

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ  
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ  
«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ  
«ВЫСШАЯ ШКОЛА ЭКОНОМИКИ»  
Международный институт экономики и финансов

Дергачев Кирилл Олегович

Модели оценки активов: эмпирические тесты  
(Asset pricing models: empirical evidence)

Курсовая работа - БАКАЛАВРСКАЯ РАБОТА

по направлению подготовки 38.03.01 «Экономика»  
образовательная программа «Программа двух дипломов по экономике  
НИУ ВШЭ и Лондонского университета»

Рецензент  
ст. преподаватель  
Н. К. Пирогов

Научный руководитель  
ст. преподаватель  
Н. К. Пирогов

Москва 2020

# Contents

1	Introduction	3
1.1	Asset pricing models . . . . .	3
1.2	Stock pricing . . . . .	3
2	Theoretical basis	5
2.1	New factors explaining returns . . . . .	5
2.2	Underlying research . . . . .	7
2.3	The model . . . . .	10
3	Empirical tests	11
3.1	Tools . . . . .	11
3.2	The data . . . . .	11
3.3	Procedure . . . . .	13
3.4	Results . . . . .	14
4	Conclusions	20
4.1	Shortcomings . . . . .	20

## Abstract

In this paper I'm using recent data to test whether the findings of three factor asset pricing model based on Fama and French "Common risk factors in the returns on stocks and bonds" (1993)[4] are still holding and if the factor relationships changed. The model is the one connecting excess returns on given portfolio with excess return on market portfolio, small minus big capitalization and high minus low BE/ME returns.

# 1 Introduction

## 1.1 Asset pricing models

Asset pricing models are models that describe fair prices and returns of different assets based on some underlying assumptions and theories. This is vital for a financial system as suppliers setting fair prices and buyers knowing how much an instrument is worth allow for proper functioning of financial markets. Most financial assets are hard to price completely fair (as for example for simple bonds some discount rate has to be obtained to use the basic NPV calculation and some estimates are usually used) and stocks are even harder to price in that respect as have hardly predictable cash flows if any and often pretty volatile prices with not much certainty if one tries to price them based on cash flows and if one tries to come up with the model with premiums over the risk proving that these premiums are true determinants of the asset's value is usually problematic as it can usually be argued that these could be proxies for the true determinant if they are significant.

## 1.2 Stock pricing

The two main models of stock valuation before the three factor one (3F) are dividend discount model (DDM) from M. J. Gordon and E. Shapiro (1956)[6]

and Capital Asset Pricing Model (CAPM) from Jack Treynor (1961)[15], William F. Sharpe (1964)[13], John Lintner (1965a,b)[7] and Jan Mossin (1966)[10]. The DDM uses one of the main principles of asset pricing - discounting to determine a stock's price. But to use it one should basically predict all the future dividends of the stock. The initial dividends impact the price much more than later ones but these still have to be predicted somehow and once again some discount factor has to be calculated. The CAPM model predicts excess return on an instrument comparing degree of its co-movement with the market. Unlike the 3F model that is to be discussed later CAPM is very theoretically based as it starts with some basic assumptions and then comes up with a mathematical form of returns calculation. This paper mainly focuses on the three factor model proposed in "Common risk factors in the returns on stocks and bonds" in which for the stock valuation part it was suggested that there are factors other than market that explain excess returns on stocks. These factors are BE/ME (book value of common equity to market capitalization ratio) and ME (market capitalization). In practical use one has to get excess market returns, SMB (size premium), HML (BE/ME premium) and regress the asset/portfolio in question on them; if the coefficients are assumed to be more or less stable in time, then the return on the portfolio can be determined using newer HML, SMB market and risk free data. The terms are to be discussed later in more depth. This model proved to be more powerful than CAPM that uses only the market factor even though the theory underlying CAPM suggests no other factors should influence returns on stocks. In most models of returns on assets the constant can be viewed as indicator of consistent out/under performance of an asset which is applicable for both CAPM and 3F models and is one of ways to test whether these models are missing some other important dependent variables (Merton 1973)[9].

## 2 Theoretical basis

### 2.1 New factors explaining returns

As mentioned before one of the main asset pricing instruments was CAPM. This model suggested that of all assets on the market and their combinations there are some that have greater returns for some given level of risk and assuming all agents are rational, risk-averse and have same correct information the agents will prefer these portfolios called efficient frontier (formulated by Harry Markowitz in 1952[8], Markowitz model). If under Markowitz model a risk free asset is available there is only one portfolio all agents buy and combine it with the risk free asset in order to maximize their utility whatever the utility function is (if it satisfies previous assumptions). That portfolio is called market portfolio and consists of all (value weighted) assets available if the market is in equilibrium. The degree of an asset co-movement with the market determines excess return on that asset. The formula for the model is:

$$R_a - r_f = \beta * (R_M - r_f)$$

where  $R_a$  is return on the asset,  $r_f$  is risk free rate,  $R_M$  is return on market portfolio and  $\beta$  is measure of asset co-movement with the market defined as  $\frac{Cov(R_M, R_a)}{Var(R_M)}$ . Later in multiple papers it was noted that market factor is not the only one determining return on an asset. For example Banz (1981)[1] found that size factor being company's market capitalization (ME) adds to the explanatory power compared with the model with only market premium being the independent variable. Stattman (1980)[14]; Rosenberg, Reid, and Lanstein (1985)[11]; Chan, Hamao, and Lakonishok (1991)[3] showed that book value of common equity is significant in regressions of excess returns. Basu (1983)[2] adds earnings-price ratio to test with size and market factor obtaining results supporting its significance. All of these factors were shown to be somehow connected to the dependent returns but it could be argued that some of them are just proxies for the others and some

additional tests should be conducted. In "The Cross-Section of Expected Stock Returns"(1992)[5] Fama and Fench combined above mentioned findings and tested which factors are robust to tests and thus are likely to explain returns rather than being significant by simple correlation. To fight high correlation between independent variables the stocks were split into groups based on parameters in question. For example to have variation in size and betas that are more or less independent there were obtained sets of stocks based on equal splits by pre-ranked betas and sets of stocks based on equal splits by size. Then portfolios were formed based on intersections of each of the sets in the first group with each of the sets in the second one. For more clarity as this split is important for testing of the 3F model as well say one wants 4 portfolios to run regressions based on betas and size. Then all stocks are sorted based on betas and split in two one group with bigger ones and one with smaller, then independently the same split (in two groups again) is made fo size. Now there are two sets of two groups each one for size and one for betas. Intersections of first group in the first set with the first group in the second set determines stocks to be included in the first portfolio, intersection of the first group of the first set with the second group of the second set determines the second portfolio and so on getting four of them. The portfolios are value weighted so proportion of a stock's market capitalization in sum of stock market capitalizations of the whole portfolio determines weight of said stock. Such construction of portfolios allow for more independent variation in parameters and thus to some extent mitigates high correlation which is present between many of the mentioned factors so there is more certainty that a factor affects the returns rather than just being correlated with the one that affects returns. Moreover use of portfolios allows for better parameter estimation as it decreases the part of volatility that is hardly explainable through diversification. Using Fama-MacBeth regressions the authors find that out of the mentioned factors size and BE/ME ratios absorb explanatory power of other ones and even the market factor. The relationship is that small cap and high BE/ME stocks have higher returns

and vice versa. They hypothesised that these are connected to dimensions of risk other than the ones connected to market and later in "Common Risk Factors in Stock and Bond Returns" (1993).

## 2.2 Underlying research

In "Common Risk Factors in Stock and Bond Returns" (1993) Fama and French after finding relationships in previous work (1992) suggested a model based on it and tried to expand it to include bonds and factors that are more connected to bonds so that they could also test whether the markets are integrated and thus one model could work for both. The new factors in question are TERM and DEF. TERM measures changes in interest rates and is constructed as monthly return of a long term government bond and one month bill, DEF is considered a measure of default risk for corporate bonds and is calculated as the difference between the return on a market portfolio of long-term corporate bonds and the long-term government bond return. As for the stock factors size risk measure is SML and HML is the BE/ME risk measure. The process is similar to the one described in the previous section for the dependent returns: all stocks are split in two groups based on size (small - S, large - L) and three groups based on BE/ME ratios (high - H, medium - M, low - L), all intersections of these groups make up six portfolios once again these portfolios are value weighted. The process is repeated each year so that as parameters of stocks change the portfolios are still properly displaying the groups. The returns used are monthly ones and the portfolios are changed yearly due to the fact that for sure companies disclose their statements once a year and it has to be filed with accordance with SEC 10-K<sup>[12]</sup> guidelines, general deadline for which did not change though some firms ("accelerated filers") are now obliged to have it earlier. It would be interesting to check how the results change with less cautious start of the year (e.g. taking data from April) but once again due to COVID-19 situation gathering data deemed problematic. Then HML and SMB are

calculated as:

$$SMB_t = \frac{R_{S|H,t} + R_{S|M,t} + R_{S|L,t}}{3} - \frac{R_{B|H,t} + R_{B|M,t} + R_{B|L,t}}{3}$$

$$HML_t = \frac{R_{S|H,t} + R_{B|H,t}}{2} - \frac{R_{S|L,t} + R_{B|L,t}}{2}$$

where  $R_{X,t}$  is the return of portfolio  $X$  for the period  $t$ . Market was also used in the model even though it seemed to have lost its power when BE/ME and ME factors were added in the previous paper. So the construction of variables allows to view them as some premiums for different risks as both of them are supposed to be some measure of how much stocks with some greater parameter outperform the ones for which this parameter is smaller and vice versa. Now that the model is not in accordance with CAPM assumptions the beta does not have a theoretically derived formula so the alternative used is just including excess market return in the regression and determining the coefficient with standard econometric methods. As there are three or more independent variables now the formula in statistical function form is complicated, but is clearly different from the original CAPM as depends on covariance of market with other variables. In standard matrix notation it will be in the  $\beta_{k \times 1}$  vector in it's respective position where  $\beta_{k \times 1}$  is defined as:

$$(X^T X)^{-1} X^T Y = \beta$$

where  $X_{n \times k}$  is matrix of  $n$  observations of independent variables with  $k$  variables including the constant,  $Y_{n \times 1}$  is vector of dependent variables. Dividing the stocks into three groups based on BE/ME ratio and only two on ME was justified by the fact that in plain parameter regressions the role of BE/ME was bigger than that of ME. The dependent returns are formed in the same way as other non market returns: all stocks are split by size and BE/ME into 5 groups for each of the parameters, then their intersections are the stocks to include in value weighted portfolios excess returns on which are to be explained, providing 25 portfolios to explain returns of. The split is once



again to make sure that returns are less dependent on other variables used in the regressions and thus to make sure no mere correlation is interpreted as causation. The split is to make sure the effects are tested properly, but the idea is that when they are tested there's no limitation to which portfolios are used in the model as dependent returns are constructed to have low correlation between each other and high correlation between dependent and independent variables does not cause problems in econometric analysis. In the tests themselves the authors running regressions with different variables find out how the asset types are connected and which model suits which asset best. Initially regression on TERM and DEF was analysed. Both factors were significant in both stock and bond cases and notably coefficients were greater in stock regression but for stocks a lot of variation was left unexplained. Then market alone was used in the tests. Market return captured more variation in stock returns than bond factors but still left some unexplained. For bonds it also worked well alone. SMB and HML worked well for stock portfolios but not that well for bond returns as one was more than 4 standard deviations from 0, 2 were more than 1 standard deviation and 4 others were closer, and all  $R^2$  being very close to 0. Using all three stock market factors (market, SMB, HML) adds some explanatory power to the model for bond returns and pushes much higher for stock returns compared with using market only. Then combinations of factors are tested. For example adding bond factors to three stock factors model kills explanatory power of stock factors. As it was noted that market return is highly correlated with other factors it was tested. To test the hypothesis that market is the reason the significance of tests is not robust to addition of other coefficients a test using orthogonalized market factor is conducted. It is defined as residuals of regression of excess market return on all other factors plus constant so it should be independent from them and using it one can test whether market returns absorb explanatory power of other variables. In the end it is shown that stock market factors have little effect on bond portfolios and effect of bond factors on stock returns is buried in market returns but these factors are in fact significant and thus

capture some common variation between stocks and bonds. So in order to model stock returns one could use RMO, SMB, HML, TERM and DEF or excess market returns, SMB and HML. Even though the first one is somewhat more correct as it shows underlying relationship the results for both are similar and the second version was more popular (most likely because of simplicity and it being somewhat of an improvement of CAPM which was popular before it) becoming the three factor model.

### 2.3 The model

So to sum up in from the discussed articles three factor model was derived that is still used in pricing of stocks. The form of the model is:

$$R_P - r_f = \alpha + \beta(R_M - r_f) + hHML + sSMB$$

where the unknowns are  $\alpha, \beta, h, s$  which are found based on historical data using econometric methods and then used in evaluation of performance and pricing.  $\alpha$  is the measure of stock's consistent over/underperformance.

## 3 Empirical tests

### 3.1 Tools

For the tests a multitude of tools were used. To get the main data tables Bloomberg terminal provided by HSE was used. Prices and risk-free rate data tables were collected through Yahoo finance API and U.S. Department of Treasury site using Python 3.7.6. All the data processing was done using Python 3.7.6 as well. The tables are stored in MS Excel's .xlsx format. Everything except for Bloomberg tables due to possible problems with terms of use is presented in Github repository at [https://github.com/kdergachev/FF3\\_CW2](https://github.com/kdergachev/FF3_CW2) with some further explanations of the process.

### 3.2 The data

The core of data used in the tests is a set of tables taken from Bloomberg terminal with yearly common equity and market capitalization from 1993 to 2019. All of them are for June 31 for the same reasons as in Fama French (1993) as some companies could be late with filing their statements and 6 months is plenty to ensure most/all of the data is taken from the statements and not somehow extrapolated from other data but at the same time relevant for the returns being explained. Once again the deadlines for K-10 filing generally did not change so the same date is used. The tables were obtained with EQS search for all common stocks on NYSE with added Tot CE LF field. On top of that using Yahoo finance were built tables with monthly (as of the 1st day of the month) prices of all stocks in Bloomberg tables that could have been found on Yahoo finance. The use of two different sources is due to the fact that taking all the data from the Bloomberg terminal seemed too greedy as there are some limitations on the terminal use and taking so much data was not desirable. The other option could be trying to random sample the stocks but this would be hard to implement with different new stocks appearing and others disappearing throughout the years and also I

hope that benefits coming from having a much bigger sample should negate the issues of combining different sources. Moreover with limited data from using random sample approach the portfolios are likely to be small or even empty because tests require multiple intersections of stock sets based on different parameters and even with about 1000 stocks available with the method used some of the portfolios were small or not present for some periods for 5x5 dependent portfolio splits. The risk free is taken from US department of the treasury's resource center interest rate statistics "Daily treasury yield curve rates" for "All" time periods <https://www.treasury.gov/resource-center/data-chart-center/interest-rates/Pages/TextView.aspx?data=yieldAll> (there is no download button so it was web-scraped, the code can be found in Github repo mentioned above). Until 07/31/01 the data for one month rates is not available so instead it was taken from longer maturity rates. In both cases it was determined as  $Y * 30/360$  where  $Y$  is the number presented on the site as the data there is annualised yield (in percent) with assumption of 360 day year thus the formula takes it back to monthly terms. The first years rates are not exactly accurate as are calculated as if the yield curve was flat which is rarely the case, but the results show that most likely the error was not that big. All the tests will be constructed around the prices data as it is a reduction of the Bloomberg dataset and only the stocks with available data are to be used. In said table there are 324 entries (monthly prices) of 1036 stocks. This will provide a good sample for the time series regressions to be conducted and enough stocks to properly assign them to portfolios to control for variables. The prices used are adjusted so they should include info other than pure price so that proportion of price to some previous price should provide return over chosen period for the stock. With this the starting dataset is complete and using it all the necessary tests can be conducted.

### 3.3 Procedure

The procedure is very similar to the FF3 one. Each year corresponding Bloomberg table is taken. Then from Tot CE LF and Market Cap fields BE/ME parameter is calculated and added to the table. Then of the total one month yield rates table the rates for the period are taken which are then used as risk free. Market returns is calculated next. As it is usually done they are calculated as returns on value weighted portfolio of all stocks in the analysis as it seemed like the size of the sample allowed to to it rather than relying on some index. Subtracting risk free from said market returns gets excess market return dependent variable. After that HML and SMB are constructed. The procedure is the same once again as I decided to vary the dependent returns portfolios rather than composition of SMB and HML ones. Thus the division is 2 by size and 3 by BE/ME ratios. With six portfolios stocks obtained the value weighted portfolio returns are computed and then the premiums are obtained. This procedure is done for all the years and all returns are concatenated into one table with monthly entries for the whole period. I would like to add that for specifically my take on the tests the subsets are named S1, S2, S3... and BE1, BE2, BE3... The portfolios of intersections are noted as BE1|S1, BE1|S2... Where Sx are size subdivisions and BEx are BE/ME subdivisions and lower number x means higher parameter so S1 are stocks with highest ME values and S5 are with lowest in cases where the portfolios were divided into 5 size groups. When the table is obtained for all variables and periods each of BEx|Sx is regressed on excess market return, SMB and HML with constant. With available data 5x5 split of dependent returns was more or less adequate with exception of BE5|S5 which only provided data for 2 years and then became empty for all other periods. Showing all the results will take up a lot of space so only few regressions will be present in this work but all results for other subdivisions are saved in .xlsx format and accessible in the Github repo referenced. Though in portfolios BE parameter stands first in the tables S is the index and BE is in columns.

### 3.4 Results

As mentioned before  $5 \times 5$  split of dependent returns gives mostly good portfolios, but BE5|S5 is present for only 2 years data, in fact in most splits tested the lowest size/lowest BE/ME portfolios are small or disappear which is somewhat expected if the parameters don't perfectly move together. So in the paper I would like to present the classic  $5 \times 5$  result and just note that for BE5|S5 the results should be given much less value due to small sample and  $3 \times 4$  to have tests with more complete samples present in this work. The results of the tests are to some extent in line with findings of Fama and French. The significance of all the models is high and the trends in coefficients is similar.

$$R_P - r_f = \alpha + \beta(R_M - r_f) + hHML + sSMB$$

One can see some parameters of the regression with  $5 \times 5$  dependent portfolio split in table 1 (Note: in all results values are rounded to 5 decimal places or presented in scientific notation with mantissa rounded to three decimal place in order to present comparable parameters even for very small numbers).

Table 1: Regression parameters of  $5 \times 5$  dependent return portfolio splits

These table contain results of regressions of 25 dependent portfolios excess returns from five by five splits by BE/ME and market cap on constant, market portfolio's excess returns, HML and SMB returns. The model is of form  $R_P - r_f = \alpha + \beta(R_M - r_f) + hHML + sSMB$ .  $p(X)$  is p-value of some X,  $\alpha, \beta, h, s$  are respective coefficient values and  $R^2$  and F-statistic are respective measures of fit of the model overall.

	BE1	BE2	BE3	BE4	BE5	BE1	BE2	BE3	BE4	BE5
	$\alpha$					$\beta$				
S1	-0.00997	-0.01171	-0.00949	-0.00405	-0.00683	0.93024	0.94538	0.94021	0.95068	0.93707
S2	0.00349	-0.00406	-0.00750	-0.00242	-0.00640	0.99263	0.945645	0.94187	0.95783	0.95046
S3	-0.01221	-0.00730	-0.00786	-0.00935	-0.01169	0.93074	0.96408	0.95839	0.97060	0.96120
S4	-0.01079	-0.00389	-0.01269	-0.00463	-0.00218	0.94474	0.96771	0.93548	0.95811	0.98837
S5	-0.02095	-0.01524	-0.00636	-0.01617	-0.16364	0.89791	0.87905	0.93801	0.94119	0.53802
	$h$					$s$				
S1	0.05563	-0.05015	-0.11597	-0.15501	-0.19078	0.18149	0.06942	0.04787	0.01306	-0.04123
S2	0.06005	-0.07260	-0.11346	-0.12469	-0.23417	-0.01967	0.07317	0.13873	0.11125	0.29743
S3	0.31444	-0.05619	-0.19558	-0.43635	-0.38930	0.31237	0.21015	0.27093	0.48014	0.44834
S4	0.05389	-0.11092	-0.22323	-0.23138	-0.33370	0.24364	0.23390	0.31946	0.24792	0.27607
S5	0.34907	0.08346	-0.07592	-0.08392	-1.25338	0.67751	0.24577	0.12980	0.34959	1.09563
	$p(\alpha)$					$p(\beta)$				
S1	0.10869	0.00575	0.01081	0.21529	0.02132	3.102e-129	2.418e-177	1.002e-192	2.782e-210	3.840e-221
S2	0.50297	0.33213	0.09771	0.59347	0.17148	1.873e-157	2.197e-178	3.492e-168	7.191e-170	2.932e-165
S3	0.11132	0.19293	0.08928	0.10275	0.08096	3.739e-106	2.819e-145	7.670e-168	1.588e-143	1.706e-124
S4	0.08675	0.52843	0.02047	0.47759	0.77468	1.696e-129	2.554e-134	8.662e-145	6.137e-127	1.044e-106
S5	0.04806	0.07493	0.34332	0.25797	0.13198	1.659e-71	3.185e-89	2.487e-121	1.120e-46	0.05838
	$p(h)$					$p(s)$				
S1	0.26700	0.14113	0.00013	1.017e-08	2.503e-14	0.00407	0.10429	0.20239	0.69258	0.16872
S2	0.15359	0.03206	0.00204	0.00074	1.693e-09	0.70885	0.08441	0.00262	0.01585	1.064e-09
S3	6.160e-07	0.21428	2.769e-07	8.376e-19	4.021e-12	6.874e-05	0.00024	1.583e-08	3.277e-15	1.459e-10
S4	0.28852	0.02658	6.738e-07	1.487e-05	1.260e-07	0.00015	0.00021	1.711e-08	0.00020	0.00043
S5	1,853E-11	0,00239	0,42193	0,99185	0,00032	1,332E-24	1,972E-09	1,038E-10	0,00070	0,00789

Table 1 continued

	BE1	BE2	BE3	BE4	BE5	BE1	BE2	BE3	BE4	BE5
	$R^2$					F-statistic				
S1	0.85772	0.92931	0.94359	0.95644	0.96279	618.919	1349.72	1717.31	2254.27	2656.59
S2	0.90451	0.93037	0.91945	0.92114	0.91771	972.454	1371.8	1171.82	1199.18	1144.97
S3	0.80902	0.88771	0.92046	0.89010	0.85228	434.913	811.614	1188.05	831.486	592.358
S4	0.85951	0.86774	0.88852	0.85238	0.81391	628.081	673.574	818.306	592.821	431.554
S5	0,40973	0,21062	0,22544	0,09738	0,69693	71,26429	27,39377	29,88198	8,66669	15,33058



As one can see in provided results the regressions proved to be powerful in explaining stock returns used and explaining most of the variation. It can be seen that both in absolute values and in standard deviations most of the constants are not significantly far from zero as indicated by high p-values, for which 15 out of 25 are higher than 0.1 meaning that most portfolios do not have returns in excess over the ones predicted by the model and thus that the model includes if not all but most or the most important parameters. Other coefficients are mostly significant. Most of the regressions the market factor coefficient is a bit smaller than 1 and very significant for all but the BE5|S5 portfolio which caused concerns before the regressions were run. HML and SMB coefficients are showing the stated before dependencies: values of  $h$  decrease with BE/ME ratio's decrease keeping size constant as high BE/ME firms are found to outperform low BE/ME and HML variable is measure of that outperformance; similarly  $s$  increases with decrease in size keeping BE/ME constant with variable once again measuring outperformance, but this time of small companies over big ones. Significance for both of them is generally high: 15 of 25 portfolios give lower than 0.01 p-value for  $h$  and 18 of 25 for  $s$ . In absolute values the coefficients are mostly in range of  $[-1, 1]$  but I wouldn't call them close to zero. As for the overall fit the models prove very good similarly to the results of Fama and French. F statistics are all higher than 8 and in fact most of them are greater than 100 meaning that overall regression is very significant.  $R^2$  are ranging from 0.21 to 0.96 but the low ones are in smallest size portfolios and in others the values are not less than 0.8 which proves that chosen factors are very powerful in explaining variation in used stocks' returns. These results prove that as I hoped the size of the sample helped provide data good enough to test the model as getting results like the ones obtained with a wrongly collected data would be very unlikely. Nevertheless to make some proper conclusions better collected (or assessed) data is required. Still these test prove that for newer data, and in fact the data used for the "Common Risk Factors in the Returns on Stocks and Bonds" are up to 1993 while for this work data was used from 1993 to

2018 as Bloomberg did not provide historical data before 1993 with EQS search, the three factor model is a very powerful tool in evaluating returns on portfolios. In "Common Risk Factors in Stock and Bond Returns" the authors used one portfolio split and mentioned that their hope was that it did not matter for the tests. The results of  $3 \times 4$  dependent portfolio splits are presented in Table 2. In this table although the dependent returns portfolios are different from the previous tests the parameters presented are very similar to the ones in  $5 \times 5$  in absolute values, distance from zero in standard variations, direction and sign of change and overall goodness of fit. In fact the results are even more in line with expectations in the  $3 \times 4$  tests but one could argue that these results are more due to the fact that  $3 \times 4$  division allows for better portfolio construction rather than  $5 \times 5$  one due to sample not being as big. Anyway generally the results are very similar so there is not much evidence that dependent portfolio splits affect quality of the tests on a significant level. The tests were also conducted for all possible splits of type  $a \times b$  where  $a \in [3, 5], b \in [3, 5]$  as more splits were certainly to provide empty portfolios and less turn out to be too closely related with one of the dependent returns. Nine tables were obtained in the end and throughout all of them the results are similar proving the authors right in believing that the choice of splits did not matter much.

Table 2: Regression parameters of  $3 \times 4$  dependent return portfolio splits  
These table contain results of regressions of 12 dependent portfolios excess returns from three by four splits by BE/ME and market cap on constant, market portfolio's excess returns, HML and SMB returns. The model is of form  $R_P - r_f = \alpha + \beta(R_M - r_f) + hHML + sSMB$ .  $p(X)$  is p-value of some X,  $\alpha, \beta, h, s$  are respective coefficient values and  $R^2$  and F-statistic are respective measures of fit of the model overall.

	BE1	BE2	BE3	BE4	BE1	BE2	BE3	BE4
	$\alpha$				$\beta$			
S1	-0.00300	-0.00808	-0.00567	-0.00640	0.98852	0.94941	0.95087	0.93991
S2	-0.01061	-0.00946	-0.00217	-0.01412	0.95288	0.94451	0.96232	0.97070
S3	-0.02196	-0.00912	-0.01053	-0.02223	0.91885	0.91351	0.95541	0.90044
	$h$				$s$			
S1	0.11866	-0.06642	-0.15711	-0.18587	0.11220	0.04414	0.03260	-0.02949
S2	0.37684	-0.14977	-0.20515	-0.43840	0.26779	0.25898	0.21515	0.48090
S3	0.57400	0.00882	-0.23891	-0.61798	0.63233	0.24918	0.32014	0.59136
	$p(\alpha)$				$p(\beta)$			
S1	0.49600	0.019543	0.08941	0.02471	2.650e-177	3.329e-203	1.136e-207	9.056e-227
S2	0.12769	0.03912	0.61386	0.01481	4.780e-119	3.975e-167	5.753e-177	1.406e-142
S3	0.02549	0.22060	0.08439	0.00233	2.180e-80	2.986e-107	2.343e-134	1.273e-97
	$p(h)$				$p(s)$			
S1	0.00095	0.01744	1.279e-08	1.133e-14	0.01233	0.20606	0.33392	0.30491
S2	8.996e-11	6.236e-05	8.989e-09	1.068e-18	0.00017	4.757e-08	1.252e-06	4.760e-15
S3	3.066e-12	0.88323	1.850e-06	3.101e-21	6.317e-10	0.00103	3.683e-07	7.321e-14
	$R^2$				F-statistic			
S1	0.93056	0.95183	0.95480	0.96583	1375.78	2028.6	2168.65	2901.82
S2	0.84339	0.91953	0.93004	0.88854	552.871	1173.16	1364.84	818.445
S3	0.74554	0.80308	0.86943	0.81589	300.804	418.7	683.641	419.522

## 4 Conclusions

In conclusion the three factor model is proven to be a powerful tool in explaining stock returns on newer data as it was when the tests were first conducted unlike for example beta alone used in CAPM. Even though the results obtained are good and explain a lot of variation in stock returns as it turned out there are some other factors that were found to add to the model later which are not discussed in this paper. They could be the ones explaining the variation left, but even though the most popular multifactor model is still the three factor one. This model is widely used due to its high explanatory power and comparative ease of use as basically it requires the same data as the one needed for the CAPM unless one uses proxies like S&P500 instead of calculating the parameters. The data required is risk free which can be easily obtained on the treasury's site and some set of stocks to represent the market with prices ME and BE which is harder to get but for those with access to services providing the data the model is not that difficult to use. The constant (as in Merton) is once again a very simple way to evaluate excess of the fair returns of a portfolio.

### 4.1 Shortcomings

Some additional tests were were not conducted to test robustness on some smaller horizon effects like January effect, September effect and some other parameters like E/P or TERM and DEF that were used in the same paper that introduced the three factor model.

## References

- [1] R Bantz. The relationship between return and market values of common stock. *Journal of Financial Economics*, 9(1):3–18, 1981.
- [2] Sanjoy Basu. Earnings' yield and the size effect. *Journal of Financial Economics*, 12:129–156, 1983.

- [3] Louis KC Chan, Yasushi Hamao, and Josef Lakonishok. Fundamentals and stock returns in japan. *The journal of finance*, 46(5):1739–1764, 1991.
- [4] Fama Eugene and French Kenneth. Common risk factors in the returns on stocks and bonds. *Journal of financial economics*, 1993.
- [5] Eugene F Fama and Kenneth R French. The cross-section of expected stock returns. *the Journal of Finance*, 47(2):427–465, 1992.
- [6] Myron J Gordon and Eli Shapiro. Capital equipment analysis: the required rate of profit. *Management science*, 3(1):102–110, 1956.
- [7] John Lintner. Security prices, risk, and maximal gains from diversification. *The journal of finance*, 20(4):587–615, 1965.
- [8] Harry Markowitz. Selection, portfolio and diversification, e. *The Journal of Finance*, 7(1):77–91, 1952.
- [9] Robert C Merton. An intertemporal capital asset pricing model. *Econometrica: Journal of the Econometric Society*, pages 867–887, 1973.
- [10] Jan Mossin. Equilibrium in a capital asset market. *Econometrica: Journal of the econometric society*, pages 768–783, 1966.
- [11] Barr Rosenberg, Kenneth Reid, and Ronald Lanstein. Persuasive evidence of market inefficiency. *The Journal of Portfolio Management*, 11(3):9–16, 1985.
- [12] SEC. Sec k-10 guidelines (expires october 31, 2022), <https://www.sec.gov/files/form10-k.pdf>. [Online; accessed 27-June-2020].
- [13] William F Sharpe. Capital asset prices: A theory of market equilibrium under conditions of risk. *The journal of finance*, 19(3):425–442, 1964.
- [14] Dennis Stattman. Book values and stock returns. *The Chicago MBA: A journal of selected papers*, 4(1):25–45, 1980.

- [15] Jack L Treynor. Market value, time, and risk. *Time, and Risk* (August 8, 1961), 1961.