Radboud University

Bachelor Thesis

Artificial Intelligence department

Beating the market, using linear regression to outperform the market average

Author:
Jelle Verstegen

Supervisors:
Marcel van Gerven
Pasi Jylänki

August 8, 2014

Abstract

In this thesis we set out to find out whether we could use linear models on financial data to make stocks picks that return above the market average. The linear models were successful in outperforming the market over different periods, though the model performed best when it picked stocks for a period of 2 years. The models also worked together in a committee to create and manage a portfolio for over a decade, the return of the portfolio was above the average return of the population. There was however a survivorship bias in the data which has a considerable effect on the average population performance. Though our models outperformed this high population average, it remains to be seen whether the models would outperform in a population without the survivorship bias.

Contents

1 Introduction

Machine learning is becoming ever more important in the financial world, which given the large amount of publicly available financial data is probably no surprise. There has been a lot of research in to applying machine learning on financial markets, with different focus points[1, 21]. If we look specifically at the stock market, this gives rise to some interesting questions. Specifically, can linear models use fundamental financial data to find stocks which perform above average. We could not find an answer to that very question from the literature though theories do give an indication on whether it is even possible.

According to the efficient market hypothesis, linear models should not be able to use fundamental financial data, or any other data, to find above average performing stocks.[8, 9] According to the efficient market hypothesis any publicly available data is always factored into the price by the market. So when a model receives the inputs, those can not be used to make predictions about the price, because the inputs are already factored into the price.

If the hypothesis holds, linear models should not be able to exploit relations between inputs and targets to make above average returns by picking stocks.

However, other research shows that there are indeed certain financial data which might have an explanatory value for predicting the future[13, 3]. This forecasting of stock prices, or stock price movements, should be possible using certain financial data[5, 4]. If this research is correct, hopefully linear models will be able to pick up on those relations and will be able to exploit the explanatory value the financial data has. In that case they might be able to pick stocks which perform above average.

Most researched focused on a more simple decision, should you invest in a stock-index or a risk-free asset like treasury bills. The models in that research had to make an A or B decision. This is very different from what we wish to find out. We wish to find out whether it is possible to find stocks which perform above the average of a population of hundreds or even thousands of stocks. Given that there does not seem to be a publicly available answer with a solid explanation of the methods used and the results obtained, we set out to find out whether certain financial data have predictive value on stock picking and whether linear models can exploit them to make above average stock picks.

1.1 Research goal

The main goal was to find out whether a linear odel could be used to make investments that could outperform the population average, which is simply the average of all the possible investments it could have picked. In our dataset, the amount of investments varies from roughly 800 to 2000 depending on the time point from which the data is taken. Given the difficulty of the goal, we are merely interested in whether it is possible, or plausible. If so, the implication would be that a linear model could be used to either make investments on its own, or help its human counterpart to make investments.

1.2 Task of the model

The model has to form a sub-population, which will be called the portfolio of the entire population. The average percentage change in stock prices of this portfolio has to be higher than the average percentage change in stock prices of the entire population. In essence, this means the model has to create an portfolio of companies in which it will invest. The performance of the model will then be determined by looking at the return of its portfolio and comparing it to the return of the entire population. It can pick any company from the NASDAQ or NYSE for which ADVFN[14] has data.

The task of the linear model is not an easy one, it is a task at which the average investor is unsuccessful[2], and there have been documented cases where experts failed[10]. Given the fact that the task is difficult even for humans, the model can only decide to buy companies, or go long. It can not go short, it can not use options, it can only go long on companies. This means the total amount of investments options it has is considerably smaller than the average investor or financial expert. It also means the model can not use any leverage.

After creating an initial portfolio, the model has to manage it throughout the time period. Every time it receives new financial data it will have to analyze the new data. Based on that new data it may have to make adjustments to its portfolio meaning that it might have to sell some of the companies in the portfolio so that it can then buy others which are suggested to be better investments by the new data. Managing the portfolio over time is a part of its task, though it should be noted it receives new data only on a quarterly basis for the ADVFN data.

1.3 Measuring performance

The performance of the model will be measured mainly by its success in creating and managing the portfolio's of companies it invests in. Specifically we want the return of the portfolio to be higher than the population (or market) average. Other assessment factors used will be how good the fit of its predictions is, how much of the movement in stock prices it can explain and how similar the companies it invests in are to the actual best companies. Lastly a list of companies it invests in the most gives a more tangible indication for what the model invests in, it will also allow us to look at how many changes it makes to its portfolio.

2 Data

The first data-set comes from Quandl[16], they have been kind enough to gather a lot of financial data and make it publicly available for free. I used Quandl's data mainly for the initial model creation. The second data-set comes from ADVFN[14], they have a lot of information available and have been kind enough to give permission to use their data during research. The data from ADVFN was used to achieve the results described in this thesis. Pricing data was obtained from Yahoo Finance[20], they provide a API to gather the pricing data I needed. The data provided by Quandl, ADVFN and Yahoo play an instrumental role in my Thesis, therefore I am very grateful for the data provided by them.

2.1 Formal representation

Let $\mathbb{D}$ be the entire data set, with inputs $\mathbb{X} = \{x_1, x_2, .., x_n\}$ and outputs $\mathbb{Y} = \{y_1, y_2, .., y_n\}$, where n is the number of companies in the data set $\mathbb{D}$.

$\mathbb{X}$ is a design matrix of n rows, and m columns, where m is the number of features.

$\mathbb{Y}$ is a column vector of n continuous values between $\{-1, \infty]$, where -1 means -100%, so this means a stock loses all of its value. Please notes that this never happens because of the survivorship bias. Each value in $\mathbb{Y}$ represents the percentage stock change of a company i over a period p. Let t be a certain point in time, than the percentage stock change is

$$y_i = \frac{\text{price}_{i,t+p} - \text{price}_{i,t}}{\text{price}_{i,t}}.$$

2.2 Time frame

The Quandl data-set has data starting from 2000 up to now, the ADVFN data-set has data starting from 1994 up to now. The Quandl data is annual, whereas the ADVFN data is quarterly. The Quandl data has 12 different time points for which there is data. The ADVFN data is a lot bigger, it has 13 years worth of data, and the data is quarterly, so about 52 time points for which there is data. It is worth noting that the quarter and or year from which the data is taken can have quite a large effect on the data. The input variables are very different during the height of the dot-com bubble when compared to most other years. The financial crisis of 2008 is another example, the data a few years before the financial data clearly has different effects than the data a few years after the crisis. Because of this, the more years available to train the model on can have real beneficial effects.

2.3 Amount of features

The amount of features is different for each data set. In the ADVFN data set, there are 250 features ($m = 250$). In the Quandl data set there are less features, 69 to be specific ($m = 69$). Though only a small selection of the variables is used for training the models. The selection used can be found in the Appendix.

2.4 Market listing

The data consists of hundreds of companies listed at the NASDAQ or NYSE. Only companies listed at the NASDAQ or NYSE are in the data set, given the fact that the United States financial market is highly transparent and thus has a lot of data which can be used. The number of companies changes over time, but for the ADVFN data the number of companies goes from roughly 800 in 1994, to 2000 companies in 2013. The Quandl data has roughly 1000 companies in the the year 2000 and roughly 2000 companies in 2013. The difference in the amount of companies is mainly due to the survivor bias in the data.

2.5 Survivor bias

There exists a survivorship bias in the data, because only companies that survived up to now are in it. This means that companies that were delisted at some point, are simply not in the data. This causes performance to be considerably higher than it should be. There can be several reasons why a company is delisted, bankruptcy, acquisition, not being in compliance with the listing requirements or voluntarily delisting. This means that the data does not contain any companies that go bankrupt, which obviously helps the model. The performance of the model is almost certainly higher than it would be if the population contained no survivorship bias. The model should not be expected to perform as well in real life, simply because it is no longer being helped by the survivorship bias.

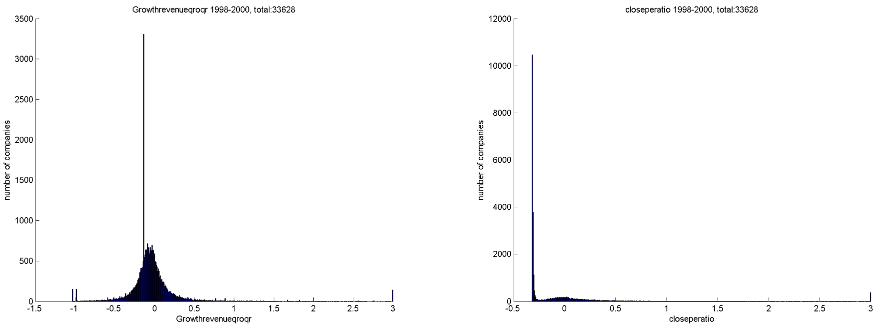
Because of the survivorship bias, the return on its own is not very meaningful. The return compared to the return of the population on the other hand is meaningful. If the model can outperform the return of the population as a whole, we might reasonably expect the model to outperform a population which is not affected by the survivorship bias. The results below generally put a lot more weight on the performance compared to the average performance of the population, simply because that should give us more information on how the model will perform in a more realistic population.

2.6 Input feature distributions

The features are usually similar to a Gaussian distribution, though some more closely resemble a inverse gamma distribution or even an exponential distribution. Some features have a large spike at '0', because for some features missing values, or nonsensical values are set to zero. Almost all features also have a large number of outliers, which may be as large as 30 standard deviations from the mean. This means that the features initially might not align with the assumption of the model that the features are Gaussian. There are however several transformations to make the features more Gaussian. Which includes removing (or moving) the outliers and removing the zero-values.

To deal with the long right-tail the exponent transformation can be used. The difficulty lies in the fact that the features are taken out of different time-points and different time-points usually require different exponent transformations. You do not know in advance what the distributions in the test set will be like.

Dealing with the zero-values is more difficult, they can not be removed because a company might have normal values for dozens of features and a zero value for one or two. Because of that removing companies which happen to have a zero value for a feature will lead to a very small data-set. A feature which has a zero-value should have no effect on the prediction



(a) Gaussian with zero-values spike (b) Exponential distribution with outliers

Figure 1: Two example distributions, the first is the distribution for the quarterly revenue growth and the second is the price to earnings growth.

for that company. Because the coefficients are multiplied with the features to form the prediction, a zero value will have no effect on the prediction. However, in our case the variables are standardized and the zero values become non-zero through the standardization. So in our case, they do effect the predicted value. For this reason another standardization which ignored the zero values was used, so that zero remained zero.

2.6.1 Gaussian with zero-values spike

Here is an example distribution in which the input variable closely resembles a Gaussian distribution, namely the growth in revenue from one quarter to the next. However, for roughly one tenth of the companies no previous quarter revenue was reported, so there is no increase or decrease in revenue and because of it the quarterly revenue growth is zero. This is basically a missing value for one tenth of the companies, and it leads to a spike in the distribution. Note that the values are standardized, so the spike is for companies that have 0 quarterly revenue growth, but standardized their value is not zero. This data distribution is common for the input variables related to growth.

2.6.2 Exponential distribution with outliers

The valuation ratios have a rather difficult distribution. They have a rather large spike at zero when the valuation ratio does not make sense. For instance, for the price to earnings, when a company has no earnings, or negative earnings (a loss) the ratio is set to zero. This means that only companies that had positive earnings have a price to earnings ratio, the rest have zero-values. Then there is the problem that there is quite a large group of companies below the mean, and a considerably smaller group above the mean. This is because the value lies in the $[0, \infty)$ interval. It can be extremely large for some companies, whereas most companies lie below the mean.

2.7 Limitations of the data

The data consists only of fundamental data, which means the model has access to limited data on any company. An annual or quarterly report consists of dozens of pages and our model has access only to the balance sheet, income statement and cash flow statement. For this reason it does not have access to a lot of other valuable information:

- Company strategy,
- Risks associated with the company,
- Risks and changes in the industry / sector,
- Market sentiment,
- Technical data,
- Macro-economic data,
- Competitive advantages of any company.

Many other sources of information are also unavailable to the model. For this reason it is plausible that the model is only able to explain very little of the stock price movements. The hope is that what it can explain is sufficient to pick stocks that will do above average in terms of return.

3 Method

To learn the relationship between the inputs and the targets linear regression was used. This assumes there is some linear relation between the inputs and targets. Two different approaches were used, a least squares approach was used to determine the coefficients and a Bayesian regression approach was also used.

3.1 Linear model

Multiple linear regression was used to calculate the coefficients for the linear model. To obtain the coefficients two methods were used: least squares and Bayesian linear regression. The model focuses on predicting percentage change in stock prices given certain inputs. It makes these predictions by first finding a set of coefficients which best fit the training data, in which the best fit is determined by minimizing a certain cost function. After it has found the coefficients (β), it makes the predictions by multiplying the coefficients with the input variables for a given company

$$y = X\beta.$$

There is however quite some noise which the model could not explain. As discussed in the previous section 2.7, the linear model only has access to very limited information on any given company. It has no access to company strategy, competitive advantages, management information, market sentiment etc. Because of this, one would expect the noise to be large and this is what we found when we applied the model on our data. To account for the noise in the observations which can not be explained, a noise factor ε can be introduced

$$y = X\beta + \varepsilon.$$

3.2 Assumptions of the linear model

- *Linearity:* The linear model assumes that the outputs are a linear combination of the coefficients and the input variables. If no such linear combination exists, the model will fail to fit even the training data properly. The input variables indicated that the effect might not be strictly linear, and some input variables seemed to have a polynomial effect on the outputs. In any case, we did find that the linear combination which minimized the cost, still had a considerably large error, which might have been smaller if no linearity was assumed.

- *Constant variance:* The model assumes the outputs have a constant variance for the error term. But this is not the case, in the stock market price changes do not have constant variance noise. There are Bayesian approaches which do not make this assumption, though we did not use them. Our approach does not implement variable noise variances, though there are approaches which do implement it[11, 19]. These approaches which implement variable noise variance might work better than the methods we proposed, because the stock market simply does not have constant noise variance. However due to lack of time we did not have the time to try models with variable noise variance, which is something that could be tried in future research.

3.3 Obtaining the coefficients

The coefficients are used to make the predictions, but they have to be learned from the training data. Two approaches were used to determine the coefficients, a least squares approach and a Bayesian approach.

3.3.1 Least squares normal

The coefficients can be calculated using a ordinary least squares approach

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

where X is the design matrix and y are the targets. After the coefficients β are calculated, the predicted outputs are calculated as follows

$$y = X\beta.$$

No regularization was used because the training fit itself was not that good to begin with. As there were a lot of ways to improve the model and its performance, regularization was not a top priority. Even without regularization the model was not necessarily over-fitting, especially if multiple quarters are used for training data. When multiple quarters are used the model can learn from different market sentiments and can find a more general way to predict changes in stock prices. If only a single quarter is used for training data the model might learn one specific market sentiment really well and not be able to explain the price changes in the future. If regularization is used however, the factor should be determined with the help of a cross-validation set. As shrinking of the weights should not be too high because that could hurt the performance of the model. regularization in the least squares approach would lead to the following method to obtain the coefficients

$$\beta = (X^{\mathrm{T}} X + \kappa I) X^{\mathrm{T}} y.$$

Here κ is the regularization parameter, which determines the amount of regularization. If we look at the training data fit, it seems that the regularization parameter should be relatively small. But this is best determined by trying several regularization parameter values and determining which provides the best fit on a cross-validation set.

3.3.2 Bayesian Linear regression

The predictions are again calculated as follows

$$y = X\beta,$$

and the observation model is defined as

$$p(y|w = \beta, x, \sigma^2) = \mathcal{N}(y|\beta^{\mathrm{T}} x, \sigma^2).$$

Where the distribution of the error term e is centered around 0 and has a standard deviation of σ^2

$$e \sim N(0, \sigma^2).$$

Here σ^2 is τ^{-1} and τ is the variance. $N(0, \sigma^2)$ is a Gaussian distribution with a mean of zero and σ^2 variance. In general $N(\mu, \sigma^2)$ is a Gaussian with a mean of μ and with variance of σ^2. The prior probability distribution of the coefficients and the variance is a conjugate normal inverse-gamma

$$p(\beta, \tau \mid \alpha) = \mathcal{N}(\beta \mid 0, (\tau\alpha)^{-1}\mathcal{I})Gam(\tau \mid a_0, b_0),$$

where $Gam()$ is the gamma distribution.The posterior probability distribution is defined as

$$p(\beta, \tau, \alpha) = \frac{p(y|\beta, \tau, X)p(\beta, \tau|\alpha)p(\alpha)}{\int p(y|\beta, \tau, X)p(\beta, \tau|\alpha)p(\alpha)d\beta d\tau d\alpha},$$

where the parameter α is assigned the hyper-prior

$$p(\alpha) = Gam(\alpha \mid c_0, d_0).$$

Because of the hyper-prior, there is no analytical solution for the posteriors. The the denominator and other expectations such as mean and covariance with respect to the posterior of Bayesian inference are analytically intractable. Therefore we use an approximation method for which we used variational Bayesian Inference. The process of the variational Bayesian Inference can be found in a paper by Drugowitsch[6].

4 Single-model experiment

This experiment encompasses the method and results of a single-model creating the portfolio. This model does not yet maintain the portfolio, it simply creates them for a specific time period. The specific task in this experiment is to train on a training-set, then to make predictions for the test-set and finally to form a portfolio of 30 stocks from the test set. Training on the training-set can be done through any of the regression methods mentioned above, the numerical linear regression or Bayesian linear regression.

A number of variables can be changed, one with a considerable impact is how many years into the future it has to make the predictions. Asking the model to predict 1-year into the future is very different from asking it to predict 5-years into the future. Predicting can be hard if the predictions have to be made on the very short-term, it seems that the error-term is considerably higher when compared to a period which better suits the model. Likewise, the error-term is considerably higher if the predictions have to be made for the very long-term. For the short-term, the market-effect might simply be very large. For the long-term the input-variables might no longer correlate enough with the targets, because there is too much time between them. In the end, a period around 8 quarters, or 2 years seems to be best for our model with the input-variables that it has.

Another variable which has a considerable impact is how many quarters of training data is taken. At the beginning, tests were done with simply 1 quarter of training data, but the model would have a considerable over-fit on a specific market sentiment. To counter the over-fitting, the model can be trained on multiple quarters so it trains on different market-sentiments. This prevents the model from over-fitting on a specific market-sentiment, and allows it to be better prepared for a different market-sentiment in the test-data. Because the fit on the training set was not that good to begin with, regularization was not a top priority. Using multiple quarters was the most important approach for reducing the over-fitting. It should be noted that regularization might still be beneficial, even though we did not take advantage of it in this experiment.

In the ADVFN-data set it creates a new portfolio each quarter then portfolio is held over the period for which the model had to predict. The average performance of the portfolio is then compared to the average of the entire population. Of interest is whether the portfolio of the model outperforms the population as a whole.

4.1 Method

The model follows a certain amount of steps that combined produce a portfolio of 30 stocks, first it prepares the training-data, then it performs the regression and computes the coefficients. It then extracts the test-data and using the coefficients it makes the predictions. After it has computed the predictions, it uses the predictions to form a portfolio.

4.1.1 Preparing the training-data

1. *Extracting the training data:* In the first quarter, only a single quarter is available, in the second quarter, two quarters become available for training purposes. As the model goes further into the future more quarters become available for training purposes. It combines a maximum of 12 quarters to form the training data. This means that only after three years the model is training with 12 quarters worth of training data.

2. *Adding growth rates:* Based on the ADVFN-data, it calculates growth rates for certain inputs. It calculates the quarterly growth, 4-quarter average growth, the annual growth, and 4-year average growth. It does so for the following variables: Revenue, Net Income, Total Assets and Free-cash-flow.

3. *Removing small companies:* Small companies tend to have extremely large outliers, and tend to be more noisy. The model removes all companies which have total assets of less than 250 million US Dollars. Re-

search shows that in a data set with survivorship bias small-companies will do better than they would without the survivorship bias[17]. Our linear model seemed to favor small companies, which was unrealistic, because the size effect is only there because of the survivorship bias. To prevent both the extreme outliers and the model from taking advantage of a size effect, small companies were simply filtered out.

4. *Adding a scoring variable:* Because the linear model uses a linear combination, certain combinations of the variables are difficult to establish for the model. To help the model, a scoring variable was introduced which scores certain inter-variable relationships. It tries to create a variable which indicates whether companies are both undervalued and high-growth. Though other variables indicate whether a company is either undervalued, or high-growth, they do not indicate whether both are the case. The scoring variable does indicate whether this is the case.

5. *Standardization:* It standardizes each feature by extracting the mean and then dividing by the standardization of each feature. Values should then be in the same scale.

6. *Moving outliers:* The data includes a large number of outliers, which has the troubling effect that companies can make it in the portfolio, simply because one variable happens to be a very large outliers. To prevent this, all features which are lower than -2.5 standard deviations from the mean, are set to -2.5 standard deviations from the mean. Likewise, all features which are higher than 2.5 standard deviations from the mean are sets to 2.5 standard deviations. Though not the best approach for dealing with the outliers, it worked for our purposes.

7. *Adding the bias term:* A column of ones is added to the design matrix, after which the features of the training-set are finished.

8. *Normalizing the targets:* The mean of the targets was subtracted from all targets. So all targets which were larger than the mean, are now positive, and all targets which were smaller than the mean are now negative. The new mean is now centered around zero, and the model only has to find the targets which are positive for its portfolio. Because all targets that are positive have a higher value than the market average. This also means negative values no longer mean the investment lost money, it simply means the stock performed below the market average.

4.1.2 Performing the regression

After the training data is ready, the features and targets can be given to the linear model. The model then calculates the coefficients, either analytically or through Bayesian inference. The exact way the coefficients are calculated are discussed in the previous section 3.3. In any case, all of the discussed methods give a vector of coefficients, which can then be used to make predictions for the test-data.

It should be noted that the regression can be done both with or without regularization. Because the fit on the training data is not very good to begin with, regularization was not implemented for the analytical solution. If regularization is used, the regularization parameter should not be set too high. The linear model will have trouble finding a good fit as is, penalizing it too much for larger coefficients might mean the model is unable to fit even the training data. In general, the numerical linear regression did not use any regularization. Though regularization might well have a beneficial effect and if done right it should at not have a negative effect. However given the large amount of ways to improve the performance of the model, regularization simply was not a priority.

4.1.3 Preparing the test-data & making the predictions

The test data is prepared in the same way as the training-data. The only difference is that the test-data is always 1 quarter. It never combines multiple quarters as is done in the training-data. We want the model to use *only the most recent* data for making the predictions. After the test-data is ready, the predictions are made by using the coefficients

$$\text{predictions} = X^{\text{test}}\beta.$$

In which β is the vector of coefficients you have learned during the regression.

4.1.4 Forming the portfolio

The portfolio is simply the 30 stocks for which the model makes the highest predictions. So the exact prediction the model makes for any stock in the portfolio is irrelevant. The only thing that matters is that the predictions were apparently higher than companies which did not make it into the portfolio. Because only the ordering of the predictions matter to the portfolio, factors such as the root mean squared error are not really relevant for the portfolio. Granted that the more accurate the predictions are, the better the portfolio will perform. But the task of making sure the best investments have a higher predicted value (whatever the value is) than the other companies might be easier than making very accurate predictions. It should be noted that the model is unable to create a portfolio of the actual top-30 investments. Even the ordering of the predictions seems to be considerably difficult. Some investments it makes are actually unprofitable investments, who should be near the bottom of the ordering instead of the top.

4.2 Results

The results show both how the model performs at different points in time and how different periods effect the performance. The performance over time has a fairly high variance, in most quarters it outperforms the market but in some it underperforms the market. The period used to make the predictions also has a substantial effect on the performance, predicting two years into the future seems to work best with our model and the data that it uses.

4.2.1 Performance at different quarters

The performance of the model changes over time, in some quarters it does better than in other quarters. It seems that market sentiments can have a considerable influence on performance. The performance of predicting two years into the future can be seen in Figure 2. What stands out is the fact

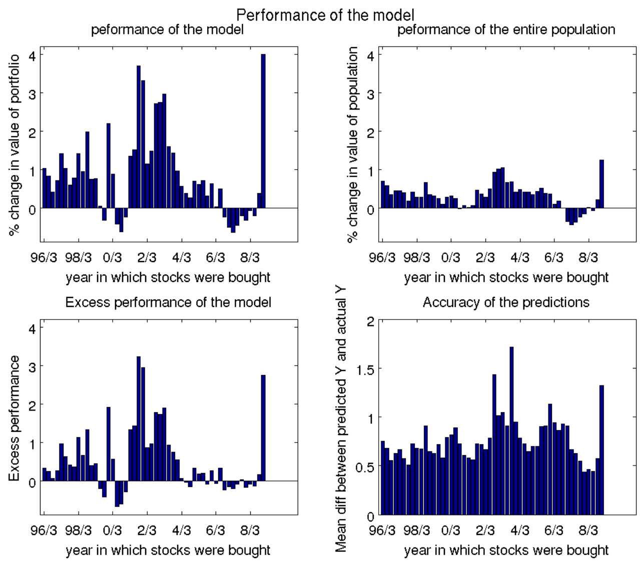


Figure 2: Performance of the model in different quarters. The model uses Bayesian linear regression, it predicts 2 years into the future and holds the stocks for 2 years after buying them. Performance is simply how much the value increased / decreased over 2 years. The accuracy is the average absolute difference between the actual targets and the predictions.

that the model can outperform considerably during certain quarters, but under performance can also be considerable. The model can outperform the market 3 to 1, it can also under perform the market 1 to 3. Further more there are more quarters in which the model outperforms, then there are quarters in which it under performs. Further indicating that the model is able to exploit the data effects in at least some quarters, though not every quarter. The model can of course be ran on different time periods as well, we can also let the model make predictions a single year, or three years into the future. The performance of different periods is reported in Figure 1. The

Table 1: Performance of the model in different quarters. The model uses Bayesian linear regression, it predicts p years into the future and holds the stocks for p years after buying them. Performance is simply how much the value increased / decreased over p years. The ratio indicates by how much the model outperforms the market, on average. Where as quarters tested is how many quarters the model bought stocks to get to this result.

Period p	mean model perf.	mean market perf.	ratio	quarters tested
1	0.4028	0.1882	2.14 : 1	60
2	0.8783	0.3216	2.73 : 1	52
3	0.7545	0.4593	1.64 : 1	44
4	1.2517	0.6930	1.83 : 1	36
5	1.2324	1.0507	1.17 : 1	28

table shows that the model can outperform the market over several periods, where $p = 2$ seems to be the best performing period to chose. Interestingly the models performance seems to shrink if p becomes larger. This might be because the inputs no longer have any explanatory value after 5 or more years have passed.

4.3 Goodness of the data fit

The fit the model achieved on both the training and test data was usually different over the years. In most cases the model could find a relatively good fit on the training data, but the test data fit varied wildly. There could be considerable over-fit in some quarters, there could also be no over-fit at all in which the fit on the training data was roughly the same as the fit on the test data. There were also some odd cases in which the slope of the regression line of the predictions and targets changed, it would be positive in the training data and negative in the test data. When the line is positive, it means on average the higher the predicted value the higher the target. When the slope is negative the opposite is true, on average the higher the predictions the lower the actual targets. For the model to perform its task preferably there should be a positive slope, as that would mean the

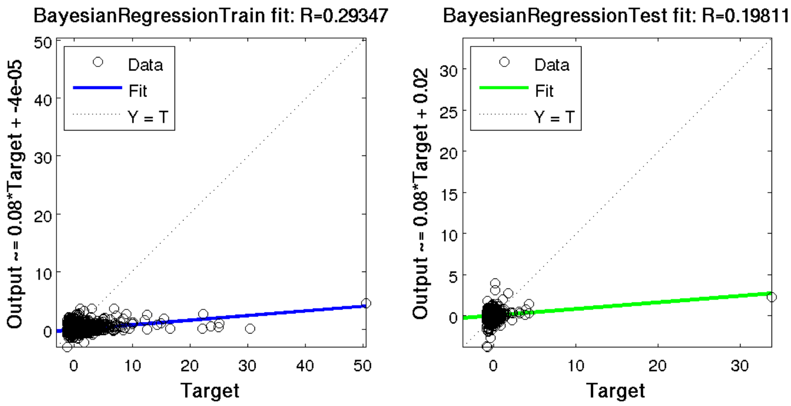


Figure 3: An example fit of the model, which is close to the fit the model usually achieved. For more examples of data-fit the model achieved view in the Appendix B. The training fit is in the left plot and the test fit is in the right plot, which indicates there is some over-fitting. The R-value indicates how much the model could explain, an R-value of 0.20 would mean the model could explain 20% of the stock price movements. The R-values achieved were usually reasonable, because the model has access to only very limited data on each company. Therefore it makes sense that the R-values are not very large, it also makes sense that there is a lot the model can not explain.

probability that the companies that make it into the portfolio are in fact good investments. Because a positive slope means the higher the predicted value, the higher the actual value, therefore the top-30 predictions that make it into the portfolio have a better chance of beating the market if the slope is at least positive in the test data. Good investments and should be as far to the right as possible. The targets are on the x-axis, a value of zero there means the target performed the market average. A negative value means the target underperformed the market and a positive value means the target outperformed the market. So all targets to the right of zero, are investments the model could make to successfully fulfill its task. All targets to the left of zero should be avoided. The further to the right of zero, the better, the largest x-axis value is the best investment and the lowest x-axis value is the worst investments. What we want is the model to find the best investments and we know the model invests in the top-30 predicted values. The top-30 predicted values are the top-30 largest y-values. So they are always the highest plotted points. We want those to be good investments, so we want them to be as far to the right as possible. What this means is that the top-30 outputs preferably should be in the top-right corner of the plot, or at least to the right of the zero on the x-axis. That would mean that the top-30 outputs do better than the market average. However as the plots show the top-30 outputs are not all too the right of market average. Plots of different fits can be found in the Appendix B.

5 Multi-model experiment

The multi-model experiment builds on top of the results achieved in the single-model experiment. In the single-model experiment a portfolio was constructed based on the suggestions of a single model. However there might be several reasons why we would want to construct a portfolio based on what multiple models think are good investments. The single-model experiment created a model for a specific period, which could be anywhere from 1 year, 2 years, 3 years etc. from now. So the portfolio was constructed to do well for a specific period, the portfolio for a 3 year period might not do so well on a single year into the future. But what we really want is a portfolio that does well on multiple time periods, instead of a previously defined period. So to get multiple models to work together the choice was made to create a scheme that allows them to construct one portfolio, together.

The basic idea is that we again use the ordering of predictions, instead of the exact values. A financial expert might well know which stock is going to perform better than another, without knowing exactly how well either is going to perform. The ordering each model produced was seen as a list that could be used to score certain stocks. So the combination of multiple models, is simply looking at which stocks seems to be favored by most models. The exact method of combining the orderings each model made will be discussed in the methods subsection 5.2 below.

It should be noted that the models now work together in what could be called a committee. Who are subordinate to the utility function that produces the eventual scores for each company. The utility function outputs a single score for each company, these scores are then used to form the portfolio. Again the top-30 scores were used for the portfolio, though other portfolio sizes might be just as reasonable.

Because the committee now predicts over multiple periods, the portfolio might well need adjustments over time. So we allowed the committee to change its portfolio on a quarterly or semi-quarterly basis. This means the committee now functions much more like a fund-manager. It takes in financial data each quarter, makes predictions for all companies then the ordering is used to score all companies. The utility function combines all scores into a single score for each company. Then we can decide whether the portfolio needs to change, is the top-30 we have now different from the one we had last quarter? If so, then we make the appropriate changes to the portfolio to make sure that the top-30 remains up-to-date on a quarterly or semi-quarterly basis.

Allowing the models to change the portfolio adds a different complexity though. What if the top-30 scoring companies are different each quarter? This would lead to a high-turnover which is not going to help. The model predicts companies to do well on a single or multiple year period. So if it buys stocks in a certain quarter, then sells them all the next quarter to buy others, no investments will be held long enough to become profitable. To prevent the model from buying and selling too much, too often, we added the historic scores to the utility function. So it will favor companies which it thought were good investments in the past and avoid ones which used to be bad investments. This considerably lowered the turn-over. The exact amount of history used can be changed by a parameter, which allows for additional tweaking.

5.1 Task

The models now have to function in a way that is more similar to an actual fund-manager. They receive new financial data each quarter, which they have to analyze. They make predictions on that new financial data based on what they have learned from past financial data. The ordering of the predictions are then used to determine the best possible investments. For the portfolio, the 30 stocks with the highest scores are selected. Here is how the process works:

1. The models are trained on past financial data
$$X^{\text{train}} \to \text{Linear Models} \to \beta.$$
2. Then the β coefficients are used to make predictions, which lead to scores
$$X^{\text{test}} \xrightarrow{\text{Apply } \beta} \text{predictions} \xrightarrow{\text{utility function}} \text{current scores.}$$
3. The model also gave scores to all stocks in the previous quarter though, so both the previous and current score are used to come up with the final score
$$\text{current score \& historic score} \xrightarrow{\text{utility function}} \text{final score.}$$

4. Which then leads to the decision on whether adjustments are needed. If a company was in the top-30 the previous quarter, but in the new final scores it no longer is in the top-30, then it needs to be sold. If a company was not in the top-30 in the previous quarter, but is in the top-30 in the new final scores, it needs to be bought.

This process is repeated every single quarter. So now, the committee of models not only create portfolios, they actively manage them. Whenever they make a investment, which with new data seems like a bad investment, the investment is sold. Whenever a stock which seemed like a reasonable investment, but with new data seems like a really good investment, it buys it when it becomes a better investment than the others. In other words, when it becomes a top-30 investment.

Performance is tracked by comparing how well the portfolio does over time compared to the population average. The performance is cumulative, so if either does really well over some period, they will have an advantage in the next period. This cumulative performance can be plotted which would give the cumulative performance of the model and the cumulative performance of the market average. The plotted line of the model should be higher than the performance of the market average. Or at least be higher over some periods, preferable the final periods.

5.2 Method

The methods section of the single-model experiment still holds here, the methods described there are still used for the models in this experiment. However, they now form a committee which is subordinate to the utility function. The utility function will be described below. How the predictions are made can be viewed in section 4.1.3.

5.2.1 From predictions to scores

New financial data arrives each quarter, which is then used to make predictions. Mind that every model m creates their own predictions. Again we are more interested in the ordering than the actual values of the predictions. To reflect this, we created a utility function to score score all companies, based on the predictions each model makes. Here the rank of the predictions is of importance, the rank of company i out of a population of size n is calculated by using tied ranks. Which means each company is given a rank based on their position on an ordered list. If predictions are the same, the ranks are averaged and every company which had the same predictions receives an average rank. So if no company has a lower predictions than company i then rank_i becomes 1. If all companies have a lower prediction than i then rank_i becomes n. Now we can move on to combining the scores of multiple models. Summarized the process has the following form:

$$X^{\text{test}} \xrightarrow{\text{Apply } \beta} \text{predictions} \xrightarrow{\text{utility function}} \text{current scores.}$$

Now if we let i be a company in the population of size n companies and let m be one of the models who made a prediction for all companies then we can define the score of company i by model m as

$$\text{bonus}_i^m = \begin{cases} \psi & \mid rank_i^m >= n - 30 \\ 0 & \mid rank_i^m < n - 30 \end{cases},$$

$$\text{score}_i^m = \begin{cases} \frac{rank_i^m}{n} * 100 + \text{bonus}_i & \mid rank_i^m >= \frac{n}{2} \\ (\frac{rank_i^m}{n} * 100) - 100 + \text{bonus}_i & \mid rank_i^m < \frac{n}{2} \end{cases}.$$

To give some intuition for the scoring, you can think of them as a percentile-score. If you receive a score of 90, it means the model thinks you will do better than 90% of all stocks in the population. If you receive a score of 60, it thinks you will do better than 60% of the population. If the model thinks you do worse than 50% of all companies, you will no longer receive a positive score but a negative score. Which reflects penalizing potentially bad investments. If the model thinks a company will do better than 40% of all companies, the score will be $40 - 100 = -60$.

The bonus is variable in this case, ψ can be any value. It basically gives an additional bonus to stocks that made it in to the portfolio of the individual model m. So if model m had constructed a portfolio, the companies who receive the bonus ψ would have been in it. This reflects that fact that the single-model experiment showed that individual models can form above average performing portfolios. We want to make sure that those stocks have a higher probability of ending up in the committee portfolio.

5.2.2 Using historic scores & final scores

The predictions from every model m have let to m scores for each company i. Now we move on to combining the scores of all models m. This is done by simply summing over the scores of every model, and then adding a fraction or multiple of the previous score. We will introduce t for this, which is the current quarter we are in, $t - 1$ is the previous quarter. Let there be n models, this leads to the following final score

$$\text{final score}_i = (\sum_{m=1}^{n} score_i^m)^t + \alpha * \text{final score}_i^{t-1}.$$

This leads to a final score for all companies, where the company liked by most of the models has the highest score. Whereas the company that receives the lowest utility, is a bad investment according to the models. It should be noted that the score gives no indication of how well the model thinks a company will do, it only reflects how well the model thinks the company will do compared to all the other companies in the population. This is reasonable because we want to outperform the population average, so as long as you do better than most companies, no matter what the exact performance is, the models are doing fine. Summarized the process looks like this:

$$\text{current score \& historic score} \xrightarrow{\text{utility function}} \text{final score.}$$

5.2.3 Transaction costs

In practice, buying and selling stocks has a transaction cost. We take this into account during simulation by setting a fixed 8% transaction cost per trade. A trade consists of selling stock A and buying stock B. This means that a trade in our simulation is basically two transaction, a sale and a buy. The 8% transaction cost is fairly high, but transaction costs have a rather small effect on performance. So setting them a little lower or a little higher does not have a considerable effect on performance.

5.3 Results

We tested our model committee from 1994 to 2012. The first stock is bought in 1996, the last time performance is measured is in 2010. This means it manages a portfolio over 14 years, a time period in which both the dot-com and financial crisis occur. This means the model will have to manage a portfolio over very different market sentiments. We both want the model committee to be able to outperform and underperform the market average, because if you have learned to win at a game you should be able to lose on purpose. Though we have focused mainly on outperforming, so the methods and tuning are mostly geared towards outperformance. Which might mean that certain steps taken in losing on purpose actually causes performance to be higher instead of lower. The performance is plotted in Figure 4. The model is able to find companies which outperform the market, and more interestingly it is able to create a portfolio that performs better than the market average. This means the companies it invests in, on average and over time, perform better than the companies it chooses not to invest in. The sub selection it chooses is indeed more profitable than investing in the entire population. This implicates that the models have found some linear relation between the inputs and the targets which can be exploited to achieve above average market returns.

Another interesting observation is that the model is unable to do well in market sentiments for which it has not trained. Especially the financial crisis in 2008-2009 has a considerable negative impact on the performance of the portfolio. It was unable to find good investments during the crisis. Further

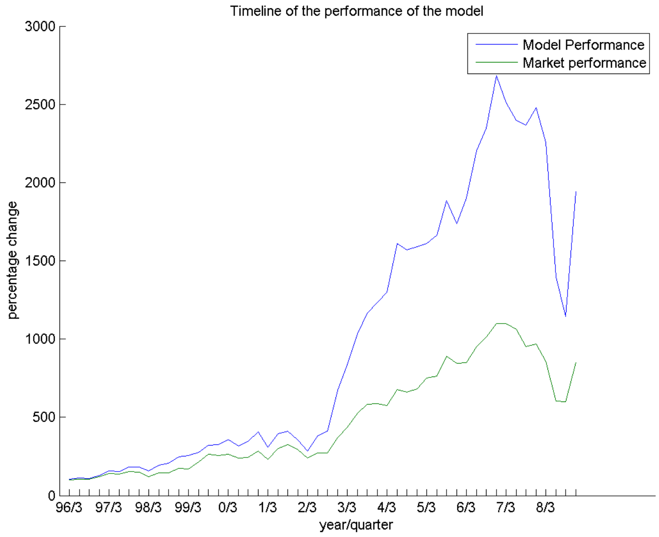


Figure 4: Performance of the model committee as compared to the market average. The committee consisted of 3 Bayesian models predicting 8, 10 and 12 quarters into the futureand 3 linear regression models predicting 8, 10 and 12 quarters into the future. The history factor was set to $\alpha = 0.5$ and the bonus was set to $\varphi = 1$. The portfolio was re-evaluated every 2 quarters.

more it performed worse than the market average during the crisis. Meaning the portfolio lost even more value than the population. As promised we will also discuss how a random model performs. The result can be seen in Figure 5. Please note that this follows every step that the individual models take in the single model experiment, and every step in the multi-model experiment. The only difference is that the coefficients β are randomized with a mean of 1 and a standard deviation of 2. This was done because any step in the process might help the model, and we specifically want to know how much training the model effects the performance. Perhaps removing the small companies and moving the outliers on their own make up for a performance boost. Or any other combination of the steps might lead to a good performing end-result. To account for this, we did an entire run with all the usual steps, except the training of the model. This should reflect how the model committee might be expected to perform if none of the models learn anything useful. If they do not learn anything useful, the coefficients will be nonsensical, or fairly random. As the results show, this leads to a considerably lower performance. It underperforms the market average by a wide margin. It underperforms the market committee which does learn their coefficients by an even wider margin. This indicates that the models do seem to learn something useful, though we can not be exactly sure that a non-

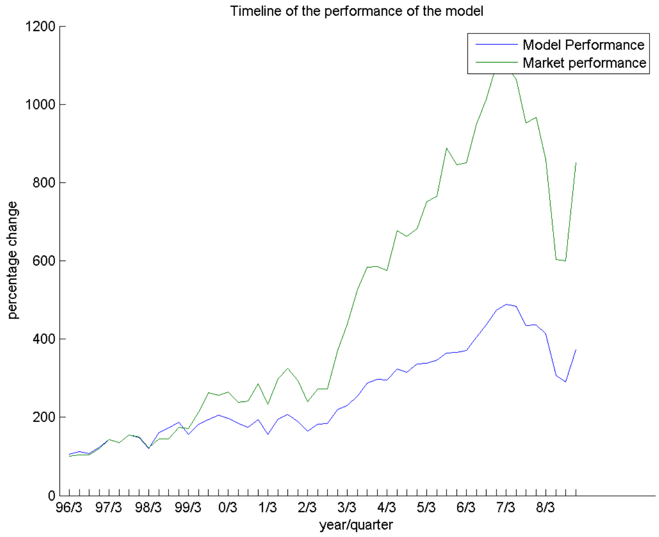


Figure 5: Performance of the random models as compared to the market average. The committee consisted of 3 Bayesian models predicting 8, 10 and 12 quarters into the future and 3 linear regression models predicting 8, 10 and 12 quarters into the future. They did not learn their coefficients β, instead the coefficients were randomized every quarter.

linear relationship might work even better, it seems that a linear relation can explain some of the price changes. The results show that the linear model can exploit some of those relationships to make some above average investments.

Losing on purpose is possible, though the difference with the market is not as substantial as the outperforming model. The performance can be seen in Figure 6, it market performs for roughly six and a half years. After six and a half years it clearly underperforms the market average. It finishes lower than both the outperforming model committee and the market average, but it also finishes higher than the random model committee. The difference between outperforming and underperforming is considerable, the outperforming model committee finishes roughly three times higher than the underperforming model committee.

An odd observation is the fact that the randomized weight model committee underperforms the average, while one might expect it to perform market-average. However, the targets are distributed with outliers on the positive values. They range from $[-1, +\infty]$, even though no value is infinite, they can be extremely large. This means the distribution leads the average to be higher than what one would achieve by randomly picking stocks. The median would be much closer to the random model performance.

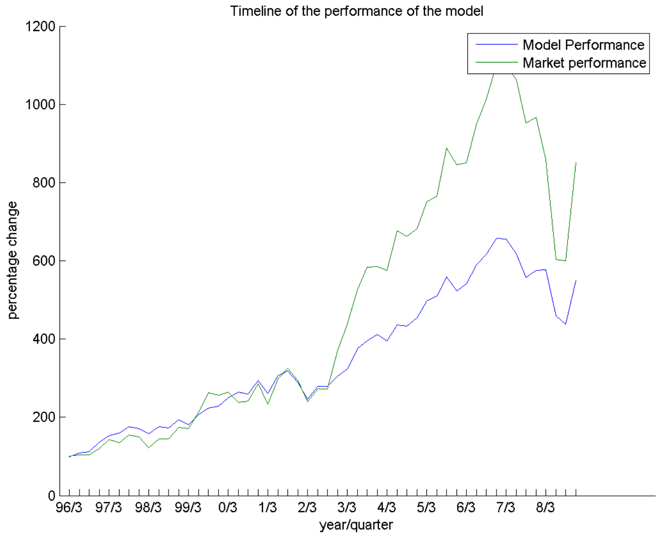


Figure 6: Performance of the model committee as compared to the market average when the committee uses the lowest-30 scores. This in a sense is the same as losing on purpose. The committee consisted of 3 Bayesian models predicting 8, 10 and 12 quarters into the future and 3 linear regression models predicting 8, 10 and 12 quarters into the future. The history factor was set to $\alpha = 0.5$ and the bonus was set to $\varphi = 1$. The portfolio was re-evaluated every 2 quarters. The only difference between the outperforming model committee is that instead of the top-30 the lowest-30 scores are used.

5.3.1 Prototype

Performance is not the only measure to determine whether the models have learned something relevant or useful. We can also look at what the average company the committee invests in looks like, and if that is similar to the average company that outperforms the market. For this we have created prototypes, which is a non-existent average member of a certain population. One of the populations is the companies that are in any of the portfolios over 1996 to 2010. If we take the average value for all values, for all companies in every portfolio, we get a prototype for a company the committee invests in. We can do the same for the actual top-30 performers, or the population the committee should have invested in to achieve the best results. We can also do it for all companies that outperform the market, or the population the committee could have picked from to achieve its goal. Those three prototypes will be discussed below. Different parts of the prototype will be discussed, and whether they make sense. If the prototype is of the committee is similar to the prototype of outperform companies, then we can derive that the model chose the right companies.

- *Size:* The prototype company our committee invested in had total assets of roughly 6885 million\$. The actual top-30 had on average had total assets of roughly 599 million\$. The average out-performing company had total assets of roughly 4330 million \$. This indicates our committee should have picked smaller companies to perform better. It should however be noted that we filtered out companies which had less than 250 million \$ in assets. So part of the bias towards larger companies is because of the removal of small companies. The median of the top-30 companies actually is 67 million \$. So even though it might perform better if it picks small companies, it should be noted that they are considerably more risky. Our population with survivorship bias has less underperforming small companies than there should be[17], so this extra risk is not reflected properly in our population.

- *Valuation:* The prototype of the committee had a price to earnings of 15.88. Where as the actual top-30 had a price to earnings of 18.14 and the outperforming prototype had a price to earnings of 22.50. The committee prototype had a price to revenue of 2.70, the top-30 had 15.25 and the outperforming had 9.50. Though the median values for the prototypes where a lot closer: 0.61 for the committee, 1.07 for the actual top-30 and 1.30 for the outperforming. Lastly, the price to tangible book value paints a similar picture as the price to earnings: 2.72 for the committee, 2.90 for the top-30 and 3.90 for the outperforming. Based on these results it seems the committee favors companies with a slightly lower valuation. It seems that for the price to revenue, the model does not pick any of the outliers, or companies with a really high valuation, even though the top-30 do include some of those companies. It seems that the committee could pick some more companies with a high valuation. But it is pretty close to the correct valuation. In terms of valuation, the companies it invests in seem to make sense.

- *Profit margins:* The gross profit margin of the committee prototype is 39.88, the actual top-30 had an average gross margin of 34.94 lastly the outperforming companies had an average gross margin of 37.76. The net-profit margin gives a more interest indication: the committee prototype had a net-profit margin of -1.79, the top-30 had -3.59 and the outperforming had 3.40. The profit margins seem to be close, the interesting thing is that the committee somehow learned that it is okay if a company does not make a profit. A negative net profit margin means the company had a loss in the most recent quarter. Somehow the committee learned that this is fine. Though the median is 0.1, so at least 50% of the companies it invests in do turn a profit. All in all the profit margins seem to make sense, the companies do generate more than enough sales to turn a profit after all costs associated with production are subtracted, because the gross margin is roughly 40%. Yet after all costs are subtracted a good part of the companies that make it into the portfolio are not profitable. Which seems to be close to the actual top-30.

- *Leverage position:* The debt to equity of the committee prototype is 3.48, of the top-30 the debt to equity is on average 2.31 and of the outperforming it is 3.12. The Leverage ratio of the committee is 3.96, of the top-30 it is 4.07 and of the outperforming it is 4.81. Further more the quick ratio for the committee is 1.31, for the top-30 it is 3.70 and for the outperforming it is 2.69. This indicates that the committee seems a little biased towards companies which take on quite some debt. A little more so than it should be. Though the leverage ratio is extremely close, the other ones indicate that the committee takes on more companies with high-debt than it should. Though the exact values are not high, the quick ratio is above one, meaning the prototype of the committee could still pay back its short-term debt without any problems. All in all, it seems the committee favors companies that finance with debt, a little more so than it should.

- *Growth:* There are many values associated with growth, for the prototype values please refer to the tables in the appendix: Table 9 and Table 10. The values indicate that our committee seems to favor companies with high revenue-growth. Even though the earnings might not grow or might even decline. More importantly it seems to favor slow and steady revenue growth, and not high quarter on quarter growth. The four-quarter average revenue growth was 5% the four year average annual revenue growth was 6%. Though the medians were 2% and 0% respectively. This conclusion of the committee is correct, but it should favor them even more than it does now. The mean and median values are higher for both the actual top-30 and the outperforming. For net-income it seems to favor companies which had a recent decline in net-income, while the other prototypes indicate it should favor companies with no growth or decline. Though the median values are closer to 0, indicating that it might mean that the company should pick less outliers when it comes to net income decline. With total-assets growth, the committee favors companies with a small and steady growth in assets, which is very close to the actual-top 30. The outperforming prototype has higher total assets growth than the committee and the actual top-30. So it seems the committee successfully learned to exploit total assets growth. The growth in free cash flow medians are again very close to the actual top-30, but the means are too negative. So the committee again seems to pick some companies with a very large decline in free cash flow, which the other prototypes indicate is a mistake. All in all, it seems to get growth mostly correct. It is only some companies with a very large decline in earnings or free cash flow, that the committee should learn to avoid. This should bring the prototype of the committee closer to where it should be.

5.3.2 Stock picks & turnover

The average hold time for stocks in the portfolio is 5.4 quarters. Whereas the minimal hold time for each stock is only 2 quarters, so the committee does not buy an entirely new portfolio every time it re-evaluates its positions. The maximum hold time is 24 quarters, though the quarters do not have to be sequential. Over the period it invests in a total of 291 different companies and 100 of them are only held for the minimal holding time of 2 quarters. A partial list of companies it invested in will can be seen in Table 2, it will be ordered on how many quarters it invested in the stock.

6 Conclusion

The results indicate that we can in fact use linear models to achieve above average returns. Though we are not the first to achieve these results[4], the uniqueness lies in the fact that our models could manage a portfolio of its own over more than a decade. It did its task well enough to beat the market average which is what we set out to achieve in the beginning. Letting models decide whether to invest in some risk-free asset or a stock index has been extensively researched [18, 7], our models however had no access to risk-free assets. Nor did it make a one or the other decision. It had to form a portfolio of 30 stocks, which meant it had a lot more options than just two. Our results indicate that linear models can indeed perform tasks more complex than just choosing between a stock index or a risk-free asset. Our models could create a stock portfolio of its own, out of a population of a few thousand stocks and outperform the market average with that portfolio. Our models can be seen as the next step, once you have decided the stock market is a better investment than a risk-free asset, a portfolio can be constructed instead of simply buying a stock index.

The performance itself is far higher than it would have been in a population without the survivorship bias which is why the population average should be considerably lower. As the population average goes down, it is highly likely that the performance of the models will go down by roughly the same amount. The average return of companies that are not in the population might be as low as roughly -50%[17]. Taking this into account we can reasonably assume the actual performance of the models in our results are meaningless. It is only the performance compared to the population average

Table 2: A portion of all the companies the committee invested in, sorted by how many quarters the company was in the portfolio. The sectors and industries are included to give an indication of the diversification the committee achieves.

Company (StockTicker)	quarters in portf.	Sector	Industry
Mentor Graphics Corp (NASDAQ:MENT)	24	Technology	Semiconductors
Cirrus Logic, Inc (NASDAQ:CRUS)	20	Technology	Integrated Circuits
DISH Network Corp (NASDAQ:DISH)	20	Cyclical Consumer Goods & Services	Cable Service Providers
Carmike Cinemas, Inc (NASDAQ:CKEC)	20	Cyclical Consumer Goods & Services	Movie Theaters & Movie Products
Bally Technologies (NYSE:BYI)	18	Cyclical Consumer Goods & Services	Games, Toys & Children Vehicles
W.R. Grace & Co (NYSE:GRA)	18	Basic Materials	Specialty Chemicals
Denny's Corporation (NASDAQ:DENN)	18	Cyclical Consumer Goods & Services	Restaurants & Bar
Gilead Sciences, Inc (NASDAQ:GILD)	16	Healthcare	Biopharmaceuticals
Earthlink Holdings Corp (NASDAQ:ELNK)	16	Technology	IT Services & Consulting
AK Steel Holding Corporation (NYSE:AKS)	16	Basic Materials	Iron, Steel Mills & Foundries
Cablevision Systems Corporation (NYSE:CVC)	16	Cyclical Consumer Goods & Services	Cable Service Providers
General Electric Company (NYSE:GE)	14	Industrials	Industrial Conglomorates
Omnicom Group Inc (NYSE:OMC)	14	Cyclical Consumer Goods & Services	Advertising & Marketig
Sinclair Broadcast Group Inc (NASDAQ:SBGI)	14	Cyclical Consumer Goods & Services	Broadcasting
GrafTech International Ltd (NYSE:GTI)	14	Industrials	Electrical Components & Equipment
Safeguard Scientifics, Inc (NYSE:SFE)	14	Industrials	Business Support Services
Rent-A-Center Inc (NASDAQ:RCII)	14	Cyclical Consumer Goods & Services	Other Specialty Retailers

that is of interest. The hope is that the models will continue to outperform the market average by a same amount if the survivorship bias disappears.

Table 2 shows some of the companies the committee invested in, some of which were really good long-term investments and some of which were really bad long-term investments. But what it also shows that the committee achieves a good diversification, the sectors and industries it invests in are diverse. Though the committee does not focus on achieving diversification, it seems that its indifference to sectors and industries allows it to achieve diversification automatically, without any other work needed. It simply does not care at all about what sector or industry a company is in, it has no information on it. So the committee does not purposefully achieve diversification, it is simply a side-effect of not knowing anything at all about sectors and industries and the fact that the population includes a vast number of sectors and industries. It is reasonable to assume that the random model committee achieved diversification as well, simply by picking stocks at random from the population some sort of diversification would be achieved.

We can also conclude that our results are not in line with the efficient market hypothesis[8, 9], according to the efficient market hypothesis our linear models should not be able to find any useful relations between the inputs and the targets. Financial variables should at all time be incorporated into the market. If new financial data becomes available it should either be already incorporated into the market or be incorporated almost instantly. Our models bought stocks a quarter after the financial data was made available to the public. To account for the delay there might be in publishing the financial data. So our models did not have early-access to financial data. All financial variables used by the models could have been incorporated into the markets. The models we used however could find some relation between the financial variables we used and the future stock price changes. It could also successfully exploit those relations to make above average investment decisions. This indicates there are some inefficiencies in the market which the linear models can exploit. Our results are more in line with other research indicating that in fact certain financial variables do have a predictive value[18, 4, 12].

The financial variables that we used as input have some meaningful predictive value on the future price changes of stocks but the model makes some of the same mistakes human investors do. Most importantly it falls prey to the recency bias[15]. Investors tend to look at the past few years and assume that the same market pattern will occur over the next few years. Which is exactly what our models do, by definition. They only use 12 quarters worth of training data, which means they will always assume the same market sentiment will occur in the near future. This short-sightedness is inherent to the techniques we used. A solution for this might be dynamic models, which change their coefficients over time and remember what they have learned in the past. Though the models do not necessarily need it, because they rebalance their portfolio quite often. So the performance increase of dynamic models might not be very high because the committee can be highly flexible to changing market environments. The committee lost roughly 60% during the financial crisis, but it adjusted and made up for most of the losses in the year after the crisis. Some of this flexibility might be lost when using dynamic models, which means dynamic models will probably have to be tested to find out whether they can solve the recency bias and yet remain flexible to changing market environments.

However our model does not fall prey to other biases commonly associated with human investors. The model does not over trade. Each trade has transaction costs associated with it, over-trading is a common problem for private investors[2]. But our models remember what they preferred in the past and it only makes changes to its portfolio on a quarterly basis. So the transaction costs the model incurs have only a negligible effect on its performance. At the end of a quarter, it usually favors companies it already invested in. Because of the way the utility function works our models are biased towards companies that were a good investment in the past. In the beginning turnover can be as high as 80%, however after a few years turnover becomes as low as 30%. Because the evaluation is done on a quarterly basis and even when it evaluates it still keeps most stocks, transaction costs that are incurred are quite low. The average hold period for an investment is 5.4 quarters, which means that on average it holds stocks for more than a year. If transaction costs are turned off during the simulation, the effect on performance is only small. Further indicating that transaction costs are not a problem for our model committee.

This only partially explains why our models outperform private investors. The recency bias described above has some negative effect on the models in that it is unable to explain certain changes in market-sentiment. However our models remain indifferent to market-sentiments altogether. When the markets were in the depths of the crisis, many private investors chose to take out all their investments. However our models remained indifferent in the midst of the crisis. It continued to invest and in a few quarters it had regained most of the loses incurred during the crisis. The point here is that our model consistently continues what it does well. In the midst of the crisis it is not afraid to lose more, when the markets go to all time highs it does not grow overconfident either. No matter what the market-sentiment, it does the same thing. The fact that it ignores market-sentiments, which it can not explain, and focuses on what it can explain, increases performance.

Given all the results we achieved, which include the performance of the model and the prototypes of its investments, we can conclude that linear models can be used to find outperforming stocks. Though the predictions are not necessarily accurate, the ordering they produce can be useful to find out which stocks might outperform the market.

7 Acknowledgments

A major part of the results described here were achieved with the advice and suggestions of Pasi, who helped me a lot throughout the process. We usually met on a weekly basis to discuss new results and to determine what to focus on next, he could give me a lot of advice on both the technical aspects and the financial implications of the results. Pasi also proofread the initial draft and gave many valuable suggestions on it. For his help in my thesis I am very grateful. Marcel for his part helped me in setting up my thesis and allowed me to start the thesis in the first place. We both understood that the task at hand would be difficult, but after incorporating his suggestions for the research goal the thesis was underway. Throughout the process he helped me with the more bureaucratic side of the thesis. I am very grateful for all the help received from Marcel and Pasi during the semester.

I would also like to thank ADVFN, Quandl, Yahoo and Google for providing me with all the data I needed to perform my research. The data played an essential part in my thesis and I am very grateful they allowed me to use it for my research.

8 References

References

[1] George S Atsalakis and Kimon P Valavanis. Surveying stock market forecasting techniques–part ii: Soft computing methods. *Expert Systems with Applications*, 36(3):5932–5941, 2009.

[2] Brad M Barber and Terrance Odean. Trading is hazardous to your wealth: The common stock investment performance of individual investors. *The Journal of Finance*, 55(2):773–806, 2000.

[3] John Y Campbell and Yasushi Hamao. Predictable stock returns in the united states and japan: A study of long-term capital market integration. *The Journal of Finance*, 47(1):43–69, 1992.

[4] John Y Campbell and Samuel B Thompson. Predicting excess stock returns out of sample: Can anything beat the historical average? *Review of Financial Studies*, 21(4):1509–1531, 2008.

[5] Louis KC Chan, Yasushi Hamao, and Josef Lakonishok. Fundamentals and stock returns in japan. *The Journal of Finance*, 46(5):1739–1764, 1991.

[6] Jan Drugowitsch. Variational bayesian inference for linear and logistic regression. *arXiv preprint arXiv:1310.5438*, 2013.

[7] David Enke and Suraphan Thawornwong. The use of data mining and neural networks for forecasting stock market returns. *Expert Systems with applications*, 29(4):927–940, 2005.

[8] Eugene F Fama. Efficient capital markets: A review of theory and empirical work*. *The journal of Finance*, 25(2):383–417, 1970.

[9] Eugene F Fama. Efficient capital markets: Ii. *The journal of finance*, 46(5):1575–1617, 1991.

[10] Philippe Jorion. Risk management lessons from long-term capital management. *European financial management*, 6(3):277–300, 2000.

[11] Pasi Jylänki, Jarno Vanhatalo, and Aki Vehtari. Robust gaussian process regression with a student-t likelihood. *The Journal of Machine Learning Research*, 12:3227–3257, 2011.

[12] Dennis Olson and Charles Mossman. Neural network forecasts of canadian stock returns using accounting ratios. *International Journal of Forecasting*, 19(3):453–465, 2003.

[13] M Hashem Pesaran and Allan Timmermann. Predictability of stock returns: Robustness and economic significance. *The Journal of Finance*, 50(4):1201–1228, 1995.

[14] ADVFN PLC. Advfn database, 2014. `http://uk.advfn.com/`.

[15] Michael M Pompian. Using behavioral investor types to build better relationships with your clients. *Journal of Financial Planning*, 21(10):64–76, 2008.

[16] Quandl. Quandl database, 2014. `http://www.quandl.com/`.

[17] Tyler Shumway and Vincent A. Warther. The delisting bias in crsp's nasdaq data and its implications for the size effect, 1998.

[18] Suraphan Thawornwong and David Enke. The adaptive selection of financial and economic variables for use with artificial neural networks. *Neurocomputing*, 56:205–232, 2004.

[19] Michael E Tipping and Neil D Lawrence. Variational inference for student-¡ i¿ t¡/i¿ models: Robust bayesian interpolation and generalised component analysis. *Neurocomputing*, 69(1):123–141, 2005.

[20] Yahoo. Yahoo finance, 2014. `http://finance.yahoo.com/`.

[21] Paul D Yoo, Maria H Kim, and Tony Jan. Machine learning techniques and use of event information for stock market prediction: A survey and evaluation. In *Computational Intelligence for Modelling, Control and*

Automation, 2005 and International Conference on Intelligent Agents, Web Technologies and Internet Commerce, International Conference on, volume 2, pages 835–841. IEEE, 2005.

A Data cheat-sheet

A.1 Quandl data-set

The date is always year-December-31, so annual data is always applicable to the entire year. Data related to a given day is always applicable to the very last day of the year. As such the stock price for instance, is the stock price at the very last day the markets were open in any given year.

X_n	Variable	Description	Start year	Unit
x_1	Shares outstanding	The total amount of shares the company has outstanding on the market.	2003	-
x_2	3-Year Regression Beta	Estimated by regressing weekly returns on stock against NYSE composite.	2000	-
x_3	3-year Standard Deviation of Stock Price	The standard deviation in monthly stock prices, annualized.	2000	-
x_4	Book Debt to Capital Ratio	The debt to capital ratio compares the overall debt that a company has to total capital on their books.	2000	%
x_5	Book Value of Equity	The total value of the equity in millions USD	2006	millions USD
x_6	Book Value of Assets	The total value of the assets in millions USD	2000	millions USD
x_7	Capital Expenditures	The total amount of capital expenditures in millions USD	2000	millions USD
x_8	Cash	The total amount of cash in millions USD	2000	millions USD
x_9	Cash as Percentage of Firm Value	Cash as Percentage of Firm Value	2000	%
x_{10}	Cash as Percentage of Revenues	Cash as Percentage of Revenues	2000	%
x_{11}	Cash as Percentage of Total Assets	Cash as Percentage of Total Assets	2005	%

x_{12}	Change in Non-Cash Working Capital	Percentage Change in Non-Cash Working Capital	2001	%
x_{13}	Correlation with the Market	This is the correlation of stock returns with the market index, using the same time period as the beta estimation.	2000	-
x_{14}	Current PE Ratio	Stock price divided by earnings per share	2000	-
x_{15}	Depreciation	The total amount of depreciation in thousands of USD	2000	millions USD
x_{16}	Dividend Yield	Dividend per share divided by the stock price	2000	%
x_{17}	Dividends	Dividends paid in thousands of USD	2000	millions USD
x_{18}	Earnings Before Interest and Taxes	Earnings Before Interest and Taxes in millions USD	2000	millions USD
x_{19}	EBIT for Previous Period	Earnings before interest and taxes in the previous year, in millions USD	2000	millions USD
x_{20}	Earnings Before Interest Taxes Depreciation and Amortization	Earnings Before Interest Taxes Depreciation and Amortization in millions USD	2000	millions USD
x_{21}	Effective Tax Rate	Effective Tax Rate	2000	%
x_{22}	Effective Tax Rate on Income	Effective Tax Rate on Income	2010	%
x_{23}	Enterprise Value	Market Value of Equity + Market Value of Debt - Cash	2000	millions USD
x_{24}	EV to Invested Capital Ratio	Enterprise value divided by invested capital	2004	-
x_{25}	EV to Trailing Sales Ratio	Enterprise value divided by the trailing sales	2004	-
x_{26}	EV to EBIT Ratio	Enterprise value divided by the EBIT	2000	-
x_{27}	EV to EBITDA ratio	Enterprise value divided by the EBITDA	2000	-
x_{28}	EV To Sales Ratio	Enterprise value divided by sales	2000	-

x_{29}	Expected Growth in Earnings Per Share	Expected Growth in Earnings Per Share	2000	%
x_{30}	Expected Growth in Revenues	Expected Growth in Revenues	2003	%
x_{31}	Free Cash Flow to Firm	FCFF = EBIT(1-t) - (Capital Expenditures - Depreciation) - Change in Non-Cash Working Capital.	2000	millions USD
x_{32}	Firm Value	Firm Value in millions of USD	2000	millions USD
x_{33}	Ratio of Fixed Assets to Total Assets	Fixed assets divided by total assets	2000	%
x_{34}	Forward Earnings Per Share	Forward Earnings Per Share, estimated earnings per share next year.	2002	USD
x_{35}	Forward PE Ratio	Forward earnings per share divided by the stock price	2001	-
x_{36}	Growth in Earnings Per Share	The change in earnings per share, last year to this year.	2002	%
x_{37}	Previous Year Growth in Revenues	Previous Year Growth in Revenues	2000	%
x_{38}	Hi-Lo Risk	?	2000	?
x_{39}	Insider Holdings	Percentage of shares owned by employees of the company	2000	%
x_{40}	Institutional holdings	Percentage of shares held by institutions	2000	%
x_{41}	Ratio of Intangible Assets to Total Assets	Intangible assets divided by total assets	2002	%
x_{42}	Invested Capital	total debt + total equity - non-operating cash and investments, in millions of dollars.	2000	millions USD
x_{43}	Market Capitalization	Amount of shares times stock price	2000	millions USD
x_{44}	Market Debt to Equity Ratio	Total debt divided by shareholders equity	2000	-
x_{45}	Market Debt to Capital Ratio	Total debt divided by shareholders equity + total debt	2000	%

x_{46}	Net Income	Net income in millions USD	2001	millions USD
x_{47}	Net Margin	Net income divided by revenue	2000	%
x_{48}	Non-Cash Working Capital	Non-Cash Working Capital in millions USD	2001	millions USD
x_{49}	Non-Cash Working Capital as Percentage of Revenues	Non-Cash working capital divided by revenue	2000	%
x_{50}	Payout Ratio	Dividends per share / earnings per share	2000	%
x_{51}	Price to Book Value Ratio	Stock price divided by book value	2000	-
x_{52}	PE to Growth Ratio	Price to earnings divided by Annual earnings per share growth	2000	-
x_{53}	Pre-Tax Operating Margin	Pre-Tax Operating Margin	2000	%
x_{54}	Price to Sales Ratio	Stock price divided by sales (revenue)	2000	-
x_{55}	Reinvestment Amount	?	2001	?
x_{56}	Reinvestment Rate	The amount of interest that can be earned when money is taken out of one fixed-income investment and put into another. The reinvestment rate is the amount of interest the investor could earn if s/he purchased a new bond, if the same investor is holding a callable bond that is called due because interest rates have declined.	2000	-
x_{57}	Revenues	Total sales in millions of USD	2004	millions USD
x_{58}	Return on capital	Net income / (Debt + Equity)	2000	-
x_{59}	Return on Equity	Net income / Equity	2000	-
x_{60}	Sales General and Administration Expenses	Sales General and Administration Expenses in millions of USD	2000	millions USD

x_{61}	Stock price	stock price in USD	2001	USD
x_{62}	Total Debt	Total debt in millions of USD	2001	millions USD
x_{63}	Trading volume	trading volume, amount of shares traded	2002	-
x_{64}	Trailing 12-month Revenues	revenue over the past 12 months	2001	millions USD
x_{65}	Trailing Net Income	Net income in millions of USD	2000	millions USD
x_{66}	Trailing PE Ratio	Price divided by earnings per share	2000	-
x_{67}	Trailing Revenues	Revenue in millions of USD	2000	millions USD
x_{68}	Value Line Beta	?	2002	-
x_{69}	EV to Book Value Ratio	Enterprise Value divided by book value	2000	-

A.2 ADVFN data-set

The ADVFN data-set had more variables than listed here, the variables listed below are the once used to obtain the results.

X_n	Variable	Description	Unit
x_1	Close P/E	The price to earnings ratio at the end of the quarter.	Valuation ratio
x_2	High P/E	The highest value the price to earnings reached during the quarter.	Valuation ratio
x_3	Low P/E	The lowest value the price to earnings reached during the quarter.	Valuation ratio
x_4	Gross profit margin	The profit margin after all costs related to production are subtracted from revenue.	Profitability ratio
x_5	Pre-tax profit margin	The profit margin before taxes.	Profitability ratio
x_6	Post-tax profit margin	The profit margin after taxes.	Profitability ratio
x_6	Net profit margin	The profit margin after all costs are subtracted from revenue.	Profitability ratio

x_7	Interest coverage (cont. operations)	The amount of times interest is covered by money coming in from continued operations	Leverage ratio
x_8	interest as % of invested capital	Interest as a percentage of all invested capital	Leverage ratio
x_9	Effective tax rate	The tax rate the company actually payed	-
x_{10}	Quick ratio	(current assets − inventories) / current liabilities	Leverage ratio
x_{11}	Current ratio	Current assets / current liabilities	Leverage ratio
x_{12}	Payout ratio	How much of earnings are payed out to shareholders	-
x_{13}	Total debt to equity	Indicates how much of the assets were bought using debt.	Leverage ratio
x_{14}	Long-term debt/total capital	Long-term debt / total capital	Leverage ratio
x_{15}	Leverage ratio	Indicates how much operations are financed with debt and how much with the companies own money.	Leverage ratio
x_{16}	Asset turnover	Revenue / total assets	Profitability ratio
x_{17}	Cash as % of revenue	Cash as % of revenue	Liquidity ratio
x_{18}	Receivables as % of revenue	Receivables as % of revenue	Liquidity ratio
x_{19}	Sg&a as % of revenue	Selling, General & Administrative Expense as % of revenue	Sales ratio
x_{20}	R&d as % of revenue	Research & development as % of revenue	Sales ratio
x_{21}	Revenue per $ cash	Revenue per $ cash	Sales ratio
x_{22}	Revenue per $ plant (net)	Revenue per $ plant (net)	Sales ratio
x_{23}	Revenue per $ common equity	Revenue per $ common equity	Sales ratio
x_{24}	Revenue per $ invested capital	Revenue per $ invested capital	Sales ratio
x_{25}	Receivables turnover	Net credit sales / Average account receivables	Sales ratio

x_{26}	Inventory turnover	Sales / inventory	Sales ratio
x_{27}	Receivables per day sales	Receivables per day sales	Sales ratio
x_{28}	Sales per $ receivables	Sales per $ receivables	Sales ratio
x_{29}	Sales per $ inventory	Sales per $ inventory	Sales ratio
x_{30}	Revenue/assets	Revenue / total assets	Sales ratio
x_{31}	Number of days cost of goods in inventory	Indicates how long inventories are kept before being sold	Sales ratio
x_{32}	Intangibles as % of book-value	Indicates how much of the assets are intangible.	-
x_{33}	Inventory as % of revenue	Inventory as % of revenue.	-
x_{34}	lt-debt to equity ratio	Long-term-debt to equity ratio.	Leverage ratio
x_{35}	lt-debt as % of invested capital	Long term debt as a percentage of invested capital.	Leverage ratio
x_{36}	lt-debt as % of total debt	Long term debt as a percentage of total debt.	Leverage ratio
x_{37}	total debt as % total assets	Total debt as % total assets.	Leverage ratio
x_{38}	Working captial as % of equity	Working captial as percentage of equity.	Liquidity ratio
x_{39}	Price/revenue ratio	Price / revenue.	Valuation ratio
x_{40}	Price/equity ratio	Price / equity.	Valuation ratio
x_{41}	Price/tangible book ratio	Price / tangible book value	
x_{42}	Working capital as % of price	Working capital as percentage of price.	Liquidity ratio
x_{43}	Return on stock equity (roe)	Return on stock equity	Profitability ratio
x_{44}	Return on capital invested (roci)	Return on capital invested	Profitability ratio
x_{45}	Return on assets (roa)	Return on assets	Profitability ratio
x_{46}	Price/cash flow ratio	Price / cash flow	Valuation ratio

x_{47}	Price/free cash flow ratio	Price / free cash flow	Valuation ratio
x_{48}	Growth Net Income qr-o-qr	Growth in net income in the most recent quarter.	Growth ratio
x_{49}	Growth Net Income four quarter average	Average quarterly growth in net income over the past 4 quarters.	Growth ratio
x_{50}	Growth Net Income y-o-y	Growth in net income in the most recent year.	Growth ratio
x_{51}	Growth Net Income 4-year-average	Average annual growth in net income over the past 4 years.	Growth ratio
x_{52}	Growth total assets qr-o-qr	Growth in total assets in the most recent quarter.	Growth ratio
x_{53}	Growth total assets four quarter average	Average quarterly growth in total assets over the past 4 quarters.	Growth ratio
x_{54}	Growth total assets y-o-y	Growth in total assets in the most recent year.	Growth ratio
x_{55}	Growth total assets 4-year-average	Average annual growth in total assets over the past 4 years.	Growth ratio
x_{56}	Growth free cash flow qr-o-qr	Growth in free cash flow in the most recent quarter.	Growth ratio
x_{57}	Growth free cash flow four quarter average	Average quarterly growth in free cash flow over the past 4 quarters.	Growth ratio
x_{58}	Growth free cash flow y-o-y	Growth in free cash flow in the most recent year.	Growth ratio
x_{59}	Growth free cash flow 4-year-average	Average annual growth in free cash flow over the past 4 years.	Growth ratio
x_{61}	Growth revenue qr-o-qr	Growth in revenue in the most recent quarter.	Growth ratio
x_{62}	Growth revenue four quarter average	Average quarterly growth in revenue over the past 4 quarters.	Growth ratio
x_{63}	Growth revenue y-o-y	Growth in revenue in the most recent year.	Growth ratio
x_{64}	Growth revenue 4-year-average	Average annual growth in revenue over the past 4 years.	Growth ratio

B Goodness of fit

The figures below give an indication for the data-fit the linear models achieve. They are basically correlation plots of the predictions made by the model and the actual targets. If the predictions match the targets exactly, there would be a diagonal regression line where the diagonal is plotted as a dotted line in the correlation plot. Both the correlation plot for the training and test data are plotted, to indicate whether the model over-fits on the training data. The fit ranges from hardly any fit, to considerable over-fit to a reasonably good fit. In almost all quarters the model could find a reasonable fit in the training data. The fit on the test data was a lot less stable, if the market sentiment was similar in the test data as it was during the training data, then the fit was relatively good in the test data. If the market sentiments were different, the test-fit was considerable worse.

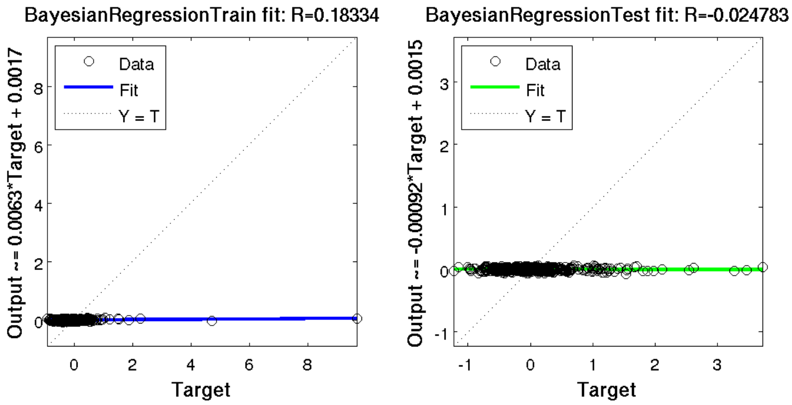


Figure 7: In some cases there does not seem to be any fit at all, it just seems to predict the average.

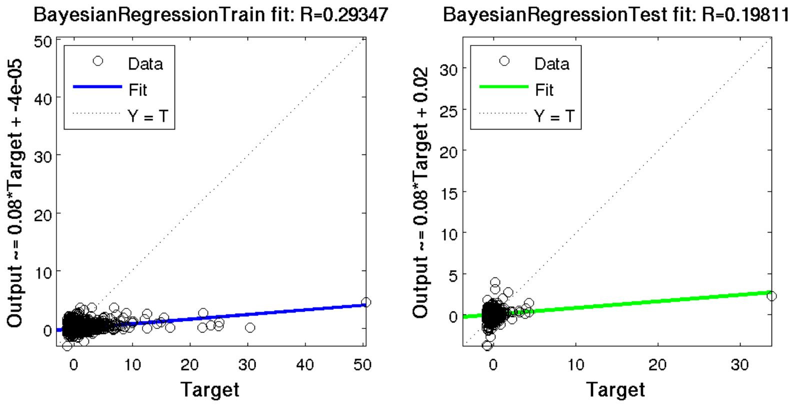


Figure 8: The fit can also be fairly reasonable, though there is some over-fit.

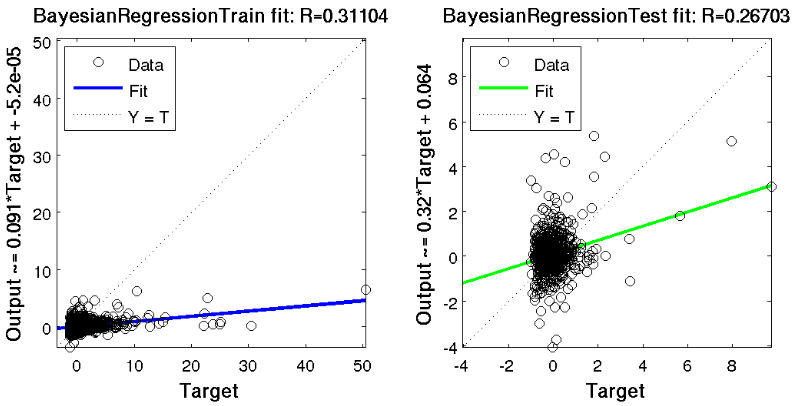


Figure 9: The regression line can become fairly steep as well, the variance in the test data can get quite high as well.

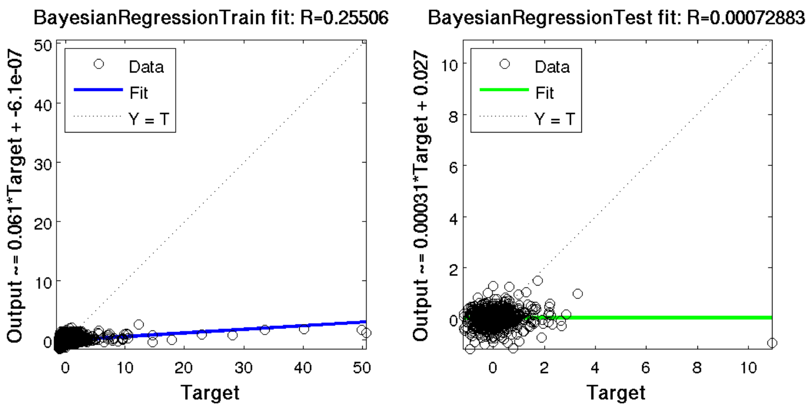


Figure 10: There can also be substantial over-fitting, which can be seen in the R-value, which is considerable higher in the training data than in the test data.

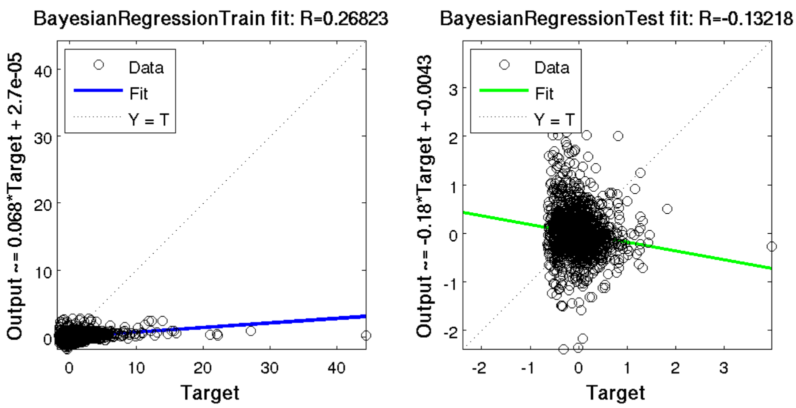


Figure 11: The regression line can get a negative slope as well, though it is rare. In this case the over-fit is especially considerable.

C Prototypes

The prototypes are the average (or median) member of a certain population. It is non-existent company of a population with all values filled in with either the average or median value. So the prototype of the committee is a non-existent company, for which all values are filled in with the average value of all companies the committee invested in. The prototype for all outperforming companies likewise is a non-existent company for which all values are filled in with the average of all companies that outperform the market. The prototypes provide valuable inside on their own, as they indicate what the average company that outperforms the market looks like. Given that we want to learn the models to find such companies, we would want the prototype the models learn to be similar to the actual prototype of outperforming companies.

Table 5: Size prototype, all values are in millions of USD. The committee prototype of the portfolios, or what the average (or median) company the committee invests in looks like. The top-30 is the average (or median) of the actual top-30 investments. The outperforming is the average (or median) of all companies that outperformed the market.

Population	Mean/Median	Earnings	Assets	Revenue
Committee	Mean	25.45	6886	1109
	Median	3.73	1106	270
Top-30	Mean	0.62	599	125
	Median	0.40	67	13
Outperforming	Mean	65.91	4330	795
	Median	5.40	449	70

Table 6: Valuation prototype. The committee prototype of the portfolios, or what the average (or median) company the committee invests in looks like. The top-30 is the average (or median) of the actual top-30 investments. The outperforming is the average (or median) of all companies that outperformed the market. P/E is the price to earnings, P/S is the price to sales (or price to revenues), P/B is the price to tangible book value and P/EQ is the price to equity.

Population	Mean/Median	P/E	P/S	P/B	P/EQ
Committee	Mean	15.88	2.70	2.72	4.91
	Median	0.00	0.61	1.62	1.64
Top-30	Mean	18.14	15.26	2.90	4.52
	Median	0.18	1.07	2.29	2.16
Outperforming	Mean	22.50	9.50	3.90	4.14
	Median	14.23	1.30	2.50	2.26

Table 7: Profit margin prototype. The committee prototype of the portfolios, or what the average (or median) company the committee invests in looks like. The top-30 is the average (or median) of the actual top-30 investments. The outperforming is the average (or median) of all companies that outperformed the market. RoA is return on assets and RoE is the return on equity.

Population	Mean/Median	Gross Margin	Net Profit Margin	RoA	RoE
Committee	Mean	39.88	-1.80	-3.69	15.60
	Median	37.03	0.10	0.13	0.00
Top-30	Mean	34.94	-3.59	-66.47	8.93
	Median	33.10	0.00	0.00	0.00
Outperforming	Mean	37.76	3.40	-12.63	12.76
	Median	36.53	4.80	4.38	10.05

Table 8: Leverage prototype. The committee prototype of the portfolios, or what the average (or median) company the committee invests in looks like. The top-30 is the average (or median) of the actual top-30 investments. The outperforming is the average (or median) of all companies that outperformed the market. QR is the quick ratio, CR is the current ratio, TD/EQ is the total debt to equity ratio and LR is the leverage ratio.

Population	Mean/Median	QR	CR	TD/EQ	LR
Committee	Mean	1.32	2.19	3.48	3.96
	Median	0.95	1.55	0.61	2.60
Top-30	Mean	3.70	4.50	2.31	4.07
	Median	1.43	2.20	0.08	1.60
Outperforming	Mean	2.69	3.54	3.12	4.81
	Median	1.20	1.93	0.29	1.90

Table 9: Growth prototype. The committee prototype of the portfolios, or what the average (or median) company the committee invests in looks like. The top-30 is the average (or median) of the actual top-30 investments. The outperforming is the average (or median) of all companies that outperformed the market.TA/4q is the four quarter average total assets growth and TA/4y is the four year average total assets growth. E/4q is the four quarter average earnings growth and E/4y is the four year average earnings growth. Zero values usually indicate missing data. If a company does not have four years worth of data, it also does not have a four year average growth value.

Population	Mean/Median	TA/4q	TA/4y	E/4q	E/4y
Committee	Mean	1.14%	4.85%	3.48%	3.96%
	Median	0.51%	0.00%	0.61%	2.60%
Top-30	Mean	6.80%	6.73%	2.31%	4.07%
	Median	0.90%	0.00%	0.08%	1.60%
Outperforming	Mean	8.40%	12.10%	3.12%	4.81%
	Median	1.89%	2.22%	0.29%	1.90%

Table 10: Growth prototype. The committee prototype of the portfolios, or what the average (or median) company the committee invests in looks like. The top-30 is the average (or median) of the actual top-30 investments. The outperforming is the average (or median) of all companies that outperformed the market.FC/4q is the four quarter average free cash flow growth and FC/4y is the four year average free cash flow growth. R/4q is the four quarter average revenue growth and R/4y is the four year average revenue growth. Zero values usually indicate missing data. If a company does not have four years worth of data, it also does not have a four year average growth value.

Population	Mean/Median	FC/4q	FC/4y	R/4q	R/4y
Committee	Mean	-53.75%	57.92%	5.36%	6.26%
	Median	8.80%	0.00%	2.10%	0.00%
Top-30	Mean	4.75%	4.93%	16.41%	43.13%
	Median	2.76%	0.00%	2.64%	0.00%
Outperforming	Mean	-12.91%	-24.95%	13.12%	10.85%
	Median	0.00%	0.00%	2.75%	0.00%