Sizing the bets in a focused portfolio

Vuko Vukčević† and Robert Keser†

†In silico Ltd., Zagreb, Croatia,
vuko.vukcevic@insilico.hr, robert.keser@insilico.hr

October 2, 2023

1 Summary

The paper attempts to provide a mathematical model and a tool for the focused investing strategy as advocated by Buffett [1], Munger [2], and others from this investment community. The approach presented here assumes that the investor's role is to think about probabilities of different outcomes for a set of businesses. Based on these assumptions, the tool calculates the optimal allocation of capital for each of the investment candidates. The model has the option to provide constraints that ensure: no shorting, no use of leverage, providing a maximum limit to the risk of permanent loss of capital, and maximum individual allocation. The software is made available for public use.

This paper consist of two parts, one without the mathematics and one with the mathematics. For a non–mathematical reader, it is best to skip the two sections related to mathematics and numerics, namely section 4 and section 5. The reader who is interested in mathematics should still read the preceding sections since they provide context within which the mathematics is built on. Moreover, for mathematically skilled reader, the problem formulated and solved here represents a non–linear constrained optimization problem.

2 Motivation

Before going on to explain the motivation about this work, a bit of context might come useful. In silico is an employee–owned engineering consultancy company that invests its excess cash into publicly traded stocks. The excess cash comes sporadically, and In silico has been a net buyer over the last four years. This is a trend that will likely continue. When the excess cash comes in, the following question arises: How much money to put in which stocks? That question had been answered by looking at our fundamental analysis for each company, and performing some vague hand calculations. That of course works well because of all the other uncertainties in the investment process, but we were still motivated to:

1. Improve the odds of us behaving rationally (i.e. minimize psychological misjudgments such as anchoring bias [3], consistency and commitment tendency [4], and others with their combined effects [2]),
2. Save time (i.e. avoid doing hand calculations).

Since mathematics is usually good at keeping people rational, and software is great for saving time, a decision has been made to write a software that answers the following question: For a set of candidate companies and their current market capitalization, each having a set of scenarios defined by a probability and intrinsic value estimate, how much of our capital to invest in each?

3 Introduction

The answer to this question has been given by Kelly with his widely known Kelly formula (sometimes called Kelly criterion) [5]. The Kelly's approach starts by maximizing long–term growth of capital when one is presented with an infinite amount of opportunities to bet on. This work extends that idea by considering:

1. Multiple companies (stocks) in parallel,
2. An arbitrary number of scenarios for each company,
3. Multiple constraints for modelling our preference towards no shorting, no use of leverage, and providing a maximum value for the risk of permanent loss of capital,
4. A fundamentals–based analysis with a very long time horizon. After all, a stock is an ownership share of business [6], in which time–scales are significantly larger than the time–scales of day–to–day stock price fluctuations.

The generalization leads to a non–linear system of equations that when solved, yields a fraction of capital to invest in each of the candidate companies. The mathematical derivation is presented in section 4, while in the next two subsections (subsection 3.1 and subsection 3.2), a discussion on assumptions in a non–mathematical way, and the margin of safety, are presented.

3.1 Disclaimers and Assumptions

The underlying philosophy is that one should spend the majority of their time analyzing investments and thinking about intrinsic values under different scenarios that might play out, independently of outside thoughts and events. However, calculating intrinsic value of a company is more of an art than a science, especially for a high-quality, growing businesses within one's circle of competence. And according to Charlie Munger, Warren Buffett, Mohnish Pabrai and the like-minded others from whom the authors got the inspiration for this work, one should focus precisely on getting such great businesses for a fair price. That means that one shouldn't take what this approach says at face value, and one should probably use its guidance infrequently.

During the mathematical derivation presented in section 4, an assumption is made that the number of bets is very high (tends to infinity). This work does not try to justify this assumption in a strong mathematical form [1]. Here's a soft, non-mathematical reasoning on why the authors think this is fine: The assumption is made in order to write the equations in terms of probabilities instead of the number of outcomes divided by the number of bets. Therefore, as long as the input probabilities are *conservatively* estimated, the framework should still be valid. This essentially represents the most important margin of safety [6].

3.2 Margin of Safety

In addition to the most important margin of safety mentioned above, there are however a couple of more margins of safety embedded in the current framework. These are:

- No shorting allowed. From a purely mathematical point of view, shorting would be allowed. Without detailed math, it is easy to see how a company with a negative expected return would result in a short position. However, due to the asymmetry of the potential losses compared to gains, coupled with the usual time–frame limit that comes with shorting, we provided a constraint for being long–only.

- No use of leverage allowed. Again, from a purely mathematical point of view, use of leverage would sometimes be useful. The thinking in avoiding to use leverage is that no–one should be in a hurry to get rich, and should avoid risking good night's sleep based on short–term market fluctuations, which are fairly hard to predict consistently (unless you work at Renaissance Technologies, in which case you can scratch this).

- No company without at least one downside scenario is allowed. By disallowing inputs without downside, the framework forces you to focus and think about what can go wrong, as opposed to dreaming about what can go right. This is in–line with thoughts by Pasad [7] about the importance of avoiding the errors of investing in a bad company. However, if one is absolutely (100%) sure that a company does not have a downside, the solution is to put all the money in that one company. It is our feeling that the best way to handle such cases is outside of this framework. Alternatively, one can always model an unknown downside scenario with a small probability (say 5–20%) and intrinsic value of zero.

[1] It is on the author's TODO list to try and prove this.

These assumptions and margins of safety are embedded into the framework as constraints in order to try and tie the rational mathematics with the real world.

4 Mathematics

The following derivation mostly follows the first part of the work by Byrnes and Barnett [8]. The problem statement is repeated here for convenience: Given a set of candidate companies, each having a set of scenarios described by the probability p and estimate of the intrinsic value $\mathcal{V}$, calculate the optimal allocation fraction f for each candidate company by maximizing the long–term growth rate of assets. After a single outcome (realization), the change in value of assets can be written as follows:

$$\mathcal{A}_{after} = \mathcal{A}_{before}\left(1 + \sum_{j}^{N_c} f_j k_j\right) \tag{1}$$

where $\mathcal{A}$ is the value of assets (capital), N_c is the number of candidate companies to consider, f_j is the allocated fraction to jth company, and k_j is a return for a company j defined as the relative difference between the estimated intrinsic value under a given scenario ($\mathcal{V}$) and the market capitalization at the time of investing ($\mathcal{M}$):

$$k_j = \frac{\mathcal{V}_j - \mathcal{M}_j}{\mathcal{M}_j} \tag{2}$$

If a significant number of (re)allocations (N_a) is performed in succession, the equation (1) can be written as follows:

$$\mathcal{A}_{N_a} = \mathcal{A}_0 \prod_{i_1, i_2, \ldots, i_{N_o}} \left(1 + \sum_{j}^{N_c} f_j k_{ij}\right)^{n_i} \tag{3}$$

where $n_i \in n_{i1}, n_{i2}, \ldots, n_{N_o}$ is the number of times the ith outcome has occurred. Note that k_{ij} represents the return of the jth company for the ith outcome. Following original Kelly's approach, a logarithmic growth function $\mathcal{G}$ is introduced:

$$\mathcal{G} = \lim_{N_a \to \infty} \frac{1}{N_a} \ln \frac{\mathcal{A}_{N_a}}{\mathcal{A}_0} \tag{4}$$

and the goal is to find its maximum with respect to allocation fractions f_j:

$$\frac{\partial \mathcal{G}}{\partial f_j} = 0 \tag{5}$$

Substituting equation (1) into equation (5) results in the following equation, after some calculus and algebra:

$$\lim_{N_a \to \infty} \frac{1}{N_a} \sum_{i}^{N_o} \frac{n_i k_{ij}}{1 + \sum_j^{N_c} f_j k_{ij}} = 0 \tag{6}$$

If one assumes an infinite number of (re)allocations N_a[2], the following relation holds:

$$\lim_{N_a \to \infty} \frac{n_i}{N} = p_i \tag{7}$$

Where p_i is the probability of the ith outcome. For example, if there are two companies, each with two 50–50 scenarios, there will be four outcomes in total with the probability of each outcome equal to 25%. Finally, substituting equation (7) into (6) results in a system of equations written in terms of probabilities p_i, expected returns for each company in each of the outcomes k_{ij}, and allocation fractions for each company f_j:

$$\sum_{i}^{N_o} \frac{p_i k_{ij}}{1 + \sum_{j}^{N_c} f_j k_{ij}} = 0 \tag{8}$$

The equation (8) represents a non–linear system of equations in the unknown fractions f_j, which when solved, should yield optimal allocation strategy for maximizing long–term growth of capital.

4.1 Constraints

Equation (8) has no constraints, meaning that after solving the system of equations, the resulting fractions f_j might be negative and greater than one. This would imply short positions and use of leverage, respectively. A general inequality constraint may be written in the following form:

$$I(f_j) \leq 0 \tag{9}$$

where, for example, $I(f_j) \equiv -f \leq 0$ models a long-only constraint that would make sure that the fractions are positive. In order to transform the inequality constraint into an equality constraint, we introduce a slack variable s that must be positive:

$$I(f_j) + s = 0 \equiv \mathcal{C}(f_j, s) = 0 \tag{10}$$

where the second part of the equation introduces a useful substitution for deriving the constrained system later on.

In this work, we define four constraints that serve as additional margins of safety in a focused investment approach:

1. Long–only constraint ensures that a fraction cannot be negative and is added to all candidate companies. Adding this constraint means: "I do not allow short positions.":

$$f \geq 0 \to -f + s = 0 \tag{11}$$

[2]The assumption regarding the infinite number of allocations is something that the authors are slightly uncomfortable with, but *feels* it is fine because of the margins of safety embedded into the thinking that goes into assessing each investment opportunity.

2. Maximum leverage constraint ensures that the leverage is limited up to L. Note that $L = 0$ implies no leverage. Adding this constraint means: "I want to limit the use of leverage (to a minimal amount).":

$$\sum_{j}^{N_c} f_j \leq 1 + L \rightarrow \sum_{j}^{N_c} f_j - 1 - L + s = 0 \tag{12}$$

3. Maximum individual allocation constraint ensures that a fraction does not exceed the specified amount. Adding this constraint means: "I feel uncomfortable putting more than M of my capital into a single company.":

$$f \leq M \rightarrow f - M + s = 0 \tag{13}$$

4. Maximum allowable permanent capital loss constraint ensures that the worst–case outcome does not exceed losing a specified amount of capital with a specified probability. Adding this constraint means: "Under a worst–case scenario, I am comfortable permanently losing K (e.g. 25%) of my capital with probability P (e.g. 0.1%).":

$$\sum_{j}^{N_c} f_j \min(p_{ij} k_{ij}) \geq P \cdot K \rightarrow -\sum_{j}^{N_c} (f_j \min(p_{ij} k_{ij})) + P \cdot K + s = 0 \tag{14}$$

   where $\min(p_{ij} k_{ij})$ is the worst–case outcome across all scenarios for the j-th candidate company. Here, the minimum returns k_{ij} and the maximum worst–case return K are both negative by convention, indicating a loss, hence the $\geq$ sign.

   Note that the maximum allowable permanent capital loss constraint given by equation (14) only works with the long–only constraint because the fraction f_j is assumed positive. It is also important to note that here we do not talk about *temporary* loss of capital due to short-term stock market fluctuations, but rather *permanent* loss of capital due to the fundamental business environment of candidate companies. In addition, it is of course possible for one to lose more than the specified maximum allowable amount of capital because of the assumptions made with respect to the inputs: After all, coming up with intrinsic values and probabilities for each scenario is more of an art than a science.

4.2 Constrained System: Putting It All Together

The growth function (4) can be constrained with an arbitrary amount of constraints (10) by introducing a Lagrangian:

$$\mathcal{L}(f_j, \lambda_l) = \mathcal{G}(f_j) - \sum_{l}^{N_l} \lambda_l \mathcal{C}_l(f_j, s_l) \tag{15}$$

where N_l is the number of constraints and l denotes the l-th constraint defined either by l-th Lagrange multiplier λ_l or the l-th slack variable s_l. There are two necessary conditions for finding a constrained maximum of the growth function:

$$\alpha_i(f_j, \lambda_l, s_l) = \frac{\partial \mathcal{L}(f_j, \lambda_l)}{\partial f_j} = \frac{\partial \mathcal{G}(f_j)}{\partial f_j} - \sum_{l}^{N_l} \lambda_l \frac{\partial \mathcal{C}_l(f_j, s_l)}{\partial f_j} = 0 \tag{16}$$

$$\beta_i(f_j, s_l) = \frac{\partial \mathcal{L}(f_j, \lambda_l)}{\partial \lambda_l} = -\mathcal{C}_l(f_j, s_l) = 0 \tag{17}$$

where α_i and β_i have been introduced as a shorthand notation that distinguishes between two vector equations: There are N_c of α equations and N_l of β equations. Therefore, there are $N_c + N_l$ equations, but $N_c + 2N_l$ unknowns, because of N_c unknown fractions, N_l unknown Lagrange multipliers λ_l and N_l unknown slack variables s_l. However, an inequality constraint cannot be active and inactive at the same time. An active constraint is characterized by $\lambda_l \neq 0$ and $s_l = 0$, whereas an inactive constraint is characterized by $\lambda_l = 0$ and $s_l > 0 \neq 0$. This means that we have to solve 2^{N_l} nonlinear systems to cover all combinations of constraints and pick the best solution. As an example, adding all constraints mentioned in subsection 4.1 for a portfolio with five candidate companies would result in having to solve $2^{12} = 4096$ systems, while having ten candidate companies would imply solving $2^{22} = 4194304$ systems, demonstrating the exponential complexity of the problem.

5 Numerics

The equations (16) and (17) can be written succinctly as:

$$\mathcal{F}_i(x_i) = 0 \tag{18}$$

where x_i is a vector of unknown fractions and unknown Lagrange multipliers or slack variables:

$$x_i = \{f_1, f_2, \ldots, f_{N_c}, \lambda_1 | s_1, \lambda_2 | s_2, \ldots, \lambda_{N_l} | s_{N_l}\} \tag{19}$$

Because $\mathcal{F}_i$ is a non–linear equation in f_j, the Newton–Raphson method is used to find a numerical solution. The method is iterative and starts by linearizing the equation around the previous solution from the previous iteration:

$$\mathcal{F}_i^o + \sum_{i}^{N_c + N_l} \mathcal{J}_{ij}^o (x_j^n - x_j^o) = 0 \tag{20}$$

where $\mathcal{J}_{ij}$ is the Jacobian of $\mathcal{F}_i$, and superscripts ^{n} and ^{o} denote the new value and the old value, respectively.

5.1 Jacobian for an Active Constraint

If a constraint is active ($\lambda_l \neq 0, s_l = 0$), the corresponding unknown is the Lagrange multiplier and the Jacobian has the following form:

$$\mathcal{J}_{ij} = \begin{pmatrix} \frac{\partial \alpha_i}{\partial f_j} & \frac{\partial \alpha_i}{\partial \lambda_j} \\ \frac{\partial \beta_i}{\partial f_j} & \frac{\partial \beta_i}{\partial \lambda_j} \end{pmatrix} = \begin{pmatrix} \frac{\partial^2 \mathcal{G}}{\partial f_i \partial f_j} - \sum_j^{N_l} \left(\lambda_j \frac{\partial^2 \mathcal{C}_i}{\partial f_j^2} \right) & -\frac{\partial \mathcal{C}_i}{\partial f_j} \\ -\frac{\partial \mathcal{C}_i}{\partial f_j} & 0 \end{pmatrix} \tag{21}$$

5.2 Jacobian for an Inactive Constraint

If a constraint is inactive ($\lambda_l = 0, s_l > 0 \neq 0$), the corresponding unknown is the slack variable and the Jacobian has the following form:

$$\mathcal{J}_{ij} = \begin{pmatrix} \frac{\partial \alpha_i}{\partial f_j} & \frac{\partial \alpha_i}{\partial s_j} \\ \frac{\partial \beta_i}{\partial f_j} & \frac{\partial \beta_i}{\partial s_j} \end{pmatrix} = \begin{pmatrix} \frac{\partial^2 \mathcal{G}}{\partial f_i \partial f_j} & 0 \\ -\frac{\partial \mathcal{C}_i}{\partial f_j} & -1 \end{pmatrix} \tag{22}$$

Partial derivatives of the constraint function can be readily obtained because all constraints presented in this work are simple analytical functions (see subsection 4.1). For completeness, the Hessian of the growth function that appears in the upper–right corner is:

$$\frac{\partial^2 \mathcal{G}}{\partial f_i \partial f_j} = -\sum_o^{N_o} \frac{p_o k_{oi} k_{oj}}{\left(1 + \sum_j^{N_c} f_j k_{ij}\right)^2} \tag{23}$$

where subscript for outcome i has been changed to o in order to be able to write the Hessian in the standard ij notation.

Note that the equations (21) and (22) are presented in a way that all constraints are either active or inactive. Since we have to solve 2^{N_l} systems characterized by active/inactive status of each constraint, that simply means that we need to add respective active/inactive contributions to the Jacobian $\mathcal{J}_{ij}$ and the right–hand–side $\mathcal{F}_i^o$ in equation (20) for a particular constraint j.

5.3 Notes on the Numerical Procedure

The numerical procedure starts by assuming the uniform allocation across all companies, i.e. $f_j = f = \frac{1}{N_c}$. Based on the f_j^o in the current iteration, the linear system in equation (10) is solved to find the new solution f_j^n. The process is repeated until sufficient level of accuracy is reached.

After solving all 2_l^N solutions, we end up with less than 2_l^N viable solutions. A solution is considered viable if the Newton–Raphson procedure managed to find a numerical solution and if all slack variables for all inactive constraints are positive. Finding the best solution out of all viable solutions would require evaluating the growth function for all solutions, which is particularly challenging due to the product series in equation (3). In this work, we take a pragmatic approach and simply choose the solution that maximizes the expected value of the portfolio out of a set of most diversified solutions, i.e. the solutions with the highest number of non–zero allocations.

6 Basic Validation Tests

In order to validate the numerical model, a basic example of five candidate companies is considered, where each candidate has the same set of scenarios (probabilities and intrinsic values) and the same market cap. The inputs that present a 50% loss and 100% gain with 50–50 probabilities are presented in Table 1.

Table 1: Company for a validation test with a market cap of 1.

Scenario	Intrinsic value	Probability
50% down with 50% probability	0.5	50%
100% up with 50% probability	2	50%

Solving the system without any constraints yields a uniform allocation of 35% of capital in each company. Note that because we considered 5 companies, that implies 75% leverage ($5 \cdot 35\% = 175\%$). Even without the constraint for maximum allocation of capital, just maximizing the long–term growth–rate of assets prefers a diversified solution, which is expected.

Adding a constraint for no leverage (given by equation (12) and setting $L = 0$), results in a uniform allocation of 20%, as expected. It is straightforward to show that the worst–case outcome in such a portfolio implies permanently losing 50% of the capital with probability of 3.125%.

The final test is done by setting the maximum allowable permanent capital loss constraint as given by equation (14). Setting $P = 5\%$ and $K = 50\%$, indicating that we are comfortable risking to lose 50% of the capital with 5% probability, results in a uniform allocation of 2%. Because only 10% of capital is invested in that case, there is a possibility of losing 5% of capital with probability of 3.125%. In the worst–case scenario, the probability–weighted return is $-0.5 \cdot 0.02 \cdot 0.5 \cdot 5 = -2.5\%$, which is equal to $P \cdot K$.

7 Example

With the basic validation of the numerical model performed, the attention is moved to a realistic example. Consider five candidate companies, each with up to three scenarios. Each scenario is represented by an intrinsic value and the probability of the scenario happening (or intrinsic value being reached at some point in the future). Note that how these numbers are obtained is outside of the scope of this work, although it is important to stress that the validity and conservative assumptions behind these numbers are probably the most important part of an investor's job. The example inputs are presented in Table 2 to Table 6.

Table 2: Company A with current market cap of 225B USD.

Scenario	Intrinsic value	Probability
Total loss	0 USD	5%
Base thesis	270B USD	60%
Bull thesis	420B USD	35%

Table 3: Company B with current market cap of 450M USD.

Scenario	Intrinsic value	Probability
Total loss	0 USD	5%
Bear thesis	350M USD	50%
Base thesis	900M USD	45%

Table 4: Company C with current market cap of 39M GBP.

Scenario	Intrinsic value	Probability
Total loss	0 GBP	10%
Bear thesis	34M GBP	40%
Base thesis	135M GBP	50%

Table 5: Company D with current market cap of 751M SGD.

Scenario	Intrinsic value	Probability
Bear thesis	330M SGD	30%
Base thesis	1B SGD	70%

Table 6: Company E with current market cap of 126B HKD.

Scenario	Intrinsic value	Probability
Total loss	0 HKD	5%
Bear thesis	50B HKD	10%
Base thesis	300B HKD	85%

Based on these inputs, with the long–only strategy, without leverage, and with maximum individual allocation of 30%, the portfolio allocation that maximizes the long–term growth–rate of capital is presented in Table 7.

Table 7: Portfolio that maximizes long–term growth–rate of capital.

Company	A	B	C	D	E
Allocation fractions	30%	8%	30%	2%	30%

With the obtained fractions, it is easy to obtain some useful statistics on the portfolio, namely:

- Expected gain of 32 cents for every dollar invested,
- Cumulative probability of loss of capital of 16%,
- Permanent loss of 60% of capital with probability of 0.008%.

The last item is particularly interesting to the authors. According to Actuarial Life Tables in [9], the probability of the (currently 34 year–old, the oldest) author dying within the next year is approximately 0.26%[3]. That is two orders of magnitude higher than the probability of the permanent loss of capital for this portfolio. Considering that the portfolio has five stocks, that is a very strong argument against excessive diversification, especially if:

- One thinks of stocks as ownership shares of businesses, which implies long–term thinking and not being bothered by market fluctuations,
- One embeds a margin of safety in different scenarios for different companies by e.g. recognizing that both unknown and known bad things may happen.

The observation about excessive diversification is inline with the thoughts from the Poor Charlie's Almanack [2] and one of the lectures from Li Lu that the authors frequently watch [10].

8 Problems, Discussion, and Future Work

There are a couple of technical problems that the authors observed:

1. It is possible that a given non–linear system for a particular combination of constraint statuses (active/inactive) does not converge. This may happen if the resulting matrix is singular, or if the Newton–Raphson algorithm does not find the solution within a prescribed number of iterations. These errors are ignored, which means that it is possible that the real best solution is not found due to numerical issues.
2. The exponential complexity of the model makes it challenging to use for more than several candidate companies without using significant compute resources. For example, having a 20 candidate companies with all

[3] This is the only thing that's probably worth measuring in the basis points.

constraints would result in around 4 trillion non–linear systems to solve. Therefore, the model is not suitable for cases with excessive diversification, although there is a possibility to filter some of these upfront without attempting to solve them, which may be one of the topics for future work.

To conclude, the authors believe that the most challenging aspect of an investor's work that might use this software is to think hard about the inputs. The authors see the usefulness of this software mainly in:

- Forcing the investors to think consistently in terms of probabilities and long–term business outcomes across the range of candidate companies,
- Calculating the optimal allocation in a short amount of time, based on the investor's inputs.

9 Access to the Software

References

[1] Warren E. Buffett. Berkshire Hathaway Shareholder Letters. https://www.berkshirehathaway.com/letters/letters.html, 2023. [Online; accessed 19 August 2023].

[2] P.D. Kaufman, E. Wexler, W.E. Buffett, and C.T. Munger. *Poor Charlie's Almanack: The Wit and Wisdom of Charles T. Munger*. Walsworth Publishing Company, 2005.

[3] D. Kahneman. *Thinking, Fast and Slow*. Farrar, Straus and Giroux, 2011.

[4] R.B. Cialdini. *Influence: The Psychology of Persuasion*. Harper Business, 2006.

[5] J.L. Kelly. A New Interpretation of Information Rate. *Bell System Technical Journal*, 35(4):917–926, 1956.

[6] B. Graham. *The Intelligent Investor*. Harper Business, 2003.

[7] P. Prasad. *What I Learned About Investing from Darwin*. Columbia Business School, 2023.

[8] T. Byrnes and T Barnett. Generalized framework for applying the Kelly criterion to stock markets. *International Journal of Theoretical and Applied Finance*, 21(5), 2018.

[9] United States Social Security Administration. Actuarial Life Table. https://www.ssa.gov/oact/STATS/table4c6.html, 2023. [Online; accessed 19 August 2023].

[10] Li Lu. Li Lu's Lecture at Professor Greenwald's 2006 Class. https://www.youtube.com/watch?v=-jF5au0-JiY, 2023. [Online; accessed 2 September 2023].