factor persists and is still statistically significant, so the diversification benefits from combining value and momentum highlighted by Asness, Moskowitz, and Pedersen (2013) are achievable for an investor who only invests in the largest and most liquid stocks.

Table 10. Asset spanning regressions in the large-cap sample
This table reports the performance of decile portfolios constructed as univariate sorts on total (WML) and residual return momentum (rWML). We include all common stocks (CRSP share code 10 or 11) that traded on the NYSE, AMEX or NASDAQ (CRSP exchange codes 1, 2 or 3) from January 1969 to December 2021 with share prices above $1 and market capitalization above the 70th percentile of NYSE listed companies. Total return momentum is calculated as the 12M-2M cumulative return. Residual momentum is the volatility scaled 12M-2M residual return over the past 36 months using the five factor model of Fama and French (2015). For each portfolio we report alphas relative to the CAPM, the three factor model of Fama and French (1993), the five factor model of Fama and French (2015) and associated t-statistics. We also show the coefficients associated with each of the five factors. To account for heteroscedasticity, t-stats are computed using White (1982) standard errors. Portfolios are equal weighted and rebalancing is done on a monthly basis. **Note:** *, ** and *** denote statistical significant results at a 5%, 1% and 0.1% level, respectively.

		CAPM α	T-stat	3FF α	T-stat	5FF α	T-stat	β_{RMRF}	β_{SMB}	β_{HML}	β_{RMW}	β_{CMA}
WML	P1	-0.08	-0.49	-0.2	-1.31	-0.05	-0.29	1.21^{***}	0.07	0.49^{***}	-0.20	-0.31^{*}
	P10	0.61	4.28	0.73	5.58	0.83	5.68	1.03^{***}	0.26^{***}	-0.39^{***}	-0.30^{**}	0.05
	P10 - P1	0.68	2.72	0.93	3.85	0.88	3.14	-0.17^{**}	0.18	-0.88^{***}	-0.01	0.36
rWML	P1	0.04	0.69	-0.02	-0.18	-0.00	-0.04	1.07^{***}	0.05	0.19^{***}	-0.08	-0.04
	P10	0.60	7.04	0.63	7.48	0.60	7.05	1.00^{***}	0.04	-0.12^{**}	0.03	0.07
	P10 - P1	0.56	3.67	0.65	4.25	0.60	3.74	-0.08	-0.01	-0.30^{***}	0.03	0.11

5.2 (J,K) momentum strategies

We will now study whether the performances of residual and total return momentum are robust to different definitions of momentum. As mentioned in section 3.2, we use a 12-month formation period (excluding the most recent month, to avoid the short-term reversal effect documented by Jegadeesh (1990) and Lehmann (1990)) because this is the most frequently used formulation of momentum in the literature. To see if residual momentum delivers superior volatility-adjusted returns in alternative approaches to momentum, we employ the overlapping portfolios approach of Jegadeesh and Titman (1993), forming portfolios on J-month lagged returns (including the most recent month) and held for K months, where $J \in \{3, 6, 9, 12\}$ and $K \in \{3, 6, 9, 12\}$. For each (J, K) combination we calculate average monthly returns, standard deviations and Sharpe ratios:

Table 11. Total return and residual return momentum performance using the *(J, K)* momentum definitions
In this table we report descriptive average monthly returns, volatilities and Sharpe ratios for the *(J, K)* momentum strategies developed by Jegadeesh and Titman (1993), where $J \in \{3, 6, 9, 12\}$ and $K \in \{3, 6, 9, 12\}$. We include all common stocks (CRSP share code 10 or 11) that traded on the NYSE, AMEX or NASDAQ (CRSP exchange codes 1, 2 or 3) from January 1969 to December 2021 with share prices above $1 and market capitalization above the 20th percentile of NYSE listed companies. For total return momentum, stocks are sorted on their past J-month cumulative return. For residual momentum, they are sorted on the sum of their J-month residual returns relative to the five-factor model of Fama and French (2015), standardized by the volatility of residuals over the same period. Portfolios are equal weighted and held for K months.

		WML				rWML			
		$J = 3$	$J = 6$	$J = 9$	$J = 12$	$J = 3$	$J = 6$	$J = 9$	$J = 12$
$K = 3$	Return	0.41	0.62	0.72	0.46	0.18	0.36	0.58	0.51
	Volatility	4.74	5.52	5.66	5.79	1.72	2.76	3.12	3.27
	Sharpe Ratio	0.30	0.39	0.44	0.28	0.36	0.45	0.64	0.54
$K = 6$	Return	0.41	0.63	0.52	0.32	0.19	0.41	0.52	0.38
	Volatility	4.06	4.85	5.14	5.37	1.43	2.44	2.81	3.03
	Sharpe Ratio	0.35	0.45	0.35	0.20	0.47	0.59	0.64	0.44
$K = 9$	Return	0.44	0.49	0.34	0.17	0.22	0.39	0.40	0.28
	Volatility	3.42	4.29	4.76	4.99	1.24	2.14	2.56	2.80
	Sharpe Ratio	0.44	0.39	0.25	0.12	0.61	0.63	0.54	0.35
$K = 12$	Return	0.28	0.27	0.15	0.06	0.18	0.25	0.26	0.18
	Volatility	3.05	3.97	4.46	4.68	1.08	1.95	2.33	2.64
	Sharpe Ratio	0.31	0.24	0.11	0.04	0.56	0.45	0.39	0.24

Interestingly our results run counter to those of Jegadeesh and Titman (1993) because we do not document any dramatic decrease in performance when using the shorter formation period of three months, indicating that the short-term reversal effect[1] of Jegadeesh (1990) and Lehmann (1990) is not as prevalent in our sample. This might be due to a variety of factors, namely the fact that we include only stocks with 36 months of return history as of the formation month or different time periods, however it is likely that the exclusion of microcaps is contributing to these results because the work of Hou, Xue, and Zhang (2020) has shown that the short-term reversal anomaly is rendered statistically insignificant when microcaps are excluded. In any case, the main takeaway from this analysis is that, no matter the formation and holding periods utilized, residual momentum strategies exhibit higher Sharpe ratios compared to their total return counterparts. The improvement in this risk-adjusted measure of performance is mainly the result of a substantial decrease in the volatility of returns, echoing the conclusions we arrived at in subsection 4.2.

[1] Technically the short-term reversal effect sorts stocks based on their performance on the most recent month, but given that three months is a very short period of time results over this time-frame tend to be dominated by short-term reversal when this phenomenon is present.

5.3 Transaction costs

The last of our robustness tests concerns transaction costs. These are of paramount importance in practice because anomalies are not really relevant from the point of view of an investor if their profitability is eliminated by the trading costs incurred to exploit them. The work of Korajczyk and Sadka (2004) studied the extent to which momentum's returns were robust to trading costs and found that these costs, both in terms of bid-ask spreads and the price impact of the trading itself, were not sufficient to explain the persistence of a winners-minus-losers strategy, however they showed that momentum strategies are capacity constrained, with returns losing their statistical significance for funds with more than $2 billion.

Estimating trading costs is quite challenging because the necessary data is hard to come by, since costs vary over time, are not the same for every stock, with smaller and less liquid stocks being more expensive to trade, and also change depending on the investor. As such, we will circumvent the difficulty of having to directly estimate trading costs by employing the same methodology as Grundy and Martin (2001) and Barroso and Santa-Clara (2015), which hinges on calculating the level of transaction costs that would result in the strategies losing their statistical significance. The first step in this process is estimating the turnover of each leg of our momentum portfolios:

$$\tau_t = 0.5 \times \Sigma_i^{N_t} |w_{i,t} - \tilde{w}_{i,t-1}|, \tag{5}$$

with

$$\tilde{w}_{i,t-1} = \frac{w_{i,t-1}}{\Sigma_i^{N_t} w_{i,t-1}(1 + r_{i,t})} \tag{6}$$

where $w_{i,t}$ is the weight of stock i in one leg of the portfolio at time t, N_t is the total number of stocks in each leg of the portfolio at time t, and $\tilde{w}_{i,t-1}$ is the weight prior to trading in the current period. Given that our baseline strategy weighs stocks equally (in accordance with most of the literature), $w_{i,t}$ is simply the reciprocal of N_t. The turnover of a long-short strategy is simply the sum of the turnovers of the short and long portfolios. Applying these formulas to our total return and residual return momentum long-short portfolios we obtain average monthly turnovers of 71.87% for WML and 70.49% for rWML[2]. To assess the impact of such large turnovers on

[2] The long leg of WML has an average turnover 36.08%, while for the long leg average turnover equals 35.80%. For rWML, these amount to 34.49% and 36.00%, respectively.

the statistical significance of each strategy, we follow Grundy and Martin (2001) in computing the level of round-trip transaction costs that would render each strategy's risk-adjusted returns (relative to the CAPM, the three factor model and the five factor model) insignificant at a 5% level. These levels of costs are computed as follows:

$$net\_performance\_measure = gross\_performance\_measure - \bar{\tau} * trading\_cost \quad (7)$$

where $\bar{\tau}$ is the average turnover of each strategy, the $net\_performance\_measure$ is chosen so as to be the largest value that is statistically insignificant at a 5% level and the trading cost that leads to statistical insignificance is simply backed out from equation 7.

Table 12. Breakeven round-trip transaction costs
This table reports the round-trip transaction costs that would render each strategy's alphas with respect to the CAPM, the three-factor model of Fama and French (1993) and the five-factor model of Fama and French (2015) statistically insignificant at the 5% level. We include all common stocks (CRSP share code 10 or 11) that traded on the NYSE, AMEX or NASDAQ (CRSP exchange codes 1, 2 or 3) from January 1969 to December 2021 with share prices above $1 and market capitalization above the 20th percentile NYSE listed companies. Portfolios are equal weighted and rebalancing is done on a monthly basis.

		CAPM α	3FF α	5FF α
WML	Round-trip cost which would remove significance	1.27%	1.80%	1.41%
rWML	Round-trip cost which would remove significance	1.29%	1.54%	1.46%

We can see from the results in Table 12 that both strategies would be insignificant at fairly close trading costs, so we can confidently say that residual momentum is as sensitive to transaction costs as total return momentum. We should emphasize that these figures are not representative of the transaction costs that would be incurred from implementing these strategies (these may be higher or lower), but are simply an estimate of the costs that would remove significance from each strategy's risk-adjusted returns. An investor wishing to build portfolios using the strategies outlined in this analysis should first estimate the costs he or she would have to pay to rebalance portfolios with the levels of turnover outlined above and compare these costs to the ones in Table 12 to ascertain whether or not it makes sense for them to add total return momentum and residual return momentum to their overall portfolio.

6 Conclusion

Our paper undertakes a comparative analysis of momentum and five-factor residual momentum strategies. We show that both strategies deliver statistically significant profits and risk-adjusted returns, regardless of the form of risk we are analyzing. In accordance with previous literature, we find that attempting to neutralize factor exposure by using real-time information does not mitigate the crash risk of momentum and that exposure to factor timing risk does not explain the profitability of total return momentum.

The conclusions we draw are handicapped in several ways, many of which are common to most work in empirical asset pricing. There is of course the obvious issue pertaining to the non-stationary nature of the distribution of most asset returns, which of course makes it very hard to have a high degree of confidence that our conclusions will hold up in the future. This is just a restatement of the classical mutual fund disclosure: ”past performance is not indicative of future results”, but in the case of empirical asset pricing we face an even more challenging issue: that of ensuring that past performance is indicative of *past* results. This challenge arises because throughout our study we (mostly) ignored issues of liquidity and other transactions costs. This is especially dangerous when documenting the performance of long-short strategies, because we are not taking into account the potentially substantial borrowing costs that are incurred when selling a stock short.

Taking stock of all of our work, and assuming that the limitations outlined above do not invalidate our main conclusions, we remain puzzled by the historical pervasiveness of the profitability of momentum and the implications that the lack of a definitive risk based explanation for this phenomenon has for the efficient market hypothesis, one of the core tenets of academic finance.

Bibliography

C.S. Asness. Momentum in japan: The exception that proves the rule. *The Journal of Portfolio Management*, 37:67–75, 2011.

C.S. Asness, J.M. Liew, and R.L. Stevens. Parallels between the cross-sectional predictability of stock and country returns. *Journal of Portfolio Management*, 23:79–87, 1997.

C.S. Asness, T.J. Moskowitz, and L.H. Pedersen. Value and momentum everywhere. *Journal of Finance*, 68:929–985, 2013.

P. Barroso. The bottom-up beta of momentum. *29th Australasian Finance and Banking Conference 2016, Available at SSRN: https://ssrn.com/abstract=2144204 or http://dx.doi.org/10.2139/ssrn.2144204*, 2016.

P. Barroso and P. Santa-Clara. Momentum has its moments. *Journal of Financial Economics*, pages 111–120, 2015.

D. Blitz, J. Huij, and M. Martens. Residual momentum. *Journal of Empirical Finance*, 18: 506–521, 2011.

D. Blitz, M.X. Hanauer, and M. Vidojevic. The idiosyncratic momentum anomaly. *International Review of Economics and Finance*, 69:932–957, 2020.

J.Y. Campbell, J. Hilscher, and J. Szilagyi. In search of distress risk. *Journal of Finance*, 63: 2899–2939, 2008.

D.B. Chaves. Idiosyncratic Momentum: U.S. and international evidence. *Journal of Investing*, 25:64–76, 2016.

K. Daniel and T.J. Moskowitz. Momentum crashes. *Journal of Financial Economics*, pages 221–247, 2016.

W.F.M. DeBondt and R. Thaler. Does the stock market overreact? *Journal of Finance*, 40: 793–805, 1985.

E.A. Dyl. Capital gains taxation and year-end stock market behavior. *Journal of Finance*, 32: 165–175, 1977.

S. Ehsani and J.T. Linnainmaa. Factor momentum and the momentum factor. *Journal of Finance*, 77:1877–1919, 2022.

C.B. Erb and C.R. Harvey. The strategic and tactical value of commodity futures. *Financial Analysts Journal*, 62:69–97, 2006.

E.F. Fama. Efficient capital markets: A review of theory and empirical work. *Journal of Finance*, 25:383–417, 1970.

E.F. Fama and K.R. French. Common risk factors in the returns of stocks and bonds. *Journal of Financial Economics*, pages 3–56, 1993.

E.F. Fama and K.R. French. Multifactor explanations of asset pricing anomalies. *Journal of Finance*, 51:55–84, 1996.

E.F. Fama and K.R. French. Dissecting anomalies. *Journal of Finance*, 63:1653–1678, 2008.

E.F. Fama and K.R. French. Size, value and momentum in international stock returns. *Journal of Financial Economics*, 63:457–472, 2012.

E.F. Fama and K.R. French. A five factor asset pricing model. *Journal of Financial Economics*, pages 1–22, 2015.

E.F. Fama and J.D. MacBeth. Risk, return, and equilibrium: Empirical tests. *The Journal of Political Economy*, 81:607–636, 1973.

M. Fan, Y. Li, M. Liao, and J. Liu. A reexamination of factor momentum: How strong is it? *Financial Review*, 57:585–615, 2020.

A. Frazzini and L.H. Pedersen. Betting against beta. *Journal of Financial Economics*, pages 1–25, 2015.

M. Grinblatt and T.J. Moskowitz. Predicting stock price movements from past returns: the role of consistency and tax-loss selling. *Journal of Financial Economics*, 71:541–579, 2004.

B.D. Grundy and J.S. Martin. Understanding the nature of the risks and the source of the rewards to momentum investing. *The Review of Financial Studies*, 14:29–78, 2001.

R.C. Gutierrez and C.A. Prinsky. Momentum, reversal and the trading behaviors of institutions. *Journal of Financial Markets*, 53:48–75, 2007.

M.X. Hanauer. Is japan different? evidence on momentum and market dynamics. *International Review of Finance*, 14:141–160, 2014.

M.X. Hanauer and S. Windmüller. Enhanced momentum strategies. *Journal of Banking and Finance*, 148:106712, 2023.

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

K. Hou, C. Xue, and L. Zhang. Replicating anomalies. *The Review of Financial Studies*, 33: 2019–2133, 2020.

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

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

N. Jegadeesh and S. Titman. Profitability of momentum strategies: An evaluation of alternative explanations. *Journal of Finance*, 56:699–720, 2001.

C.P. Jones, D.K. Pearce, and J.W. Wilson. Can tax-loss selling explain the january effect? a note. *Journal of Finance*, 42:453–461, 1987.

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

B.N. Lehmann. Fads, martingales and market efficiency. *The Quarterly Journal of Economics*, 105:1–28, 1990.

R.A. Levy. Relative strength as a criterion for investment selection. *Journal of Finance*, 22: 595–610, 1967.

J. Lintner. Security prices, risk, and maximal gains from diversification. *Journal of Finance*, 20:587–615, 1965.

H. Markowitz. Portfolio selection. *Journal of Finance*, 7:77–91, 1952.

R.C. Merton. On the pricing of corporate debt: The risk structure of interest rates. *Journal of Finance*, 29:449–470, 1974.

T.J. Moskowitz and M. Grinblatt. Do industries explain momentum. *Journal of Finance*, 54: 1249–1290, 1999.

T.J. Moskowitz, Y.H. Ooi, and L.H. Pedersen. Time series momentum. *Journal of Financial Economics*, pages 228–250, 2012.

J. Okunev and D. White. Do momentum-based strategies still work in foreign currency markets? *Journal of Financial and Quantitative Analysis*, 38:425–447, 2003.

M.R. Reinganum. The anomalous stock market behavior of small firms in january: Empirical tests for tax-loss selling effects. *Journal of Financial Economics*, 12:89–104, 1983.

R. Roll. Vas ist das? the turn-of-the-year effect and the return premia of small firms. *Journal of Portfolio Management*, 9:18–28, 1983.

K.G. Rouwenhorst. International momentum strategies. *Journal of Finance*, 53:267–284, 1998.

K.G. Rouwenhorst. Local return factors and turnover in emerging stock markets. *Journal of Finance*, 54:1439–1464, 1999.

W.F. Sharpe. Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance*, 19:425–442, 1964.

J.L. Treynor. Toward a theory of market value of risky assets. 1962. Unpublished manuscript.

F. Wang and X.S. Yan. Downside risk and performance of volatility-managed portfolios. *Journal of Banking and Finance*, 131:106198, 2021.

H. White. Maximum likelihood estimation of misspecified models. *Econometrica*, 50:444–455, 1982.

A Appendix

In this appendix we show the results of the total return momentum and five-factor residual return momentum strategies without excluding microcaps (companies with a market capitalization below that of the 20th percentile NYSE listed company).

Table 13. Descriptive statistics of decile portfolios

This table reports descriptive statistics of decile portfolios constructed as univariate sorts on total (WML) and residual return momentum (rWML). We include all common stocks (CRSP share code 10 or 11) that traded on the NYSE, AMEX or NASDAQ (CRSP exchange codes 1, 2 or 3) from January 1969 to December 2021 with share prices above $1. Total return momentum is calculated as the 12M-2M cumulative return. Residual momentum is the volatility scaled 12M-2M residual return over the past 36 months using the five factor model of Fama and French (2015). For each portfolio we report the maximum and minimum monthly returns, the average annualized return, annualized volatility, skewness, kurtosis and Sharpe ratio. Returns are presented in percentage and portfolios are equal weighted with rebalancing done on a monthly basis.

		Maximum	Minimum	Average Return	Volatility	Skewness	Kurtosis	Sharpe Ratio
WML	P1	64.62	-35.88	5.79	31.42	1.09	9.90	0.18
	P2	39.41	-33.66	11.50	23.31	0.25	8.17	0.49
	P3	35.20	-26.47	13.75	21.05	0.16	7.80	0.65
	P4	30.88	-26.46	13.50	19.05	-0.18	7.49	0.71
	P5	29.17	-27.05	14.05	18.28	-0.31	7.60	0.77
	P6	28.01	-27.98	14.88	17.60	-0.45	7.26	0.85
	P7	24.56	-27.63	15.03	17.43	-0.61	7.03	0.86
	P8	22.69	-29.09	15.74	17.72	-0.69	7.11	0.89
	P9	20.74	-29.71	16.51	18.71	-0.81	6.73	0.88
	P10	28.66	-33.40	17.94	23.03	-0.60	6.12	0.78
	P10 - P1	18.05	-57.59	12.15	22.00	-2.30	17.90	0.55
rWML	P1	38.00	-26.37	8.78	22.00	0.18	6.69	0.40
	P2	33.24	-26.12	12.26	21.17	-0.01	6.49	0.58
	P3	28.96	-26.92	11.73	20.19	-0.26	6.53	0.58
	P4	32.12	-27.45	13.13	19.84	-0.20	7.34	0.66
	P5	31.31	-27.78	13.58	20.00	-0.09	7.29	0.68
	P6	28.84	-29.18	15.02	19.75	-0.32	7.24	0.76
	P7	32.72	-30.90	15.38	19.59	-0.35	7.49	0.79
	P8	26.82	-29.32	15.44	19.62	-0.39	6.69	0.79
	P9	25.46	-29.82	16.56	19.41	-0.64	6.64	0.85
	P10	23.57	-30.16	16.98	19.73	-0.57	6.17	0.86
	P10 - P1	23.26	-28.64	8.19	11.66	-0.92	15.90	0.70

Table 14. Asset spanning regressions of decile portfolios

This table reports the performance of decile portfolios constructed as univariate sorts on total (WML) and residual return momentum (rWML). We include all common stocks (CRSP share code 10 or 11) that traded on the NYSE, AMEX or NASDAQ (CRSP exchange codes 1, 2 or 3) from January 1969 to December 2021 with share prices above $1. Total return momentum is calculated as the 12M-2M cumulative return. Residual momentum is the volatility scaled 12M-2M residual return over the past 36 months using the five factor model of Fama and French (2015). For each portfolio we report monthly alphas (in percentage) relative to the CAPM, the three factor model of Fama and French (1993), the five factor model of Fama and French (2015) and associated t-statistics. We also show the coefficients associated with each of the five factors. To account for heteroscedasticity, t-stats are computed using White (1982) standard errors. Portfolios are equal weighted and rebalancing is done on a monthly basis. **Note:** *, ** and *** denote statistical significant results at a 5%, 1% and 0.1% level, respectively.

		CAPM α	**T-stat**	**3FF α**	**T-stat**	**5FF α**	**T-stat**	β_{RMRF}	β_{SMB}	β_{HML}	β_{RMW}	β_{CMA}
WML	P1	-0.37	-1.65	-0.63	-3.66	-0.58	-2.89	1.34^{***}	1.15^{***}	0.72^{***}	-0.07	-0.12
	P2	0.27	1.75	0.05	0.46	0.02	0.87	1.12^{***}	0.83^{***}	0.63^{***}	0.18^{*}	-0.14
	P3	0.50	3.92	0.31	3.51	0.21	2.20	1.08^{***}	0.73^{***}	0.50^{***}	0.27^{***}	-0.01
	P4	0.53	4.76	0.35	4.72	0.23	2.94	1.02^{***}	0.63^{***}	0.44^{***}	0.35^{***}	-0.00
	P5	0.59	5.70	0.42	6.46	0.29	4.53	1.00^{***}	0.58^{***}	0.37^{***}	0.31^{***}	0.10
	P6	0.67	7.13	0.52	8.46	0.37	6.29	0.99^{***}	0.53^{***}	0.32^{***}	0.35^{***}	0.11^{*}
	P7	0.68	7.52	0.56	8.87	0.39	6.88	1.00^{***}	0.48^{***}	0.23^{***}	0.32^{***}	0.20^{***}
	P8	0.73	7.66	0.62	8.92	0.44	6.74	1.00^{***}	0.52^{***}	0.19^{***}	0.33^{***}	0.23^{***}
	P9	0.77	7.29	0.67	8.60	0.52	6.59	1.02^{***}	0.56^{***}	0.17^{***}	0.29^{***}	0.20^{**}
	P10	0.79	5.33	0.72	6.63	0.61	5.61	1.14^{***}	0.78^{***}	0.09	0.13	0.28^{*}
	P10 - P1	1.16	5.02	1.35	6.11	1.18	4.80	-0.20^{*}	-0.37^{*}	-0.71^{***}	0.21	0.40
rWML	P1	0.06	0.42	-0.11	-1.07	-0.16	-1.40	1.11^{***}	0.72^{***}	0.42^{***}	0.14^{*}	-0.00
	P2	0.35	2.94	0.18	2.19	0.10	1.12	1.12^{***}	0.67^{***}	0.41^{***}	0.19^{*}	0.05
	P3	0.34	2.95	0.17	2.14	0.08	0.36	1.08^{***}	0.65^{***}	0.42^{***}	0.24^{**}	0.02
	P4	0.46	4.14	0.30	4.30	0.18	2.62	1.06^{***}	0.68^{***}	0.40^{***}	0.26^{***}	0.06
	P5	0.49	4.50	0.32	4.84	0.21	2.96	1.08^{***}	0.67^{***}	0.39^{***}	0.27^{***}	0.08
	P6	0.62	5.71	0.46	6.92	0.33	5.04	1.08^{***}	0.66^{***}	0.35^{***}	0.27^{***}	0.13^{*}
	P7	0.66	5.88	0.49	7.50	0.37	5.50	1.05^{***}	0.70^{***}	0.39^{***}	0.27^{***}	0.09
	P8	0.66	6.04	0.50	7.34	0.38	5.49	1.06^{***}	0.66^{***}	0.37^{***}	0.31^{***}	0.07
	P9	0.76	6.86	0.62	9.49	0.50	7.59	1.04^{***}	0.69^{***}	0.26^{***}	0.24^{***}	0.16^{*}
	P10	0.79	6.90	0.68	8.50	0.54	7.00	1.05^{***}	0.65^{***}	0.16^{**}	0.26^{***}	0.19^{*}
	P10 - P1	0.73	5.50	0.79	6.03	0.70	5.03	-0.07	-0.06	-0.26^{**}	0.12	0.20