Evaluating the violation probability of data-driven robust constraints of financial risk on portfolio optimization experiments

Tomás Gutierrez*
Bruno Reis*
Eduardo Bolonhez*
Davi Valladão*
Fernando Luiz Cyrino Oliveira*

* Departamento de Engenharia Industrial, Pontifícia Universidade Católica do Rio de Janeiro

ABSTRACT

This work proposes to evaluate two investment strategies that can be solved under a robust linear optimization approach to deal with the uncertainty. It is of desire to quantify how the out-of-sample result violates the expected results and understand its relation to the parameters specification. These models constitute a different approach to financial risk management, and is of academic and applied interest. The historical set of values is known, and in order to evaluate the performance more appropriately maintaining the assets intrinsic relationship, a bootstrap method is applied to generate several samples.

KEYWORDS

Assets, Bootstrap, Optimization, Robust, Worst Case, Risk Management.

1. INTRODUCTION

Robust optimization is a method of solving optimization problems that contain uncertainties related to the parameters. Instead of assuming that there are known probability distributions of the problem, like stochastic optimization solving methods does, it takes on that the uncertainties are within a previously defined set. The conservatism of each method is reflected by the treatment given in each set, mainly regarding the number of elements that can deviate from the nominal case.

Uncertainty sets for robust optimization were first introduced by (Soyster 1973), that proposed a linear optimization problem where its optimal value is viable in all possible scenarios inside a convex set. This model, due to its conservatism, traded risker optimal solutions (in a practical sense) for solutions in all possible scenarios. Years later, (A Ben-Tal and Nemirovski 1998; Aharon Ben-Tal and Nemirovski 1999; Ghaoui and Lebret 1997; Ghaoui, Oustry, and Lebret 1998) researched deeply into linear optimization and general convex optimization. (Aharon Ben-Tal and Nemirovski 1999) introduced a less conservative model, considering an ellipsoidal uncertainty set problem with a convex robust counter-part that could be treated computationally and using second-order conic programming. (Dimitris Bertsimas and Sim, 2004), presented a new methodology to control the conservatism of such problems, keeping the advantages of the linear formula of (Soyster, 1973). They also presented the concept of uncertainty "budget", that controls the level of conservatism of the solution, by limiting the number of parameters that can deviate from the nominal case concomitantly.

Robust portfolios can be modeled to respond more steadily to variations of average returns and covariance. Average returns are used to solve the optimization problems, but they can diverge from the real returns of the assets due to the uncertainties and risks associated with investments. The riskier the investment, greater the uncertainty of its return. The uncertainty sets describe a geometrical structure surrounding the values of the future expected returns (Kim, Kim, and Fabozzi, 2013).

Due to the practicality of the method, many have studied the application of such to optimize assets portfolio. (Lobo and Boyd, 2000) provide an introduction to portfolio optimization formulations from the perspective of robustness, listing possible sets of uncertainties that are convex and treatable to model returns on assets, while (Halldórsson and Tütüncü, 2003) introduced a robust formulation for the mean-variance model, in which the optimal solution considers the worst case within the set of values for the returns and the covariance matrix. (Fernandes et al., 2016) proposed a new robust data-driven adaptive portfolio model. This model uses data-driven sets of polyhedral uncertainties to build loss constraints in a rolling horizon scheme.

(Bjerring et al, 2017) conducted tests to evaluate the best selection portfolio for a robust optimization, showing that a more careful selection of the assets data can improve the model results, and (Lotfi et al, 2017) evaluate the level of conservatism of robust portfolios, aiming for more riskier approaches.

By dealing with uncertainty, robust optimization arise as an alternative to risk management. Vastly spread in the literature, risk measures seek to address the "risk" problems of undesirable results in a direct and pragmatic way, whereas the distribution of financial values would be directly monitored and "controlled" through these measures and exposure limit values (which can be stipulated a priori by agents). As such measures, are commonly used VaR (Value-at-Risk, the least financial deposit necessary so that the probability of a debt, or negative values for payments flow, is not above a significance ($1-\alpha$) value), and CVaR (Conditional Value-at-Risk, or the metric that measures the financial flow of values lower than the quantile of ($1-\alpha$)% of the distribution of investment probabilities). Those are measures that are used commonly in stochastic programming. (Sarykalin et al, 2008) give a detailed insight in such measures, and provide direct application of such methods in risk management and optimization.

The simulation method here is used to generate random samples that will be used to evaluate an empirical probability. The advantage of this method in the context here introduced is that it maintains the structural correlation between the assets. Thus, we do not impose any prediction method, and use only the available data to investigate the behavior of the models proposed.

The main contributions of this work are the study of a probabilistic behavior of a method already proposed in the literature (Fernandes, 2016) and of an extended method based on (Bertsimas, 2001). This way, we provide an insightful evaluation of risk management, from a robust perspective.

This paper is structured as follows: Section 2 shows the methodology used to generate the data and simulate the scenarios of the robust optimization, also showing the equations used to solve the problem, section 3 contains the results and analysis of the simulations and finally section 4 concludes this work, also providing insights to future studies in the area.

2. METODOLOGY

In this study, two different models are evaluated, with similar approaches to deal with risk. Under a financial perspective, risk is frequently associated with possible loss of a portfolio. To evaluate the exposure to possible losses, a usual approach is to estimate a distribution function and its parameters from the data available and use stochastic optimization methods combined with some risk measure, such as variance or Conditional Value at Risk (Sarykalin et al, 2008). In this work, we propose a different approach to risk management, and we evaluate its performance and sensitivity to the parameters variation.

In order to better understand its dynamical properties, our scope is as follows: at a given time t, we use the last B days to, under a certain criterion, allocate our wealth W_t to the available N assets. To each asset, it will be associated a percentage x_i, that must be a positive value due to short sale constraints, that affect the optimal portfolio. An implied restriction is that the sum of those percentages must add 1, representing the whole portfolio. After we allocate our wealth, we evaluate the out of sample performance, by calculating the actual return that the portfolio established at t would have at time $t + 1$.

Despite the returns achieved by the portfolio optimization method itself, we are interested in understanding the two risk management approaches. To do so, we calculate at each step whether the constraints were violated in an out-of-sample analysis. After applying the method to the whole data set, we calculate the ratio between how many times the constraints were violated and how many times (steps) the method was applied. We obtain different samples rather than just the historical data by simulating scenarios using a resampling method, preserving assets' interdependencies. We also vary the dataset to which the method is applied, containing different assets and structures to avoid a sample bias.

The following subsections discuss our approach to generate the datasets, and the two models evaluated.

2.1 Simulation approach

As briefly mentioned, we generate different samples using past data, following an bootstrap method. The idea here is that this way we keep the correlation structure between the assets, while considering several possible sequences of scenarios. One could argue that there is no prediction model used, but as seen in (Black, 1993), returns predictability are very low, if not zero. That way, we use real data, and resample its values several times, as described above, to generate different samples with similar structures, as desired.

In numerical terms, we have a full year of returns (252 days). To avoid sample bias, we generate different samples, each with 252 days, by a bootstrap method. Also, we repeat this procedure in different datasets, with different assets within. The main purpose is to provide different conditions onto which the model is applied, in order to actually evaluate an empirical performance without restraining to a specific condition.

2.2 First model (Fernandes et al., 2016)

The first method we present was originally developed by (Fernandes et al., 2016). The intuition is that at a specific time t, we use the last B trading days to allocate our portfolio, as previously mentioned. The reasoning here is to create a portfolio that would have performed better than a minimum return constraint value R in all of the last B days that simultaneously maximizes the expected value for the following day. Defining R as the maximum loss tolerated (or even as minimum gain desired), we develop an allocation that would have achieved this goal in all of the previous B days. The intuition is that we assume that our future return in $t + 1$ will most likely fall within a convex hull generated by the past returns. In other words, we assume that with high probability the future return will be some linear combination of the past B returns. By ensuring a minimum return of R in the vertices of the generated convex hull (the past B days), we are obtaining a portfolio that maximizes the expected value with the restriction that the worst scenario, considered the same as the worst case when applied to the last B days, will have a return of at least R.

In this method, we can create different risk profiles by defining different pairs (R, B) of return required and past days used. The lower the minimum return required, the easier it is to respect the worst-case constraint while selecting assets with higher expected values. With regard to the last B days, the longer the period the more conservative we are. The logic is that a bigger window most likely contains lower returns than a smaller one, so the allocation has to be robust to a higher number of parameters.

R: *required rate of return*; *constraint value*;
B: *number of past days used at each point in time*;
$r^*_{i,t+1}$: *expected return for* $t+1$ *calculated at* t
$r_{i,t}$: *realized return of asset* i *at time* t
$x_{i,t}$: *percentual allocation in asset* i *at time* t

For a specific time t, we can write our model as:

$$Max_x \sum_{i=1}^{N} x_{i,t}\, r^*_{i,t+1} \tag{1a}$$

$$s.t.$$

$$\sum_{i=1}^{N} x_{i,t}\, r_{i,j} \geq R, \qquad \forall j = t-B, \dots, t \tag{1b}$$

$$\sum_{i=1}^{N} x_{i,t} = 1$$

$$x_{i,t} \geq 0, \qquad \forall i = 1, \dots, N$$

(1c)

That is, we aim to maximize the expected value of the portfolio at time $t+1$, with the constraint that this portfolio would have a return of at least R if applied to the returns of the last B trading days. In Figure 1, we see that we are actually guaranteeing that, if the return in $t+1$ falls within the convex hull generated by the past data, we will obtain a return of at least R, since the solution of this linear optimization problem is at its vertices, given by the actual returns evaluated in the past.

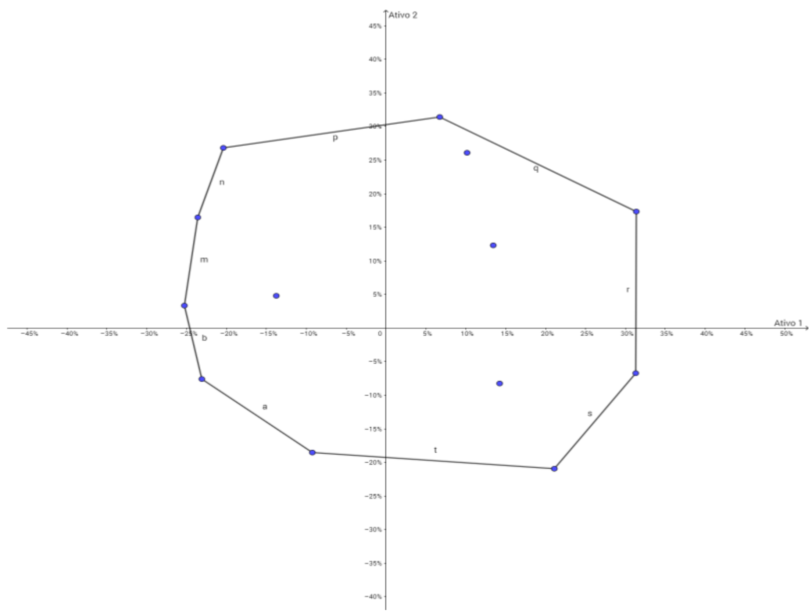

Figure 1: Convex hull formed by the last B trading day's returns

We can also see that the past returns have a big influence on the hull itself. Consider two scenarios, one with high volatility (crisis period) and one more controlled (non-crisis period). As we can observe in Figure 2, the period denoted "*crisis period*" generates a bigger figure, covering worst cases than a "*regular period* ". It is important to notice that the axis x and y represent two risky assets, with the z axis (not shown here) representing the risk-free asset (with less expected return). In crisis periods, a bigger allocation will be in the risk-free asset since the worst case during this period usually is more damaging. That way, during crisis period the model acts in a more conservative way since a worst case during a crisis usually is more harmful than during a regular period.

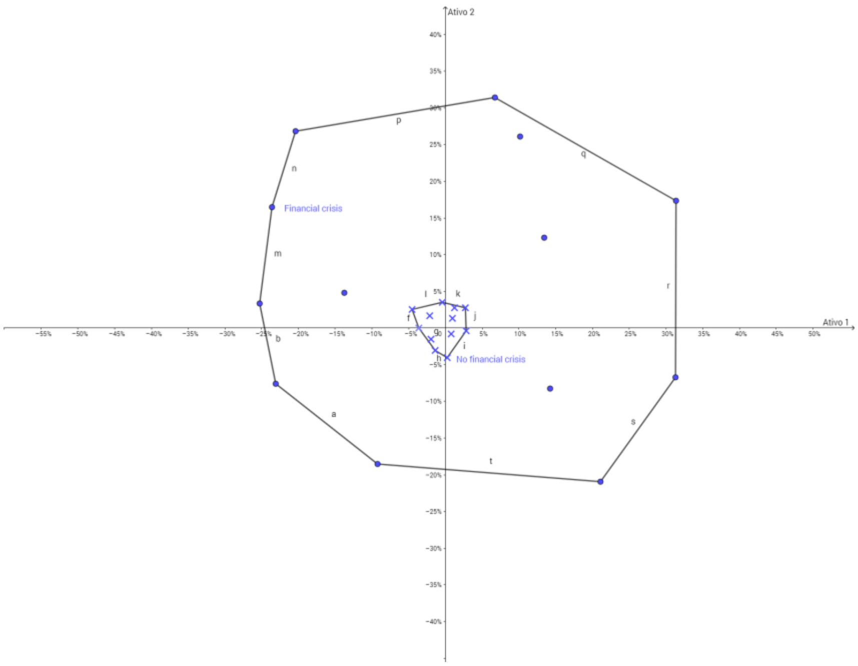

Figure 2: Convex hull and returns distributed in financial and non-financial crisis

We now turn our focus to evaluating how efficient is to define our risk management that way. By varying the values in our pair that defines our risk exposure, (R, B), we calculate how many times the actual return in $t + 1$ was less than the required rate of return R comparing to the number of repetitions. Using the bootstrap methodology presented, we conduct the experiment several times, and assume that the obtained value is a proxy for a probability of violation for each specification (R, B), as it is shown in the section of numerical experiments.

2.3 Second model (Bertsimas, 2004)

The second method presented is derived from the method derived by (Bertsimas, 2004), using a similar reasoning to the previous method (Section 2.2). However, a key difference is how we evaluate the worst-case scenario. Previously, the worst-case was exactly the same as the worst-day from the last B trading days, i.e., from the historical data. Now, we allow the possibility of a combination of returns from each asset, including the worst return from each one, despite the absence directly in the data.

A simple numerical example can be very helpful in understanding the idea. Suppose we use two days as historical data, for three assets. The returns are shown in table 1.

Table 1: 2 day/2 asset random returns

	Asset 1	Asset 2	Asset 3	Risk-Free
Day 1	10%	-8%	-2%	1%
Day 2	-15%	6%	4%	1%

Using the model presented in section 2.2, the portfolio allocation would have to be robust to the 2 days used. In other words, the vector allocation x_t would be such that, for a given minimum return R, the following constraint are not violated:

$$10\% * w_1 + (-8\%) * w_2 + (-2\%) * w_3 + 1\% * w_4 \geq R \tag{2a}$$

$$(-15\%) * w_1 + 6\% * w_2 + 4\% * w_3 + 1\% * w_4 \geq R \tag{2b}$$

Now, we allow the model to consider all possible combinations of returns. Using 2 days and 3 assets, we now have 2^3 constraint to consider. Due to the exponential growth of the number of constraints (for T days and N assets we would have to evaluate T^Nconstraints, which quickly becomes unfeasible due to computational time), we use a technique from robust optimization to solve the problem. The idea here is that in order to respect all constraints imposed, it would be better to respect the more restrictive one, since all others would be contained from there. In other words, in our context, if we identify the worst possible day from each asset, we will be condensing all restrictions into this one, since all of the others are less restrictive than the worst one. In our numerical example, this constraint would be:

$$(-15\%) * w_1 + (-8\%) * w_2 + (-2\%) * w_3 + 1\% * w_4 \geq R \tag{2c}$$

Note that this constraint is not directly referenced to a particular day, but to a combination of all possible cases considered. In mathematical terms, we are now considering as a constraint, the following:

$$\min r^T . w \geq R \tag{2d}$$

Where r^T is the transposed vector of returns, w represents the portfolio and R the minimum return required.

In addition, a feature derived from (Bertsimas, 2001) is considered. One could argue that it would result in a solution too conservative if it's allowed to all assets to go to the worst case, with too much allocation to the risk-free asset. To add more judgment to the model, we expand our idea by determining, instead, how many assets can be considered as a worst scenario by a parameter Γ. The value of Γ, which can be even a non-integer number, but has to be positive, determines how many assets we are allowing to go simultaneously to the worst case. The remaining ones are considered to stay on a mean value, calculated as the mean of the last B days.

Still considering our simple numerical example, assume we choose $\Gamma = 2$. That way, we have to consider the following constraint, where μ_i is the mean value for asset i (since the risk-free asset does not variate, we don't have to consider a mean value):

$$(-15\%) * w_1 + (-8\%) * w_2 + \mu_3\% * w_3 + 1\% * w_4 \geq R \tag{2e}$$

$$(-15\%) * w_1 + \mu_2\% * w_2 + (-2\%) * w_3 + 1\% * w_4 \geq R \tag{2f}$$

$$\mu_1\% * w_1 + (-8\%) * w_2 + (-2\%) * w_3 + 1\% * w_4 \geq R \tag{2g}$$

To give a graphical perspective, we show in Figure 3 how the region of considered values would look like in a problem with 2 risky assets and $\Gamma = 2$. The z axis, not shown here, represents the risk-free asset.

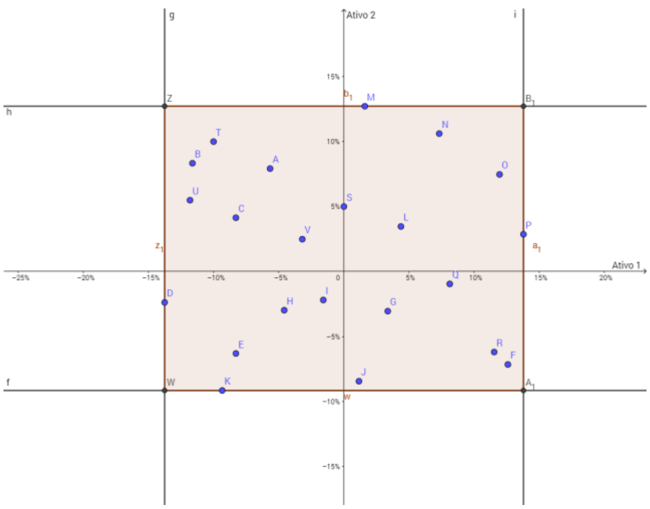

Figure 3: Region considered in a problem a problem with 2 risky assets and $\Gamma = 2$

Note that the points A_1, B_1, W and Z do not represent actual returns observed, but a combination of worst/best values, as desired. It is important to mention that even though there is a combination of best scenarios, it is not relevant to the solution since we are robust to all scenarios, and the more restrictive is the worst (in the image, it's the point W.

Thus, our model is primarily written, to a specific point in time t:

$$\max_{w} \mu_t^T . w \tag{3a}$$

$s.a$

$$\mathbf{1}^T = 1 \tag{3b}$$

$$\min_{r} r^T . w \geq R \tag{3c}$$

$s.a$

$$\sum_i Z_i \leq \Gamma \qquad :\lambda \tag{3d}$$

$$r_i \leq \mu_i + Z_i.\delta_i\,, \quad \forall i \qquad :\pi_1 \tag{3e}$$

$$r_i \geq \mu_i - Z_i.\delta_i\,, \quad \forall i \qquad :\pi_2 \tag{3f}$$

$$0 \leq Z_i \leq 1, \quad \forall i \qquad :\theta \tag{3g}$$

Where μ_t represent the vector of expected values (mean values) calculates in t using the last B days. The variables λ, π_1, π_2 and θ are used to solve the robust problem, and each one is linked to the respective constraint. A good reference for the method of solving robust optimization formulations, as is the case, can be found in (Soyster, 1973). The new formulation, called robust-counter-part, is given by:

$$max_w \quad \mu^T.w \tag{3h}$$

$$\lambda, \pi_1, \pi_2, \theta \geq 0 \tag{3i}$$

$$s.a \tag{3j}$$

$$\mathbf{1}^T = 1 \tag{3k}$$

$$-\lambda.\Gamma + \sum_i \mu_i(\pi_{2i} - \pi_{1i}) - \sum \theta_i \geq R \tag{3l}$$

$$\pi_{2i} - \pi_{1i} - w_i = 0 \tag{3m}$$

$$\lambda - \pi_{1i}\,\delta_i - \pi_{2i}\,\delta_i + \theta_i \geq 0 \tag{3n}$$

Now we have a linear programming model to be solved, much faster than the original formulation. The key feature is that we eliminated the constraint with a min argument, replaced by linear constraints.

With the addition of the flexibility of assets to be considered at the worst scenario, we now have a new way to manage the risk. Instead of the pair of values used in the previous model, we now have to choose a tuple of values, (R, B, Γ), namely the minimum return, the number of past trading days considered and the parameter of flexibility.

The influence on the probability violation of the variation of these parameters are the focus on the section of numerical experiments, as is the case with the other model.

An interest aspect to be considered is that when we allow all assets to go to the worst case, we have a direct relationship with the model presented by (Fernandes et al. 2016). While we now consider all combinations, the previous model considers only the real data. By doing so, we are actually creating a hull outside the convex hull generated in (Fernandes et al. 2016). We can see this relation explicitly when we consider 2 risky assets, as is shown in the figure 4. Given this configuration, we can clearly see that our approach gives more conservative solutions.

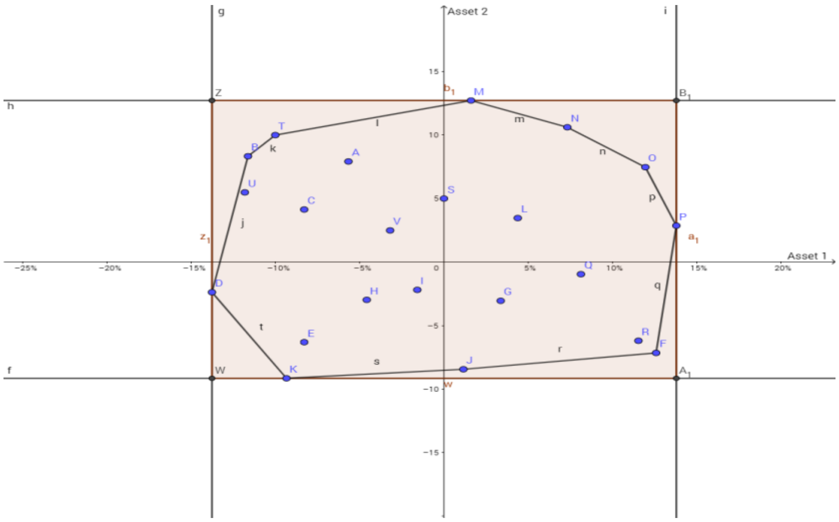

Figure 4: Comparison between the two models' uncertainty set figure

3. NUMERICAL EXPERIMENTS

We first evaluated the model described in section 2.2. Our specification was to use a full year of trading days (252 observations), evaluating to the whole set whether the return on $t + 1$ fell inside the convex hull specified at t, using past data. We repeated this procedure 10 times generating different samples using a bootstrap simulation method, and calculated the mean empirical probability of each parameters combination (B, R). The results here refer to a dataset that is composed by different assets. In the Appendix section, we show results for a different database, to imply that there is no data bias affecting the analysis. In figure 5, we observe the parameter influence on the probability itself. As expected, the more restrictive the minimum return requirement, the higher the probability violation. Also, the more days that were used to construct the convex set, the smaller the probability violation, since more days were evaluated as constraints. It is interesting to notice that after 50 days used as past data (B), the sensitivity to this parameter almost disappears. Also, when the time window is smaller, the sensitivity to the return constraint is higher.

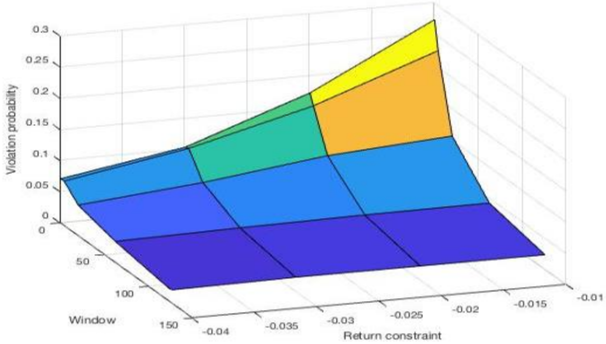

Figure 5: Violation probability x specification parameter of return constraint and past data.

After that, we evaluated the model presented in section 2.3 regarding its parameters. Since we now have a tuple of parameters (B, R, Γ), we have to present 3 different graphics, each representing the joint relationship effect on the probability violation empirically calculated. We repeated the bootstrap approach, and displayed the mean results obtained. The results are showed below.

First, we see in figure 6 that the return constraint is responsible for the highest impact on the probability violation when combined with different time window constraints. Also, when compared to the constraint of worst cases Γ, we see the higher influence of the return constraint (Figure 7). When comparing the effect of the constraint of worst cases Γ with the time constraint, we notice a higher influence of the last (Figure 8). Interestingly, the effect of the variation of Γ is small when compared to the others, even though different portfolio allocations are built. Further studies are necessary to understand more precisely why this happens, but a first investigation suggests that there is not enough gain related to the uncertainty modelling when considering only a few assets to go to the worst case while the others remain in the historical mean.

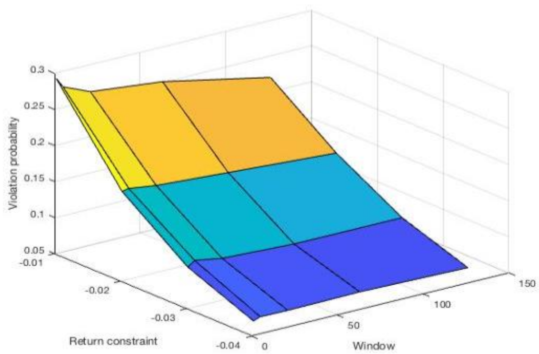

Figure 6: Violation probability for each parameter combination of time and return constraint.

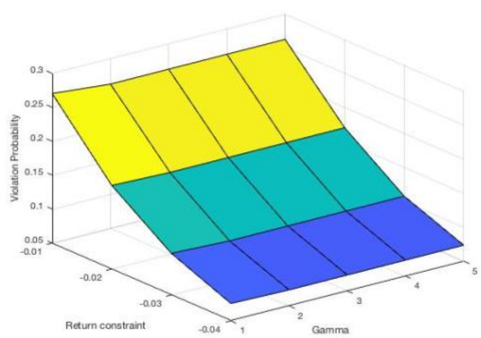

Figure 7: Violation probability each parameter combination of Gamma and return constraint.

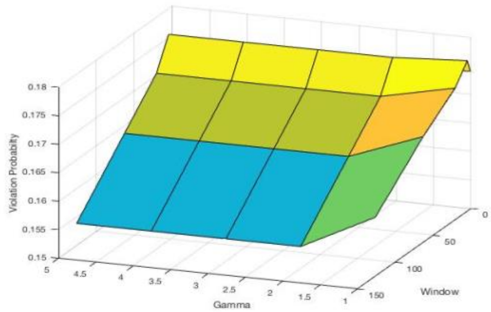

Figure 8: Violation probability for each parameter combination of time and Gamma

4. CONCLUSION

Our aim in this work was to analyze the sensitivity of two portfolio optimization models, with a robust to risk management. Since robust optimization deal with uncertainty within a specific interval of values, it is of interest to evaluate how appropriate this set of values actually is. To do so, we varied the parameters used to build those sets, calculating an empirical probability of violation. In both models, we observed that the parameter related to the number of days and the parameter related to the minimum return required (maximum loss tolerated) are more violated the more restrictive they are. These insights are valid in all datasets considered, using a bootstrap methodology in all of them. That way, we testes in several samples and obtained similar results. This behavior was expected on a qualitative basis but now has quantitative and empirical foundations. Regarding the second model, we also observed an expected relationship between our out-of-sample results and the parameter Γ, responsible to manage the level of conservatism expected.

It is important to notice that although those results were expected in an in-sample analysis, the out-of-sample results were not obvious. As previously cited, expected returns are very difficult to predict. That way, since we did not use any statistical method to predict future returns despite the mean (expected return directly using historical data), we can associate the results to the robust specification with the respective parameters. Thus, we showed a sensitivity analysis to an empirically robust procedure to manage risk that can be valuable in defining risk profiles to investors.

Further investigation will aim the cumulative performance of those models for each parameter combination. One of the questions that we aim do answer is related to whether exists an expected trade-off between risk and return on a cumulative basis. Since there is a higher constraint violation associated with riskier parameters profiles, there is no assurance that they will outperform more conservative choices in a longer period cumulative returns comparison.

REFERENCES

Ben-Tal, A, and A Nemirovski. 1998. “Robust Convex Optimization.” Article. Mathematics of Operations Research 23 (4): 769–805.

Ben-Tal, Aharon, and Arkadi Nemirovski. 1999. “Robust Solutions of Uncertain Linear Programs.” Article. Operations Research Letters 25 (1): 1–13. doi:10.1016/S0167-6377(99)00016-4.

Bertsimas, Dimitris and N. Tritsiklis, John. 1997. Introduction to Linear Optimization. Athena Scientific.

Bertsimas, Dimitris, and Melvyn Sim. 2004. “The Price of Robustness.” Operations Research 52 (1). INFORMS: 35–53. doi:10.1287/opre.1030.0065.

Bjerring, Thomas Trier; Ross, Omri; Weissensteiner, Alex. 2017. “Feature selection for portfolio optimization”. ANNALS OF OPERATIONS RESEARCH 256 (3): 21-40

Black, Fisher. 1993. Fynancial Analists Journal, Vol. 49, No. 5 (Sep. - Oct., 1993), pp. 36-38

Carpenter, James; Bithell, John; 2000. “Bootstrap confidence intervals: when, which, what? A practical guide for medical statisticians”. Statist. Med. 19:1141-1164

Fernandes, Betina; Street, Alexandre; Valladão, Davi, and Fernandes, Cristiano. 2016. “An Adaptive Robust Portfolio Optimization Model with Loss Constraints Based on Data-Driven Polyhedral Uncertainty Sets.” European Journal of Operational Research 255 (3): 961–70.

Ghaoui, Laurent El, and Hervé Lebret. 1997. “Robust Solutions to Least-Squares Problems with Uncertain Data.” Article. SIAM Journal on Matrix Analysis and Applications 18 (4): 1035–64.

Ghaoui, Laurent El, Francois Oustry, and Hervé Lebret. 1998. “Robust Solutions to Uncertain Semidefinite Programs.” Article. SIAM Journal on Optimization 9 (1): 33–52.

Halldórsson, B V, and R H Tütüncü. 2003. “An Interior-Point Method for a Class of Saddle-Point Problems.” Article. Journal of Optimization Theory and Applications 116 (3): 559–90.

Kim, Jang Ho, Woo Chang Kim, and Frank J Fabozzi. 2013. “Recent Developments in Robust Portfolios with a Worst-Case Approach.” Article. Journal of Optimization Theory and Applications 161 (1): 103–21.

Lobo, M S, and S Boyd. 2000. “The Worst-Case Risk of a Portfolio.” Article. Technical Report, Stanford University.

Lotfi, S.; Salahi, M.; Mehrdoust, F. 2017. “Adjusted robust mean-value-at-risk model: less conservative robust portfolios”. OPTIMIZATION AND ENGINEERING 18 (2): 467-497.

Soyster, A L. 1973. “Technical Note—Convex Programming with Set-Inclusive Constraints and Applications to Inexact Linear Programming.” Article. Operations Research 21 (5): 1154–57.

Uryasev, S; Sarykalin, S and Serraino, G (2008). “Value-at-Risk vs. Conditional Value-at-risk in Risk Management and Optimization”. INFORMS Tutorial.