The Predictive Ability of Stock Market Factors

Mohammed Elgammal: Finance and Economics Department, College of Business and Economics, Qatar University
Fatma Ahmed: School of Management, Swansea University
David G. McMillan: Division of Accounting and Finance, University of Stirling

September 2021

Purpose
This paper asks whether a range of stock market factors contain information that is useful to investors by generating a trading rule based on one-step-ahead forecasts from rolling and recursive regressions.

Design/methodology/approach
Using USA data across 3256 firms, we estimate stock returns on a range of factors using both fixed-effects panel and individual regressions and, using rolling and recursive approaches, generate time-varying coefficients. Subsequently, we generate one-step ahead forecasts for expected returns, simulate a trading strategy and compare its performance with realised returns.

Findings
Results from the panel and individual firm regressions show that an extended Fama-French five-factor model that includes momentum, reversal and quality factors outperform other models. Moreover, rolling based regressions outperform recursive ones in forecasting returns.

Research limitations/implications
Our results support notable time-variation in the coefficients on each factor, while suggesting that more distant observations, inherent in recursive regressions, do not improve predictive power over more recent observations. Results support the ability of market factors to improve forecast performance over a buy-and-hold strategy.

Practical implications
The results presented here will be of interest to both academics in understanding the dynamics of expected stock returns and investors who seek to improve portfolio performance through understanding which factors determine stock return movement.

Originality/value
We investigate the ability of risk factors to provide accurate forecasts and thus have economic value to investors. We conducted a series of moving and expanding window regressions to trace the dynamic movements of the stock returns average response to explanatory factors. We use the time-varying parameters to generate one-step-ahead forecasts of expected returns and simulate a trading strategy.

Keywords: Stock Returns, Stock Market Factors, Predictability, Panel, Trading Rule
JEL Codes: C22, G12

Acknowledgments: The authors gratefully acknowledge financial assistance from the Qatar National Research Fund (QNRF), project number NPRP8-601-5-074. The authors would like to thank and appreciate the comments and suggestions of SEF editor Professor Niklas Wagner and the two anonymous referees, which have greatly contributed to improve the article.

1. Introduction.

A major strand of the current asset pricing literature is the search for factors that explain the behaviour of stock returns. This approach largely began with Fama and French (1992, 1993), who identify their three-factor model. Subsequently, this line of research has developed towards the five-factor model (Fama and French, 2015, 2016). Concurrently, research has identified further factors not included in the five-factor model,[1] while Harvey et al. (2016) present a comprehensive discussion and note over 300 factors identified within the literature.

Much of this work involves the construction of portfolios of stocks based on a given firm characteristic. For example, such portfolios can be formed on firm size (e.g., large and small firms, Fama and French 1993), valuation (e.g., value and growth firms, Fama and French 1993) or recent performance (e.g., winning and losing firms based on profit or stock performance, Carhart, 1997). These portfolios are then regressed on identified explanatory factors, which are often hedged portfolios of the same firm characteristics. The validity of a given factor is then determined by its statistical significance in a regression of the unhedged portfolio (e.g., small stocks) on the hedged portfolio (e.g., small minus large stocks). This approach leads to a race between researchers to introduce new factors to explain stock return behaviour and has become known as a factor zoo. While this approach may identify cross-sectional significance, it is unclear whether it contains information relevant for investors in guiding future decisions.

In assessing the performance of a given asset-pricing model and the accompanying set of factors, the traditional approach is to examine how much of the variation in asset returns the model explains. For example, Fama and French (2015, 2016), Hou et al. (2015, 2017) and

[1] Examples include DeBondt and Thaler (1985), Jegadeesh (1990), Carhart (1997), Pastor and Stambaugh (2003), Vassalou and Xing (2004), Ang et al. (2006), Asness et al. (2013), Wagner and Winter (2013) Norvy-Marx (2013), Hou et al. (2015, 2017), Stambaugh and Yuan (2016), Bollerslev et al. (2016), Cederburg and O'Doherty (2016), Chai et al. (2017), Nartea et al. (2017), Detzel and Strauss (2018), Vidal-García et al. (2019), Shi and Li (2020), Salisu et al. (2020), Elgammal et al. (2020), Liang et al. (2020), Jun Xie et al. (2021) and Stereńczak (2021).

Stambaugh and Yuan (2016) compare the performance of different asset pricing models using the Gibbons et al. (1989) F-statistic for the zero-alpha restriction. In a different approach, Barillas and Shanken (2017) argue that if the factors contain independent information for stock returns, they should be able to price both asset and other factor returns.[2] For example, Fama and French (2016) highlight that the value factor (HML) is redundant as it is explained by other factors in the Five-Factor model. In a related example, Wagner and Winter (2013) add idiosyncratic risk and liquidity to the Fama and French (1992) and Carhart (1997) factors. Their results show that while market excess returns and size can explain variation in mutual fund returns, other factors only explain the performance of a subset of funds.

However, a key drawback of this approach is that factors can be identified to model specific portfolio types rather than their ability to model and predict stock returns as a whole. As an indication of this, Harvey et al. (2016) note over 300 factors. Further, Fama and French (1996) argue that a stock market factor based on one type of characteristic (e.g., size) should be able to explain the behaviour of stocks sorted on a different characteristic for the factor to have useful explanatory power. Equally, any identified factor should have explanatory power in a regression model that uses individual stock returns as the dependent variable. Moreover, this approach does not provide any information regarding the predictive ability of the factors for subsequent stock returns, which is a more pertinent question for investors.

This paper, therefore, addresses these concerns by estimating firm-level stock returns on a range of factors and examining, first, whether the factors are significant and, second, whether these regression models can be used to generate accurate forecasts of subsequent returns. Hence, our dependent variable is firm-level stock returns and not portfolios created according to a firm characteristic. Our primary contribution is to investigate the economic value

[2] The origin of this framework appears in the work of Fama (1998) and Asness and Frazzini (2013).

of risk factors to investors. That is, our analysis regarding the significance of factors is based on their ability to provide accurate forecasts for investors.

This line of research builds on the work of, notably, Lewellen (2015) and Dickson (2016) and more generally that of Cochrane (2011) and Harvey et al. (2016) who both question research that seeks to find an ever-increasing number of stock market factors. Of notable relevance, Lewellen (2015) considers whether adding factors to a model of expected returns improves the model fit and the ability to generate a profitable investment strategy. Dickson (2016) continues this approach and examines whether a range of factors can be used to build an investment strategy. However, our approach also differs from that taken by Lewellen (2015) and Dickson (2016) who consider the predictive power for individual stocks using individual firm-specific values of the stock market factors. For example, Dickson (2016) uses firm-level size, book-to-market ratio, profitability, investment and past returns to predict returns. In contrast, we consider the ability of the market-wide risk factors, e.g., Fama-French style factors, to explain individual firm stock market movement.

Following the general idea of Lewellen (2015) and Dickson (2016) and considering predictive power for individual stocks, this paper uses a series of rolling (fixed window) and recursive (expanding window) regressions for such stocks. From these regressions, we obtain the time-varying coefficients, generate out-of-sample forecasts for returns, construct a simple trading rule and compare the performance of different models against realised returns. We differ from the above-cited work as we focus on the forecast performance for individual stock returns, as opposed to forming portfolios, of market factors, rather than firm-specific factors. Thus, while one approach uses market factors to examine portfolio behaviour and a second approach uses individual factors to examine individual stocks, this paper considers market factors to model individual stocks and thus sits between, and extends, these two strands.

A key distinction with our paper in comparison to much of the literature is the use of individual stocks as opposed to portfolios. While the majority of the literature does use portfolios, following the approach initially suggested by Blume (1970) and popularised in Fama and French (1992, 1993),[3] as noted by Ang et al. (2018) the use of portfolios leads to a loss of information. A further consideration concerns the fixed effects panel model and statistical inference. While the Fama-MacBeth approach is the dominant one for firm-level data, as shown by Petersen (2009), the standard errors obtained from such an approach in the face of firm (cross-sectional) effects are not robust. Thus, adopting a fixed-effects panel approach with correctly adjusted standard errors will allow for accurate inference regarding whether stock market factors have statistical power in predicting returns and provides a point of comparison to previous work. Hence, the approach is taken here, while different from much of the literature, is empirically robust.

Notwithstanding the above, an obvious drawback of the fixed effects panel approach is that only a single coefficient value is obtained for each factor across all firms. This approach provides information effectively representing the average coefficient value across all the cross-sections and thus average forecast ability. As noted, this differs from Lewellen (2015) and Dickson (2016) who use individual firm-specific values of the stock factors. Hence, to consider individual effects, we also generate forecasts for individual firm-level stock returns, again using rolling and recursive regressions. Having obtained the expected return series, we again, generate a trading strategy and compare the expected return model with realised stock returns.

To examine the ability of factors models to provide forecast power for stock returns, we compare the performance of a baseline autoregressive model against, a one-factor (CAPM) model, Fama-French three- and five-factor models, the Cahart four-factor model as well as

[3] Blume (1970) argues that the use of portfolios leads to more precise estimation (reduces the standard error of estimates) as the construction of portfolios reduces idiosyncratic volatility.

models that include short- and long-term reversals and quality factors.[4] We report that the factor models that include momentum perform better than those that do not, while the rolling model approach, which includes fewer in-sample observations, provides forecasts that are more accurate compared to the recursive approach. This suggests that observations that are more distant do not add predictive power and indicates notable time-variation in parameter values. The results presented here will be of interest to both academics in understanding the dynamics of expected stock returns and investors who seek to improve portfolio performance through understanding which factors determine stock return movement.

The rest of the paper is organised as follows. Section 2 describes the data and the methodology used in this paper and Section 3 presents the empirical results. Section 4 concludes the paper with a summary of the implications of the work.

2. Data and Methodology.

2.1 Data

We obtain monthly data for 3256 US firms over the sample period 1990:1 to 2016:10. We include all firms and thus account for births and deaths to avoid survivorship bias. The sample is all common stocks from the Center for Research in Security Prices (CRSP) monthly files including NYSE, AMEX, and NASDAQ stocks as obtained from Bloomberg.

In modelling the behaviour of firm stock returns, we consider ten factors. Harvey et al. (2016) show a wide range of factors can be used to model stock returns. However, it will be difficult for any research to consider all suggested factors (more than 300 factors). Therefore, our selection is motivated by those factors that are commonly used within the literature and

[4] As noted by Harvey et al. (2016), there is a large (over 300) number of potential factors. We consider reversals as previously well-established factors (Jegadeesh, 1990; De Bondt and Thaler, 1985) and quality as a more recent factor that is found to have some success (Asness et al, 2013).

have some theoretical base and economic explanation.[5] We include the Fama and French (2015) five factors: market risk premium; small minus big firms (SMB); high minus low book-to-market firms (HML); conservative minus aggressive investing firms (CMA); profitable minus unprofitable firms (PMU). The Fama-French five factors are obtained from the data library of Kenneth French.[6,7] We construct the following risk factors: the stock price continuation or momentum effect (MOM, Carhart, 1997); the reversal of stocks over the short-run (previous month, Jegadeesh, 1990) and the long-run (between one and five years and sometimes referred to as the overreaction effect, De Bondt and Thaler, 1985); high volatility minus low (tranquil) volatility stocks (VMT, Ang et al., 2006).[8] We also obtain the high quality minus low quality (junk) firms factor (QMJ, Asness et al., 2013).[9]

Insert Table 1 Here

Table 1 reports summary statistics for monthly returns and the ten risk factors defined above. The numbers represent time-series averages of the monthly mean, standard deviation, skewness and kurtosis for each variable. We restricted the sample to firms with at least five years of information. Table 1 shows that the momentum strategy has the highest average return among risk factors, while the volatility investment style has a negative sign and confirms the puzzle that low volatility stocks have higher returns than high volatility stocks. The volatility factor also has the largest standard deviation in the sample.[10]

Table 1 also presents the correlation coefficients between the factors. This indicates some high correlations, for example, the correlation coefficient is over 0.5 for six pairs of

[5] A significant element in our factor selection is data availability that constrains our ability to collect a wider set of factors.
[6] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html
[7] We also consider the alternative profit factor of Novy-Marx (2013), with qualitatively similar results.
[8] We construct the momentum factor based on performance over the previous six months, short-term reversal is based on the previous month's performance, long-term reversal involves return performance between past months 13 to 60, and volatility is based on the previous 30-day standard deviation. In each case, stocks are ranked according to the criteria and a hedged portfolio formed as the difference between the top and bottom declines.
[9] The quality factor is obtained from the AQR website (https://www.aqr.com/library/data-sets/quality-minus-junk-factors-monthly).
[10] In the reported results, the stock return data is winsorised at the 1% level.

variables. This includes the quality factor four times (with the market, size, profit, and volatility factors), and between the size and profit and the value and investment factors. Further, the correlation coefficient is above 0.4 and 0.3 for a further four and seven pairs respectively. These values highlight two issues. First, the factors exhibit a degree of commonality, and this sits consistently with the factor zoo argument that a multitude of reported factors may not contain independent information. Second, that any in-sample modelling would have to be cognizant of potential multicollinearity, that the estimated coefficients of a model that includes all these variables may be dependent on each other and change when variables are included or excluded. Notwithstanding this, such multicollinearity issues are often minimised by the large number of observations used in this type of analysis.[11] Moreover, the primary purpose of this paper is to consider the forecast potential of the different stock market factors and thus, issues surrounding multicollinearity and in-sample significance are of lesser importance.

2.2 Methodology

To investigate the investment value of the above stock market factors, we predominantly use a fixed-effects panel approach, but also consider individual firm-level regressions. To examine time-variation in the regression coefficients and to generate out-of-sample forecasts, we use rolling and recursive regressions. From these regressions, we construct a simple trading rule and compare the performance of different models against the realised returns.

Our predictive regression model is a fixed-effects panel regression for all stocks on the above-listed factors:

$$r_{it} = \alpha + \Sigma_j \beta_j x_{j,t-1} + \gamma_i + \varepsilon_{it} \tag{1}$$

[11] In the case of this paper, we use over 500,000 observations.

where r refers to the stock return on firm i and at time t. The term x refers to the values of j factors. The individual (firm) fixed effects are given by γ_i, while ε_{it} is the error term.

To consider time-variation within the stock return and factor relations, we re-estimate equation (1) using both a twelve-month fixed window-rolling regression and an expanding window where the first twelve months serve as the starting point. Having obtained the time-varying coefficients, and following the general approach in Dickson (2016), we use this information to compute expected returns as such:

(2) $$E_t(r_{t+1}) = \Sigma_j \beta_{jt} x_{j,t} .$$

As can be observed, we derive the expected return for r_{t+1}. This ensures we are only using information that would be available in real-time to an investor to compute the future expected return. A further point to note is that we are only obtaining a single expected return per period, rather than for each firm (which we consider below).

To consider the ability of different factor models to forecast returns, we obtain the expected returns from equation (2) using alternative subsets of the full set of (ten) factors outlined above. Specifically, we obtain expected returns using a one-factor (CAPM) model (i.e., only the market risk premium), using the Fama and French three- (market, size and value) five- (market, size, value, profit and investment) factors and using the Cahart four- (market, size, value and momentum) factor model. We also consider all the above variables (ALL (3)). As the market and volatility factors are likely to capture similar information, we also exclude the volatility factor (ALL(1)) and alternatively the market factor (ALL(2)).

As an alternative to including the factors in equation (1), we estimate returns using a simple autoregressive model of order one. This approach acts as a baseline model, together with the realised returns themselves, and will allow us to consider whether the stock factors contain any useful predictive information. To examine the economic content of the factors, we design a simple trading rule that states if the expected return is positive then we buy the stock, while if the expected return is negative then we sell the stock. As we are buying the individual stocks, we can compare the outcome of the trading rule with that of the realised returns.[12]

The approach taken above examines the predictive ability of the asset pricing models utilising a panel regression for all stocks. This produces a common coefficient for each variable across all stocks. Of course, individual stocks are likely to have different loadings across the different factors. That is, we would expect every stock to respond differently to size, value, profitability and so forth. Notwithstanding this, the above analysis demonstrates whether the asset-pricing model can generate successful predictive ability. However, we can also estimate the asset pricing models individually for each firm in the sample. Again, in order to generate time variation, we use both a rolling fixed window and recursive expanding window regression approach, although now the rolling window length is five years (60 observations) to ensure a sufficient number of observations. We also only consider the ALL expected return model. As before, we generate the expected return according to equation (2).

3. Empirical Results.

3.1 Panel Regression

We begin our analysis using a fixed-effects panel regression for all stocks on the first lag of the above-listed factors as shown in equation (1). The estimates from this regression are reported in Table 2 and, as we can observe, all the coefficients are statistically significant. The coefficient signs represent the average relation across all firms to the stock market factors. We can observe a positive relation between the market and volatility premium and subsequent stock

[12] This exercise ignores the role of transaction costs as its purpose is to determine whether the factors have any information content for the movement of returns rather than as a full-scale investment strategy replication.

returns. This is consistent with the view that these factors represent an overall market risk, such that an increase in volatility leads to higher future (expected) returns.[13,14]

In terms of the other risk factors, we find a positive relation between stock returns and both the value (HML) and quality (QMJ) premiums. In addition, the results reveal a negative relation between stock returns and the size (SMB), investment (CMA) and profit (PMU) factors. For the factors examining trends within stock price movement, there is a positive relation between stock returns and short-term reversals and negative relation with momentum and long-term reversals. Over a large number of firms, it is difficult to provide a simple rationale for the sign of these coefficients, although it perhaps hints at the general characteristics of the sample of firms e.g., a tendency for the sample to contain larger, value firms. Moreover, it is our view that there is time variation within the estimated coefficients and so the nature of the relations is likely to change.

Insert Table 2 Here

3.2 Time-Variation, Forecasts and Trading Rule

3.2.1 Time-Varying Coefficients

We re-estimate equation (1) using a twelve-month fixed window-rolling regression to examine the time-varying nature of the relation between stock return and factors. Figure 1 presents graphical evidence that demonstrates movement in the coefficients across the sample period. Evident in each of these graphs is noticeable movement in the coefficient values around a fixed mean, with a small number of very large values, which can arise from using a one-year window. A key point of interest in these rolling coefficients is that the values switch between positive

[13] As an aside, Perras and Wagner (2020) report a negative association between stock returns and realised volatility for the S&P 500 index. They explained their results by the 'fear-of-missing-out' assumption. This result differs from ours and suggests that using different measures of volatility may yield different conclusions.

[14] Arguably, the market and volatility factors capture similar information. Experimenting by successively including one and then the other variable makes very little difference to the estimated coefficient values and significance.

and negative over the sample period. This switching behaviour casts doubt in regarding these variables as systematic risk factors as they do not exhibit a consistent relation with stock returns over time.[15]

Insert Figure 1 Here

3.2.2 Statistical Forecast Results

Having obtained the time-varying coefficients, we use this information to forecast the next period expected return as in equation (2). Table 3 shows the values of the mean error (ME), mean absolute error (MAE) and root mean squared error (RMSE) for the forecasts of expected returns estimated by the different predictive models as outlined above. The forecasting models include a simple autoregressive, AR(1), model as a baseline model, the CAPM and Fama and French three (FF3) and five-factor (FF5) models. Furthermore, we estimate expected returns using the Cahart four factor model, and three comprehensive models that consider all factors, including market, size, value, investment profit, momentum, short-term reversal, long-term reversal, quality minus junk, volatility. We consider three variants: ALL(1) that excludes the volatility factor; ALL(2) that excludes the market factor; ALL(3) that includes all factors. The rationale behind this choice is that the market and volatility factors are likely to capture similar information.

Panel A of Table 3 presents the results based on a 12-month fixed-window rolling regression, while in Panel B the results are based on an expanding window recursive regression. We can see in Table 3 that the ALL(1) model generally has lower forecast errors compared to other models (although not without exception) and the expanding window recursive regression has lower errors compared to fixed-window rolling regression. The three, four and five factor models generally perform at a similar level across the MAE and RMSE statistics, and with the

[15] An equivalent figure for the recursively estimated coefficients is available upon request.

rolling approach preferred (for the ME measure, the recursive approach performs better). For the CAPM, the rolling approach is also preferred.

3.2.3 Trading Rule Results

Table 4 presents the success ratio, trading rule returns and Sharpe ratio based on whether the predicted returns from equation (2) are positive (buy) or negative (sell) using the previously explained eight models [AR(1); CAPM; FF3; FF5; C4; ALL(1); ALL(2); ALL(3)].[16] Panel A presents the results based on a 12-month fixed-window rolling regression, while in Panel B the results are based on an expanding window recursive regression.

Insert Tables 3 and 4 Here

Examining the results in Panel A, under the heading ‘Rolling’, we can see that the expected returns obtained using the AR(1) model are poor compared to the realised returns, for which the average value is 0.433, with a Sharpe Ratio of 0.107. Thus, the AR(1) approach achieves lower values than these, being -0.209 and -0.051 respectively. The CAPM approach in contrast provides an improvement over both the realised returns and the AR(1) modelled expected returns. Notably, the monthly returns from the CAPM approach are 1.130, with a Sharpe ratio of 0.288. The success ratio of the AR(1) model is also noticeably lower.

Expanding the expected return model beyond CAPM through the inclusion of additional factors leads to a noticeable improvement in the forecast measures. Across the six expanded expected return models, we can see that each of the trading returns and Sharpe ratios are higher by at least 10% compared to the CAPM expected return, while the success ratio is highest for the ALL models. Between these expanded models, the ALL(1) model achieves the highest return and Sharpe ratio, while the FF3 model achieves the lowest. The ALL(1) model

[16] The success ratio is the proportion of correctly forecast returns signs (positive or negative returns) and thus a measure of directional accuracy. The Sharpe ratio is the return from the simulated trading strategy in excess of a 3-month Treasury bill divided by the variance of the return.

also has the highest success ratio. The performance of the ALL(2) model and ALL(3) model are only slightly better than the FF5 and are subordinate to the ALL(1) model. Thus, the ALL(1) model appears superior to the other expected return models. This suggests two broad points. First, is the relative importance of the market factor compared to the volatility (VMT) factor as the ALL(1) expected return model outperforms the ALL(2) model. Second, the importance of several factors in addition to the FF5 model in modelling expected returns, notably those capturing price trends. However, the model with the most number of factors [ALL(3)] does not outperform all other models, suggesting that not all factors are helpful.

In the above analysis, we obtain the time-varying coefficients using a rolling fixed window. Following Dickson (2016), we consider an alternative approach using a recursive expanding window. Therefore, Table 4 also reports the trading rule-based results obtained recursively. We will only briefly discuss these results as two pertinent points stand out. First, the ordering of the different expected return models is the same as for the rolling regressions, with one exception. Thus, the AR(1) approach provides the worst performance, and the ALL(1) provides the best performance. The ALL(2) and ALL(3) models perform at a similar level and slightly better than the other expected return models (although the Sharpe ratio of the recursive ALL(2) model is poor). Second, the performance of all recursively obtained expected returns is noticeably worse than those obtained from the rolling window approach. As both the rolling and recursive approaches include the new observations as they move through the sample, the difference in performance of these two approaches must lie in the discarding of old observations by the rolling approach. Thus, in modelling expected returns, this exercise states that old observations (in this case older than one year) provide no information beneficial in modelling expected returns. The advice for investors is to use a rolling modelling method.

3.3 Firm-Level Regression

In this section, we present the results of estimating the ALL factors expected return model individually for each firm in the sample. Again, in order to generate time variation, we use both a rolling fixed window and recursive expanding window regression approach, although now the rolling window length is five years (60 observations). The results in terms of the trading return and the Sharpe ratio are reported in Table 4 Panel C. A key observation emanating from this exercise is that the rolling regression approach continues to outperform the recursive approach. This continues to support the above results and suggests that dropping old observations when predicting the subsequent expected return provides superior forecasts. Moreover, the difference in values (0.750 compared to 0.119, and 0.087 compared to 0.022 for the trading returns and Sharpe ratio respectively) is noticeably greater using individual firm-level estimates than the differences from using the panel regression approach.

4. Summary and Conclusion.

In considering whether stock factors have any information content for investors, we examine their ability to predict future returns. We estimate expected returns in a panel regression setting and consider whether the predicted values can generate a successful trading strategy in comparison to realised returns. In obtaining the predicted values, we consider both rolling and recursive regressions and we compare the performance of these two methods.

In estimating expected returns, we consider a range of models, including a simple AR(1), as well as the CAPM, Fama-French three and five-factor and further models that incorporate additional factors including, momentum, reversal, quality and volatility. The results clearly point towards several conclusions. First, the range of stock market factors does provide useful information in respect of predicting future returns. Moreover, a model that uses multiple factors outperforms both the more restricted CAPM and Fama-French five-factor model. However, there may be a limit to the number of factors to be included, as the expected return model that contains all the factors (the ALL(3) model) does not achieve the best performance across the full range of measures (although it should be noted that it does not produce the worst forecasts). This result speaks to the earlier work (notably Cochrane, 2011; Harvey et al, 2016) that (implicitly) complains about the growing number of identified factors (or factor zoo).

Second, in obtaining the trading returns, we use both a rolling and recursive regression approach. Our results indicate that rolling regressions outperform recursive regressions based on the success ratio (directional accuracy), simulated trading return and Sharpe ratio. As the two approaches differ in how they treat old observations, this suggests that in predicting future returns, investors should drop older values as doing so improves performance.

In sum, we advance the literature by considering whether stock market factors contain useful information for investors. We do this by generating a trading rule based on one-step-ahead forecasts from rolling and recursive regressions. Thus, we provide an out-of-sample dimension, which contrasts with much of the existing literature that focuses on the cross-section dimension. We further contribute to the literature by considering the usefulness of all factors for subsequent stock returns and thus, the burgeoning factor zoo literature. As a final contribution, we compare the performance of trading rules obtained rolling and recursive regressions, with preference found for the former.

References

Ang A., Hodrick R.J., Xing Y. and Zhang X. (2006). The cross-section of volatility and expected returns. *Journal of Finance*, 61, 259-299.

Ang, A., Liu, J. and Schwarz, K. (2018). Using stocks or portfolios in tests of factor models. Available at SSRN: https://ssrn.com/abstract=1106463 or http://dx.doi.org/10.2139/ssrn.1106463

Asness, C. and Frazzini, A. (2013). The devil in HML's details. *Journal of Portfolio Management*, 39, 49-68.

Asness, C.S., Frazzini, A. and Pedersen, L.H. (2013). Quality minus junk. SSRN: https://ssrn.com/abstract=2312432

Barillas, F. and Shanken, J. (2017). Which alpha? *Review of Financial Studies*, 30, 1316-1338.

Blume, M.E. (1970). Portfolio theory: A step toward its practical application. *Journal of Business*, 43,152-73.

Bollerslev, T., Li, S.Z. and Todorov, V. (2016). Roughing up beta: Continuous versus discontinuous betas and the cross section of expected stock returns. *Journal of Financial Economics*, 120, 464-490.

Carhart, M.M. (1997). On persistence in mutual fund performance. *Journal of Finance*, 52, 57-82.

Cederburg, S. and O'Doherty, M.S. (2016). Does it pay to bet against beta? On the conditional performance of the beta anomaly. *Journal of Finance*, 71, 737-774.

Chai, D., Limkriangkrai, M. and Ji, P.I. (2017). Momentum in weekly returns: the role of intermediate-horizon past performance. *Accounting & Finance*, 57, 45-68.

Cochrane, J.H. (2011). Discount rates. *Journal of Finance*, 66, 1047-1108.

De Bondt, W. and Thaler, R. (1985). Does the stock market overreact? *Journal of Finance*, 40, 793-808.

Detzel, A. and Strauss, J. (2018). Combination return forecasts and portfolio allocation with the cross-section of book-to-market ratios. *Review of Finance*, 22, 1949-1973.

Dickson, M. (2016). Quantitative style investing. SSRN: https://ssrn.com/abstract=2781560

Elgammal, M.M., Ahmed, F.E. and McMillan, D.G. (2020). The information content of US stock market factors. *Studies in Economics and Finance*, 37, 323-346.

Fama, E.F. and French, K.R. (1992). The cross-section of expected returns. *Journal of Finance*, 47, 427-465.

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

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

Fama, E.F. and French, K.R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, 116, 1-22.

Fama, E.F. and French, K.R. (2016). Dissecting anomalies with a five-factor model. *Review of Financial Studies*, 29, 69-103.

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

Gibbons, M., Ross, S. and Shanken, J. (1989). A test of the efficiency of a given portfolio. *Econometrica*, 57, 1121-1152.

Harvey, C.R., Liu, Y. and Zhu, H. (2016). … and the cross-section of expected returns. *Review of Financial Studies*, 29, 5-68.

Hou, K., Xue, C. and Zhang, L. (2015). Digesting anomalies: An investment approach. *Review of Financial Studies*, 28, 650-705.

Hou, K., Xue, C. and Zhang, L. (2017). A comparison of new factor models. Fisher College of Business Working Paper No. 2015-03-05; Available at SSRN: https://ssrn.com/abstract=2520929 or http://dx.doi.org/10.2139/ssrn.2520929

Jegadeesh, N. (1990). Evidence of predictable behavior of security returns. *Journal of Finance*, 45, 881-898.

Lewellen, J. (2015). The cross section of expected stock returns. *Critical Finance Review*, 4, 1-44.

Liang, C., Ma, F., Li, Z. and Li, Y. (2020). Which types of commodity price information are more useful for predicting US stock market volatility? *Economic Modelling*, 93, 642-650.

Nartea, G.V., Kong, D. and Wu, J. (2017). Do extreme returns matter in emerging markets? Evidence from the Chinese stock market. *Journal of Banking & Finance*, 76, 189-197.

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

Pastor, L. and Veronesi, P. (2003). Stock valuation and learning about profitability. *Journal of Finance*, 58, 1749-1789.

Perras, P. and Wagner, N. (2020). Pricing equity-bond covariance risk: Between flight-to-quality and fear-of-missing-out. *Journal of Economic Dynamics and Control*, 121, 104009.

Salisu, A.A. and Vo, X.V. (2020). Predicting stock returns in the presence of COVID-19 pandemic: The role of health news. *International Review of Financial Analysis*, 71, 101546.

Shi, Q. and Li, B. (2020). The evaluation and comparison of three benchmark asset pricing models with daily data: Supplementary evidence. *Asia-Pacific Journal of Accounting & Economics,* (in press).

Stambaugh, R.F. and Yuan, Y. (2016). Mispricing factors. *Review of Financial Studies*, 30, 1270-1315.

Stereńczak, S. (2021). Conditional stock liquidity premium: Is Warsaw stock exchange different? *Studies in Economics and Finance,* 38, 67-85.

Vassalou, M. (2003). News related to future GDP growth as a risk factor in equity returns. *Journal of Financial Economics*, 68, 47-73.

Vassalou, M. and Xing, Y. (2004). Default risk in equity returns. *Journal of Finance*, 59, 831-868.

Vidal-García, J., Vidal, M., Boubaker, S. and Manita, R. (2019). Idiosyncratic risk and mutual fund performance. *Annals of Operations Research*, *281*, 349-372.

Wagner, N. and Winter, E. (2013). A new family of equity style indices and mutual fund performance: Do liquidity and idiosyncratic risk matter? *Journal of Empirical Finance*, 21, 69-85.

Xie, J., Hu, N., Gao, B. and Tan, CZ. (2021). Representativeness heuristic in stock market: Measurement and its predictive ability. Emerging Markets Finance and Trade, (in press).

Table 1. Summary Statistics 1990:1 to 2016:10.

	Mean	Std Dev	Skew	Kurt
Returns	0.308	9.241	-0.131	5.588
Market	0.604	4.348	-0.687	4.378
SMB	0.158	2.745	0.715	5.065
HML	0.135	2.720	0.372	5.237
CMA	0.248	2.086	0.569	5.317
PMU	0.338	2.747	-0.453	13.827
MOM	0.627	4.776	-1.813	15.046
STR	0.218	3.547	0.242	8.154
LTR	0.280	2.471	0.576	4.543
QMJ	0.442	2.860	0.19 1	5.414
VMT	-0.743	8.667	-0.346	10.167

Correlation Matrix

	Mkt	SMB	HML	CMA	PMU	MOM	STR	LTR	QMJ	VMT
Market	-	0.260	-0.218	-0.361	-0.466	-0.267	0.302	0.037	-0.688	0.479
SMB		-	-0.172	-0.138	-0.512	-0.183	0.220	0.318	-0.547	0.384
HML			-	0.704	0.389	-0.086	-0.042	0.424	0.091	0.019
CMA				-	0.250	0.053	-0.153	0.452	0.188	-0.124
PMU					-	0.100	-0.091	-0.227	0.796	-0.315
MOM						-	-0.229	0.065	0.306	-0.236
STR							-	0.021	-0.240	0.217
LTR								-	-0.293	0.152
QMJ									-	-0.510
VMT										-

The sample includes all common stocks on CRSP with current-month returns (Return, %). The numbers represent the time-series averages of the cross-sectional mean, standard deviation (‘Std Dev’), Skewness (‘Skew’), and Kurtosis (‘Kurt’) for each variable. Market = Market risk premium, SMB = the return on small minus big firms, HML= the returns on high minus low book-to-market firms; CMA= conservative minus aggressive investing firms; PMU= profitable minus unprofitable firms; MOM = the stock price continuation or momentum effect; STR= the reversal of stocks over the short-run (previous month); LTR= the long-run reversal (between one and five years returns); QMJ = high quality minus low quality (junk) firms; VMT = high volatility minus low (tranquil) volatility stocks. The individual return series is winsorised at 1%.

Table 2. Predictive Fixed Effects Panel Regression Results.

$$r_{it} = \alpha + \Sigma_j \beta_j x_{j,t-1} + \gamma_i + \varepsilon_{it}$$

Variable	CAPM	FF3	FF5	C4	All
Market	0.224	0.237	0.212	0.219	0.180
	(39.48)	(40.49)	(35.49)	(37.59)	(25.75)
SMB		-0.051	-0.091	-0.070	-0.147
		(-7.16)	(-11.35)	(-9.50)	(-16.2)
HML		0.062	0.115	0.040	0.169
		(8.16)	(12.19)	(5.24)	(13.48)
CMA			-0.054		-0.062
			(-4.84)		(-4.81)
PMU			-0.106		-0.313
			(-10.73)		(-21.6)
MOM				-0.058	-0.053
				(-14.30)	(-12.7)
STR					0.012
					(2.16)
LTR					-0.041
					(-4.20)
QMJ					0.165
					(10.26)
VMT					0.037
					(12.74)
Adj R sq.	0.008	0.008	0.008	0.008	0.011

Notes: entries are coefficient values and *t*-statistics that are robust to heteroscedasticity and autocorrelation from equation (1). Market = Market risk premium, SMB = the return on small minus big firms, HML= the returns on high minus low book-to-market firms; CMA= conservative minus aggressive investing firms; PMU= profitable minus unprofitable firms; MOM = the stock price continuation or momentum effect; STR= the reversal of stocks over the short-run (previous month); LTR= the long-run reversal (between one and five years returns); QMJ = high quality minus low quality (junk) firms; VMT = high volatility minus low (tranquil) volatility stocks. The individual return series is winsorised at the 1%.

Table 3. Statistical Forecast Results.

	A: Rolling			B: Recursive		
	ME	MAE	RMSE	ME	MAE	RMSE
	Panel Regression Based Results					
AR(1)	0.557	3.031	4.082	0.502	3.003	4.087
CAPM	0.498	2.799	3.721	0.316	2.886	3.905
FF3	0.736	2.686	3.464	0.298	2.837	3.829
FF5	0.957	2.610	3.446	0.324	2.835	3.842
C4	0.761	2.599	3.311	0.424	2.878	3.859
ALL(1)	0.347	3.212	4.217	0.259	2.750	3.723
ALL(2)	1.770	4.746	12.699	0.654	3.040	4.027
ALL(3)	0.347	4.626	7.708	0.439	3.007	3.963

Notes: Entries are the values of the mean error (ME), mean absolute error (MAE) and root mean squared error (RMSE) for the forecasts from each model. All the values are multiplied by 100. The models are: AR(1) = returns estimated with a single lag of returns; CAPM = returns estimated using the market factor only; FF3 = returns estimated using the market, size and value factors; FF5 = returns estimated using the market, size, value, investment and profit factors; C4 = returns estimated using the market, size, value and momentum factors. ALL refers to all factors considered = market, size, value, investment profit, momentum, short-term reversal, long-term reversal, quality minus junk, volatility. We consider three variants: ALL(1) exclude the volatility factor; ALL(2) exclude the market factor; ALL(3) includes all factors. The rationale behind this choice is that the market and volatility factors are likely to capture similar information. Panel A presents the results based on a 12-month fixed-window rolling regression, while in Panel B the results are based on an expanding window recursive regression.

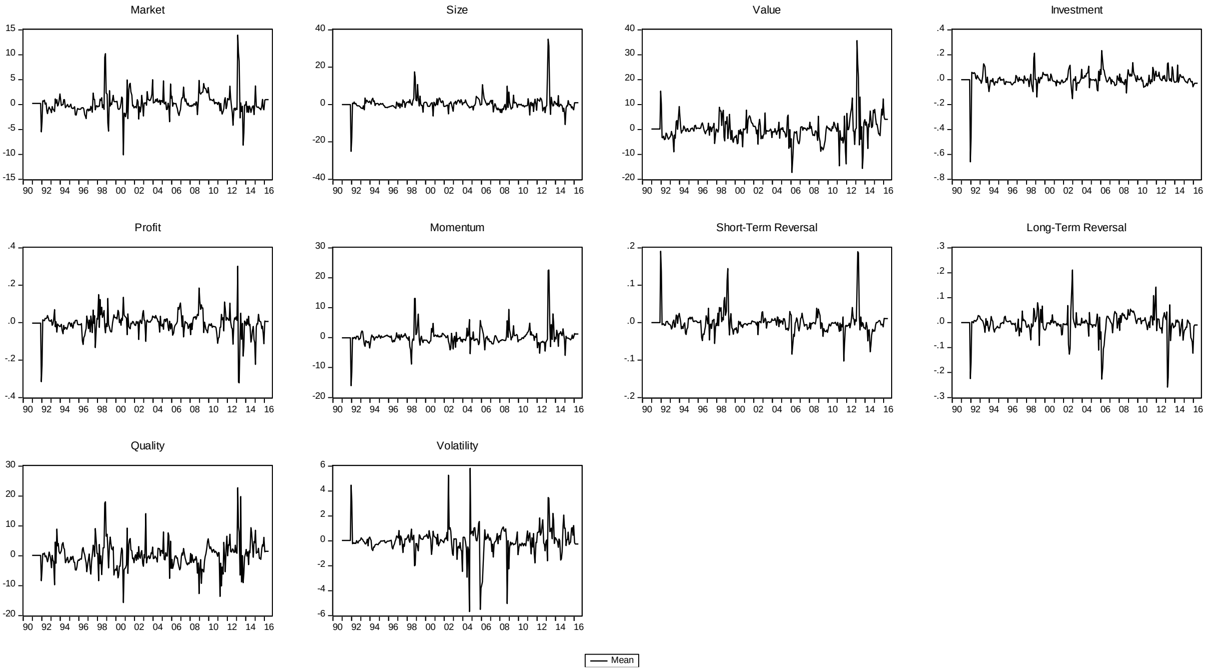
Table 4. Trading Rule Results

$$E_t(r_{t+1}) = \Sigma_j \beta_{jt} x_{j,t}$$

	A: Rolling			B: Recursive		
	Success Ratio	Return	Sharpe Ratio	Success Ratio	Return	Sharpe Ratio
	Panel Regression-Based Results					
Realised Rets	-	0.433	0.107	-	0.433	0.107
AR(1)	44%	-0.209	-0.051	41%	-0.508	-0.125
CAPM	62%	1.130	0.288	61%	0.714	0.178
FF3	63%	1.275	0.329	61%	0.824	0.206
FF5	67%	1.641	0.439	61%	0.866	0.217
C4	67%	1.609	0.429	64%	1.049	0.266
ALL(1)	72%	1.912	0.530	63%	1.218	0.313
ALL(2)	70%	1.785	0.486	52%	0.362	0.089
ALL(3)	70%	1.787	0.486	55%	0.563	0.139
	C: Individual Return Regression Results					
ALL	-	0.750	0.087	-	0.119	0.022

Notes: Entries are the success ratio, trading rule returns and Sharpe ratio based on whether the predicted returns from equation (2) are positive (buy) or negative (sell). The success ratio is the proportion of the return signs (direction) that are correctly forecast and the Sharpe ratio is calculated as the return from the simulated trading strategy over a 3-month Treasury bill dividend by the standard deviation of the trading return. Realised Rets= Realised Returns, AR(1) is estimate returns using a simple autoregressive model of order one, CAPM= estimate returns using Market factor only, FF5= the Fama-French five-factor model of the market risk premium, the SMB, HML, CMA and PMU premiums. ALL (1) is the Fama-French five-factor model plus MOM, STR, LTR, and QMJ. ALL (2) is the same as the ALL (1) model except we replace the market risk premium with the volatility premium. ALL(3) is the ALL (1) model in addition to VMT. Panel A shows 12 months of rolling regression where panel B displays recursive regression.

Figure 1. Time-Varying Coefficients.



Notes: Entries are graphical evidence that demonstrates movement in the coefficient values (vertical axis) across the sample period (horizontal axis) from re-estimating equation (1) using a twelve-month fixed window-rolling regression to examine the time-variation nature of the relation between stock return and factors.