Explainable data-driven portfolio construction with conditional bootstrapped Shapley values

Emiel Lemahieu[1 2 3]

Ghent University, InvestSuite

Abstract

Data-driven portfolio construction is non-parametric in the sense that it does not impose its objective on a parametric representation of input paths, such as a variance-covariance matrix, but rather imposes its objective directly on these paths without having to make assumptions on their underlying data generating process (DGP). The advantage is thus that one can optimize on much richer dynamics than typical DGP assumptions, e.g. Gaussian assumptions behind mean-variance optimization. The main disadvantage is that one can seemingly interpret the parameters that gave rise to the optimal portfolios not that easily anymore, and for instance do sensitivity analysis, i.e. how will the portfolio change when parameter x or y in- or decreases? This is conducive to the *black box* perception of increasingly more data-driven portfolio construction tools. However, tools from explainable AI such as Shapley values, which are mathematically principled contribution estimators of features to their output, can be used to overcome this issue. We advocate a novel approach that combines the powerful statistical approach of conditional bootstrapping with Shapley values that attribute changing optimal portfolio weights to changing market conditions. The empirical usefulness is shown on a US equity portfolio backtest, where the portfolio manager gets a better insight in his or her portfolio composition and its sensitivity to changing market conditions.

Keywords: Data-Driven Portfolio Construction, Bootstrapping, Shapley Values
MSC: 91B28, 91B84

1. Introduction

1.1. Problem setting

Data-driven portfolio construction comes down to imposing less to no parametric restrictions on the behaviour of the investible instruments and their comovements, but rather performing an optimization on their paths data directly. It is non-parametric in the sense that we do not assume a data generating process (DGP) that generated these input sample paths, such as the assumed Gaussianity of mean-variance optimization, or assumed DGP parameters in mean-variance-skew-kurtosis or higher DGP-moments extensions. We do not rely on a parametric representation of these input sample path such as a variance-covariance matrix or higher order

[1]Ghent University, Sint-Pietersplein 6, 9000 Gent, Belgium
[2]InvestSuite, `www.investsuite.com`
[3]Corresponding author at emiel.lemahieu@ugent.be

cumulants, but use sample paths as input in our optimization problem. Moreover, the optimization tool we propose is also non-parametric in the sense that it does not rely on optimization hyperparameters like the thresholds in (conditional) value-at-risk (VaR) or expected shortfall (ES) optimizations.

The advantage of such approaches is that one is more loyal to the actual DGP by not imposing a restrictive view on the DGP params (such as Gaussianity) and one finds more robust portfolios as one is not prone to noise- and signal-induced instabilities of parametric models. The disadvantage is that the quantitative finance literature has built a large knowledge on top of these standard assumptions (such as mean-variance quadratic utility, portfolio *betas*, related computational considerations etc.) among which the experience to interpret the estimated parameters and leverage them for sensitivity analysis and model interpretability in general. This is something that is often claimed to be lost with purely data-driven models.

In the context of portfolio optimization, we want to argue against this common critique of data-driven optimizers being *black boxes*. Something as simple as a covariance matrix can be used as a black box, if one just calibrates it to historical data, plugs it into a linear-quadratic program and uses the optimal portfolio as some ground truth, without evaluating concentration and stability, estimation error or shrunk covariance, non-stationarity or autoregressive covariance, and so and so forth.

In other words, even the value of a seemingly simple model is in the eye of the beholder. It is true that for a non-parametric model one can not evaluate the impact of the parameters on the outcome portfolio. E.g. in a mean-variance optimizer one would be interested in the sensitivity of a return estimate (i.e. a view) on the resulting portfolio in a Black-Litterman-like way. Alternatively, one would be interested in the impact of a change in volatility estimate (e.g. empirical versus GARCH or factor models) on the outcome portfolio.

Admittedly, with non-parametric optimization this is far less trivial. However, this paper argues that one can still investigate the impact of a change in input samples on a change in resulting portfolio by using a combination of two statistical techniques:

1. **A conditional bootstrap procedure** (Hinkley [1]): resample historical scenarios with replacement and condition them on some variable, e.g. macro indicators.

2. **SHAP or Shapley values** (Shapley [2], Lundberg and Lee [3]): a mathematically principled tool of attributing outcome function values to changes in their input based on the principles of game theory (coalition games).

We argue that a combination of these two established approaches is a relatively easy but nonetheless powerful technique to help 'white-boxing' the increasingly more popular data-driven optimizer techniques.

1.2. Contribution

To the best of our knowledge we are the first to use a combination of bootstrap methods and Shapley values for enhancing the explainability of portfolio construction methods that are purely data-driven as per our above definition. Data-driven asset allocations are becoming more popular in the industry, and the combination of bootstrap and SHAP outlined below is not specific to the optimization context of min drawdown optimization (e.g. neural network optimization, reinforcement learning, etc.), hence it could be a useful addition to the general quant or asset manager's toolbox.

1.3. Related literature

The idea of a Shapley value, essentially evaluating the marginal contribution of a feature to a prediction, is such a universal concept that it is not surprising that since the breakthrough of the concept in the machine learning literature (Lundberg and Lee [3]) many authors have applied it to finance. These include Hagan et al. [4], Colini-Baldeschi et al. [5], Shalit [6], [7], [8], Tarashev et al. [9], Moehle et al. [10], Simonian [11], Babaei et al. [12], Ohana et al. [13], Benhamou et al. [14] and Kimura et al. [15]. Closest related to this paper is the works of Jaeger et al. [16] which apply bootstrapped Shapley values for explainable portfolio construction, but they focus on inverse volatility (IVP), equal risk contribution (ERC) and hierarchical risk parity (HRP), which are essentially just variance-based portfolio constructors and are as white-box as what we first labeled the 'traditional' techniques. They are only data-driven in the sense that HRP uses single linkage clustering to solve an inverse variance problem on clusters of the original universe. This still assumes Gaussianity over these clusters. Colini-Baldeschi et al. [5] describe the SHAP values of holdings to utility of a variance-based utility investor. In a concise but powerful paper Hagan et al. [4] use SHAP values as alternatives to the *greeks* (risk sensitivities) in a derivative portfolio with stochastic volatility dynamics. They first establish (from Colini-Baldeschi et al. [5]) that for a mean-variance investor the Shapley values are very closely related to *portfolio betas* (i.e. the covariance of instrument returns with the portfolio returns divided by the instrument volatility). Next, they do portfolio VaR and ES attribution to risk factors, but rely on elliptical assumptions to do so. Shalit [6] first answered the question how much each security in a portfolio contributes to the risk-reward tradeoff using Shapley value regression, using a different interpretation of SHAP, namely a local Shapley regression. The latter method will also be used in this paper. Shalit [8] follows up with generalized risk decomposition with SHAP and Shalit [7] focuses on systematic risk identification. Similarly Tarashev et al. [9] performs systematic risk identification with SHAP, but focusing on banks. Moehle et al. [10] offers a more theoretical treatment with digressions on tax management and factor investing. Kimura et al. [15] use Shapley values for factor investing as well. Simonian [11] interestingly uses SHAP for portfolio selection itself, but not for explaining existing optimizer tools. Babaei et al. [12] use Shapley values to explain traditional Markowitz mean-variance models by observing only its outcome, without knowing the input parameters (e.g. by observing an unknown strategy returns and assuming it uses mean-variance optimization). Finally, Ohana et al. [13] and Benhamou et al. [14] also include macro conditions as features, but for the prediction of regime switches and not for white-boxing data-driven portfolio construction. There has been no paper that started from a fundamentally non-parametric (no assumed DGP) and used Shapley values as a demystifier.

1.4. Organization of results

In the next section 2 we discuss data-driven portfolio optimization as optimizers that do not require a DGP, a parametric specification of sample paths, or hyperparameters, to optimize. As an example, we use portfolio drawdown optimization, but it should be stressed that the bootstrap and SHAP approach is general enough to apply to other data-driven optimizers. Next, we discuss the conditional bootstrap as a specific form of weighted historical simulation that constitutes the conditional expectation of the risk functional one optimizes for. In section 3 we delve deeper into the definition of Shapley values, the link with conditional bootstrap and the link with efficient implementations of local Shapley regression. Section 4 discusses the application and backtest findings. Data-driven methods offer the possibility to better accommodate the true DGP, be more predictive in what they try to predict (i.e. have lower out-of-sample error), generate better risk-adjusted returns, exhibit more stability, have less concentrations. However, they are seemingly less interpretable. Therefore, we show in section 3 that by using relatively simple methods one can still do sensitivity analysis on the portfolio and formulate expectations on how the model is going to behave when the market conditions change. For these conditions we built a toy model of the US economy based on FRED data and LASSO regression. Section 5 concludes.

2. Data-driven portfolio optimization

For an N-dimensional universe of investible instruments, let us denote by $\mathbf{w}$ the vector of portfolio weights $w_i, i \in \{1, ..., N\}$. Further, $\mathbf{\Sigma}$ is the sample variance-covariance matrix of their historical return timeseries $\mathbf{X} : [0, T] \to \mathbb{R}^N$, where $x_{i,t} = s_{i,t}/s_{i,t-1} - 1$, and $\mathbf{S} : [0, T] \to \mathbb{R}^N$ is the T by N matrix of historical spot prices or index levels. For convenience, let us also write $\boldsymbol{\sigma} = diag(\mathbf{\Sigma})$, where σ_i corresponds to individual asset i's volatility. We introduce the minimum drawdown portfolio in terms of these notations below[4].

2.1. Portfolio drawdown optimization

The minimum drawdown portfolio is the solution to the following linear optimization problem:

$$\begin{aligned} \min_{\mathbf{w}} \quad & \mathbb{E}(\xi(\mathbf{w})) \\ \text{s.t.} \quad & \xi_{\mathbf{t}} = \mathbf{m_t} - \mathbf{w}\mathbf{S}_t \\ & \mathbf{m}_t \geq \mathbf{m}_{t-1} \\ & \mathbf{m}_t \geq \mathbf{w}\mathbf{S}_t \\ & \mathbf{w}\mathbf{I}^N = 1 \\ & \mathbf{w} \geq 0 \end{aligned} \tag{1}$$

where we minimize the expected drawdown ξ as a function of portfolio weights $\mathbf{w}$. The drawdown ξ is a non-linear function of the portfolio path $\mathbf{P_t} = \mathbf{w}\mathbf{S}_t$, $\xi_{\mathbf{t}} = \max(\max_{k<t}(\mathbf{P_k}) - \mathbf{P_t}, 0)$, but can hence be written as a linear problem by instrument variable $\mathbf{m_t}$ which denotes the monotonic growth of the portfolio value $\mathbf{m}_t \geq \mathbf{w}\mathbf{S}_t$. Chekhlov et al. [17] show that the minimum drawdown measure satisfies the properties of a deviation measure[5] and generalizes them to a dynamic conditional.

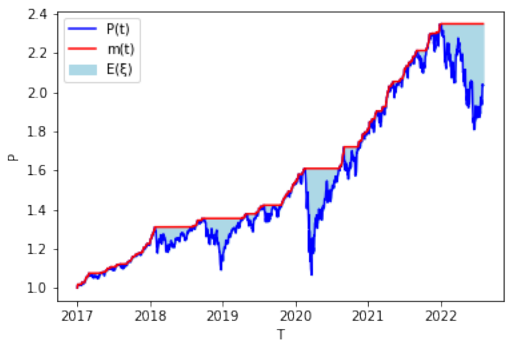

Figure 1: Example of ξ, m_t and S_t for a US Equity index (S&P500)

[4]The benchmark portfolios - minimum volatility, the market cap weighted index, equal weighting - are introduced in terms of the above notation in Section 4.2.

[5]More specifically, (1) non-negativity, (2) insensitivity to a constant shift, (3) positive homogeneity and (4) convexity.

It is crucial to understand that we do not rely on a parametric representation of S, need no additional regularity like elliptic assumptions, hence make no assumptions about the DGP underlying S. Moreover, the model does not require any hyperparameters such as VaR or ES confidence levels.

The focal element of (1) is the expectation $\mathbb{E}(\xi)$, which can be taken over time $[0, T]$ as an (ergodic) time-integral, or integrated over a (non-ergodic) ensemble of paths. The next section delves deeper into the difference between the two and the importance of simulation methods.

2.2. Expected portfolio drawdown and the conditional block bootstrap

The standard go-to implementation for the input path space $\mathbf{S} : [0, T] \to \mathbb{R}^N$ is to use the historical price paths [18][19][20]. This makes a few (most often implicit) assumptions: (a) the scenario length τ is equal to the amount of historical observations T, (b) all historically observed drawdowns are signal for expected drawdown (i.e. no noise), (c) all possible drawdown states were ever realized. All three assumptions are problematic.

The true expected drawdown is the probability-weighted drawdown integrated over all possible drawdown states Ω. Of course, we do not know the probabilities and tend to estimate them from the historical sample as some time-integrated mean:

$$\mathbb{E}(\xi) = \int_{\Omega} \xi dp = \frac{1}{T} \int_0^T \xi(P_t) dt \tag{2}$$

where the time-integrated mean is assumed equal to the true expectation. This is the common ergodicity assumption linked to assumption (b) and (c)[6]. This implies that over many repeated samples $j \in [0, N_s]$:

$$\mathbb{E}(\xi) = \frac{1}{N_s \tau} \int_0^{N_s} \int_0^{\tau} \xi(P_{t,j}) dt dj \tag{3}$$

the two integrals are commutative and equal to (2). This assumption is clearly flawed. By contrast, the non-ergodic alternative implies that the integrals in Eq. (3) are non-commutative nor equal to time-integrated mean and Eq. (3) is closer to the true $\mathbb{E}(\xi)$. Eq. (3) is essentially an ensemble mean of many possible scenarios that can be noisy, where the noise is removed by additional integration.

Further relaxing assumption (a), we naturally arrive at the historical simulation methods proposed in [23][24]. When $\tau < T$, $N_s = \lfloor T/\tau \rfloor$ non-overlapping blocks of data exist and $N_s = T - \tau$ overlapping blocks can be used (e.g. Figure 2):

$$S = (S_1, S_2, ..., S_{N_s}) \tag{4}$$

Moreover, weighted historical simulation can be used:

$$\mathbb{E}(\xi) = \frac{1}{N_s \tau} \int_0^{N_s} \int_0^{\tau} b_j \xi(P_{t,j}) dt dj \tag{5}$$

Examples, most noteworthy, include:

[6] Ergodicity is a concept from statistical physics that essentially states that for a system in equilibrium the expectation over time will equal the ensemble expectation. This is generally not observed in finance and economics (Peters [21] and Taleb [22]). Processes are just one realization of what could have happened and they never achieve all their possible states. This questions the use of historical data in general, but unfortunately it is all we have for quantitative optimization. The implications here are that, in any case, we are advised to take an ensemble average rather than a single time average.

- *Exponentially-weighted moving average (EWMA)*: b_j is an exponential function of j, giving higher weight to more recent observations that are considered more relevant today.

- *Volatility-filtered sampling*: $b_j = \frac{\sigma_t}{\sigma_j}$, the historical samples are divided by their historical volatilities and multiplied with contemporaneous volatility.

- *Conditional sampling*: $b_j = 0$ for sequences not satisfying a historical condition (e.g. VIX > 40%, CPI growth > 5%, etc.) and $b_j = 1$ if they do, only including information of historical episodes considered more relevant today (e.g. market turmoil).

We propose a latter bootstrap procedure where the weight b_j is attributed to the resulting portfolio weights **w**. A *block bootstrapped* minimum drawdown problem [23][24] is one where for a large number of draws R, e.g. R = 1000, we pick a random index j in $\{1, 2, ..., N_s\}$ and add S_j from Eq. (4) to the problem (1). We do this with *replacement*, meaning an index j can be used multiple times (duplicate trajectories), otherwise R would be limited to N_s. The word *block* refers to the fact that every S_j is of length τ such that individual S_t at time t are not assumed i.i.d.

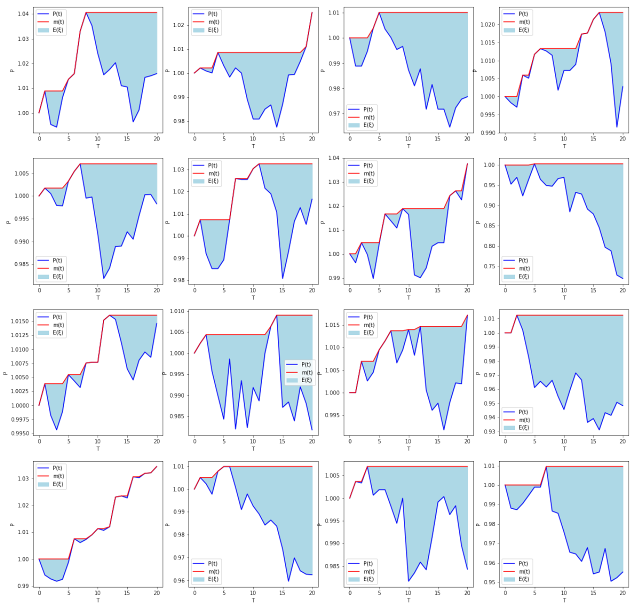

Figure 2: Example of S for $\tau = 20 < T = 1300$, with $N_s = 1280$ $(= T - \tau)$ overlapping blocks for the US Equity index (S&P500). I.e. multi-scenario drawdown plots for 20 day (+- 1 month) scenarios.

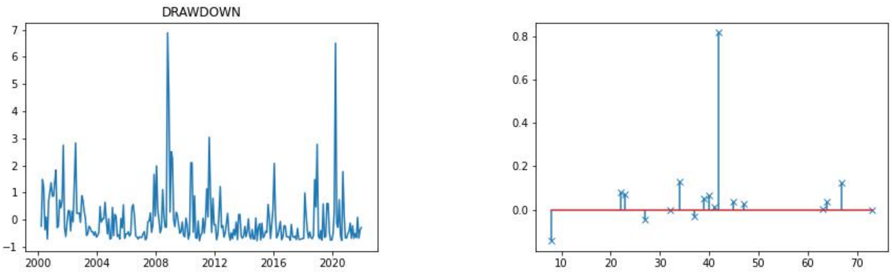

Figure 3: The evolution of ξ of US Stock market index (Wilshire, left), the LASSO coefficients of the conditions (right)

3. Conditional attribution: portfolio sensitivity to underlying conditionalities

Conditions C_i are economic priors for the model. Like views, they let us generate scenarios under the prevalent market circumstances, such as contemporaneous volatility, but they also alow us to vary these exogenous factors and evaluate what that implies for our the generated paths and the resulting optimal portfolios. Whereas traditional simulation techniques have focused on the DGP, data-driven optimization does not need assumptions on the DGP and recently machine learning has even offered tools to approximate a DGP using flexible mappings (such as neural networks) and has allowed us to shift the focus of the modeling exercise to the economic priors. It allows one to essentially train a model on historical conditions, while predicting on current conditions. This also gives one leeway for introducing nowcasting timeseries, such as real-time macro data or ESG data into the model.

In this paper we will focus on the US market and build a toy model of the US economy, but this model would be applicable to real-time sentiment, ESG, etc. as well. We collect approximately 100 conditions C_i from the Federal Reserve Economic Data (FRED) database, including credit and monetary data, interest rates, employment, commodity prices, stress indicators, volatility indices, and consumer sentiment. Figure 4 gives an overview of the high-level categories. Table A.2 in Appendix Appendix A gives a list of all the indicators we have considered.

To get a first idea of the most apposite conditions, we look at the total drawdown path of the total US stock market (Wilshire) and do a LASSO[7] regression to select the most relevant (linear) features. This gives the weights displayed in Table 1. As one would expect, the CBOE volatility index dominates the other coefficients. Other stress indicators, such as the FED St. Louis Financial Stress Index [25] and declining Consumer Sentiment as measured by the University of Michigan [26] contribute significantly to general market drawdown. Moreover, time series of manufacturing, export, currency and long-term mortgage rate data were found significant. Only 4 factors had a significant negative impact on the historical market drawdown.

The aim of our analysis is to introduce appropriate C_i to our optimization model, such that we can evaluate $\mathbb{E}(\xi|C_i)$ at the current level of C_i (e.g. $b_j = 1$) as well as for our own scenarios of C_i (e.g. $b_j = 0$, or the other way around).

For instance, given the current level of volatility, what do paths and the optimal portfolio look like, and which positions are most affected if one gradually increases the volatility to levels seen

[7]Least Absolute Shrinkage and Selection Operator, a simple linear regression with a L1-norm penalty on the coefficients, `https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.Lasso`. We used 10-fold cross-validation to find the optimal penalty hyperparameter.

Largest positive contributors to ξ		**Largest negative contributors to ξ**	
CBOE Volatility Index	0,815680	US Gov't Securities at All Com. Banks	-0,142223
Avg Weekly OT Hours: Manufacturing	0,129751	Long Term Unemployment: 27 WKS	-0,043401
Exports to Mexico	0,126585	JPN/USD Currency Exchange Rate	-0,029723
Univ. of Michigan: Consumer Sentiment	0,079161	Avg Hourly Earnings: Manufacturing	-0,001523
St. Louis Financial Stress Index	0,072814		
CNY/USD Currency Exchange Rate	0,068154		
CAD/USD Currency Exchange Rate	0,053743		
Imports from UK	0,038683		
30-yr Conventional Mortgage Rate	0,037272		
Effective Federal Funds Rate	0,029571		

Table 1: Lasso coefficients of C_i to ξ

during the GFC or the Covid-19-induced March 2020 meltdown? What does one's portfolio look like with current market sentiment, and which positions are likely to be first and mostly affected when sentiment turns sour gradually? This is what we conceptually discuss here and numerically investigate in the next section.

As a tool to evaluate changing paths and portfolios to changing conditions, we use Shapley (SHAP) values [3]. Given one set of n_{cond} conditions $C = (C_i)_{i=\{1,\dots,n_{cond}\}}$, an optimal portfolio can be seen as a linear combination w_n^*, for $n \in N$, where the weights reflect some contribution (of risk, return, drawdown) to the optimal portfolio timeseries w^*S.

Given a set of N_p condition sets $C = (C^k)_{k=\{1,\dots,N_p\}}$, each set corresponding to a C that generates $J = n_{cond}$ sequences S_j, each C will thus correspond to a unique portfolio that is optimal over this subset of sequences, i.e. for each k one has a different portfolio. Now we can see the w_k^* as the output, and evaluate the contribution of each condition C_i in C^k to the optimal portfolio. The SHAP values to each w_d^* can then formally be defined as:

$$\Phi_i(w_d^*) = \sum_{S \subset [N_s \setminus \{i\}]} \frac{|S|!(N_s - |S| - 1)!}{N_s!} (w_d^*(S \cup \{i\}) - w_d^*(S))) \tag{6}$$

This is the SHAP Φ_i for condition C_i in terms of the resulting optimal weight w_d^*.

Intuitively, for the N_p optimal portfolios we evaluate all the subsets of S where condition i was not active ($b_j = 0$) and compare with the optimal portfolios where it was $w_d^*(S \cup \{i\})$, or $b_j = 1$. The average contribution of this condition to the optimal weight thus constitutes the SHAP value.

This allows for visualizations of the conditional optimal portfolios, such as waterfall and beeswarm plots ([3] and below), that are popular explainable machine learning tools for applications in deep learning and computer vision. Efficient implementations that rely on local approximation using linear regression can be found on the original Shap repo page. These are so-called kernel Shapley values, which is simply a special weighted linear regression to approximate Eq. (6). We do not reiterate the details here but refer the interested reader to Moehle et al. [10] and Shalit [6].

4. Applications and backtests

4.1. Set-up and data

For a point-in-time DOW universe, we collect 15 years of data from 31th Dec 2007 up until 1st Jan 2023, which includes the 2007-2008 Great Financial Crisis (GFC), and the 2009-2022 bull market (including the 2020 Covid collapse and sharp recovery). In the backtest, portfolios were rebalanced monthly using a recursive historical window of three years with no look-ahead biases.

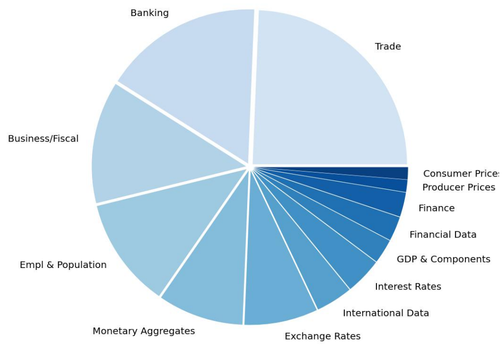

Figure 4: Macro condition high-level categories

4.2. Benchmark models

We include three simple benchmarks: a minimum variance portfolio, an equally weighted portfolio and the market cap weighted index (not computed but directly fetched index levels from the market data provider).

Minimum volatility portfolio. The minimum volatility portfolio is the solution to the following quadratic optimization problem:

$$\begin{aligned} \min_{\mathbf{w}} \quad & \mathbf{w}\mathbf{\Sigma}\mathbf{w}' \\ \text{s.t.} \quad & \mathbf{w}\mathbf{I}^N = 1 \\ & \mathbf{w} \geq 0 \end{aligned} \tag{7}$$

In other words, we pick the portfolio weights that minimize portfolio volatility, subject to making sure that the weights add up to one and a long-only constraint. Compared to traditional mean-variance optimization the returns are implicitly assumed to be symmetric, hence dropped from the objective.

Equally weighted portfolio. The equally weighted portfolio allocates equal proportions to all the assets in the investible universe:

$$\mathbf{w} = (1/N, ..., 1/N) \tag{8}$$

The '1/N'-portfolio minimizes model risk and assumes symmetrical returns, volatilities and correlations.

4.3. Results and discussion

4.3.1. Equity backtest

We perform an equity backtest over the five year 2017-2022 period. Figures B.11, B.12, B.13, B.14, 5, 6 and 7 show the results of the backtest. As from portfolio value point of view, Figure B.11 shows that our data-driven optimizer is on par with the index and an equally weighted portfolio. However, annual returns (Figure B.12) are much more consistent in up and down years, and max and mean drawdowns (Figures B.14 and B.13) are substantially reduced. Note that as expected the minimum drawdown portfolio minimizes the latter. Most remarkably, minimum variance portfolio does not participate in the recovery after the 2020 Covid-induced downfall while minimum drawdown explicitly optimizes for recovery. Therefore, the number of days where a particular drawdown was prevalent (Figure 5) is maybe the most important figure. While a non-optimized market or equally weighted portfolio has many +25 and +30% drawdowns, this is not the case for minimum variance and minimum drawdown portfolio. Minimum variance has low max drawdowns but high moderate drawdowns or mean drawdown because of its slow recovery. Moreover, our data-driven optimizer excels in terms of stability and lack of concentration. The instability of variance-covariance matrices can be brought back to *correlation breakdown*, or episodes where volatility spikes, all correlations go up and the conditioning of a variance-based problem goes south. The latter means that when the average correlation of a correlation matrix goes up, its concentration of eigenvalues (condition number) goes up, and consequently the precision of the inverted covariance matrix goes down. This inverted matrix is required for solving the problem and determines its statistical accuracy, therefore it is called the *precision matrix*. Figure 6 plots the turnover (half the sum of absolute values of change in portfolio weights) of Min Drawdown and Min Variance respectfully, and puts them next to the rolling two-month volatility, the average correlation of the DOW universe and the condition number of the rolling two-month correlation matrix. It is clear that on average Min Drawdown is more stable than Min Variance, and less prone to spikes in the three above-mentioned indicators of so-called correlation breakdown. In terms of concentration Figure 7 shows the maximum weights of the portfolios and their Herfindahl indices (sum of squared weights). Minimum Variance likes to invest in individual low variance instruments and takes on positions up to 40% in individual instruments, while this is less than half for Minimum Drawdown. The effect is even more pronounced in the Herfindahl index. From these figures one can argue that a minimum drawdown portfolio can be constituted by combining instruments into a joint low drawdown path, while minimum variance is often a product of individual low-vol holdings. This concentration clearly has an iteraction effect with stability: the more extreme the portfolios at one point in time, the more extreme the rebalance can be at the next. In summary, minimum variance seems to reduce portfolio risk versus the market cap or equally weighted portfolio in exchange for underperformance, instability and concentration. The minimum drawdown portfolio seems to resolve these issues in exchange for non-interpretable optimal portfolio solutions. There is no parametrized version of the input samples nor hyperparameters, so in order to evaluate the effect of a variable, such as a risk factor or macro indicator, on the resulting optimal portfolio, one needs to resort to explainable AI. In the next section we describe the findings for the methodology outlined in Section 3.

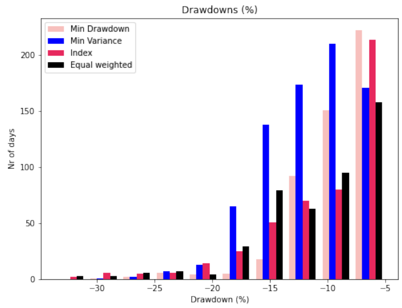

Figure 5: Days in large drawdown

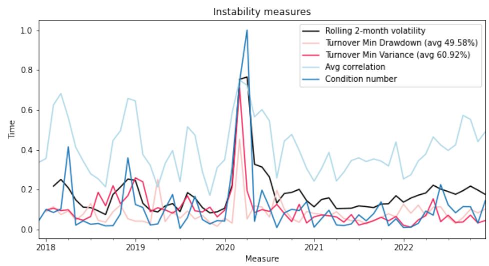

Figure 6: Portfolio instability

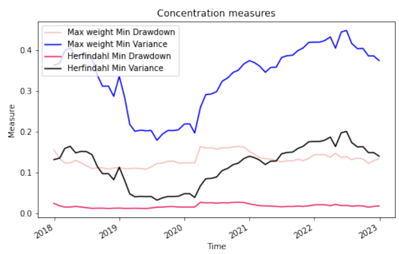

Figure 7: Portfolio concentration

4.3.2. *Portfolio explainability*

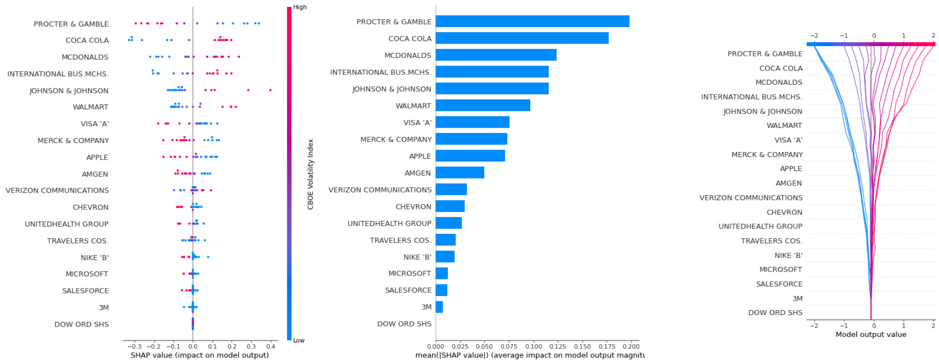

Figure 8: Shapley values for CBOE Volatility index on the optimal minimum drawdown portfolio. Left: beeswarm plot which indicates the impact on the portfolio weights per instrument and the level of VIX as color. Middle: bar chart which is the average absolute value of the average impact of VIX on the position (SHAP-value). Right: a decision chart or the gradual impact on the portfolio weights by increasing/decreasing VIX.

As an illustration we first take the most obvious macro-economic indicator for economic stress: the CBOE Volatility Index (VIX) indicator.

How sensitive are the positions in our portfolio for a new shock in volatility? Figure 8 gives an overview of the results: the instruments with a high Shapley value have a high sensitivity to the VIX assumption. In case the world changes abruptly, such as a global pandemic or a war, when the VIX spikes, the asset manager might want to know which of his positions are likely to be first and most impacted and likely to be bought or sold by the optimizer. Positions with small or close to zero Shap values are positions that were obtained under any conditional historical scenario and are thus expected to be robust over new scenarios of market turmoil[8]. In casu, P&G and Coca-Cola stock are most sensible to volatility spikes as from the beeswarm plot it is clear that these positions have the biggest deviations from their unconditional values. As much as 20% of the position is affected by the VIX condition. Figure 10 shows the full sensitivity analysis. There it is clear that the delta is approximately 10% for both, with with high VIX meaning higher positions in Coca-Cola and lower position in P&G, and vice versa. The opposite can be said for a position like Microsoft or Nike. For both high and low VIX we find similar positions which are thus relatively insensitive to the VIX condition.

This was just one example that we could reiterate for the 100 features of our toy macro model of the US economy. We could now build dashboards that explain the sensitivity of one's current portfolio to macro factors and the current state of the factors. Figure 9 is a simple example. On the left we see the current portfolio (i.e. on the final backtest date), on top we find the ten selected predictors of market drawdown from the LASSO selection. The top rows contain the

[8] As with the ergodicity assumption, this assumes history tells us something about the future. Unfortunately it is all we quants have. And why would one use quant methods if these scenarios do not contain information for the future? In terms of market efficiency, we are not predicting returns but risk, and these sensitivities are essentially drawdown sensitivities to the market conditions. There might be good economic reasons (this is not in the scope of this paper) as to why a low drawdown instrument (like a cashcow safe haven) might be a low drawdown instrument again in a future period of economic distress.

	Current Holdings	Univ of Michigan: Consumer Sentiment	US Treasuries Held by the Fed	US Gov't Securities at All Com. Banks	Long Term Unemployment: >27 WKS	JPN/USD Currency Exchange Rate	Exports to Mexico	Civilian Total Unemployment Rate	CNY/USD Currency Exchange Rate	CBOE Volatility Index	Avg Weekly OT Hours: Manufacturing
Last value		59.7	5435582	4432.808	1069	129.97	25596.236	3.5	6.782	18.51	3.6
Previous		56.8	5436722	4439.887	1215	128.48	27915.721	3.6	6.774	18.73	3.7
Change		5.10%	0.00%	-0.20%	-12.00%	1.20%	-8.30%	-2.80%	0.10%	-1.20%	-2.70%
AMGEN	7.8%	0.052986258	0.052802903	0.064200948	0.077048502	0.056876176	0.064109974	0.03578008	0.030630963	0.046127068	0.034533547
APPLE	0.8%	0.060811318	0.048817234	0.070668376	0.059235012	0.075496029	0.052866806	0.039849038	0.034835337	0.063317381	0.045299441
CHEVRON	5.3%	0.05544251	0.054247714	0.06762688	0.039611005	0.050165616	0.058903097	0.053001852	0.044836211	0.038043243	0.047353964
COCA COLA	5.9%	0.042207028	0.053037406	0.059611334	0.066422289	0.047743565	0.079417614	0.09266578	0.045628382	0.180406373	0.037008844
JOHNSON & JOHNSON	10.2%	0.092732803	0.085386992	0.082653512	0.074167847	0.072164284	0.122412389	0.073350676	0.057214875	0.121289631	0.11907102
MCDONALDS	9.0%	0.021489415	0.018355678	0.042233716	0.059096289	0.065351926	0.041457831	0.030147139	0.028740571	0.119804306	0.049674787
MERCK & COMPANY	6.0%	0.057017213	0.049976877	0.057004532	0.096488593	0.049294659	0.08539876	0.069782638	0.026796147	0.076440345	0.050035384
MICROSOFT	3.8%	0.036461939	0.036479167	0.047138099	0.021903417	0.033883353	0.026542922	0.013347832	0.0471462	0.01249175	0.053637885
NIKE 'B'	2.3%	0.031337197	0.019478244	0.025028338	0.024553256	0.041202481	0.039872278	0.031600402	0.014046177	0.02191613	0.028046689
PROCTER & GAMBLE	13.4%	0.074973713	0.045062375	0.050334948	0.150795901	0.039793467	0.081165363	0.216342725	0.025691296	0.202577083	0.026208027
TRAVELERS COS.	4.4%	0.030693917	0.007588248	0.012600206	0.093026386	0.014307226	0.012643239	0.063997145	0.015702092	0.01291851	0.062671172
UNITEDHEALTH GROUP	3.5%	0.036663961	0.028029084	0.04359054	0.027339529	0.018009922	0.041863718	0.03981607	0.018623258	0.028587745	0.035476698
VERIZON COMMUNICATIONS	8.9%	0.027278608	0.02601565	0.02716867	0.071636266	0.043777217	0.032751525	0.045683405	0.019457459	0.032208955	0.033845752
VISA 'A'	0.1%	0.072777085	0.052815853	0.060736867	0.058056613	0.040564638	0.073538785	0.067579037	0.038748602	0.074601816	0.051777059
WALMART	8.4%	0.0246824	0.01858237	0.053806642	0.069252639	0.041596232	0.054871841	0.047000199	0.03978451	0.098605233	0.047263486

Figure 9: Macro dashboard

last observed value of these indicators, the value before that, and thus the observed last change. The remainder of the heatmap are the Shapley values. For instance, we see a decrease in US long-term unemployment at the end of 2022, where the P&G position is most sensitive to such a shock. The second to last column contains the CBOE Volatility index Shapley values we know from Figure 8. We find that the last update was a small decrease in volatility, such that we do not expect the sensitive Coca-Cola and P&G positions to move much.

Such dashboards could be synchronized with the portfolio manager's economic calendar such that s/he can anticipate the effects on the data-driven optimizer beforehand. Moreover, the changes in conditions should not only be considered a given by the market, but could also be based on stresstesting values or views by the asset manager, and risk factors and nowcasting timeseries such as environmental, social, governance and news sentiment data could be included. Therefore, although the model seemed initially not interpretable, its value lies in the eyes of its beholder.

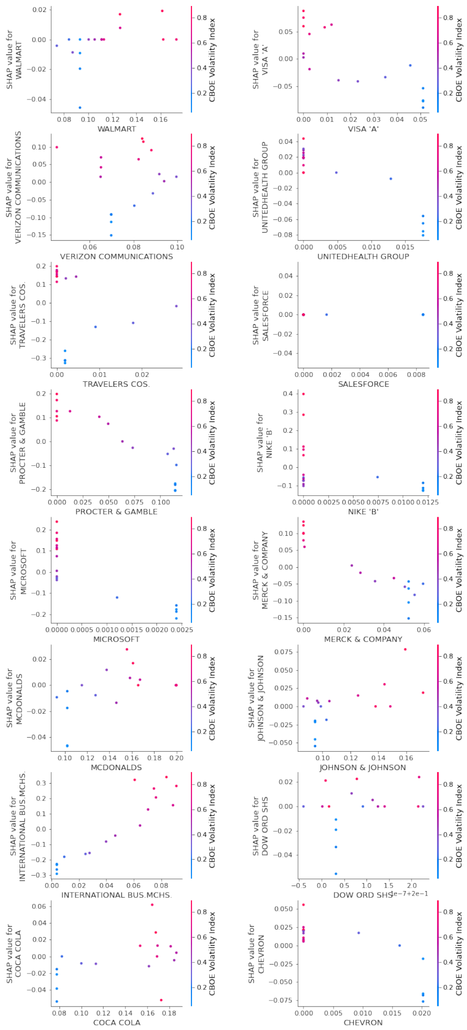

Figure 10: Sensitivity to CBOE Volatility of optimal minimum drawdown portfolio positions

5. Conclusion

A portfolio manager's job is mainly to deal with the uncertainty of markets. The stochasticity of markets is different than e.g. random processes in physics. In physics, if a DGP is unrealistic, one can add complexity, i.e. more mathematics, to make it more realistic. In finance, extra complexity often creates more room for error, as there are no theories and only models. Simpler often means better as long as it is not too simplistic. This has been one of the longstanding under- or overtones in the quantitative financial literature. Doing exact derivations of optimal allocations under wrong DGP assumptions has dominated the literature since Markowitz introduced the mean-variance paradigm as Modern Portfolio Theory in 1952. Therefore, people have focused on combatting its limitations: specification issues and overfitting, concentration, instability, etc.

Another way of looking at the allocation problem is not focusing on the DGP but doing optimization directly on unparametrized versions of the data, which we called data-driven models. Data-driven models potentially have the advantage of better accommodating the implicit DGP, lack of concentration, stability, and more. However, they are often called complex and black boxes, which does not fit the usual paradigm that, if one deals with a peculiar kind of uncertainty like financial markets, simpler is better than complex.

This paper essentially uses a combination of very simple and established methods, albeit in a data-driven way. We propose one specific data-driven optimizer, the minimum drawdown portfolio which is an objective that is imposed directly on input samples. We discussed the implications of the conditional expectation of portfolio drawdown for a conditional bootstrap setup. Next, we were the first to link such a bootstrap procedure conditional on macro economic data to the resulting optimal portfolio outcomes using the concept of Shapley values. Shapley values are a universal approach to estimate the contributions of individual features to an outcome function. Through this technique one can assess the sensitivities of individual positions to shocks in exogenous macro economic factors. Moreover, one can build holistic dashboards that display the sensitivities and changes in a host of factors. In addition, one could go from a low-frequency macro approach into a nowcasting approach with higher frequency proxies of these factors and also include environmental, social and governance factors or news.

Future work will include this analysis, synchronize our approach with the economic clock and evaluate whether this could be conducive to more effective asset management. Moreover, assessing the portfolio's sensitivities to news or ESG data is a popular topic that could be considered from this angle.

Appendix A. Macro conditions details

Table A.2: Macro-economic conditions

ID	FRED ID	FRED Cat.	Detailed Cat.	Indicator
0	TREAST	Finance	Monetary Data	US Treasuries Held by the Fed
1	MBST	Finance	Monetary Data	Mortgage Backed Sec Held by the Fed
2	WALCL	Banking	Monetary Factors	All Fed Reserve Banks - Total Assets
3	TLAACBW027SBOG	Banking	Monetary Factors	All Commercial Banks - Total Assets
4	BOPBCA	Banking	Conditions	Number of US Banks
5	USNUM	Banking	Conditions	Number of US Commercial Banks
6	EQTA	Banking	Conditions	Equity/Asset Ratio
7	TOTBKCR	Banking	Commercial Credit	Bank Credit of All Commercial Banks
8	TOTALSEC	Banking	Commercial Credit	Securitized Total Consumer Loans
9	TOTALSL	Banking	Commercial Credit	Total Consumer Credit Outstanding
10	INVEST	Banking	Investment	Total Investments All Commercial Banks
11	USGSEC	Banking	Investment	US Gov't Securities at All Com. Banks
12	CONSUMER	Banking	Loans	Total Consumer Loans
13	BUSLOANS	Banking	Loans	Total Commercial/Industrial Loans
14	DALLCACBEP	Banking	Delinquencies	Delinquencies On All Loans And Leases
15	T10Y2Y	Banking	Interest Rates	US 10-YR / 2-YR Spread
16	TB3MS	Banking	Interest Rates	3-Month T-Bill: Secondary Market Rate
17	DGS10	Banking	Interest Rates	10-Yr Treasury Const. Maturity Rate
18	GFDEBTN	Business/Fiscal	Federal Government	Federal Government Debt (Public)
19	FYOINT	Business/Fiscal	Federal Government	Interest on National Debt
20	FYONET	Business/Fiscal	Federal Government	Federal Spending
21	FYFR	Business/Fiscal	Federal Government	Federal Receipts
22	FYFSD	Business/Fiscal	Federal Government	Budget Deficit/Surplus
23	CDSP	Business/Fiscal	Household Sector	Consumer Debt/Income Ratio
24	PERMIT	Business/Fiscal	Household Sector	New Home Permits
25	HSN1F	Business/Fiscal	Household Sector	New Home Sales
26	CMDEBT	Business/Fiscal	Household Sector	Outstanding Mortgage Debt
27	DGORDER	Business/Fiscal	Ind. Production	Manufacturers' New Orders
28	TCU	Business/Fiscal	Ind. Production	Capacity Utilization: Total Industry
29	TTLCONS	Business/Fiscal	Construction	Total Construction Spending
30	BUSINV	Business/Fiscal	Other	Total Business Inventories
31	ALTSALES	Business/Fiscal	Other	Light Weight Vehicle Sales
32	UMCSENT	Business/Fiscal	Other	Univ of Michigan: Consumer Sentiment
33	STLFSI	Business/Fiscal	Other	St. Louis Financial Stress Index
34	OILPRICE	Business/Fiscal	Other	Spot Oil Price - West Texas Intermediate
35	CPIAUCSL	Consumer Prices	CPI	Consumer Price Index: Seasonally Adj.
36	UNRATE	Empl & Population	Household Survey	Civilian Total Unemployment Rate
37	UEMP27OV	Empl & Population	Household Survey	Long Term Unemployment: 27 WKS
38	UEMPMED	Empl & Population	Household Survey	Length of Unemployment
39	CE16OV	Empl & Population	Household Survey	Total US Workforce
40	EMRATIO	Empl & Population	Household Survey	US Employment/Population Ratio
41	POP	Empl & Population	Population	US Population
42	AHEMAN	Empl & Population	Est. Survey	Avg Hourly Earnings: Manufacturing
43	AWHMAN	Empl & Population	Est. Survey	Avg Weekly Hours: Manufacturing
44	AWOTMAN	Empl & Population	Est. Survey	Avg Weekly OT Hours: Manufacturing
45	DEXUSUK	Exchange Rates	Daily Rates	USD/GBP Currency Exchange Rate
46	DEXUSEU	Exchange Rates	Daily Rates	USD/EUR Currency Exchange Rate
47	DEXJPUS	Exchange Rates	Daily Rates	JPN/USD Currency Exchange Rate
48	DEXMXUS	Exchange Rates	Daily Rates	MXP/USD Currency Exchange Rate
49	DEXCAUS	Exchange Rates	Daily Rates	CAD/USD Currency Exchange Rate
50	DEXCHUS	Exchange Rates	Daily Rates	CNY/USD Currency Exchange Rate
51	COMPOUT	Financial Data	Monetary	Commercial Paper Outstanding
52	VIXCLS	Financial Data	Volatility Indexes	CBOE Volatility Index
53	GDP	GDP & Components	GDP/GNP	US Gross Domestic Product
54	GNP	GDP & Components	GDP/GNP	US Gross National Product
55	NETFI	GDP & Components	Imports & Exports	US Current Account Balance
56	EXPGS	GDP & Components	Imports & Exports	US Exports Goods & Services
57	IMPGS	GDP & Components	Imports & Exports	US Imports Goods & Services
58	DGI	GDP & Components	Govt Accounting	Fed Govt: Defense Budget
59	FGRECPT	GDP & Components	Govt Accounting	Fed Govt: Tax Receipts
60	TGDEF	GDP & Components	Govt Accounting	Fed Govt: Budget Deficit
61	CP	GDP & Components	Industry	Corporate Profits After Tax
62	DIVIDEND	GDP & Components	Industry	Corporate Dividends
63	PI	GDP & Components	Personal	Personal Income
64	PSAVE	GDP & Components	Savings & Inv.	Personal Savings
65	PSAVERT	GDP & Components	Savings & Inv.	Personal Savings Rate
66	MORTGAGE30US	Interest Rates	30yr Mortgage	30-yr Conventional Mortgage Rate
67	DPCREDIT	Interest Rates	FRB Rates	Discount Rate
68	FEDFUNDS	Interest Rates	FRB Rates	Effective Federal Funds Rate
69	GRCPROINDMISMEI	International Data	Indicators	Production of Total Industry in Greece
70	GRCSARTMISMEI	International Data	Indicators	Total Retail Trade in Greece
71	GRCURHARMMDSMEI	International Data	Indicators	Unemployment Rate - Greece
72	M1	Monetary Aggregates	M1	M1 Money Supply
73	M2	Monetary Aggregates	M2	M2 Money Supply
74	MZM	Monetary Aggregates	MZM	MZM Money Supply
75	M1V	Monetary Aggregates	M1	Velocity of M1 Money Stock
76	M2V	Monetary Aggregates	M2	Velocity of M2 Money Stock
77	MZMV	Monetary Aggregates	MZM	Velocity of MZM Money Stock
78	MULT	Monetary Aggregates	M1	M1 Money Multiplier
79	PPIACO	Producer Prices	PPI	Producer Price Index: All Commodities
80	IMPCH	Trade	Imports	Imports from China
81	IMPJP	Trade	Imports	Imports from Japan
82	IMPMX	Trade	Imports	Imports from Mexico
83	IMPCA	Trade	Imports	Imports from Canada
84	IMPGE	Trade	Imports	Imports from Germany
85	IMPUK	Trade	Imports	Imports from UK
86	EXPCH	Trade	Exports	Exports to China
87	EXPJP	Trade	Exports	Exports to Japan

Continued on next page

Table A.2 – continued from previous page

ID	FRED ID	FRED Cat.	Detailed Cat.	Indicator
88	EXPMX	Trade	Exports	Exports to Mexico
89	EXPCA	Trade	Exports	Exports to Canada
90	EXPGE	Trade	Exports	Exports to Germany
91	EXPUK	Trade	Exports	Exports to UK
92	BOPGEXP	Trade	Exports	Exports: Goods
93	BOPGIMP	Trade	Imports	Imports: Goods
94	BOPGTB	Trade	Balance	Balance: Goods
95	EXPGS	Trade	Exports	Exports: Services
96	BOPSIMP	Trade	Imports	Imports: Services
97	BOPSTB	Trade	Balance	Balance: Services
98	BOPGSTB	Trade	Balance	Balance: Goods & Services

Appendix B. Backtest figures

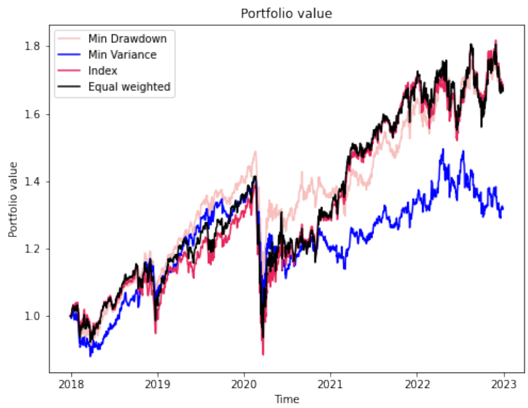

Figure B.11: Portfolio values backtest

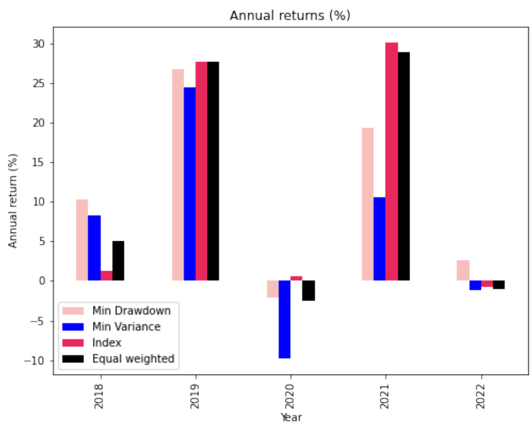

Figure B.12: Returns backtest

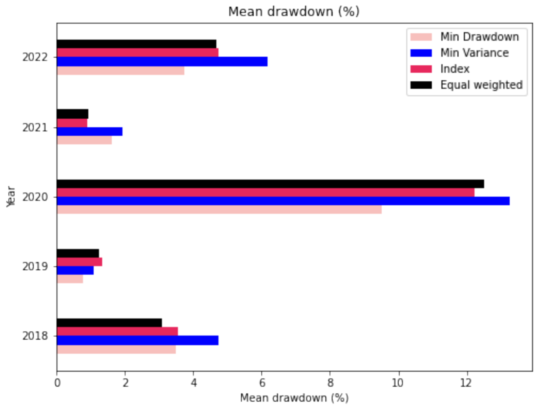

Figure B.13: Mean drawdowns backtest

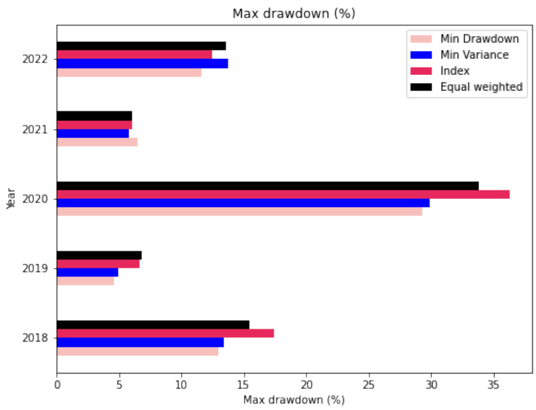

Figure B.14: Max drawdowns backtest

References

[1] D. V. Hinkley, Bootstrap methods, Journal of the Royal Statistical Society: Series B (Methodological) 50 (1988) 321–337.

[2] L. S. Shapley, Notes on the N-person Game, Rand Corporation, 1951.

[3] S. M. Lundberg, S.-I. Lee, A unified approach to interpreting model predictions, in: Proceedings of the 31st international conference on neural information processing systems, pp. 4768–4777.

[4] P. S. Hagan, A. Lesniewski, G. E. Skoufis, D. E. Woodward, Portfolio risk allocation through shapley value, arXiv preprint arXiv:2103.05453 (2021).

[5] R. Colini-Baldeschi, M. Scarsini, S. Vaccari, Variance allocation and shapley value, Methodology and Computing in Applied Probability 20 (2018) 919–933.

[6] H. Shalit, The shapley value of regression portfolios, Journal of Asset Management 21 (2020) 506–512.

[7] H. Shalit, Using the shapley value of stocks as systematic risk, The Journal of Risk Finance 21 (2020) 459–468.

[8] H. Shalit, The shapley value decomposition of optimal portfolios, Annals of Finance 17 (2021) 1–25.

[9] N. Tarashev, K. Tsatsaronis, C. Borio, Risk attribution using the shapley value: Methodology and policy applications, Review of Finance 20 (2016) 1189–1213.

[10] N. Moehle, S. Boyd, A. Ang, Portfolio performance attribution via shapley value, arXiv preprint arXiv:2102.05799 (2021).

[11] J. Simonian, Portfolio selection: a game-theoretic approach, The Journal of Portfolio Management 45 (2019) 108–116.

[12] G. Babaei, P. Giudici, E. Raffinetti, Explainable artificial intelligence for crypto asset allocation, Finance Research Letters 47 (2022) 102941.

[13] J. J. Ohana, S. Ohana, E. Benhamou, D. Saltiel, B. Guez, Explainable ai (xai) models applied to the multi-agent environment of financial markets, in: Explainable and Transparent AI and Multi-Agent Systems: Third International Workshop, EXTRAAMAS 2021, Virtual Event, May 3–7, 2021, Revised Selected Papers 3, Springer, pp. 189–207.

[14] E. Benhamou, J.-J. Ohana, D. Saltiel, B. Guez, Explainable ai (xai) models applied to planning in financial markets (2021).

[15] K. Kimura, K. Schwaiger, D. Sharma, A. Ang, Factors with style, The Journal of Investing 30 (2021) 21–46.

[16] M. Jaeger, S. Krügel, D. Marinelli, J. Papenbrock, P. Schwendner, Interpretable machine learning for diversified portfolio construction, The Journal of Financial Data Science 3 (2021) 31–51.

[17] A. Chekhlov, S. Uryasev, M. Zabarankin, Drawdown measure in portfolio optimization, International Journal of Theoretical and Applied Finance 8 (2005) 13–58.

[18] A. Chekhlov, S. Uryasev, M. Zabarankin, Portfolio optimization with drawdown constraints, in: Supply chain and finance, World Scientific, 2004, pp. 209–228.

[19] L. R. Goldberg, O. Mahmoud, On a convex measure of drawdown risk, Archive preprint archive 1404 (2014).

[20] L. R. Goldberg, O. Mahmoud, Drawdown: from practice to theory and back again, Mathematics and Financial Economics 11 (2017) 275–297.

[21] O. Peters, The ergodicity problem in economics, Nature Physics 15 (2019) 1216–1221.

[22] N. N. Taleb, Fooled by randomness: The hidden role of chance in life and in the markets, volume 1, Random House Trade Paperbacks, 2005.

[23] J. James, L. Yang, Stop-losses, maximum drawdown-at-risk and replicating financial time series with the stationary bootstrap, Quantitative Finance 10 (2010) 1–12.

[24] O. Van Hemert, M. Ganz, C. R. Harvey, S. Rattray, E. S. Martin, D. Yawitch, Drawdowns, The Journal of Portfolio Management 46 (2020) 34–50.

[25] K. L. Kliesen, D. C. Smith, et al., Measuring financial market stress, economic synopses (2010).

[26] R. T. Curtin, Indicators of consumer behavior: The university of michigan surveys of consumers, Public Opinion Quarterly 46 (1982) 340–352.