Parametric Portfolio Selection: Evaluating and Comparing to Markowitz Portfolios

(Seleção Paramétrica de Portfólios: Avaliação e Comparação com Portfólios de Markowitz)

Marcelo C. Medeiros*
Artur M. Passos**
Gabriel F. R. Vasconcelos***

Abstract

In this paper we exploit the parametric portfolio optimization in the Brazilian market. Our data consists of monthly returns of 306 Brazilian stocks in the period between 2001 and 2013. We tested the model both in and out of sample and compared the results with the value and equal weighted portfolios and with a Markowitz based portfolio. We performed statistical inference in the parametric optimization using bootstrap techniques in order to build the parameters empirical distributions. Our results showed that the parametric optimization is a very efficient technique out of sample. It consistently showed superior results when compared with the VW, EW and Markowitz portfolios even when transaction costs were included. Finally, we consider the parametric approach to be very flexible to the inclusion of constraints in weights, transaction costs and listing and delisting of stocks.

***Keywords*:** parametric portfolio; portfolio optimization; portfolio policies.

JEL codes: G12; G11.

Submitted 3 June 2014. Reformulated 13 August 2014. Accepted 21 August 2014. Published on-line 6 October 2014. The article was double blind refereed and evaluated by the editor. Supervising editor: Ricardo P. C. Leal. The views expressed in this paper are those of the authors and not necessarily reflect those of the Central Bank of Brazil or the Brazilian Development Bank.

*Department of Economics, Pontifical Catholic University of Rio de Janeiro, Rio de Janeiro, RJ, Brazil. E-mail: mcm@econ.puc-rio.br

**Department of Economics, Pontifical Catholic University of Rio de Janeiro, Rio de Janeiro, RJ, Brazil. E-mail: artur.passos@gmail.com

***Department of Electrical Engineering, Pontifical Catholic University of Rio de Janeiro, Rio de Janeiro, RJ, Brazil. E-mail: gabrielrvsc@yahoo.com.br

ISSN 1679-0731, ISSN online 1984-5146

Resumo

Este trabalho tem como objetivo explorar o método de otimização paramétrica de portfólios no mercado brasileiro. O banco de dados utilizado consiste de retornos mensais de 306 ações brasileiras no período de 2001 à 2013. O ajustamento do modelo foi testado dentro e fora da amostra e os resultados foram comparados com o portfólio igualmente balanceado e o portfólio otimizado via Markowitz. A inferência estatística foi feita utilizando técnicas de *bootstrap* para construir distribuições empíricas dos parâmetros. Os resultados mostraram que o método paramétrico é uma técnica muito eficiente fora da amostra. Ela mostrou resultados superiores quando comparada com as demais técnicas mesmo na presença de custos de transação. Por último, pode-se concluir que o método paramétrico é muito flexível para a inclusão de restrições em pesos, custos de transação e entrada e saída de ações no mercado.

***Palavras-chave*:** portfólio paramétrico; otimização de portfólio; política de portfólio.

1. Introduction

The research of portfolio selection is motivated by two factors. First, the increase of market complexity and computing power pose both challenges and opportunities to the investor's decision problem. Second, the traditional mean-variance approach of Markowitz (1952), despite several recent improvements, deals with a big number of arguments to be optimized given that we need to estimate all portfolio weights directly as arguments in the objective function. The Markowitz approach can be viewed as the minimization of the portfolio expected conditional volatility given a target conditional expected return and asset's covariances. The application of this approach goes beyond the simple minimization of the expected conditional volatility, it considers several other restrictions and fixes such as constraining portfolio weights, imposing transaction costs, shrinkage of the estimates and other adjustments that ensures the model is as close as possible to the reality of the market and the estimates are adequate. However, the complexity of the model increases significantly when the number of restrictions grows bigger. A survey of these studies is found in Brandt (2010).

Our objective in this paper is to evaluate the parametric portfolio optimization approach proposed by Brandt *et al.* (2009) using monthly data of Brazilian stocks. We used as characteristics (explanatory variables) the book-to-market ratio (BTM), the market equity (ME), defined as the number of stocks of a given company times its price, and the one year momentum (MOM). We compared the results of the parametric portfolio to the value-weighted portfolio, equal-weighted portfolio, and the traditional Markowitz-based portfolio. The base parametric model refers to the model with a simple linear restriction on weights. The parametric model is extended to include short sale constraints[1], maximum absolute weight on individual stocks, transaction costs, and the inclusion of a risk-free asset in the investable set.

The optimized portfolios consistently show risk-adjusted returns above all other portfolios. These results remain after imposing weight constraints and market costs. The optimized parametric portfolios are also superior to Markowitz based portfolios. We find the parametric approach to yield better results, be computationally simpler, it makes easier to handle changes in the investable set through time, and market costs are easy to implement. We used the constant relative risk aversion (CRRA) utility function to estimate the optimal portfolios, we tested the model for several levels of risk aversion, and even for extreme risk aversions $\gamma = 100$, the investor chooses to keep some of his wealth in stocks when the risk free asset is available. Nevertheless, the certainty equivalent, in this case, converges to the risk free rate as the relative aversion grows bigger. Furthermore, we tested the parametric portfolio optimization out of sample performance and it yielded higher returns than the market and the out of sample Markowitz based, even when transaction costs are included. Although the parametric optimization is described by Brandt *et al.* (2009) as a method of moments estimator from Hansen (1982), the estimation of the covariance matrix of the parameters may be troublesome when facing nonlinear constraints, to solve this issue we used bootstrap techniques and the parameters were estimated through nonlinear optimization methods. Additionally, the risk-adjusted returns above markets obtained using a few publicly available data suggest the Brazilian stock market is still inefficient. Finally, the sample period we used involves several changes in the economical and financial environment in Brazil, which are captured in all portfolios. The fact the parametric portfolios are formed by constant coefficients across the entire period is a strong argument in favor of (i) the parametric approach and (ii) the inefficiency of Brazilian stock market.

[1]This extension also follows Brandt *et al.* (2009)

This paper is organized as follows: section 1 describes the parametric portfolio optimization and its extensions. In section 2 we discuss the bootstrap for statistical inference, section 3 describes the data and an empirical application. Our final remarks are in section 4.

2. Methodology

2.1 Parametric approach

This section aims to present the basic parametric portfolio optimization, originally and better described in Brandt *et al.* (2009).

Let N_t be the number of stocks in the investable set at each date t. Each stock i has a return $r_{i,t+1}$ from period t to period $t+1$ and is associated to a vector of characteristics[2] observed at period t.

A portfolio is a vector of weights $\{w_{i,t}\}_{i=1}^{N_t}$ which has a return $r_{p,t+1}$. The investor's problem is to choose the portfolio that maximizes his conditional expected utility,

$$\max_{\{w_{i,t}\}_{i=1}^{N_t}} E_t[u(r_{p,t+1})] = E_t\left[u\left(\sum_{i=1}^{N_t} w_{i,t} r_{i,t+1}\right)\right] \quad (1)$$

The parametric approach has such name because it parametrizes the optimal portfolio weights as a functions of stocks' characteristics. The parametrization may be very general, but in this paper we used the weights as a linear function of the characteristics from Brandt *et al.* (2009), such that:

$$w_{i,t} = \bar{w}_{i,t} + \frac{1}{N^t}\theta^\top \hat{x}_{i,t} \quad (2)$$

where $\bar{w}_{i,t}$ is the weight of stock i at time t on a benchmark [3] portfolio, θ is the vector of coefficients to be estimated, $\hat{x}_{i,t}$ are the observed characteristics of stock t, they are standardized cross-sectionally to have zero mean and unit standard deviation across all stocks at date t. The symbol $\top$ indicates that the vector is transposed.

The zero mean cross-section ensures the weights always sum one, regardless the values of θ. The unit standard deviation makes sure $\hat{x}_{i,t}$ is

[2] Characteristics are variables containing information regarding the stocks, e.g. market capitalization, book-to-market, dividend yield and the lagged twelve-month return

[3] A benchmark portfolio is, for example, the equal-weighted portfolio or the value-weighted portfolio.

stationary trough time, while the raw $x_{i,t}$ may be non-stationary. Finally, the term $\frac{1}{N^t}$ is a normalization that allows the portfolio weight function to be applied to any time varying number of stocks. Without this normalization, an increase in the number of stocks would result in increased leverage with larger long and short positions.

The idea of the linear parametrization of equation (2) is that of a portfolio management relative to a performance benchmark. Since the characteristics are cross-sectionaly standardized, the term $\theta^\top \hat{x}_{i,t}$ will have a zero cross-sectional mean, and consequently, the deviations from the optimal portfolio wights from the benchmark will sum to zero, and this ensures the optimal portfolio weights sum one.

In the traditional Markowitz approach, one would have to estimate portfolio weights for each stock for each period of time. In the parametric portfolio optimization we estimate weights as a single function of characteristics, the parameters are the same for all stocks and for every period of time.

The key to understand Brandt *et al.* (2009) parametrization is the constant coefficients of θ for all stocks and for the entire period. Constant coefficients across stocks imply that the portfolio weights depends only on stocks' characteristics, i.e. stocks with similar characteristics must have similar weights even if their past returns are different. The implications of the constant θ through time are even stronger, the coefficients that maximize the investor's conditional expected utility at a given date are the same for all dates, and therefore they also maximize the investor's unconditional expected utility. If a set of stocks is described by five parameters, they will be the only five parameters for all assets in the entire period. Thus, the coefficient of time can be excluded from the expectation of equation (1):

$$\max_{\theta} E[u(r_{p,t+1})] = E\left[u\left(\sum_{i=1}^{N_t} w_{i,t} r_{i,t+1}\right)\right] \tag{3}$$

The coefficients that maximizes the investor's unconditional utility are:

$$\hat{\theta} = \arg\max_{\theta} E[u(r_{p,t+1})] = \arg\max_{\theta} E\left[u\left(\sum_{i=1}^{N_t} w_{i,t} r_{i,t+1}\right)\right] \tag{4}$$

Thus, the sample analogue of (4) for some given utility function and the portfolio police of equation (2) is:

$$\begin{aligned}\hat{\theta} &= \arg\max_{\theta} \frac{1}{T}\sum_{t=0}^{T-1} u(r_{p,t+1}) \qquad (5)\\ &= \arg\max_{\theta} \frac{1}{T}\sum_{t=0}^{T-1} u\left(\sum_{i=1}^{N_t}(\bar{w}_{i,t} + \frac{1}{N^t}\theta^\top \hat{x}_{i,t}) r_{i,t+1}\right)\end{aligned}$$

Due to the relatively low dimensionality of the parameter vector, the suggested approach is computationally simple. The complexity of the problem increases only if we choose to add more characteristics to the model, the increase of the number of assets has nearly no effect in estimating $\hat{\theta}$. Besides, the use of few parameters increases numerical robustness and may reduce the risk of in-sample overfitting.

Regarding the utility function, we chose to use the constant relative risk aversion[4] (CRRA), with a base relative risk aversion γ of five. This function has three desirable characteristics. First, they incorporates preferences towards higher-order moments without introducing additional preference parameters. Second, the function is twice continuous differentiable, which helps the optimization process. Finally, the CRRA function is optimal under a partial myopic behavior (e.g. Mossin (1968)).

The optimal coefficients $\hat{\theta}$ defined in equation (5) can be interpreted as a method of moments estimator. However, the constraints imposed to the problem in the next section prevent us from using the asymptotic properties of the covariance matrix $\hat{\Sigma}_\theta$. Instead, $\hat{\Sigma}_\theta$ is estimated by bootstrapping techniques.

2.2 Extensions

This section aims to find the adequate extensions of the base model described in 2.1 for the Brazilian stock market. First, the absence of a liquid market for short positions during most of the evaluated period requires limits on negative weights. Second, local regulations imposes maximum weights on single stocks. Third, we discuss the inclusion of a risk free asset in the investable set. Last, the parametric approach works with high turnover portfolios, requiring the implementation of transaction costs.

[4]The CRRA utility function has the advantage of having the relative risk aversion measure as a parameter of the function (see Ljungqvist & Sargent (2004)).

$$u(x) = \begin{cases} \frac{x^{1-\gamma}-1}{1-\gamma} \gamma > 0, \;\; \gamma \neq 1 \\ \log(x) \quad \gamma = 0 \end{cases}$$

The general idea of these extensions are also demonstrated in Brandt *et al.* (2009). Our goal here is to adapt them to the Brazilian framework.

Weight constraints

The most common weight constraint applied to portfolios is to forbid short-sales. We follow Brandt *et al.* (2009) and simply truncate negative weights. Unfortunately this restriction generates two issues:

- The nonlinearity prevent us from using the asymptotic covariance matrix of coefficients derived from the method of moments estimator.
- If negative weights are not allowed, the sum of the optimal portfolio weights becomes more than one.

The use of bootstrap techniques solves the first issue. To solve the second issue, Brandt *et al.* (2009) created a renormalization of the portfolio weights as follows:

$$w_{i,t}^{+} = \frac{\max[0, w_{i,t}]}{\sum_{j=1}^{N_t} \max[0, w_{i,t}]} \tag{6}$$

The renormalization above only works because 0 is division invariant. The idea behind equation (6) is to turn every negative weight into 0 and divide all weights by the sum of all non-negative weights. Unfortunately the solution after the normalization may not be optimal because when weights are truncated in 0 the optimal solution may have a positive weight that would be negative in the optimal unconstrained solution.

Normally, a short seller sells stocks that have been borrowed from an other investor who owns them. The alternative, know as naked short selling, is selling stocks without previously borrowing them. The Brazilian short sale market has the following characteristics:

- Naked short selling is forbidden.
- The investor who originally owns the stock has its devolution guaranteed by the stock exchange. He also remains receiving dividends from the rented stock.
- The contract may allow anticipated devolution, with the borrowing investor paying taxes proportional to the time remaining to the contract to end.

- The market liquidity is expanding and the interest rate is becoming less prohibitive. However, during most of the sample period, a small number of stocks had an active market for borrowing.

Another weight restriction to be considered when using Brazilian data is the maximum portfolio weights imposed by Securities and Exchange Commission of Brazil (CVM) to investment funds. To satisfy this imposition we use a maximum weight of 20% of the investor's wealth in a single asset. When short sales are allowed, this constraint is also applied to short positions.

In order to apply the maximum absolute weight constraint of 20% we subtract a penalty function from the objective function to be optimized. The penalty is presented as follows:

$$penalty = \alpha \sum |w_{i,t}^{>r^*}| \tag{7}$$

where $\sum |w_{i,t}^{>r^*}|$ is the sum of all weights that are bigger in absolute value than r^* (20% in our case) and α is a positive number which forces the restriction to be satisfied, in our[5] case it was 1. The new objective function is:

$$\max_{\theta} E[u(r_{p,t+1})] = E\left[u\left(\sum_{i=1}^{N_t} w_{i,t} r_{i,t+1}\right)\right] - \alpha \sum |w_{i,t}^{>r^*}| \tag{8}$$

The penalty function works similarly to a Lagrange constrained optimization. If the penalty is significantly big compared to the rest of the function, it will be prioritized in the optimization. When the minimum possible value of the penalty function is achieved the restriction is satisfied, this minimum value is 0, i.e. the sum of all absolute weights above r^* is null. The only way this is true is if no weight is bigger in absolute value than r^*. To input the no short-sale constraint we used the Brandt *et al.* (2009) normalization from equation 6 in the optimization problem.

2.2.1 Risk-free Bond

We extend the base model to include a generic risk-free bond in the investor's problem. The inclusion of such bond allows us to verify the

[5]We had to use higher values for α only in the $\gamma = 100$ optimizations, this is because the utility over the relative risk aversion of 100 has a very big absolute value, which makes the $\alpha = 1$ penalty despicable and the restriction is not satisfied.

optimal position for investors with different coefficients of risk aversion by using the same CRRA utility function with different values for the relative risk aversion γ.

In order to extend the problem, we include a factor θ_{rfw}. This factor defines the weight in the risk-free bond: $\omega_{rf} = \max\{\theta_{rfw}, 0\}$ for all date t. Therefore, $1 - \theta$ will be the proportion of wealth invested in the risky portfolio. This way, the sum of weights will remain 1.

This approach makes the weight of the risk-free bond constant through time. Besides, no original parameters changes the weights in such bond, i.e. the investor's decision to put some of his wealth in the risk-free bond does not depend on the characteristics used for stocks. Similarly, the risk-free bond weight does not change the relative stock weights. The investor's decision can be divided in two steps: first, one must choose an optimal portfolio from the set of available stocks; second, once the investor has chosen his optimal portfolio, he must divide his wealth between the optimal portfolio and the risk-free bond according to his preferences for risk.

2.2.2 Transaction Costs

In this section we present how Brandt *et al.* (2009) deal with transaction costs and how we adjust their model to our problem. The first portfolios we generated using the parametric approach, with no transaction costs, have shown risk-adjusted returns above the market, which may be caused by an arbitrage over market imperfections. The inclusion of transaction costs makes this arbitrage harder to be exploited. Therefore, we aim to verify if the inclusion of transaction costs changes the portfolio's core strategy and if the risk-adjusted returns above the market[6] are persistent under these costs.

For a given portfolio policy, the turnover at each period t is the sum of all absolute changes in weights from $t-1$ a t such that,

$$T_t = \sum_{i=1}^{N_t} |w_{i,t} - w_{i,t-1}| \tag{9}$$

A turnover of x at date t indicates a change of $x{*}100\%$ of the portfolio's wealth in the positions from $t-1$ to t. Note that this change is on the absolute sum of portfolio wealth, which sums more than 100% in cases when short-sales are allowed.

[6]We do not show results for the IBOVESPA portfolio, instead we show the performance of the equal Weighted Portfolio and the value Weighted because they performed better than IBOVESPA in our sample.

Therefore, the return of the portfolio net of trading costs is

$$r_{p,t+1} = \sum_{i=1}^{N_t} (w_{i,t} r_{i,t+1} - c_{i,t} |w_{i,t} - w_{i,t-1}|) \tag{10}$$

where $c_{i,t}$ represents the proportional transaction cost for stock i at time t. It is well known that transaction costs show considerably variation across stocks, being larger for small caps than for large caps. It is also known that transaction costs have been decreasing over time due to increased liquidity, but since our sample period is relatively short, we do not include this trend. The variable $c_{i,t}$ from equation (10) may be, for example, a function of the portfolio turnover or stocks' market equity.

Finally, we do not consider taxes on profits from stock trading because Brazil has an homogeneous tax across it, given no strategy we use involves day-trade. The tax is currently 15%, only the net profit and losses can be used to mitigate taxes from future profits.

3. Statistical Inference

Brandt *et al.* (2009) show that the parametric optimization can be interpreted as a method of moments estimation as in Hansen (1982). However, the covariance matrix of the parameters $\hat{\Sigma}_\theta$ depends on asymptotic results to be properly estimated. In developing countries like Brazil it is not very good to count on asymptotic results given that the samples are not very big, in our case, the sample[7] has 316 stocks if we consider the entire sample, but some periods have less than 100 stocks and the period with more stocks has 201. To bypass this issue Brandt *et al.* (2009) use bootstrap[8] techniques, which do not rely on asymptotic results.

[7]There are two dimension for the asymptotic properties of the model to be considered. First we would need the parameters to be consistent in the cross-sectional dimension, i.e. the number of stocks. Second, we also need the parameters to be consistent in the time. In our sample we have 132 periods of time (months) and only 39 stocks cover it all. Therefore, even if we consider 132 periods enough to rely on asymptotic results, some stocks cover much smaller periods.

[8]Even considering that their sample has thousands of stocks and covers a longer period, they still choose to use bootstrap techniques to make inference

These techniques are described in Efron & Tibshirani (1994), Davison *et al*. (1986) and Davison (1997). They consist on generating many random[9] sub-samples and estimating the optimization parameters on these sub-samples. As a result, we will have an empirical distribution of each parameter, the inference is then made on these distributions. Given that we know the entire distribution of the parameters, deviations from the normality[10] can be handled easier. To test a hypothesis, for example, that $x = k$ we analyzed directly if the value k is in the percentile of 5% of the distribution of x.

Since we used monthly returns, it is reasonable to assume that they are not autocorrelated. However, we tested them for both autocorrelation and heteroskedasticity. Since we have one time serie for each stock, the test results are presented in figures 1 and 2. They show the autocorrelations of lags one to four of the returns and the squared returns respectively. In both cases most of the autocorrelations are smaller than 0.2 in absolute value. In fact, 78% of the autocorrelations of the returs are below 0.2 in the first lag, 90% in the second lag, 93% in the third lag and 95% on the fourth lag. For the square returns these values are 83% in the first lag, 89% in the second lag, 91% in the third lag and 88% in the fourth lag. As for the significance level of these autocorrelations, less than 5% are significant at 5% and none at 1% significance level. The results are similar for the partial autocorrelation function.

[9]In our case, the sub-samples were randomly generated across the time, since it makes no sense to sort the characteristics and the returns cross-sectionally given that it would eliminate all relation between the characteristics and the returns

[10]Most of the empirical distributions of parameters we estimated were asymmetric

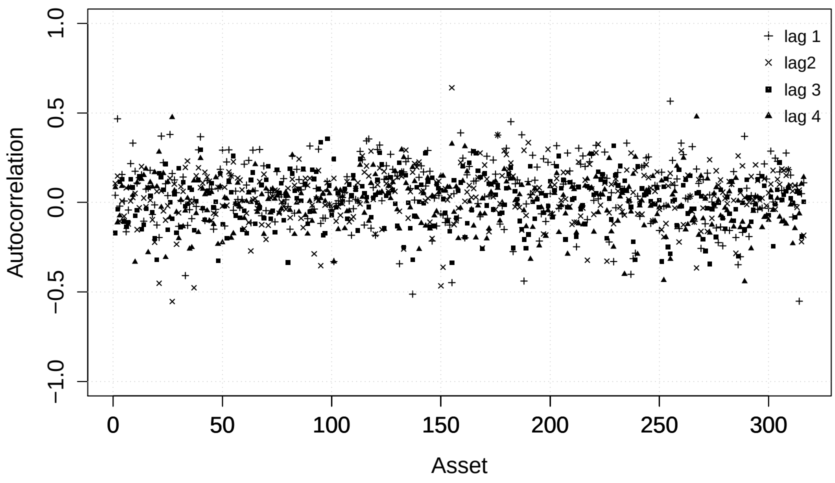

Figure 1
Autocorrelation of retuns – lags 1 to 4

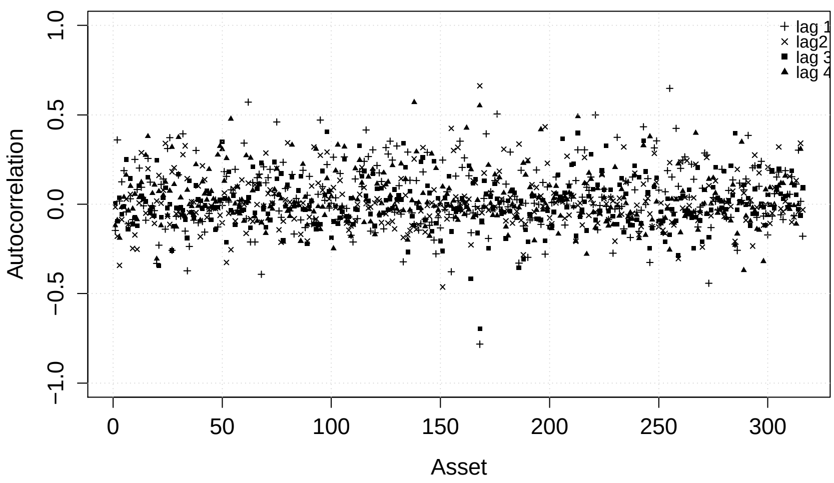

Figure 2
Autocorrelation of squared retuns – lags 1 to 4

In this paper we used 1000 bootstrap samples across t to calculate the percentiles and the standard deviation of the parameters. In some cases, there were outliers in the parameters distributions that pulled up the standard deviation, this happened when for some bootstrap sample the optimization was not feasible. Although some of the outliers affect the standard deviation, they have no effect when we test the hypothesis using the percentiles. However, when some bootstrapped parameter was more than 50 times bigger (in absolute value) than the one that came immediately before, we excluded it from the bootstrap distribution and generated an extra random sub-sample to replace it.

4. Empirical Application

4.1 The data

The data is composed of monthly returns from 2001 to 2013 obtained trough the Economática historical data. Initially we extracted the entire Bovespa data set, but most of the stocks were very illiquid and had no price for a large number of months. We created a criteria to extract only the fairly liquid stocks described as follows. We selected all stocks with returns available form July 2001 to June 2003, then we did the same from July 2002 to June 2004 and the same procedure were repeated until July 2013. In the end we had 316 stocks in the investable set, the number of stocks across the period is shown in figure 3. By following Fama & French (1996) and Brandt *et al.* (2009), we used the BTM from December $t-1$ to explain the returns from July t to June $t+1$. This six month gap is commonly used as the time needed for the information to be public. The book value we used is the company's total assets minus liabilities, companies with negative BTM were excluded from the data. The market equity is a measure of firm's size, it is defined as the log of the price per share times the number of shares. We used the ME from July t to e explain returns from January t to December t. Last, the one year momentum was calculated as the compounded return from $t-13$ to $t-1$. Moreover, we selected only stocks with 24 continuous returns to certify that all stocks would have an one year momentum. In fact, most investment funds do not invest in stocks at the moment they are listed, on the contrary, they wait for some time to see how the new stock behaves.

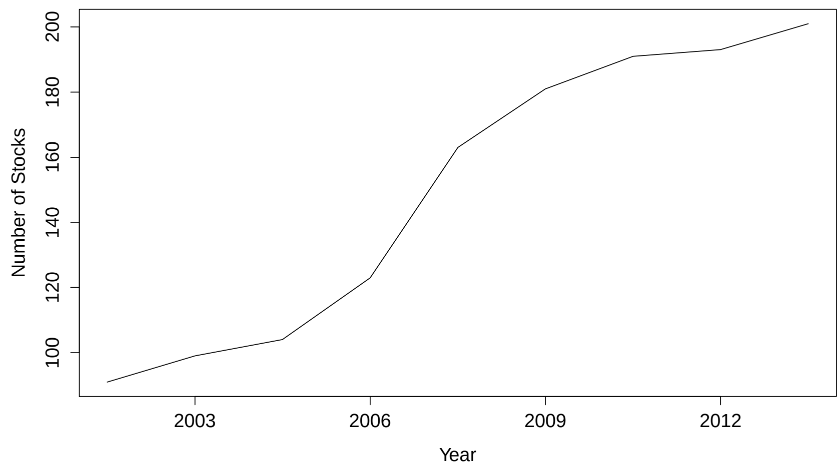

Figure 3
Number of stocks over the years

4.2 Simple linear restriction

Table 1 shows the results for the base case, the maximum weight for both long and short positions is 20%. This restriction is due to legal and contractual conditions that forbid investment funds to have very high positions in a single asset. There is no risk free asset, no transaction costs and no short sale restrictions. We used the CRRA utility function with relative risk aversion γ of 5. We also display the performance of the value weighted portfolio and the equal weighted portfolio in the first and second columns. The third column shows the results for the parametric case in sample and the last column shows its out of sample performance.

The rows of table 1 are divided into three categories. The first category of rows shows the parameters of the optimization, their standard deviation and the percentile containing the value of the null hypothesis ($\theta = 0$), the last two were calculated using 1000 bootstrap samples. In order to obtain out of sample results, we optimized on the first three years of the sample and used the parameters to estimate the following year portfolio. Then, the procedure was repeated and the first four years were used to estimate the portfolio on the fifth year. This was done until the last year by enlarging the sample. The out of sample parameters are the mean of all sub-samples optimization parameters. The same is valid for the standard deviation and the percentile. We also used 1000 bootstrap samples for each sub-sample during the out of sample estimation.

The second set of rows describes the distribution of weights, the items are, respectively, the average weight on each stock, the maximum average weight on a single stock, the minimum average weight on a single stock, the average sum of short positions, the average fraction of stocks with weight lesser or equal than 0, and the monthly turnover of the portfolio. Note that these averages are across t.

The last rows shows some properties of the optimized portfolio returns. They are, the certainty equivalent, the mean, the standard deviation, the Sharpe ratio with $rf = 0$ and the Sharpe ratio with $rf = 10\%$ per year. All values are annualized for better interpretation. We used two different Sharpe ratios because the Brazilian proxies of the risk free rate are very high compared to U.S. and European rates.

Table 1
Simple linear restriction

Variable	EW	VW	In Sample PPP	Out of Sample PPP
θ_{btm}	-	-	1.199	1.586
Std.	-	-	(0.315)	(0.603)
Per.	-	-	(0.006)	(0.037)
θ_{me}	-	-	-0.162	-1.319
Std.	-	-	(0.592)	(0.969)
Per.	-	-	(0.367)	(0.094)
θ_{mom}	-	-	2.492	1.696
Std.	-	-	(0.358)	(0.842)
Per.	-	-	(0.000)	(0.095)
$\lvert w_i\rvert \times 100$	0.786	0.786	1.734	1.585
$\max\ w_i \times 100$	0.786	1.102	10.701	11.432
$\min\ w_i \times 100$	0.786	0.485	-3.657	-3.287
$\sum w_i I(w_i < 0)$	0	0	-0.599	-0.601
$\sum I(w_i \leq 0)/N_t$	0	0	0.398	0.379
$\sum \lvert w_{i,t} - w_{i,t-1}\rvert$	0.005	0.009	0.713	0.547
CE	0.001	-0.002	0.237	0.182
$\bar{r}$	0.154	0.149	0.425	0.359
$\sigma(r)$	0.219	0.219	0.236	0.227
SR	0.701	0.679	1.799	1.584
SR$_{rf=0.1}$	0.245	0.222	1.376	1.143

The parameters we obtained are consistent with the theory, θ_{btm} and

θ_{mom} are positive and θ_{me} is negative. However, θ_{me} was not significant in sample and marginally significant out of sample, most likely because the base portfolio $\bar{w}_{i,t}$ we used is the value portfolio, which already has ME information. Regarding the second set of rows, the average absolute weight of the optimal portfolio (1.734%) is more than twice that of the equal and value portfolios (0.786%). However, considering the number of stocks available each period, the positions of the optimal portfolio are not extreme. The average maximum and minimum weights support this fact, they are respectively, 10.7% and -3.66%. The average sum of negative weights is -0.60%, this implies that the average sum of positive weights is 160%. The average fraction of negative weights of the optimized portfolio is 0.40%. The parametric portfolio did not bet on extreme positions even considering the fact that the investable set is significantly smaller for the Brazilian market when compared to U.S. markets. The last item of this set of rows is the turnover, which is of 71% per month. The turnover is high because this base case has no transaction costs and because short sales are allowed. When we allow short sales the positions of the portfolio are higher, which makes the turnover also high. Note that the average sum of all absolute positions is 220%, the 71% turnover is 32% of the total position, which is not very high.

The certainty equivalent of the optimized portfolio is very high when compared to the value and equal weighted portfolios, the first is 24% in sample and 18% out of sample and the last two are -0.2% and 0.1%. The volatility is similar amongst all portfolios, but the average return of the optimized portfolio (42.5%) is nearly three times higher than the EW (14.9%) and VW (15.4%) portfolios. The out of sample average return is 35.9%. The two last items in the table are the Sharpe ratio and the Sharpe ratio with a 10% risk free rate, which are much higher for the optimized portfolio both in and out of sample.

Not only the parametric portfolio performed much better than the EW and VW portfolios, it kept the high performance out of sample. The main Brazilian market index (IBOVESPA) were not considered due to its poor performance along the sample period.

4.3 Long only

Table 2
Long only

Variable	EW	VW	In Sample PPP	Out of Sample PPP
θ_{btm}	-	-	2.717	3.570
Std.	-	-	(1.275)	(2.235)
Per.	-	-	(0.002)	(0.084)
θ_{me}	-	-	-1.624	-2.551
Std.	-	-	(1.529)	(2.553)
Per.	-	-	(0.173)	(0.064)
θ_{mom}	-	-	3.896	3.865
Std.	-	-	(2.500)	(3.021)
Per.	-	-	(0.002)	(0.045)
$\lvert w_i \rvert \times 100$	0.786	0.786	0.786	0.689
$\max\ w_i \times 100$	0.786	1.102	8.642	9.055
$\min\ w_i \times 100$	0.786	0.485	0	0
$\sum w_i I(w_i < 0)$	0	0	0	0
$\sum I(w_i \leq 0)/N_t$	0	0	0.490	0.475
$\sum \lvert w_{i,t} - w_{i,t-1} \rvert$	0.005	0.009	0.286	0.277
CE	0.001	-0.002	0.122	0.067
$\bar{r}$	0.154	0.149	0.286	0.225
$\sigma(r)$	0.219	0.219	0.222	0.222
SR	0.701	0.679	1.289	1.016
$\text{SR}_{rf=0.1}$	0.245	0.222	0.839	0.565

Most equity portfolios are faced with short-sale constraints, specially when considering small funds and the low liquidity for short-sales of the Brazilian market. We present the results for a long only portfolio in table 2, but the reader should have in mind that the methodology is very flexible, and we can, for example, optimize a portfolio with a 20% restriction for long and 5% restriction for short just as easily. In order to guaranty that all the weights were less than 20% out of sample, we had to use a stronger restriction in sample of 15% maximum weight. The parameters of the optimized portfolio have the same pattern as the base case regarding size, sign and significance, the θ_{me} is also not significant in sample and all parameters are only marginally significant out of sample. The parameters standard deviation are greater than the base case, but this does not affect their significance since the bootstrap distribution of parameters are skewed in favor of the rejection of the null hypothesis. Figure 4 shows the parameters distributions. In the BTM and the MOM histograms the distributions are highly concentrated on positive values, on the other hand the ME distribution is concentrated on negative values, since the ME parameter itself is negative, but there is a considerable part of the distribution in positive numbers, which is why the ME requires a higher significance level to reject the null. This behavior of the bootstrapped parameters is persistent in cases of high standard deviation.

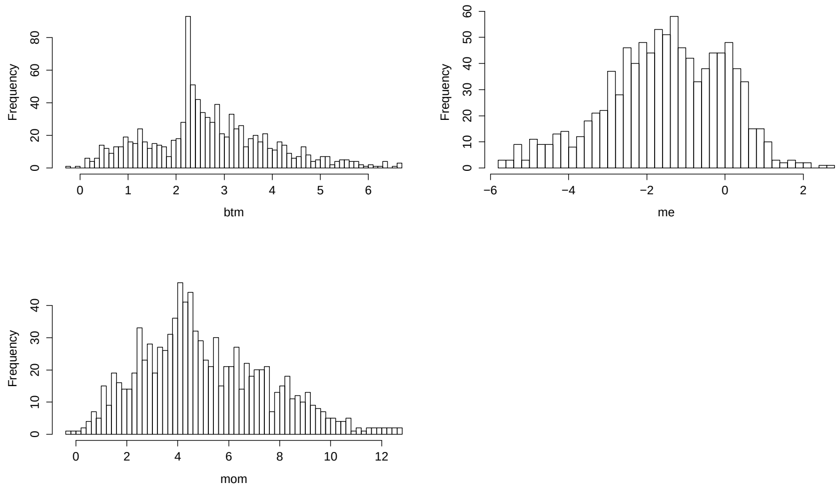

Figure 4
Histograms of the bootstrapped parameters for the long only portfolio

The portfolio weights on the other hand, are very different from the base case. The weights are smaller since the investor cannot use short positions to have a larger exposure to long positions. The average maximum weight is 8.64% for the long only against 10.7% for the base case. An interesting fact is that the mean absolute weight is the same for the optimal portfolio, the EW and the VW, this is due to the fact that the long portfolio simply divides the wealth amongst all assets, and in this case the average weight will always be the position of the equal weighted portfolio or, in a multi-period case, it will be the average of all equal weighted positions across t. The the average fraction of stocks with weight lesser or equal than 0 is, in this case, the proportion of stocks in which no wealth were invested. Last in this section of the table, the turnover in the long only portfolio (28.6%) is significantly less than that of the base case (71.3%). However, in this case, the sum of all absolute weights is 100%, which makes the proportional turnover equal 28.6% against 36% in the base case.

Due to the no short sales restriction, the certainty equivalent is 12.2% in sample and 6.7% out of sample. The average returns are also less than the base case, but they still bigger than the EW and VW examples. The in sample average return is 28.6% per year and the out of sample is 22.5% per year. Although the long only portfolio is not as good as the base case when we compare their returns, Sharpe ratios and certainty equivalent, it still better than the EW and VW portfolios.

There is a greater discrepancy between in sample and out of sample results for the Brazilian case when compared with the NYSE empirical application of Brandt *et al.* (2009), which had nearly 40 years of data and 8 years for the first out of sample sub-sample. In our case the data (after filtering) has only 11 years and if we try to extend it the number of assets decreases significantly in the years preceding. Moreover, we used only 3 years to the first out of sample sub-sample, otherwise we would not have a reasonable number of sub-samples. However, even with this greater discrepancy, the out of sample results still superior to the EW and VW portfolios with an acceptable margin.

4.4 Varying risk aversion

Table 3
Varying risk aversion – simple linear restriction

Variable	VW	In Sample PPP $\gamma=2$	Out of Sample PPP $\gamma=2$	In Sample PPP $\gamma=5$	Out of Sample PPP $\gamma=5$	In Sample PPP $\gamma=100$	Out of Sample PPP $\gamma=100$
θ_{btm}	-	1.301	1.648	1.119	1.586	0.941	0.471
Std.	-	(1.823)	(0.661)	(0.315)	(0.603)	(0.428)	(0.212)
Per.	-	(0.001)	(0.018)	(0.006)	(0.037)	(0.048)	(0.254)
θ_{me}	-	-0.374	-1.382	-0.162	-1.319	-0.226	-0.113
Std.	-	(1.936)	(2.589)	(0.592)	(0.969)	(0.243)	(0.112)
Per.	-	(0.140)	(0.040)	(0.367)	(0.094)	(0.309)	(0.173)
θ_{mom}	-	2.307	1.642	2.492	1.696	2.095	1.047
Std.	-	(1.833)	(1.764)	(0.358)	(0.842)	(0.939)	(0.466)
Per.	-	(0.001)	(0.047)	(0.000)	(0.095)	(0.023)	(0.224)
$\lvert w_i \rvert \times 100$	0.786	1.682	1.518	1.743	1.585	1.478	0.910
$\max\ w_i \times 100$	1.102	11.011	11.898	10.701	11.432	8.886	3.465
$\min\ w_i \times 100$	0.485	-3.375	-3.336	-3.657	-3.287	-2.923	-0.723
$\sum w_i I(w_i < 0)$	0	-0.703	-0.561	-0.559	-0.601	-0.437	-0.212
$\sum I(w_i \leq 0)/N_t$	0	0.401	0.372	0.398	0.379	0.366	0.174
$\sum \lvert w_{i,t} - w_{i,t-1} \rvert$	0.009	0.665	0.531	0.713	0.547	0.598	0.316
CE	-0.002	0.346	0.266	0.237	0.182	-0.859	-0.967
$\bar{r}$	0.149	0.420	0.335	0.425	0.359	0.376	0.222
$\sigma(r)$	0.219	0.235	0.231	0.236	0.227	0.220	0.219
SR	0.679	1.791	1.454	1.799	1.548	1.705	1.014
$SR_{rf=0.1}$	0.222	1.341	1.020	1.376	1.143	1.251	0.558

This section aims to compare the parametric portfolio optimization under different relative risk aversion γ. Table 3 shows the base case with $\gamma = 2$, $\gamma = 5$ and $\gamma = 100$ and the value weighted portfolio, which is the base portfolio in the policy described by equation 2. The VW portfolio in table 3 is for $\gamma = 5$, we do not show its results for other relative risk aversion because the only value that respond to a variation in γ is the certainty equivalent, which will be shown during the discussion.

The first part of table 3 shows that the parameters tend to loose significance as the relative risk aversion increases. All the out of sample parameters[11] for $\gamma = 100$ are not significant. Although the average returns for the $\gamma = 100$ portfolio is higher than the VW returns, the investor with this relative risk aversion is indifferent between the out of sample parametric portfolio and the VW, the first has a CE of -96.7% and the second -96.5%.

The out of sample portfolio for $\gamma = 100$ also has very small positions when compared to all other parametric portfolios. While all portfolios have more than 35% of short positions, the $\gamma = 100$ out of sample portfolio has only 17.4% of its stocks short. Its average absolute weight, average maximum and minimum weights and its turnover are also smaller.

There is a very small difference between the returns and Sharpe ratios when we compare the $\gamma = 2$ and the $\gamma = 5$ portfolios. However, their certainty equivalent is different since the first accepts more risk. The CE for the $\gamma = 2$ VW portfolio is 9.2%. The $\gamma = 100$ has smaller returns and smaller risk than the other two, its in sample CE is -85.9%. Although the $\gamma = 100$ CE is negative, the investor dislikes the VW even more, it has a CE -96.5%. In the out of sample case the investor is indifferent between the VW and the $\gamma = 100$ portfolios, as pointed out before.

[11]The out of sample parameters are the mean of each sub sample parameters that were used to estimate each out of sample portfolio.

Table 4
Varying risk aversion – long only

Variable	VW	In Sample PPP $\gamma = 2$	Out of Sample PPP $\gamma = 2$	In Sample PPP $\gamma = 5$	Out of Sample PPP $\gamma = 5$	In Sample PPP $\gamma = 100$	Out of Sample PPP $\gamma = 100$
θ_{btm}	-	3.209	6.454	2.717	3.570	0.591	0.090
Std.	-	(1.873)	(4.116)	(1.275)	(2.235)	(0.414)	(0.276)
Per.	-	(0.001)	(0.015)	(0.002)	(0.085)	(0.172)	(0.415)
θ_{me}	-	-1.727	-5.499	-1.624	-2.551	-1.663	-1.412
Std.	-	(1.873)	(5.448)	(1.529)	(2.553)	(1.789)	(1.034)
Per.	-	(0.110)	(0.033)	(0.173)	(0.064)	(0.297)	(0.174)
θ_{mom}	-	4.428	7.018	3.896	3.865	2.298	2.234
Std.	-	(4.265)	(4.578)	(2.500)	(3.021)	(2.169)	(1.322)
Per.	-	(0.000)	(0.038)	(0.002)	(0.045)	(0.023)	(0.227)
$\lvert w_i \rvert \times 100$	0.786	0.786	0.689	0.786	0.689	0.786	0.689
$\max\ w_i \times 100$	1.102	8.977	9.487	8.642	9.055	5.554	2.634
$\min\ w_i \times 100$	0.485	0	0	0	0	0	0
$\sum w_i I(w_i < 0)$	0	0	0	0	0	0	0
$\sum I(w_i \leq 0)/N_t$	0	0.503	0.478	0.490	0.475	0.390	0.211
$\sum \lvert w_{i,t} - w_{i,t-1} \rvert$	0.009	0.288	0.267	0.286	0.277	0.272	0.180
CE	-0.002	0.224	0.156	0.122	0.067	-0.919	-0.967
$\bar{r}$	0.149	0.288	0.217	0.286	0.225	0.260	0.153
$\sigma(r)$	0.219	0.223	0.224	0.222	0.222	0.214	0.218
SR	0.679	1.291	0.970	1.289	1.016	1.216	0.704
$\text{SR}_{rf=0.1}$	0.222	0.843	0.523	0.839	0.565	0.748	0.225

Table 5 shows the long only portfolio under different relative risk aversion. An interesting fact regards the $\gamma = 100$ portfolio, it has the same behavior of the base case in the out of sample tests. The certainty equivalent is basically the same for the $\gamma = 100$ portfolios and the value weighted portfolio for the risk aversion of 100.

4.5 Risk free asset

In this section we test the parametric portfolio with a risk free asset that yields 0.79% per month, which is approximately 10% per year. The inclusion of a risk free asset gives the investor the option of investing some of his wealth in the risky portfolio and the rest of it in r_f. Table 5 summarizes the results, the proportion of risky assets in the portfolio decreases with the relative risk aversion and the certainty equivalent increases with the inclusion of the risk free asset. The $\gamma = 2$ investor is 34.7% short in r_f, his certainty equivalent and the portfolio average returns increases considerably, the first is 42% against 34% without r_f and the second is 57% against 42%. The $\gamma = 5$ investor has 9.1% of his wealth invested in r_f and 80.9% invested in the risky portfolio. His certainty equivalent also increases with the inclusion of the risk free asset, from 23.7% to 24.4%. The average return on the other hand decreases from 42.5% to 40.2%. Since we are optimizing the utility function it is natural to have lower returns when long positions on r_f are taken. However, the standard deviation decreases from 23.6% to 21.8% and the Sharpe ratio increases. Last, the $\gamma = 100$ investor has 98.4% of his wealth in the risk free asset, the parameters of the risky portfolio are also very small. This kind of extremely risk averse investor is very sensitive to losses, which makes the risk free asset very attractive to him, his certainty equivalent increases from -85.9% to 10%, the last is the return of the risk free asset. Since the portfolio has basically no variation (standard deviation of 0.3%) it is natural to have the certainty equivalent equal the returns, which is also the risk free rate. Despite his extreme risk aversion, the investor still chooses to keep some of his wealth in risky assets, but due to the small parameters the risky portfolio is very similar to the value weighted portfolio (base portfolio).

Table 5
Risk free asset under different relative risk aversion

Variable	VW	In Sample PPP $\gamma = 2$	In Sample PPP $\gamma = 5$	In Sample PPP $\gamma = 100$
θ_{btm}	-	1.396	1.382	0.057
Std.	-	(7.228)	(1.114)	(1.021)
Per.	-	(0.096)	(0.113)	(0.073)
θ_{me}	-	-0.419	-0.338	-0.014
Std.	-	(0.954)	(0.747)	(0.420)
Per.	-	(0.296)	(0.357)	(9.281)
θ_{mom}	-	2.406	2.442	0.127
Std.	-	(3.895)	(1.612)	(1.035)
Per.	-	(0.088)	(0.111)	(0.030)
θ_{rf}	-	-0.347	0.091	0.984
Std.	-	(5.789)	(0.537)	(2.298)
Per.	-	(0.148)	(0.362)	(0.002)
$\lvert w_i \rvert \times 100$	0.786	1.761	1.768	0.786
$\max\ w_i \times 100$	1.102	11.636	11.565	1.217
$\min\ w_i \times 100$	0.485	-3.570	-3.616	0.392
$\sum w_i I(w_i < 0)$	0	-0.610	-0.614	0
$\sum I(w_i \leq 0)/N_t$	0	0.410	0.409	0
$\sum \lvert w_{i,t} - w_{i,t-1} \rvert$	0.009	0.695	0.704	0.042
CE	-0.002	0.421	0.244	0.100
$\bar{r}$	0.154	0.573	0.402	0.100
$\sigma(r)$	0.219	0.323	0.218	0.003
SR	0.701	1.771	1.842	29.590
$\text{SR}_{rf=0.1}$	0.245	1.462	1.384	0

4.6 Transaction costs

Since the parametric optimization is based on a portfolio police such as equation (2), a very high turnover is possible because the portfolio weights may change every period accordingly to variations in the characteristics $x_{i,t}$. So far, all the results are for optimizations with no transaction costs, however this assumption is even more unrealistic in the parametric case.

This section shows the results with transaction costs, we assumed that they are a function of the market equity, following evidence in the literature that smaller companies have higher transaction costs (see Hasbrouck (2009)). Another evidence points out that transaction costs have a decreasing trend, but due to our small period sample we did not use it in our function. We used the same function for the transaction costs as Brandt *et al.* (2009), which is $c_{i,t} = 0.006 - 0.0025me_{i,t}$, where $me_{i,t}$ is rescaled to be between 0 and 1. Thus, the transactions costs must be between 0.35% and 0.6% of the company's turnover.

Table 6
Transaction costs

Variable	VW	In Sample Without TC	Out of Sample Whithout TC	In Sample With TC	Out of Sample With TC
θ_{btm}	-	1.199	1.586	1.631	1.988
Std.	-	(0.315)	(0.603)	(0.737)	(0.816)
Per.	-	(0.006)	(0.037)	(0.019)	(0.139)
θ_{me}	-	-0.162	-1.319	-0.386	-1.500
Std.	-	(0.592)	(0.969)	(0.861)	(1.255)
Per.	-	(0.367)	(0.094)	(0.258)	(0.292)
θ_{mom}	-	2.492	1.696	2.121	1.154
Std.	-	(0.358)	(0.842)	(0.800)	(0.976)
Per.	-	(0.000)	(0.095)	(0.028)	(0.126)
$\lvert w_i \rvert \times 100$	0.786	1.764	1.585	1.656	1.551
$\max\ w_i \times 100$	1.102	10.701	11.432	12.337	13.645
$\min\ w_i \times 100$	0.485	-3.657	-3.287	-3.177	-3.317
$\sum w_i I(w_i < 0)$	0	-0.599	-601	-0.542	-0.581
$\sum I(w_i \leq 0)/N_t$	0	0.398	0.379	0.407	0.372
$\sum \lvert w_{i,t} - w_{i,t-1} \rvert$	0.009	0.713	0.547	0.622	0.497
CE	-0.002	0.237	0.182	0.190	0.114
$\bar{r}$	0.154	0.425	0.359	0.372	0.296
$\sigma(r)$	0.219	0.236	0.227	0.236	0.237
SR	0.701	1.799	1.584	1.578	1.251
$SR_{rf=0.1}$	0.245	1.376	1.143	1.154	0.829

Table 6 shows the results. When transaction costs are included, the turnover reduces from 71.3% to 62.2% (14% decrease) in sample and from 54.7% to 49.7% (9% decrease) out of sample. The average return decreased both in and out of sample and the standard deviation had little change, which made the Sharpe ratios and also the certainty equivalent smaller. However, the parametric portfolio still much superior to the value and the equal weighted portfolios. Although there is a significant certainty equivalent loss, the results remain good because investor minimizes the difference between the portfolios by adjusting the parameters to accommodate the transaction costs. Comparing in and out of sample results, the transaction costs case has a bigger loss than the base case due to the fact that the investor adjust the parameters with the in sample turnover, which may have a very different structure out of sample. This also have an impact on the smaller proportional decrease in the out of sample turnover.

4.7 The Markowitz case

In this section we compare the parametric optimization with the Markowitz portfolios. The in sample Markowitz portfolio was estimated with a 24 months window to allow listing and delisting of stocks, i.e. we took the first 24 months of the sample and optimized them saving the weights, then we took the next 24 months and so on until the end of the sample. The target returns in the Markowitz optimization are the average returns of the parametric portfolio for each 24 months window. We used the estimated weights to compute the results regarding the distribution of weights and the portfolio returns. Since the portfolio is rebalanced every 24 months, the turnover is null in most of the sample and there is only a small adjustment in the weights at the first month of each 24 months sub sample, for this reason we choose for simplicity to ignore transaction costs. Weights are restricted to an absolute value of 20% and short sales are allowed. The out of sample Markowitz portfolio was calculated using a rolling window of 24 months to estimate the weights to be used in the next 12 months, by doing this we make sure the model depends only on information that is available before the beginning of the out of sample period.

Table 7
Markowitz and parametric portfolio optimization

Variable	VW	In Sample Without TC	Out of Sample Whithout TC	In Sample With TC	Out of Sample With TC	In Sample Markowitz	Out of Sample Markowitz
θ_{btm}	-	1.199	1.586	1.631	1.988	-	-
Std.	-	(0.315)	(0.603)	(0.737)	(0.816)	-	-
Per.	-	(0.006)	(0.037)	(0.019)	(0.139)	-	-
θ_{me}	-	-0.162	-1.319	-0.386	-1.500	-	-
Std.	-	(0.592)	(0.969)	(0.861)	(1.255)	-	-
Per.	-	(0.367)	(0.094)	(0.258)	(0.292)	-	-
θ_{mom}	-	2.492	1.696	2.121	1.154	-	-
Std.	-	(0.358)	(0.842)	(0.800)	(0.976)	-	-
Per.	-	(0.000)	(0.095)	(0.028)	(0.126)	-	-
$\lvert w_i \rvert \times 100$	0.786	1.764	1.585	1.656	1.551	2.331	2.264
$\max\ w_i \times 100$	1.102	10.701	11.432	12.337	13.645	7.704	7.947
$\min\ w_i \times 100$	0.485	-3.657	-3.287	-3.177	-3.317	-6.678	-6.352
$\sum w_i I(w_i < 0)$	0	-0.599	-601	-0.542	-0.581	-0.880	-0.868
$\sum I(w_i \leq 0)/N_t$	0	0.398	0.379	0.407	0.372	0.333	0.344
$\sum \lvert w_{i,t} - w_{i,t-1} \rvert$	0.009	0.713	0.547	0.622	0.497	-	-
CE	-0.002	0.237	0.182	0.190	0.114	0.323	0.080
$\bar{r}$	0.154	0.425	0.359	0.372	0.296	0.351	0.202
$\sigma(r)$	0.219	0.236	0.227	0.236	0.237	0.091	0.198
SR	0.701	1.799	1.584	1.578	1.251	3.854	1.021
$SR_{rf=0.1}$	0.245	1.376	1.143	1.154	0.829	2.761	0.516

The results are in table 7. The parametric portfolio with and without transaction costs is also in the table for better comparison. The average absolute weight for the Markowitz portfolio is bigger than the parametric case, however it takes less extreme positions[12]. The parametric portfolio takes long positions bigger than the short ones, but in the Markowitz case the long and short position are more or less the same. Last on the distribution of weights, the average sum of negative positions is also higher in the Markowitz portfolio. Regarding the returns, the first thing worth mentioning is that the Markowitz portfolio had very big certainty equivalent (32%) and Sharpe ratios (3.85 and 2.76) in sample, they were even bigger than the ones in parametric case without transaction costs, specially because of the very small standard deviation (9.1%) of the Markowitz portfolio. However, the Markowitz portfolio had a poor performance out of sample, its out of sample certainty equivalent decreased 76.7% and the average return decreased from 35.1% to 20.2%. The parametric portfolio was considerably superior than the Markowitz case out of sample, even when transaction costs are included, its certainty equivalent was 11.4% against 8% of the Markowitz case.

[12]The Markowitz portfolio average maximum weight is considerably smaller than the parametric one, they are, respectively 7.7% and 11.4%

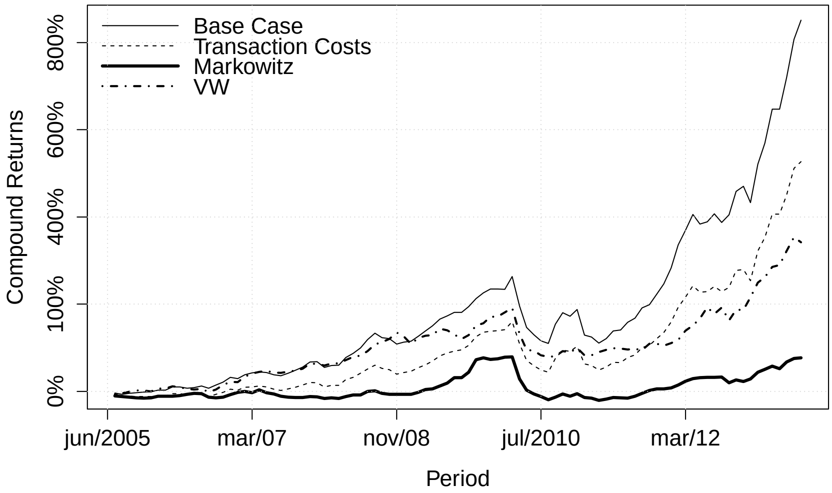

Figure 5
Out of Sample Compound Monthly Returns

Figure 5 helps us to understand some of the dynamics of both models, it shows the compound out of sample returns of the base parametric case, the parametric case with trasaction costs, the Markowitz case and the value weighted portfolio. An interesting fact is that the Markowitz portfolio was better than the parametric with transaction costs until the end of 2011, but after that the parametric case accumulated twice as much than the Markowitz. In the end the first had a compound return of 334% and the second 527%. We can see some tendency similarities between all the time series, which is natural since they are in the same economy and are made of the same stocks. In fact, the two parametric series (with and without transaction costs) are nearly the same differing only on a drift, this is not exactly true since their parameters are not the same.

5. Final Remarks

This paper aimed to apply the parametric optimization techniques in the Brazilian market. We presented the method and its extensions, such as weighs constraints and transaction costs and the way we find to be the best to estimate the optimal parameters. We parametrized the portfolio weights using three characteristics: the book-to-market ratio, the market equity and the one year momentum.

Our empirical application had 316 companies listed in the Brazilian stock exchange BOVESPA in the period that goes from 2001 to 2013. We performed several tests including linear weight constraints, long only constraints, transaction costs and a comparison with a Markowitz based portfolio.

The parametric optimization has shown a very good out of sample performance even when transaction costs were included. It was consistently superior to the VW and EW portfolios and although the Markowitz case had a better performance in sample, it performed poorly out of sample. The out of sample parametric portfolio was superior to all other portfolios, even the out of sample Markowitz.

Moreover, we conclude that the parametric optimization is a very good option for quantitative funds. Besides, the consistent elevated risk adjusted returns provided by the parametric optimization are evidence that the Brazilian market still inefficient.

Finally, It should be noted any excess gain related to the parametric strategy on a given set of cross-sectionally variant characteristics disappears after multiple agents start exploiting the same approach and same data.

References

Brandt, Michael W. 2010. Portfolio choice problems. *Pages 269–336 of:* Ait-Sahalia, Y., & Hansen, L. P. (eds), *Handbook of Financial Econometrics, Volume 1: Tools and Techniques*.

Brandt, Michael W., Santa-Clara, Pedro, & Valkanov, Rossen. 2009. Parametric portfolio policies: Exploiting characteristics in the cross-section of equity returns. *Review of Financial Studies*, **22**, 3411–3447.

Davison, Anthony C. 1997. *Bootstrap methods and their application*. Vol. 1. Cambridge university press.

Davison, Anthony C., Hinkley, David V., & Schechtman, Edna. 1986. Efficient bootstrap simulation. *Biometrika*, **73**, 555–566.

Efron, Bradley, & Tibshirani, Robert J. 1994. *An introduction to the bootstrap*. Vol. 57. CRC press.

Fama, Eugene F., & French, Kenneth R. 1996. Multifactor explanations of asset pricing anomalies. *The journal of finance*, **51**, 55–84.

Hansen, Lars Peter. 1982. Large sample properties of generalized method of moments estimators. *Econometrica*, **50**, 1029–1054.

Hasbrouck, Joel. 2009. Trading costs and returns for US equities: Estimating effective costs from daily data. *The Journal of Finance*, **64**, 1445–1477.

Ljungqvist, Lars, & Sargent, Thomas J. 2004. *Recursive macroeconomic theory*. MIT press.

Markowitz, Harry. 1952. Portfolio selection. *The journal of finance*, **7**, 77–91.

Mossin, Jan. 1968. Optimal multiperiod portfolio policies. *The Journal of Business*, **41**, 215–229.