A Contemporary Analysis on the Empirical Performance of Equity Premium Prediction

Jason Ho
University of Toronto St. George

Yifan Fang
University of Toronto St. George

Abstract

Equity investors hold the belief that certain fundamental ratios or economic factors can predict equity premium, with a plethora of literature attesting for any and every predictor imaginable. Through re-examining the methods proposed by Welch & Goyal (2008), Rapach et al. (2009) and Fabian & Lukas (2016) using recent data up to December 2018, we confirm the results from previous studies and found structural instabilities in well documented economic and technical variables. Grounded on economic reasoning proposed by literature, we extended the literature by employing machine learning techniques to enhance our predictive regression framework: Regularized Regressions (Ridge and Lasso) and Artificial Neural Networks. These methodologies yield more stable models and better out-of-sample performances than traditional econometrics models over the same period.

JEL: **C53, C58, G12, G17**

1. Introduction

1.1 Research Question

Traditional time series analysis consistently reveals that equity premiums have insignificant autocorrelations (any significant autocorrelations are merely an artefact of measurement errors and nonsynchronous trading) and are likened to a White Noise stochastic process.

Instead, equity investors hold the belief that certain fundamental ratios or economic factors can predict equity premium, with an abundance of literature attesting for any and every predictor available. The predictability of equity returns is perhaps best summarized by Welch and Goyal (2008), "‘‘prediction works’’—though it is unclear exactly what works". The wealth of conflicting literature proves to be of little use for an actual investor wishing to "time the markets" by strategically allocating into a broad equity market index, after accounting for forecast uncertainties and transaction costs.

Our article thus focuses on forecasting equity premium for an investor using well documented predictors, emphasizing economic significance and out-of-sample performance. We scour the existing literature for valid models and build on them with a machine learning implementation.

1.2 Literature Review

It was not until the 1990s when researchers started to examine predictive models of equity premium both in-sample (IS) and out-of-sample (OOS). Using historical data up to 2005, Welch and Goyal (2008) studied 14 predictors proposed by literature as good predictors of US equity premiums and found poor IS and OOS performance for the three decades leading up to 2005. These models performed worse statistically and economically in predicting equity premiums for a market timing investor, than a prevailing historical mean, and were often unstable with predictability only in contiguous unusual years. Rapach et al. (2009) built on this and found that averaging over these individual forecasts delivered significant OOS gains relative to an unconditional historical average equity premium consistently over time. The resulting model was closely tied to the real economy. Later, Ferreira and Santa-Clara (2011) proposed a Sum-Of-Parts (SOP) approach, which decomposed and predicted the components of equity premium separately: capital gains (price-earnings multiple growth and earnings growth), dividend yield and risk-free rate. Fabian and Lukas (2016) extended the SOP approach by using the economic variables in Welch and Goyal to predict price-earnings multiple and incorporating additional technical indicators in Neely et al (2014) to predict earnings growth. In fact, Bartsch et al. (2017) found that after accounting for data snooping and transaction costs, many models fail to generate robust results, except for certain sum-of-parts configurations using both economic and technical variables. The success of complementary economic and technical variables provides intuition on fundamental and technical analysis used by practitioners.

Built upon this literature, our paper aims to explore a machine learning approach beyond traditional econometrics models. We first reexamine the empirical evidence using data up to 2018, and later search for parsimonious models using machine learning techniques including regularized regressions and artificial neural networks. All our models will be assessed based on OOS performance which is most relevant to a market timing investor.

1.3 Summary of our Methods

This paper uses the same set of economic and technical variables as in Welch and Goyal (2008) and Fabian and Lukas (2016), but with a monthly dataset spanning from January 1950 to December 2018. We first updated results from Welch and Goyal (2008) by employing a simple Ordinary Least Square (OLS) linear regression model for each individual economic variable and a multiple linear model using all economic variables, and then repeated this procedure with the inclusion of technical indicators. Using economic and technical variables, a total of 28 predictions are computed for each month. We then average over these individual predictions for each month by applying forecast combinations models in Rapach et al. (2010). To better understand the drivers of equity premium, we also decompose equity premium into its component drivers and predict them separately as in Fabian and Lukas (2016). Moreover, we tested forecasting performance of predictors after potential structural break in 2008 Global Financial Crisis (GFC) and compared it with previous structural breaks like the 1973 Oil Crisis and the Dot-com Bubble from 1995 to 2001.

To extend the literature, we employ regularized regressions (Ridge and Lasso) on the entire set of predictors on equity premium instead of averaging over individual predictions. We then attempt to use regularized regressions to predict component drivers of equity premium using an extended sum-of-parts approach, based on their robust performances as in Bartsch et al. (2017). We select optimal tuning parameters with the objective of maximizing statistical (R^2_{OOS}) and economic (ΔCEV) significance. Finally, we again predict equity premium and separately via its component parts using a series of Artificial Neural Networks to hopefully account for any nonlinearities among the predictors.

For each model, we use the first 240 months (20 years between 1950-1970) as the initial estimation sample and a recursively expanding estimation window to predict one-step-ahead equity premiums (beginning from January 1971). To assess the performance of each model, we primarily compute OOS statistics because we are interested in the OOS performance of the models on previously unseen test data, as would a market timing investor. An unconditional historical average equity premium is used as the benchmark in which all models are evaluated against. In effect, we attempt to answer, whether market-timing investors are better off expecting equity premium to be "as it always has been". We also explicitly consider the economic significance for a mean-variance investor by calculating the difference in their Certainty Equivalence Value (ΔCEV) and assuming a risk aversion factor of 3, as in Campbell and Thompson (2005).

1.4 Main Findings and Limitations

Overall, we find individual economic and technical variables to be instable in predicting equity premiums using OLS and forecast combinations. However, technical indicators provide complementary information to economic variables and lead to significant OOS performances when predicting using an extended sum-of-parts configuration combined with Ridge regression and Artificial Neural Networks.

Our machine learning models are reliant on an iterative specification search for an optimal tuning parameter. However, we circumvent the issue of overfitting by relying on OOS statistics to evaluate and compare models. Furthermore, our models may not be robust to future structural breaks. A structural break of the magnitude in 2008 GFC may potentially wipe out any profits from these models for a market timing investor.