Factor Timing via Machine Learning

Xingkong Wei[*], Huihang Wu[*], Xiaoyan Zhang[*] and Xinran Zhang[*]

April 4, 2019

Abstract

Factor timing is a very important topic in both academia and industry. However, factor timing is very challenging for three reasons. First, factor returns are very volatile across periods; second, it is not clear what predictors are useful in factor timing; third, traditional financial econometric models may fail to capture non-linear relationship between predictors and factor returns under the high-dimensional data structure. In this paper, we try to shed light on above issues via machine learning.

1 Research Objectives

- How to dynamically allocate assets across factors?
- What variables can predict factors portfolio returns? Macro variables, factor momentums or other factors?
- Can machine learning outperform tradition financial econometrics models in predicting factor returns?

[*]PBC School of Finance, Tsinghua University

2 Introduction

Factor long-short strategy products have more than five hundred billion assets under management. Academic research has shown that some factors generate positive returns across time and industries. However, factors, like value and growth, exhibit highly volatile behaviour and suffer from significant drawback during bad times. As a result, investors need to diversify factors investment portfolios across periods, raising the question of how to allocate assets across factors dynamically. There is a recent debate about whether factor timing is possible. Arnott et al. (2016) find that performance can easily be improved by emphasizing the factors or strategies that are trading cheap relative to their historical norms and by deemphasizing the more expensive factors or strategies. However, Asness et al. (2017) examine tactical style timing strategies in a multi-style framework, asking whether dynamic allocations can improve the performance of a diversified multi-style portfolio. They compare the impact of value timing to that of strategic exposure to value itself and find that strategic diversification turns out to be a tough benchmark to beat. Our paper aims to answer this debate with more predictors and more complex models from machine learning literatures.

If we want to time factors, what are the most useful variables? The premiums of factor strategies incorporate three sources of predictability. First, factor returns might associate with macro variables. Bender et al. (2018) revisit the promises of factor timing, documenting the historical linkages between equity factor performance and different groupings of macro predictors. Second, Gupta and Kelly (2019) show that, in general, individual factors can be reliably timed based on their own recent performance. Third, the expected return of one factor may relate to other factor returns(Hodges et al., 2017; Lee,2017). We try to exhaust potential predictors to time factors across periods.

Hodges et al. (2017) argue that combining information from several predictors such as business cycle indicators, valuation, relative strength, and dispersion metrics is more effective than using individual predictors. With so many potential predictors in previous literature, traditional financial econometrics methods may fail to provide stable predictions due to the number of predictors approaching the total observations, serious muti-collinearity and non-linear relationship. Gu et al. (2018) find that machine learning offers an improved description of expected return behavior relative to traditional forecasting methods. Chinco et al. (2018) apply LASSO model to deal with sparse signals in financial markets. In this paper, we try to apply a bunch

of machine learning techniques to predict factors portfolio returns and build up factors investment strategies based on expected returns.

3 Research Design

Following Gu, Kelly, and Xiu (2018) setups, we describe an factor returns as an additive prediction error model:

$$r_{i,t+1} = g^*(r_{-i,t}, mom_{i,t-\tau}, macro_t) + \epsilon_{i,t+1} \tag{1}$$

where factors are indexed as $i = 1,\ldots,$ N and months by $t = 1,\ldots,$ T. $r_{i,t+1}$ is factor i realized return in time t+1. $r_{-i,t}$ are lag periods returns of other factors, and $mom_{i,t-\tau}$ are factors momentums over 2-3, 2-6, 2-12 periods. f_t are the macro variables at time t. $g^*(.)$ is flexible function of these predictors.

We try to perform a comparative analysis of methods in the machine learning repertoire, including principal component analysis, partial least squares regression, support vector regression, boosted regression trees, random forests, and neural networks.

Programming Languages We build our models on PYTHON, an interpreted, high-level, general-purpose programming language. We use scikit-learn package for linear models and tree models and tensorflow for neural network models.

4 Data

Stock Returns We obtain monthly total individual equity returns from CRSP for all firms listed in the NYSE, AMEX, and NASDAQ. Our sample begins in June 1967 and ends in December 2016, totaling 50 years.

Macro Variables We construct eight macroeconomic predictors following the variable definitions detailed in Welch & Goyal (2008), including dividend-price ratio (dp), earnings-price ratio (ep), book-to-market ratio (bm), net equity expansion (ntis), Treasury-bill rate (tbl), term spread (tms), default spread (dfy), and stock variance (svar). We also merge the 120 monthly macro variables spanning 1967:07 to 2016:12 from the FRED-MD database.

5 Factor construction

Our work extend the empirical literature on factor timing, which comes in two basic strands. The first strand uses MSCI smart beta indices as definitions for the factors: value, quality, momentum, size, and minimum volatility(Hodges et al.,2017; Arnott et al.,2016). This method lacks sufficient choices to time factors. The second strand constructs factor portfolios based on firm characteristics, aiming to cover the expanse of factors proposed in the academic literature that studies the cross section of stock returns(Gupta and Kelly, 2019). They form size neutral value-weighted portfolios and then combine these into an ultimate long–short factor portfolios. This method fails to consider that the short side of strategies may not necessarily perform bad all the times. In fact, factors perform badly when the short side portfolio outperforms the long side portfolio.

In order to overcome the shortcomings of above methods, we construct 94 long-short strategies based on firm characteristics by following Green et al. (2017). Then we use the underlying 188 portfolios from the long and short sides of these strategies as our factor timing candidates(Kozak et al., 2018). This third method generates enough factor candidates and make use of the short side portfolios when they are expected to perform well. Our factor list includes a variety of factors such as size, book-to-market, momentums, profitability, liquidity and beta.

References

[1] Arnott, R. D., Beck, N. and Kalesnik, V., 2016, "Timing 'Smart Beta' Strategies? Of Course! Buy Low, Sell High!", Working Paper.

[2] Asness, C., Chandra, S., Ilmanen, A. and Israel, R., 2017, "Contrarian Factor Timing is Deceptively Difficult", The Journal of Portfolio Management, 43(5): 72.

[3] Bender, Jennifer, Jerry Le Sun, and Ric Thomas.,2018, "Asset Allocation vs. Factor Allocation—Can We Build a Unified Method?." The Journal of Portfolio Management 45(2): 9-22.

[4] Chinco, A., Clark Joseph, A. D. and Ye, M., 2018, "Sparse Signals in the Cross-Section of Returns", The Journal of Finance, 74: 449-492.

[5] Green, J., Hand, J. R. M. and Zhang, X. F., 2017, "The Characteristics that Provide Independent Information About Average U.S. Monthly Stock Returns", The Review of Financial Studies, 30(12): 4389-4436.

[6] Gupta, T. and Kelly, B., 2019, "Factor Momentum Everywhere", The Journal of Portfolio Management, 45(3): 13.

[7] Gu, S., Kelly, B. T. and Xiu, D., 2018, "Empirical Asset Pricing Via Machine Learning", Working Paper

[8] Hodges, P., Hogan, K., Peterson, J. R. and Ang, A., 2017, "Factor Timing with Cross-Sectional and Time-Series Predictors", The Journal of Portfolio Management, 44(1): 30.

[9] Kao, Duen-Li, and Robert D. Shumaker. ,1999,"Equity Style Timing (corrected)." Financial Analysts Journal 55(1): 37-48.

[10] Kozak, Serhiy, Stefan Nagel, and Shrihari Santosh.,2018, "Interpreting factor models." The Journal of Finance 73(3): 1183-1223.

[11] Lee, W., 2017, "Factors Timing Factors", The Journal of Portfolio Management, 43(5): 66.

[12] Welch, I. and Goyal, A., 2008, "A Comprehensive Look at the Empirical Performance of Equity Premium Prediction", The Review of Financial Studies, 21(4): 1455-1508.

[13] McCracken, Michael W., and Serena Ng.,2016, "FRED-MD: A monthly database for macroeconomic research." Journal of Business & Economic Statistics 34(4): 574-589.