NORTHERN TRUST

ASSET MANAGEMENT

RESEARCH REPORT

DYNAMIC FACTOR TIMING

ROB LEHNHERR
Head of Quantitative Equity Research

MANAN MEHTA
Head of Quantitative Fixed Income Research

KEY FINDINGS

- We leverage a regularization approach at the intersection of machine learning, financial economics, and portfolio theory to construct a novel factor timing portfolio.
- We apply a shrinkage penalty to historical estimates of mean and covariance which governs the extent of factor timing vis-à-vis an equally weighted factor portfolio.
- Our empirical analysis of six factor combinations over the past 18 years shows that the out-of-sample Sharpe ratios of the long/short factor timing portfolios are higher than naïve equally weighted factor portfolios by an average value of 0.29 in the large-cap US equity market.

SUMMARY

In this paper, we assess the performance of a machine learning-based, long/short factor timing strategy in the large-cap US equity market. The portfolios are constructed via mean-variance optimization with regularization applied to historical estimates of risk and return. We construct a timeseries of factor interactions by multiplying the long/short factor returns by a set of predictor variables including macroeconomic (market) data and factor characteristics. We then conduct a backtest in which the data is split into training, validation and testing on an annual basis. Using the training data, we estimate the mean and covariance of the factor interactions. During the validation period, a range of shrinkage penalties are applied to the estimates to generate a set of mean-variance optimal (factor interaction) weights, and then aggregated for each factor. The shrinkage penalty that maximizes the Sharpe ratio of the factor timing portfolio during the validation period is selected for the testing period, thus adjusting the influence of the predictor variables over time. As the shrinkage penalty approaches its maximum value, the aggregate factor weights converge to an equally weighted factor portfolio. Over the past 18 years, the factor timing strategy improved the out-of-sample Sharpe ratio of the equally weighted portfolio by an average value of 0.29 across six factor combinations.

The growing popularity of systematic investing has led to increased adoption of style factor strategies across asset classes and market segments. While style factors have been shown to outperform over the long run, they are susceptible to prolonged periods of underperformance. Factor cyclicality is not a new phenomenon and has been observed globally. Factor diversification is the primary approach to mitigating cyclicality but its effectiveness varies with market conditions. Another way to combat cyclicality is to introduce factor timing where factor exposures are dynamically adjusted to improve return or reduce risk. However, successful factor timing strategies are notoriously difficult to develop. The identification of a set of timing signals that works well across all regimes is particularly challenging. In this paper, we apply a novel approach to factor timing that leverages a well-known machine learning regularization technique to adjust the influence of a given set of timing signals in response to changing market conditions. This framework separates the factor timing problem into two distinct parts, namely 1) the identification of a robust set of factor timing signals, and 2) a mechanism for detecting and adjusting to new market dynamics.

In this research report we provide an overview of the methodology and apply it using a basic set of factor definitions and predictor variables within the large-cap US equity market. We evaluate the framework empirically by way of a backtesting process, wherein it consistently generates higher out-of-sample Sharpe ratios than equally weighted baseline portfolios. We then analyze the framework from the perspective of a long-only investor and observe similar efficacy when compared to the long/short results.

METHODOLOGY (FRAMEWORK) OVERVIEW

This section describes how the model inputs are constructed, the process to derive the optimal factor timing weights, and the backtesting procedure to perform out-of-sample (OOS) testing. Our approach follows Kozak, Nagel and Santosh (2020) which proposed a high dimensional characteristics-based factor pricing model.

Consider a set of N long/short factor portfolio returns at period t, F_t, a set of K factor characteristics at period $t\text{-}1$, X^C_{t-1}, and a set of J macroeconomic and market predictors at period $t\text{-}1$, X^M_{t-1}. The long/short return for the n^{th} factor portfolio at time t, $F_{n,t}$, represents the return of the top factor portfolio minus the return on the bottom factor portfolio.[1] Factor portfolios are sorted, grouped, and market capitalization (value) weighted at time $t\text{-}1$.

The timeseries of factor interactions based on the factor characteristics are constructed as the product of the long/short factor portfolio returns and their characteristics lagged by 1 month. More precisely, the factor interactions based on the factor characteristics can be expressed as:

$$FX^C_{n,k,t} = F_{n,t}\, X^C_{n,k,t-1} \qquad \text{Eq. [1]}$$

for $n = 1, 2, \ldots N$ and for $k = 1, 2, \ldots K$

The factor interactions based on the macroeconomic and market predictors can be expressed as:

$$FX^M_{n,j,t} = F_{n,t}\, X^M_{n,j,t-1} \qquad \text{Eq. [2]}$$

for $n = 1, 2, \ldots N$ and for $j = 1, 2, \ldots J$

The timeseries of factor interactions are augmented by including the long/short factor portfolio returns without interaction, $F_{n,t}$, for $n = 1, 2, \ldots N$. The total number of timeseries is therefore $M = N^*(1 + K + J)$. As a shorthand, we refer to the M timeseries as the factor interactions henceforth, although it includes the long/short factor return for each factor.[2]

The timing weights for each of the M factor interactions are determined by an optimization process with l_2 regularization. The first step estimates the $M \times 1$ mean vector, μ, and $M \times M$ covariance matrix, Σ, during the training period. The second step computes a set of timing weights for a range of shrinkage parameter values, λ, in accordance with the following:

$$w = (\Sigma + \lambda\, I)^{-1}\, (\mu + \lambda\, w_0) \qquad \text{Eq. [3]}$$

where w is the $M \times 1$ vector of timing weights, I is the identity matrix, and w_0 is an $M \times 1$ vector that represents the default weighting scheme. In our implementation, the components of w_0 that are associated with the N long/short factor portfolio returns are assigned a value of $1/N$, and the components of w_0 that are associated with the $N^*(M+K)$ factor interactions are assigned a value of zero. The optimal timing weights under no shrinkage (i.e., $\lambda = 0$) are equal to the conventional mean-variance optimal (maximum Sharpe ratio) output $w = \Sigma^{-1}\mu$, whereas the optimal timing weights under extreme shrinkage (i.e., $\lambda \to \infty$) converge to the default weighting scheme, $w \to w_0$. The third step transforms the $M \times 1$ factor interaction weights into N factor weights by applying a factor rotation:

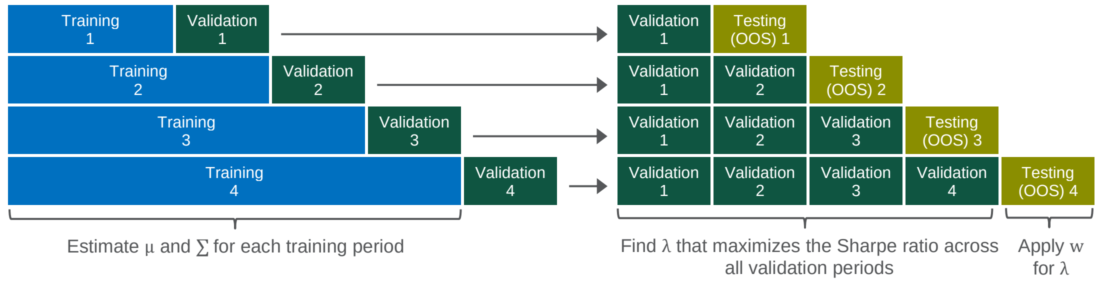
$$w_f = (w^T R)^T \qquad \text{Eq. [4]}$$

where w_f is the $N \times 1$ vector of factor timing weights, and R is an $M \times N$ matrix that contains the ending training period values for the factor characteristics, macroeconomic, and market predictors.[3,4] The fourth step rescales the factor timing weights by dividing them by the sum of the absolute weights, such that $-1 <= w_f <= 1$, and $\text{sum}(\text{abs}(w_f)) = 1$. This convenience allows the optimal factor weights to be interpreted in a long-only context, where a negative weight is implemented as a long position in the bottom portfolio. The final step selects the shrinkage parameter value λ that maximizes the Sharpe ratio across all validation periods.

Exhibit 1 summarizes the backtesting procedure used to test the framework empirically.

[1] We group using a traditional 30/40/30 split.
[2] This is technically correct if we add a constant interaction term $(=1)$ to the set of $(K+J)$ predictors.
[3] Each column in R will have $(1+K+J)$ non-zero elements to apply the relevant predictors for each factor.
[4] All factor characteristics, macroeconomic, and market predictors are standardized relative to the training period. Thus, the M non-zero elements in the R matrix are comprised of $N^*(K+J)$ z-scores, and N constants $(=1)$ related to the long/short factor returns.

Exhibit 1: Backtesting process – training, validation, and testing (OOS)



At each iteration, the average (μ) and covariance (Σ) of the M factor interactions are estimated using the training dataset.[5] The set of factor timing weights associated with a range of shrinkage parameter values (λ) are derived during the subsequent validation period. The shrinkage penalty (λ) that maximizes the Sharpe ratio across all validation periods is selected for the out-of-sample (OOS) testing period.

DEFINITIONS AND CONFIGURATIONS

Factor definitions and portfolios

We construct portfolios for four factors: value, profitability, investment, and momentum. For the value factor, we evaluate three popular metrics: book-to-market ratio, earnings yield, and cash flow yield. For the profitability factor, we test two variants: operating profitability and gross profitability. For the investment and momentum factors, we use the standard definitions, which are changes in total assets and total stock return over the last 12 months excluding the most recent month, respectively. We rank and sort all constituents in the Russell 1000 Index by each respective factor and bucket all stocks into three groups by count (30% of index constituents for top and bottom portfolios and 40% for the middle portfolio). Portfolios are market capitalization (value) weighted and rebalanced monthly. The long/short factor returns are calculated as the difference between the top and bottom portfolio returns.

Factor characteristics (predictors)

To create the timeseries of factor interactions based on factor characteristics we include six characteristics for each factor. They are: 3-month factor return, 12-month factor return, 3-month daily factor volatility, value, profitability, and investment. The 3-month and 12-month factor returns aim to capture factor momentum at different horizons. The 3-month daily factor volatility is included to account for the relationship between volatility and future returns. The value, profitability, and investment factors are used to introduce the interaction effects among factors in the dynamic framework. The value, profitability, and investment characteristics are computed as a spread (top portfolio minus the bottom portfolio), using the same definitions as those used to form the factor portfolios. All factor characteristics are standardized over the training period.

Macroeconomic and market predictors

We selected 5 macroeconomic and market predictors: 1-year real yield, yield slope, yield change, 3-month market excess return, and credit spread. The 1-year real yield is defined as the 1-year US treasury yield[6] minus the year-on-year percentage change in the US Consumer Price Index for all Urban Consumers. The yield slope is defined as the 5-year US treasury yield minus the 1-year US treasury yield. The yield change is computed as the year-on-year change in the 1-year US treasury yield. The 3-month market excess return is defined as the difference between the value weighted 3-month return of all CRSP US stocks minus the 3-month return of the 1-month US treasury bill. The credit spread is defined as Moody's seasoned Baa corporate bond yield minus the 20-year US treasury yield.[7] Each of the macroeconomic and market predictors are standardized over the training period.

Backtesting configuration

For the backtest analysis, the data starts in December 1984 and ends in December 2023. The minimal training period size is 240 months and expands at each iteration (every 12 months). The backtesting validation and out-of-sample testing periods are 12 months in length.

[5] We use an expanding window for the training period. The z-scores of the predictor variables are recomputed over the entire training period at each iteration.
[6] All US treasury yields refer to their constant maturity yields.
[7] The long-term US government yield is substituted in periods when the 20-year US treasury yield is not available.

BACKTEST RESULTS

The summary results for all factor combinations under consideration are reported in **Exhibit 2**. The optimal timing portfolio achieves a higher Sharpe ratio than the equally weighted factor portfolio in each combination tested (refer to the last column).

Exhibit 2: Out-of-sample (OOS) summary statistics

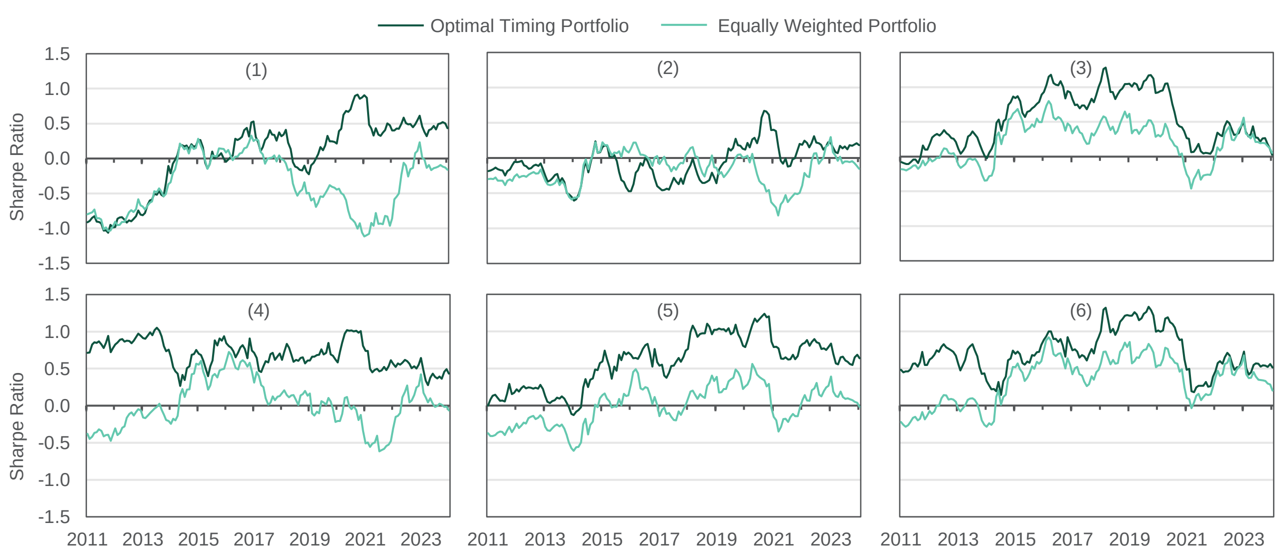
	Factors	Portfolio	Return (%)	Std. Dev. (%)	Sharpe Ratio	Δ Sharpe Ratio
(1)	Book-to-Market, Investment, Operating Profitability, Momentum	Equally Weighted	-1.35	5.46	-0.25	
		Optimal Timing	-0.53	6.75	-0.08	0.17
(2)	Earnings Yield, Investment, Operating Profitability, Momentum	Equally Weighted	-0.84	6.89	-0.12	
		Optimal Timing	-0.25	6.47	-0.04	0.08
(3)	*Cash Flow Yield, Investment, Operating Profitability, Momentum	Equally Weighted	0.93	7.97	0.12	
		Optimal Timing	2.35	8.10	0.29	0.17
(4)	*Book-to-Market, Investment, Gross Profitability, Momentum	Equally Weighted	-0.12	3.96	-0.03	
		Optimal Timing	3.27	7.41	0.44	0.47
(5)	*Earnings Yield, Investment, Gross Profitability, Momentum	Equally Weighted	-0.30	6.52	-0.05	
		Optimal Timing	3.99	8.35	0.48	0.53
(6)	*Cash Flow Yield, Investment, Gross Profitability, Momentum	Equally Weighted	1.56	6.88	0.23	
		Optimal Timing	4.67	8.30	0.56	0.32

*Excludes Financials, Real Estate, and Utilities[8]

Source: Northern Trust Quantitative Research, FTSE Russell, FactSet, CRSP,[9] FRED.[10] Data from 12/31/2005 to 12/31/2023. For illustrative purposes only. Actual results may vary.

Exhibit 3 plots the 5-year rolling out-of-sample Sharpe ratio for each backtest. Among the six formulations reported, only the second (2) underperforms the equally weighted baseline for a meaningful duration (from 2015 to 2018). In all other cases, the optimal timing portfolio either tracks or plots above the equally weighted portfolio over the entire backtest period. The performance of the optimal timing portfolios do not follow a discernable pattern across backtests, indicating that the success of the framework is not related to a single period or episode. Relative to the equally weighted baseline, the optimal timing portfolios perform best at different points of the backtest horizon. The performance gap is widest at the beginning of backtests four (4) and six (6), near the mid point of backtest three (3), and towards the end of all others (1, 2, and 5).

Exhibit 3: Five-year rolling out-of-sample (OOS) Sharpe ratio



Source: Northern Trust Quantitative Research, FTSE Russell, FactSet, CRSP, FRED. Data from 12/31/2010 to 12/31/2023. For illustrative purposes only. Actual results may vary.

[8] Cash flow yield and gross profitability have little-to-no relevance to these sectors.

[9] Center for Research in Security Prices (CRSP).

[10] Federal Reserve Economic Data (FRED) published by the Federal Reserve Bank of St. Louis.

In order to compare the influence of the two sets of predictor variables, three variations of the optimal timing strategy were run: (1) both factor characteristics and macro (and market) predictors, (2) macro and market predictors only, and (3) factor characteristics only. The summary results are reported in **Exhibit 4**.

Exhibit 4: Comparison of factor characteristics and macroeconomic predictors

Factors	Predictor Set	Return (%)	Std. Dev. (%)	Sharpe Ratio	Factors	Predictor Set	Return (%)	Std. Dev. (%)	Sharpe Ratio
(1) Book-to-Market, Investment, Operating Profitability, Momentum	Both	-0.53	6.75	-0.08	(2) Earnings Yield, Investment, Operating Profitability, Momentum	Both	-0.25	6.47	-0.04
	Macro	0.26	7.09	0.04		Macro	0.21	6.63	0.03
	Characteristics	0.32	6.29	0.05		Characteristics	-1.41	8.13	-0.17
(3) *Cash Flow Yield, Investment, Operating Profitability, Momentum	Both	2.35	8.10	0.29	(4) *Book-to-Market, Investment, Gross Profitability, Momentum	Both	3.27	7.41	0.44
	Macro	3.12	7.59	0.41		Macro	0.94	6.62	0.14
	Characteristics	0.39	8.76	0.04		Characteristics	4.47	7.65	0.58
(5) *Earnings Yield, Investment, Gross Profitability, Momentum	Both	3.99	8.35	0.48	(6) *Cash Flow Yield, Investment, Gross Profitability, Momentum	Both	4.67	8.30	0.56
	Macro	0.33	6.96	0.05		Macro	3.73	8.13	0.46
	Characteristics	4.77	9.08	0.53		Characteristics	3.60	8.31	0.43

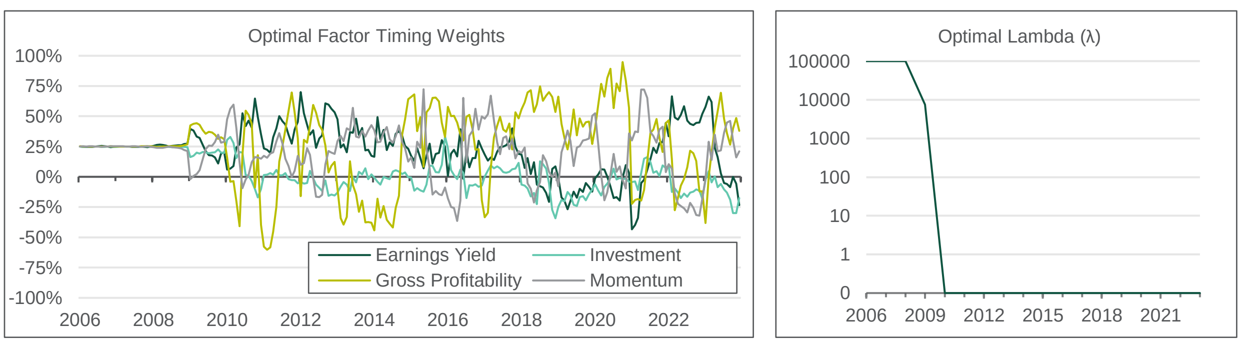
*Excludes Financials, Real Estate, and Utilities

Source: Northern Trust Quantitative Research, FTSE Russell, FactSet, CRSP, FRED. Data from 12/31/2005 to 12/31/2023. For illustrative purposes only. Actual results may vary.

In terms of relative importance, the results in Exhibit 4 are mixed. The number of best (worst) performing predictor sets are as follows: both 1 (1), macro and market predictors only 2 (2), and factor characteristics only 3 (3). The performance differential among the three variations is also difficult to generalize, as the range of Sharpe ratios is as low as 0.13 in the first (1) backtest (-0.08 vs. 0.05), and as high as 0.48 in the fifth (5) backtest (0.53 vs. 0.05). From these results, we do not find one predictor set to be more or less influential than the other in terms of performance impact.

For the purpose of illustrating specific aspects of the framework, we limit our analysis to the best performing factor combination (5) for the rest of this section. **Exhibit 5** plots the optimal factor timing weights and shrinkage penalty (lambda) over the backtest horizon. The impact of lambda is clear in this example as the optimal weights assume their default values (i.e. are equally weighted) from the outset of 2006 to the end of 2008, when the highest possible value (100,000) is selected for lambda. The optimal weights begin to diverge modestly at the beginning of 2009 as lambda begins to fall, with the dispersion increasing markedly in 2010 as lambda approaches zero.

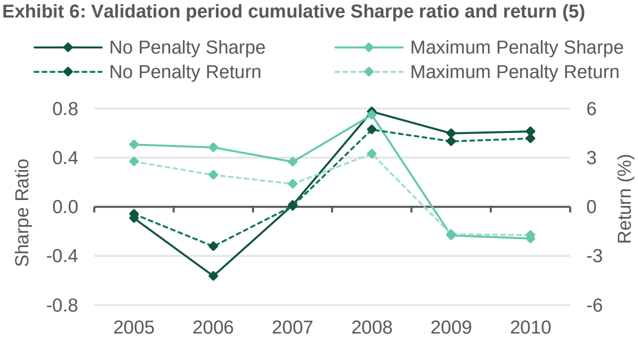
Exhibit 5: Optimal factor timing weights and lambda (5)



Source: Northern Trust Quantitative Research, FTSE Russell, FactSet, CRSP, FRED. Data from 12/31/2005 to 12/31/2023.

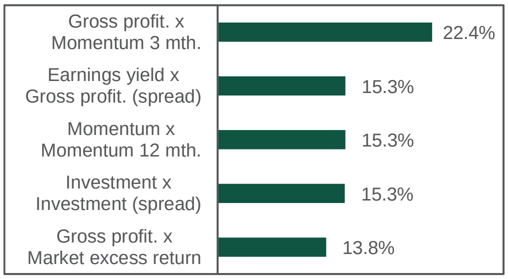
Recall that lambda is chosen such that it maximizes the Sharpe ratio across all validation periods. **Exhibit 6** shows the cumulative validation period performance under two extremes: “No Penalty” (λ=0), and “Maximum Penalty” (λ=100,000). The chart shows the “Maximum Penalty” Sharpe ratio dominating the “No Penalty” Sharpe ratio through 2007, resulting in the maximum optimal lambda (λ=100,000) being selected in years 2006 through 2008. The cumulative Sharpe ratios converge sharply in 2008, coinciding with the optimal lambda falling in 2009. Beginning in 2009, the cumulative “No Penalty” Sharpe ratio dominates the “Maximum Penalty” Sharpe ratio, with the optimal lambda approaching zero thereafter.

Exhibit 6: Validation period cumulative Sharpe ratio and return (5)



Source: Northern Trust Quantitative Research, FTSE Russell, FactSet, CRSP, FRED. Data from 12/31/2005 to 12/31/2010.

Exhibit 7: Highest average factor interaction weights (5)



Source: Northern Trust Quantitative Research, FTSE Russell, FactSet, CRSP, FRED. Data from 12/31/2005 to 12/31/2023.

While the influence of the individual factor interactions changes over time, their relative importance may be determined by comparing the average *absolute* optimal weights. The five most impactful factor interactions are listed in **Exhibit 7**, with the three month momentum of the gross profitability factor receiving the highest average optimal weight. Interestingly, all four factors are listed in the top five, and none of the predictors are repeated, suggesting a balanced representation among the framework inputs. Although the determination of the optimal factor interaction weights is multi-dimensional (means, variances, and covariances), the linear nature of the factor rotation allows for transparent attribution of changes to factor weights during the out-of-sample testing periods. **Exhibit 8** illustrates this over an eventful two month period during the backtest, which is summarized in the top section. In April of 2015, the optimal timing weight for momentum increased by almost 50% (from 24.0% to 72.3%), while the optimal timing weight for gross profitability fell by over 40% (from 49.3% to 8.3%). In the following month, the optimal weights for both factors reversed entirely. The bottom section of the exhibit itemizes the factor rotation effect associated with each factor predictor. The first column lists the optimal factor interaction weights derived at the end of the validation period (December 2014) for both factors. The next two columns show the change in the factor predictors (z-scores) during the months of April and May, respectively. The final two columns report the "factor rotation effect",[11] which is computed by mulitplying the optimal factor interaction weight (first column) by the change in the factor predictor. The sum of all factor rotation effects is reported at the bottom ("Total factor rotation effect"). The difference between the monthly change in the factor weight and the total factor rotation effect is attributable to the final step wherein the (absolute) factor weights are rescaled to sum to 1 ("Factor scaling effect").

Exhibit 8: Factor rotation analysis (5)

Momentum			Gross Profitability		
	April 2015	May 2015		April 2015	May 2015
Beginning period factor weight	24.0%	72.3%	Beginning period factor weight	49.3%	8.3%
Ending period factor weight	72.3%	23.1%	Ending period factor weight	8.3%	53.3%
Δ Factor weight	48.3%	-49.2%	Δ Factor weight	-41.1%	45.0%
Factor rotation effect	45.4%	-41.8%	Factor rotation effect	-47.0%	62.0%
Factor scaling effect	2.9%	-7.4%	Factor scaling effect	6.0%	-17.0%

	Momentum					Gross Profitability				
	Optimal	Δ Predictors (z-scores)		Factor rotation effect		Optimal	Δ Predictors (z-scores)		Factor rotation effect	
Factor Predictors	w	April 2015	May 2015	April 2015	May 2015	w	April 2015	May 2015	April 2015	May 2015
Factor Characteristics										
Earnings yield spread	-15.9%	0.00	-0.06	0.0%	1.0%	14.3%	0.02	-0.02	0.3%	-0.2%
Investment spread	-14.8%	-0.21	-0.05	3.2%	0.7%	-11.0%	-0.02	-0.19	0.2%	2.1%
Gross profitability spread	-18.7%	0.01	0.04	-0.1%	-0.8%	-24.9%	-0.02	-0.02	0.5%	0.5%
Momentum 3 mth.	-22.1%	**-1.11**	**0.87**	**24.5%**	**-19.2%**	54.1%	**-0.66**	**0.71**	**-36.0%**	**38.4%**
Momentum 12 mth.	-35.0%	**-0.20**	**0.18**	**7.0%**	**-6.4%**	28.0%	0.04	**0.33**	1.0%	**9.4%**
Volatility 3 mth.	-21.2%	-0.03	-0.01	0.6%	0.3%	18.6%	-0.11	-0.09	-2.0%	-1.7%
Macro and Market										
Real interest rate	13.1%	0.05	-0.07	0.6%	-0.9%	9.6%	0.05	-0.07	0.5%	-0.6%
Yield slope	-7.4%	-0.21	0.25	1.6%	-1.9%	-17.1%	-0.21	0.25	3.6%	-4.3%
Yield change	-27.2%	0.00	0.01	0.0%	-0.4%	-31.1%	0.00	0.01	0.0%	-0.4%
Market excess return	20.5%	**0.44**	**-0.57**	**9.0%**	**-11.8%**	-34.1%	**0.44**	**-0.57**	**-14.9%**	**19.5%**
Credit spread	-30.3%	0.03	0.08	-1.0%	-2.5%	-6.7%	0.03	0.08	-0.2%	-0.5%
Total factor rotation effect				45.4%	-41.8%				-47.0%	62.0%

Source: Northern Trust Quantitative Research, FTSE Russell, FactSet, CRSP, FRED. Data from 12/31/2014 to 6/30/2015.

In addition to transparency, the detail listed in Exhibit 8 allows for interpretability, which is an important consideration for models of higher complexity. For example, the optimal weights for the 3-month and 12-month momentum interactions are negative for the momentum factor, which indicates that momentum of momentum is a contrarion signal (at both horizons). By contrast, the optimal weights for the momentum interactions are positive for gross profitability, implying that positive (negative) momentum is a bullish (bearish) indicator for the gross profitability factor. The sign of the optimal interaction weights for market excess return reveals that the model treats strong market performance as a buy signal for momentum and a sell signal for gross profitability.[12] To highlight a specific example, in April of 2015 the trailing 3-month gross profitability factor return (i.e. the "Momentum 3 mth." factor characteristic) fell by 3.9% (from 0.7% to -3.2%), resulting in a

[11] Factor rotation effects in excess of 5% have been bolded for emphasis.

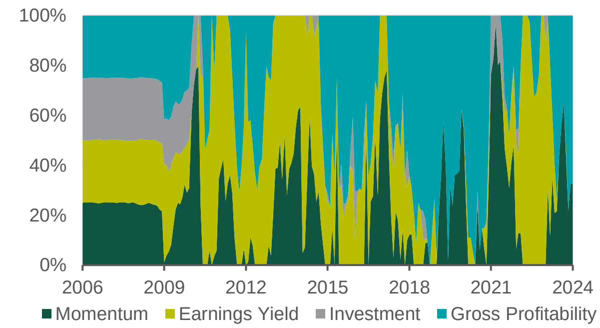
[12] As mentioned previously, optimal factor interaction weights are a function of the covariance matrix as well as the mean vector. However, in the examples highlighted here the optimal weights are in fact directionally aligned with the factor interaction averages.

z-score change of -0.66. The following month, the 3-month gross profitability return increased by 4.2% (from -3.2% to 1.0%), for a z-score change of +0.71. The impact of these two changes to the optimal gross profitability factor weight was -36.0% and 38.4%, respectively, representing over 60% of the total factor rotation effect in both periods. This type of attribution is useful for intuiting model output and understanding market dynamics.

IMPLEMENTATION CONSIDERATIONS

While the backtest results demonstrate the potential of the framework, a number of implementation considerations arise. Many investors establish factor programs within a long-only allocation, and most prefer to maintain positive factor exposure. In this section, we evaluate the framework under these constraints using two popular implementation methods: 1) bottom-up, and 2) top-down. In a bottom-up implementation, a single multi-factor score is assigned to every stock in accordance with a linear weighting scheme (in this case, the dynamic factor weights). A single portfolio is then formed from the stocks with the highest ranking multi-factor scores. A top-down implementation resembles a "fund of funds", where single factor portfolios are formed independently and then aggregated to form the multi-factor portfolio. In order to accommodate the constraints with minimal change to the framework, we assigned zero weight to all negative factor weights prior to the final rescaling step. This satisfies both constraints as all resulting factor exposures and weights are positive. **Exhibit 9** plots the long-only factor timing weights with this modification applied. Though not enforced by the framework, we see that the factor content is generally diversified, with a single factor receiving the full (100%) allocation in only 7 of the 216 months.

Exhibit 9: Optimal long-only factor timing weights (5)

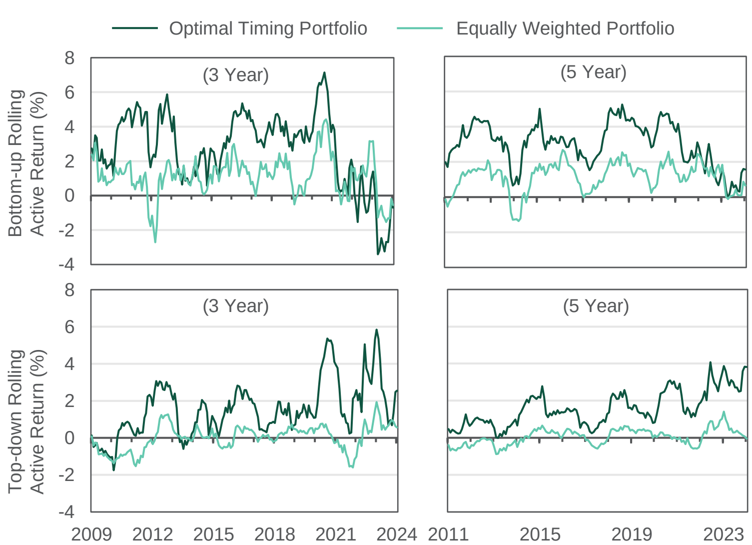


Source: Northern Trust Quantitative Research, FTSE Russell, FactSet, CRSP, FRED. Data from 12/31/2005 to 12/31/2023.

Exhibit 10: Long-only out-of-sample (OOS) summary statistics and rolling active returns (5)

	*Russell 1000 Ex	Bottom-up Eq. Weight	Bottom-up Optimal	Top-down Eq. Weight	Top-down Optimal
Return (%)	11.1	11.9	13.4	11.1	12.7
Std. Dev. (%)	15.7	14.9	15.3	15.4	15.9
Sharpe Ratio	0.66	0.75	0.82	0.67	0.75
Active Return (%)		0.9	2.4	0.0	1.6
Tracking Error (%)		4.4	5.1	1.9	3.7
Information Ratio		0.19	0.46	0.00	0.44

*Excludes Financials, Real Estate, and Utilities



Source: Northern Trust Quantitative Research, FTSE Russell, FactSet, CRSP, FRED. Data from 12/31/2005 to 12/31/2023. For illustrative purposes only. Actual results may vary.

Exhibit 10 compares the results of the long-only factor portfolios against the market capitalization weighted universe (Russell 1000 excluding financials, real estate, and utilities), and plots the rolling 3-year[13] and 5-year active returns. To summarize, the results in Exhibit 10 are similar to those reported in the previous section. The optimal timing portfolio dominates the equally weighted baseline in both bottom-up and top-down implementations, and consistently outperforms it on a 3-year and 5-year horizon. From the perspective of a benchmark-aware investor the results are equally encouraging. On a rolling 5-year basis, the active returns of the optimal timing portfolios are negative in only 1 of the 156 months over the backtest period in both implementations. On a rolling 3-year basis the hit rate declines but is still promising, as the active returns of the optimal timing portfolios are negative in only 39 of the 360 months combined across both backtests. This type of consistency is helpful for mitigating the divestment pressures that tend to build with sustained underperformance.

Although the performance benefits of the framework appear to translate in a long-only context, an obvious area of concern that remains is that of transaction costs. One can reasonably assume that prohibiting negative factor exposure significantly reduces the natural turnover of the strategy, as the

[13] Three years is a common look-back period for evaluating active strategies.

portfolio is prevented from oscillating between positive and negative positions (e.g. reversing from quality to junk). However, the potential for significant turnover remains even after limiting the portfolio to positive factor exposure. While the inclusion of turnover constraints within the backtesting framework is beyond the scope of this paper, a comparison of the transition matrix between the (bottom-up) multi-factor scores and those of the individual factors offers some insight. **Exhibit 11** shows the quarter-over-quarter transition statistics for the (top 30%) multi-factor and single factor portfolios. Specifically, Exhibit 11 reports the average ending period distribution of the top portfolio, where the first column represents the average number of stocks that remain in the top 30% on a quarter-over-quarter basis. Perhaps surprisingly, the difference in the retention rate between the naïve equally weighted baseline and the optimal factor timing portfolio is only 6.3% (75.6% vs. 69.3%) despite the dynamic nature of the optimal factor weights. When compared to the individual factors, the stability of the top multi-factor portfolios falls between earnings yield (81.8%) and momentum (64.3%). While the summary statistics reported in Exhibit 11 do not supplant the need for robust transaction cost analysis, they do offer hope that the benefits of the framework may be captured with modern porfolio construction techniques.

Exhibit 11: Top portfolio quarterly transition analysis (5)

	Top	Middle	Bottom
Bottom-up Equally Weighted	75.6%	21.9%	2.5%
Bottom-up Optimal Timing	69.3%	24.8%	5.9%
Gross Profitability	96.3%	3.6%	0.0%
Earnings Yield	81.8%	14.8%	3.4%
Investment	75.2%	21.6%	3.2%
Momentum	64.3%	30.1%	5.6%

Source: Northern Trust Quantitative Research, FTSE Russell, FactSet, CRSP, FRED. Data from 12/31/2005 to 12/31/2023.

CONCLUSION

Factor cyclicality is often cited as the biggest challenge to factor investing. Factor diversification is the most common method of mitigating cyclicality, but even multi-factor portfolios are prone to periods of sustained underperformance or steep drawdowns. The proliferation of factor timing research offers promise to address this problem, though success remains elusive as strategies must adapt to ever-changing market conditions if they are to consistently add value.

In this paper we applied a novel approach to factor timing that leverages foundational aspects of machine learning to adjust the influence of factor timing signals in response to new market regimes. In doing so, we effectively separated the factor timing problem into two distinct parts, namely 1) the identification of a robust set of factor timing signals, and 2) a mechanism for detecting and adjusting to new market dynamics. We evaluated the framework using a basic set of long/short factor returns and predictors, including macroeconomic (market) data and factor characteristics. In an emprical backtest spanning six different factor combinations, the optimal factor timing strategy outperformed the naïve equally weighted baseline in every instance, improving the Sharpe ratio by 0.29 on average. We then analyzed the framework from the perspective of a long-only investor with positive factor exposure constraints, and observed broad consistency when compared to the long/short results. The quarterly transition matrix of the optimal factor timing portfolio implies turnover that is comparable to single factor portfolios, suggesting practical implementation may be feasible.

We find these results to be encouraging, and believe this line of research has potential to enhance multi-factor strategies. Moreover, the framework established herein lends itself to broad applicability and extensibility, and may therefore have efficacy beyond traditional style factor investing.

References

Ang, A. (2014). Asset management: A systematic approach to factor investing. Oxford University Press.

Ang, A., Nabar, N., & Wald, S. (2013). Search for a Common Factor in Public and Private Real Estate Returns. The Journal of Portfolio Management, 39, 120–133.

Arnott, R. D., Clements, M., Kalesnik, V., & Linnainmaa, J. T. (2021). Factor momentum. Available at SSRN 3116974.

Campbell, J. Y. (1996). Understanding risk and return. Journal of Political Economy, 104(2), 298-345.

Ehsani, S., & Linnainmaa, J. T. (2022). Factor momentum and the momentum factor. The Journal of Finance, 77(3), 1877-1919.

Estrella, A., & Trubin, M. (2006). The yield curve as a leading indicator: Some practical issues. Current issues in Economics and Finance, 12(5).

Falck, A., Rej, A., & Thesmar, D. (2020). Is Factor Momentum More than Stock Momentum? Available at SSRN 3688983.

Fama, E. F., & French, K. R. (1988). Dividend yields and expected stock returns. Journal of Financial Economics, 22(1), 3-25.

Fama, E. F., & French, K. R. (1989). Business conditions and expected returns on stocks and bonds. Journal of Financial Economics, 25(1), 23-49.

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. Journal of Financial Economics, 33(1), 3-56.

Fama, E. F., & Gibbons, M. R. (1984). A comparison of inflation forecasts. Journal of Monetary Economics, 13(3), 327-348.

Fama, E. F., & Schwert, G. W. (1977). Asset returns and inflation. Journal of financial economics, 5(2), 115-146.

Flögel, V., Schlag, C., & Zunft, C. (2022). Momentum-managed equity factors. Journal of Banking & Finance, 137, 106251.

Gupta, T., & Kelly, B. (2019). Factor momentum everywhere. The Journal of Portfolio Management, 45(3), 13-36.

Hahn, J., & Lee, H. (2006). Yield spreads as alternative risk factors for size and book-to-market. Journal of Financial and Quantitative Analysis, 41(2), 245-269.

Hong, H., & Yogo, M. (2012). What does futures market interest tell us about the macroeconomy and asset prices? Journal of Financial Economics, 105(3), 473-490.

Keim, D. B., & Stambaugh, R. F. (1986). Predicting returns in the stock and bond markets. Journal of Financial Economics, 17(2), 357-390.

Kozak, S., Nagel, S., & Santosh, S. (2018). Interpreting factor models. The Journal of Finance, 73(3), 1183-1223.

Kozak, S., Nagel, S., & Santosh, S. (2020). Shrinking the cross-section. Journal of Financial Economics, 135(2), 271-292.

Kozak, S., & Nagel, S. (2023). When do cross-sectional asset pricing factors span the stochastic discount factor? (No. w31275). National Bureau of Economic Research.

Longstaff, F. A. (2004). The Flight-to-Liquidity Premium in US Treasury Bond Prices. Journal of Business, 77(3).

Ma, T., Liao, C., & Jiang, F. (2023). Factor momentum in the Chinese stock market. Available at SSRN 4148445.

Zaremba, A., & Shemer, J. (2018). Is there momentum in factor premia? Evidence from international equity markets. Research in International Business and Finance, 46, 120-130.

Important Information

Northern Trust Asset Management (NTAM) is composed of Northern Trust Investments, Inc., Northern Trust Global Investments Limited, Northern Trust Fund Managers (Ireland) Limited, Northern Trust Global Investments Japan, K.K, NT Global Advisors, Inc., 50 South Capital Advisors, LLC, Northern Trust Asset Management Australia Pty Ltd, and investment personnel of The Northern Trust Company of Hong Kong Limited and The Northern Trust Company.

NTAM products and services are issued in the United Kingdom by Northern Trust Global Investments Limited, issued in the European Economic Association ("EEA") by Northern Trust Fund Managers (Ireland) Limited, issued in Australia by Northern Trust Asset Management (Australia) Limited (ACN 648 476 019) which holds an Australian Financial Services Licence (License Number: 529895) and is regulated by the Australian Securities and Investments Commission (ASIC), and issued in Hong Kong by The Northern Trust Company of Hong Kong Limited which is regulated by the Hong Kong Securities and Futures Commission.

This information is directed to institutional, professional and wholesale current or prospective clients or investors only and should not be relied upon by retail clients or investors. This document may not be edited, altered, revised, paraphrased, or otherwise modified without the prior written permission of NTAM. The information is not intended for distribution or use by any person in any jurisdiction where such distribution would be contrary to local law or regulation. NTAM may have positions in and may effect transactions in the markets, contracts and related investments different than described in this information. This information is obtained from sources believed to be reliable, its accuracy and completeness are not guaranteed, and is subject to change. Information does not constitute a recommendation of any investment strategy, is not intended as investment advice and does not take into account all the circumstances of each investor.

This report is provided for informational purposes only and is not intended to be, and should not be construed as, an offer, solicitation or recommendation with respect to any transaction and should not be treated as legal advice, investment advice or tax advice. Recipients should not rely upon this information as a substitute for obtaining specific legal or tax advice from their own professional legal or tax advisors. References to specific securities and their issuers are for illustrative purposes only and are not intended and should not be interpreted as recommendations to purchase or sell such securities. Indices and trademarks are the property of their respective owners. Information is subject to change based on market or other conditions.

All securities investing and trading activities risk the loss of capital. Each portfolio is subject to substantial risks including market risks, strategy risks, advisor risk, and risks with respect to its investment in other structures. There can be no assurance that any portfolio investment objectives will be achieved, or that any investment will achieve profits or avoid incurring substantial losses. No investment strategy or risk management technique can guarantee returns or eliminate risk in any market environment. Risk controls and models do not promise any level of performance or guarantee against loss of principal. Any discussion of risk management is intended to describe NTAM's efforts to monitor and manage risk but does not imply low risk.

Past performance is not a guarantee of future results. Performance returns and the principal value of an investment will fluctuate. Performance returns contained herein are subject to revision by NTAM. Comparative indices shown are provided as an indication of the performance of a particular segment of the capital markets and/or alternative strategies in general. Index performance returns do not reflect any management fees, transaction costs or expenses. It is not possible to invest directly in any index. Net performance returns are reduced by investment management fees and other expenses relating to the management of the account. Gross performance returns contained herein include reinvestment of dividends and other earnings, transaction costs, and all fees and expenses other than investment management fees, unless indicated otherwise. For US NTI prospects or clients, please refer to Part 2a of the Form ADV or consult an NTI representative for additional information on fees.

Forward-looking statements and assumptions are NTAM's current estimates or expectations of future events or future results based upon proprietary research and should not be construed as an estimate or promise of results that a portfolio may achieve. Actual results could differ materially from the results indicated by this information.

Hypothetical portfolio information provided does not represent results of an actual investment portfolio but reflects representative historical performance of the strategies, funds or accounts listed herein, which were selected with the benefit of hindsight. Hypothetical performance results do not reflect actual trading. No representation is being made that any portfolio will achieve a performance record similar to that shown. A hypothetical investment does not necessarily take into account the fees, risks, economic or market factors/conditions an investor might experience in actual trading. Hypothetical results may have under- or over-compensation for the impact, if any, of certain market factors such as lack of liquidity, economic or market factors/conditions. The investment returns of other clients may differ materially from the portfolio portrayed. There are numerous other factors related to the markets in general or to the implementation of any