
Hierarchical Portfolio Stress-Testing with Flexible Probabilities



UNIVERSITY OF CAPE TOWN

A research dissertation submitted in fulfillment of the requirements for the degree of
MASTERS OF SCIENCE IN MATHEMATICAL STATISTICS
in the

DEPARTMENT OF STATISTICAL SCIENCES

AUTHOR: Nina Matthews

SUPERVISOR: A/Prof. Tim Gebbie

November 6, 2025

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

Declaration of Authorship

I, Nina Matthews, declare that this project titled, “Hierarchical Portfolio Stress-Testing with Flexible Probabilities” and the work presented in it is my own.

I confirm that:

- ▶ This work was done wholly or mainly while in candidature for a Masters degree at the University of Cape Town.
- ▶ Where any part of this dissertation has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- ▶ Where I have consulted the published work of others, this is always clearly attributed.
- ▶ Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this dissertation is entirely my own work.
- ▶ I have acknowledged all main sources of help.

Name: Nina Elizabeth Matthews

Date: 06/11/2025

Signed:

Signed by candidate

Reproducible Research

The MATLAB files used to produce the results in this work along with the data can be found in the public GitHub repository: [Meucci-HS-FP-framework-SA-context](https://github.com/nematthews/Meucci-HS-FP-framework-SA-context) [75]. Details of the various functions are provided in the attached Appendix, along with specific information on their usage within the individual files.

To access:

<https://github.com/nematthews/Meucci-HS-FP-framework-SA-context>

Files that were used to produce selective plots in this proposal can be found in the public GitHub repository created for a 2021 honours thesis project: [Honours-Thesis-Cluster-Analysis-for-Portfolio-Selection](https://github.com/nematthews/Honours-Thesis-Cluster-Analysis-for-Portfolio-Selection) [77].

To access:

<https://github.com/nematthews/Honours-Thesis-Cluster-Analysis-for-Portfolio-Selection>

UNIVERSITY OF CAPE TOWN

Abstract

Department of Statistical Sciences

Master of Science

Hierarchical Portfolio Stress-Testing with Flexible Probabilities

by NINA MATTHEWS

Traditionally, portfolio optimisation has heavily relied on historical data to predict future outcomes, with the embedded assumption that historical distributional properties persist over time. However, observations have shown that many of these properties are not necessarily preserved within financial time series. This research challenges these assumptions by adopting a non-parametric, scenario-based approach using the Historical Simulation with Flexible Probabilities (HS-FP) framework (Meucci [89]). The study extends this framework by integrating the Hierarchical Risk Parity (HRP) strategy (De Prado [29]) and the Mean-Variance (MV) approach (Markowitz [69]) to enhance portfolio resilience within a South African financial market context. The HS-FP methodology quantifies future market risk by mapping drivers into state- and time-conditioned probabilities, generating forward-looking estimates of returns and volatilities. This study reveals that while the Sharpe-Ratio maximising MV portfolio demonstrated stronger cumulative performance, HRP achieved superior risk-adjusted returns, highlighting its robust diversification and resilience to extreme market volatility. The integration of the HS-FP framework, while not consistently improving overall performance, illuminated the sensitivity of traditional methods to underlying assumptions and underscored the value of incorporating market regime considerations. The findings advocate for adaptive approaches that move away from rigid assumptions, proposing a systematic, scenario-based approach to regime capturing in order to enhance the real-world applicability of stress tests.

Contents

1	Introduction	1
1.1	Context: time series analysis	1
1.1.1	Historical development of TSA: probability theory and stochastic processes	3
1.1.2	Market dynamics: systemic and idiosyncratic shocks	8
1.1.3	Quantitative portfolio construction: a cautionary tale	10
1.1.3.1	The selection problem: stress-testing	10
1.1.3.2	Structuring a quantitative portfolio process	12
1.2	Fundamentals: assumptions made for portfolio theory	14
1.2.1	Statistical considerations	14
1.2.1.1	Independently and identically distributed	14
1.2.1.2	The Law of Large Numbers	15
1.2.1.3	Normality via the Central Limit Theorem	16
1.2.1.4	Stationarity	17
1.2.2	Financial data considerations	17
1.2.2.1	Returns	18
1.2.2.2	Return Averages	19
1.2.2.3	Volatility as Measures of Risk	20
1.2.2.4	Sharpe Ratio	22
1.2.3	Practical considerations	24
1.3	Relevance of research	25
1.4	Objective and structure of dissertation	26
2	Conventional approach & practical implementation	28
2.1	Portfolio optimisations	28
2.1.1	Comparison Mean-Variance & Hierarchical Risk Parity	28
2.1.2	Sharpe Ratio maximising Mean-Variance optimisation	31
2.1.3	Hierarchical Risk Parity	32
2.1.3.1	Step 1: tree clustering	32
2.1.3.2	Step 2: quasi-diagonalization	34
2.1.3.3	Step 3: recursive bisection	34
2.2	Additional performance measures for advanced comparisons	36
2.2.1	Extending the Sharpe Ratio	36
2.2.1.1	Probabilistic Sharpe Ratio	38
2.2.1.2	Deflated Sharpe Ratio	39

2.2.2	Further metrics	44
2.3	Simulation design	44
2.3.1	Asset class data selection	44
2.3.1.1	Equity	44
2.3.1.2	Bonds	46
2.3.1.3	Cash	46
2.3.2	Data processing & exploration	47
2.3.3	Benchmarks	49
2.3.4	Dynamic back-test design	50
2.3.4.1	Capturing the effect of time	51
2.4	Preliminary results	51
2.4.1	Assessing inputs	52
2.4.2	Assessing structural concerns	54
2.4.2.1	Turnover constraint	55
2.4.2.2	Cash constraint	55
2.4.2.3	Underlying data concerns	55
2.4.2.4	Comparison of structural “fixes”	59
2.5	Model selection	62
2.5.1	The Deflated Sharpe Ratio and multiple trials	62
2.5.2	Results	64
2.5.2.1	Implication of results	68
2.6	Summary	68
3	Forward looking simulations	69
3.1	Intro to scenario analysis	70
3.2	Mapping drivers into the market model: the HS-FP framework	70
3.2.1	Risk drivers	72
3.2.2	Invariants	73
3.2.2.1	Testing for invariance	73
3.2.3	Forward looking distribution	76
3.3	Market state signal preparation	77
3.3.1	Market state variable selection	77
3.3.2	Market signal processing	78
3.4	Flexible probabilities methods	80
3.4.1	Time conditioned	81
3.4.1.1	Rolling window	81
3.4.1.2	Exponential decay	81
3.4.2	Single market state conditioned	82
3.4.2.1	Crisp probabilities	83
3.4.2.2	Kernel smoothing	83
3.4.3	Joint time & state conditioning	84
3.4.3.1	Entropy pooling	84
3.4.4	Multi-market state & time conditioned	87

3.4.4.1	Equally weighted ensemble	87
3.4.4.2	Conditional Bayesian ensemble	87
3.5	Modelling forward-looking estimates of risk drivers	89
3.6	Application	92
3.7	Results	94
3.7.1	Base model and HS-FP methods comparison	94
3.7.2	MV <i>vs.</i> HRP comparison within methods	97
3.7.3	HS-FP signals comparison within methods	99
3.7.4	Case studies exploring underlying simulation dynamics	101
3.8	Results discussion	106
4	Concluding remarks	109
Appendix		115
A1.	Algorithms	115
A1.1	Double exponential decay	115
A2.	Proofs	117
A2.1	Finite-ensemble average and the expected values	117
A2.2	Sharpe Ratio–Maximising Mean–Variance Portfolio with a budget constraint	118
A2.3	Inverse-variance allocation	121
A3.	Additional simulation designs	122
A3.1	Benchmark portfolios for overlapping rolling window simulations	122
A3.2	Comparative back-test: growing window simulation	122
A4.	Data engineering	123
A5.	Reproducible research	126
A5.1	Systems specification	127
A5.2	Data management	127
A5.3	Data structures	127
A5.4	Toolbox construction	128
A5.4.1	Class construction	129
A5.4.2	Class properties	129
A5.4.3	Class user example	132
A5.5	Test case documentation	133
A5.6	Sourcing data: a step-by-step guide to Bloomberg	133
A5.7	Sourcing bond data: additional steps needed	134
A5.7.1	Accessing additional bonds data through Bloomberg	135
A5.7.2	Accessing additional bonds data: contacting the JSE	136
A6.	Additional Software Documentation	138
A6.1	Multiple simulation test cases	138
A6.2	Code base file structure and descriptions	140
A6.2.1	Class definition files	140
A6.2.2	Toolbox definition files	141
	References	145

List of Abbreviations

ABSA	Amalgamated Banks of South Africa Limited
ALBI	All Bond Index
ALSI	All Share Index
API	Application Programming Interface
BB	Bloomberg
BEASSA	Bond Exchange of South Africa
BH	Buy-Hold portfolio
BL	Black and Litterman
CDF	Cumulative Distribution Function
CILI	Composite Inflation-Linked Index
CLA	Critical Line Algorithm
CM	Constant Mix portfolio
CPU	Central Processing Unit
CSCV	Combinatorially Symmetric Cross Validation
DLI	Domestic Leading Indicator
DSR	Deflated Sharpe Ratio
EDA	Exploratory data analysis
ENS	Effective Number of Scenarios
ESG	Environmental, Social, and (Corporate) Governance
ETF	Exchange Traded Fund
EVT	Extreme Value Theory
EWMA	Exponentially Weighted Moving Average
FI	Fixed Income
FIW	Fixed Income Worksheet on Bloomberg
FPrs	Flexible Probabilities
FSAP	Financial Sector Assessment Program
FTSE	The Financial Times Stock Exchange
GFC	Global Financial Crisis
GPD	Generalised Pareto Distribution
GPU	Graphics Processing Unit
HP	Historical price table on Bloomberg terminal
HRP	Hierarchical Risk Parity
HS-FP	Historical Simulations with Flexible Probabilities
ICB	Industry Classification Benchmark
<i>i.i.d.</i>	Independent and Identically Distributed
IMF	International Monetary Fund
IS	In-sample
JIBAR	Johannesburg Interbank Average Rate
JIBA3M	3 month Johannesburg Interbank Average Rate
JIT	Just-In-Time (compiler)
JSE	Johannesburg Stock Exchange
LLN	Law of Large Numbers
MPT	Modern Portfolio Theory
MTM	Mark To Market
MV	Mean-Variance Optimisation
OOS	Out-of-sample
PBO	Probability of Back-test Overfitting
PCH	Percentage price change
PSR	Probabilistic Sharpe Ratio
PX_LAST	Last traded price for a security on Bloomberg
QE	Quantitative Easing
RV	Random Variable
SR	Sharpe Ratio
SRCH	Search function on Bloomberg terminal
SQP	Sequential Quadratic Programming
TAA	Tactical Asset Allocation
TRI	Total Return Index
UCT	University of Cape Town
USD	United States Dollar
YoY	Year-on-year
ZAR	South African Rand

Notation

Variable	Definition
Chapter 1 & 2	
t	generic time point within a time series at a regular time interval (defined as monthly from Chp.2 onward)
J	number of instruments in the investment universe being considered for optimisation
T	terminal time point of data. In the context of back-testing - time point at which allocation decisions are made
τ	an investors investment horizon
$\mathbf{P}_t$	vector of closing price of all J instruments in the given market at a generic time t
$p_{j,t}$	price of the j -th instrument at time $t(t = 1, \dots, T)$
σ_P	standard deviation of the portfolio returns (estimation denoted as $\hat{\sigma}_P$)
$\boldsymbol{\omega}$	portfolio weight vector for a given strategy (optimal weights given by $\boldsymbol{\omega}^*$)
R_P	portfolio return
R_f	risk-free rate of return
ρ	$J \times J$ correlation matrix of the J instruments
$\rho_{i,k} = \rho[X_i, X_k]$	correlation between the i -th and k -th asset for $i, k = 1, \dots, J$
m	number of clusters formed by the HRP method $m = 1, \dots, (J - 1)$
L	sorted list of instruments created in Step 2 of HRP
L_0	parent cluster made of initial sorted list L
$\tilde{V}_i^{(k)}$	variance of the k -th sub-cluster of parent cluster L_i
$\Sigma_i^{(k)}$	covariance between the constituents of the k -th sub-cluster in the i -th parent cluster
$\tilde{\boldsymbol{\omega}}_i^{(k)}$	vector of weights generated by HRP for the constituents in k -th sub-cluster in the i -th parent cluster
SR	Sharpe Ratio (estimation denoted as $\widehat{SR}$)
$\widehat{PSR}$	estimated Probabilistic Share Ratio
SR^*	hypothetical predefined benchmark SR used to calculate $\widehat{PSR}$
$\widehat{DSR}$	estimated Deflated Share Ratio
M	total number of simulated trials (strategy configurations) generated for the DSR
K	estimated number of independent simulated trials
s_{Winsor}	Winsorise Threshold stated in number of sample standard deviations
λ_{Reg}	covariance Ridge regularisation parameter
Chapter 3	
$\mathbf{X}_t$	vector of D risk driver at time $t(t = 1, \dots, T)$
D	number of risk drivers
$\boldsymbol{\varepsilon}_t$	matrix of V invariant
V	number of invariants
$\mathcal{T}$	Effective number of scenarios
λ	rolling window length for HS-FP time conditioned method
τ	half-life for HS-FP exponential decay
$\mathbf{S}_t$	vector of N state variables at time $t(t = 1, \dots, T)$
N	number of state variables
$\mathbf{Z}_t$	vector of Q state signals at time $t(t = 1, \dots, T)$
Q	number of state signals
$\mathbf{p}_t$	vector of flexible probabilities at time $t(t = 1, \dots, T)$
$\mathbf{r}_t$	vector of returns at time $t(t = 1, \dots, T)$
z^*	target value for HS-FP crisp method
α	probability for symmetric band around z^* for the HS-FP Crisp method
γ	kernel used for the Kernel smoothing HS-FP method where $\gamma = 1$ for an exponential kernel, $\gamma = 1$ for a Guassian kernel
h	bandwidth value for Kernel smoothing HS-FP method

Chapter 1

Introduction

“Science progresses best when observations force us to alter our preconceptions.” – Vera Rubin [32]. The problem is that Modern Portfolio Theory (MPT) is built on the back of assumptions that rarely hold in reality. We aim to address some of these assumptions to bolster the trust investors have in the outcomes of their analysis.

In this introduction, we establish an understanding of the assumptions underlying MPT, detailing how its reliance on historical averages and the presumption of stable, well-behaved distributions may not hold when confronted with evolving market conditions. We experiment with an example that uses South African (SA) market data to examine how these limitations intersect with non-ergodic behaviour and regime shifts in financial time series—particularly when using popular optimisations like Mean Variance (MV) theory—highlighting the challenges they pose for traditional portfolio optimisation. We then outline these key assumptions and their theory to emphasise the implications for model reliability and predictive accuracy.

This lays the groundwork for investigating scenario-based approaches, which offer a more adaptive framework capable of capturing the complexities of market behaviour and acknowledge that real-world data rarely conforms to neat statistical assumptions. Ultimately, each concept discussed, brings us closer to our overarching objective: reducing generalisation errors and enhancing the robustness of portfolio construction methods in dynamic, ever-changing financial environments.

1.1 Context: time series analysis

At the heart of the MPT sits the intersection between financial time series data and the optimisation of investment portfolios. The evolution of the theoretical frameworks used to explain financial markets and investment theory was driven by early preconceptions of dynamic economic environments, advancements in mathematics and statistics, and early forms of Time Series Analysis (TSA). Traditionally, the optimisation of these portfolios relied heavily on historical data to predict future outcomes. Embedded is the assumption that historical distributional properties such as means and variances can be used as reliable estimators for future returns and risk, yet when we let our preconceptions give way to what can actually be observed, these properties are not necessarily preserved within financial time series.

One of the fundamental assumptions is the ergodic hypothesis established in 1930 by George Birkhoff [16]. It assumes that the ensemble average over the accessible states of the system of an observable *e.g.* the expected value or mean, is equivalent to the times series average along a particular path taken by a process through time. However, financial time series data has the tendency to be non-ergodic, meaning the expected value across the ensemble in cross section does not always equate to the time average down a path due to the nature of system dynamics *e.g.* the geometric compounding of wealth, or the measurement process itself [46, 104]. This can be thought of as the case where the average over multiple paths, all sampled at the same time—many different realisations from random scenarios—are not always equivalent to the average outcome along a particular single observed sample path.

Importantly, as Peters [103] pointed out, this only holds true (at best) for equilibrium systems. In practice, complications are encountered when we merely have access to finite historical samples. For example, this misunderstanding appears to be a key component for why the MV portfolio of Markowitz [69] has been incorrectly challenged as an appropriate portfolio selection approach in asset management. Yet, the Markowitz [69] theory is careful to be clear that forward varying return and risk expectations are prudent, and that scenarios need to be reasonably independent, representative of possible futures; this implies that how forward varying return and risk are computed is crucial for the meaningfulness of the theoretical framework to be robust.

This also implies that naively relying on the Kolmogorov formulation of probability theory—where all possible future outcomes are reliably encoded and probabilistically weighted (by fiat) in the construction of the underlying probability space for estimation—can also be problematic [33]. Naively using historic data—perhaps with arithmetic means (that are not ergodic measures) and covariance estimates that are based on data that has not been statistically cleaned to ensure realisations of scenarios are sufficiently *i.i.d.* (due to incorrect treatment of noise or missing measurements and other lead-lag sampling artifacts, or simply just using raw historic data)—has resulted in claims of the need to modify the MV theory, rather than thinking more carefully about the data-science surrounding the calibration, testing, and deployment of the theory itself.

Peters and Adamou [104] described ergodicity in financial times series analysis as “*the assumption that the distribution (of the financial time series) converges in the long run to a unique and stable form, regardless of initial conditions.*”. By assuming ergodicity it allows one to say that if convergence were to take place, then the (finite historical) time series average of the observable path would equate to its ensemble average at any time point with respect to the stable distribution—thereby making the historical time average a reliable estimate of future expected values. This simplification is convenient for analysis purposes allowing classical methods (such as those that the MV theory is based on) to be applied but the question should be asked: “If a stable distribution does exist, how long would it take to converge?”. As Atkinson [4] aptly put it “*the speed of convergence makes a great deal of difference to the way in which we think about the model*”. Convergence may occur in an infinitely distant future but that is no use if we are optimising a portfolio over an investment horizon of five years. This naive use of ergodicity and historical data is attributed to the history of probability theory and stochastic processes.

1.1.1 Historical development of TSA: probability theory and stochastic processes

Probabilistic statistical reasoning was originally designed in cases where time was frozen.¹ Concepts such as frequency distributions, correlations, least squares regression, and the probabilistic practice of comparing deviations from the mean were carried out in studies in which the history of the system was irrelevant [55]. This is because the foundation of probabilistic statistical reasoning was introduced in the 17th century when the impact of time was still thought to be random and unpredictable.

Pascal and De Fermat pioneered probability theory when they introduced the notion of the Expected Value in 1654 [55, 104]. Their contribution was the first step in the development of mathematical handling of uncertainty and risk. The advances that followed, such as the Law of Large Numbers (LLN), first proved by Jacob Bernoulli [13], and the expected utility theory put forward by Nicholas Bernoulli, began equipping individuals with novel decision-making tools when faced with risk and uncertainty, evaluating the utility given a certain level of risk [1]. When given a sample of observations such as multiple outcomes of a game of chance individuals could use the LLN to approximate Pascal and De Fermat's expected value when assessing the sample average. The LLN provides the mathematical grounding that the sample mean of an experiment should converge to the expected mean.

This early formulation of the expected value was appropriate for independent events, such as tossing a coin or throwing dice, but neglected the time domain inherent in certain sequential data—such as financial time series—due to its perceived randomness. Yet, using historical financial data to form predictions of future events is not as reliable as it is highly unlikely that historical performance of certain assets would be indicative of future performances attributed to the occurrence of market disruptions and changing market regimes [41, 93]. In reality, financial data such as return series should be modelled as a stochastic process consisting of both a time domain and an amplitude domain.

A stochastic process can be understood as a function that depends both on a random variable and on time. It can have multiple realisation paths, each representing a distinct possible sequence of outcomes over time driven by random variations effectively serving as different scenario outcomes. Fig.1.1 on the next page illustrates these two domains. 1.1a illustrates the amplitude domain and 1.1b shows various paths of a process over time. Following Peters and Adamou [104], given a random variable Z , any function such as $Y(Z)$ is itself also a random variable. A stochastic process is a special case in which this function also depends on time, such that:

$$Y_Z(t) = f(Z, t). \quad (1.1)$$

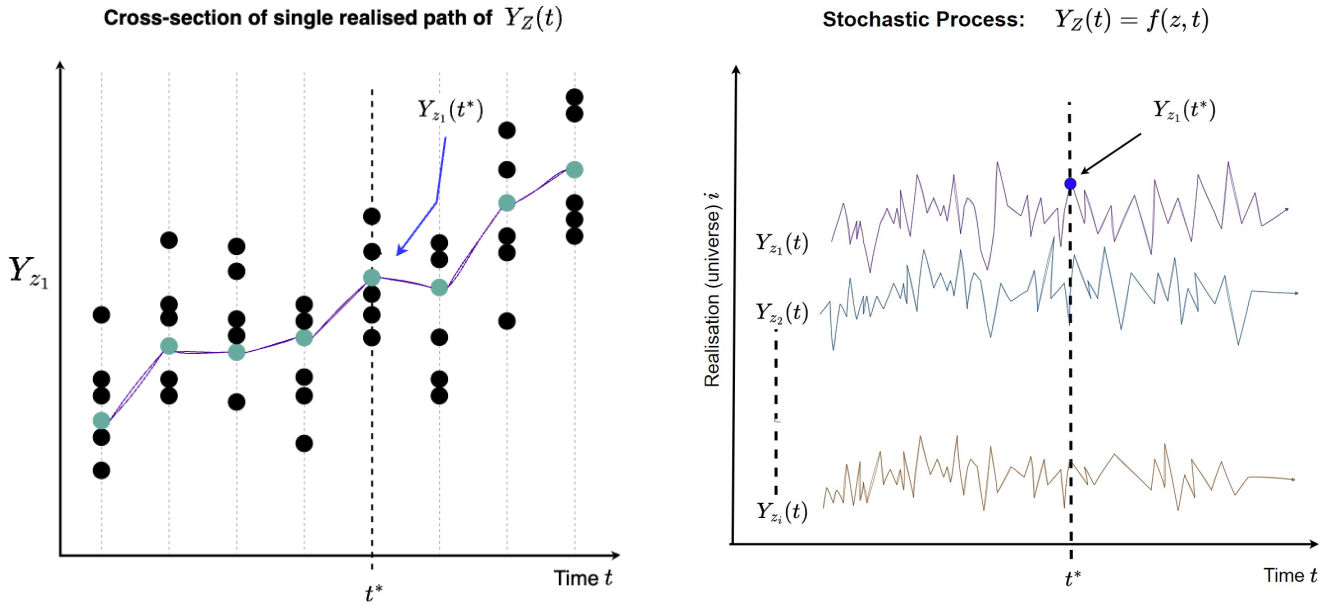
In simulations, stochastic processes are often generated iteratively: starting from an initial value, the process is updated at each step by drawing a new random number, producing a trajectory that represents one possible evolution of the process.

¹A popular example of this is a game of chance: one person repetitively throwing a die five times technically takes place over time but the process is treated no different from five individuals simultaneously throwing a die. This example was used by Fermat and Pascal to develop their theory of stable probability ratios which later led to the use of expected values—the average over a finite number (n) trials, *i.e.* the finite-ensemble average.

In the representation above, however, a single realisation z can be drawn from Z once and substituted into the function $f(Z, t)$, producing a deterministic function of time:

$$Y_z(t) = f(z, t), \quad (1.2)$$

where all randomness is captured by the single draw z . Once z is specified, the trajectory $Y_z(t)$ is fully determined for all t . This trajectory is called a realisation or sample path of the stochastic process. To illustrate, for any realisation i of a stochastic process $Y_Z(t)$ at any given time point t , there exists a distribution in cross section that defines the amplitude domain as seen at t^* in sub-figure 1.1a. Eight different observations of a single realised path Y_{z_1} at different time points (indicated in green - left to right) are shown along with their distributions consisting of alternative possible outcomes (indicated in black - top to bottom). The cross-sectional ensemble of points that make up the distributions at each t can be summarised by a probability density function (PDF) $\mathcal{P}_Z(z)$, created by observations over several realisations, $Y_{z_1} \dots Y_{z_i}$, at a single point in time t^* . Each of the black points indicated, therefore, have their own probability of occurrence attached to them, creating a probability distribution in cross-section.



(a) A single realisation of a stochastic process illustrates the amplitude domain that is present at each time point. (b) Multiple realisations of a stochastic process illustrate the time domain.

Figure 1.1: A stochastic process $Y_Z(t)$ consists of both an amplitude domain (a function of multiple realisations i) and a time domain (a function of time t). At any given time point t , there exists a distribution in cross section that defines the amplitude domain. Fig.1.1a represents a single realisation Y_{z_1} with cross sections at fixed time points t . The observed realisations of a single realised path of the stochastic process $Y_Z(t)$ are indicated in green, while the remaining potential outcomes are represented in black. Each outcome within the ensemble at a given time point is associated with its own probability within the cross-sectional distribution, where the probability of observing a specific outcome can vary over time based on available information. Fig.1.1b demonstrates, that in reality, it is possible that multiple paths of $Y_Z(t)$ exist across all time points. The ensemble of the observed realisations of these multiple paths define the amplitude domain.

This example of the amplitude domain *vs.* the time domain, showcases why investors should not solely rely on historical observations of time series data, as it is effectively a single measure of an ensemble distribution of possible outcomes at a given time point t . What determines which observation is realised out of an ensemble is the probability of that outcome occurring within its distribution $\mathcal{P}_Z(z)$. Yet, distributional properties (such as μ and σ) are not preserved over time as there is no $t-1$ to t mapping, making previous ensemble distribution unreliable predictors for future time points. Consequently, an investor effectively has uncertainty in both the amplitude and the time domain.

MPT largely relies on concepts based in the amplitude domain, such as the expected value and the covariance that form its inputs. In the setting of TSA there are three different expressions of the expected value used. No version is incorrect, however, they need to be interpreted and applied in the correct context. In the amplitude domain, two types of averages can be distinguished: 1) the finite ensemble average, and 2) the ensemble average in the limit, both of which are time independent. In contrast, when operating in the time domain, the time series average can be used as a corresponding metric. Below we walk through the various forms of the expected value and place the theory in the context of MPT. Each definition below should be applied based on the nature of the data, if it is ergodic, stationary, *etc.*:

Definition 1 (Finite-ensemble average). The finite-ensemble average in a stochastic setting is time-dependent and is the average over a finite number of known realisations in the amplitude domain at a fixed point in time t^* , *e.g.*, the average outcome from several simultaneous coin tosses:

$$\mathbb{E}[Y_Z(t^*)] = \frac{1}{n} \sum_i^n Y_{z_i}(t^*), \quad (1.3)$$

where $Y_Z(t^*)$ is the stochastic process observed at a single point in time t^* , i represents a particular realisation $z(t)$ of the stochastic process $Y_Z(t)$ and n is the finite number of realisations averaged over *i.e.* the sample size.

If we suppress Y in the notation to remove the stochastic process component, and remove the dependency on time t , we can obtain a formula that more closely resembles the conventional time-independent representation of the finite-ensemble average (sample mean):

$$\mathbb{E}[Z] = \frac{1}{n} \sum_i^n z_i. \quad (1.4)$$

Creating this distinction between time-independent and time-dependent calculations will help us highlight the importance of not neglecting the time domain when assessing time series data and applying expected values.

The expected value of the random variable Z , however, can only be approximated using the finite-ensemble average as in Eq.1.4 if there are enough realisations, which is rarely the case. Alternatively to represent the expected value we use the large-ensemble limit of the finite-ensemble average².

²The ensemble average over infinitely many realisations as $n \rightarrow \infty$ relates to the Law of Large Numbers. See Sec.1.2.1.2 for details.

Definition 2 (Ensemble average (in the limit)). The Ensemble average (representative of the expected value) is the large-ensemble limit of the finite-ensemble average:

$$\mathbb{E}[Z] = \lim_{n \rightarrow \infty} \mathbb{E}[z_i]. \quad (1.5)$$

Eq.1.5 can further be shown to take the probabilistic form (see Apx 4):

$$\mathbb{E}[Z] = \sum_j p_j z_j. \quad (1.6)$$

Noting that the subscript i indexes a particular realisation of z while j indexes a possible value of z . The sample average of Z , that takes on discrete values z_j , is the sum of all possible j values weighted by their probabilities p_j .^a

^aIn the classical application of the ensemble average—*i.e.* when calculating the expected value—it is common to assume that p follows a uniform distribution with $p_j \equiv \frac{1}{n}$ as is in the finite-ensemble average Eq.1.4.

If the realisations z are viewed as continuous rather than discrete, then the continuous ensemble average can be used.

Definition 3 (Continuous ensemble average (in the limit)). The continuous ensemble average of a function $f(Z)$ of the system state z is given by

$$\mathbb{E}[f(Z)] = \int_{\Omega} f(z) \mathcal{P}_Z(z) dz, \quad (1.7)$$

where $\mathcal{P}_Z(z)$ is the time-independent PDF of the random variable Z at state z , and Ω denotes the state space.

However, as we are working in the context of financial time series data, the time domain cannot be forgotten, therefore we reintroduce the time domain t :

Definition 4 (Time series average). The time series average is defined as the long-term mean of a function $f(z(t))$, as if we are averaging from left to right across the time domain down the realised path z :

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T f(z(t)) dt. \quad (1.8)$$

Here the state depends on time t and integrates to the divergent averaging time T .

The combination of definitions 3 and 4 is essential to examining ergodicity. They enable the comparison of the time-independent ensemble average (Eq.1.7) with the time-dependent average (Eq.1.8), providing a framework to determine whether the average across the amplitude domain corresponds to the average along a single realisation path—the criterion to be met for ergodicity.

In line with Peters [103], if we equate the two above as in Eq.1.9 and it holds, then the expected value equates to the time average, satisfying the Birkhoff–Khinchin condition [16]. We can therefore theoretically conclude that the observed realisation is ergodic and avoid integrating over time which is divergent when averaging to a distant time T . We can instead integrate over the state space of the amplitude domain Ω .

Theorem 1 (Birkhoff–Khinchin Ergodic Theorem). Let $z(t)$ be the state of a system evolving in a state space Ω , and let $f : \Omega \rightarrow \mathbb{R}$ be an integrable function. If the system is *ergodic*, then the long-term time average of f along the trajectory of the system equals the ensemble (space) average of f over the state space:

$$\underbrace{\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T f(z(t)) dt}_{\text{Time average of } f} = \underbrace{\int_{\Omega} f(z) \mathcal{P}_Z(z) dz}_{\text{Ensemble average of } f}. \quad (1.9)$$

Here, $z(t)$ varies with time, and $\mathcal{P}_Z(z)$ is the stationary probability density of the random variable Z over Ω . The state on the left depends on time t and integrates to the divergent averaging time T , while the right integrates over the space of all states, Ω .

Placing this theory backing into context, likely contributors to the probability distributions $\mathcal{P}_Z(z)$ in financial time series data (for example, asset returns) would be a blend of effects of free agents, the speed and efficiency with which information disseminates through markets, as well as how effectively market participants respond to the information. We know that a probability distribution must exist for each realisation across each time point but there are clear limitations on estimating the distributions, and along with them the future path of returns. The significance of highlighting this characteristic of time series data is twofold: Point 1.) It illustrates that more than one path is possible for any given time series *i.e.*, different scenarios are possible but only one is historically realised. Point 2.) It speaks to the importance of being aware of the predominant assumptions that underlie TSA. If the idea is that one wants to predict the trend of the time series into the future using extrapolation of the historical time series, the time series needs to be stationary and ergodic. This is necessary so that the time series average converges to the cross-sectional ensemble average in the limit [40] as shown in Eq.1.9.

Key Takeaway: In MPT, two prominent issues arise concerning the expected value and ergodicity. First, due to the limited supply of historical data, we can only compute the finite ensemble average in Eq.1.4 (*i.e.*, the sample arithmetic mean). This is often assumed to approximate the expected value described by Eq.1.5 or Eq. 1.7 (the ensemble average in the limit), which is a common simplification when naively applying MV theory. Second, these expected values are then used for predictions or optimisations that inform future decisions, essentially assuming that the Birkhoff–Khinchin theorem [16] applies, *i.e.*, that the system is ergodic. This creates estimation error as there is not enough data to assume convergence to the ensemble average and also assumes all possible outcomes in the ensemble distribution are equally probable.

Recent work by both Hitt et al. [46] and Peters [102] have highlighted the limitations of applying ergodic economic theories to non-ergodic financial data, particularly in a post-Covid-pandemic world, where economic conditions and variables have exhibited heightened levels of systemic and idiosyncratic shocks, unpredictability, and structural breaks. The section to follow intends to showcase the impact that these idiosyncratic shocks can have on an investment portfolio.

1.1.2 Market dynamics: systemic and idiosyncratic shocks

Historically, global markets have experienced cyclical patterns. Therefore, markets have been argued to be predictable in the cyclical sense but not in the time series sense [78, 80]. Investors never know when markets might be at a point of inflection. It is unlikely that the features that materialise at a future inflection point within the more predictable “cycles” would be present in your historical dataset. Hence, how are models meant to produce reliable predictions of future events if the information is not available in your historical data for it to build on?

For example, unforeseen events such as the Dotcom Crash in the early 2000s and the Global Financial Crisis (GFC) around 2009 caused widespread market crashes, generating significant repercussions across all major markets globally. Fig.1.2 on the next page, employs equity indices as proxies for the United States (US), Europe, and South Africa (SA), illustrating the impact of three major market crashes—The Dotcom crash, The GFC and the onset of the Covid-19 pandemic—on world markets. Although the extent of the impact varies among markets, the repercussions are seen across all proxies. Notably, the sudden emergence of the COVID-19 pandemic led to a sharp decline and subsequent recovery in markets, making it somewhat obscured by the third vertical line. The S&P 500 serves as a representation of the US market, and The Financial Times Stock Exchange 100 Index and The STOXX Europe 600 act as proxies for the European market. In the case of the SA market, the FTSE/JSE Africa All Share Index (ALSI) was chosen as the SA equity proxy.

Fig.1.2 below illustrates that global markets are inherently susceptible to unexpected turning points, capable of catching market participants off guard. When thinking of these events in the context of portfolio construction, it becomes important to consider that not all assets react uniformly to systemic shocks. Equity and equity-related assets exhibit greater exposure to global economic risk, while real assets and nominal bonds (or similar) may derive returns from alternative sources, making them less dependent on the equity market [108]. This suggests that an effective strategy for building a resilient portfolio against the impact of systemic shocks would involve diversifying with a mix of relatively independent high-risk/high-reward assets and lower-risk assets.

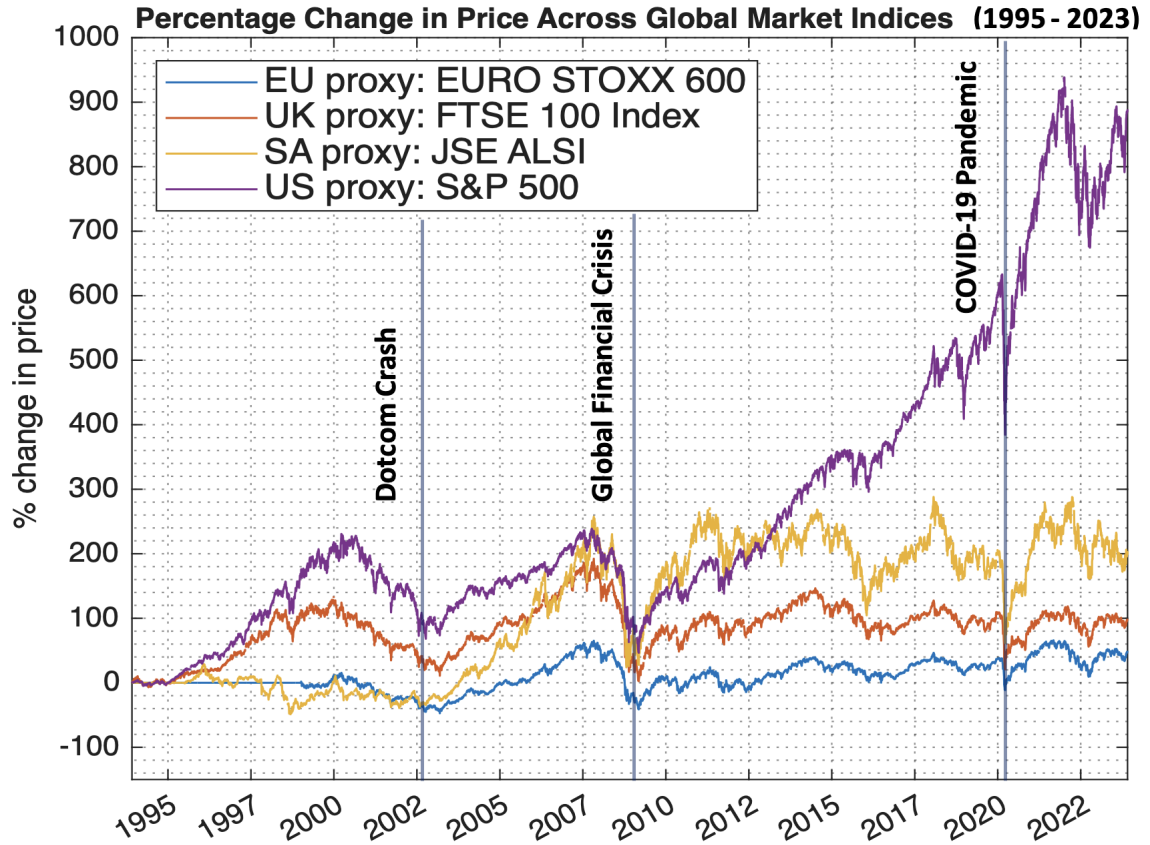


Figure 1.2: The data presented is the percentage change in daily prices for four major global market indices over twenty-eight years, denominated in USD. There are significant events marked by vertical lines: the late 2002 line signifies the end stages of the Dotcom crash, the 2009 line corresponds to the trough of the Global Financial Crisis (GFC), and the early 2020 line marks the COVID-19 pandemic. These three events, which occur roughly once per decade, represent major systemic shocks, each unique in nature and cause. The impact of these kinds of events is clearly felt across global markets to varying degrees, prompting a global need and desire to anticipate them. However, the occurrence of the shocks themselves serves as evidence that the market, at any given time, can be at an unanticipated and unpredictable point of inflection which is comprised of a time and a state component. Investors acknowledge that these tipping points exist, yet as a majority, we still lack adequate tools to safeguard our assets against them. Generated by: `daily_market_data_cleaning.1a.mlx`

These events are widespread and are comprised of a *time* and a *state* component - “When in history did it occur?” and “What was the state of the market in the run-up to, during, and following the event?”. As stated by Sebastian and Gebbie [116], lack of considering both the time and state of the market suggests that models making use of financial time series data may suffer from generalisation errors. A more systematic approach to portfolio construction and asset allocation decisions would involve considering multiple key inputs rather than relying on a single source of information such as return paths. However, most approaches such as MV rely heavily on historical returns and not the broader category of market data which would include information on risk drivers *etc.*. Bringing together what we have discussed thus far, we can consider using the concept of multiple paths in historical market data to our advantage. We can use the presence of a probability distribution $\mathcal{P}_Z(z)$ at each time stamp t to manipulate the historical times series to fit current market regimes [103]. Instead of approaching it from a parametric TSA approach, we can rather focus on using a scenario-based approach. This was the sentiment of Meucci [85] in 2010 when outlining the Historical Simulation with Flexible Probabilities (HS-FP) framework [89].

The HS-FP framework plays a central role in this dissertation and will be thoroughly examined in the chapters ahead. However, before diving into its adaptation, we first review the foundational principles of portfolio construction and the techniques that underpin it.

1.1.3 Quantitative portfolio construction: a cautionary tale

Multiple strategies exist for building an investment portfolio: Active *vs.* passive management, discretionary *vs.* quantitative, a tactical asset allocation (TAA) or a strategic asset allocation (SAA) strategy, *etc.*, but each of these can be linked to research topics of their own. This dissertation zeroes in on the quantitative methods, which are gaining increasing popularity in practical applications.³

The inception of MPT and portfolio construction can be traced back to the 1930s when economists such as Williams [127] began the development of fundamental analysis of stock prices. Williams [127] made significant contributions with his seminal work, “*The Theory of Investment Value*”, published in 1938. During this period, investors predominantly applied a discretionary approach to asset selection for a given portfolio. The mentality was to buy up a selection of sound stocks at the best prices, often neglecting to consider risk management or diversification [71].

Through the work of Williams [127], Marschak [72], and later Leavens [58] (in his 1945 article “*Diversification of Investments*”) the thinking behind how markets operate began to shift. Since the early 1950s, observed “expected returns”, and “variance” of returns have become crucial criteria in portfolio selection—in attempts to achieve reasonable risk-adjusted returns. Researchers such as Markowitz [69] and Roy [113] pioneered the in-depth exploration of these concepts in 1952 and beyond. The use of these quantitative measures saw a significant change in how historical data played a role in investor decision-making. The nature of observed historical time series data, however, presents a challenge as it is derived from non-stationary underlying processes and is subject to market fluctuations [38]. Therefore, a more systematic approach, grounded in rule-based methods, is needed to navigate complex market dynamics influenced by both systematic and idiosyncratic shocks. Recognizing the continuously changing market landscape is crucial to effectively navigate the intricacies of portfolio construction. Part of attempting to immunise a portfolio requires stress-testing to build confidence in the robustness of portfolios in dynamic markets.

1.1.3.1 The selection problem: stress-testing

The practice of stress-testing gained prominence in the mid to late 20th century, notably through the Financial Sector Assessment Program (FSAP) initiated by the International Monetary Fund and the World Bank in 1999 [73]. Defined as “*forward-looking exercises that aim to evaluate the impact of severe but plausible adverse scenarios (on the resilience of financial institutions.)*” [12].

Up until the mid-90’s, data supply was limited and it was challenging to model market dynamics. However, in the 30 years following there has been a rise of easy access to extensive data, facilitated by the emergence of the internet, improved data storage capabilities, and enhanced computational power [114].

³Even when quantitative methods are employed, a manager retains the ability to express their perspectives by selecting the methods and inputs for the models. While this may not be synonymous with having full discretion, it nevertheless has a significant influence on the final portfolio.

The influx of data resources prompted the development of further quantitative approaches to portfolio construction, not necessarily moving away from the fundamentals, but rather creating an alternative or supplementary path to achieve the same objectives. These advances in quantitative strategies began to allow investors to navigate market shocks in a more informed and responsive manner, as well as to better stress-test their portfolios. Additionally, data informed methods that utilise historical data are thought to be beneficial as they avoid the need for data simulations that require assumptions of dependency structures, data distributions, and other features [116]. Quantitative strategies have also become more favoured by some as they decrease the time and manpower needed to assess the right asset allocation required to meet a mandate or a client’s return objective [60].

In the early stages, particularly in the context of portfolio theory, stress-testing heavily relied on parametric methods. Asset risk drivers like returns were modelled through parametric distributions, such as f_X^θ (where θ represents parameters). However, limitations arose for a variety of assets as the relationship between their profits & losses (P&L) and their risk drivers are non-linear. Estimation risk became more prominent in this case when trying to establish the distribution of yet-to-be realised P&L as the distribution of the P&L cannot be computed exactly [85]. A classic example is the modelling of returns using the normal distribution, with applications to the Sharpe Ratio (SR) or Value at Risk (VaR), which, if used carelessly and in the incorrect context, have been proven to yield unreliable results [66]. Although these methods became more accessible and computationally cheaper with the increasing availability of financial data, the GFC highlighted the necessity for a substantial update in the approach to risk management. The average investment firm at the time was ill-equipped to minimise the downside risk against the systemic shock of the crash.

Basic logic needs to be applied to the stress-testing process. Currently, models with strong results undergo testing, modification, and retesting, while those with poor results are promptly discarded. This creates two flaws in logic. Firstly, some effective models may fail in the test period, leading to their unintended exclusion, despite it potentially being purely an artefact of a unique dataset—a false negative. Secondly, researchers may construct narratives to rationalise under performing models that appear successful during the test period, often for reasons unrelated to the model’s future efficacy—a false positive. Sadly, an exaggerated positive is even more common than a false positive, making the outcome seem disproportionately stronger than it is feasible to be in the future [3, 30].

It is important to recognise that, irrespective of a model’s foundation and sophistication, we can never definitively prove it’s correctness or verify it. Instead, we can corroborate it over time by building up historical evidence that suggests it is the most robust model available—until new evidence comes along that requires you to re-evaluate it. Stress-testing and continuous monitoring of current models become key in this setting. There is a need for more discussion on how to identify when a model might be incorrect, to foster an environment for ongoing development, improvement, and to refine our thought processes. This highlights that focusing solely on the math isn’t enough; we must incorporate a degree of subjective judgment to boost confidence in stress-test outcomes [110].

1.1.3.2 Structuring a quantitative portfolio process

Constructing a trading or investment strategy can be categorised into two key domains: risk allocation and alpha generation (*i.e.* excess return generation) [25, 36, 113]. Extracting excess returns primarily hinges on factors such as the trading frequency or trading horizons and, in the case of a quantitative portfolio, it can entail the application of statistical techniques. The allocation of associated risk predominantly relies on the selection of an appropriate model and objective criteria. Fig.1.3 lays out the most basic structure for the portfolio construction process in industry. It requires two inputs, the chosen securities and a set of objectives and constraints put forward by a manager. These are then fed into a chosen optimisation method, which should then ideally be stress-tested appropriately.

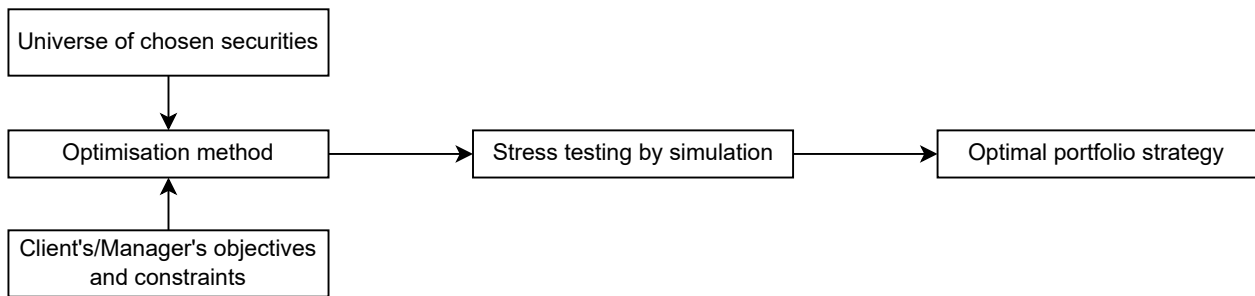


Figure 1.3: The conventional quantitative portfolio structuring process.

In recent years diverse approaches have been employed for quantitative portfolio construction, focusing on risk management, the maximisation of returns, or a calculated combination of both objectives. Some strategies—such as a Global Minimum Variance Portfolio (GMV), a Risk Parity Portfolio (RPP), a Hierarchical Risk Parity (HRP) approach, or an Inverse Volatility Portfolio (IVP)—place a distinct emphasis on risk management. In contrast, others—such as a SR Maximising, a Classical MV, or a Balanced Portfolio (60:40 mix)—are designed to place precedence on risk-adjusted returns [22, 29, 54, 69, 118]. However, varying methods frequently yield vastly different outcomes over time based on each strategy's priority, some perform well in the short term while others perform better over a longer time horizon. The outcomes are contingent on the constituents that a specific method would select for a given portfolio type, their reactions to the described market behavior, and the imposed constraints. Fig.1.4 on the next page illustrates this, showcasing the comparative performance of various approaches that aim to balance the risk-reward trade-off throughout a business cycle.

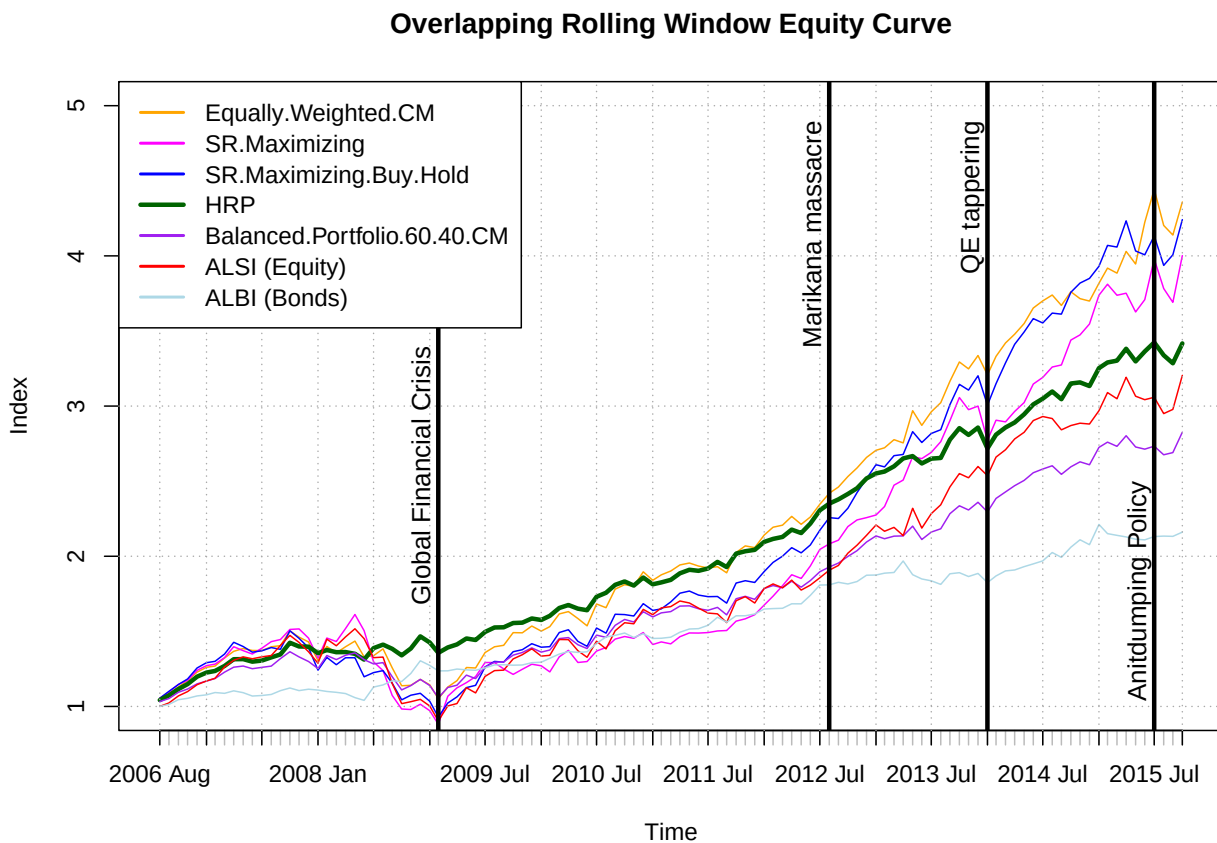


Figure 1.4: A diverse range of portfolio construction techniques are available today, each with distinct objectives such as risk management or maximising excess returns α . Their effectiveness varies based on market states. Stress-testing a selection of portfolios⁴ in a local market, like South Africa, using a dynamic back-test technique reveals assorted equity curve trajectories. By using an overlapping rolling three-year window of monthly SA balanced fund data a substantial variation in performance is shown between August 2006 and November 2015⁵. Similar to the shocks seen in Fig.1.2, market-specific events, notably the GFC, are evident. The GFC impacted the α -maximising portfolio (SR maximising MV) more than the risk management portfolio (HRP). Nevertheless, over the investment horizon, the α -maximising portfolio outperformed. No single portfolio dominates under all market states, emphasising the importance of efficiently adapting models to market changes. Therefore, the ability to model the market is essential when stress-testing portfolio P&L under abnormal market conditions. In Chp.3, the HS-FP framework is introduced to enhance the conventional stress-testing technique.

Through the dynamic back-test performed on SA sector data in Fig.1.4, an overlapping rolling window method was applied rebalancing monthly with a fixed in-sample window length of three years worth of monthly historical data. Similar to events marked in Fig.1.2, points of SA-specific interest are indicated by vertical lines in Fig.1.4. The primary line of interest would be at the start of 2009, indicating the trough of the GFC. We see that the relative performance of the different portfolios follow a similar path except for the HRP portfolio. HRP being a diversification-focused algorithm, outperformed during the idiosyncratic shock brought on by the GFC. The strength of diversification is illustrated here as the HRP portfolio was barely weakened during the crash.

⁴“CM” denotes a constant mix portfolio, “SR Maximizing” refers to a SR maximising MV portfolio, and the “Balanced Portfolio.60.40.CM” is a balanced fund with a 60:40 equity to bond constant mix ratio.

⁵Fig.1.4 was generated as part of an honours-level thesis by myself alongside Siphesihle Cele for the degree of honours in statistics in the Department of Statistical Science at the University of Cape Town (UCT) [76]. Refer to Apx. Tab.1 for details on the benchmark portfolios and Apx. Fig.2 for a comparison with an alternative form of dynamic back-testing, a growing window simulation.

Two of the top performing strategies outside of the GFC—the SR maximising and the SR maximising Buy-Hold (BH) strategy—outperformed both prior as well as following the crisis. The performance of these particular optimisations appeal to many as they promise enticing returns during bullish markets, but no single portfolio dominates through all market states. These are by and large uninformed quantitative models with no foresight or awareness of market shifts and are built upon several potential error-introducing assumptions such as ergodicity, normality, or stationarity.

1.2 Fundamentals: assumptions made for portfolio theory

An appropriate concern would be to what extent should a manager trust in these quantitative models. This section tackles this question and underscores the necessity to approach quantitative models in their present state with a measure of caution due to their reliance on several fundamental statistical and financial assumptions.

1.2.1 Statistical considerations

The parameters of the statistical processes that govern financial data may not remain constant over time. Sec.1.1.2 reminded us that not only is the data regime dependent and non-ergodic, but there is also a potentially strong dependence on the measurement process that samples the data from the real world system, and this does not necessarily faithfully represent the underlying system itself, which is in turn non-linear, adaptive and complex. As a result, the use of different epochs or historical periods to calculate statistical metrics might not be warranted, or should be approached with care, as the underlying statistical properties of financial data are widely subject to change over time.

While the theoretical formulation of Modern Portfolio Theory (MPT) can be justified under either quadratic utility or normally distributed returns, its practical application typically relies on additional statistical assumptions for parameter estimation. In particular, the estimation of expected returns and covariances from historical data often presumes that the data are (here in particular to the total returns data) [30, 69]:

1. Identically and independently distributed *i.i.d* data (Sec.1.2.1.1),
2. Normal distribution of returns (Sec.1.2.1.3),
3. Stationarity (Sec.1.2.1.4), and
4. Ergodicity (Sec.1.1).

1.2.1.1 Independently and identically distributed

A sample of random variables $X_1, X_2, \dots, X_n$ is said to be *i.i.d*. if the following is met:

1. **Independence:** Each random variable X_i is independent of the others, meaning the occurrence of one event does not affect the outcome probability of another.
2. **Identical Distribution:** All X_i within the sample are taken from the same probability distribution with the same distributional parameters.

Theorem 2 (Independently and Identically Distributed data). Let $F_{X_1, \dots, X_n}(x_1, \dots, x_n) = P(X_1 \leq x_1 \wedge \dots \wedge X_n \leq x_n)$ denotes the joint cumulative distribution function (CDF) of random variables $X_1, \dots, X_n$ within a sample. $X_1, \dots, X_n$ are *i.i.d.* if and only if:

$$\begin{aligned} F_{X_1}(x) &= F_{X_k}(x) & \forall k \in \{1, \dots, n\} \text{ and } \forall x \in I, \\ F_{X_1, \dots, X_n}(x_1, \dots, x_n) &= F_{X_1}(x_1) \cdot \dots \cdot F_{X_n}(x_n) & \forall x_1, \dots, x_n \in I. \end{aligned} \quad (1.10)$$

In time series, *i.i.d.* implies that observations are not influenced by preceding values, which is unrealistic in real-world data where autocorrelation is common. However, by assuming data is *i.i.d.* it can simplify analysis and allows the use of laws like the LLN and the Central Limit Theorem (CLT).

1.2.1.2 The Law of Large Numbers

Under the LLN formalised by Bernoulli [13] it is stated that as the number of *i.i.d.* observations of a random variable $X_1, X_2, \dots, X_n$ increases, the sample mean $\bar{X}$ approaches the expected value μ of the random variable itself. This allows one to define the sample mean denoted by $\bar{X}$ as:

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n}. \quad (1.11)$$

Two different modes of convergence can be considered for the LLN, the “weak” and the “strong” form. Both versions of the LLN assert that:

$$\bar{X} \rightarrow \mu \quad \text{when } n \rightarrow \infty. \quad (1.12)$$

The Weak LLN guarantees convergence in probability, while the Strong LLN guarantees almost sure convergence⁶. It is known here that the strong form implies the weak form, however, the reverse is not true.

Theorem 3 (Weak LLN). Let $X_1, X_2, \dots, X_n$ be a sequence of *i.i.d.* random variables with expected value μ and finite variance σ^2 . Then, for any $\varepsilon > 0$,

$$\lim_{n \rightarrow \infty} \Pr(|\bar{X} - \mu| < \varepsilon) = 1. \quad (1.13)$$

This implies that $\bar{X} \xrightarrow{P} \mu$ as $n \rightarrow \infty$. *i.e.* the sample mean converges in probability to the expected value.

⁶The strong LLN is a particular example of the point-wise ergodic theorem [105].

Theorem 4 (Strong LLN). Let $X_1, X_2, \dots, X_n$ be a sequence of *i.i.d.* random variables with expected value μ . Then,

$$\Pr \left(\lim_{n \rightarrow \infty} \bar{X} = \mu \right) = 1. \quad (1.14)$$

This means that $\bar{X} \xrightarrow{\text{a.s.}} \mu$ as $n \rightarrow \infty$. *i.e.* the sample mean converges almost surely to the expected value.

These theorems imply that the accuracy of an estimation process improves as the number of data points increase in the sample, assuming the data is *i.i.d.* However, financial data can exhibit heteroscedasticity with time-varying distributions, and are frequently non-ergodic and non-stationary. Structural breaks, and regime changes make LLN-based estimates potentially misleading as the LLN may not apply directly. This implies that time averages may not reflect the true expected value. The LLN is further discussed in Sec.3.2 where it is used with the generalised empirical distribution to derive the HS-FP framework.

1.2.1.3 Normality via the Central Limit Theorem

Given $X_1, X_2, \dots, X_n$ as a sequence of *i.i.d.* random variables, if the strong LLN is invoked then $\bar{X} \xrightarrow{\text{a.s.}} \mu$ as $n \rightarrow \infty$. While the strong LLN describes convergence in the mean, it does not specify the distributional shape of $\bar{X}$ around its limit. The Central Limit Theorem (CLT) addresses this by describing the asymptotic distribution of the sample mean. If a population has a finite mean μ and finite variance σ^2 , then as n increases, the distribution of the sample means (from an infinite number of independent samples, each with n independent observations) converges to a normal distribution with mean μ and variance $\frac{\sigma^2}{n}$, irrespective of the population's original distribution shape.

Theorem 5 (Lindeberg–Lévy Central Limit Theorem). Let $X_1, X_2, \dots, X_n$ be a sequence of *i.i.d.* random variables with finite mean $\mathbb{E}[X_i] = \mu$ and finite variance $\text{Var}(X_i) = \sigma^2 < \infty$. Define the sample mean as in Eq.1.11

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i. \quad (1.15)$$

Then, as $n \rightarrow \infty$, the normalised sum converges in distribution to a standard normal variable:

$$\sqrt{n} \left(\frac{\bar{X} - \mu}{\sigma} \right) \xrightarrow{d} \mathcal{N}(0, 1). \quad (1.16)$$

Equivalently,

$$\sqrt{n}(\bar{X} - \mu) \xrightarrow{d} \mathcal{N}(0, \sigma^2). \quad (1.17)$$

A standard proof of this result can be found in Feller [34].

As n increases, the distribution of $\sqrt{n}(\bar{X} - \mu)$ approaches a normal distribution with mean 0 and variance σ^2 . This implies that for large n , the sample mean $\bar{X}$ will become approximately normally distributed with mean μ and variance $\frac{\sigma^2}{n}$, regardless of the original distribution of the data. This assumption of normality (supported by the CLT) is a common assumption in TSA. It allows for the application of more conventional statistical methods as it creates a known distribution with known properties. This assumption, however, often fails due to the presence of autocorrelation, non-stationarity, and other dependencies that violate the *i.i.d.* conditions.

1.2.1.4 Stationarity

If a process is stationary then the statistical properties of the series do not change over time. A stochastic process is strictly stationary in both the amplitude and the time domain if the joint distribution of a stochastic process $Y_Z(t)$ remains invariant under time shifts, meaning the statistical properties are identical regardless of the time at which they are observed.

Definition 5 (Strict Stationarity). A stochastic process $\{Z(t)\}$ is strictly stationary if the joint distribution of $Z(t_1), Z(t_2), \dots, Z(t_k)$ is the same as the joint distribution of $Z(t_1 + \tau), Z(t_2 + \tau), \dots, Z(t_k + \tau)$ for all time points $t_1, t_2, \dots, t_k$ and for all time shifts τ . Formally,

$$F_Z(z_{t_1+\tau}, z_{t_2+\tau}, \dots, z_{t_k+\tau}) = F_Z(z_{t_1}, z_{t_2}, \dots, z_{t_k}) \quad \text{for all } \tau, t_1, t_2, \dots, t_k \in \mathbb{R} \text{ and for all } k \in \mathbb{N}_{>0},$$

where $F_Z(z_{t_1+\tau}, z_{t_2+\tau}, \dots, z_{t_k+\tau})$ denotes the cumulative distribution function (CDF) of the joint distribution of the stochastic process $\{Z(t)\}$ at times $t_1 + \tau, t_2 + \tau, \dots, t_k + \tau$.

A less stringent version of stationarity is wide-sense stationarity as defined below in Def.6. It requires that only the mean, variance, and autocovariance of the process remain constant over time and depend only on the time difference τ and not on the specific time points t .

Definition 6 (Wide-Sense Stationarity). A stochastic process $\{Z(t)\}$ is wide-sense stationary if it satisfies the following conditions:

1. The expected value $\mathbb{E}[Z(t)]$ is some **constant** μ for all $t \in \mathbb{R}$: $\mathbb{E}[Z(t)] = \mu$.
2. The variance is some **constant** σ for all $t \in \mathbb{R}$: $\text{Var}(Z(t)) = \mathbb{E}[(Z(t) - \mu)^2] = \sigma^2$.
3. The auto-covariance function where $\gamma(\tau)$ depends only on the lagged time difference τ , and not on the specific times t and $t + \tau$:

$$\text{Cov}(Z(t), Z(t + \tau)) = \mathbb{E}[(Z(t) - \mu)(Z(t + \tau) - \mu)] = \gamma(\tau) \quad \text{for all } t, \tau \in \mathbb{R}.$$

A time series $\{X_t\}$ is then said to be wide-sense (or weakly) stationary if:

1. The mean $\mathbb{E}[X_t]$ is constant for all t .
2. The variance $\text{Var}(X_t)$ is constant for all t , and
3. The autocovariance function $\text{Cov}(X_t, X_{t+\tau})$ depends only on the time lag τ .

1.2.2 Financial data considerations

Beyond the statistical assumptions underlying traditional MPT-based portfolio construction, it is equally important to evaluate how portfolio performance is assessed. As demonstrated in Sec.1.1.1, performance metrics must be selected with careful consideration of their intended application and the nature of the underlying data, as not all metrics are equally reliable across varying contexts.

The evaluation of portfolios within the MPT framework necessitates performance measures that include (but are not limited to):

1. Returns (Sec.1.2.2.1),
2. Return averages (Sec.1.2.2.2),
3. Volatility as measures of risk (Sec.1.2.2.3), and
4. The SR as a measure of risk adjusted return (Sec.1.2.2.4).

1.2.2.1 Returns

When measuring the performance of an asset or a portfolio of assets, the increase or decrease in their value reflects the change in wealth over a specific period. The gain or loss in value *i.e.* the return can be deduced from the change in price relative to a starting value. Which returns are used as well as the method by which they are averaged should be determined by the use case. For instance, when optimising monthly data over a year, one should rather use the geometric average due to the non-ergodic nature of longer time series data. This approach is essential in MV optimisation, where the expected mean is used to make optimal decisions. While arithmetic averages suffice for short-term monthly optimisations, annual decisions require compounding monthly returns to reflect the expected annual return accurately.⁷ The alternative return calculations are laid out below.

Definition 7 (Price Relatives). The price relative from time $t - 1$ to t is given by:

$$x_t = \frac{P_t}{P_{t-1}}, \quad (1.18)$$

where P_t is the price at time t and P_{t-1} is the price at time $t - 1$.

$x_t < 1$ indicates a decrease in value while a $x_t > 1$ indicates an increase. The price relative is also known as the “wealth ratio” which can be used to calculate the simple return value as a discrete measure of the absolute change in value of an asset.

Definition 8 (Simple Returns). The simple return R_t from time $t - 1$ to t is defined using the price relative as:

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}} = x_t - 1 = \frac{P_t}{P_{t-1}} - 1, \quad (1.19)$$

where P_t is the price at time t and P_{t-1} is the price at time $t - 1$.

⁷Note for this work that the geometric average of the time series is not being used directly in the optimisations but rather the arithmetic average of simple returns is converted into the geometric expectation. However, when attributing performance, geometric compounding through time is necessary.

If we were to take the log of x_t we can define the continuous return. Unlike simple returns which can exhibit positive bias due to their arithmetic nature, continuously compounded returns (log returns) handle the compounding effect more accurately [5].

Definition 9 (Continuous Returns). The continuous return μ_t from time $t - 1$ to t is given by:

$$\mu_t = \ln(P_t) - \ln(P_{t-1}) \iff P_t = e^{\mu_t} P_{t-1} \iff x_t = e^{\mu_t}. \quad (1.20)$$

This implies that $P_t = P_0 e^{\mu_t t}$, given an initial price P_0 .

One perceived benefit of continuous returns is that if prices follow a log-normal distribution (though not always the case), the continuous returns will exhibit a normal distribution, thereby simplifying subsequent analysis. Log returns are also additive which makes it easier to calculate the overall performance of a portfolio over time as the logarithmic return of a portfolio composed of multiple assets is simply the sum of the logarithmic returns of each individual asset [123].

Definition 10 (Geometrically Compounded Returns). Geometric returns compound over discrete time intervals for times $t = 0, 1, 2, \dots, t, \dots, T$, where there are $T + 1$ prices and T returns:

$$(1 + R_1)(1 + R_2) \dots (1 + R_t) \dots (1 + R_T) = \frac{P_1}{P_0} \frac{P_2}{P_1} \dots \frac{P_t}{P_{t-1}} \dots \frac{P_T}{P_{T-1}} = \frac{P_T}{P_0}. \quad (1.21)$$

Here R_t are simple returns and P_t is the price at time t .

1.2.2.2 Return Averages

The correct application of arithmetic and geometric means as well as the appropriate returns calculations play a big role in the accuracy of performance evaluation and optimisation. The arithmetic mean treats all data points as identical in terms of weighting, while the geometric mean considers each data point as a scaling factor, combining them multiplicatively [5].

Definition 11 (Arithmetic Average). The arithmetic mean treats each data point equally within a sample. It is the summation of all values divided by the number of observations. This method is commonly used for calculating simple returns.

$$\text{Arithmetic average or average return} \quad R_A = \frac{f}{T} \times \sum_{t=1}^T R_t, \quad (1.22)$$

where T is the number of periods under analysis and f is the number of periods within the year (monthly $f = 12$, quarterly $f = 4$).

The arithmetic mean is appropriate for calculating average returns over a single period but can be misleading for multi-period returns due to the compounding effect. When used in conjunction with simple returns, which can be positively biased, simple returns tend to overestimate the true average return of an investment over multiple periods. This means that the use of the arithmetic mean of simple returns can lead to inflated optimistic expectations of future performance. For this reason, geometric returns are often preferred for long-term investment analysis and performance evaluation.

A second noteworthy aspect of the arithmetic mean is that it is mathematically equivalent to the finite-ensemble average defined in Eq.1.6 when all observations are assigned equal probability, $p_j \equiv \frac{1}{n}$ *i.e.* equivalent to Eq.1.4. Both represent an equal-weighted summation of outcomes, but they differ in what is being averaged. The finite-ensemble average aggregates across distinct realisations of the system at a given point in time, whereas the arithmetic mean aggregates sequentially across time for a single realised path. This distinction is subtle but conceptually important: when the arithmetic mean is used in the analysis of time series data, it implicitly assumes that averaging over time is equivalent to averaging over the ensemble of possible realisations. As discussed in detail in Sec.1.1, this assumption corresponds to the property of ergodicity, which holds only when time averages converge to ensemble averages. In practice, this means that applying the arithmetic mean to financial returns quietly assumes that the underlying stochastic process governing returns is ergodic.

The geometric mean, on the other hand, involves multiplying all the values together and then taking the n -th root (where n is the number of values). This method is particularly useful for calculating compounded returns over multiple periods, as it accounts for the multiplicative nature of returns.

Definition 12 (Geometric Average).

$$\text{Geometric average or annualised return } R_G = \left(\prod_{t=1}^T x_t \right)^{\frac{1}{T}} - 1 = \left(\prod_{t=1}^T (1 + R_t) \right)^{\frac{1}{T}} - 1, \quad (1.23)$$

where T is the number of periods under analysis and R_t are simple returns.

The calculation of returns, whether arithmetic or geometric, more often than not relies on the assumptions of: ergodicity, stationarity, *i.i.d.* data, and log-normal distribution for continuous returns.

1.2.2.3 Volatility as Measures of Risk

In finance, volatility—defined as the standard deviation of asset returns over a specified period—is a common metric used to assess risk by measuring the variability or dispersion of returns relative to their average. It captures the magnitude of fluctuations in asset returns and makes it a popular, intuitive input in risk management and portfolio optimisation strategies [50]. Volatility is derived from the variance of returns.

Definition 13 (Sample Variance and Volatility). Variance quantifies the dispersion of data points around their mean. For returns, the sample variance measures how returns deviate from the average return, providing a risk measure. It is calculated as:

$$s^2 = \frac{1}{T-1} \sum_{t=1}^T (R_t - \bar{R})^2, \quad (1.24)$$

where T is the number of observations, R_t represents each individual return (usually simple returns as in Def.8), and $\bar{R}$ is the sample mean (usually the arithmetic mean as in Def.11).

The sample standard deviation s , also known as **volatility**:

$$s = \sqrt{s^2} = \sqrt{\frac{1}{T-1} \sum_{t=1}^T (R_t - \bar{R})^2}. \quad (1.25)$$

The factor $\frac{1}{(T-1)}$ (Bessel's correction) adjusts for the fact that the sample mean $\bar{R}$ is estimated from the same data, reducing one degree of freedom. Volatility measures the typical magnitude of deviations from the mean in the same units as the data, making it more interpretable in practice.

The sample variance defined above uses observed data and corrects for the estimation of the mean, whereas the population variance assumes knowledge of all possible outcomes (or equal likelihood of each outcome).

Definition 14 (Population Variance and Volatility). If the population consists of outcomes $\{R_i\}$ with probabilities $\{p_i\}$, the population mean is

$$\mu = \sum_i p_i R_i, \quad (1.26)$$

and the population variance, which measures the average squared deviation from the mean, is

$$\sigma^2 = \sum_i p_i (R_i - \mu)^2. \quad (1.27)$$

In the special case where all outcomes are equally likely, $p_i = \frac{1}{n}$, this simplifies to^a

$$\sigma^2 = \frac{1}{n} \sum_{i=1}^n (R_i - \mu)^2. \quad (1.28)$$

The population standard deviation (volatility) is

$$\sigma = \sqrt{\sigma^2}. \quad (1.29)$$

^aUsing n in the denominator assumes that the dataset represents the full population of returns (or an equiprobable discrete population). If the data were only a sample, the unbiased sample variance would divide by $(T-1)$ to account for the degrees of freedom lost in estimating the mean. When all assets or funds are measured over the same horizon, using T instead of $T-1$ has a negligible effect on relative magnitudes of the resulting statistics.

If asset returns are normally distributed, the entire distribution can be described by the mean μ and the variance σ^2 (or equivalently, the standard deviation σ) [19], making variance and volatility sufficient statistics for risk assessment. However, as previously emphasised, normality in time series data is not guaranteed, which can make risk estimates less accurate when used in portfolio optimisation.

1.2.2.4 Sharpe Ratio

In 1952, Roy [113] introduced the concept of a risk-reward ratio to evaluate portfolio performance. Building on this idea, Sharpe [117, 118, 119] formalised the measure by applying Roy's logic to Markowitz [69]'s MV optimisation framework, leading to the development of the Sharpe Ratio (SR). This innovation enabled the evaluation of portfolio performance in terms of risk-adjusted returns and has since become one of the most widely used metrics for portfolio comparisons [7].

In Sharpe's initial 1966 and 1975 papers he discusses the concept of "excess returns" defining it as "*difference between the actual return and the return on a riskless security*". Later in 1998 he broadens the application and defines it in the context of general Fund F and a benchmark Fund B. It is however still more common practice to make the benchmark fund a risk-free asset to assess the "active return" whereas if an alternative benchmark is selected as Fund B the excess return could be better termed as the "selection return" [119].

For the purposes of the analysis to follow we will be using the SR as the former, "active return" as each portfolio will be assessed on its own merit, and whether it is worthwhile to add an active approach to investing. In order to differentiate between a predictive measure and a historical performance measure Sharpe [119] defines both the *ex ante* method of calculation and the *ex post*.

The *ex ante* SR is the *expected* differential return given the associated *risk*, computed using historical realisations but for the purpose of prediction:

Definition 15 (Ex Ante Sharpe Ratio). If R_p represents the return of a portfolio in the *forthcoming* period and R_f the return on a risk-free asset, then $\tilde{d}$ ^a will denote excess return *i.e.* $\tilde{d} \equiv \tilde{R}_p - \tilde{R}_f$. The SR can then be defined as the expectation for the differential per unit of associated risk:

$$SR \equiv \frac{\tilde{d}}{\sigma_d} \equiv \frac{E[R_p - R_f]}{sd[R_p - R_f]}. \quad (1.30)$$

Here σ_d is the predicted standard deviation of $\tilde{d}$, and $\tilde{d}$ the expected value of $\tilde{d}$.

^aThe use of tilde in this case is an indication that the values are not known with certainty in advance.

If the metric is used as an indicative measure of historical performance, the *ex post* formulation can be used:

Definition 16 (*Ex Post Sharpe Ratio*). Observing the historical realisation within a period t , let $D_t = R_{p,t} - R_{f,t}$ be difference between the return on the portfolio $R_{p,t}$ and the return on the risk-free asset $R_{f,t}$ in time period t :

$$SR_h \equiv \frac{\bar{D}}{\sigma_D} \equiv \frac{\frac{1}{T} \sum_{t=1}^T D_t}{\sqrt{\frac{\sum_{t=1}^T (D_t - \bar{D})^2}{T}}}, \quad (1.31)$$

where $\bar{D} \equiv \frac{1}{T} \sum_{t=1}^T D_t$, the average D_t over the time period T . The *Ex Post* SR, SR_h is expressed as a function of the *average* differential (excess) return and the associated *variability*.

In Eq.1.31, Sharpe [119] used the arithmetic average of the returns, however in alignment with his opening statement in [118] “A common human failing is the desire for simple answers to difficult questions”, he details how the formulation may require an increase in sophistication. Eq.1.31 can be built onto to account for the compounding of multi-period returns and that the SR is not independent of the time period T over which it is calculated.

Using the geometric average from Eq.1.23, which is more appropriate when attempting to account for the compounding effect over time, Eq.1.31 can be restated as⁸:

Definition 17 (*Geometric Ex Post Sharpe Ratio*). Here the differentials are geometrically averaged to better account for the compounding effect of time series data:

$$SR_h \equiv \frac{\bar{D}}{\sigma_D} \equiv \frac{\left(\prod_{t=1}^T (D_t + 1)\right)^{\frac{1}{T}} - 1}{\sqrt{\frac{\sum_{t=1}^T (D_t - \bar{D})^2}{T}}}. \quad (1.32)$$

Here as in the standard *Ex post* SR: $D_t = R_{p,t} - R_{f,t}$ and $\bar{D} \equiv \frac{1}{T} \sum_{t=1}^T D_t$.

The use of unadjusted historical *ex post* SR as proxies for unbiased predictions of *ex ante* ratios can be unreliable [119], however, if the intention is to use it for decision-making one would prefer to use the *ex ante* SR, which can provide valuable inputs for various investment decisions [119]. Regardless of the application, it is crucial to remember that the SR does not account for correlations. When a decision could impact significant correlations with other assets in an investor’s portfolio, such information should be considered in addition to comparisons based on SRs [119].

Going forward, the *ex post* SR, as defined in Def.17, will be used for all reported performance metrics in both chapters 2 and 3. This choice reflects its role as an indicative measure of historical performance, rather than a tool for forecasting.

⁸In Eq.1.30 through 1.32, the formula for σ assumes that the length of T is equal for all assets/funds being measured. This assumption results in the denominator being T and not $(T - 1)$, despite the fact the historical data is technically a sample of the population (all possible values of the excess returns) in which case the use of $(T - 1)$ would account for degrees of freedom. However in this case T suffices as the relative magnitudes of the resulting measured would be similar if not the same [119].

Sharpe Ratio shortfalls: As much as the SR is a widely used measure for assessing the performance of investment strategies, it has limitations due to the reliance on the assumptions of the statistical behaviour of the underlying return used for the calculation *i.e.* *i.i.d.* returns and normal distribution. The accuracy of the estimated SR is heavily impacted by these underlying statistical properties of the returns as they impact the distribution of the SR estimates themselves. Estimation errors are inherent in SR as the expected returns and variances need to be estimated as they are not generally observable. While we employ the SR as an indicative measure to gauge relative performance, its use needs to be applied cautiously. For instance, non-normal distributions with very diverse risk profiles can all have the same SR as shown by Bailey and Lopez de Prado [7], making the metric potentially misrepresentative of risk-adjusted returns. We will assess alternative performance measures in Sec.2.2 to account for the possibility of comparing different SR point estimates with potentially vastly different measures of uncertainty measured by their confidence bands [7].

1.2.3 Practical considerations

Statistical assumptions and model selection are not the only concerns when constructing portfolios. Consideration needs to be given to practical aspects of trading such as transaction costs, after-tax allocation, and the distinction between gross- and net-of-fee calculations. Depending on the type of assets and portfolios, transaction costs, including brokerage fees, bid-ask spreads, and other trading expenses, can significantly erode the net returns of an investment. This has a pronounced impact with higher frequencies of trading or regular portfolio rebalancing, where the costs accumulate over time [56].

Different types of investments are also subject to varying tax treatments and there is an important distinction between gross- and net-of-fee calculations when evaluating portfolio performance. For instance, dividends, interest income, and capital gains may be taxed at different rates, affecting one's net income and gross-of-fee returns represent the portfolio's performance before deducting management fees, performance fees, and other expenses [52]. Clients only realise net-of-fee returns which reflect the actual returns received after all fees and expenses have been deducted [5]. Understanding this distinction is vital because fees can significantly impact the net returns of a portfolio. High fees can erode the gains from investments, making it crucial to consider both gross and net returns when assessing the performance and cost-effectiveness of investment strategies.

1.3 Relevance of research

All investors—governments, asset managers, and individuals—aim to achieve favourable returns over a given horizon to meet their objectives. Yet, implementing the MPT framework to achieve this is challenging. As discussed, traditional portfolio optimisation relies heavily on historical data, which is both imperfect and laden with assumptions about means, covariances, and other key inputs. These assumptions (Sec.1.2), while necessary under conventional methods, often fail to capture the dynamic, non-ergodic nature of financial markets as shown in Sec.1.1, making historical performance a poor indicator of future outcomes. This reliance limits the realism of stress-testing portfolios for future shocks and introduces errors from overgeneralisation.

To build a more realistic and responsive portfolio, we want to reduce this dependence on historical data while also addressing the assumptions embedded in conventional optimisation inputs. We also need to ask: *“Is the sample data that we are using to represent some future scenario able to reliably represent that scenario?”* Scenarios here refer to the other possible paths that a time series may have taken *e.g.*, how would asset X have behaved if we were in world (scenario) “A” where inflation is high? *vs.* how asset X would have behaved if we were in world (scenario) “B” where inflation is low?

The non-parametric scenario-based approach of HS-FP, touched on in Sec.1.1.2 offers a practical solution by simultaneously achieving two key objectives: 1) it improves the realism of future scenario predictions by mapping historical data onto current market states using time- and market-conditioned probabilities. This transformation allows investors to evaluate how assets might behave under various plausible scenarios, and 2) reducing reliance on the assumptions required by traditional approaches. This shift from a purely historical perspective to a scenario-based framework allows for more intuitive and actionable portfolio insights. The goal is to stress-test portfolios against potential future outcomes in a way that aligns with the realities of today’s markets.

An added benefit of the framework is that it promotes client engagement and buy-in by incorporating investors’ views into the modelling process. This approach reframes portfolio strategies as a coherent narrative grounded in economic principles, making them more accessible and actionable for investors. For instance, clients can actively contribute by identifying state variables that align with their market outlook, creating a sense of ownership, and promoting transparency in the decision-making process.

1.4 Objective and structure of dissertation

This dissertation aims to address the limitations of traditional historical back-tests and portfolio construction methods by evaluating the benefits of reducing the number of assumptions and generating forward-looking estimates for portfolio optimisation inputs. It adopts a scenario-based, non-parametric approach, the HS-FP framework [89], to capture regime changes and enhance portfolio responsiveness to market shocks. This is conducted within a South African context, using Johannesburg Stock Exchange (JSE) equity and bond indices as the basis for portfolio construction and evaluation. The core of the analysis centres on two portfolio construction methods: MV optimisation [69] and HRP [29]. Additional benchmarks, such as balanced fund strategies and market indices, are incorporated to provide a comparative perspective.

The MV optimisation requires the estimation of the expected value and the variance of returns. In contrast, the HRP method only requires the variance estimation, reducing the overall potential estimation error. Using performance metrics such as the SR [119], the Probabilistic Sharpe Ratio (PSR) [7], and the Deflated Sharpe Ratio (DSR) [6], I assess whether the relative reduced number of parameter estimations required for the HRP results in improved OOS performance of the HRP relative to MV within a conventional back-test framework.⁹ I then introduce the HS-FP framework and review it to see if it improves either optimisation in any meaningful way.

A significant portion of this dissertation is dedicated to establishing a rigorous back-testing framework. This involves carefully addressing critical aspects such as the treatment of the risk-free asset, accurate return calculations, and performance measurements—both of which are crucial under geometric compounding over the longer periods that differentiate investment strategies from trading strategies. A comprehensive Matlab codebase is developed to facilitate reproducibility and wider applicability, reinforcing the practical value of this research.

Using the DSR and multiple simulations, I evaluate the reliability of key data processes, such as Winsorising returns [44], implementing cash constraints, and regularising covariance matrices. These steps ensure that results are dependable and applicable in practical contexts. The broader purpose of my research is to highlight the importance of, and acknowledge the limitations of many quantitative methods that are currently in use, specifically those that rely too heavily on historical realisations and limiting assumptions about how this can be resampled, sliced-and-diced for strategy-proving, testing and validation.

⁹While the two methods differ in their approaches—HRP focusing on diversification and mean-variance on risk-adjusted return—the reduction in estimated parameters also influences relative performance. The interplay of these factors will be discussed in Sec.2.5.2.

The dissertation is structured around two key objectives:

1. First, to assess the relative performance of MV optimisation *vs.* HRP, focusing on the HRP's reduced estimation requirements, which theoretically reduces compounded estimation errors.
2. Second, to evaluate whether incorporating the non-parametric scenario-based HS-FP framework enhances the effectiveness of either optimisation method by reliably capturing regime changes and reducing reliance on restrictive assumptions placed on financial time series data.

The structure of the dissertation aligns with these objectives. In Chp.2, the theoretical foundations of the MV and HRP methods are established, along with a discussion of additional performance measures, building on the theory introduced in Sec.1.2. This chapter transitions into practical implementation, covering the selection and processing of asset classes and data. Common benchmarks, such as equally weighted and balanced fund portfolios, are introduced, followed by a detailed explanation of the simulation design. Initial results are obtained, and the back-test framework is further refined before a conclusive model is selected, providing a robust basis for comparing the performance of the MV and HRP methods under a conventional back-testing approach.

Chp.3 transitions to the scenario-based HS-FP framework, where its performance is compared to the conventional methods presented in Chp.2. Scenario analysis is highlighted as a valuable tool to enhance transparency in market risk assessment and improve the modelling of portfolio P&L under abnormal market conditions [12]. To apply the HS-FP framework effectively, I outline the process of deriving invariants (*e.g.*, asset returns) from market drivers (*e.g.*, asset prices) and projecting these invariants to the investment horizon using flexible probabilities (FPrs) as defined by Meucci [85, 89]. To condition the FPrs on market state information such as macroeconomic variables, I establish a distinction between market state variables (*e.g.*, inflation) and market state signals (*e.g.*, inflation after undergoing signal processing).

The chapter then explores the various HS-FP methods to obtain FPrs, including time-conditioned, state-conditioned, combined time-and-state-conditioned, and multi-signal-conditioned approaches. Each method is defined and discussed in terms of its practical application within the stress-testing process. All these theoretical components are then integrated into a unified framework for stress-testing. This involves deriving forward-looking estimates of μ (*e.g.*, expected returns) and Σ (*e.g.*, covariance matrix) using FPrs that aim to capture market regimes effectively. These forward-looking estimates are then applied within the MV and HRP optimisations and the back-test frameworks.

Lastly in Chp.4, I consolidate what was achieved in each proceeding chapter by providing a chapter breakdown. I close by providing additional thoughts on possible future extensions of the research laid out in my dissertation.

Chapter 2

Conventional approach & practical implementation

As detailed in [Objective and structure of dissertation](#), this chapter focuses on the first objective of this dissertation: to compare the relative performance of MV optimisation and the HRP method. This comparison is made within the context of a conventional backtesting framework, allowing for an assessment of how differences in their underlying assumptions might influence estimation errors. To facilitate this analysis, a conventional backtest framework—a simulation applying strategies to historical data to evaluate their past performance—will be used. Before presenting these results, we review supporting theory for both methods, providing a detailed understanding of their optimisation methodologies and the evaluation metrics used.¹ The Probabilistic Sharpe Ratio (PSR) and the Deflated Sharpe Ratio (DSR) are introduced as supplements to the classic SR, with a view to enhancing the reliability of our performance evaluation by accounting for statistical properties and potential biases.

The results from this conventional back-test will then be analysed and documented, serving as a base comparison for Chp.3's scenario-based simulations.

2.1 Portfolio optimisations

Understanding the mechanisms behind the risk-adjusted-return-targeting MV and the diversification-focused HRP will provide us with valuable insights when interpreting their simulation results.

2.1.1 Comparison Mean-Variance & Hierarchical Risk Parity

The portfolio construction methods of MV and the HRP algorithms each have their strengths and weaknesses in the pursuit to achieve desirable risk adjusted returns. MV can be seen as an extension of Harry Markowitz's Critical Line Algorithm (CLA), first described in his 1952 paper "Portfolio Selection" [69]. His work earned him a Nobel Prize in 1990 along with the reputation as the pioneer of MPT at the time [70]. Several years later, the HRP algorithm was initially proposed by De Prado [29] in 2015 as a response to the limitations [De Prado](#) identified in the MV approach. This work was subsequently published in The Journal of Portfolio Management in 2016 [29]. In previous research I conducted in conjunction with

¹Certain components of theory within this chapter build upon and extend my previous Honours thesis work [76].

Siphehile Cele in 2021, two major contributing factors were observed to be central in the relative performance of the MV and the HRP algorithms: “1) The structure and quality of the data being used and 2) the level of risk present in the market at the time of implementation.” [76]. These two factors stem from the respective shortfalls within each algorithm.

As MV is a quadratic optimisation problem built around stringent linear and quadratic constraints, it results in a control problem where the optimal portfolio weights are proportional to the inverse covariance matrix, Σ^{-1} multiplied by said constraints. A positive-definite and consequently invertible covariance matrix Σ , is therefore vital to the MV framework. Yet when the quality of the data is poor an inverse covariance matrix will become unstable with an increased condition number leading to the phenomenon that Michaud and Michaud [90] refers to as “error-maximisation”. Added difficulties arise during market shocks as higher correlations and spiking volatilities increase a covariance matrix’s condition number making it more ill-conditioned. As the condition number of a covariance matrix increases within the unconstrained quadratic portfolio optimisation, the more the optimal weights will diverge and seem less and less intuitive [39]. This speaks to the importance of the data structure and quality in MV’s relative performance. MV optimisation is frequently solved as a deterministic problem, which effectively disregards uncertainty inherent in the inputs and the potential estimation error [56]—relying solely on point estimates for expected returns and the covariance matrix of returns, assuming these estimates are without error in the portfolio allocation process.

However, HRP, which incorporates unsupervised learning techniques through hierarchical clustering, excels in uncovering complex structures in high-dimensional and highly correlated data, relying only on a single-parameter estimate of the covariance matrix [98]. As HRP employs cluster analysis, it does not rely on the actual entries in the covariance matrix, but rather on the structure of the matrix. Consequently a portfolio constructed using HRP has far greater diversification than a MV portfolio which tends to overweight a select few assets. This increased diversification gives an HRP portfolio a noticeable advantage over MV as it provides a cushioning effect during unfavourable market conditions, leading to out-performance during times of systemic or idiosyncratic shocks. Nevertheless, HRP is a risk management approach; therefore, maximising returns is not a clear benefit of HRP.

Extensive efforts have been made to address the shortcomings of both the MV optimisation and HRP. Subsequent approaches sought to address the sensitivity of MV optimisation to estimation error. A particularly influential extension is the Black and Litterman [17] model, which combines equilibrium returns implied by market capitalisation weights with investor views to generate more stable and intuitive expected returns. This framework remains widely used in practice by asset managers and provides a more tractable approach to portfolio optimisation than relying solely on raw sample estimates.

An early contribution to improving covariance conditioning was provided by Higham [45], who proposed a method for finding the nearest positive semidefinite matrix to a given (possibly indefinite) covariance estimate. Higham’s algorithm ensures that an estimated covariance matrix is mathematically admissible, which is particularly important in high-dimensional settings where noisy estimates can easily lose positive definiteness. This work was the precursor to later approaches, such as Ledoit and Wolf [59] who’s work addressed the related but distinct issue of stability and estimation error by introducing shrinkage esti-

mators. Their method constructs a linear combination of the sample covariance matrix and a structured target, typically the identity matrix, to reduce the condition number and improve the stability of the inverse. More recently, Ollila et al. [95] proposed non-linear eigenvalue transformations that further adjust the sample eigenvalues to mitigate the effects of noisy estimation, providing a more flexible approach to covariance regularisation.

Building on the geometric framework introduced by Golts and Jones [39] regarding α -weight angles, Dalmeyer and Gebbie [28] further explore the mathematical relationship between the condition number of a covariance matrix and the α -weight angle. This angle measures the divergence between an investor's expected returns (α) and the optimised portfolio weights generated by unconstrained MV optimisation. If you have a vector α of expected returns and the optimal weight allocation vector ω^* , one wants the angle between the two vectors to be within tight bounds as it reflects that the optimisation outcomes are relatively aligned with an investor's insights. If the angle is tending towards orthogonality, where the angle is widening towards 90 degrees, the relationship between the α 's and the portfolio positions diverge. The numerical value of the α -weight angle was shown by Golts and Jones [39] to have an upper bound contingent on the condition number of the matrix.

Dalmeyer and Gebbie [28] make a point of showing that it is not only the concern of error maximisation that is caused by covariance matrix inversion that needs to be addressed, but that one needs to focus on the direction in which the optimal portfolio is pushed away from the expected returns with increasing condition numbers. They demonstrate how to improve the robustness of MV portfolio construction when reducing the α -weight angle by improving the conditioning of the covariance matrix. Golts and Jones [39] proposed the separation of the *magnitude* and *direction* of the positions vector ω^* . By doing so they were able to isolate the *direction* component of the vector and implement robust optimisation procedures in order to generate a better-conditioned covariance matrix. This had the impact of constraining the α -weight angle, bringing the optimisation outcome closer to the investor's expected returns. Taking this a step further, Dalmeyer and Gebbie [28] extended the collection of robust optimisation to incorporate gearing constraints and looked at several other permutations of the optimisation.

The covariance matrix is not the only concern of MV. Paskaramoorthy et al. [99] explored the effect of distribution misspecification on the MV efficient frontiers, highlighting the problems that accompany re-sampling and cross validation processes. They show that the empirical performance of a MV that comprises of predictions and an optimisation is dependent of the prediction process. Specifically they show the improvement of the covariance matrix is of greater importance than the prediction of the means. This would link back to the consensus of fellow researcher above who focus primarily on the covariance matrix when addressing MV under performance. We therefore see that MV is highly dependent on the "information" it is fed, hence we want to enrich the data inputs by aligning them more with current market information - but this we will keep for Chp.3.

Remark 1. Chopra and Ziemba [23] have previously shown that the errors created from estimating the sample mean has a larger impact on the OOS performance than that resulting from estimating the sample covariance. This suggests that the HRP may be more reliable OOS as it only needs the covariance estimate and MV requires the estimate of both.

2.1.2 Sharpe Ratio maximising Mean-Variance optimisation

Introduced as part of Markowitz [69]’s seminal work on efficient frontier theory in 1952, the MV portfolio represents one of the many possible optimisations along the frontier, balancing risk and return. By extending the MV framework to include additional objective functions, such as maximising the SR, it is possible to derive a portfolio with the highest risk-adjusted returns. This dissertation adopts this specific formulation—the SR-maximising MV portfolio—as the foundation for analysis.

Below in Def.18 we define the classical MV portfolio before defining the objective function for the SR maximising version of it.

Definition 18 (Markowitz Mean-Variance (MV) Portfolio Selection). For the MV algorithm, it will be assumed that assets are fully invested and that the mean $\boldsymbol{\mu}$ and the covariance Σ of the returns are known through sample estimation of whatever time-series input is provided: $\hat{\boldsymbol{\mu}}, \hat{\Sigma}$. The aim of the MV portfolio selection problem is to minimise the overall portfolio risk at a fixed level of return given the expected portfolio return $\mathbb{E}[R_P]$:

$$\mathbb{E}[R_P] = \boldsymbol{\omega}'\boldsymbol{\mu} = \sum_{t=1}^T \omega_t \mu_t. \quad (2.1)$$

Then the MV optimisation objective function can be defined as:

$$\operatorname{argmin}_{\boldsymbol{\omega}} \left\{ \boldsymbol{\omega}'\hat{\Sigma}\boldsymbol{\omega} \right\} \quad \text{s.t } \boldsymbol{\omega}'\hat{\boldsymbol{\mu}} = \mathbb{E}[R_P] \text{ and } \boldsymbol{\omega}'\mathbf{1} = 1, \quad (2.2)$$

where $\boldsymbol{\omega}$ is a vector of portfolio weights ($\boldsymbol{\omega}'$ denoting the transpose) and $\mathbf{1}$ is a vector of ones of equal length. The set of weights for different return expectations then define the Pareto optimal set known as the efficient frontier as a function of return and portfolio variance.

We can now extend the classical MV in Eq. 2.2. It is an optimisation based on maximising Sharpe [119]’s popular portfolio performance measure, the SR (as introduced in Sec.1.2.2.4). As previously discussed, this statistic is defined as the ratio between a portfolio’s average excess return and the risk of the portfolio. The risk of the portfolio is captured by the standard deviation of the returns, σ_P :

$$\hat{\sigma}_P = \sqrt{\boldsymbol{\omega}'\hat{\Sigma}\boldsymbol{\omega}}. \quad (2.3)$$

The portfolio’s average excess return is a function of $\mathbb{E}[R_P]$ and the rate of return paid by a risk-free asset, known as the risk-free rate, R_f . This can then be coupled with Eq.2.3 to define the SR:

$$SR = \frac{\mathbb{E}[R_P] - R_f}{\hat{\sigma}_P}. \quad (2.4)$$

By formulating the optimisation in terms of the maximisation of the SR we get the overall portfolio objective function².

Definition 19 (SR Maximising MV Portfolio Objective Function).

$$\max_{\omega} \left(\frac{\mathbb{E}[R_P] - R_f}{\hat{\sigma}_P} \right) = \max_{\omega} \left(\frac{\omega' \hat{\mu} - R_f}{\sqrt{\omega' \hat{\Sigma} \omega}} \right) \text{ s.t } \omega' \mathbf{1} = 1 \text{ and } \omega \geq 0. \quad (2.5)$$

The use of the SR in performing MV makes intuitive sense as the portfolio defined by the weights obtained from Eq.2.5 would be maximising returns subject to a level of risk, making it a member of the efficient frontier as put forward by Markowitz [69].

It is important to note that from here on out, throughout the dissertation, reference to the MV portfolio pertains exclusively to this SR-maximising variant of the MV.

2.1.3 Hierarchical Risk Parity

HRP uses the information held in the covariance matrix regarding the assets relationships with one another to generate clusters based on similarities. Therefore, it is not as reliant on the covariance matrix as the MV approach as it is not specifically using the empirical values housed in the matrix (this point will become clear in a moment). The key to the algorithm is that it is primarily concerned with the **order** in which the linkages are formed when clustering rather than the solution clusters themselves. This ordering is instrumental in assigning weights to assets further on in the algorithm. The algorithm flows through 3 steps that have been well defined by De Prado [29] in his 2016 paper. These 3 steps are executed in the following order: Tree Clustering, Quasi-diagonalization and Recursive Bisection.

2.1.3.1 Step 1: tree clustering

The first step follows an agglomerative clustering method and begins with computing an $J \times J$ correlation matrix ρ . The correlations are computed from an $T \times J$ dataset, consisting of the returns of J assets observed over T time points. The entries of ρ are indexed as $\rho_{i,k} = \rho[X_i, X_k]$ which is the correlation between the i -th and k -th asset for $i, k = 1, \dots, J$. The initial $J \times J$ correlation distance matrix D is constructed capturing the distances between assets based on their correlation coefficients. If the Cartesian product of assets $\{1, \dots, J\}$ is B , then the distance measure can be defined as $d : (X_i, X_k) \in B \rightarrow \mathbb{R} \in [0, 1]$. These distances are given by:

$$d_{i,k} = \sqrt{\frac{1}{2}(1 - \rho_{i,k})}. \quad (2.6)$$

$D = \{d_{i,k}\}_{i,k=1,\dots,J}$ is therefore a correlation-distance matrix.

²Note that additional constraints such as no-short-selling can be added as is the case in Eq.2.5 where $\omega \geq 0$. However, when a no-short-selling constraint ($\omega \geq 0$) is imposed, the problem becomes a quadratic program with inequality constraints, for which no general closed-form solution exists. In these cases, numerical optimisation techniques (e.g., quadratic programming or interior-point methods) would need to be used to obtain the optimal weights. See Apx.4 for a mathematical proof of the SR-maximising portfolio weights under a budget constraint, allowing short-selling. For the treatment of more complex optimisations, see Dalmeyer and Gebbie [28] for further derivations.

Part 2 of this step is pairing the column vectors of $\mathbf{D}$ and calculating the Euclidean distances between pairs where the distances between any two columns vectors of $\mathbf{D}$ is:

$$\tilde{d}_{i,k} = \sqrt{\sum_{j=1}^J (d_{j,i} - d_{j,k})^2}. \quad (2.7)$$

Here the new distances are $\tilde{d} : (D_i, D_k) \subset \mathbf{B} \rightarrow \mathbb{R} \in [0, \sqrt{J}]$. These form the final distance matrix $\tilde{\mathbf{D}}$ which is effectively a distance-of-distances matrix where $\tilde{d}_{i,k} = \tilde{d}[D_i, D_k]$. $\tilde{\mathbf{D}}$ gives an indication of the closeness in similarity of the assets with the rest of the portfolio as each $\tilde{d}_{i,k}$ is a function of the correlation matrix in its entirety [125].

With the distance measures now defined, the agglomerative clustering can begin to form hierarchical clusters in a bottom-up manner—starting with J individual assets and clustering them until they all form part of a single parent cluster. Columns of $\tilde{\mathbf{D}}$ are clustered into pairs (i^*, k^*) , with the first one formed being $u[1]$ such that $(i^*, k^*) = \operatorname{argmin}_{(i,k)_{i \neq k}} \{\tilde{d}_{i,k}\}$. Taking the minimum distance in a hierarchical cluster analysis is referred to as the single-linkage criterion. The distances between clusters are defined by individual asset pairs that are strictly not from the same clusters but are closest to one another given this condition. The distance between $u[1]$, the newly formed cluster, and the remaining unclustered assets is established using:

$$\dot{d}_{i,u[1]} = \min \left[\left\{ \tilde{d}_{i,k} \right\}_{k \in u[1]} \right], \quad (2.8)$$

where i denotes the i -th item in $\tilde{\mathbf{D}}$. In order to update $\tilde{\mathbf{D}}$, the rows and columns corresponding to $u[1]$ are dropped and $\dot{d}_{i,u[1]}$ is appended on both dimensions of the matrix. To continue the cluster formation we used the updated $\tilde{\mathbf{D}}$ and once again select the asset with the smallest distance from the cluster to form part of a new cluster. This is recursively done until all assets are grouped under a single cluster which causes the clustering portion of the algorithm to terminate.

A linkage matrix $\mathbf{Y}$ can now be defined by a $(J - 1) \times 4$ matrix. $\mathbf{Y}$ houses information on all m clusters, $m = 1, \dots, (J - 1)$. Each row contains a 4-tuple corresponding to each cluster formation *i.e.* each row corresponds to the merging of two branches into a single branch. A 4-tuple is structured as: $(y_{m,1}, y_{m,2}, y_{m,3}, y_{m,4})$. Here the first two items for the m -th cluster is $y_{m,1}$ and $y_{m,2}$, and they record the constituents of the cluster. The third item $y_{m,3}$ is the distance between $(y_{m,1}, y_{m,2})$ as captured by $\tilde{d}_{y_{m,1}, y_{m,2}}$. This linkage matrix structure was defined by De Prado [29], however, in his paper he neglected to explicitly define what information is contained in the the final entry of the tuple, $y_{m,4}$. When looking at the Python code snippet he provides (*“Code snippet 1 - Quasi-diagonalization”*), however, it seems that the answer to this is that $y_{m,4}$ may hold information on the number of original items present in the m -th cluster.

2.1.3.2 Step 2: quasi-diagonalization

The content of the linkage matrix can now be put to use. In order to diagonalise the original covariance matrix the columns and rows need to be rearranged in such a way that the resulting matrix has the large values along the diagonal and the smaller values on the off-diagonal. The effect this has on the assets is to organise them according to their similarity/dissimilarity—dissimilar assets are positioned furthest apart and most similar assets are placed nearest to one another.

For this to be achieved the algorithm uses each row of $\mathbf{Y}$, replacing the clusters in $(y_{J-1,1}, y_{J-1,2})$ with the constituents until all clusters are dissolved and only individual items remain. This results in a sorted item list consisting of the N original assets which can now be used to rearrange the covariance. Visualising the final quasi-diagonalised matrix in a heatmap gives a visual representation of the inherent clusters within the data when using the hierarchical order of clusters. An example of this can be found in De Prado's paper "Building diversified portfolios that outperform out-of-sample" [29].

2.1.3.3 Step 3: recursive bisection

The final yet arguably the most important step is that of the recursive bisection as it is where assets are assigned weights that will characterise the end composition of the HRP portfolio. With a diagonalised covariance matrix the optimal way to assign weights is by using the inverse-variance allocation.³ The inverse-variance weighting approach is advantageous as it can be applied at each branching stage in a top-down fashion, splitting weight allocations between sub-clusters of a parent cluster in proportion to the cluster's variance (using a splitting factor). It can also be taken advantage of in a bottom-up manner, defining the variance of a sub-cluster using the inverse-variance weighting of the constituents in said sub-cluster.

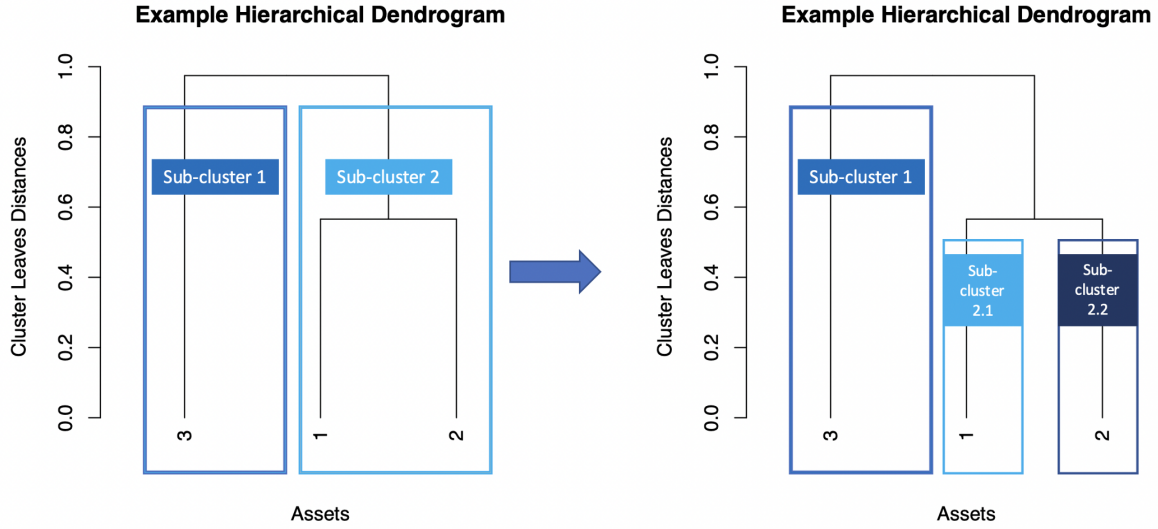
De Prado [29] implements the algorithm by making use of the sorted list created at the end of *Step 2* and denotes it as $L = \{L_0\}$ where $L_0 = \{j\}_{j=1,\dots,J}$. All J assets are assigned an initial weight of 1 to begin the algorithm:

$$w_j = 1, \forall j = 1, \dots, J. \quad (2.9)$$

He recursively bisects the list and implements the weight allocation at each iteration.

Mirroring De Prado [29]'s ordered list and recursive bisectioning in a visual approach we use the results from *Step 1* to construct the hierarchical dendrogram in a 3 asset case in Fig. 2.1 below. For the equations to follow a tilde will be placed above variables when referring to components of sub-clusters, such as a vector that contains weights for the k -th sub-cluster of the i -th parent cluster: $\tilde{\omega}_i^{(k)}$.

³Supporting theory for inverse-variance allocation can be found in A2.2.1 Inverse-variance allocation.



(a) First iteration of the recursive bisection.

(b) Second iteration of the recursive bisection.

Figure 2.1: A visual representation of the Recursive Bisectioning algorithm and the partitioning that takes place at each iteration. Fig. 2.1a illustrates the first partition of the largest parent cluster L_0 into its two sub-clusters. Fig. 2.1b shows the second iteration of the algorithm where the sub-clusters from 2.1a become the parent clusters that are bisected. Sub-cluster 1 consists of only 1 asset therefore sub-cluster 2 is the only one bisected.

In Fig. 2.1 the L_0 parent cluster containing all the assets is composed of nested sub-clusters forming a binary tree where each parent cluster L_i has two branches $L_i^{(1)}$ and $L_i^{(2)}$. In Fig. 2.1a we have the first bisection, splitting L_0 into its two sub-clusters $L_0^{(1)}$ and $L_0^{(2)}$ where $L_0^{(1)}$ consists of asset 3 and $L_0^{(2)}$ of 1 and 2. Within the iteration the variance of each sub-cluster needs to be established. The variance of the k -th sub-cluster of parent cluster L_i is denoted as $\tilde{V}_i^{(k)}$ and takes the quadratic form:

$$\tilde{V}_i^{(k)} = \tilde{\omega}_i^{(k)'} \Sigma_i^{(k)} \tilde{\omega}_i^{(k)}. \quad (2.10)$$

Here $\Sigma_i^{(k)}$ is the covariance matrix between the constituents of the k -th sub-cluster in the i -th parent cluster and $\tilde{\omega}_i^{(k)}$ is a vector of weights for the constituents in said sub-cluster. $\Sigma_i^{(k)}$ is sliced from the diagonalised covariance matrix created in *Step 2* and contains only the covariances between the assets in the sub-cluster of interest, $L_i^{(k)}$.

$\tilde{\omega}_i^{(k)}$ in Eq. 2.10 is defined as:

$$\tilde{\omega}_i^{(k)} = \frac{\text{diag}[\Sigma_i^{(k)}]^{-1}}{\text{trace}(\text{diag}[\Sigma_i^{(k)}]^{-1})}. \quad (2.11)$$

Using sub-cluster 2, $L_0^{(2)}$, as an example, $\tilde{V}_0^{(2)}$ can be established in a bottom-up manner by using the weights vector $\tilde{\omega}_0^{(2)}$ which contains the weights of asset 1 and 2:

$$\tilde{\omega}_0^{(2)} = \frac{\text{diag}[\Sigma_0^{(2)}]^{-1}}{\text{trace}(\text{diag}[\Sigma_0^{(2)}]^{-1})}. \quad (2.12)$$

This is used along with the covariance matrix of assets 1 and 2 to calculate sub-cluster 2's variance:

$$\tilde{V}_0^{(2)} = \tilde{\omega}_0^{(2)'} \Sigma_0^{(2)} \tilde{\omega}_0^{(2)}. \quad (2.13)$$

The same process is conducted for sub-cluster 1.

Moving on to distributing the variance between sub-clusters of L_0 in a top-down manner a splitting factor is computed. This split factor allows for the re-scaling of the asset weights in the final phase of each iteration. The splitting factor for a binary case is defined by Eq.2.14:

$$\alpha_i = 1 - \frac{\tilde{V}_i^{(1)}}{\tilde{V}_i^{(1)} + \tilde{V}_i^{(2)}}, \quad 0 \leq \alpha_i \leq 1. \quad (2.14)$$

We re-scale all weights dependent on their sub-cluster membership, scaling the weights of all the assets in sub-cluster $L_0^{(1)}$ by a factor of α_0 and those in $L_0^{(2)}$ by $(1 - \alpha_0)$. Updating the weights as follows:

$$\tilde{\omega}_0^{(1)} = \alpha_0 * \tilde{\omega}_0^{(1)}, \quad (2.15)$$

$$\tilde{\omega}_0^{(2)} = (1 - \alpha_0) * \tilde{\omega}_0^{(2)}. \quad (2.16)$$

With the weights now adjusted using a single hierarchical level the procedure can be iterated, moving down the tree at each iteration. Fig.2.1b represents the second iteration for this example case. We see sub-cluster 2 being further partitioned while sub-cluster 1 consisting of a single asset is not. With each round of partitioning, weights are distributed according to the method above until all partitioning ceases due to all clusters consisting solely of a single asset. Therefore, the second iteration of the example would be the final iteration.

These steps ensure that $0 \leq \omega_j \leq 1, \forall j = 1, \dots, J$, and $\sum_{j=1}^J \omega_j = 1$.

2.2 Additional performance measures for advanced comparisons

To accurately evaluate the performance of modern portfolio strategies under real-world conditions, we need to consider advanced performance measures beyond the classical SR. In this section we define the theory of the PSR, the DSR, as well as discuss further metrics that can enhance our evaluation of portfolios when stress-testing.

2.2.1 Extending the Sharpe Ratio

The SR, as introduced in Sec.1.2.2.4, is a widely used metric when evaluating the performance of investment portfolios due to its simplicity and intuitive appeal. However, the measure of uncertainty around the SR estimator $\widehat{SR}$ is determined by the statistical properties of the returns R_t [6, 7, 8, 66]. Sharpe [119] himself states in his initial paper debuting the SR that “*comparisons based on the first two moments of a distribution do not take into account possible differences among portfolios in other moments or in distributions of outcomes across states of nature that may be associated with different levels of investor utility*”. The limiting assumptions detailed in Sec.1.2—namely that the underlying returns data is *i.i.d.*, normally distributed, stationary, and ergodic—are often violated in real-world financial data, potentially

leading to misleading conclusions about portfolio risk-adjusted performance. To address these challenges, this section introduces additional performance metrics designed to mitigate the impact of these limitations and serve as intuitive guardrails when interpreting simulation results.

Lo [66] discusses the SR distribution under the assumptions of normality and *i.i.d.* data in-depth, showing that under the assumption of *i.i.d.* normal returns the $\widehat{SR}$ estimate can be reduced to:

Definition 20 (Asymptotic distribution of $\widehat{SR}$ assuming *i.i.d.* normal returns).

$$(\widehat{SR} - SR) \xrightarrow{a} \mathcal{N}\left(0, \frac{1 + \frac{1}{2}SR^2}{T}\right), \quad (2.17)$$

where $\xrightarrow{a}$ denotes asymptotic convergence via the CLT (see Sec.1.2.1.3), $\mathcal{N}$ denotes the normal distribution, and SR is as defined in Eq.1.30.

Efforts by Mertens [81] and separately by Christie [24] (proven to be equivalent by Opdyke [96]) showed that the distribution of the SR can still be found to be normal if one were to only assume stationary and ergodic returns. The implication of these lesser assumptions would allow for non-*i.i.d.* non-normal returns, time-varying conditional volatilities, and serial correlation.

Definition 21 (Asymptotic distribution of $\widehat{SR}$ assuming *i.i.d.* non-normal returns).

$$(\widehat{SR} - SR) \xrightarrow{a} \mathcal{N}\left(0, \frac{1 + \frac{1}{2}SR^2 - \mu_3SR + \frac{\mu_4 - 3}{4}SR^2}{T}\right). \quad (2.18)$$

Here $\mu_1 = \mu, \mu_2 = \sigma, \mu_3 = \frac{\mathbb{E}[r - \mu]^3}{\sigma^3}$ denotes skewness, and $\mu_4 = \frac{\mathbb{E}[r - \mu]^4}{\sigma^4}$ is kurtosis of the strategy's realised returns R_t .

Eq.2.18 implies that $\widehat{SR}$ will still follow a normal distribution even if R_t does not. However, by adjusting the distributional properties of the SR to account for non-normal returns as in Eq.2.18 the standard deviation and measure of uncertainty of $\widehat{SR}$ is significantly impacted by the incorporation of skewness and kurtosis of R_t .

The standard deviation of the point estimate when assuming *i.i.d.* normal returns is given by:

Definition 22 (Standard deviation of $\widehat{SR}$ assuming *i.i.d.* normal returns).

$$\sigma(\widehat{SR}) = \sqrt{\frac{1}{T-1} \left(1 + \frac{1}{2}\widehat{SR}^2\right)}. \quad (2.19)$$

Note that $T - 1$ is accounting for Bessel's correction [7].

When accounting for the skewness and kurtosis of R_t this becomes:

Definition 23 (Standard deviation of $\widehat{SR}$ assuming non-*i.i.d.* non-normal returns).

$$\hat{\sigma}_{\widehat{SR}} = \sqrt{\frac{1 - \widehat{\mu}_3 \widehat{SR} + \frac{\widehat{\mu}_4 - 1}{4} \widehat{SR}^2}{T - 1}}. \quad (2.20)$$

This suggests that non-normal distributions can substantially exaggerate $\widehat{SR}$. Consequently, a high SR alone may not reliably indicate its statistical significance. This impact can be seen if we define the measurement of uncertainty using the confidence band around the point estimate of SR.

Definition 24 (Confidence band around $\widehat{SR}$ with significance level α).

$$\text{Prob} \left[SR \in \left(\widehat{SR} - Z_{\alpha/2} \hat{\sigma}_{\widehat{SR}}, \widehat{SR} + Z_{\alpha/2} \hat{\sigma}_{\widehat{SR}} \right) \right] = 1 - \alpha, \quad (2.21)$$

where $\hat{\sigma}_{\widehat{SR}}$ is defined by Eq. 2.20 and α is a chosen significance level.

The PSR and the DSR have been developed to provide more robust performance evaluations. The PSR seeks to measure the confidence investors can have in the observed SR while the DSR is an adjusted metric of the SR accounting for non-*i.i.d.* returns.

2.2.1.1 Probabilistic Sharpe Ratio

The PSR statistic proposed by Bailey and Lopez de Prado [7] provides a probabilistic framework to assess the statistical significance of the SR and aims to account for possible inflationary effects caused by the skewness and/or kurtosis of returns. It calculates the probability that the observed SR $\widehat{SR}$ exceeds a benchmark SR^* under non-Normal returns, considering the variability and distribution of returns. This approach helps in understanding the level of confidence one can have in the performance measure. By incorporating the variability and distribution of returns, the PSR offers a more robust evaluation of performance in the presence of non-normal returns and serial correlations.

The PSR allows investors to ascertain the probability of a false positive discovery occurring *i.e.* to assess whether a chosen portfolio strategy produces a favourably high SR because it is genuinely a top performing strategy or whether the reported SR is purely by chance as may be the case when an increasing number of back-tests are performed [67]. The SR defined in Eq 2.18 can be written into a probabilistic form using the expression in Def. 24:

Definition 25 (Probabilistic expression of the Sharpe Ratio given a benchmark).

$$\widehat{PSR}(SR^*) = \text{Prob}[SR \leq \widehat{SR}] = \int_{-\infty}^{\widehat{SR}} z(SR | SR^*, \hat{\sigma}_{\widehat{SR}}) \cdot dSR, \quad (2.22)$$

where $z(SR | SR^*, \hat{\sigma}_{\widehat{SR}})$ is the pdf of the Normal distribution with the first two moments being SR^* and $\hat{\sigma}_{\widehat{SR}}$ respectively.

To assess the probability that the estimated SR is greater than an hypothetical predefined benchmark SR, $\widehat{PSR}(SR^*)$ we formally define the PSR:

Definition 26 (Probabilistic Sharpe Ratio (PSR)).

$$\widehat{PSR}(SR^*) = Z \left[\frac{(\widehat{SR} - SR^*) \sqrt{T-1}}{\sqrt{1 - \hat{\mu}_3 \widehat{SR} + \frac{\hat{\mu}_4 - 1}{4} \widehat{SR}^2}} \right], \quad (2.23)$$

where Z is the CDF of standard normal distribution. See Bailey and Lopez de Prado [7] for further discussion.

2.2.1.2 Deflated Sharpe Ratio

The DSR was introduced by Bailey and De Prado [6] in 2014 after their contribution of the PSR in 2012 to adjust the SR for the non-*i.i.d.* nature of the underlying returns by incorporating the effects of skewness and kurtosis of R_t . The hope was that it would provide a more accurate measure of performance by considering the actual distribution of returns rather than assuming normality.

When performing and reporting on the back-testing process, investment managers tend to report outcomes that are favourable to their objective which is considered as selection bias and fail to report the process in which unfavourable results are discarded throughout the process of performing back-testing. The introduction of the DSR as a performance measure assists in mitigating the effects of the two leading sources of performance inflation: Selection bias under multiple testing and non-normally distributed returns [6]. Even though back-testing is an important tool for measuring the strength of an investment strategy, one of its caveats is that it can be strategically manipulated to reflect supposedly favourable outcomes [8]. Since critical information about the strategy is withheld, the probability of executing a strategy that should not have been picked (type 1 error) increases and leaves the decision maker vulnerable. This is a false positive result [8]. Back-test overfitting happens when the optimisation of a set of parameters occurs in order to maximise the performance of a back-test [9].

Bailey et al. [8] argues that due to the competitive environment amongst investment managers, the ratio of signal to noise in financial series is low. This increases the probability of discovering a chance configuration rather than an actual signal making it hard to avoid back-test overfitting. Journals and other publications often claim profitable strategies but only report performance metrics from in-sample (IS) testing. This can lead to performance inflation, as multiple tests are conducted and only the best-performing strategies are shared, leading to selection bias. These top-performing strategies often suffer OOS due to the performance inflation observed in IS [8]. The DSR accounts for the multiple testing problem, which arises when evaluating multiple strategies or portfolios. By deflating the SR, it reduces the likelihood of falsely identifying a strategy as superior due to random chance.

The DSR attempts to correct for two sources of performance inflation:

1. **Selection Bias:** This is addressed by considering the number of independent trials conducted in the investment strategy selection process (N).
2. **Non-Normal Returns:** These are accounted for by incorporating information about the length of returns (T), the expected variance of the Sharpe ratio estimates from N trials, and the skewness and kurtosis of the returns.

To calculate the DSR we build on the initial theory laid out for the PSR but this time the benchmark SR^* is defined by the expected maximum estimated SR, $\mathbb{E} \left[\max \{ \widehat{SR}_k \} \right]$ out of K independent simulated trials. To obtain SR^* , suppose there are K independent trials, where each trial k would have an associated SR estimate, $\widehat{SR}_k$.⁴ Each trial consists of alternative hyper-parameter configurations that need to be generated. However, not all configurations will be independent when simulated, hence K would need to be estimated from a total of M trials. For this, we will use the unsupervised learning methods of López de Prado and Lewis [67] to estimate $\mathbb{E}[K]$ as the optimal number of clusters from M simulated trials. We begin with defining the False Strategy Theorem to estimate $\mathbb{E} \left[\max \{ \widehat{SR}_k \} \right]$ as the expected SR under multiple trials.

False strategy theorem: expected Sharpe Ratio under multiple trials: As defined by Bailey and De Prado [6] each element of the set of K independent simulations is called a trial, and it is associated with a SR estimate, $\widehat{SR}_k$, with $k = 1, \dots, K$.

Theorem 6 (Expected Maximum Sharpe Ratio under Multiple Trials (False Strategy Theorem)). Let $\{\widehat{SR}_k\}_{k=1}^K$ denote a set of K independent SR estimates obtained from K independent trials, where each estimate is assumed to be an *i.i.d.* draw from a normal distribution,

$$\widehat{SR}_k \sim \mathcal{N} \left(0, \text{Var} \left[\{ \widehat{SR}_k \} \right] \right).$$

Then, the expected value of the collection of maximum SRs across these K trials is approximately given by

$$\mathbb{E} \left[\max_k \{ \widehat{SR}_k \} \right] \approx \sqrt{\text{Var} \left[\{ \widehat{SR}_k \} \right]} \left[(1 - \gamma) Z^{-1} \left(1 - \frac{1}{K} \right) + \gamma Z^{-1} \left(1 - \frac{1}{Ke} \right) \right], \quad (2.24)$$

where $Z^{-1}[\cdot]$ denotes the inverse of the standard Gaussian CDF, e is Euler's number, and $\gamma \approx 0.5772$ is the Euler-Mascheroni constant.

Eq.2.24 implies that unless $\max_k \{ \widehat{SR}_k \} \gg \mathbb{E} \left[\max_k \{ \widehat{SR}_k \} \right]$ the discovered strategy that is being evaluated is likely to be a false positive *i.e.* the strategy is likely to be a “False Strategy”. See Bailey and De Prado [6] for derivation of Thm.6.

⁴Here, each $\widehat{SR}_k$ is an estimator computed from the random returns in trial k , making it a random variable itself. Consequently, the maximum over these random variables, $\max \{ \widehat{SR}_k \}$, is also a random variable as it depends on the particular draw of all $\widehat{SR}_k$. The expectation operator $\mathbb{E}[\cdot]$ is used to define the average value of this maximum across many hypothetical sets of K trials. Please see Bailey and De Prado [6] for a more in-depth discussion.

Defining the DSR: To account for the estimation impact of multiple trials in the DSR a new benchmark $SR^* = \widehat{SR}_0$ is defined for the PSR and substituted into the PSR Eq. 2.23. Here we set $\widehat{SR}_0 = \mathbb{E} \left[\max_k \{ \widehat{SR}_k \} \right]$ in Eq.2.24 making the expected maximum SR the new benchmark that is used to assess the probability of observing a truly outperforming SR.

Theorem 7 (Deflated Sharpe Ratio Benchmark SR^*). Let $\{\widehat{SR}_k\}_{k=1}^K$ denote a set of K independent SR estimates obtained from different trials or hyperparameter configurations, each assumed to be an *i.i.d.* draw from a normal distribution,

$$\widehat{SR}_k \sim \mathcal{N}\left(0, \text{Var}\left[\{\widehat{SR}_k\}\right]\right).$$

Then, following the expected maximum SR shown in Thm.6, the benchmark SR^* is given by

$$\widehat{SR}_0 = \sqrt{\text{Var}\left[\{\widehat{SR}_k\}\right]} \left[(1 - \gamma)Z^{-1}\left(1 - \frac{1}{K}\right) + \gamma Z^{-1}\left(1 - \frac{1}{Ke}\right) \right], \quad (2.25)$$

Here, $\sqrt{\text{Var}\left[\{\widehat{SR}_k\}\right]}$ represents the variance of SRs across K independent hyper-parameter configurations. As before, γ is the Euler-Mascheroni constant, e is Euler's number, $Z[\cdot]$ is the CDF of the standard Normal distribution, and $k = 1, \dots, K$.

The expression in Thm.7 follows directly from the expected maximum SR under multiple trials (Thm.6) as derived in Bailey and De Prado [6]. Substituting in the newly defined benchmark from Eq.2.25 into the PSR of Eq.2.23 we define the DSR:

Theorem 8 (Deflated Sharpe Ratio). The DSR provides an adjusted measure of statistical confidence in an observed SR by correcting for both non-normality in returns and the inflationary effect of multiple testing. It is obtained by substituting the benchmark $SR^* = \widehat{SR}_0$ from Thm.7 into the PSR formulation in Eq.2.23. The resulting expression is

$$\widehat{DSR} = \widehat{PSR}[\widehat{SR}_0] = Z \left[\frac{(\widehat{SR} - \widehat{SR}_0) \sqrt{T-1}}{\sqrt{1 - \widehat{\mu}_3 \widehat{SR} + \frac{\widehat{\mu}_4 - 1}{4} \widehat{SR}^2}} \right], \quad (2.26)$$

Here $\mu_3 = \frac{\mathbb{E}[r - \mu]^3}{\sigma^3}$ denotes skewness and $\mu_4 = \frac{\mathbb{E}[r - \mu]^4}{\sigma^4}$ is kurtosis of the strategy's realised returns R_t .

Ultimately in this context of the DSR, we are assessing the probability of performance inflation, which can stem not only from the non-normal distribution of asset class returns but also from potential selection bias introduced during multiple testing. The DSR is a modification of the PSR that adjusts the rejection threshold to account for the multiplicity of trials. It provides the probability that is complementary to the false positive rate *i.e.* the confidence level of the SR estimate.

Estimating $\mathbb{E}[K]$ the number of independent clustered trials: There are two suggested ways to estimate the number of independent trials K from a larger sample of M trials. The first is the initial solution put forward by Bailey and De Prado [6] that accounts for the average correlation between the trials and the second is López de Prado and Lewis [67] unsupervised learning approach using clustering

methods. We will place our focus on the latter as it is a more robust approach⁵.

The unsupervised approach to estimating K is similar to the HRP clustering methods in Sec.2.1.3⁶. Let ρ be the correlations matrix of all M simulated trials, where $\rho_{i,j}$ is the correlation of the returns observed between strategies i and j . The intention is to identify clusters of trials with high intra-cluster correlations but low inter-cluster correlations. Therefore we make use of a distance matrix D to identify trials that are highly correlated being those with smaller distances:

$$D_{i,j} = \sqrt{\frac{1}{2}(1 - \rho_{ij})}, \quad (2.27)$$

for $i, j = 1, \dots, M$ trials. To assess the distances of the trials not only between one another but rather with information about the global system the Euclidean distance matrix can be defined as:

$$\tilde{D}_{i,j} = \sqrt{\sum_k (D_{ik} - D_{jk})^2}. \quad (2.28)$$

By using $\tilde{D}$, which is a global “distance of distances” measure, the metric becomes more robust and noise is reduced in the clustering process. The clustering algorithm applied is the K-means algorithm using $\tilde{D}$ as the input, however, it requires a user defined K , therefore we modify the clustering process to make it iterative, trying different values of K . To do so we use the silhouette scores S_i as an objective function to obtain the optimal K (number of clusters) [112]:

$$S_i = \frac{b_i - a_i}{\max\{a_i, b_i\}}, \quad (2.29)$$

where a_i is the average distance between trial i and all other points within the same cluster, and b_i is the shortest average distance between i and all points in the next nearest cluster. The silhouette score ranges from -1 to 1, with higher values indicating better-defined clusters. In order to quantify the quality of the clusters a measure of quality q is given to each cluster:

$$q = \frac{\mathbb{E}[\{S_i\}]}{\sqrt{\text{Var}[\{S_i\}]}} \quad (2.30)$$

where $\mathbb{E}[\{S_i\}]$ is the mean silhouette score and $\text{Var}[\{S_i\}]$ is the variance of the silhouette scores. López de Prado and Lewis [67] uses this quality metric in the “optimal number of clusters (ONC)” algorithm to iteratively assess the clusters generated. If the quality of a cluster is not satisfactory, the clustering process is repeated with different parameters until an optimal clustering is achieved thereby settling on the optimal K . When implementing the K-mean clustering a user will need to specify the initial maximum number of clusters that the algorithm can select as well as the number of iterations that the algorithm goes through in order to optimise the final value of K ⁷.

⁵For further detail see Bailey and De Prado [6] Apx. A.3. ESTIMATING THE NUMBER OF INDEPENDENT TRIALS.

⁶Here a correlation matrix of M strategies is created as apposed to the correlation matrix of assets J seen previously in Sec.2.1.3

⁷See A6.1 Multiple simulation test cases for details on implementing the algorithms in practice.

Estimation of the expected variance of the Sharpe Ratio estimate: To estimate the expected variance of the SR estimate, we use minimum variance allocation for assigning weightings to the return time series for aggregation in the time series:

$$w_k = \frac{\sum_k^{-1} 1}{1' \sum_k^{-1} 1}, \quad (2.31)$$

where w_k are the weights assigned to each return time series within a cluster. The aggregated time series of returns for each cluster is then calculated as:

$$S_{k,t} = \sum_{i \in C_k} w_{k,i} r_{i,t}, \quad (2.32)$$

where $S_{k,t}$ is the aggregated return for cluster k at time t , and $r_{i,t}$ are the returns of the individual strategies within the cluster.

Next, we adjust the frequency of trade to make the cluster SRs comparable:

$$\text{Years}_k = \frac{\text{Last Date}_k - \text{First Date}_k}{365.25 \text{ days}}, \quad (2.33)$$

where Years_k is the duration of the time series for cluster k . The frequency of observations for cluster k is then calculated as:

$$\text{Frequency}_k = \frac{T_k}{\text{Years}_k}, \quad (2.34)$$

where T_k is the total number of observations for cluster k . The annualised Sharpe Ratio for each cluster is then given by:

$$\widehat{\text{aSR}}_k = \frac{\mathbb{E}[\{S_{k,t}\}] \text{Frequency}_k}{\sqrt{\text{Var}[\{S_{k,t}\}] \text{Frequency}_k}} = \widehat{\text{SR}}_k \sqrt{\text{Frequency}_k}. \quad (2.35)$$

As the estimated SRs are now comparable since they are all annualised, we can estimate the variance of the clustered trials⁸:

$$\mathbb{E}[\text{Var}[\{\widehat{\text{SR}}_k\}]] = \frac{\text{Var}[\{\widehat{\text{aSR}}_k\}]}{\text{Frequency}_{k^*}}, \quad (2.36)$$

where Frequency_{k^*} is the frequency of the cluster with the highest SR. This provides an estimate of the expected variance of the SR across the clustered trials, which is crucial for calculating the DSR.

The DSR, with its theoretical foundation in addressing non-linear and non-normal return distributions, directly responds to the practical challenges faced by traditional metrics, making it especially compatible with the adaptive nature of HRP and the flexibility of HS-FP, which both aim to handle real-world market complexities more effectively.

⁸From a theoretical perspective, the variance of the SR across K clustered trials can be viewed as a random variable, as it depends on the particular clustering of M simulated trials. The expectation operator $\mathbb{E}[\cdot]$ in Eq. 2.36 represents the expected variance across repeated hypothetical clusterings. In practice, we approximate this expectation by calculating the variance of the observed annualised cluster SRs and adjusting for the frequency of observations in the cluster with the highest SR.

2.2.2 Further metrics

When time permits, it is beneficial to consider additional statistical and financial metrics to gain a more comprehensive understanding of a strategy's performance. Further statistical metrics include Bailey and Lopez de Prado [7]'s Minimum Back-test Track Record Length (MinBTL), which helps ensure the robustness of the back-test results, and the Probability of Back-test Overfitting (PBO) by Bailey, Borwein, Lopez de Prado, and Zhu [9], which assesses the likelihood that the strategy's performance is due to overfitting rather than genuine skill. Additional financial metrics would be the Maximum Drawdown (MDD) which provides insight into the potential risk and worst-case scenario of the strategy, and the portfolio Turnover that measures the frequency of trading, which can impact transaction costs and overall strategy efficiency.

In summary, the advanced performance measures discussed, such as the PSR, DSR, and the PBO, provide crucial enhancements over the classic SR by accounting for the complexities of non-normal and non-linear return distributions that are often overlooked. These metrics are not just theoretical improvements but practical tools that address real-world portfolio challenges, particularly under dynamic and stressed market conditions. This provides the basis for the upcoming simulation design, where these advanced metrics will be used to evaluate the performance of the MV, HRP, and the HS-FP models across a range of market conditions.

2.3 Simulation design

With the base theory in place for the MV and the HRP algorithms, and supporting advanced performance metrics, we can construct a simulated back-test to assess relative performance. We select assets available to SA investors and representative of each major asset class. Additional benchmarks are included to provide a fair comparison of the portfolio strategies against common portfolio types such as the $\frac{1}{N}$ portfolio.

2.3.1 Asset class data selection

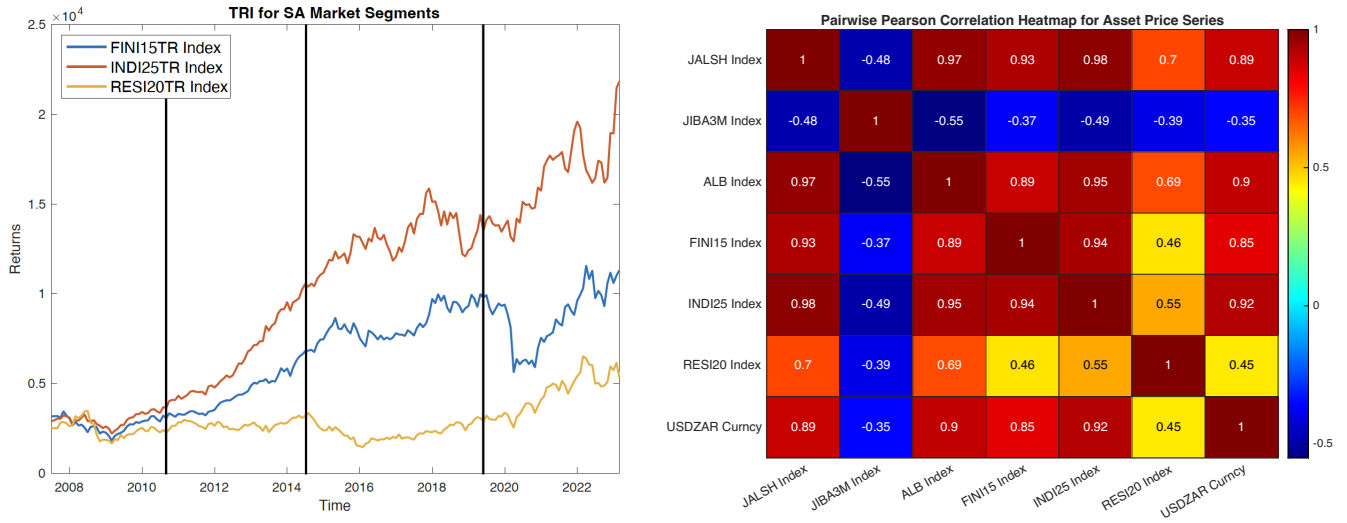
This section outlines the asset data used in the subsequent simulation sections. It discusses the selection of asset classes along with a description of the preprocessing steps performed prior to simulation. For the purpose of constructing a well-rounded onshore sample portfolio, asset data representative of the three major classes are considered: equity, bonds and cash.

2.3.1.1 Equity

South Africa's equity market has been shown to be unique in that if a principle component analysis is performed, two dominant principle components emerge: the JSE Financial-Industrial market and the Resource market. This observation was supported by the work of Wilcox and Gebbie [126] as well as van Rensburg [124], noting that these two sectors, JSE Financial-Industrial market and the Resource market, create a clear dichotomy in the SA equity space. Illustratively, Fig.2.2a visualises periods of market movement that show separate rallies observed across the JSE Financial-Industrial segment and the Resource segment using the FINDI15TR Index, INDI25TR Index and the RESI20TR Index as proxies for each respectively⁹.

⁹See Tab.2.1 for the asset class and proxy definitions.

The FINI15TR and INDI25TR display positive correlation with one another while the RESI20TR exhibits negative correlation with the others.



(a) Periods of market movement show separate rallies over the period June'07 to Feb'23 across the JSE Financial-Industrial segment and the Resource segment using the FINI15TR Index, INDI25TR Index and the RESI20TR Index as proxies for each respectively. Generated by: `data_cleaning_1b.mlx` (b) Correlation heatmap of South African asset class prices over the time period of June'07 to Feb'23. The JALSH, FINI15, INDI15 AND RESI20 are equity indices and the JIBA3M is representative of a cash asset while the ALBI represents bonds¹⁰. Generated by: `tri_invariant_2.mlx`

Figure 2.2: Fig.2.2a illustrates distinct rally periods for the JSE Financial-Industrial segment and the Resource segment, reinforcing the notion that the SA equity market exhibits a dichotomous nature when subjected to principle component analysis [124, 126]. Correlations between SA asset class prices in Fig.2.2b supportingly reveal two distinct correlation patterns 1.) The RESI20 Index stands out for having a lower level of correlation within the equity assets, reinforcing the notion that it occupies a distinct branch of the equity dichotomy. 2.) The JIBA3M—later used to create a proxy for cash—exhibits a prominent negative correlation with all other assets. This inverse relationship is crucial as cash assets typically function as a haven during periods of overall market volatility.

Correlations between SA asset class prices in Fig.2.2b supportingly reveal two distinct correlation patterns 1.) The RESI20 Index stands out for having the lowest correlation with other equity assets amongst all other equity assets, reinforcing the notion that it occupies a distinct branch of the equity dichotomy. 2.) The 3 month Johannesburg Interbank Average Rate (JIBA3M)—later used to create a proxy for cash—exhibits a prominent negative correlation with all other assets. This inverse relationship is important as cash assets typically function as a haven during periods of overall market volatility.

Given the correlation structures present amongst SA equity markets indices as captured by Fig.2.2, rather than relying on a general proxy for SA equity in the portfolios to be constructed, such as the FTSE/JSE All Share Index (ALSI), a discerning approach using the three sectors identified in Fig.2.2a is used to create depth in the equity class. Consequently, the three sector proxies selected representative of the equity class are the FINI15, INDI25, and RESI20.

¹⁰See Tab.2.1 for asset class and proxy definitions.

2.3.1.2 Bonds

When considering the bond asset class we need to bear in mind that all assets selected need to be usable in the HS-FP framework to come in Chp.3 where the aim is to extract invariants from the data. Therefore, special care needs to be taken when examining individual bonds. Unlike equity, returns cannot be used as invariants for bonds, particularly those with short maturities. The prices of short-maturity bonds soon converge to their face value over time, posing a challenge as their returns cease to be invariants [86]. This convergence would imply that their past history is not representative of their future behaviour. To overcome this, the JSE/FTSE All Bond Index (ALBI) is used. Since it represents a basket of bonds, the ALBI avoids the predicament of price convergence, allowing the inclusion of bonds in the portfolio¹¹.

2.3.1.3 Cash

The Short-Term Fixed-Interest Index (STEFI) and the Johannesburg Interbank Average Rate (JIBAR) are commonly used cash proxies. The STEFI used to be a popular choice for a cash proxy but important limitations need to be recognised that may influence the intended use of the proxy. To represent cash in a portfolio, the variance of the instrument should ideally be set to zero, considering that cash yields a known return. However, this is challenging with the STEFI, given its nature as a collection of money market indices. Inevitably the purchase and sale of units within the instrument introduces a degree of variance.¹²

The reported value of the JIBA3M—on the other hand—represents the yield an investor would achieve if the instrument were held from the start date to maturity at three months, which is a known value. The mid-yield of the JIBAR is derived as the mean value between the bid and offer rates provided by contributing South African banks for their Negotiable Certificate of Deposits (NCDs) with a discount term of three months [35]. This mid-yield is a representative estimate, reflecting an average of borrowing and lending rates prevalent in the South African market, making it a prominent benchmark for short-term interest rates within the South African financial sector [11], as well as a promising candidate for a cash proxy.

In the context of portfolio construction, the historical means of equity and bond yields can be utilised and projected forward, which would introduce uncertainty linked to the historical variance. Unlike equities and bonds, the historical mean of cash does not contribute to the variance of cash, given the certainty of the return. As the JIBA3M is reported as an *annualised* yield to maturity forming the basis of Negotiable Certificates of Deposit (NCDs) yields, a monthly interest rate needs to be computed to establish *monthly* returns as our cash proxy. This is dealt with in Sec.2.3.2 during asset data processing.

¹¹In the concluding chapter, the section on [Potential future work](#) discuss incorporating individual bonds into the framework and creating our own state variables.

¹²Including the STEFI as the cash asset may offer the advantage of capturing genuine correlations between the bonds and equity assets with a risk-free asset, the introduction of variance to a cash asset may not be worth the trade-off.

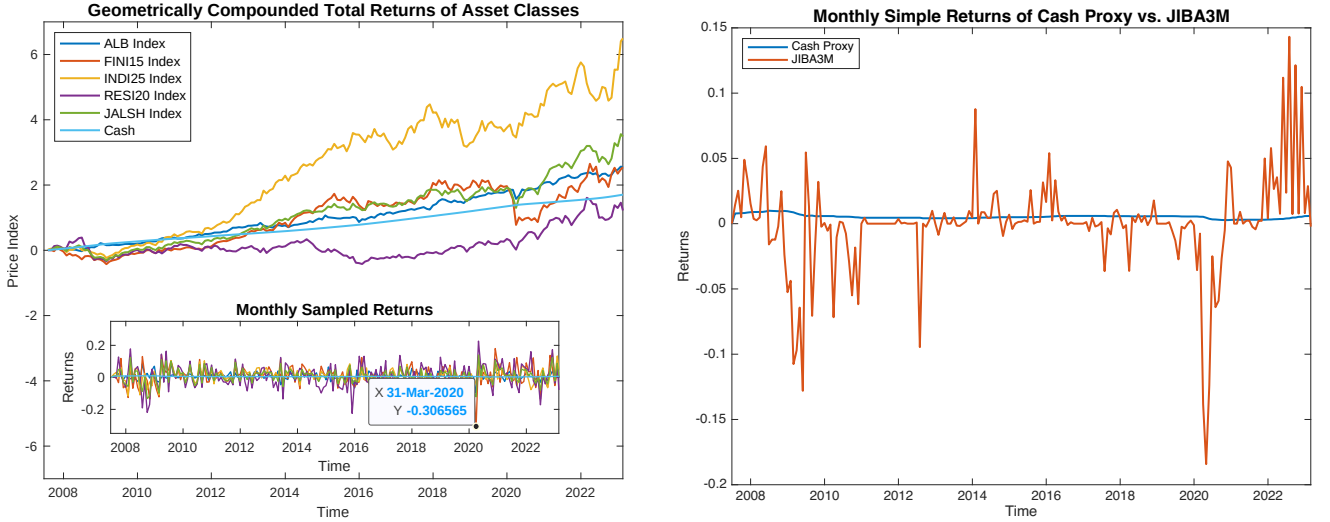
Tab.2.1 below summarises the asset classes and respective proxies along with the corresponding Bloomberg tickers and descriptions. Note that although the ALSI will not be used in the focus portfolios, it is included as a point of comparison during the analysis. For a detailed guide on sourcing the correct data refer to [A5.6 Sourcing data: a step-by-step guide to Bloomberg](#).

Table 2.1: For the purpose of constructing a well-rounded onshore sample portfolio to be used within the portfolio simulation work, asset data representative of the three major classes are considered: equity, bonds and cash. Three sectors: industrials, financials, and resources (INDI, FINDI and RESI) are used to represent the SA equity market. Bonds will be represented by the ALBI and the cash asset will be constructed by discounting the JIBA3M. The asset classes and respective proxies are listed below along with the corresponding Bloomberg tickers and descriptions. Note that although the ALSI will not be used in the focus portfolios, it is included as a point of comparison during the analysis.

Asset Class	Ticker	Description
SA Market Equity	JALSHTR Index	FTSE JSE All Share Index. Sectors below will be used to represent a 3 sector equity market.
Industrials	INDI25TR Index	25 largest JSE listed (based on market cap) companies in the industrials sector.
Financials	FINI15TR Index	15 largest JSE listed (based on market cap) companies in the financials sector.
Resources	RESI20TR Index	20 largest JSE listed (based on market cap) companies in resources sector.
Bonds	ALBITR Index	FTSE JSE All Bond Index
Cash	JIBA3M Index	Johannesburg Interbank Average Rate (JIBAR): 3 month money market rate.

2.3.2 Data processing & exploration

The initial monthly price (PX_LAST) data extracted from Bloomberg is comprised of a 23-year period spanning 29/02/2000 to 28/02/2023, however, this needed to be truncated to ~ 16 year period from 30/06/2007 to 28/02/2023 as not all tickers had data covering the full initial window. This is discussed in more detail in [Apx.4](#). For the assets listed in [Tab.2.1](#) we computed simple monthly returns as defined in [Def.8](#) for each asset class, excluding the cash asset. Notably, extreme outliers were not removed, as we observed that certain shocks were exogenous to the system and represented genuine market disturbances, distinct from shocks resulting from system artefacts, such as corporate events. Retaining information on significant market events is essential, as they represent extreme market states that hold significance when utilising the HS-FP framework. Examples of legitimate outliers are shown in [Fig.2.3a](#) below.



(a) Returns are centred around zero, with a longer left tail driven by major shocks, such as the COVID-19 pandemic in Q1 2020. Retaining such events is crucial for capturing extreme market states in the HS-FP framework for regime identification. (b) A comparison of the proxy and original JIBA3M simple monthly returns. The reported JIBA3M is annualised and exhibits variability unsuitable for a risk-free asset. Monthly returns are computed using Eq.2.37 to realistically represent a cash asset with minimal variability.

Figure 2.3: The historical returns of equity and bonds seen in Fig.2.3a can appropriately be projected forward to introduce uncertainty into the future behaviour of the assets that would be linked to the historical variance. Unlike equities and bonds, in practice the historical mean of cash should not contribute to the variance of cash, given the certainty of the return as a risk-free asset. As the JIBA3M is reported as an annualised yield to maturity forming the basis of Negotiable Certificates of Deposit (NCDs) yields, a monthly interest rate is computed and shown in Fig.2.3b, removing significant uncertainty from the risk-free asset as the proxy for holding cash in the portfolio. Both sub-figures were generated by: `tri_returns_prep.3a.mlx`

In Fig.4, heightened volatility is evident during late 2008, reflecting the market turmoil caused by the GFC. Similarly, Q1 of 2020 is characterised by extreme outliers, marking the onset of the COVID-19 pandemic. Sec.2.3.1.3 emphasises that the JIBA3M is reported as an annualised rate. In order to obtain a simple return for cash similar to that of the equity assets we calculate the monthly forward rate of interest earned. The quoted JIBA3M changes daily therefore a new monthly rate is calculated at each time point (t). We can define the monthly compound interest rate formula as:

Definition 27 (Monthly compound interest rate). When an interest rate is reported on an annual bases r_{annual} , the monthly rate r_{monthly} can be obtained using:

$$r_{\text{monthly}} = (1 + r_{\text{annual}})^{\frac{1}{12}} - 1. \quad (2.37)$$

r_{monthly} is the monthly interest rate in decimal form and r_{annual} is the annualised interest rate in decimal form.

We will preprocess the JIBA3M to have the realised rate one month in advance for all the simulation work. In the preprocessing, we use a Càdlàg approach. Using the reported three-month rate at time (t) and calculating the forward rate as in Eq.2.37 and lagging it to time ($t + 1$), using that as the realised rate of return one would earn over the period $t \rightarrow (t + 1)$. This produces a form of right-continuity in the sense that the rate at ($t + 1$) depends on the information available at (t). However, it should be noted that if one were to implement optimisations in real-time (*i.e.* not in back-tested simulations but on a real-time month-to-month basis) one would need to calculate the one month ahead rate at each iteration of the monthly rebalancing.

2.3.3 Benchmarks

Within the back-tests to come, a two-branched approach is employed, comparing the “Conventional” MV and HRP portfolios generated in Sec.2.4 with those generated using the FPrs from the HS-FP framework in Chp.3. To increase the understanding of the relative performance of the optimisations, benchmarks will be incorporated. To ensure comparability, assets and indices that are locally investable are used, making them viable alternatives for South African investors. Three classes of benchmarks are used for comparison: an Equally Weighted (EW) portfolio, Balanced Funds (BF), and Market Indices, as listed in Tab.2.2 below, along with details on how each benchmark was defined during the construction of the back-tests.

Table 2.2: The six benchmarks used in the rolling window dynamic back-tests are given below. The assets that are available to each benchmark are recorded under “Investment Universe”. The rebalancing strategies that are applied to each actively managed portfolio on a Monthly (M) basis, this would either be: i.) a Constant Mix (CM) strategy, or ii.) a Buy-and-Hold (BH) strategy (which employs passive management).

#	Benchmark Type	Investment Universe ¹³	Rebalancing Strategy	Rebalancing Freq.
Equally Weighted				
1	Equally Weighted (EW)	FINI, INDI, RESI, ALBI, Cash	CM	M.
Balance Funds				
2	Balance Fund Type 1 (BF-CM)	ALSI, ALBI, Cash	60 : 35 : 5 CM	M.
3	Balance Fund Type 2 (BF-BH)	ALSI, ALBI, Cash	60 : 35 : 5 BH	-
Market Indices				
4	SA Equity Proxy	ALSI	-	-
5	SA Bonds Proxy	ALBI	-	-
6	SA Cash Proxy	JIBA3M forward rate	-	-

Equally weighted portfolio and balanced funds: Benchmark 1 is a simple yet effective uniform Constant Mix (CM) portfolio, conventionally referred to as an EW or a $\frac{1}{N}$ portfolio ($\frac{1}{J}$ in this case). Each month, the weight allocation is adjusted to maintain a constant mix, ensuring that each of the J investable assets receives an equal weight of $\frac{1}{J}$ throughout the simulation. Benchmarks 2 and 3 are constructed around a BF strategy, with 60% of the portfolio invested in equities, 35% in bonds, and 5% in cash¹⁴. The distinction between the two lies in their approach. Benchmark 2 actively rebalances to maintain the CM on a monthly basis, while Benchmark 3 adopts a passive Buy-and-Hold (BH) strategy. The ALSI is used as equity data and the ALBI as data on the bond market. Like the equally weighted CM, the balanced-portfolio rebalances monthly, this time to a 60:40 mix, ensuring that the composition is adjusted monthly back to 60% equity (ALSI) and 40% bonds (ALBI).

¹³See Sec.2.3.1 for full description of each asset selected.

¹⁴See [Market indices](#): for detail of asset class proxies.

Market indices: By comparing the MV and HRP strategy returns to those of benchmark indices, we can better assess whether the portfolio adds value beyond what could have been achieved through passive investment in a market index. One can then isolate the strategy's alpha, or excess return, and better evaluate its risk-adjusted performance. By using market indices as benchmarks it allows for a more objective and standardised evaluation. To this end, the following market indices were selected as benchmark proxies specific to the SA market:

1. SA Equity: FTSE/JSE All Share Index (ALSI)[62],
2. SA Bonds: FTSE/JSE All Bond Index (ALBI)[63],
3. SA Cash: adjusted¹⁵ 3 month Johannesburg Interbank Average Rate (JIBA3M)¹⁶[11].

2.3.4 Dynamic back-test design

A dynamic back-test in this context is defined as a framework that consists of IS data windows that are adjusted over time, incrementing across the OOS data, simulating how an investment strategy would have performed under defined historical scenarios. There are alternative methods like a growing window approach, an exponentially weighted window or in the case of this research, a rolling/overlapping window approach. In an overlapping window back-test the IS window of set length rolls incrementally month by month across the OOS data. At each increment, time (t), each portfolio strategy is implemented using the data that falls within the IS window. The optimal IS portfolios are then identified and the associated optimal weights are in turn applied in the subsequent month OOS, to ascertain OOS performances. A time series of strategy performance for each portfolio optimisation method is generated to create a simulated historical performance path.

For the MV strategy, this involves computing the expected returns over the IS period as well as the associated covariance matrix. For the HRP algorithm, only the covariance matrix is required to find the optimal portfolio weights. The weights are then applied to the returns over the month at time ($t+1$) OOS. As the window moves forward by a month, at time ($t+1$), the oldest month (historically) that formed part of the IS window at time (t) is dropped, maintaining the window length at each iteration. Over each month the realised returns (over that month) per portfolio are stored to be used after the simulation to plot an equity curve that represents the portfolio's OOS performance.

¹⁵The JIBAR rate needs to be adjusted and discounted to correctly reflect the cash asset compounding in the simulations as like-for-like relative to the equity and bond assets. This adjustment of the cash asset is detailed in Sec.2.3.2

¹⁶The JIBAR is being phased out and replaced by Zaronia, the new SA overnight index average rate by 2026. The benefit of using Zaronia over the JIBAR is that it is based on historical transactions and calculated in arrears rather than the forward-looking rate of the JIBAR, this removed the need to discount the rate [51].

A 36 month fixed IS window length is set to ensure that there exists sufficient data for portfolio strategies to adapt within a business cycle over the period following the initial IS window, but not an excessive amount of IS data such that the models overfit OOS. The OOS period of this dataset captures the after effect of the GFC and more significantly the impact of the COVID-19 pandemic, as well as several economic and political events such as the tapering of quantitative easing in 2014. With the inclusion of such events an understanding can be gained as to how each portfolio performs relative to one another during periods of market flux. The flow of data from the source to the end of the back-testing framework is represented schematically in Fig.6 in [A5.4 Toolbox construction](#)¹⁷.

2.3.4.1 Capturing the effect of time

We need to be aware of how we handle the return and performance metric calculations within the back-test to account for the returns accumulating over the time of investment. Before performing the back-test, arithmetic month-to-month data *i.e.* simple returns (as defined in Def.8) are calculated from the price data. Geometric processes are applied in two instances, within the back-test the geometric mean (as defined in Def.12) is passed as input to the optimisations after which the geometric mean is used again to compound the returns for TRI series when evaluating historical performance.

2.4 Preliminary results

In this section we run the initial first pass of the back-test framework to obtain preliminary results. After which we utilise the DSR and multiple simulations to evaluate the reliability of data processes such as: winsorising returns [44], the implementation of cash constraints, and the regularisation of covariance matrices to obtain more dependable results. The purpose is to obtain a well balanced conventional back-test to fairly compare the results of the MV and the HRP optimisations and to have a strong enough model to take into further analysis in Chp.3.

Running the initial 36 month rolling window back-test for the MV and HRP models with the benchmark portfolios we use the assets in Tab.2.1 to generate the simulation summary figures in Fig.2.4 on the next page.

¹⁷The diagram shows the progression through the full project workflow including the signal processing and application of the HS-FP methods applied in Chp.3. To apply only the conventional approach to back-testing, exclusively follow the rightward (downward) arrows out of each node.

In the cumulative relative performance plot above (top left) we see the most performant profile is the ALSI benchmark market equity index with the second-best performance produced by both BF benchmark portfolio strategies outperforming the EW and all other benchmarks. Studying the optimal weight surfaces of each strategy, we observe expected results for the EW portfolio (top middle), as well as from the BF-BH (bottom left) and the BF-CM (bottom right). However, the MV (top right) and HRP (bottom middle) portfolios have failed in this instance, as both have succumbed to the cash asset dominating the optimal asset allocation. The HRP exhibits near zero variability but the MV shows extreme swings in decision changes in and out of the cash asset. In this case the optimisations were unconstrained with regard to the max and min optimal ω per asset, therefore we can speculate that there are corner solutions present in the MV optimisation. The resulting optimal weight allocation for the cash asset of the HRP produces an $\omega \sim 1$ throughout the simulation, a clearly undesirable none-diversified portfolio.

These initial results obtained do not allow for a fair comparison of the MV and the HRP or reliable results to use in Chp.3. Therefore we need to implement interventions to assess the cause of the failures. We examine both the inputs to the back-test as well as the underlying data and apply various data processes in hopes of obtaining more dependable results.

2.4.1 Assessing inputs

We ask three questions to assess the reliability of the optimisation inputs:

1. Were the cash forward rates calculated correctly?
2. Were all asset class returns calculated correctly?
3. Were the volatilities calculated correctly?

To eliminate possibility 1.), we perform a sanity check by observing the annual compounded rate of cash over the simulated period (approximately 12 years), which is calculated to be around 6%. Comparing this with the money market and easy access account options available at major banks such as Nedbank, Discovery Bank, and ABSA, we find potential returns ranging from 5% to 7% [65] [10] [61]. This suggests that the cash forward rate was correctly calculated as set out in Sec.2.3.2. To address concern 2.), we use the theory established in Sec.1.2 to ensure that we compute the annualised returns with the most appropriate method, which in this case is the geometrically calculated average. Similar to our approach with cash, we apply a sanity check to the realised annualised returns of the portfolios and benchmarks. Lastly, for point 3.) we observe the risk/returns in Fig.2.5, the volatilities are within a reasonable range, with cash having an annualised volatility close to 0% and the equity assets ranging between 7% and 14%.

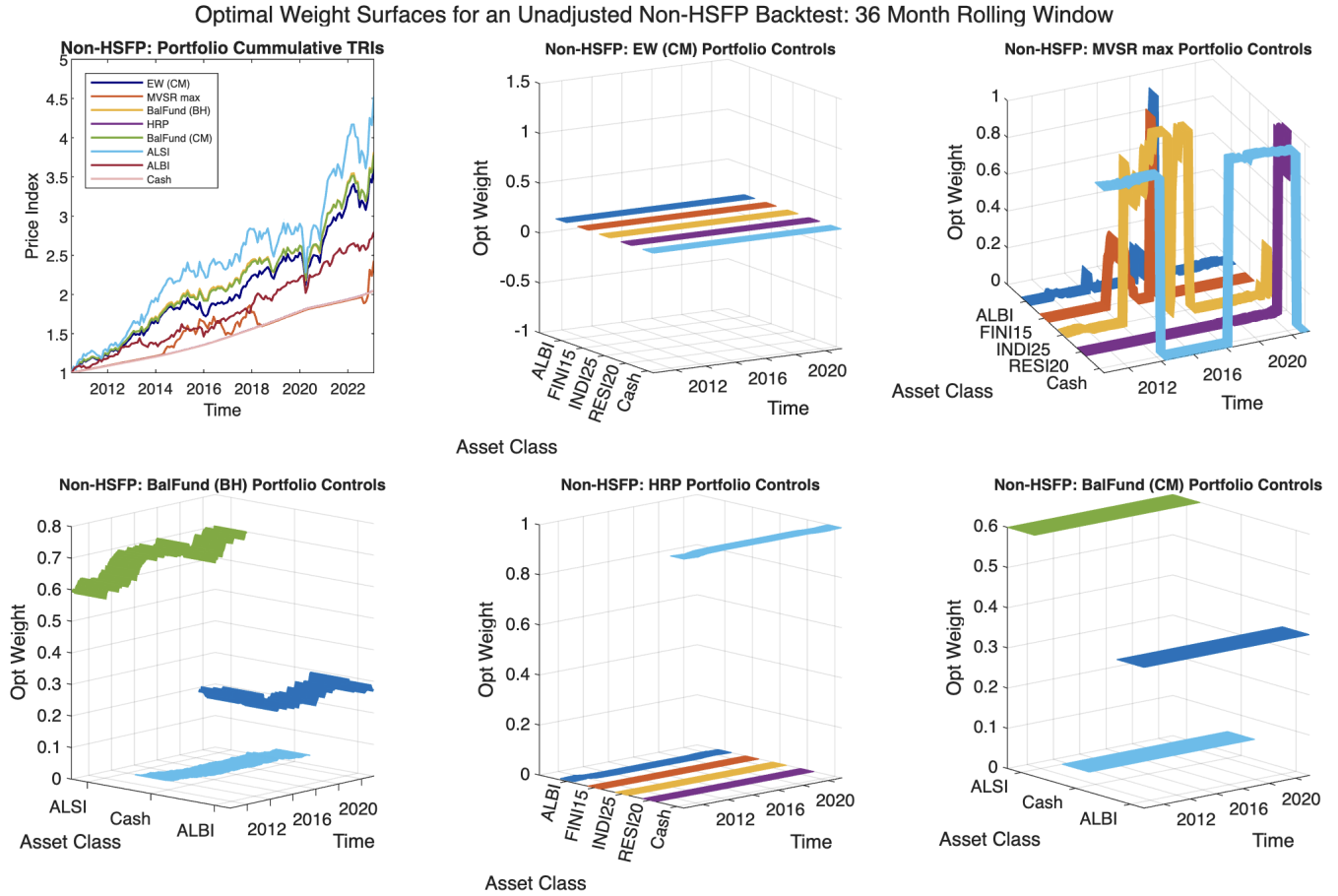


Figure 2.4: A 36 month rolling window back-test is run for the MV and HRP models and the benchmark portfolios using the investable universe as listed in Tab.2.1. In the cumulative relative performance plot (top left) the most performant profile is realised from the ALSI benchmark market equity index with the second best performance produced by both balance fund benchmark portfolio strategies, outperforming the EW and all other benchmarks. The MV and HRP portfolios seem to have failed in this instance, both succumbing to the cash asset dominating the optimal asset allocation. Studying the optimal weight surfaces of each strategy we observe expected results for the EW portfolio (top middle), the BF-BH (bottom left) and the BF-CM (bottom right) However, cash shows undesirable behavior in both the MV (top right) and the HRP (bottom middle) strategies. In this case the optimisations were unconstrained with regard to the max and min optimal ω per asset, therefore for MV we can speculate that there are corner solutions present in the optimisation potentially due to constraints dominating. The resulting optimal weight allocation for the cash asset of the HRP produces an $\omega \sim 1$ throughout the simulation. Generated by: `backtest_prep_2.mlx`.

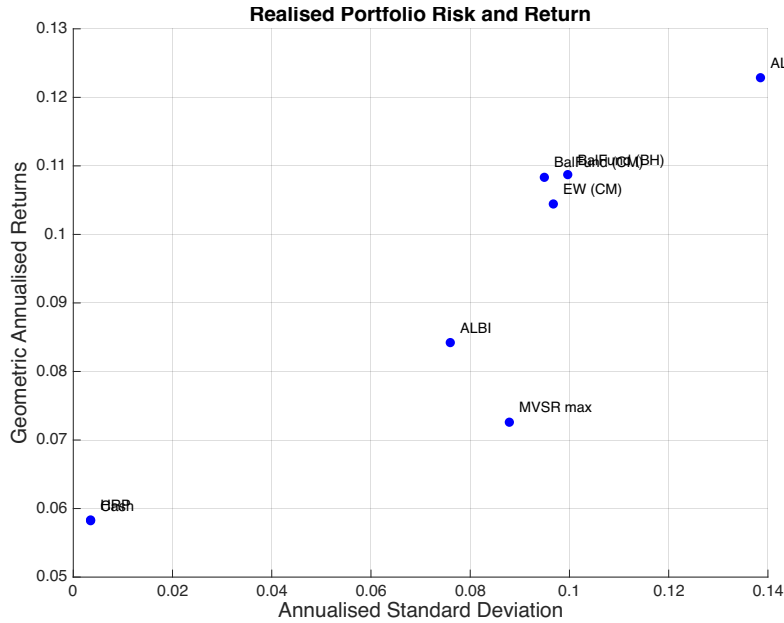


Figure 2.5: Realised annualised risk and returns generated by each portfolio (and market benchmarks) when using a traditional back-testing approach *i.e.* without incorporating the HS-FP framework. The back-test was run with monthly sampled returns with a rolling window of 36 months. Generated by: `signal_selection_backtests.mlx`.

Since we have confirmed that the inputs were correctly computed, we now turn our attention to examining the structure of the back-test and the optimisations themselves.

2.4.2 Assessing structural concerns

We refer back to Fig.2.4 which visualises the portfolio controls moving over the OSS data. It is suggestive of possible corner solutions as we see a very high turnover¹⁸. As the back-test progresses we observe spikes due to abrupt changes at each step as data falls in and falls out of the IS windows, exposing the optimisations to outliers. There are three primary solutions to structural causes:

1. Should a global turnover constraint be applied?
2. Should an inequality/equality constraint be placed on cash for MV and HRP?
3. Should the underlying returns be Winsorised or the covariance matrix regularised?

¹⁸Apart from the over allocations to cash we would also be concerned about stability and the high transaction cost that result from the high turnover seen in the MV portfolio weights.

2.4.2.1 Turnover constraint

For option 1.) the optimal decision rapidly flipping in the MV could be avoided by implementing a turnover constraint. The portfolio turnover constraint would take the form $h(\omega_{t+1}, \omega_t) \leq \bar{h}$, ensuring that the change in asset weight from the time point t in the current period to $t + 1$ in the next period is bound by $\bar{h}$. Constraining the change mitigates the significant decision changes that occur due to the optimisation being overly sensitive to small deviations in problematic data (potentially due to a near singular covariance matrix). However, this may not solve the issue of corner solutions for both MV and HRP. It would only avoid big changes during the continuous re-optimisation of the portfolios and would not necessarily address the extreme risk aversion of the HRP [101, 128].

2.4.2.2 Cash constraint

An inequality or an equality constraint can be implemented for the cash ω in both the HRP and MV methods to control for over allocation. A practical approach is needed to determine the level of the cash constraint to impose. Limiting the range could mean specifying the following fixed or varying ranges: 0%, 2-5%, or 5% only. In practice, running a cash account with a treasury management function allows for strategic management of the cash component, cash flows, and liquidity. Therefore, enforcing a constraint of 0% is not recommended, as including the cash component creates a more indicative measure of how the portfolio would truly stack up in reality. For MV it might make sense to use 2%-5%, however, due to the structure of the HRP optimisation, implementing a self-regulated inequality cash constraint is challenging and would interfere with the efficiency of the optimisation¹⁹.

When testing the various levels of cash constraint, we observed that if MV is given the self-regular inequality constraint of $0.00 < \omega_{cash} < 0.05$, it often swings to 0%, which is not conducive to a true investable portfolio due to the need for cash for transnational purposes. Consequently, a flat 5% constraint is a reasonable level and will be used for comparative purposes²⁰. In practice, for the MV portfolio, an equality constraint of $\omega_{Cash} = 0.05$ is specified for the cash asset, leaving the allocation of the remaining assets to be freely optimised. For HRP, a hold out method is used, using HRP to optimise the equity assets and then subsequently adding 5% cash back on to it, down-weight the others accordingly. This is necessary as the HRP algorithm does not easily accommodate the use of constraints.

2.4.2.3 Underlying data concerns

If the underlying data is causing the asset allocation issues, we need to determine whether the problem lies in the return series distributions exhibiting outliers or in the resulting covariance matrix. To identify the source, we implement a two-pronged approach for process of elimination: 1.) Winsorise the returns to see if managing the outliers results in decreased turnover [44][97], or 2.) Regularise the covariance matrix to flatten it, as correlations within it could cause corner solutions. Regularisation helps control correlations. We know that if the HRP does not improve using either Winsorisation or regularisation exclusively, a combined Winsorisation and regularisation approach may be needed as the issue might

¹⁹A potential solution to this could be to implement Pfitzinger and Katzke [107]'s convex hierarchical optimisation framework.

²⁰We acknowledge that 5% hard cash constraint may dilute the portfolios, however, in this particular case it is the most practical solution.

actually lie with variances. If shocks are still affecting the return series, they will impact the variances, making Winsorisation essential²¹.

Winsorising the returns: Winsorising is a transformation technique that forces data to converge towards a more normal distribution bringing in extreme outliers to Gaussianise the distribution. Named after Charles P. Winsor, it was designed to mitigate the influence of outliers by limiting extreme data points to a specified range [111]. This is achieved by setting extreme data points to a specified percentile value (as in Eq.2.38 we will use two standard deviations), effectively capping the data at these bounds. It will be a recursive process, calculating the upper and lower bounds and replacing values that fall outside these bounds with the bound values itself. This process is repeated until the data converges more closely to the normal distribution. When convergence for one asset is reached, the algorithm repeats the same steps for the next asset until all assets have converged. To calculate the lower and upper bounds for each asset, the following equation was applied:

$$\text{Upper/Lower Bounds} = \hat{\mu} \pm 2 \times \hat{\sigma} \quad (2.38)$$

In Fig.2.6a and Fig.2.6b to follow on the next page, we have the pre-winsorised and the post 2 sd. winsorised returns respectively. When applying winsorisation using 2 standard deviations we note the upper and lower extremes of the returns are pulled towards the center of the distribution, creating heavier tails and a more Gaussianised distribution.

²¹Note that for the sequence of preprocessing steps we will first Winsorise the returns and then apply regularisation.

Regularising the covariance matrix: As we have shown in Sec.2.4.2.3 with the use of Winsorisation, the distribution of the returns does not seem to be the primary force dictating the simulation performance, we now examine the covariance structure of the data. As discussed in Sec.2.1.1: “*MV is a quadratic optimisation problem built around stringent linear and quadratic constraints, ... A positive-definite and consequently invertible covariance matrix Σ is therefore vital to the framework.*” This is potentially why the covariance component of the optimisation inputs may be the cause of the unstable MV back-test solution. To assess the state of the covariance matrix we use the condition number of the matrices as a measure of noise present in the underlying data.

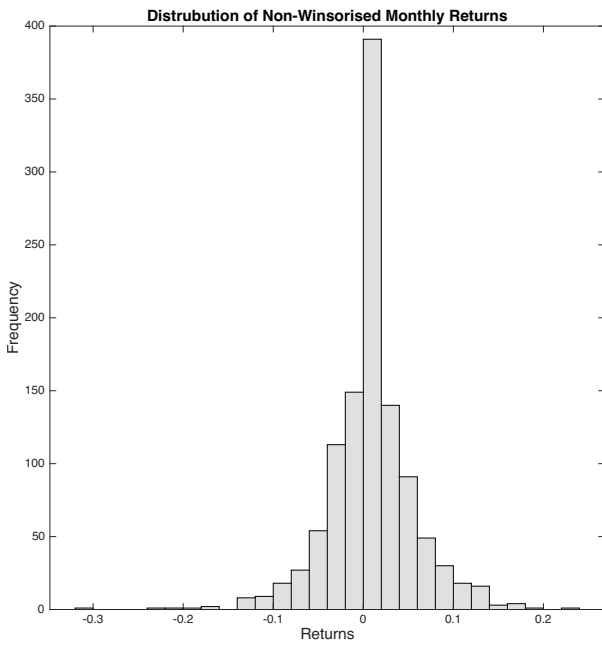
A popular way to reduce noise is to regularise the covariance matrix. The approach that we implement is an adaptation of the Ridge regression estimator, first introduced to address issues of multicollinearity in regression settings. The concept was initially proposed by Hoerl and Kennard [47] in the 1970s in the context of solving ill-posed problems and popularised for its application in linear regression models. In the setting of regression, Ridge regularisation mitigates multicollinearity by adding a penalty to the regression coefficients, to improve the invertibility of singular matrices. To regularise the covariance matrix of asset returns, we adapt the Ridge regression approach. The sample covariance matrix Σ of asset returns can be ill-conditioned or singular in high-dimensional settings. We can generate a regularised covariance matrix Σ_{ridge} by adding a small positive constant λ ²² to the diagonal elements of Σ :

$$\Sigma_{\text{ridge}} = \Sigma + \lambda \mathbf{I} \quad (2.39)$$

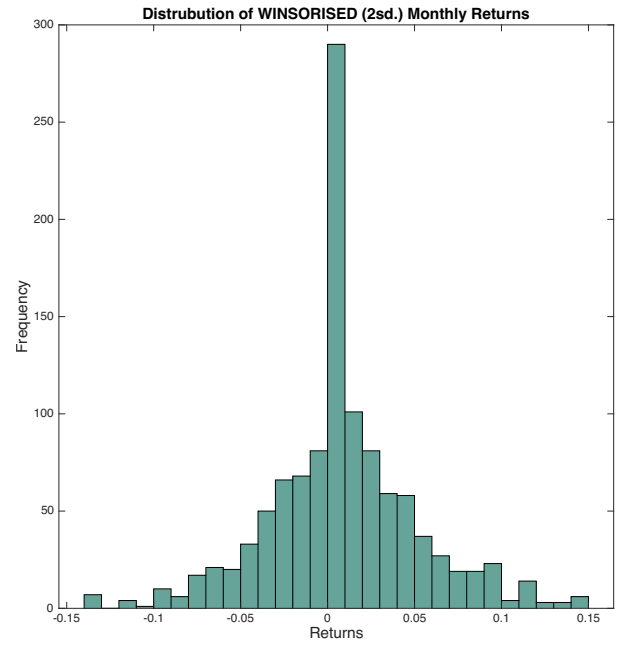
If this process of regularisation is implemented it should improve the numerical stability of the covariance matrix, making it more reliable during the simulations. In Fig.2.6c using the condition number of covariance matrices at each iteration of the back-test we get a sense of how ill-conditioned they are and if they may be leading to “error-maximisation” as described by Michaud and Michaud [90].

In both the non-regularised and the regularised simulation in Fig.2.6 the condition numbers of covariance matrices are calculated at each time step t during the rolling back-test simulation. As the condition number of a covariance matrix indicates the level of noise within the data we would expect high noise during market shocks such as Covid in 2020 and low otherwise. The non-regularised condition numbers in Fig.2.6c exhibit counter-intuitively low values during high market uncertainty, including the entire year of 2020 and beyond. After regularising the covariance matrix there is a decrease in the condition numbers seen in Fig.2.6d, displaying more intuitive results with peaks during market shock and cool offs in between. Consequently, this implies that there was a decrease in the noise of the data due to regularisation.

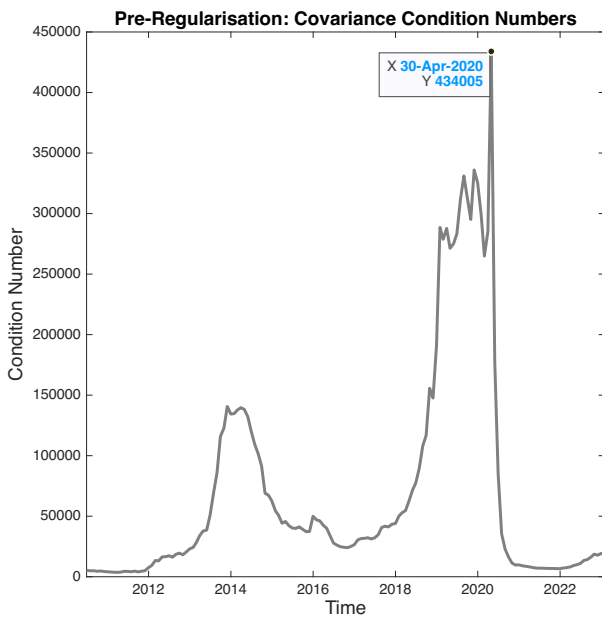
²²The selection of λ is crucial and can be done using cross-validation or other model selection techniques to balance the trade-off between bias and variance.



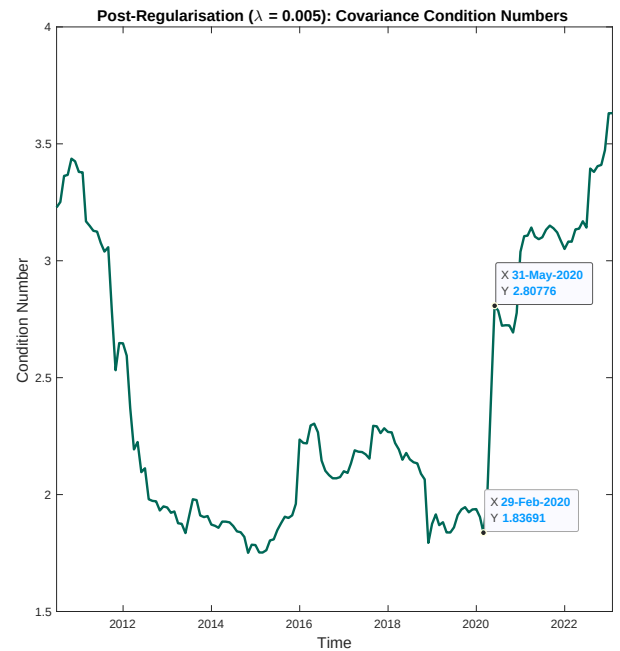
(a) Distribution of monthly return data across all asset classes. Through the distribution we note that the bulk of the distribution lies around zero as expected however, there are outliers seen on both the left and right tails.



(b) The distribution of monthly returns becomes more Gaussian as they are Winsorised [44]. This process recursively addresses outliers within 2 standard deviations. This aims to mitigate corner solutions in back-test simulations.



(c) Condition numbers of covariance matrices are calculated at each time step t during the rolling back-test simulation, and they exhibit counter-intuitively low values during high market uncertainty, including the entire year of 2020 and beyond.



(d) Condition numbers of the covariance matrices calculated at each time step t of the rolling window back-test simulation after ridge regularisation [49] is implemented to decrease the noise in the covariance matrix.

Figure 2.6: Structural data changes made to mitigate corner solutions in back-test optimisations: Winsorisation & Ridge Regularisation of the covariance matrices. After the regularisation the condition numbers reflect a more intuitive result consistent with real life market events that would introduce uncertainty such as the start of the Covid-19 pandemic in 2020. All sub-figures generated by: `backtest_prep_1.mlx`

2.4.2.4 Comparison of structural “fixes”

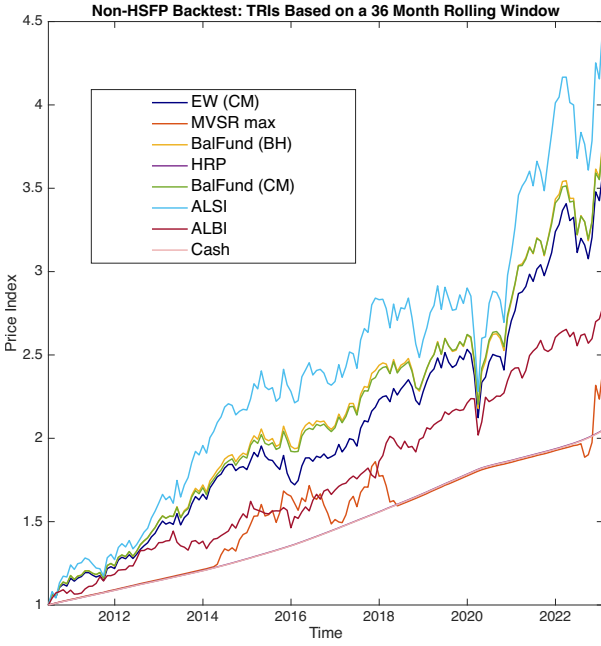
Over the next few pages, a series of model configurations are tested, and their cumulative performance curves are plotted to facilitate comparison. Two sets of comparisons are given. The first, Fig.2.7 consists of the performance curves of four distinct simulation configurations: Config.1) unadjusted simulation as seen in Fig.2.4, Config.2) applying a 5% cash constraint, Config.3) Winsorised returns to 2sd., and Config.4) regularising using $\lambda = 0.005$. The second, in Fig.2.8 is various combinations of the first four configurations in Fig.2.7: Config.5) 5% cash constraint²³ with winsorised returns, Config.6) 5% cash constraint with regularisation, Config.7) Winsorised returns and regularisation, and Config.8) 5% cash constraint with winsorised returns as well as regularisation. Tab.2.3 summaries observations of each configuration relative to the base Config.1.

Table 2.3: Performance overview of different back-test simulation configurations.

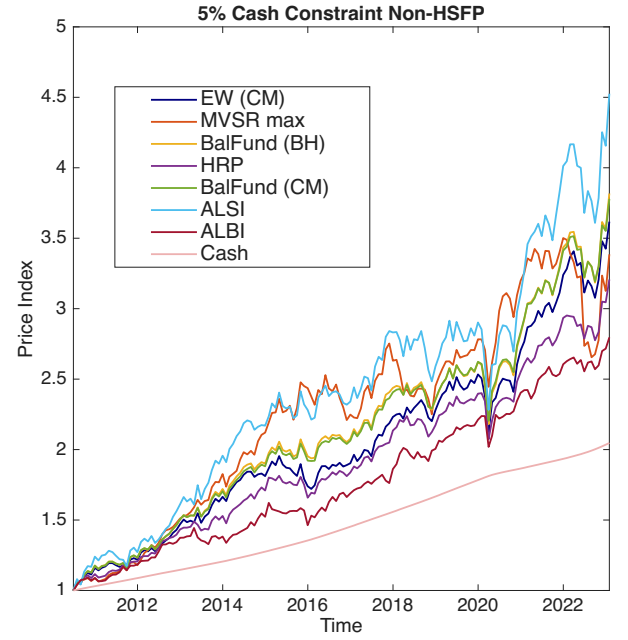
Configuration	Description	Observations
Config.1	Base configuration with no additional constraints.	Shown in Fig.2.7a,
Config.2	Equality constraint $\omega_{\text{cash}} = 0.05$ applied to both MV and HRP portfolios.	Shown in Fig.2.7b, both portfolios improve from the base case (Fig.2.7a). MV shows strong performance but dips around 2022, while HRP maintains steady performance and has the second lowest performance (excluding cash).
Config.3	Winsorised returns to 2 standard deviations applied in simulation.	As seen in Fig.2.7c, MV had slight improvement between 2014 and 2018 but experienced a negative impact afterward. The effect on HRP is negligible with no visual improvement.
Config.4	Regularisation applied to MV and HRP portfolios.	Fig.2.7d shows a positive impact, suggesting that optimal weights were diverging in the non-regularised simulation [39]. HRP, despite having advantages in noisy data and ill-conditioned covariance matrices, still faces limits (this finding is supported by previous research [76]).
Config.5	Combined 5% cash constraint with Winsorised returns.	Shown in Fig.2.8a, this configuration produces improved portfolios for both MV and HRP. The dip after 2022 experienced by the MV in the cash constraint-only simulations is rectified by removing extreme outliers using Winsorisation.
Config.6	Regularisation combined with a 5% cash constraint.	Fig.2.8b shows both MV and HRP portfolios exhibit improved performance. However, MV shows signs of performance inflation suggesting potential overfitting, while HRP is less affected.
Config.7	Regularisation combined with Winsorised returns.	Fig.2.8c shows similar results to Config.6
Config.8	Regularisation, Winsorisation, & a 5% cash constraint.	Fig.2.8d shows similar results to Config.6 and 7.

Visually, the use of a cash constraint and the regularisation suggest that the underlying structural problem may stem from noisy data and not from an excessive amount of outliers as Winsorisation made little improvement on performance. The application of regularisation, however, showed signs of overfitting, risking poor OOS performance. To quantitatively compare configurations we summarise each configuration’s statistics in Tab.2.4. In the section to follow we make use of the DSR (as defined in Sec.2.2.1.2) to account for potential selection bias and to provide a more robust measure of risk-adjusted returns with which to select the final model for this chapter of “conventional approach”.

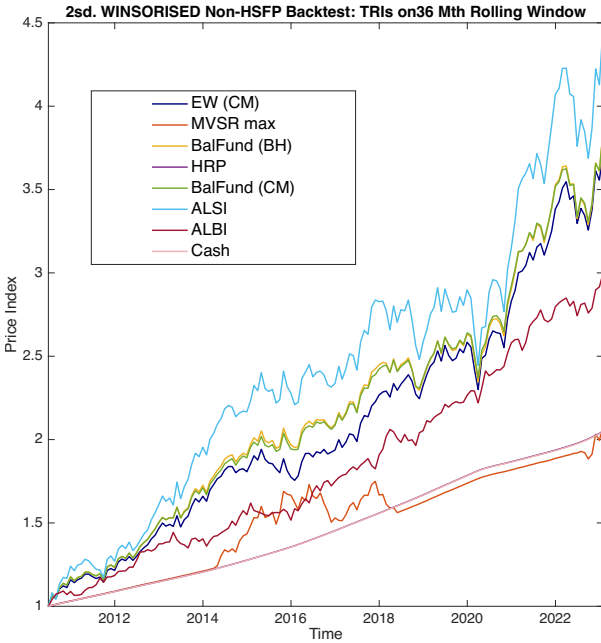
²³See Sec.2.4.2.2 for details on the implemented constraints for MV and HRP.



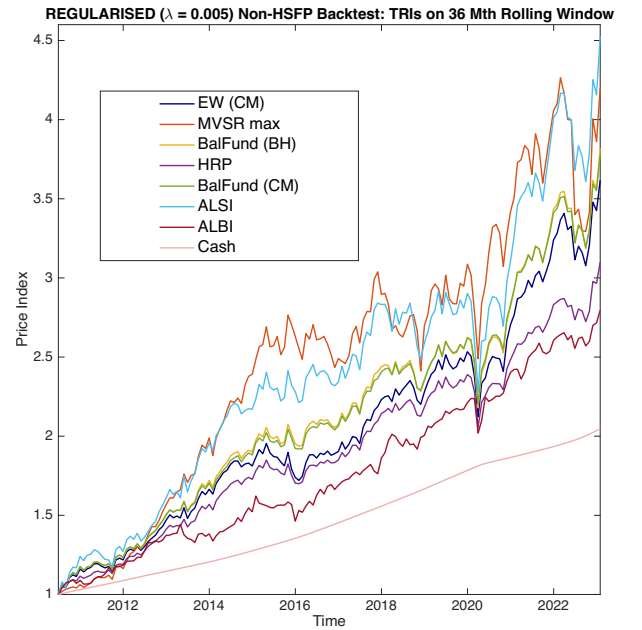
(a) Config.1 is the original 36 month rolling window back-test with no structural changes made to mitigate potential corner solutions. This uses the distribution of returns seen in Sub-Fig.2.6a.



(b) Config.2 results with a 5% cash constraint on MV and HRP portfolios show improved performance curves compared to the base case in Config.1.

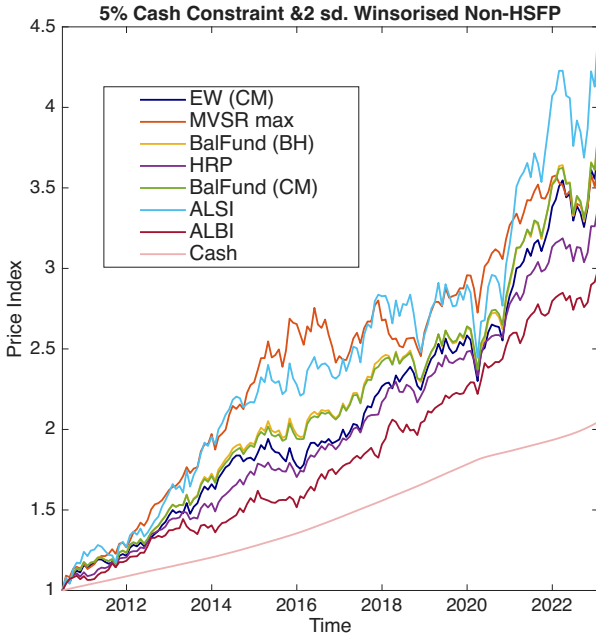


(c) Config.3 results using Winsorised returns (Sub-Fig.2.6b) show a negative impact on MV optimisations and no change for HRP compared to the base case.

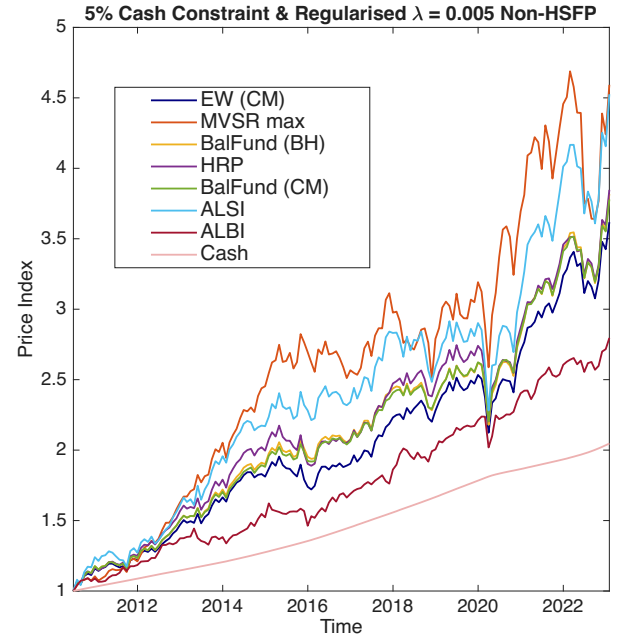


(d) Config.4 results from a regularizing penalty term $\lambda = 0.005$ show significant improvement for both MV and HRP portfolios compared to the base case.

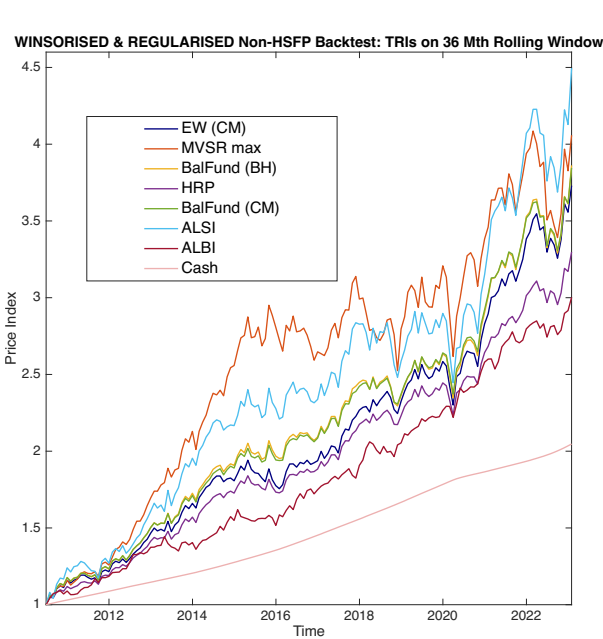
Figure 2.7: Above are the performance curves of four different simulation configurations: Config.1) unadjusted simulation as seen in Fig.2.4, Config.2) applying a 5% cash constraint²⁴, Config.3) Winsorised returns to 2sd., and Config.4) regularising using $\lambda = 0.005$. The purpose of the simulation comparison is to ascertain whether there are underlying structural concerns with the data fed to the MV and HRP optimisation methods in the unadjusted back-test where the optimisations perform poorly. For all simulations all other parameters are kept consistent with the base case (Fig.2.7a) only the parameters mentioned in the configurations are varied. A notable improvement is observed when implementing a cash constraint (Fig.2.7b) and when regularising the covariance matrices (Fig.2.7d). Summary statistics for all configurations are found in Tab.2.4. All sub-figures generated by: `dataprocessing_dsr.5.mlx`.



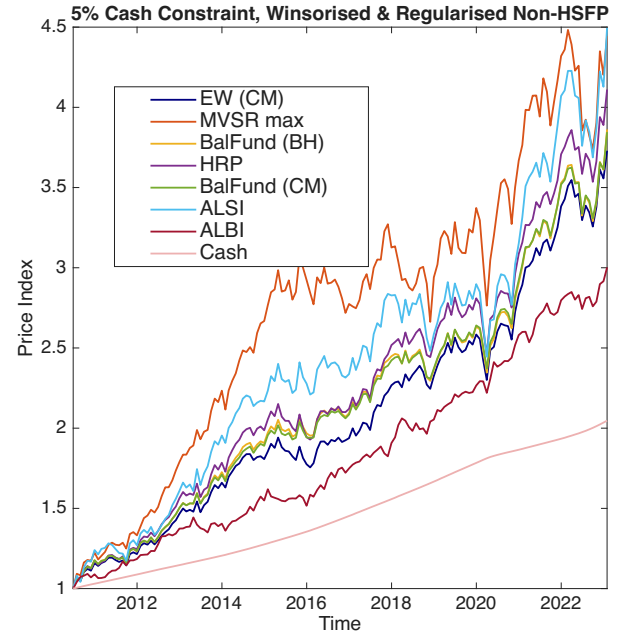
(a) Config.5 uses a 5% cash equality constraint applied to the MV and HRP portfolios, and winsorised returns in the rolling window optimisation, results in a noticeable yet not exaggerated performance improvement.



(b) Config.6 with a 5% cash constraint imposed and regularisation of the covariance matrix, the performance of the MV portfolio exhibits potential inflation, suggestive of possible overfitting.



(c) Config.7 combines the winsorisation and the regularisation approach results in improved performance from the unadjusted back-test.



(d) Config.8 uses winsorised returns and regularisation. The end performance resembles the exclusively regularised back-test. The only slight improvement is in the MV performance over the midterm 2014 - 2018.

Figure 2.8: The simulations above are the a result of various combinations of the configurations in Fig.2.7: Config.5) 5% cash constraint with winsorised returns, Config.6) 5% cash constraint with regularisation, Config.7) Winsorised returns and regularisation, and Config.8) 5% cash constraint with winsorised returns as well as regularisation. All models show an improvement from the base case, however, Config.6 and Config.8 are showing signs of overfitting, specifically for the MV portfolio. Summary statistics for all configurations are found in Tab.2.4. All sub-figures generated by: `dataprocessing_dsr.5.mlx`.

²⁴See Sec.2.4.2.2 for details on the implemented constraints for MV and HRP.

2.5 Model selection

After visually comparing the various configurations of the conventional approach to portfolio simulations in Sec.2.4.2.4 we use a quantitative approach to make the final model selection. The model that is chosen in this section should allow us to fairly compare the MV and HRP optimisations and will be the model compared to the HS-FP scenario based approach in Chp.3. For selection, we use the PSR and the DSR discussed in sections 2.2.1.1 and 2.2.1.2 as additional performance metrics along with the standard metric such as the annualised returns, annualised risk and the SR.

2.5.1 The Deflated Sharpe Ratio and multiple trials

Before assessing all metrics we need to generate the multiple trials needed to calculate the DSR. As a reminder to calculate the DSR we build on to the initial theory laid out for the PSR but this time the benchmark SR^* is defined by the expected maximum estimated SR, $\mathbb{E} \left[\max \{ \widehat{SR}_k \} \right]$ out of K independent simulated trials such that $SR^* = \widehat{SR}_0 = \mathbb{E} \left[\max_k \{ \widehat{SR}_k \} \right]$ in Eq.2.24 making the expected maximum SR the new benchmark that is used to assess the probability of observing a truly outperforming SR. As not all simulations will be independent when simulated, K is estimated from a total of M trials using López de Prado and Lewis [67]’s unsupervised learning methods.

M simulated trials are required from which K independent trials are extracted to estimate $\mathbb{E}[K]$ needed as input into the DSR benchmark $\widehat{SR}_0$ formula needed for Eq.2.26. To create the initial M we generate alternative scenarios by varying the parameters used to specify the level of cash to be included in the MV and HRP portfolios, the level of winsorisation to apply as well as the regularisation parameter selected. For each parameter below 15 evenly spaced values are stipulated within the ranges specified:

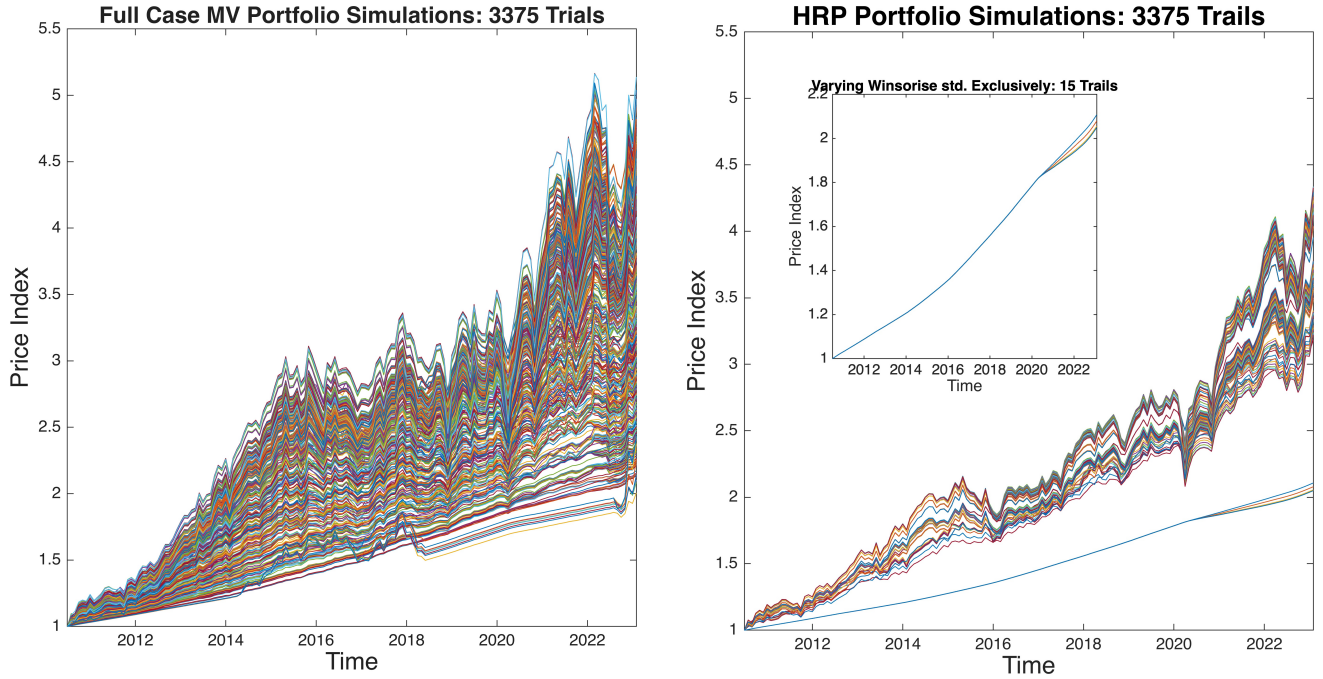
1. Cash Constraint²⁵: $\omega_{\text{Cash}} \in [0, 1]$.
2. Winsorisation standard deviation $s_{\text{Winsor}} \in [1, 5]$.
3. Regularisation $\lambda_{\text{Reg}} \in [0, 1]$.

Using the discrete parameter values in each range a 3375×3 configuration matrix is created that serves as an input to the back-testing framework. The resulting simulated performance paths are seen in Fig.2.9. The multiple trial simulations were conducted for both MV and HRP portfolios, using the varied hyper-parameters²⁶. The results highlight that there is greater sensitivity of MV portfolios to varied inputs as appose to HRP portfolios. Despite the fact that the HRP simulations have a lower degree of variability a lower overall average performance is seen, with the best performing HRP model only reaching a level of 4.25 *vs.* the 5 and above achieved by the highest performing MV portfolio.

It is worth noting that the smoothing parameters (τ half-life) applied to the signal data, as discussed in Sec.3.3, could also have been subject to variation to introduce even more variety. However, it is essential to distinguish that the signals only come into play when applying the HS-FP framework, rendering them redundant in this context as the framework is yet to be applied. In Chp.3.6 the framework is applied and assessed in its entirety and undergoes its own performance evaluation of which the DSR forms a part of.

²⁵The cash allocation is permitted to range from 0 to 1, where 0 signifies complete exclusion of the cash asset, and 1 implies that the model possesses full autonomy to allocate assets to the cash component as it deems appropriate.

²⁶The hyper-parameters were consistently varied across the simulations in Fig.2.9a and Fig.2.9b.



(a) Over 3375 simulated trials of an MV portfolio, a large degree of variability is visible in the cumulative performance curves. This is suggestive of a higher probability of obtaining a false positive considering the number of trials conducted. (b) The HRP portfolio shows a narrower spread of performance curves with two distinct path trajectories. The upper group, showing larger variation *vs.* the lower group, which has slight variation and results from varying only the Winsorisation sd. parameter.

Figure 2.9: A total of 3375 trial simulations were conducted for both MV and HRP portfolios, using varying hyper-parameters. These simulations included adjustments to data processing parameters, detailed in Tab.2.4, such as applying an equality constraint to the cash asset, regulating λ in the Ridge regularisation of the covariance matrix, and adjusting the standard deviation for winsorisation of the return series. Hyper-parameters were consistently varied across the simulations in Fig.2.9a and Fig.2.9b. The results highlight the greater sensitivity of MV portfolios to varying inputs when compared to HRP portfolios. Despite the fact that the HRP simulations have a lower degree of variability it does indicated a lower overall average performance, with the best performing HRP model only reaching a level of 4.25 versus 5 and above achieved by the highest performing MV portfolio. It is evident from the lower paths in the HRP generated from only varying the sd. for winsorisation that winsorisation does not improve the HRP method, as the cash asset still dominates cumulative performance curves. This is discussed in depth in Sec.2.4.2.3. Above plots generated by: `dataprocessing.dsr_5.mlx`.

From the 3375 simulation we apply the unsupervised learning approach of López de Prado and Lewis [67] to establish $\mathbb{E}[K]$, the optimal number of independent clusters indicating the number of independent simulated trials. In Fig.2.10 the reduction from $M = 3375$ simulations return series to K independent clustered simulations show that for the MV method $\mathbb{E}[K] = 4$ and for the HRP approach $\mathbb{E}[K] = 29$ ²⁷.

²⁷To arrive at these optimal number of clusters we specified the number of iterations as 30 and the maximum number of clusters for the K-Means algorithm as 30. See A6.1 Multiple simulation test cases for details on implementing the simulation in practice.

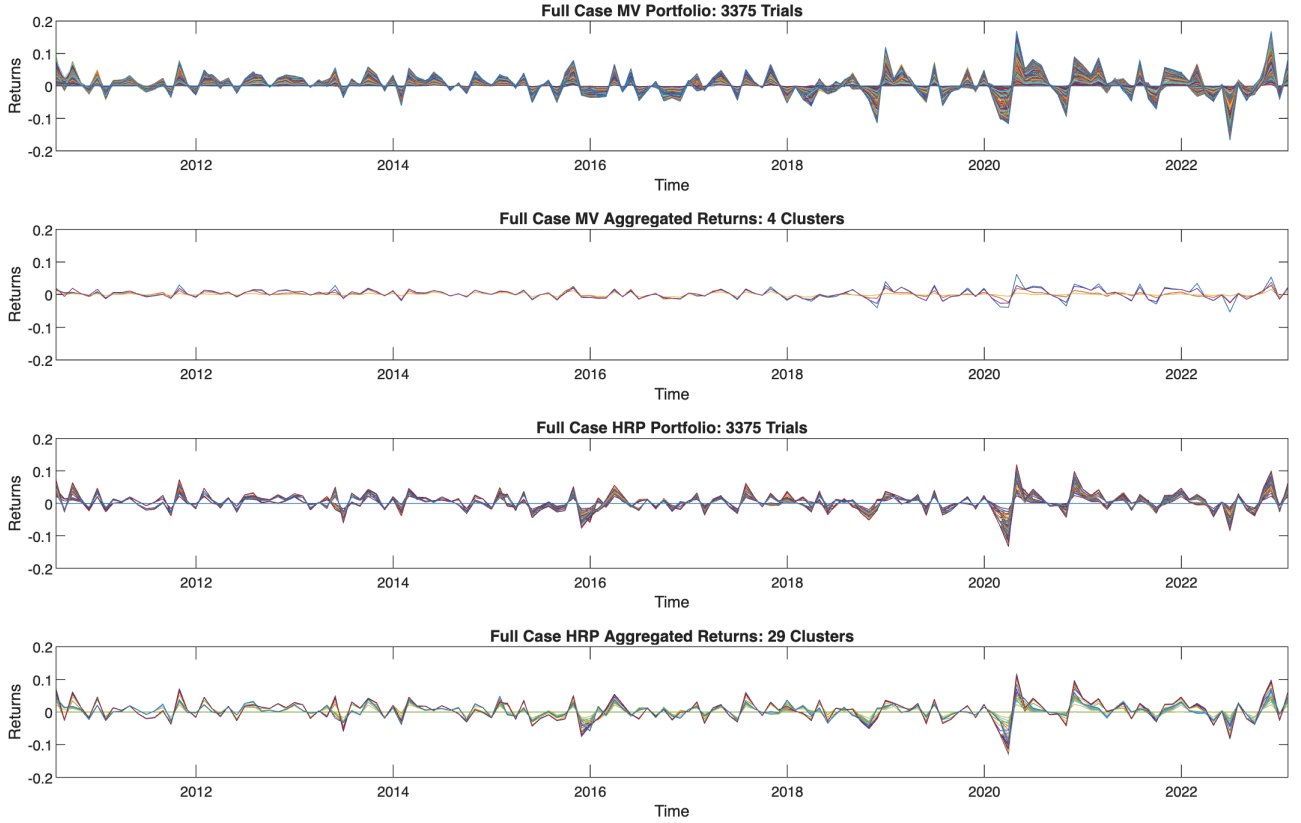


Figure 2.10: To calculate the DSR we build on to the initial theory laid out for the PSR but this time the benchmark SR^* is defined by the expected maximum estimated SR, $\mathbb{E}[\max\{\widehat{SR}_k\}]$ out of K independent simulated trials. Each trial consists of alternative hyper-parameter configurations that need to be generated. However, not all configurations will be independent when simulated hence K would need to be estimated from a total of M trials. For this López de Prado and Lewis [67]’s unsupervised learning methods is use to estimate $\mathbb{E}[K]$ as the optimal number of clusters from M simulated trials. Above we show $M = 3375$ trials being reduced through López de Prado and Lewis [67]’s iterative clustering approach. For the MV portfolio the algorithm results with $\mathbb{E}[K] = 4$ and for HRP $\mathbb{E}[K] = 29$. Generated by: `dataprocessing.dsr.5.mlx`

2.5.2 Results

Using the $\mathbb{E}[K]$ calculated from the 3375 simulations we calculate the DSR for each configuration in Sec.2.4.2.4. In assessing all metric results²⁸ in Tab.2.4 below we do not want to select a model that overfits to the data but we need enough “give” such that the models can execute successfully. We should view the metrics in Tab.2.4 in conjunction with the insights from the visual analysis in Sec.2.4.2.4 to form a well rounded opinion of the configuration performances. For instance, even though the SR for the HRP in the base case (Config.1) seems reasonable at 0.17 relative to the other configurations (with SR ranging from 0.13 - 0.18), we know that the portfolio consisted solely of cash and that the portfolio’s cumulative returns are the lowest performing in Fig.2.7a.

The SR assesses standard deviation as a measure of risk associated with returns, but it is not a well suited metric for return distributions where there is skewness lesser or greater than zero *i.e* distributions that deviate from a normal distribution [66]. In the case of the cash asset which dominates the base case HRP there is significant rightward skewness of 0.83 (See Fig.3.4). The standard deviation used in

²⁸In Tab.2.4, the PSR uses $SR^* = 0$ as the benchmark portfolio indication no investment skill comparison (see Eq.2.23) and all measures are calculated on excess returns (realised portfolio returns less cash), $\omega_{\text{Cash}} = 0.05$, $\lambda_{\text{Reg}} = 0.005$, and $s_{\text{Winsor}} = 2$.

the SR denominator takes into account both positive and negative deviations from the mean, therefore treating upside volatility and downside volatility equally²⁹. As the SR can be misleading we will place more emphasis on the PSR and the DSR for models selection, taking into account selection bias and non-normal returns.

Table 2.4: The columns of the table are performance metrics from each configuration in Sec.2.4.2.4. The processes applied in each configuration are listed (where Y = process was applied, and N = process not applied) as well as the figure number that shows their cumulative performance. Relative performance is assessed for the MV and HRP using the SR, the annualised SR, PSR, and DSR across the rows. The SR is calculated monthly and the PSR uses $SR^* = 0$ as the benchmark portfolio indication no investment skill comparison (see Eq.2.23). The DSR is calculated as set out in Sec.2.2.1.2 and Sec.2.5.1. The MV portfolio under performs for all metrics across all configurations. When results are considered in conjunction with figures 2.7 and 2.8 Config.5 with a cash constraint and Winsorised returns shows the most well balanced results for both the MV and the HRP and will therefore be selected as the final model to compare with the HS-FP approach in Chp.3. All measures are calculated on excess returns (realised portfolio returns less cash), $\omega_{\text{Cash}} = 0.05$, $\lambda_{\text{Reg}} = 0.005$, and $s_{\text{Winsor}} = 2$.

DATA PROCESSING: RELATIVE PERFORMANCE OF CONVENTIONAL BACK-TEST CONFIGURATIONS								
	Data Configurations							
Processes	Config 1: Unadjusted	Config 2: CC	Config 3: W	Config 4: R	Config 5: CC& W	Config 6: CC & R	Config 7: W & R	Config 8: All
Cash Constraint (CC)	N	Y	N	N	Y	Y	N	Y
Winsorise (W)	N	N	Y	N	Y	N	Y	Y
Regularise (R)	N	N	N	Y	N	Y	Y	Y
See Figure:	2.7a, 2.4	2.7b	2.7c	2.7d	2.8a	2.8b	2.8c	2.8d
Performance Metrics								
SR ³⁰								
MV	0.0450	0.0943	0.0067	0.1136	0.1469	0.1271	0.1230	0.1375
HRP	0.1677	0.1268	0.1804	0.1288	0.1742	0.1345	0.1649	0.1667
Annualised SR								
MV	0.5438	1.1532	0.0799	1.3993	1.8010	1.5715	1.5124	1.6976
HRP	2.0125	1.5464	2.1648	1.5688	2.1297	1.6519	2.0138	2.0518
PSR								
MV	0.7179	0.8732	0.5325	0.9165	0.9641	0.9386	0.9311	0.9520
HRP	0.9832	0.9283	0.9870	0.9343	0.9823	0.9444	0.9775	0.9792
DSR (Calculated using $M = 3375$ simulated paths)³¹								
MV	0.6229	0.8140	0.4323	0.8712	0.9392	0.9021	0.8917	0.9215
HRP	0.9688	0.8899	0.9758	0.8976	0.9683	0.9114	0.9604	0.9630

Across the configurations in Tab.2.4 we note large fluctuations in the PSR and DSR for the MV portfolio while the HRP remains fairly consistent in comparison. This suggests that the OOS performance of the HRP portfolio is more reliable and less sensitive than MV to changes in the data. The under performance of the MV relative to the HRP in the base case (Config.1) across all metrics can be attributed to the volatility spike in its cumulative performance curves between 2014-2018 (and slight spike at the end of 2022) worsening the risk/reward ratio of the SR. As intervention processes are introduced (in the table left to right) the performance of both HRP and MV show improvement, bar Config.3 where only Winsorisation was implemented. Given Winsorisation shows no model performance improvement we conclude that the presence of outliers is not the exclusive cause of under performance of MV and HRP in Config.1. The cash constraint and regularisation result in increased performance, however in conjunction with the cumulative performance curves and the DSR results we opt to exclude regularisation as a solution to the structural concerns as the results suggest overfitting.

²⁹If assessing risk is the primary focus, this limitation can be addressed by using alternative risk measures such as Value at Risk (VaR), Conditional Value at Risk (CVaR).

²⁷All SRs and derivatives are calculated using excess returns as defined in the *ex post* calculation set out in Sec.1.2.2.4. It makes use of the geometric averages of the risky asset and a risk-free asset given realised returns of both. The historical realised excess returns are used to calculate an indicative measure of historical performance.

³¹See Sec.2.5.1 for calculation details.

By introducing a cash constraint combined with Winsorisation in Config.5 the probability of outperforming a zero investment skill benchmark portfolio increases from a PSR of 0.72 to 0.96 for the MV portfolio, however for the HRP portfolio the probability remains stable at ~ 0.98 . Accounting for multiple testing bias, non-normal returns, and a short sample length in Config.5 yields a DSR of 0.94 for the MV and 0.97 for the HRP, giving us greater confidence that this configuration is likely reliable and not a false positive. As a result, Config.5 appears to deliver the most balanced portfolio outcomes for both methods. Therefore, Config.5, as summarized in Fig.2.11 and Tab.2.5, is selected as our final model for comparison with the HS-FP approach in Chp.3.

Optimal Weight Surfaces for Non-HSFP Backtest with 5% Cash Constraint & 2sd. Winsorised Returns: 36 Month Rolling Window

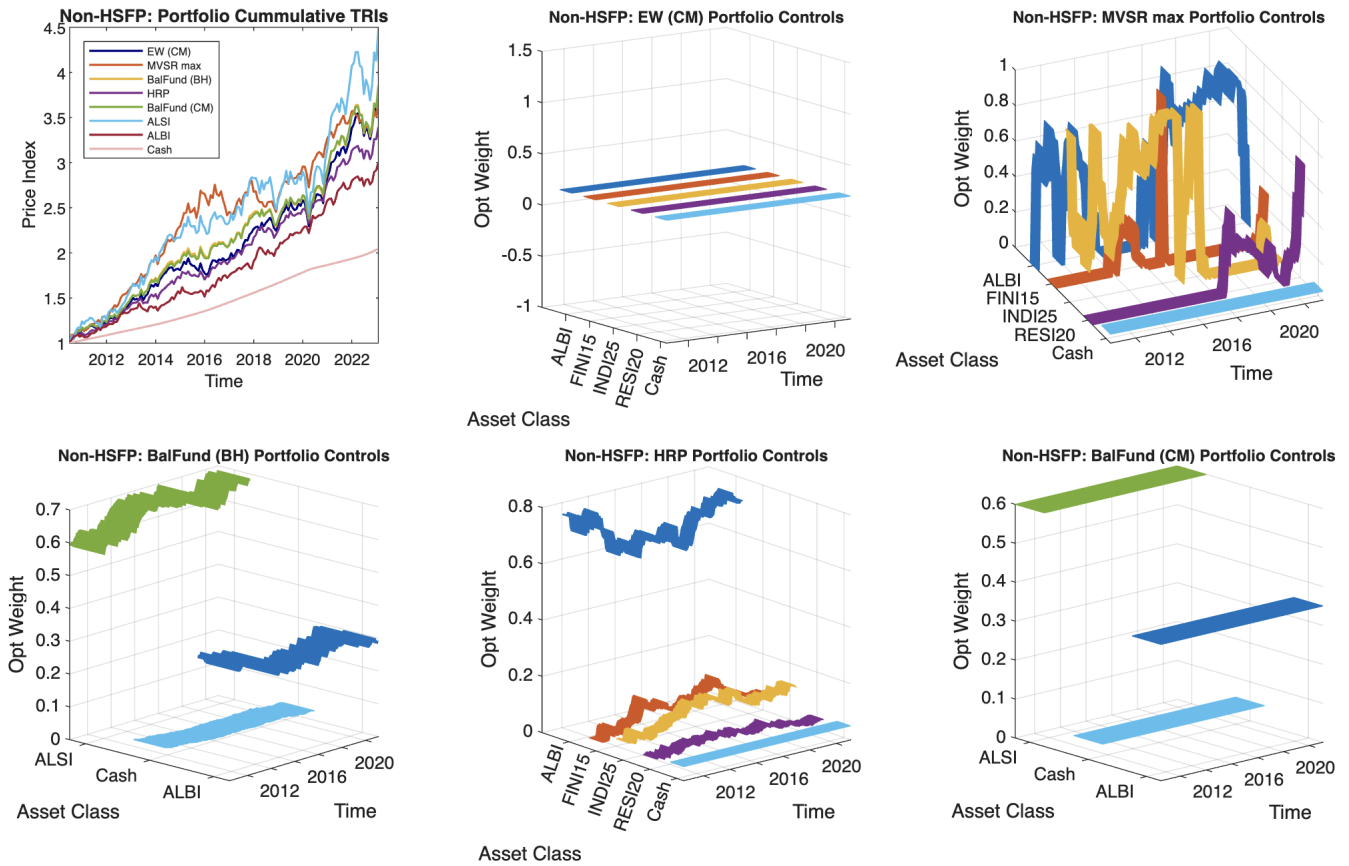


Figure 2.11: Portfolio weight surfaces generated for Config.5 seen as seen in Tab.2.4 which results from applying a 5% cash equality constraint for the MV and HRP portfolio types as well as Winsorisation of the returns using 2 standard deviations. In comparison to Fig.2.4, both the MV and HRP portfolios show improvement in their cumulative performances curves (top left). Comparing the weight surfaces of the MV (top right) with the HRP (bottom middle) we see that the MV has a high turnover rate *vs.* the HRP portfolio. Generated by: `dataprocessing_dsr.5.mlx`.

Table 2.5: Conventional back-test results of Config.5 as seen in Tab.2.4 and figures 2.8b & 2.11 using a 5% cash equality constraint³² for MV and HRP as well applying a 2 sd. Wisorisation to the returns prior to simulation. HRP is the weakest performer on an annualised return basis but due to it having the lowest volatility it is the best performing SR, annualised SR, PSR and DSR. Of all portfolios the MV has the highest maximum drawdown and the lowest PSR and DSR suggestive of having the highest chance of overfitting OOS relative to all other portfolios and benchmarks.

Configuration 5: Conventional Application Portfolio Results								
		Performance Metrics						Additional Metrics
	Portfolio Type	Ann.R(%)	Ann. σ (%)	SR ³³	Ann.SR ³³	PSR	DSR	MaxDD(%)
1	Equally Weighted (CM)	11.0280	8.6647	0.1592	1.9534	0.9745	0.9553	10.9730
2	MV SR Max	10.8760	9.1301	0.1469	1.8010	0.9641	0.9392	12.3480
3	Balanced Fund (BH)	11.3360	9.1410	0.1597	1.9619	0.9743	0.9552	10.9160
4	HRP	10.1980	6.6871	0.1742	2.1297	0.9823	0.9683	6.2954
5	Balanced Fund (CM)	11.2980	8.6065	0.1685	2.0690	0.9801	0.9644	10.1540
	Market Benchmarks							
6	ALSI	12.6870	12.8550	0.14035	1.7332	-		15.9820
7	ALBI	9.1281	6.4686	0.13626	1.6581	-		6.4854
8	Cash	5.8572	0.3497	-	-	-		0

We can now assess the relative performance in a conventional back-test setting by using Config.5 where both optimisations of interest, MV and HRP, are subject to the same treatment allowing for a fair comparison. Upon initial view of the results of the portfolio strategy and benchmarks in Tab.2.5, the diversification focused HRP strategy performs the worst on an annualised return basis but has the lowest volatility and best SR, Ann.SR, PSR and DSR. The results of this conventional simulation leads us to believe that the HRP produces more reliable results and will perform better OOS. The question is whether this is a result of its diversification focus or could it be an artefact of it requiring fewer historically dependent estimated inputs as initially hypothesised (relative to the MV).

Further observations:

1. The BF-CM appear to be achieving volatility capture due to its monthly rebalancing relative to the BF-BH. This implies that there is in fact volatility to be capture.
2. As the HRP outperforms the EW on a risk adjusted return basis, with an Ann.SR of 2.13 *vs.* 1.95. This suggests that there is a degree of correlation signal that the HRP is able to use to achieve correlation diversification.

The HRP's out performance may not be due to reduced estimation inputs as initially hypothesised, but rather its ability to achieve correlation diversification. This is evident as the EW portfolio—which would typically outperform in cases of covariance misestimation [31]—did not surpass the BF-CM or HRP portfolios. The fact that BF-CM outperforms the EW portfolio implies the covariance matrix is well-estimated (consistent with findings that regularisation led to overfitting). If diversification alone were the reason for out performance, the EW portfolio would have beaten HRP by maximising diversification across uncorrelated assets.

²⁹See Sec.2.4.2.2 for details on the implemented cash constraint for MV and HRP.

³³All SRs and derivatives are calculated using excess returns as defined in the *ex post* calculation set out in Sec.1.2.2.4. It makes use of the geometric averages of the risky asset and a risk-free asset given realised returns of both. The historical realised excess returns are used to calculate an indicative measure of historical performance.

However, HRP's focus on two types of diversification — number of assets and, more importantly, correlation diversification — allows it to hedge risks more effectively, as seen in improved maximum drawdown metrics.

2.5.2.1 Implication of results

Considering the extreme volatility in the SA market between 2010 to 2023, the results from this conventional back-test suggest that managing for correlation diversification could prove to be more effective than focusing on return expectations. The SA market's concentrated sectors and correlated noise make asset-level return predictions unreliable. Instead, predicting underlying market states or sectoral movements, though challenging due to sector concentration on the JSE, could provide better outcomes. This insight supports our scenario-based approach using market state conditioning based on indicators linked to specific sectors, introduced in Chap.3, Tab.3.1. It suggests that we are on the right track in providing a more robust approach to stress-testing a portfolio, particularly in a South African market context.

2.6 Summary

This chapter fulfils the first objective of the dissertation: assessing the relative performance of MV optimisation and the HRP method. We systematically constructed a conventional back-test framework by building up its key components: beginning with the theory for the optimisations and benchmark, asset class selection, data processing, and simulation design. We highlight key points illustrated by figures: Fig.2.2 shows the composition of the SA Equity market and the rationale for using the FINDI, INDI, and RESI representative of equity sectors, while Fig.2.3 demonstrates the treatment of the cash asset by transforming the JIBAR into a cash proxy by removing extreme variability.

Initial back-test results revealed that both the MV and HRP unconstrained optimisations over-allocated to cash, which prompted an exploration into the underlying data structure in Fig.2.6. To improve on the optimisation's performance and control overfitting, we compared alternative model configurations, as shown in Fig.2.7 and Fig.2.8. Through visual and quantitative analysis of the configurations, we methodically eliminate pure Winsorisation and regularisation as methods of resolving the over-allocation to cash in the base case as they resulted in decreased performance and overfitting respectively. Through a process of elimination and considering performance metrics captured in Tab.2.4 we concluded that Config.5 produced a balanced simulation, incorporating a 5% cash constraint and applying Winsorisation to monthly returns as shown in Fig.2.11. Therefore, Config.5 was selected as the final conventional back-test model to compare with the HS-FP framework in Chp.3.

Using Config.5, we compared the MV and HRP strategies under identical conditions. The HRP strategy, despite having the lowest annualised return, achieved superior risk-adjusted performance metrics relative to the MV and other benchmarks. We deduced that the HRP's outperformance is due to its ability to achieve correlation diversification rather than reduced estimation inputs (counter to our initial hypothesis). These results imply that managing for correlation diversification is more effective than focusing on return expectations. Asset-level return predictions are unreliable; predicting market states or sector movements would yield better outcomes. This supports our scenario-based approach using market state conditioning linked to specific sectors going into Chp.3.

Chapter 3

Forward looking simulations

The results obtained from the conventional approach of back-testing the MV and HRP methods in Chp.2 do not adequately address the goal of stress-testing posed in the Introduction Sec.1.1.3.1: “...to enhance transparency in market risk by assessing potential portfolio P&L under abnormal market conditions” [12]. This is the result of both an explicit and implicit assumption that limits its effectiveness. The explicit assumption: the underlying data is *i.i.d.* normally distributed, and the implicit assumption: sampled historical market data encompasses all possible future market conditions. Yet, in practice, market data is known to violate these distributional assumptions, and the use of historical data in optimisations is known to introduce generalisation errors into the models [30, 66, 116]. It is not realistic to assume that all possible abnormal market conditions can be contained within a single historical snapshot. Models exclusively relying on historical data such as those in Chp.2 are therefore predisposed to compromised OOS performance.

As emphasised in [Context: time series analysis](#) in the discussion on ergodic theory, historical paths are a single observed sequence of historical events; we lack definitive knowledge of the probability of this sequence, how market participants will adapt to new data, or the rate at which information dissipates [100, 103]. Consequently, the components of a back-test that are most affected by the omission of future market information are the predictors μ and Σ . These inputs need to be updated to reflect *forecasted* measures of risk and return rather than historical measures.

This chapter addresses the second aim of the dissertation: to evaluate whether the incorporation of the non-parametric scenario-based HS-FP framework enhances the effectiveness of the MV or HRP methods by reliably capturing regime changes and reducing reliance on restrictive assumptions about financial time series data. The objective of implementing the framework is to enhance the ability to capture market regimes, increase the models’ responsiveness to evolving market conditions, minimise the likelihood of extreme drawdowns, and ultimately deliver improved risk-adjusted returns. To begin, we outline the HS-FP framework, focusing on the derivation of invariants from market drivers and exploring the various methods proposed by Meucci [85, 89] for defining flexible probabilities. We also look at market state variable selection and the process of deriving effective state signals or indicators from these variables for use in state-conditioned FPrs. The objective is to establish a coherent approach that bridges “what we have” — historical market data—with “what we want”—forward-looking estimates suitable for stress-testing. The back-test design is updated accordingly, incorporating HS-FP forward-looking estimates of μ and Σ into the MV and HRP optimisations, replacing the conventional reliance on historical predictors.

3.1 Intro to scenario analysis

Scenario analysis is a more intuitive approach to portfolio stress-testing as it does not require the specification of statistical features. The burden on the investor is lighter, only requiring that a small set of possible outcomes are provided as projections of market state variables such as inflation. Investors want to be able to test different portfolios under different sets of subjective views such as “What would be the impact of an inflation rate increase?” More than a decade ago, Black and Litterman [18] (BL) introduced their approach to generalised scenario analysis that allowed investors to incorporate and account for possible uncertainty in their market views or the reference risk model used. However, it was limited as it was only capable of working with normal markets and linear views.

Researchers such as Almgren and Chriss [2], Pezier et al. [106], and Qian and Gorman [109] have since worked on individual aspects of the BL model for improvement. In 2005, Meucci [82] attempted to extend the BL methodology and generalise it to non-normal markets and non-normal views using copula and opinion pooling principles. This approach, however, did not allow for non-linear views or correlation stress-testing. Further generalisation still needed to be achieved. In 2008, Meucci [84] posed the question of how to input investors’ non-linear views into a non-normal market. This resulted in his entropy pooling (EP) methodology, which can handle views on non-linear combinations of market state variables/risk drivers that influence returns and therefore allowed for non-normal markets.

Using scenario analysis in this way allows investors to determine how a set of subjective views on possible market realisations would impact a given portfolio. With this, Meucci further enhanced scenario-based risk management by increasing its flexibility. By changing the probabilities associated with each scenario outcome, he derived a resulting forward-looking return distribution for a portfolio. This became his HS-FP framework [85].

From 2008 to 2013, Meucci developed a comprehensive non-parametric framework that culminated in his paper, “Estimation and Stress-Testing via Time- and Market-Conditional Flexible Probabilities” [89]. The HS-FP framework evolved into the Time- and Market-Conditional Flexible Probabilities framework, used for estimation and stress-testing. The value of HS-FP lies in the flexible probabilities (FPrs) conditioning future asset behaviour on time and market state variables [89]. Meucci reviewed common approaches to time and state-conditioned probabilities and proposed a method for joint conditioning using entropy pooling, as well as a method to consider multiple variables of time and state conditioning [84, 89]. By incorporating these approaches into a back-testing framework, asset allocation models should be more responsive to changes in market conditions and hence should result in better OOS performance.

3.2 Mapping drivers into the market model: the HS-FP framework

Both the MV and HRP offer a normative theoretical framework. However, to make it more relevant in practice, we introduce the HS-FP framework. The framework incorporates a forward-looking component in portfolio optimisation and the back-test framework. Forward-looking estimates of $\hat{\mu}$ and $\hat{\Sigma}$ are provided to the MV and HRP optimisations encoded with information of the market-state. To understand how the HS-FP framework is built into the back-test design we refer to Fig.3.1 which diagrammatically shows

how HS-FP processes are fed into the back-testing framework. **Fig.3.1** should be used as a guide to this chapter as each component comprises a different section. See caption for details.

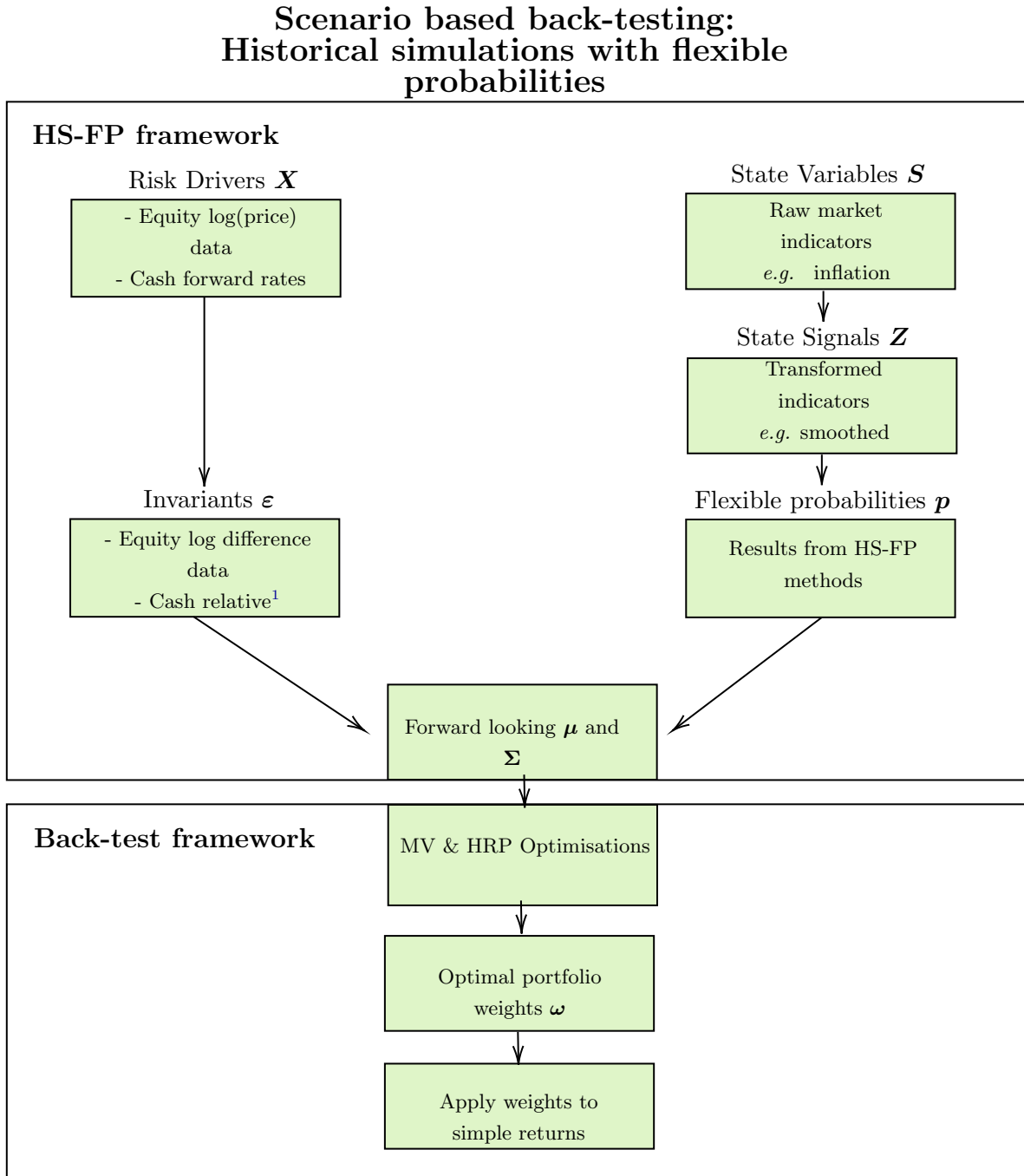


Figure 3.1: To implement Meucci's HS-FP framework in a back-test framework the flow of data requires careful management as there are different transformations needed for the asset data at different points in the data flow. Notably for equity data, the risk drivers $\mathbf{X}$ need to be homogeneous across time therefore the log of the price data is needed, and for equity invariants ε the series needs to be *i.i.d.* so the difference of the logs prices are used. For the cash asset forward rates of the JIBA3M are used as risk drivers, and for the invariants, a ratio of the forward rates is used to transform the data into *i.i.d.* observations. Refer to sections 3.2.1 and 3.2.2 for explanations of the assumptions and requirements of risk drivers and invariants. These invariants ε are then used in conjunction with the HS-FP FPrs $\mathbf{p}$ derived from state signals $\mathbf{Z}$ to generate forward-looking estimates of optimisation inputs $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$, which are then passed into the back-test framework. The optimal weights $\boldsymbol{\omega}$ are therefore generated with a forward looking view which allows us to introduce a scenario-based approach to the back-test. The processing of obtaining the state signals $\mathbf{Z}$ from the raw state variables $\mathbf{S}$ is found in Sec.3.3, while alternative methods for obtaining $\mathbf{p}$ are detailed in Sec.3.4, and the result of applying the scenario based back-test is discussed in Sec.3.6.

The data used in these frameworks need to be handled with care as different forms of the data are needed at varying points in the processes. The HS-FP framework aims to reduce distributional assumptions and incorporate market information drawn from state variables that are known key macroeconomic indicators such as inflation or GDP. The next sections walk through the process of getting from raw market drivers and state variables to forward looking market estimates of $\hat{\mu}$ and $\hat{\Sigma}$.

3.2.1 Risk drivers

To model financial markets effectively, we need to identify market invariants – variables that are stable and predictable over time. Given that financial time series data is known to be unpredictable we need to find an underlying thread that is stable enough and has *i.i.d* observations to be able to pull it forward. Therefore, these invariants should be time homogeneous **and** *i.i.d*. To be able to fit the HS-FP framework and model the market we define a set of risk drivers, $\mathbf{X}$, from which to extract the invariants. Consider the set of equally spaced time points $\mathcal{D}_{\tilde{t}, \tilde{\tau}}$ where $\tilde{t}$ is the starting point of the set and $\tilde{\tau}$ are the intervals at which these invariants are observed:

$$\mathcal{D}_{\tilde{t}, \tilde{\tau}} \equiv \{\tilde{t}, \tilde{t} + \tilde{\tau}, \tilde{t} + 2\tilde{\tau}, \dots\}. \quad (3.1)$$

In our case, the time interval $\tilde{\tau}$ would be 1 month increments, with $\tilde{t}$ as the starting point of the sample dataset as shown in Sec.2.3.2:

$$\mathcal{D}_{\tilde{t}=29/02/2000, \tilde{\tau}=1_{\text{mth}}} \equiv \{\tilde{t}, \tilde{t} + 1_{\text{mth}}, \tilde{t} + 2_{\text{mth}}, \dots\}. \quad (3.2)$$

However, note the use of the tilde is to indicate that the series can begin at any random starting point as the series is only time-homogeneous if its distribution is not dependent on the reference point $\tilde{t}$.

Definition 28 (Risk drivers). Over a finite time frame $(t, \dots, T)$, we can define the set of D risk drivers:

$$\mathbf{X}_t \equiv (X_{t,1}, \dots, X_{t,D})'. \quad (3.3)$$

To qualify as a risk drivers the series needs to follow a stochastic process that is homogeneous across time (but not *i.i.d*) and needs to fully describe the price of an asset at any given time (along with potential asset terms and conditions) [86].

From Meucci [83]’s work we know that asset price data $\mathbf{P}$ itself cannot be a risk driver as they are not homogeneous through time, however, log-prices can be used:

$$\mathbf{X}_t \equiv \ln(\mathbf{P}_t). \quad (3.4)$$

We will use the history in a stochastic environment to learn about the future behaviour by extracting a set of “invariants” from the risk driver time series. Invariants represent those patterns that we observe to repeat themselves *i.i.d*. over time.

¹Ratio of monthly forward rates see [83] pg. 109-114.

3.2.2 Invariants

Definition 29 (Invariants). The invariant series ε_t should encompass the disruptive shocks that steer $\mathbf{X}_t$'s stochastic process over a given time step $(t) \rightarrow (t+1)$. A set of V joint invariants is represented by:

$$\varepsilon_t \equiv \varepsilon_{t \rightarrow t+1} \equiv (\varepsilon_{1,t \rightarrow t+1}, \dots, \varepsilon_{V,t \rightarrow t+1})'. \quad (3.5)$$

To qualify as an invariant the series are *i.i.d.* across time steps^a and they are known at each time step $(t+1)$. Here, the notation $\varepsilon_{t \rightarrow t+1}$ is used to emphasise the time series nature of the invariants. The importance of invariants is that they allow for market randomness to be projected forward as they are the distilled repeated patterns underlying the risk drivers.

^aImportantly, Meucci [89] notes that the weaker assumption of ergodicity is sufficient.

3.2.2.1 Testing for invariance

We consider the set of random variables $\mathbf{X}_t$ as our risk drivers defined in Eq.3.4 over the evenly spaced time points $\mathcal{D}_{\tilde{t}, \tilde{\tau}}$:

$$\mathbf{X}_t \equiv \ln(\mathbf{P}_t), \quad t \in \mathcal{D}_{\tilde{t}, \tilde{\tau}}, \quad (3.6)$$

and denote the historical realisations with a lower case $\mathbf{x}_t$ and the most recent observation as T :

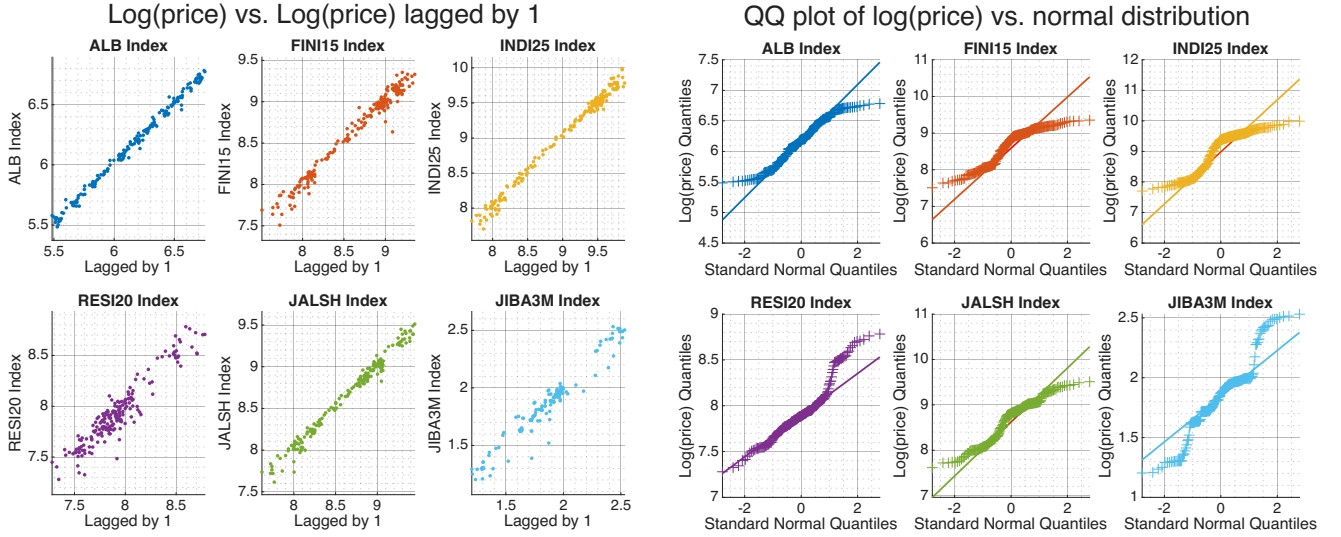
$$\mathbf{x}_t, \quad t = \tilde{t}, \tilde{t} + \tilde{\tau}, \dots, T. \quad (3.7)$$

Meucci [83] uses two tests² to graphically determine if the time series is *i.i.d.* 1.) a scatter plot of the series $\mathbf{x}_t$ against its lagged series $\mathbf{x}_{t-\tilde{\tau}}$ helps assess whether the observations are independent of each other and 2.) splitting $\mathbf{x}_t$ into two segments allows for the comparison of two separate distributions within $\mathbf{x}_t$ to observe whether they are identically distributed or not:

$$\begin{aligned} \text{Segment 1: } \mathbf{x}_t, \quad t = \tilde{t}, \dots, \tilde{t} + \left\lceil \frac{T - \tilde{t}}{2\tilde{\tau}} \right\rceil \tilde{\tau}, \\ \text{Segment 2: } \mathbf{x}_t, \quad t = \left(\left\lceil \frac{T - \tilde{t}}{2\tilde{\tau}} \right\rceil + 1 \right) \tilde{\tau}, \dots, T. \end{aligned} \quad (3.8)$$

If the data is approximately *i.i.d.* the scatter plot will display symmetry with respect to the reference axes and will roughly resemble a circular scatter of points. The two histograms of the segmented $\mathbf{x}_t$ should appear closely similar to each other to indicate that the data is identically distributed. As a precursor to extracting the invariants, we show in Fig.3.2 below with the use of the lagged scatter plot, that the risk drivers do not meet the invariant criteria of being *i.i.d.*. We also use a QQ plot to assess deviation from the normal distribution to get a better understanding of the distributional properties of the drivers and invariants.

²These tests are not definitive however Meucci argues that they are sufficient to fulfill our argument. See Hamilton [43], Campbell et al. [21] and Campbell and Viceira [20] for further support.



(a) A linear relationship is observed between all assets' monthly risk drivers $\mathbf{X}_t \equiv \ln(\mathbf{P}_t)$ and their 1-month lag, indicated by the dependency of x_t on the y-axis and its lag x_{t-1} on the x-axis, where $\mathbf{P}_t$ represents asset prices. (b) The tails of the QQ plots indicate deviation from normality with asymmetric heavy tails for all asset class risk drivers $\mathbf{X}_t \equiv \ln(\mathbf{P}_t)$ where $\mathbf{P}_t$ are the asset prices.

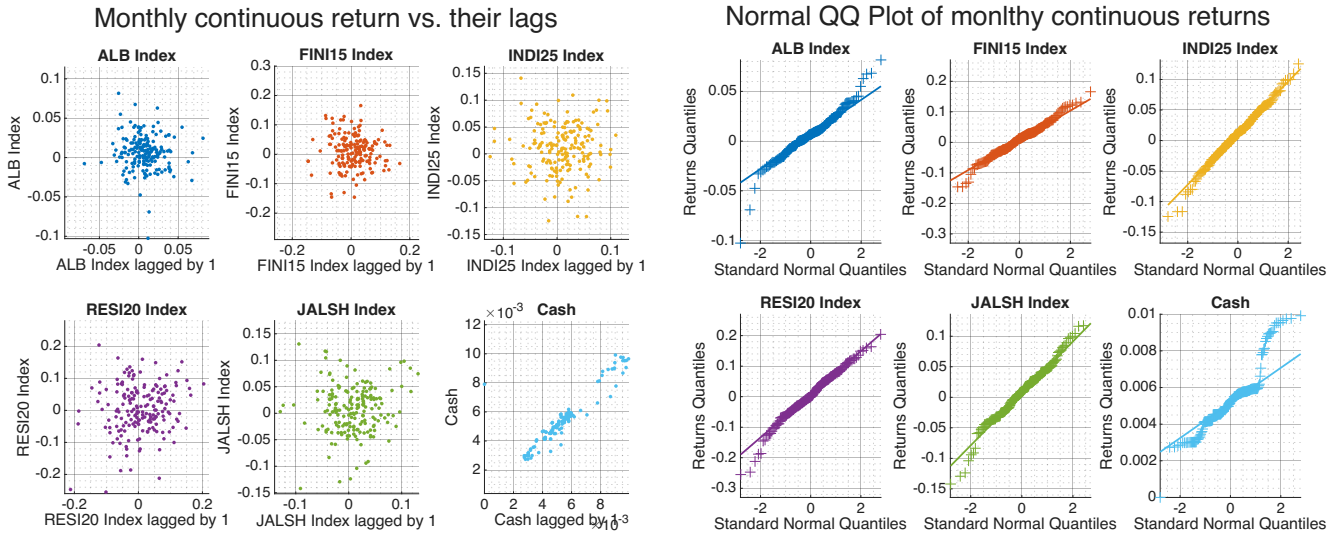
Figure 3.2: The purpose of the above plots is to show that the distribution of risk drivers is not yet suitable as invariants for the HS-FP framework, as invariants need to be *i.i.d.* The historical series of asset log price data $\ln(\mathbf{P})$ is used as the risk drivers $\mathbf{X}_t$. To assess independence, x_t is plotted against its lagged series $x_{t-\tau}$. The linear dependency shown in Fig.3.2a confirms that log price data are not independent and must be transformed to be valid invariants as defined in Def.29. Both figures are generated by: `invariant_returns_prep_3b.mlx`.

Fig.3.2 shows that the risk drivers $\mathbf{X}_t \equiv \ln(\mathbf{P}_t)$ are not yet suitable for invariants. We, therefore, transform them into invariants following Meucci [86] by obtaining the continuously compounded returns $\ln\left(\frac{\mathbf{P}_{t+1}}{\mathbf{P}_t}\right)$ and put it in the form of the log differenced asset prices:

$$\varepsilon_{t \rightarrow t+1} \equiv \ln(\mathbf{P}_{t+1}) - \ln(\mathbf{P}_t). \quad (3.9)$$

Using Eq.3.9, we transform the price data (risk drivers) for equity assets into continuous returns $\mathbf{r}_t$, obtaining the invariants shown in Fig.3.3. To maintain consistency, we use returns data across all asset classes. The forward interest rates introduced in Sec.2.3.1.3 are therefore used as a proxy for cash returns.

By converting the data into returns, we observe that equity assets exhibit greater independence between observations, as shown in Fig.3.3a. However, a distinct linear trend remains for the cash asset in the lagged scatter plot. This suggests that while continuous equity returns align well with the definition of invariants by Meucci [83, 86], the cash asset requires a different treatment.

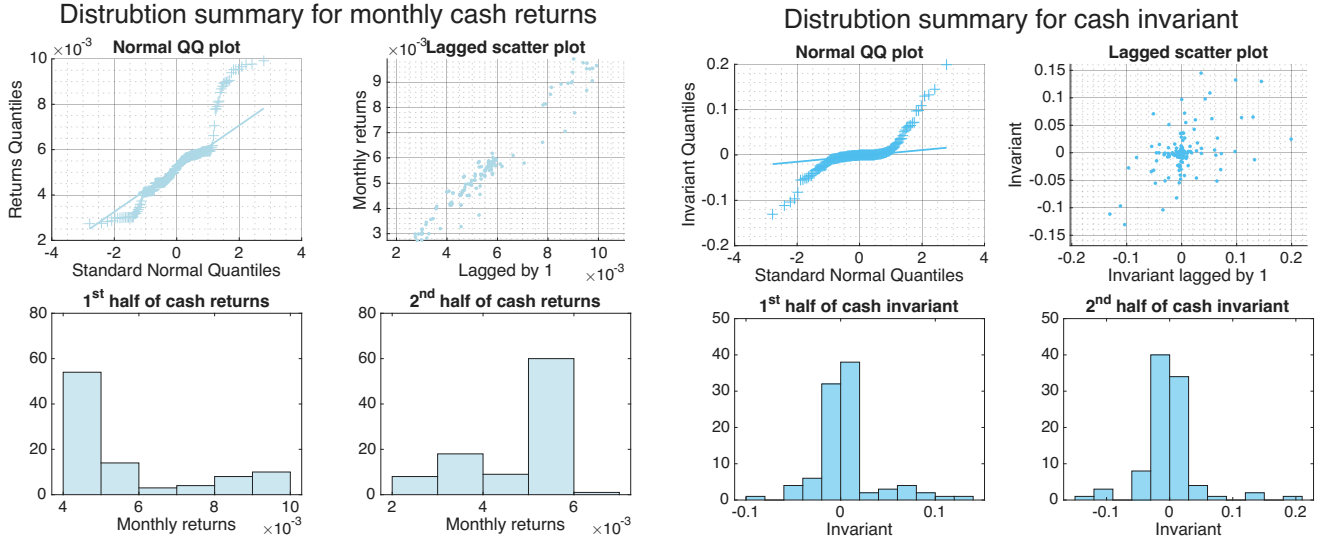


(a) In comparison to the log prices in Fig.3.2a, the continuous returns of equity assets show reasonable independence between the observations and their lag, however, a linear dependency persists for cash returns proxy.

(b) The QQ plots of the equity assets show improvement from Fig.3.2b however there is still a degree of asymmetry. The cash asset still indicates significant deviation from the normal distribution.

Figure 3.3: By transforming the equity monthly price series into continuous returns we obtain distributions that are more closely *i.i.d* and more normally distributed. Both figures are generated by: `invariant_returns_prep_3b.mlx`.

Deriving the cash invariant: Following Meucci [83], the cash asset will be treated as a fixed income instrument with a known maturity of 3 months. In line with Meucci's analysis of fixed income instruments with short maturities, we know that the cash asset value is known at the end of a 3 month period and the forward interest rates are akin to a zero coupon bond price converging to the redemption value as the maturity approaches. As the forward rates are non-overlapping but have the same time to maturity we can use the change in yield to maturity as the cash invariant. By making this transformation we see in Fig.3.4 below that the invariant for cash is closer to being *i.i.d*.



(a) The comparison of time series segments as formulated in Eq.3.8 show visually low similarity. This suggests that the data is not identically distributed. (b) The transformation of the forward rate cash proxy to the change in yield to maturity shows increased independent of observations and increased similarity in the time series segment.

Figure 3.4: As the cash uses the forward rate as a return proxy the extraction of invariants was different to that of the equity assets. Using Meucci [83]’s approach we transform the forward rates into the change in yield to maturity to extract a more *i.i.d.* distribution that can be used as an invariant in the HS-FP framework. Both figures are generated by: `invariant_prep_3b.mlx`

The lagged scatter plot result showing an increase in independence from Fig.3.4a to Fig.3.4b and the comparison of segments as discussed in Sec.3.2.2.1 (Eq.3.8) suggested that the data is closer to identically distributed than before. These results are not perfect but it does provide a better invariant candidate for the cash asset to take into the HS-FP framework.

3.2.3 Forward looking distribution

If the time series of the market invariants $\varepsilon_{V,t}$ contains a large number of historical observations, then by the LLN we can estimate the future distribution ε_{T+1} of the invariant non-parametrically. To obtain a forward-looking distribution of the invariants the HS-FP framework applies FPrs $\mathbf{p}$ (that are either time or market state conditioned) to the historical realisations of the invariants as shown in Def.30.

Definition 30 (Future distribution of market invariants). Using the invariants $\varepsilon_{V,t}$ as defined in Def.29 the future **yet-to-be realised** set of invariants ε_{T+1} can be represented via the HS-FP framework by applying unequal time and state conditioned probabilities $\{\mathbf{p}_t\}_{t=1}^T$ to a time series of historical observations $\{\varepsilon_t \equiv (\varepsilon_{1,t}, \dots, \varepsilon_{V,t})\}_{t=1}^T$. The future distribution of V market invariants ε_{T+1} are estimated by:

$$\varepsilon_{T+1} \equiv \begin{pmatrix} \varepsilon_{1,T+1} \\ \vdots \\ \varepsilon_{V,T+1} \end{pmatrix} \sim \left\{ \varepsilon_t \equiv \begin{pmatrix} \varepsilon_{1,t} \\ \vdots \\ \varepsilon_{V,t} \end{pmatrix}, \mathbf{p}_t \right\}_{t=1}^T. \quad (3.10)$$

The probabilities $\mathbf{p}_t$ are structured to reflect anticipated market conditions by conditioning on time, market states, or a combination of both.

Sec.3.3 below introduces the concept of state signals, which are used in Sec.3.4 to create state-conditioned FPrs. Sec.3.4 also outlines alternative approaches for deriving forward-looking probabilities $\mathbf{p}$ through both time and state conditioning methods.

3.3 Market state signal preparation

To condition probabilities $\mathbf{p}$ on market states, we first derive a set of Q state signals $\{Z_{q,t} = (Z_{1,t}, \dots, Z_{Q,t})\}$ from a set of N raw state variables $\{S_{n,t} = (S_{1,t}, \dots, S_{N,t})\}$. State signals are designed to encode information about market conditions, *e.g.* given by financials, risk measures or macroeconomics. However, raw indicators like the Consumer Price Index (CPI) often contain noise that obscures the true signal. For this reason, raw data such as inflation is referred to as a state variable, while the transformed version that distils the essential information becomes the state signal. This transformation process, which aims to filter out noise and enhance the reliability of the signal, is detailed below along with the selection of the raw indicators used as state variables.

3.3.1 Market state variable selection

The approach considers a set of state variables $\{S_{n,t} \equiv (S_{1,t}, \dots, S_{N,t})\}$, which are recognised as key indicators for each asset class and sector in the SA market. Tab.3.1 provides a selection of SA-specific market indicators, along with certain indicators related to a broader market. Given SA's participation in the African Growth and Opportunity Act (AGOA) with the US, we include US indicators as additional measures of market movement [94]. Several of the indicators used are self-defined. For example, the Equity Risk Premium is approximated as the return in excess of the risk-free rate which one can earn if invested in equity between the two markets. Taking the differential between the US and SA we see how much more (or less) one would earn if invested in US equity versus SA.

Number	State Variables by Sector	Description	Bloomberg Ticker	Freq.
Industrials				
1	Domestic Leading Indicator (DLI)	SA Domestic Leading Indicator	OEZAKLAP Index	M.
2	SA PMI	SA Manufacturing Purchasing Managers' Index (PMI)	NAPMPMI Index	M.
Financials				
3	JPM EMB Index	JP Morgan Emerging Markets Bond Index (EMBI)	JPEIGLBL Index	M.
4	Market Momentum	The 12 month rolling return of the FTSE JSE All Share Index is used as proxy	JALSHTR Index	M.
Bonds				
5	SA Money Supply	M3 SA Money Supply (includes M2 money and larger institutional money)	SAMYM3Y Index	M.
Volatility Indices				
6	US Volatility Index	Measure of US stock market's expectation of volatility	VIX Index	M.
7	SA Volatility Index	SAVI Top 40 SA Volatility Index: forecast of Equity Market risk in SA	SAVIT40 Index	M.
Shared Indicators				
8	Equity Risk Premium Dif	USA risk premium less SA risk premium	(SPX - US003M) - (JALSHTR - JIBA3M)	M.
9	YoY change in SA CPI	SA Consumer Price Index	SACPIYOY Index	M.
10	US/SA Currency Exchange Rate	US Dollar to South African Rand exchange rate	USDZAR Curncy	M.
11	SA Economic Growth Rate	SA Real Gross Domestic Product	SAGDPANN Index	Q.

Table 3.1: A selection of 11 SA market state variables is categorised by industry. These indicators were chosen for use in the state-conditioned HS-FP models, as detailed in Sec.3.4.2 and Sec.3.4.4. While this is not an exhaustive list of all possible SA market indicators, it includes well-known indicators that are readily available. The indicators are listed along with their descriptions, Bloomberg tickers, and frequencies. Indicators such as “YoY change in SA CPI” are not specific to any particular sector and are therefore grouped under “Shared Indicators”.

In addition to asset and sector specific indicators we consider volatility indices both locally and for the US. One commonly used measure for assessing volatility in the US is the VIX, which reflects the volatility

of the S&P 500, a prominent US equity index. Similarly, the SA equivalent, the SAVI Top 40, is designed based on the VIX and serves as an indicator of fear and market sentiment in the local equity market [64]. For a more detailed discussion on selecting market state variables in a South African context, please see Sebastian and Gebbie [116]. Data beginning in 2000 was obtained from Bloomberg for all selected state variables. See appendix A5. [Data engineering](#) for details pertaining to the signal data preparation.

3.3.2 Market signal processing

All variables listed in Tab.3.1 were obtained with monthly frequencies, however, SAGDPANN - SA Real Gross Domestic Product (GDP) is only reported quarterly. Consequently, the market only knows what the estimated GDP is when it becomes public *i.e* at the reporting date. We therefore apply zero-order hold interpolation to convert the quarterly GDP series to monthly³. We lag GDP by 3 months to account for the delay in information. Likewise, both CPI and money supply was lagged by 1 month and the DLI by 6 months. When using JALSH as a proxy for market momentum we calculate the 12 month rolling returns.

The quality of the signals present within the state variables needed to be enhanced to better guide the scenario analysis to come in Sec.3.6. One of the simpler methods to decrease the impact of “false” signals is to smooth the data to reduce the degree of noise in the time series. Therefore, a double exponential decay is used to obtain the state signals $\{Z_{q,t} = (Z_{1,t}, \dots, Z_{Q,t})\}$, accounting for both a trend and a level component within the time series. The full algorithm should be referred to in [A1.1 Double exponential decay](#), however, here we provide a summarised description of the process:

Given a set of raw state variables $\{S_{n,t} \equiv (S_{1,t}, \dots, S_{N,t})\}$ the first round of exponential decay is performed using $\tau = \alpha$ to calculate the first set of FPrs $p_t^{\text{exp}}(\alpha)$ given a fast half-life α :

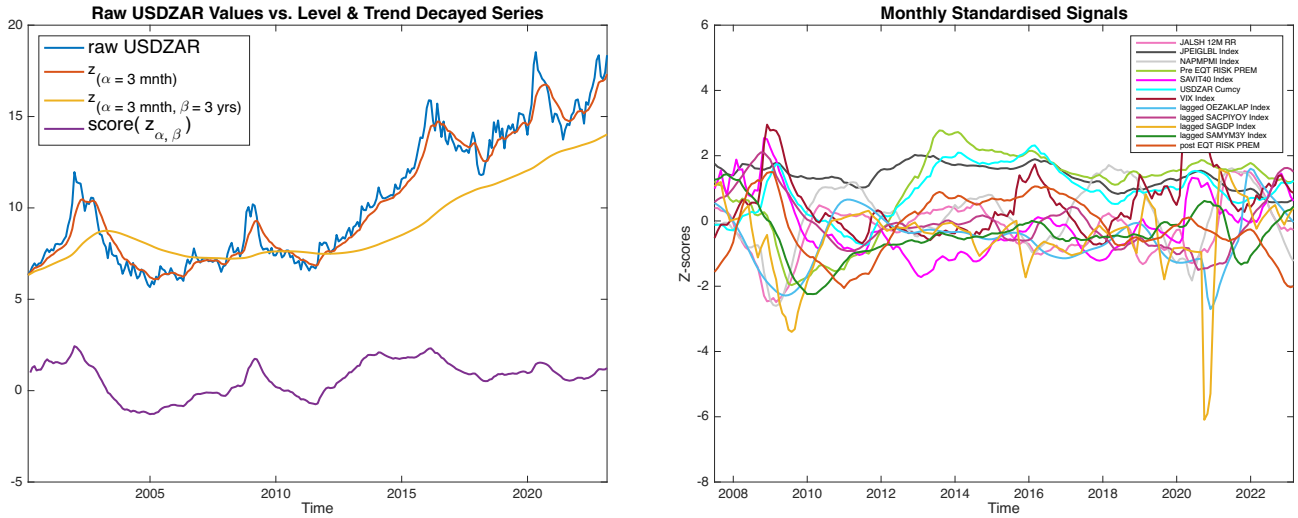
$$p_t^{\text{exp}}(\alpha) \propto e^{-\frac{\ln 2}{\alpha}|t-T|}. \quad (3.11)$$

By smoothing with a fast half-life of 3 months (as shown in Fig.3.5a) we capture the level component of the state variable, specifically targeting the short-term volatility of the series.

When $p_t^{\text{exp}}(\alpha)$ is applied to the n -th *state variable* $S_{n,t}$ using the weighted average operator in Apx. Eq.3, the result is the initial n -th *state signal* time series $\hat{\mu}_n(\alpha)$. A second set of FPrs, $p_t^{\text{exp}}(\beta)$, is calculated based on a slow half-life $\tau = \beta$, targeting the longer trend and correlation component. Applying this to $\hat{\mu}_n(\alpha)$ yields the double-decayed signal series $\hat{\mu}_n(\alpha, \beta)$.

Finally, $z_n(\alpha, \beta)$ is calculated according to Eq.10, which standardises $\hat{\mu}_n(\alpha)$ using the moments of $\hat{\mu}_n(\alpha, \beta)$. This allows for evaluating the significance of short-term variations within the context of the long-term trend. Fig.3.5a illustrates this process using the USD/ZAR time series, showcasing the smoothing effect on short-term noise while reducing the influence of the longer-term trend component.

³As appose to a cubic-spline approach as seen in [116].



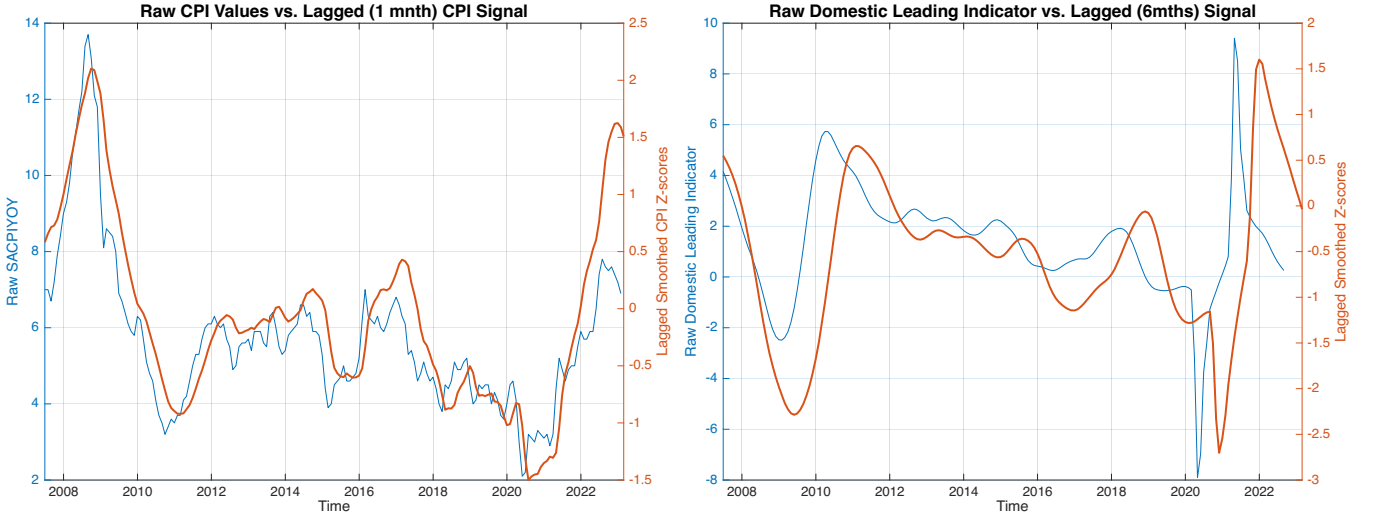
(a) The raw USDZAR currency exchange rate indicated in blue exhibits a fair amount of noise. z_{α} applies an exponential decay to the raw data using a fast half-life $\alpha = 3$ months. $z_{\alpha, \beta}$ follows on by applying an exponential decay to z_{α} using a slow half-life $\beta = 3$ years. The result is the double exponentially decayed series that is then standardised as prior description⁴.

(b) Z-scores are calculated for the state signals over the period from 30-06-2007 to 28-02-2023 for each selected state variable listed in Tab.3.1. These Z-scores will be used within the HS-FP framework to condition the FPrs.

Figure 3.5: By applying a double exponential decay to the state variables the effect of false signals is reduced. These standardised state signals are therefore more reliable indicators of market states which can be used to generate forward-looking estimates of $\hat{\mu}$ and $\hat{\Sigma}$ to be provided to the MV and HRP optimisations in the back-testing framework. Both figures above are generated by: `state_signal_prep_2.mlx`

When smoothing each variable using the above described method we preserve as much history as possible as the weighted average of the past values along with historical trends are used for the prediction of future time points. The more history available the better the smoothing. Once all variables are smoothed with their longest histories available, all smoothed signals are synchronised and cut at the same time points to ensure no missing data remains post smoothing. Fig.3.5b shows all state variables smoothed over the time period of end June'07 to end February'23 using a fast half-life of 3 months and a slow half-life of 3 years. It is important to note that no single pair of slow and fast half-lives work equally well for all time series in our state variable timetable as each series has varying types and degrees of trends. This will be discussed at a later stage. Below in Fig.3.6 we have the raw CPI and DLI data plotted on the left y-axis with the processed and lagged standardised z-score signals on the right. The effect of the double exponential decay smoothing as well as the 1 month and 6 month lag respectively was applied.

⁴Double Exponential Decay applied as discussed in Sec.3.3.



(a) The raw CPI is smoothed and then lagged by 1 month to account for delay of information to the market. The short term volatility has been reduced whilst still maintaining the important shifts and the overall trend. (b) The raw DLI is smoothed and lagged by 6 months to account for the fact that it is a leading indicator. As the raw variable displayed low short term volatility, the resulting state signal does not apply significant further smoothing.

Figure 3.6: The raw state variables, CPI and DLI, are plotted on the left y-axis of Fig.3.6a and Fig.3.6b respectively, with the processed z-score signals on the right axis. Here the impact of the double decay exponential smoothing was applied with $\alpha = 3$ months and $\beta = 3$ years⁵. Both figures are generated by: `state_signal_prep_2.mlx`.

With the raw state variables transformed into state signals, the data quality is enhanced, helping to map future uncertainties of market states onto historical invariants using state-conditioned FPrs. In the next section, both time and state FPrs are defined for the HS-FP framework.

3.4 Flexible probabilities methods

The HS-FP framework is broken down into four parts based on forms of probability conditioning: time-conditioned, state-conditioned, joint time and state, and multi-state time conditioning. By mapping the future randomness of the next periods' state variables (inflation, GDP, *etc.*) onto past time series of the invariants (as obtained in Sec.3.2.2) a forward-looking distribution of a portfolio's potential returns can be derived [84, 85, 89].

⁵As set out in Sec.3.3.2 and seen in Fig.3.5a.

3.4.1 Time conditioned

The most basic approach to obtain a more forward looking distribution of an invariant is to place emphasis on the latest data we have available, thereby conditioning the probabilities on time. This can be achieved using a naive approach of applying an arbitrary window length within the Rolling Window method or by using a slightly more sophisticated Exponential Smoothing process. Both methods are detailed below.

3.4.1.1 Rolling window

Definition 31 (Rolling window probabilities). If a scenario falls within a defined window length λ with an upper bound (usually selected as the latest observation T) and a lower bound of $T - \lambda$ we assign constant probability $\mathbf{p}_t$ to it. If it occurs outside the window it is not acknowledged, therefore assigned a probability of zero.

$$\mathbf{p}_t(T, \lambda) \equiv \mathbf{p}_t^{\text{roll}} \propto \begin{cases} 1 & \text{if } t > T - \lambda \\ 0 & \text{otherwise,} \end{cases} \quad (3.12)$$

where “ $\propto$ ” indicates that a proportional constant is needed to rescale $\mathbf{p}_t$ ensuring that $\sum_{t=1}^T \mathbf{p}_t = 1$.

A smoother and more responsive alternative for time conditioning is using exponential decay as defined below.

3.4.1.2 Exponential decay

Definition 32 (Exponential decay probabilities). To place high importance on more recent observations past values can be weighted less heavily with a lower probability, causing older data points to have diminishing influence over time:

$$\mathbf{p}_t(T, \tau) \equiv \mathbf{p}_t^{\text{exp}} \propto e^{-\frac{\ln 2}{\tau}|t-T|}, \quad (3.13)$$

here τ denotes the half-life of the exponential decay, representing the time it takes for the value of the time series to decrease by half, determining the rate at which past data points lose their influence over time. Note that $\tau > 0$ and the larger τ the more emphasis is placed on more recent observations.

Fig.3.7 compares the time conditioned FPrs where the shaded areas represent the magnitude of the probabilities applied to the observations at any given time point t . The rolling window example uses a window length $\lambda = 60$ months and the exponential decay has a half-life $\tau = 36$ months.

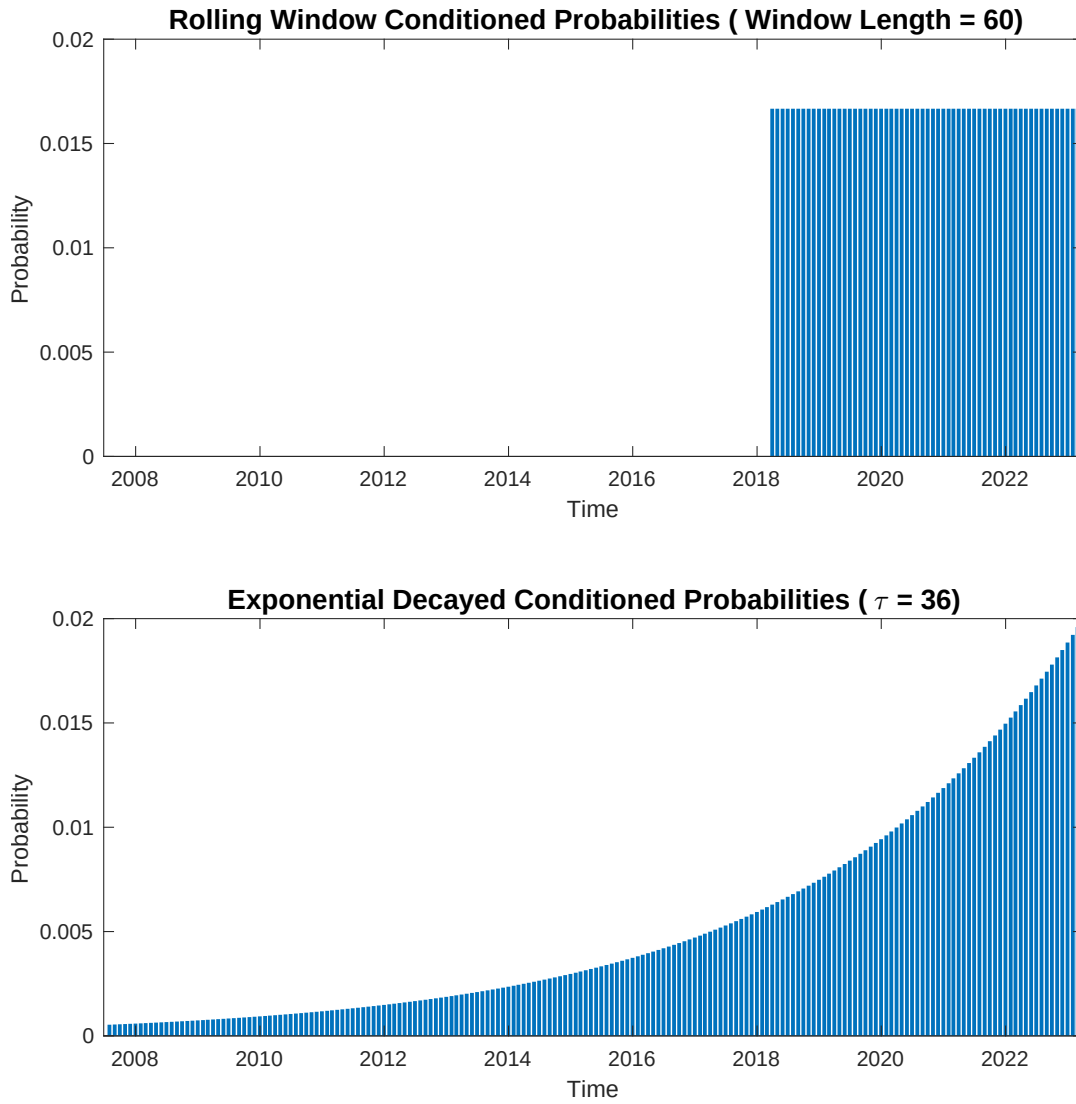


Figure 3.7: Shown in the top image is the time conditioned probabilities with a rolling window where $\lambda = 60$ months and in the bottom is the exponential decay with $\tau = 36$ months. The shaded areas represent the probability p_t at each time point t with the cumulative probabilities in each plot summing to 1. The probabilities follow the formulations as seen in sections 3.4.1.1 and 3.4.1.2 respectively. Generated by: `hsfp_tests.1.mlx`.

3.4.2 Single market state conditioned

State conditioning involves identifying and conditioning on key points within state signals (such as inflation or volatility indices) that correspond to market conditions of interest regardless of their position in the time sequence.⁶ This is the approach that allows for greater flexibility of scenario analysis. It allows historical data to be manipulated into a form that better represents future scenarios or specific abnormal market conditions for stress-testing portfolios. If we refer to single-state signals as $\mathbf{Z}_t$, we can condition probabilities based on the state of $\mathbf{Z}_t$ compared to a given target value z^* .

⁶If indicators need to be refined to remove the impact of “false” signals the indicators can be treated as a set of raw state variable $\{S_{n,t} \equiv (S_{1,t}, \dots, S_{N,t})\}$ which can undergo smoothing or signal processing to extract the state signals $\{Z_{q,t} \equiv (S_{1,t}, \dots, S_{Q,t})\}$. See Sec. 3.3.2 for an example.

3.4.2.1 Crisp probabilities

An abrupt approach to implementing state conditioning is to create crisp bands of probabilities either fully including an observation or excluding it. No smoothing is used, rather, every observation that falls within a band surrounding the target is equally weighted. The idea being that $\mathbf{z}_t$ will be “used” if it fits the profile of the market state of interest *e.g.* if $\mathbf{Z}_t$ were inflation and the portfolio needed to be tested in a high inflationary environment z^* would be specified as a high inflation value with a certain level of leeway α .

Definition 33 (Crisp conditioned probabilities). “Crisp” symmetrical bands are cut around a user specified target z^* , creating a target range $\mathcal{R}(z^*)$ based on a given range of probability α which acts as the leeway parameter forming $\mathcal{R}(z^*)$. If an element of $\mathbf{z}_t \in \mathcal{R}(z^*)$, then a probability of 1 is assigned, if not, zero. With this in mind the Crisp Flexible Probabilities can be formalised as^a:

$$\mathbf{p}_t(z^*) \equiv \mathbf{p}_t^{\text{crisp}} \propto \begin{cases} 1 & \text{if } \mathbf{z}_t \in \mathcal{R}(z^*) \\ 0 & \text{otherwise} \end{cases}. \quad (3.14)$$

Once again $\propto$ indicates the need for a proportionality constant such that the probabilities sum to 1. To ensure the band is symmetric around z^* we define the following:

$$\int_{\underline{z}}^{z^*} F_{\mathbf{Z}}(z)dz = \frac{\alpha}{2} = \int_{z^*}^{\bar{z}} F_{\mathbf{Z}}(z)dz. \quad (3.15)$$

With $F_{\mathbf{Z}}(z)$ as the cumulative distribution implied by $\mathbf{Z} \sim \{\mathbf{z}_t, \mathbf{p}_t\}_{t=1}^T$, additional conditions must be imposed on $\mathcal{R}(z^*)$. Establishing $\mathcal{R}(z^*)$ requires remaining within a feasible area such that it does not exceed the overall range of $\mathbf{Z}_t$. If $\mathcal{R}(z^*)$ extends into the lower or upper quantiles of $\mathbf{Z}$ based on α (*i.e.*, $Q_{\frac{\alpha}{2}}$ or $Q_{1-\frac{\alpha}{2}}$ of $\mathbf{Z}$), then $\underline{z} \leftarrow \min\{\mathbf{z}_t\}$ or $\bar{z} \leftarrow \max\{\mathbf{z}_t\}$, respectively.

^aThe notation of $\mathbf{p}_t(z^*)$ is taken to mean the flexible probabilities $\mathbf{p}_t$ given the target value of z^* .

Crisp conditioned FPRs are highly sensitive to the selected target value z^* as well as the choice of α . Its lack of smoothing results in harsh exclusions of observations resulting in very few time points receiving weightings, consequently compromising the method’s effectiveness.

3.4.2.2 Kernel smoothing

Instead of hard crisp band probabilities around a target range, the distance between observations and a target can be used to smooth the state conditioned probabilities. Kernel smoothing often used to obtain weighted averages is adapted in this case but still requires the selection of h , the kernel bandwidth (or radius) and kernel type γ . Meucci [87, 89] simplifies the selection of γ by limiting it to the set of $\gamma = 1$ resulting in an exponential kernel or $\gamma = 2$ for a Guassian kernel, but the choice of h still needs to be user defined⁷.

⁷See Scott [115]’s well cited (~ 9000 cites) “Multivariate Density Estimation: Theory, Practice, and Visualization” work for discussion on the selection of the tails of the kernel γ as well as the appropriate bandwidth h .

Definition 34 (Kernel conditioned probabilities). Probabilities are weighted based on the state variable $\mathbf{z}_t$ at a given t and its distance away from a specified target $\mathbf{z}^*$:

$$\mathbf{p}_t(\mathbf{z}^*, \gamma, h) \equiv \mathbf{p}_t^{\text{ker}} \propto e^{\frac{-|\mathbf{z}_t - \mathbf{z}^*|^\gamma}{h}}, \quad (3.16)$$

where $\gamma \in \{1, 2\}$, $\gamma = 1$ is an exponential kernel and $\gamma = 2$ is a Gaussian kernel. The larger the kernel bandwidth h , the wider the radius of the smoothing *e.g.* a large h in the case of $\gamma = 2$ uses a fatter tailed Gaussian and will result in smoother FPrs. A smaller bandwidth leads to a more localised weighting scheme.

3.4.3 Joint time & state conditioning

The goal of stress-testing a portfolio for market uncertainty is typically to apply it in the near future. Therefore, capturing the influence of time when generating scenarios is favourable for investors. A dual time and state conditioned approach should theoretically improve the precision of risk estimates when using the HS-FP framework by giving more weight to recent data and similar historical conditions, making models more responsive to current market conditions and relevant historical patterns.

Meucci [84] initially presented Entropy Pooling (EP) as a way to input investor non-linear market views into a non-normal market, generalising the way in which investor views can be incorporated in stress-testing. Using EP to do so allows the combining of investor views and the market model prior with minimal imposition of spurious structure. By minimising the relative entropy between the prior and posterior distributions, a resulting posterior can be obtained that is the least distorted from the prior—allowing for the integration of subjective views without undermining the historical data’s relevance. Minimising the relative entropy minimises the uncertainty within distributions that exist due to inherent randomness that is present in disorganised unpredictable systems such as markets.

3.4.3.1 Entropy pooling

Meucci [89] shows that EP can be used in a similar approach to combine a time conditioning exponential distribution with the crisp conditioning bands as the market views. By adjusting FPrs in alignment with investor views, EP can enhance the scenario analysis by capturing market conditions more accurately to support a more adaptive/responsive model. This is particularly valuable for assessing tail risks and stress-testing portfolios under varied market conditions.

Definition 35 (Entropy time and state conditioned probabilities). The prior is defined as the exponential decay distribution in Eq.3.13 to place more trust in recent events. The crisp probability distribution in Eq.3.14 sets the market views to be imposed on the market model prior. To obtain market views we define the crisp moments given a state conditioning target z^* :

$$\mu(z^*) = \sum_{t=1}^T z_t \mathbf{p}_t^{\text{crisp}}, \quad \sigma(z^*) = \left[\sum_{t=1}^T z_t^2 \mathbf{p}_t^{\text{crisp}} - \mu_{z^*}^2 \right]^{\frac{1}{2}}, \quad (3.17)$$

and set these views $\mathcal{V}(z^*)$ as constraints for the yet-to-be realised FPrs to meet when subjected to EP:

$$\mathcal{V}(z^*) : \begin{cases} \sum_{t=1}^T \mathbf{p}_t z_t = \mu(z^*) \\ \sum_{t=1}^T \mathbf{p}_t z_t^2 \leq \mu(z^*)^2 + \sigma(z^*)^2. \end{cases} \quad (3.18)$$

With the exponential decay (Eq.3.13) as the starting prior we apply EP to compute the FPrs that satisfy $\mathcal{V}(z^*)$:

$$\mathbf{p}(z^*) \equiv \underset{\mathbf{p} \in \mathcal{V}(z^*)}{\text{argmin}} \mathcal{E}(\mathbf{p}; \mathbf{p}^{\text{exp}}), \quad (3.19)$$

where $\mathbf{p}(z^*)$ is the relative entropy (or the Kullback and Leibler [57] divergence) being minimised in its discrete form:

$$\mathcal{E}(\mathbf{p}; \mathbf{p}^{\text{exp}}) = \sum_{t=1}^T \mathbf{p}_t \ln \left(\frac{\mathbf{p}_t}{\mathbf{p}_t^{\text{exp}}} \right). \quad (3.20)$$

The output distribution $\mathbf{p}(z^*)$ serves as a posterior distribution integrating all provided inputs into a joint time and state conditioned posterior.

The EP probabilities with these inputs yield a mixture of the exponential decay prior and crisp moment views, resembling the kernel method in Eq.3.14. This method allows the kernel estimator to flexibly switch between the exponential and Gaussian kernels, optimising the centre and bandwidth. The crisp views distort the exponential prior, but by using the EP approach, the posterior distribution minimises the entropy relative to the exponential prior, ensuring minimal imposition of spurious structure. This results in a distribution that incorporates desired views while maintaining the integrity of the prior with the least unnecessary deviation.

In Fig.3.8 an example comparison of the state conditioning methods is shown. The top panel shows a standardised smoothed monthly CPI state signal⁸ on which the state-conditioned probabilities are conditioned. Crisp bands (blue - -) indicating the target region around the user-defined z^* (red - -) are depicted.

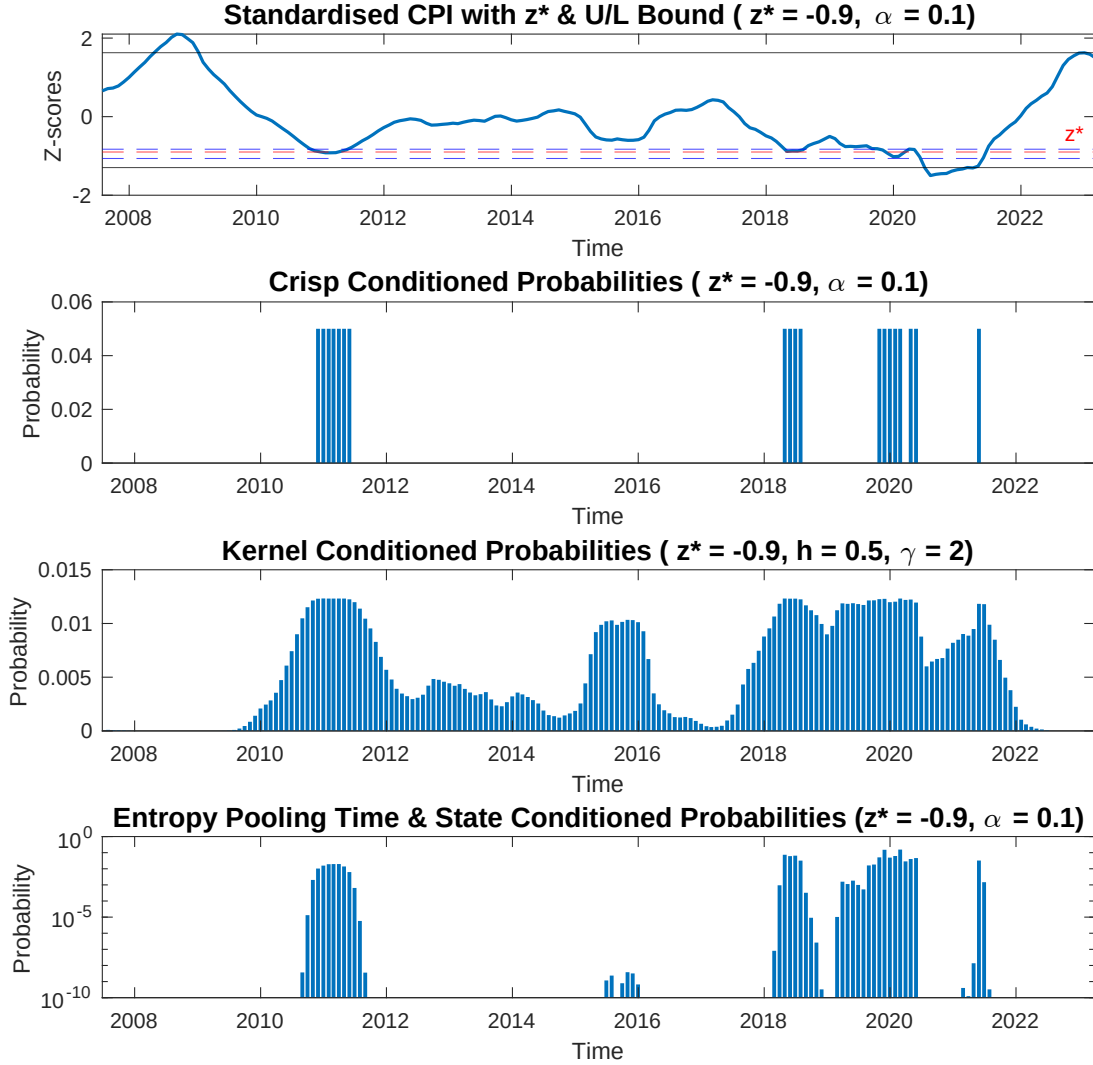


Figure 3.8: To show a test case for time & single state conditioned probabilities we use monthly CPI as a market state signal and apply the [crisp](#), [kernel](#) and [entropy pooling](#) HS-FP framework. In panel 1 (top) is the smoothed standardised time series⁴ with a target value $z^* = -0.9$ (red - -). The shaded areas in panels 2-4 represent the FPrs (weights) at any given time that will later be applied to asset returns within the HS-FP framework⁹. The crisp probabilities in panel 2 uses a range of probability $\alpha = 0.1$ resulting in an $\bar{z}$ and $\underline{z}$ bounds of $[-0.83; -1.07]$ (blue - -). In panel 3 the smoother kernel probabilities are seen using the same z^* , $h = 0.5$ and $\gamma = 2$ (defining the use of a Gaussian kernel). Lastly in panel 4 the time conditioned exponential decay probabilities (Fig.3.7) are combined with the crisp probabilities in panel 2 to condition by both time and state using entropy pooling. Generated by: `hsfp_tests.1.mlx`.

⁸State variable smoothing and standardisation is discussed in detail in Sec.4.

⁹Probabilities are normalised across each panel *i.e* they sum to 1.

3.4.4 Multi-market state & time conditioned

In practice, market behaviour is never driven by one single factor at any given time, therefore, we introduce the ensemble methods to incorporate multiple states and time conditioning.

3.4.4.1 Equally weighted ensemble

Definition 36 (Equally weighted ensemble probabilities). Given Q conditioning signals^a of interest $\mathbf{Z}_{1,t}, \dots, \mathbf{Z}_{Q,t}$ and we compute their EP (Eq.3.19) time- and state-conditioned FPrs as:

$$\mathbf{p}_q \equiv \mathbf{p}(\mathbf{z}_q^*), \quad q = 1, \dots, Q, \quad (3.21)$$

where the collection of $\mathbf{p}_q$ is a function of q target values $\mathbf{z}^*$. To create ensembles we use linear mixing to pool all the information in $\mathbf{p}_q$, weighting sets of FPrs by ω_q :

$$\mathbf{p}_{\text{ensemble}} \equiv w_1 \mathbf{p}_1 + \dots + w_Q \mathbf{p}_Q. \quad (3.22)$$

To obtain the Equally Weighted ensemble each q is weighed by $\omega_q = \frac{1}{Q}$.

^a $Q \subseteq N$, where N is the original number of signal variables. See Sect.3.3.

The time and state conditioned FPrs in Eq.3.21 are conditioned on an individual basis therefore not all estimates are equally trustworthy due to different degrees of conditioning as well as possible correlations that may exist between sets of FPrs. An alternative form of an ensemble can be created by unequally weighting different sets of FPrs using Bayesian principles.

3.4.4.2 Conditional Bayesian ensemble

For the Conditioned Bayesian Ensemble (CBE)¹⁰ we use the Bhattacharyya coefficient as a measure of how “close” two probability distributions ($\mathbf{p}_q, \mathbf{p}_r$) are, effectively measuring the amount of overlap between two samples:

$$b_{q,r} \equiv \sum_{t=1}^T (p_{t,q} p_{t,r})^{\frac{1}{2}} \quad (3.23)$$

The coefficient ranges from 0 to 1 where a value of 0 indicates no overlap or similarity between the distributions, while a value of 1 indicates complete overlap or similarity [15, 42]. These coefficients are calculated pairwise across signals and used as an input to the Hellinger distance which allows for the quantification of the similarity of the probability distributions [122]:

$$d_{q,r} \equiv \sqrt{1 - b_{q,r}} \quad (3.24)$$

¹⁰In [116] this is also referred to as Degree of Conditioning and Correlation (DCC).

A diversification indicator $\mathcal{D}_q$ can be calculated as a measure of individual diversity of a set of FPrs $\{\mathbf{p}_q\}$ in comparison to every other using the Unweighted Pair Group Method with Arithmetic Mean (UPGMA) type hierarchical clustering:

$$\mathcal{D}_q \equiv \frac{1}{Q-1} \sum_{r \neq q} d_{q,r} \quad (3.25)$$

Combining the UPGMA measure with a given signal's effective number of scenarios (ENS)¹¹ as defined by Meucci [88]: $\mathcal{T} = e^{-\sum_{t=1}^T \mathbf{p}_t \ln \mathbf{p}_t}$, we can compute the CBE weights, if a particular state signal shows low correlation with other signals or imposes a lesser degree of conditioning on numerous historical scenarios indicated by the ENS then it will be assigned a greater weight:

$$\omega_q \equiv \frac{\mathcal{T}_q \mathcal{D}_q}{\sum_{r=1}^Q \mathcal{T}_r \mathcal{D}_r} \quad (3.26)$$

If we consider the sets of FPrs in the ensemble $\mathbf{p}_q$ in Eq.3.21 as scenarios and their corresponding weights ω_q as defined above, we can represent the distribution $f_{\mathbf{P}}$ of the T -dimensional random variable $\mathbf{P}$ as the Bayesian ensemble posterior:

$$f_{\mathbf{P}} \sim \{\mathbf{p}_q, \omega_q\}_{q=1}^Q. \quad (3.27)$$

Using the same linear mix approach defined in Eq.3.22 the resulting CBE FPrs are given by:

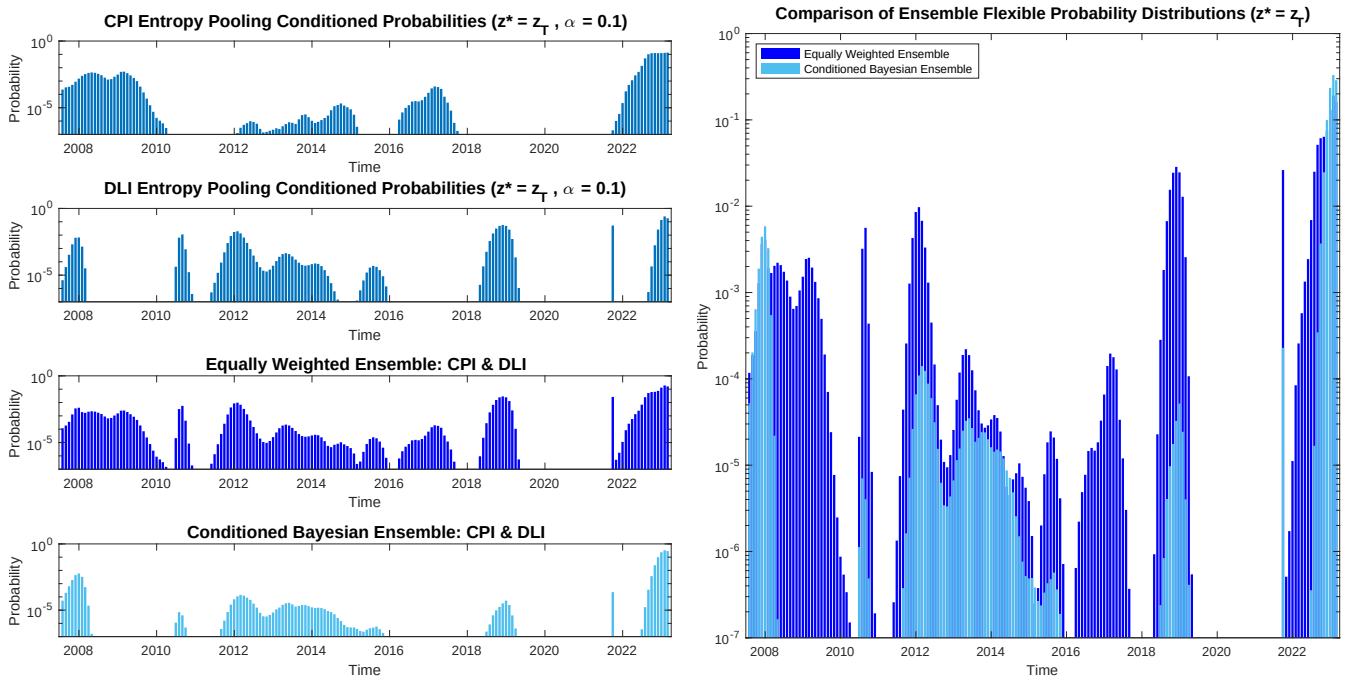
$$\mathbf{p}_{\text{CBE}} \equiv w_1 \mathbf{p}_1 + \cdots + w_Q \mathbf{p}_Q, \quad (3.28)$$

where w_q is as defined in Eq.3.26 for $q = 1, \dots, Q$ and can be interpreted as probability weightings of probabilities. Alternatively a log-linear weighted average can also be used as it aligns more closely with the principles of Bayesian combination [89, 91]:

$$\ln(\mathbf{p}_{\text{CBE}}) \stackrel{1}{\leftarrow} w_1 \ln(\mathbf{p}_1) + \cdots + w_Q \ln(\mathbf{p}_Q), \quad (3.29)$$

here $\stackrel{1}{\leftarrow}$ indicates that $\mathbf{p}_{\text{CBE}}$ is rescaled to sum to 1.

¹¹The ENS assesses the statistical power of a stress-test when using the HS-FP framework. If all $\mathbf{p}_t$'s are equal then $\mathcal{T} = T$, *i.e.* all time points (scenarios) are effectively being used, however, if only a single time point t is used then $\mathcal{T} = 1$ as all other FPrs equal zero.



(a) Panel 1 & 2 displays the entropy pooled FPRs using CPI and the DLI signals respectively with $z_{\text{CPI}}^* = 1.51$ and $z_{\text{DLI}}^* = -0.04$ (both the latest value of the signals). Panel 3 shows the ensemble of the two signals when using an equally weighted approach (Sec 3.4.4.1) and the 4th panel is the Conditioned Bayesian ensemble FPRs (Sec 3.4.4.2). Note the single spike in late 2021 is due to the DLI being relatively close in value to z_{DLI}^* .

(b) The comparison of the two ensemble approaches is seen here for CPI and the DLI. In the Equally Weighted ensemble each q is weighed by $\omega_q = 1/Q$. Whereas in the Conditioned Bayesian ensemble the ENS ($\mathcal{T}_q$) and the correlation of the signals captured by $\mathcal{D}_q$ are accounted for with $\omega_q \equiv \frac{\mathcal{T}_q \mathcal{D}_q}{\sum_{r=1}^Q \mathcal{T}_r \mathcal{D}_r}$ where r represents all remaining set of Flexible Probabilities such that $r \neq q$. We see that both posterior distributions are relatively similar when $z_1^*, \dots, z_Q^*$ is set $z^* = z_{Q,T}$ of $q = 1, \dots, Q$.

Figure 3.9: Flexible probabilities assigned to historical observations under alternative signal specifications and ensemble constructions. The figure illustrates how different approaches alter the weighting of past data points, emphasising scenarios that resemble the market conditions of interest. The panels illustrate how different methods reweight history, showing which periods are leaned on more heavily when constructing the desired distribution. Both figures are generated by: `hsfp_tests_1.mlx`.

3.5 Modelling forward-looking estimates of risk drivers

By combining the concepts of invariants and flexible probability theory discussed in sections 3.2 and 3.4, we can model the forward-looking distribution f_ϵ of the invariants using non-uniform probabilities $\{p_t\}$. In a standard empirical distribution, equal probabilities are assigned to all historical observations, implying that each observation is equally likely or carries equal importance (a concern we discussed in Chp.1 and briefly in Sec.1.2.2.2). However, to incorporate varying likelihoods and relevance of observations for a forward-looking approach, we adopt a generalised empirical distribution. This allows us to assign different weights to historical data using the HS-FP forward-looking probabilities, better reflecting their significance in predicting future outcomes and capturing market regimes. The generalised empirical distribution assumes that the observations are *i.i.d.*, a condition satisfied by the invariants. Under this framework, f_ϵ can be formally defined as a PDF:

$$f_\epsilon(\mathbf{x}) \equiv \sum_{t=1}^T p_t \delta(\mathbf{x} - \epsilon_t). \quad (3.30)$$

The term δ denotes the Dirac delta distribution [83], a generalised function that concentrates all proba-

bility mass at a single point. It satisfies what is known as the sifting property:

$$\int_{\mathbb{R}^n} g(\mathbf{x}) \delta(\mathbf{x} - \boldsymbol{\varepsilon}_t) d\mathbf{x} = g(\boldsymbol{\varepsilon}_t), \quad (3.31)$$

for any test function g . In this context, the Dirac delta is used to represent a discrete distribution where the probability mass p_t is assigned to the scenario $\boldsymbol{\varepsilon}_t$. Eq.3.30 therefore represents a multivariate histogram with its peaks centred around the observations $\boldsymbol{\varepsilon}_t$.

By considering returns as functions of the risk drivers, we can compute their statistical properties directly using the properties of the Dirac delta function. The distribution of returns $\mathbf{R}$ can be expressed as:

$$f_{\mathbf{R}}(\mathbf{r}) = \sum_{t=1}^T p_t \delta(\mathbf{r} - \mathbf{r}_t), \quad (3.32)$$

where $\mathbf{r}_t = \mathbf{R}_t$ are the historical asset returns at time t . This formulation allows us to compute statistics such as the mean and covariance directly from the weighted return scenarios.

Lemma 1 (HS-FP Expected Returns). *Let $\{\mathbf{r}_t\}_{t=1}^T$ denote a sequence of observed historical asset returns, and let $\{p_t\}_{t=1}^T$ be the corresponding set of HS-FP flexible probabilities, satisfying $\sum_{t=1}^T p_t = 1$ and $p_t \geq 0$ for all t . Then, the expected return vector under the HS-FP measure is given by:*

$$\mathbb{E}[\mathbf{R}] = \sum_{t=1}^T p_t \mathbf{r}_t. \quad (3.33)$$

Proof. The probability distribution of $\mathbf{R}$ under the HS-FP measure is given by:

$$f_{\mathbf{R}}(\mathbf{r}) = \sum_{t=1}^T p_t \delta(\mathbf{r} - \mathbf{r}_t). \quad (3.34)$$

The expectation of $\mathbf{R}$ is then:

$$\mathbb{E}[\mathbf{R}] = \int_{\mathbb{R}^n} \mathbf{r} f_{\mathbf{R}}(\mathbf{r}) d\mathbf{r}. \quad (3.35)$$

Substituting in the expression for $f_{\mathbf{R}}$, we obtain:

$$\mathbb{E}[\mathbf{R}] = \sum_{t=1}^T p_t \int_{\mathbb{R}^n} \mathbf{r} \delta(\mathbf{r} - \mathbf{r}_t) d\mathbf{r}. \quad (3.36)$$

We now apply the *sifting property* of the Dirac delta function stated in Eq.3.31, which states that for any test function $g(\mathbf{r})$,

$$\int_{\mathbb{R}^n} g(\mathbf{r}) \delta(\mathbf{r} - \mathbf{r}_t) d\mathbf{r} = g(\mathbf{r}_t). \quad (3.37)$$

In our case, the function being integrated is simply $g(\mathbf{r}) = \mathbf{r}$, and so the integral evaluates to:

$$\int_{\mathbb{R}^n} \mathbf{r} \delta(\mathbf{r} - \mathbf{r}_t) d\mathbf{r} = \mathbf{r}_t. \quad (3.38)$$

Intuitively, the delta function spikes at the point $\mathbf{r} = \mathbf{r}_t$, and is zero elsewhere. Since it integrates to 1, it effectively “picks out” the value of the integrand at $\mathbf{r}_t$.

Substituting this back, we obtain:

$$\mathbb{E}[\mathbf{R}] = \sum_{t=1}^T p_t \mathbf{r}_t, \quad (3.39)$$

as required. □

Remark 2. Lemma 1 shows that the HS-FP expected return is a probability-weighted average of historical returns. In the special case where $p_t = \frac{1}{T}$ for all t , this expression reduces to the standard empirical mean. The HS-FP framework thus generalises the empirical mean by allowing for flexible, non-uniform weights that reflect time- or state-dependent probabilities.

Similarly, the covariance matrix of the returns is calculated as:

Lemma 2 (HS-FP Covariance). *Let $\{\mathbf{r}_t\}_{t=1}^T$ denote the sequence of observed historical asset returns, and let $\{p_t\}_{t=1}^T$ be the associated HS-FP flexible probabilities satisfying $\sum_{t=1}^T p_t = 1$ and $p_t \geq 0$ for all t . Then, the covariance matrix of $\mathbf{R}$ under the HS-FP measure is given by:*

$$\begin{aligned} \text{Cov}[\mathbf{R}] &= \mathbb{E}[\mathbf{R}\mathbf{R}'] - \mathbb{E}[\mathbf{R}]\mathbb{E}[\mathbf{R}]' \\ &= \sum_{t=1}^T p_t \mathbf{r}_t \mathbf{r}_t' - \left(\sum_{t=1}^T p_t \mathbf{r}_t \right) \left(\sum_{t=1}^T p_t \mathbf{r}_t \right)', \end{aligned} \quad (3.40)$$

where $\mathbb{E}[\mathbf{R}]$ is the expected return vector as defined in Eq. (3.33).

Proof. By definition, the covariance matrix of $\mathbf{R}$ is

$$\text{Cov}[\mathbf{R}] = \mathbb{E}[\mathbf{R}\mathbf{R}'] - \mathbb{E}[\mathbf{R}]\mathbb{E}[\mathbf{R}]'. \quad (3.41)$$

The first term, the second moment matrix, can be expressed as

$$\mathbb{E}[\mathbf{R}\mathbf{R}'] = \int_{\mathbb{R}^n} \mathbf{r} \mathbf{r}' f_{\mathbf{R}}(\mathbf{r}) d\mathbf{r}. \quad (3.42)$$

Substituting the HS-FP probability distribution defined in Eq. 3.32 results in

$$\mathbb{E}[\mathbf{R}\mathbf{R}'] = \sum_{t=1}^T p_t \int_{\mathbb{R}^n} \mathbf{r} \mathbf{r}' \delta(\mathbf{r} - \mathbf{r}_t) d\mathbf{r}. \quad (3.43)$$

We then apply the sifting property of the Dirac delta (Eq. 3.31),

$$\int_{\mathbb{R}^n} \mathbf{r} \mathbf{r}' \delta(\mathbf{r} - \mathbf{r}_t) d\mathbf{r} = \mathbf{r}_t \mathbf{r}_t'. \quad (3.44)$$

Thus,

$$\mathbb{E}[\mathbf{R}\mathbf{R}'] = \sum_{t=1}^T p_t \mathbf{r}_t \mathbf{r}_t'. \quad (3.45)$$

Recall from Lemma 1 that

$$\mathbb{E}[\mathbf{R}] = \sum_{t=1}^T p_t \mathbf{r}_t. \quad (3.46)$$

Substituting these expressions into the covariance definition, we obtain

$$\text{Cov}[\mathbf{R}] = \sum_{t=1}^T p_t \mathbf{r}_t \mathbf{r}_t' - \left(\sum_{t=1}^T p_t \mathbf{r}_t \right) \left(\sum_{t=1}^T p_t \mathbf{r}_t \right)', \quad (3.47)$$

which completes the proof. $\square$

Remark 3. This result generalises the classical empirical covariance matrix by incorporating flexible, non-uniform probabilities p_t that reflect the relative importance of each historical return vector. When $p_t = \frac{1}{T}$ for all t , Eq.3.40 reduces to the conventional empirical covariance matrix.

By using FPrs and the Dirac delta function to model forward-looking distributions of returns and reflect the current market regime, we avoid relying on parametric assumptions or specific distributional forms. Historical time series data can now be weighted according to time- or state-conditioned FPrs to reflect future market behaviour, enabling portfolio allocation strategies that are more responsive to current conditions and potentially improve risk-adjusted performance by capturing regime changes. The HS-FP framework refines the estimation of expected returns and covariances, allowing MV and HRP methods to better capture market dynamics and adjust portfolio allocations based on probabilistic insights.

With all components of the HS-FP framework established, the focus turns to its implementation within a back-testing. The following sections detail how the HS-FP methods are incorporated into the back-test and how they influence the performance and decision-making processes.

3.6 Application

In Chapter 2, we implemented a conventional back-test with eight simulation configurations (Tab.2.3) with varying levels of Winsorisation, Regularisation, and cash constraints. Config.5 (Fig.2.11 and Tab.2.5), featuring winsorised returns and a 5% cash equality constraint for both HRP and MV portfolios, was selected as the base model. This chapter builds on that base model by developing a stress-testing framework designed to better evaluate portfolios under abnormal market conditions.

Recent applications of MPT often interpret the Markowitz [69] MV framework as reliant on historical estimates of μ and Σ . However, Markowitz [69] emphasises using a sound cost function, such as the SR, without prescribing a specific data generation process. Leveraging this, we can use the HS-FP framework to replace the historical estimates with forward-looking μ and Σ , enabling the MV + HS-FP portfolio to align with future market realisations. Similarly, the HRP + HS-FP portfolio incorporates forward-looking Σ estimates.

Given that there are several forms of the HS-FP models, *e.g.*, time conditioning, state conditioning, joint time and state conditioning, as well as multi-state conditioning, we chose to limit our application by excluding the multi-state methods. We number the HS-FP methods 1 to 5, which will be used for comparison in the results section below:

1. Time-Conditioned Rolling Window
2. Time-Conditioned Exponential Decay

3. State-Conditioned Crisp Bands (using signals in Tab.3.1)
4. State-Conditioned Kernel Smoothing (using signals in Tab.3.1)
5. Joint-Time & State Conditioned Entropy Pooling (using signals in Tab.3.1)

To generate a scenario-based back-test using the HS-FP models we follow the flow in Fig.3.1¹². Tab.3.2 below lists all parameters that need to be defined through the back-test framework. Each HS-FP method requires its own combination of parameters and/or state signals. We have used our domain knowledge to select default values for each parameter. However, in our closing thoughts on potential future work, we note that it would be best to move past this exploratory data analysis (EDA) approach by conducting a sensitivity analysis where we would assess how varying parameters and combinations of state conditioning signals affect the outcome performance metrics across different HS-FP methods.

Table 3.2: A summary of all parameters, their default values, and the corresponding discussion sections required for the HS-FP back-test framework is provided below. Parameters listed under headings 1–3 below are discussed in Chp.2, focusing on the conventional back-test model. Parameters under headings 4 and 5 pertain specifically to this chapter. The default values are used for the simulations presented in the final results in Sec.3.7.

HS-FP Back-test Parameter Definitions					
Notation	Parameter	Default Values	Process	Section	Figures & Tables
1. Data Processing					
s Winsor	Winsorise Threshold	2 sd.	Winsorisation	2.4.2.3	Fig.2.6, Fig.2.7, 2.8 & Tab.2.4
λ_{Reg}	Regularisation	0.00	Covariance Regularisation	2.4.2.3	Fig.2.6, Fig.2.7, 2.8 & Tab.2.4
2. Portfolio Construction					
ω_{MV}	MV Weight constraints	5% Cash Constraint	Portfolio Optimisations	2.4.2.2	Fig.2.7, 2.8 & Tab.2.4
ω_{HRP}	HRP Weight constraints	5% Cash Constraint	Portfolio Optimisations	2.4.2.2	Fig.2.7, 2.8 & Tab.2.4
3. Back-test Setup					
T	Window length of BT	3 years	Back-test loop	2.3.4	-
Rf_T	Risk-free rate over window	μ_T of Cash Asset	Back-test loop	2.3.4	-
4. Signal Data Preparation					
$\tau = \alpha$	Short half-life	3 months	Double Exponential Decay	3.3, 4	Fig.3.3.2 & Fig.3.6
$\tau = \beta$	Slow half-life	3 years	Double Exponential Decay	3.3, 4	Fig.3.3.2 & Fig.3.6
	Potentially the lags of signals	<i>e.g.</i> 1 mnth CPI	Temporal Signal Adjustment	3.3, 4	Fig.3.3.2 & Fig.3.6
5. Historical Simulation with Flexible Probabilities Framework					
λ	Rolling Window Length	12 mnths	Time-Conditioning HS-FP	3.4.1.1	Fig.3.7 & Tab.3.10 - Tab.3.12
τ	Half life of Exponential Decay	12 mnths	Time-Conditioning HS-FP	3.4.1.2	Fig.3.7 & Tab.3.10 - Tab.3.12
z^*	Crisp Target Value	z_T	State-Conditioning HS-FP	3.4.2.1	Fig.3.8 & Tab.3.10 - Tab.3.12
α	Crisp Range of Probability	0.1	State Conditioning HS-FP	3.4.2.1	Fig.3.8 & Tab.3.10 - Tab.3.12
h	Kernel bandwidth	0.5	State-Conditioning HS-FP	3.4.2.2	Fig.3.8 & Tab.3.10 - Tab.3.12
γ	Kernel type	2 (Gaussian)	State Conditioning HS-FP	3.4.2.2	Fig.3.8 & Tab.3.10 - Tab.3.12

Building on the configurations introduced in Chp.2, we extend our analysis by running configurations 9–43, which incorporate various permutations of the HS-FP back-tests. To evaluate the results, we focus on three key comparisons:

1. Sec.3.7.1: Comparison of the base model and the HS-FP methods (Fig.3.10). This aims to:
 - (a) Evaluate if incorporating forward-looking estimates of μ and Σ using HS-FP methods improves stress-testing compared to the conventional base model which relies solely on historical data.
 - (b) Determine whether any HS-FP method consistently outperforms the others and identify if specific methods are better suited for the MV or HRP portfolios.
2. Sec.3.7.2: Comparison of MV and HRP performance across all methods (Fig.3.11). This evaluates: Whether HS-FP scenario-based stress-testing can effectively capture regime changes, reduce volatil-

¹²The structure and code flow diagram can be followed in the appendix Fig.6. The code diagram resembles that of the logic illustrated in Fig.3.1.

ity, and improve the risk-adjusted returns of the MV portfolio, giving it an edge over the diversification-focused HRP portfolio.

3. Sec.3.7.3: Comparison of signals within the state-conditioned HS-FP models (Fig.3.12). This assess: Whether specific signals consistently perform well, serving as reliable market state indicators, or consistently underperform, revealing potential weaknesses.

In the upcoming sections and figures, we analyse each comparison, using Excel conditional formatting to highlight the relative performance of the configurations. Although the same set of results—comprising the base model and all permutations of the five HS-FP methods—is used throughout, the conditional formatting will be adjusted to emphasise the specific focus of each comparison. The metrics provided are the annualised returns, annualised volatility, the SR, annualised SR and the PSR¹³.

3.7 Results

Each subsection will describe how the comparative conditional formatting was applied, followed by a corresponding figure. A discussion of the results for the respective comparison will then conclude the subsection. The conditional formatting scheme applied to the results uses a colour gradient to clearly indicate the performance of each outcome. The most favourable outcomes are shaded with the darkest green, while the least favourable outcomes are shaded with the darkest red. We define favourable results as those with the highest values of each metric, except for annualised volatility ($\text{Ann.}\sigma$), where lower values are more preferable. All other results are displayed along the red-to-green gradient, reflecting their relative favourability compared to other outcomes in the comparison.

3.7.1 Base model and HS-FP methods comparison

In Fig.3.10 below, we apply the conditional formatting down the columns to compare each metric across methods. Using the annualised SR (Ann.SR) of the MV optimisations as an example: the most favourable annual SR is 1.82 from Config.24 within “Method 4: State-Conditioned Kernel Smoothing” conditioned on the JPM emerging market index and is therefore shaded darkest green. The least favourable annualised SR for the MV optimisation is generated by Config.41 and is -0.03 from “Method 5: Joint Time & State Conditioned Entropy Pooling” conditioned on YoY change in CPI and shaded darkest red.

The same principle is applied to the HRP optimisation: the most favourable and darkest green value for the Ann.SR is 2.92 from Config.19 within “Method 3: State-Conditioned Crisp Bands” conditioned on the YoY change in CPI. The least favourable annualised SR of 1.21 is generated by Config.20 also in “Method 3: State-Conditioned Crisp Bands” conditioned on the USD/ZAR exchange rate and is consequently shaded in the darkest red for that column. Working top-down, we will discuss each method’s results in comparison with others.

¹³We had hoped to calculate the DSR for all simulations, however, due to time constraints and computational challenges we were not able to do so.

1. Method Comparison

Base Model											
		MV					HRP				
Config		Ann.R(%)	Ann. σ (%)	SR	Ann.SR	PSR	Ann.R(%)	Ann. σ (%)	SR	Ann.SR	PSR
5		10.876	9.130	0.147	1.801	0.964	10.198	6.687	0.174	2.130	0.982

HS-FP Method 1: Time-Conditioned Rolling Window

		MV + HS-FP					HRP + HS-FP				
Config	Static parameter value	Ann.R(%)	Ann. σ (%)	SR	Ann.SR	PSR	Ann.R(%)	Ann. σ (%)	SR	Ann.SR	PSR
9	$\lambda = 12$ months	9.165	10.998	0.081	0.988	0.844	10.436	6.796	0.181	2.212	0.985

HS-FP Method 2: Time-Conditioned Exponential Decay

Config		MV + HS-FP					HRP + HS-FP				
	Static parameter value	Ann.R(%)	Ann. σ (%)	SR	Ann.SR	PSR	Ann.R(%)	Ann. σ (%)	SR	Ann.SR	PSR
10	$\tau = 12$ months	10.848	10.206	0.131	1.601	0.947	10.279	6.732	0.176	2.155	0.983

HS-FP Method 3: State-Conditioned Crisp Bands

Config	Sector Signal	MV + HS-FP					HRP + HS-FP				
	Industrials	Ann.R(%)	Ann. σ (%)	SR	Ann.SR	PSR	Ann.R(%)	Ann. σ (%)	SR	Ann.SR	PSR
11	SA DLI	8.958	11.526	0.073	0.884	0.819	9.462	7.838	0.124	1.506	0.935
12	SA PMI	7.061	13.020	0.025	0.304	0.621	10.132	7.301	0.157	1.922	0.971
	Financials										
13	JPM EMB Index	10.970	11.748	0.116	1.427	0.920	10.450	7.495	0.165	2.013	0.979
14	Market Momentum	11.575	12.803	0.119	1.463	0.930	9.914	8.191	0.133	1.626	0.946
	Bonds										
15	SA Money Supply	11.325	11.379	0.129	1.578	0.945	10.004	7.806	0.143	1.749	0.960
	Volatility Indices										
16	US Volatility Index	8.804	11.775	0.067	0.818	0.796	10.145	8.385	0.137	1.678	0.958
17	SA Volatility Index	8.035	11.267	0.052	0.632	0.736	11.262	8.071	0.180	2.206	0.986
	(Partially) Shared Indicators										
18	Equity Risk Premium Dif	9.315	10.771	0.087	1.054	0.856	9.815	7.771	0.137	1.668	0.951
19	YoY change in SA CPI	8.908	12.935	0.064	0.772	0.781	12.651	7.638	0.236	2.920	0.998
20	US/SA Currency Exchange Rate	11.674	10.754	0.145	1.778	0.973	8.758	7.872	0.100	1.209	0.890
21	SA Economic Growth Rate	10.343	11.070	0.109	1.328	0.907	8.922	7.590	0.109	1.322	0.911

HS-FP Method 4: State-Conditioned Kernel Smoothing

Config	Sector Signal	MV + HS-FP					HRP + HS-FP				
	Industrials	Ann.R(%)	Ann. σ (%)	SR	Ann.SR	PSR	Ann. R(%)	Ann. σ (%)	SR	Ann.SR	PSR
22	SA DLI	8.792	10.989	0.072	0.877	0.810	10.518	7.062	0.177	2.164	0.985
23	SA PMI	10.716	11.064	0.118	1.444	0.922	10.626	6.589	0.194	2.376	0.990
	Financials										
24	JPM EMB Index	11.505	10.198	0.148	1.818	0.966	10.237	6.710	0.175	2.141	0.983
25	Market Momentum	9.842	10.151	0.106	1.290	0.900	9.904	6.875	0.158	1.933	0.972
	Bonds										
26	SA Money Supply	10.854	9.229	0.145	1.776	0.962	10.408	6.706	0.182	2.227	0.986
	Volatility Indices										
27	US Volatility Index	11.361	10.493	0.140	1.716	0.956	10.297	6.736	0.177	2.164	0.984
28	SA Volatility Index	9.647	10.931	0.093	1.139	0.868	10.197	6.721	0.173	2.119	0.982
	(Partially) Shared Indicators										
29	Equity Risk Premium Dif	10.597	9.798	0.130	1.588	0.946	10.700	6.845	0.189	2.321	0.989
30	YoY change in SA CPI	8.119	10.346	0.059	0.717	0.764	10.231	6.853	0.171	2.095	0.981
31	US/SA Currency Exchange Rate	10.900	9.572	0.141	1.729	0.959	10.354	6.769	0.178	2.179	0.985
32	SA Economic Growth Rate	9.421	10.155	0.095	1.152	0.872	9.946	6.863	0.160	1.957	0.975

HS-FP Method 5: Joint Time & State Conditioned Entropy Pooling

Config	Sector Signal	MV + HS-FP					HRP + HS-FP				
	Industrials	Ann. R(%)	Ann. σ (%)	SR	Ann.SR	PSR	Ann. R(%)	Ann. σ (%)	SR	Ann.SR	PSR
33	SA DLI	7.671	12.883	0.038	0.463	0.681	9.742	7.175	0.146	1.776	0.962
34	SA PMI	7.081	13.539	0.025	0.296	0.618	11.289	7.161	0.203	2.488	0.993
	Financials										
35	JPM EMB Index	10.079	11.012	0.103	1.258	0.898	10.182	7.270	0.160	1.953	0.974
36	Market Momentum	11.962	12.525	0.130	1.600	0.945	10.073	7.501	0.151	1.845	0.965
	Bonds										
37	SA Money Supply	10.245	11.713	0.101	1.232	0.895	10.900	7.297	0.185	2.273	0.989
	Volatility Indices										
38	US Volatility Index	11.503	12.632	0.119	1.468	0.930	10.273	7.528	0.157	1.924	0.973
39	SA Volatility Index	9.142	10.991	0.081	0.981	0.839	10.492	7.472	0.167	2.040	0.978
	(Partially) Shared Indicators										
40	Equity Risk Premium Dif	9.272	10.723	0.086	1.046	0.854	11.003	7.370	0.187	2.287	0.988
41	YoY change in SA CPI	5.741	11.714	-0.003	-0.034	0.486	11.086	7.070	0.198	2.428	0.991
42	US/SA Currency Exchange Rate	9.491	11.188	0.087	1.065	0.864	10.400	7.230	0.169	2.062	0.983
43	SA Economic Growth Rate	8.444	12.069	0.058	0.704	0.762	9.296	7.348	0.126	1.533	0.939

Figure 3.10: Performance evaluation of a conventional back-test model (Config.5 from Chp.2) and five time- and state-conditioned HS-FP methods under simulated stress-tests. Metrics are compared across methods using conditional formatting: favourable values are shaded darker green and less-favourable values darker red (reversed for annualised volatility). *E.g.* the highest Ann.SR (1.82) is from the MV + HS-FP optimisation in “Method 4: State-Conditioned Kernel Smoothing” conditioned on the JPM emerging market index, shaded darkest green; the lowest (-0.03) is from “Method 5: Joint Time & State Conditioned Entropy Pooling” conditioned on YoY change in CPI, shaded darkest red. Notably, “Method 4” demonstrates the highest proportion of outperformance among all methods. Results use default HS-FP hyper-parameter values listed in Tab.3.2.

The **base model** (Config 5) shows strong performance across all metrics, with an Ann.SR of 1.80 for MV optimisation and 2.13 for HRP, indicating solid risk-adjusted returns. High PSRs of 0.96 for MV and 0.98 for HRP further boost confidence in the results. However, since the model relies entirely on historical data, there's a clear risk of overfitting, and its lack of adaptability raises questions about how it would handle different or changing market conditions. Without testing it across multiple simulated regimes, there's no guarantee the performance would hold up outside the historical dataset.

Time-conditioned HS-FP: The exponential decay method outperforms the rolling window approach across multiple metrics, supporting [85]'s argument that exponential weights better capture time-dependency. For MV optimisation, the exponential decay method achieves higher Ann.R (10.85%) compared to the rolling window (9.17%). Similarly, it provides better volatility control, with Ann. σ at 10.21% for exponential decay versus 10.998% for the rolling window. These results highlight that smoother time-conditioning, like exponential decay, is more effective than abrupt adjustments like rolling windows, especially for achieving stronger risk-adjusted performance in MV strategies. However, for HRP optimisation, the difference between the two time-conditioned HS-FP methods is minimal, providing little evidence to favour one approach over the other for incorporating time-conditioning into stress-testing.

Single-state conditioned HS-FP: Crisp band conditioning struggles with consistency as it relies on fixed thresholds that often fail to capture the fluid nature of market regimes. The results highlight this issue, *e.g.* in Config.20, when conditioned on the USD/ZAR exchange rate, it performs strongly with an Ann.SR of 1.78 (MV + HS-FP), while when the SA PMI is used, it significantly underperforms with an Ann.SR of just 0.30. This variability reflects the rigidity of crisp bands, where regimes don't often fit neatly within the predefined cut-offs.

Kernel smoothing, on the other hand, offers a more flexible approach by assigning weights continuously, making it better at capturing subtle shifts in market conditions. This is evident in Config.24 when conditioned on the JPM emerging markets index, the kernel smoothing achieves the highest Ann.SR of 1.82 for the MV + HS-FP across all methods. It also consistently delivers strong results for HRP + HS-FP, making it the most reliable method overall.

The difference becomes clearer when we consider that kernel smoothing is effectively a softened version of crisp conditioning, applying an exponential decay to the bands. The smoothing not only improves consistency but also resolves issues related to covariance estimation. For example, in Config.20, crisp bands deliver strong results for MV + HS-FP but poor performance for HRP + HS-FP with the same signal. With kernel smoothing, both the MV + HS-FP and HRP + HS-FP perform well, likely because smoothing addresses the challenges HRP + HS-FP faces when relying solely on covariance matrices.

Overall, kernel smoothing stands out for its ability to adapt to changing market conditions and consistently outperform crisp band conditioning. This comparison is investigated in greater detail in the [Case studies exploring underlying simulation dynamics](#) below.

Joint time and state-conditioned HS-FP: Despite the flexibility that entropy pooling is meant to display—given its design allows it to mimic a kernel estimator with an optimal bandwidth and centre by switching between a Gaussian or a exponential kernel (Sec.3.4.3.1)—the time and state conditioned method shows mixed results, particularly for the MV + HS-FP. The EP is meant to leverage exponential decay to emphasise data over the last 12 months as $\lambda = 12$ for the prior (the same λ used in Method 2), while incorporating the latest values of the conditioning signal as $z^* = z_T$ for the crisp views. This dual emphasis aims to balance responsiveness with stability. However, while this responsiveness can be advantageous in capturing swift market movements, it also introduces significant sensitivity, especially when the conditioning signals reflect transient or short-lived market phenomena (see Fig.3.14 below in the case study).

The lowest simulation performance obtained from any of the methods is seen in Config.41 from MV + HS-FP when conditioned on YoY change in CPI, with the lowest $\text{Ann.R}(\%) = 5.74$, a negative SR of -0.03 , and notably the lowest PSR across the methods all of 0.49 . The PSR in these simulations represents the probability of outperforming a zero investment skill benchmark portfolio (Sec.2.2.1.1). Therefore, a PSR of 0.49 says that there is a 49% probability that the portfolio's true SR is greater than zero. Conversely, there is a 51% probability that the SR is less than or equal to zero. This implies that while the method may occasionally produce poor results, these instances serve as warnings rather than reflections of typical performance, cautioning against over-interpretation and potential dismissal of the EP approach.

Interestingly, within the MV + HS-FP framework, two signals—market momentum and the US Volatility Index—achieved strong annual returns. This unexpected outcome suggests that these signals may capture persistent market trends or volatility patterns that EP can effectively exploit when combined with MV optimisation. Further investigation is necessary to understand the underlying mechanisms that enable these signals to perform well within this context (This is where sensitivity analysis can provide insight into whether the EP approach could be effective with proper calibration, provided it avoids significant overfitting).

3.7.2 MV vs. HRP comparison within methods

In Fig.3.11 below, we highlight the relative performance of the MV and the HRP optimisations within each method. The conditional formatting is applied down the columns **within** each method, comparing the metrics of the MV optimisation to those of the HRP. As before, higher values are shaded darker green and lower values darker red (reversed for annualised volatility). Using $\text{Ann.R}(\%)$ as an example: within “Method 3: State-Conditioned Crisp Bands”, the $\text{Ann.R}(\%)$ for configurations 11 to 21 are selected for the MV simulations as well as for the HRP, shading the results for both the MV and HRP collectively to see which method has the more favourable $\text{Ann.R}(\%)$ across the configurations. In this example, Config.19 under the HRP + HS-FP optimisation has the highest $\text{Ann.R}(\%) = 12.65$ and Config.12 under the MV + HS-FP optimisation has the lowest $\text{Ann.R}(\%) = 7.06$.

2. MV & HRP Comparison Within Methods

Base Model											
		MV					HRP				
Config.		Ann.R(%)	Ann. σ (%)	SR	Ann.SR	PSR	Ann.R(%)	Ann. σ (%)	SR	Ann.SR	PSR
5		10.876	9.130	0.147	1.801	0.964	10.198	6.687	0.174	2.130	0.982

HS-FP Method 1: Time-Conditioned Rolling Window

		MV + HS-FP					HRP + HS-FP				
Config.	Static parameter value	Ann.R(%)	Ann. σ (%)	SR	Ann.SR	PSR	Ann.R(%)	Ann. σ (%)	SR	Ann.SR	PSR
9	$\lambda = 12$ months	9.165	10.998	0.081	0.988	0.844	10.436	6.796	0.181	2.212	0.985

HS-FP Method 2: Time-Conditioned Exponential Decay

		MV + HS-FP					HRP + HS-FP				
Config.	Static parameter value	Ann.R(%)	Ann. σ (%)	SR	Ann.SR	PSR	Ann.R(%)	Ann. σ (%)	SR	Ann.SR	PSR
10	$\tau = 12$ months	10.848	10.206	0.131	1.601	0.947	10.279	6.732	0.176	2.155	0.983

HS-FP Method 3: State-Conditioned Crisp Bands

Config.	Sector Signal	MV + HS-FP					HRP + HS-FP				
		Ann.R(%)	Ann. σ (%)	SR	Ann.SR	PSR	Ann.R(%)	Ann. σ (%)	SR	Ann.SR	PSR
11	Industrials	8.958	11.526	0.073	0.884	0.819	9.462	7.838	0.124	1.506	0.935
12	SA DLI	7.061	13.020	0.025	0.304	0.621	10.132	7.301	0.157	1.922	0.971
	SA PMI										
	Financials										
13	JPM EMB Index	10.970	11.748	0.116	1.427	0.920	10.450	7.495	0.165	2.013	0.979
14	Market Momentum	11.575	12.803	0.119	1.463	0.930	9.914	8.191	0.133	1.626	0.946
	Bonds										
15	SA Money Supply	11.325	11.379	0.129	1.578	0.945	10.004	7.806	0.143	1.749	0.960
	Volatility Indices										
16	US Volatility Index	8.804	11.775	0.067	0.818	0.796	10.145	8.385	0.137	1.678	0.958
17	SA Volatility Index	8.035	11.267	0.052	0.632	0.736	11.262	8.071	0.180	2.206	0.986
	(Partially) Shared Indicators										
18	Equity Risk Premium Dif	9.315	10.771	0.087	1.054	0.856	9.815	7.771	0.137	1.668	0.951
19	YoY change in SA CPI	8.908	12.935	0.064	0.772	0.781	12.651	7.638	0.236	2.920	0.998
20	US/SA Currency Exchange Rate	11.674	10.754	0.145	1.778	0.973	8.758	7.872	0.100	1.209	0.890
21	SA Economic Growth Rate	10.343	11.070	0.109	1.328	0.907	8.922	7.590	0.109	1.322	0.911

HS-FP Method 4: State-Conditioned Kernel Smoothing

Config.	Sector Signal	MV + HS-FP					HRP + HS-FP				
		Ann.R(%)	Ann. σ (%)	SR	Ann.SR	PSR	Ann. R(%)	Ann. σ (%)	SR	Ann.SR	PSR
22	Industrials	8.792	10.989	0.072	0.877	0.810	10.518	7.062	0.177	2.164	0.985
23	SA DLI	10.716	11.064	0.118	1.444	0.922	10.626	6.589	0.194	2.376	0.990
	SA PMI										
	Financials										
24	JPM EMB Index	11.505	10.198	0.148	1.818	0.966	10.237	6.710	0.175	2.141	0.983
25	Market Momentum	9.842	10.151	0.106	1.290	0.900	9.904	6.875	0.158	1.933	0.972
	Bonds										
26	SA Money Supply	10.854	9.229	0.145	1.776	0.962	10.408	6.706	0.182	2.227	0.986
	Volatility Indices										
27	US Volatility Index	11.361	10.493	0.140	1.716	0.956	10.297	6.736	0.177	2.164	0.984
28	SA Volatility Index	9.647	10.931	0.093	1.139	0.868	10.197	6.721	0.173	2.119	0.982
	(Partially) Shared Indicators										
29	Equity Risk Premium Dif	10.597	9.798	0.130	1.588	0.946	10.700	6.845	0.189	2.321	0.989
30	YoY change in SA CPI	8.119	10.346	0.059	0.717	0.764	10.231	6.853	0.171	2.095	0.981
31	US/SA Currency Exchange Rate	10.900	9.572	0.141	1.729	0.959	10.354	6.769	0.178	2.179	0.985
32	SA Economic Growth Rate	9.421	10.155	0.095	1.152	0.872	9.946	6.863	0.160	1.957	0.975

HS-FP Method 5: Joint Time & State Conditioned Entropy Pooling

Config.	Sector Signal	MV + HS-FP					HRP + HS-FP				
		Ann. R(%)	Ann. σ (%)	SR	Ann.SR	PSR	Ann. R(%)	Ann. σ (%)	SR	Ann.SR	PSR
33	Industrials	7.671	12.883	0.038	0.463	0.681	9.742	7.175	0.146	1.776	0.962
34	SA DLI	7.081	13.539	0.025	0.296	0.618	11.289	7.161	0.203	2.488	0.993
	SA PMI										
	Financials										
35	JPM EMB Index	10.079	11.012	0.103	1.258	0.898	10.182	7.270	0.160	1.953	0.974
36	Market Momentum	11.962	12.525	0.130	1.600	0.945	10.073	7.501	0.151	1.845	0.965
	Bonds										
37	SA Money Supply	10.245	11.713	0.101	1.232	0.895	10.900	7.297	0.185	2.273	0.989
	Volatility Indices										
38	US Volatility Index	11.503	12.632	0.119	1.468	0.930	10.273	7.528	0.157	1.924	0.973
39	SA Volatility Index	9.142	10.991	0.081	0.981	0.839	10.492	7.472	0.167	2.040	0.978
	(Partially) Shared Indicators										
40	Equity Risk Premium Dif	9.272	10.723	0.086	1.046	0.854	11.003	7.370	0.187	2.287	0.988
41	YoY change in SA CPI	5.741	11.714	-0.003	-0.034	0.486	11.086	7.070	0.198	2.428	0.991
42	US/SA Currency Exchange Rate	9.491	11.188	0.087	1.065	0.864	10.400	7.230	0.169	2.062	0.983
43	SA Economic Growth Rate	8.444	12.069	0.058	0.704	0.762	9.296	7.348	0.126	1.533	0.939

Figure 3.11: Performance evaluation of the MV optimisation *vs.* the HRP optimisation within simulated stress-tests. Conditional formatting is applied to compare results of MV *vs.* HRP within each method. *E.g.* within “Method 3: State-Conditioned Crisp Bands” we compare the performance of all MV + HS-FP annual returns *vs.* all HRP + HS-FP annual returns to assess within method relative performance. The highest Ann.R is the darkest green (12.65% generated by the HRP + HS-FP when conditioning using YoY change in CPI) and the worst is the darkest red (7.06% from the MV + HS-FP conditioned using SA PMI). All results are generated using default HS-FP hyper-parameter values as listed in Tab.3.2.

The comparison of the MV and HRP portfolio construction techniques reveals a significant outperformance of HRP over MV across majority of the simulations. The initial expectation was that the regime-capturing, scenario-based approach to stress-testing would help the SR-focused MV strategy adapt better to changing market conditions by reducing volatility and by responding to captured regime changes. Unfortunately, this hasn't materialised, as it's consistently outperformed by the diversification-focused HRP strategy. In 94.4% (34 out of 36) of the back-test simulations, HRP achieved better annualised risk-adjusted returns, as indicated by the Ann.SR. The average Ann.SR for MV across the methods is 1.16 and for the HRP it is 2.02. Furthermore, even though the MV is more focused on maximising returns relative to HRP, its average Ann.R(%) of 9.73 also falls short of HRP's which is 10.23.

The MV also does not see improvement from the base model into the HS-FP simulations for the Ann.SR, whereas the HRP sees several of the simulations realising Ann.SR greater than the base case —and those that are not, find themselves not far below it. The FPrs in HS-FP simulations allow certain market events to have disproportionate impacts on portfolio returns, mimicking stressed market environments. HRP, with its structure that naturally minimises overexposure to any single asset or cluster, is less sensitive to these probabilistic shifts than MV, may be the cause of the underperformance.

The PSR of the simulations also provide further support for the reliability of the HRP relative to the MV portfolios. The HRP portfolios show an average PSR of 0.97 implying that on average across the stress-tests HRP there is 97.2% probability that the portfolio's true SR is greater than a zero skill portfolio, whereas the MV only shows an average probability of 85.4%.

As noted by Best and Grauer [14], covariance estimates are generally more reliable than mean estimates. If the HS-FP methods are not effective in capturing regime changes, the forward-looking estimates of μ and Σ would suffer from increased estimation errors. As the MV depends on both mean and covariance estimates, it is exposed to a dual estimate error which could have a larger impact on OOS performance [23].

3.7.3 HS-FP signals comparison within methods

In Fig.3.12, each configuration row is compared with all other configurations within the method. Using “Method 3: State-Conditioned Crisp Bands” as an example again, all metrics of Config.11 are compared (for both the MV + HS-FP and the HRP + HS-FP) to those of Config.12 through to Config.21. Our objective is to assess whether any given signal(s) produces constantly reliable/unreliable results across all state-conditioned scenario methods.

3. HS-FP Signal Comparison Within Methods

Base Model											
		MV					HRP				
Config.		Ann.R(%)	Ann. σ (%)	SR	Ann.SR	PSR	Ann.R(%)	Ann. σ (%)	SR	Ann.SR	PSR
5		10.876	9.130	0.147	1.801	0.964	10.198	6.687	0.174	2.130	0.982

HS-FP Method 1: Time-Conditioned Rolling Window											
		MV + HS-FP					HRP + HS-FP				
Config.	Static parameter value	Ann.R(%)	Ann. σ (%)	SR	Ann.SR	PSR	Ann.R(%)	Ann. σ (%)	SR	Ann.SR	PSR
9	$\lambda = 12$ months	9.165	10.998	0.081	0.988	0.844	10.436	6.796	0.181	2.212	0.985

HS-FP Method 2: Time-Conditioned Exponential Decay											
Config.		MV + HS-FP					HRP + HS-FP				
	Static parameter value	Ann.R(%)	Ann. σ (%)	SR	Ann.SR	PSR	Ann.R(%)	Ann. σ (%)	SR	Ann.SR	PSR
10	$\tau = 12$ months	10.848	10.206	0.131	1.601	0.947	10.279	6.732	0.116	2.155	0.983

HS-FP Method 3: State-Conditioned Crisp Bands											
Config.	Sector Signal	MV + HS-FP					HRP + HS-FP				
		Ann.R(%)	Ann. σ (%)	SR	Ann.SR	PSR	Ann.R(%)	Ann. σ (%)	SR	Ann.SR	PSR
11	Industrials SA DLI	8.958	11.526	0.073	0.884	0.819	9.462	7.838	0.124	1.506	0.935
12	SA PMI	7.061	13.020	0.025	0.304	0.621	10.132	7.301	0.157	1.922	0.971
	Financials										
13	JPM EMB Index	10.970	11.748	0.116	1.427	0.920	10.450	7.495	0.165	2.013	0.979
14	Market Momentum	11.575	12.803	0.119	1.463	0.930	9.914	8.191	0.133	1.626	0.946
	Bonds										
15	SA Money Supply	11.325	11.379	0.129	1.578	0.945	10.004	7.806	0.143	1.749	0.960
	Volatility Indices										
16	US Volatility Index	8.804	11.775	0.067	0.818	0.796	10.145	8.385	0.137	1.678	0.958
17	SA Volatility Index	8.035	11.267	0.052	0.632	0.736	11.262	8.071	0.180	2.206	0.986
	(Partially) Shared Indicators										
18	Equity Risk Premium Dif	9.315	10.771	0.087	1.054	0.856	9.815	7.771	0.137	1.668	0.951
19	YoY change in SA CPI	8.908	12.935	0.064	0.772	0.781	12.651	7.638	0.236	2.920	0.998
20	US/SA Currency Exchange Rate	11.674	10.754	0.145	1.778	0.973	8.758	7.872	0.100	1.209	0.890
21	SA Economic Growth Rate	10.343	11.070	0.109	1.328	0.907	8.922	7.590	0.109	1.322	0.911

HS-FP Method 4: State-Conditioned Kernel Smoothing											
Config.	Sector Signal	MV + HS-FP					HRP + HS-FP				
		Ann.R(%)	Ann. σ (%)	SR	Ann.SR	PSR	Ann.R(%)	Ann. σ (%)	SR	Ann.SR	PSR
	Industrials										
22	SA DLI	8.792	10.989	0.072	0.877	0.810	10.518	7.062	0.177	2.164	0.985
23	SA PMI	10.716	11.064	0.118	1.444	0.922	10.626	6.589	0.194	2.376	0.990
	Financials										
24	JPM EMB Index	11.505	10.198	0.148	1.818	0.966	10.237	6.710	0.175	2.141	0.983
25	Market Momentum	9.842	10.151	0.106	1.290	0.900	9.904	6.875	0.158	1.933	0.972
	Bonds										
26	SA Money Supply	10.854	9.229	0.145	1.776	0.962	10.408	6.706	0.182	2.227	0.986
	Volatility Indices										
27	US Volatility Index	11.361	10.493	0.140	1.716	0.956	10.297	6.736	0.177	2.164	0.984
28	SA Volatility Index	9.647	10.931	0.093	1.139	0.868	10.197	6.721	0.173	2.119	0.982
	(Partially) Shared Indicators										
29	Equity Risk Premium Dif	10.597	9.798	0.130	1.588	0.946	10.700	6.845	0.189	2.321	0.989
30	YoY change in SA CPI	8.119	10.346	0.059	0.717	0.764	10.231	6.853	0.171	2.095	0.981
31	US/SA Currency Exchange Rate	10.900	9.572	0.141	1.729	0.959	10.354	6.769	0.178	2.179	0.985
32	SA Economic Growth Rate	9.421	10.155	0.095	1.152	0.872	9.946	6.863	0.160	1.957	0.975

HS-FP Method 5: Joint Time & State Conditioned Entropy Pooling											
Config.	Sector Signal	MV + HS-FP					HRP + HS-FP				
		Ann.R(%)	Ann. σ (%)	SR	Ann.SR	PSR	Ann.R(%)	Ann. σ (%)	SR	Ann.SR	PSR
	Industrials										
33	SA DLI	7.671	12.883	0.038	0.463	0.681	9.742	7.175	0.146	1.776	0.962
34	SA PMI	7.081	13.539	0.025	0.296	0.618	11.289	7.161	0.203	2.488	0.993
	Financials										
35	JPM EMB Index	10.079	11.012	0.103	1.258	0.898	10.182	7.270	0.160	1.953	0.974
36	Market Momentum	11.962	12.525	0.130	1.600	0.945	10.073	7.501	0.151	1.845	0.965
	Bonds										
37	SA Money Supply	10.245	11.713	0.101	1.232	0.895	10.900	7.297	0.185	2.273	0.989
	Volatility Indices										
38	US Volatility Index	11.503	12.632	0.119	1.468	0.930	10.273	7.528	0.157	1.924	0.973
39	SA Volatility Index	9.142	10.991	0.081	0.981	0.839	10.492	7.472	0.167	2.040	0.978
	(Partially) Shared Indicators										
40	Equity Risk Premium Dif	9.272	10.723	0.086	1.046	0.854	11.003	7.370	0.187	2.287	0.988
41	YoY change in SA CPI	5.741	11.714	-0.003	-0.034	0.486	11.086	7.070	0.198	2.428	0.991
42	US/SA Currency Exchange Rate	9.491	11.188	0.087	1.065	0.864	10.400	7.230	0.169	2.062	0.983
43	SA Economic Growth Rate	8.444	12.069	0.058	0.704	0.762	9.296	7.348	0.126	1.533	0.939

Figure 3.12: The conditioning was applied within each state-conditioned HS-FP method across the signals to draw a comparison of the one signal's performance relative to another within the same method. The purpose of the comparison is to see if any of the signals display consistent performance across methods, suggestive of it being a better overall market state indicator. These results could be used to inform the selection of the top performing single-state signals as listed in Tab.3.1 which can later be incorporated into ensemble approaches (Sec.3.4.4). All results are generated using default HS-FP hyper-parameter values as listed in Tab.3.2.

At a glance, the results in Fig.3.12 above indicate that no single set of signals consistently outperforms across methods. Despite the lack of standout signals, some clusters of results show desirable outcomes. Among these, SA Money Supply is the only signal that performs relatively adequately for both MV and HRP portfolios across three state-conditioned methods. While this does not represent a breakthrough, it highlights a potential area for further investigation.

In configurations 19, 30, and 41, the YoY change in CPI as a state signal produces particularly unfavourable results for MV optimisation but works relatively well for HRP. We speculate that this may be due to a mismatch in the timing of data releases and market responses. CPI is announced monthly with a one-month lag, while the South African Reserve Bank’s Monetary Policy Committee (MPC) reviews and potentially adjusts the repo rate every two months.

By lagging the CPI measure in our data preprocessing by one month to allow for market response, we’re basing our optimisation at time t on CPI data from $t - 2$. This could create a timing mismatch because the repo rate can change at $t + 1$, affecting bond prices and realised returns then. Consequently, MV optimisation might be making allocation decisions at t based on outdated market conditions, without accounting for potential repo rate changes at $t + 1$. This misalignment may lead to poor performance for MV, which relies heavily on accurate estimates of expected returns and covariances that become less reliable when based on lagged information.

In contrast, HRP might handle this timing issue more effectively. By focusing on diversification and the hierarchical relationships between assets, HRP could be less sensitive to the lag in economic indicators like CPI, maintaining better performance even with data mismatches. The poor performance of MV with the lagged CPI signal might highlight the importance of aligning data timing with market events in portfolio construction. Adjusting for these timing mismatches or incorporating more timely indicators could potentially enhance MV’s ability to capture current market conditions and improve overall performance.

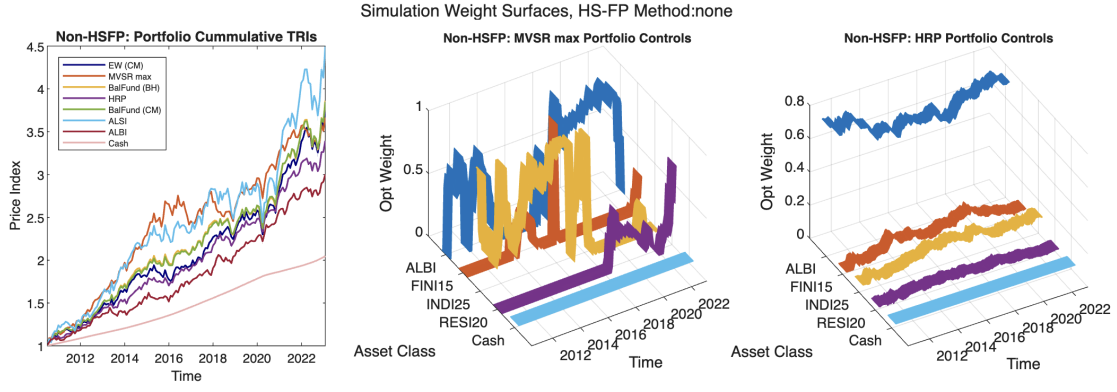
3.7.4 Case studies exploring underlying simulation dynamics

To better understand the methods and underlying simulation dynamics, we examine two cases that highlight how the HS-FP methods function and respond to the underlying data.

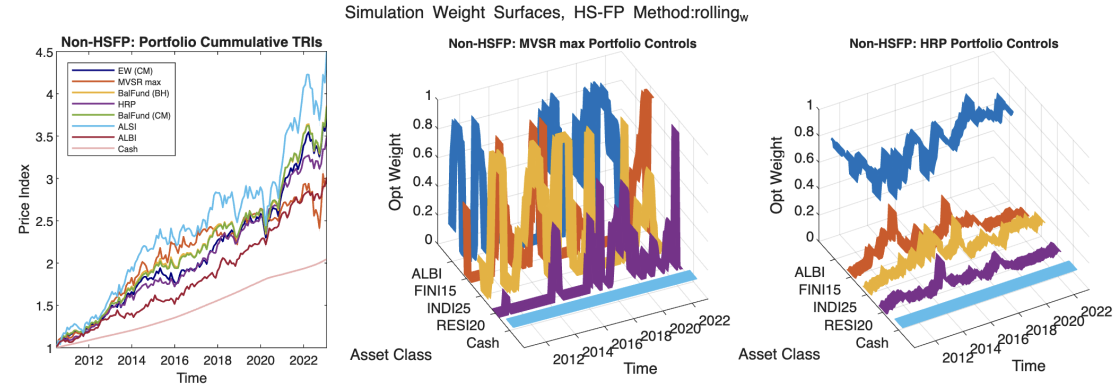
1. Observing the dynamics of all methods, using SA Money Supply for the state-conditioned methods.
2. Investigate the difference between abrupt crisp bands and kernel smoothing methods by exploring the contrasting results from Config. 20 seen in Sec.3.7.1.

Example 1: Comparing the dynamics between all methods

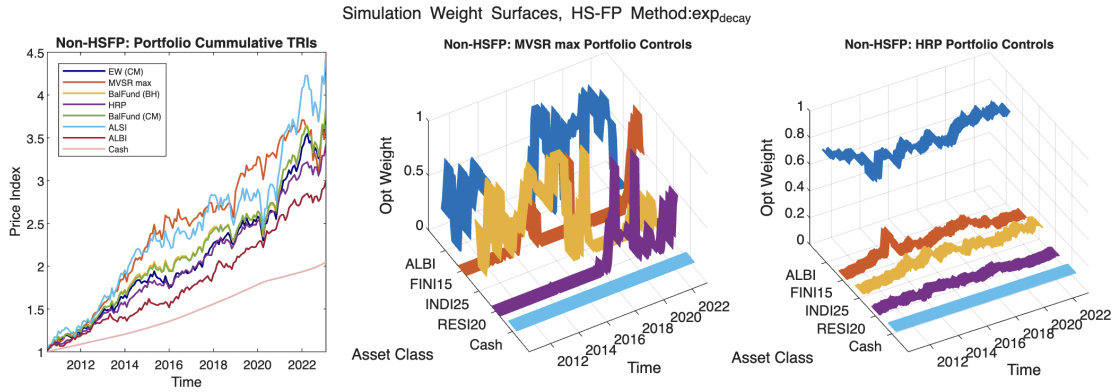
Based on the signal comparison in Fig.3.12, SA Money Supply showed relatively robust results across the three state-conditioned methods. Therefore, we will use it for our first example as it “puts all methods on an even playing field” to allow us to analyse the optimal weight surfaces of the simulations to compare the performance of conventional and HS-FP methods.



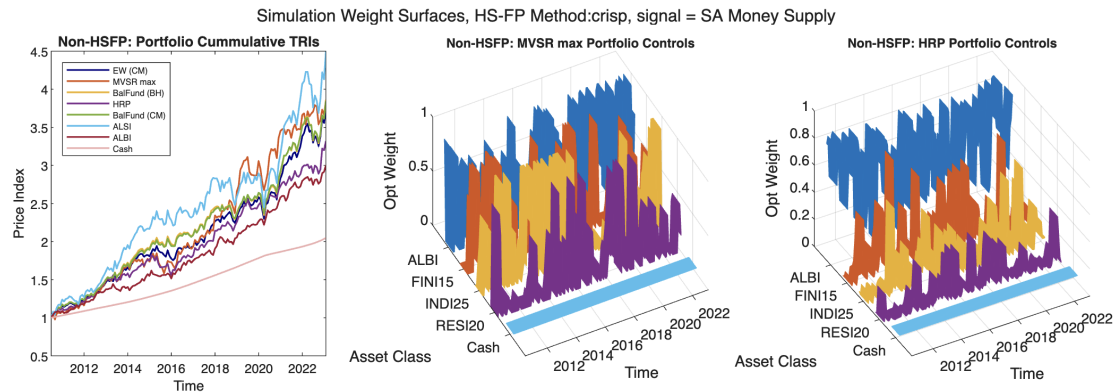
(a) **Base Model - Config.5:** conventional back-test simulation (as seen in Chp.2). **Ann. metrics:** $MV[R = 10.88\%, \sigma = 9.13\%, SR = 1.8, PSR = 0.96]$, $HRP[R = 10.20\%, \sigma = 6.69\%, SR = 2.13, PSR = 0.98]$.



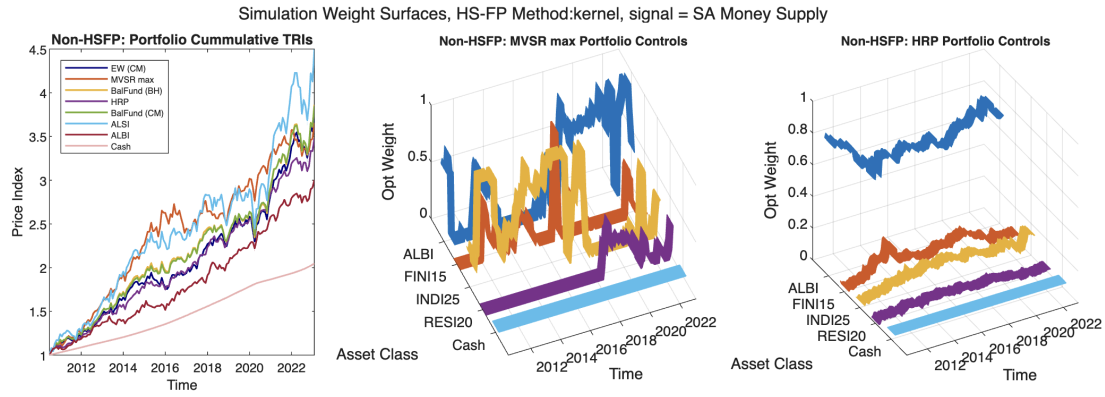
(b) **Rolling Window - Config.9:** HS-FP time conditioned simulation using a rolling window with $\lambda = 12$ months. **Ann. metrics:** $MV[R = 9.17\%, \sigma = 11\%, SR = 0.99, PSR = 0.84]$, $HRP[R = 10.44\%, \sigma = 6.80\%, SR = 2.21, PSR = 0.99]$.



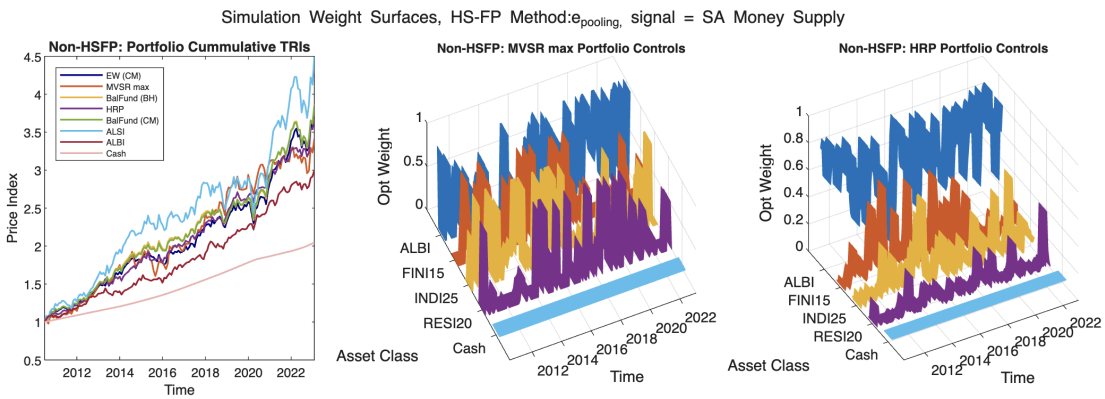
(c) **Exponential Decay - Config.10:** HS-FP time conditioned simulation using an exponential decay with $\tau = 12$ months. **Ann. metrics:** $MV[R = 10.85\%, \sigma = 10.21\%, SR = 1.60, PSR = 0.95]$, $HRP[R = 10.28\%, \sigma = 6.73\%, SR = 2.16, PSR = 0.98]$.



(d) **Crisp Bands - Config.15:** HS-FP state conditioned simulation using crisp bands conditioned on a 3 month lagged SA money supply with $z^* = z_T$ and $\alpha = 0.1$. **Ann. metrics:** $MV[R = 11.33\%, \sigma = 11.38\%, SR = 1.58, PSR = 0.95]$, $HRP[R = 10\%, \sigma = 7.81\%, SR = 1.75, PSR = 0.96]$.



(a) **Kernel Smoothing - Config.26:** HS-FP state conditioned simulation using kernel smoothing conditioned on a 3 month lagged SA money supply with $z^* = z_T$, $h = 0.5$ and $\gamma = 2$. **Ann. metrics:** MV[R = 10.85%, $\sigma = 9.23\%$, SR = 1.78, PSR = 0.96], HRP[R = 10.41%, $\sigma = 6.71\%$, SR = 2.23, PSR = 0.99].



(b) **Entropy Pooling - Config.37:** HS-FP time & state conditioned simulation using entropy pooling conditioned on a 3 month lagged SA money supply with $z^* = z_T$, $\alpha = 0.1$ and $\gamma = 2$. **Ann. metrics:** MV[R = 10.25%, $\sigma = 11.71\%$, SR = 1.23, PSR = 0.90], HRP[R = 10.90%, $\sigma = 7.30\%$, SR = 2.27, PSR = 0.99].

Figure 3.14: Visual comparison of optimal portfolio results of back-test simulation methods using the SA Money Supply as a state signal in the state-conditioned HS-FP methods for a constant comparative basis.

Fig.3.14 shows the cumulative performance curves and optimal weight surfaces for the base case and all five HS-FP methods using SA Money Supply as the conditioning signal. The initial observations suggest increased sensitivity to market changes when conditioned on either time or state (or both) compared to the base case, consistent with the HS-FP framework's aim to create more responsive portfolio stress-tests.

More abrupt conditioning methods, such as rolling window time conditioning, state-conditioned crisp bands, and entropy pooling, lead to a significant increase in portfolio turnover, particularly for MV but also for HRP. This is likely due to signal changes as data points enter and exit the 36-month rolling window in the back-test as it increments over the historical data. Such increased turnover is problematic as it raises trading costs, reducing the portfolios' realised returns. Future work should address this by incorporating turnover constraints into the portfolio optimisations and calculating more realistic realised returns that account for trading costs. Given that MV optimisation consistently results in higher turnover than HRP across simulation types, it is reasonable to assume that without turnover constraints, the realised returns for MV would suffer more from trading costs. This would likely widen the performance gap between MV and HRP, as these costs would further erode MV's annualised returns.

Example 2: Investigate the difference between crisp bands and kernel smoothing

From the first results comparison in Sec.3.7.1 we noticed that the HRP performs poorly under the crisp method relative to other methods when conditioned on the USD/ZAR exchange rate, while MV performs exceptionally well (Config. 20 *vs.* Config. 31). In contrast, with the kernel smoothing method, both MV and HRP demonstrate relatively satisfactory results. These results have been extracted into a small sub-table below as a reminder:

HS-FP State-Conditioned Method Comparison											
Config	Sector Signal	MV + HS-FP					HRP + HS-FP				
		Ann.R(%)	Ann. σ (%)	SR	Ann.SR	PSR	Ann. R(%)	Ann. σ (%)	SR	Ann.SR	PSR
Method 3: Crisp Bands											
20	US/SA Currency Exchange Rate	11.674	10.754	0.145	1.778	0.973	8.758	7.872	0.100	1.209	0.890
Method 4: Kernel Smoothing											
31	US/SA Currency Exchange Rate	10.900	9.572	0.141	1.729	0.959	10.354	6.769	0.178	2.179	0.985

As the backtest consists of a rolling window that shifts over historical time series data, we can't easily examine the FPrs that are impacting the simulations as a new set is generated at each time step. Therefore, we will use the full time period from late 2007 to early 2023 to generate state-conditioned FPrs as an example to see if it provides some insight into the HS-FP method differences. To better understand the optimisation dynamics created by the abrupt crisp method versus the smoother kernel approach, we compare “HS-FP Method 3: State-Conditioned Crisp Bands” and “HS-FP Method 4: State-Conditioned Kernel Smoothing”, both conditioned on the USD/ZAR exchange rate *i.e.* we are hoping to capture similar dynamics to those that generated the results for configurations 20 and 31. The comparison shown in Fig.3.15 below.

To try to understand what is driving these results we use Fig.3.15 and work backward through the simulation inputs, isolating sections of historical data selected by each method to illustrate what data was used to calculate the HS-FP moments in these configurations.

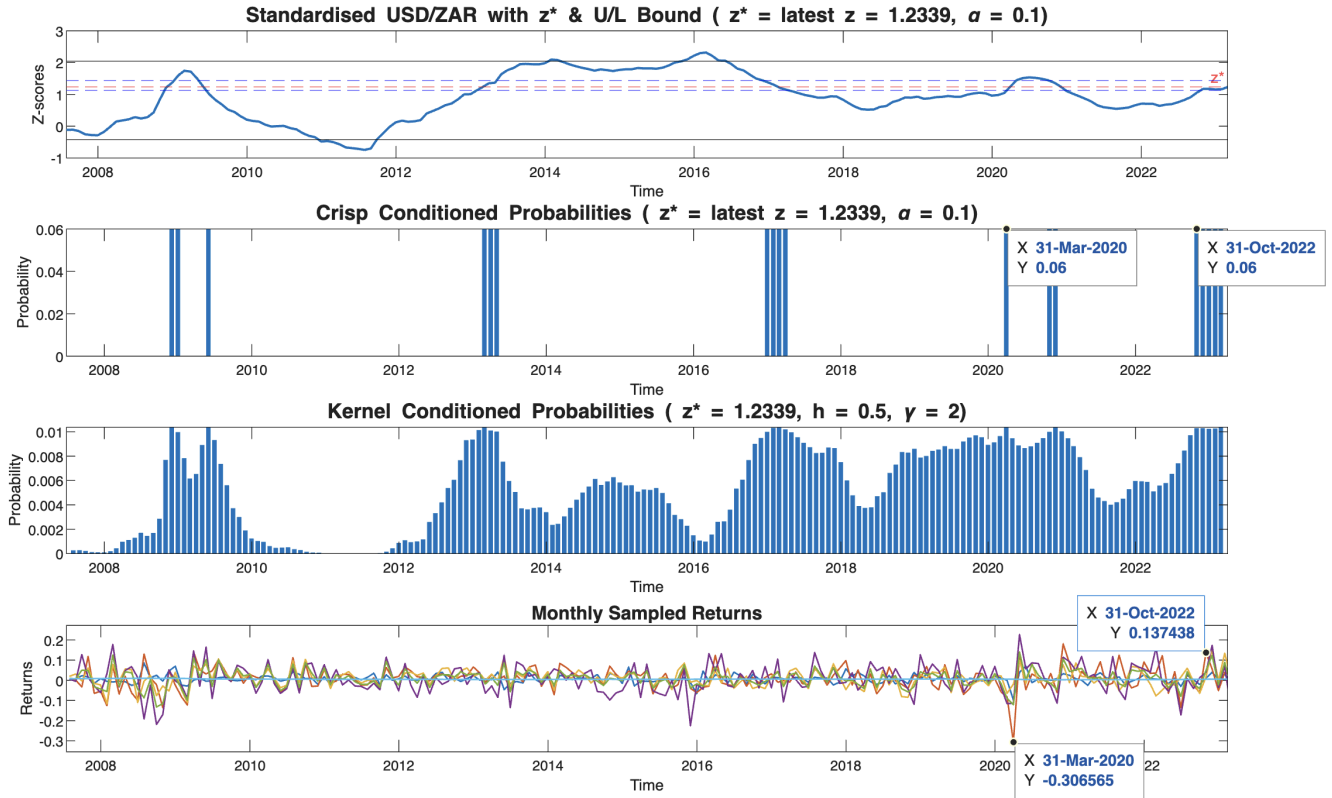


Figure 3.15: FPrs generated by “HS-FP Method 3: State-Conditioned Crisp Bands” and “HS-FP Method 4: State-Conditioned Kernel Bands,” conditioned on USD/ZAR from late 2007 to early 2023, along with monthly returns. Key dates selected by the crisp FPrs method, 31 March 2020 and 31 October 2020, are mapped onto the returns series to highlight volatility that would affect the HS-FP covariance calculation. In contrast, the kernel method smooths FPrs around these dates, reducing sensitivity to extreme values. This provides insight into why even when conditioned on the same signal, the results produced by the two HS-FP are vastly different.

The indicated dates on the crisp FPrs, 31 March 2020 and the 31st October 2020, are mapped onto the monthly returns series to highlight the volatility seen by the HS-FP covariance calculation under the crisp method. Using the crisp method with the target value set to the latest observation happened to correspond to two time points in the historical data with an extreme low (31 March 2020) and relatively high (31st October 2020) monthly returns. Given these contrasting values the volatility seen by the optimisations would be significant.

We are reminded that, the MV method requires estimation of both the expected return and the covariance, while the HRP method relies solely on covariance. What makes this interesting is that although there is benefit for the HRP in not be exposed to the double estimation error associated with needing to estimate both the expected return as well as the covariance matrix, it is a key example that if it solely relies on the covariance, the covariance estimate needs to be representative of a broader market context as that IS estimate is the only information that the optimisation receives and can heavily lead to poor performance OOS.

This example demonstrates that, when an abrupt cut-off is applied using crisp bands, the optimisation model receives highly selective information about the market, potentially resulting in misleading inputs due to the narrow scope of the state signal used. In contrast, the kernel method still emphasises the key time points identified by the crisp bands as relevant under the conditioning state signal but, by using

a bandwidth, it captures a broader range of surrounding data. By smoothing over states, the kernel method mitigates the risk of abrupt changes in probability estimates that may not accurately reflect market realities. This approach also reduces the impact of noise and outliers by distributing weights across a broader range of scenarios, thereby enhancing the robustness of the estimates. Consequently, the kernel method provides a more comprehensive and stable view of the market, leading to more reliable inputs for optimisation.

3.8 Results discussion

From Comparison 1—comparing the base model to the forward-looking HS-FP simulations—it remains inconclusive whether incorporating forward-looking estimates of μ and Σ improves the performance and reliability of MV or HRP portfolio backtests. The HS-FP models yield mixed results across simulations, providing no clear advantage. However, the inclusion of forward-looking estimates has highlighted the sensitivity of these backtests to their inputs, offering valuable insight into the range of possible outcomes when market regimes are explicitly considered in portfolio stress-testing. While it's unclear if this approach improves performance, it highlights the value of regime-aware methods in capturing diverse market conditions and helping investors understand the sensitivity of optimisations under abnormal market scenarios.

Despite the lack of overall improvement from the base case to the forward-looking HS-FP models, the state-conditioned kernel smoothing demonstrated tangible benefits by generating more consistent performance across simulations. Taking a more stable approach to state-conditioning, “Method 4: State-Conditioned Kernel Smoothing” enhanced the model’s ability to adjust to changing market conditions without being overly reactive. As the HRP is solely reliant on the covariance matrix the stable approach of the kernel smoothing resulted in more robust outcomes relative to the other methods. When explored in the case study above (Sec.3.7.4), the MV portfolio showed a strong ability to adapt to market changes, however, the turnover and the volatility that that brings, seems to be working against it as it shows inconstant results regardless of the method applied.

What we can take away from this analysis is that the benefit of introducing the HS-FP methods into the simulations results in greater responsiveness of both the MV and HRP when subjected to abnormal market conditions. If we can improve the model’s ability to correctly capture regimes and shifts, the responsiveness to these changes could prove to be extremely beneficial to the stress-testing framework.

From Comparison 2—comparing the performance of MV and HRP within methods—it is evident that the HS-FP framework failed to deliver the expected improvements in MV’s risk-adjusted returns. While HS-FP was intended to enhance MV by capturing regime changes and reducing volatility, it did not provide the anticipated stability or performance advantage over the diversification-focused HRP. Instead, the SR-maximising MV optimisation, explicitly designed to improve risk-adjusted returns, ended up being overly sensitive to HS-FP forward-looking estimates. This sensitivity not only amplified MV’s limitations under non-normal returns and estimation error but also highlighted HS-FP’s shortcomings in effectively identifying and characterising regime changes.

Regime identification is critical for forward-looking frameworks, as it underpins the validity of probabilis-

tic adjustments to historical data. If regime changes are not accurately identified or if the weights applied to historical data fail to reflect actual shifts in market conditions, the forward-looking estimates lose their intended predictive value. Without this clear regime demarcation, the MV's optimisation process became reactive to noisy as shown in the case study, undermining its ability to achieve stable risk-adjusted returns.

Another consideration is estimation error that may be present due to the limited sample sizes used in HS-FP, created by the monthly frequency of the input data. Several factors may explain the superior performance of HRP relative to MV. One key factor is the estimation error associated with the mean and covariance, which affects MV more significantly than HRP. MV optimisation requires accurate estimates of both expected returns and the covariance matrix, increasing the potential for estimation errors. As noted by Best and Grauer [14], covariance estimates tend to be more reliable than mean estimates. Moreover, errors in mean estimation can have a larger impact on OOS performance than errors in covariance estimation [23]. In contrast, the non-parametric approach of the HRP only leverages the correlation structures present within the covariance matrix making it even further removed from estimation error as stated in Sec.2.1.3. Therefore, it is not as reliant on the covariance matrix as the MV approach as it is not specifically using the empirical values housed in the matrix.

These estimation errors are further exacerbated by the assumptions underlying the returns distributions used in the calculations. Under the LLN, greater accuracy in estimation requires a larger sample size. However, the HS-FP method effectively reduces the sample size by selectively assigning varying probabilities to historical data points, increasing estimation error—particularly for mean estimates due to their higher variance. Additionally, the SR, central to MV optimisation, relies on the assumption of *i.i.d* normally distributed returns. The HS-FP models, however, operate under the assumption of ergodicity, a weaker condition than strict *i.i.d*, as suggested by [85]. This discrepancy means that the SR-maximising MV optimisation may not perform optimally under the HS-FP framework because its underlying assumptions are not fully met.

Another critical factor is the optimisation mechanisms used for asset selection: MV focuses on SR maximisation, while HRP emphasises correlation diversification. MV's emphasis on maximising the SR can lead to portfolios that are highly sensitive to estimated returns and thus more volatile. This sensitivity may result in overexposure to certain assets based on historical return patterns, which might not hold under stressed or abnormal market conditions introduced in the HS-FP simulations. In contrast, HRP constructs portfolios based on the hierarchical clustering of assets according to their correlations, achieving diversification by minimising concentrated risk rather than simply maximising returns. This approach is particularly beneficial in markets like SA's, where concentrated sectors and correlation noise make asset-level return predictions unreliable.

In the HS-FP simulations, which introduce varying probabilities to historical events and stress-test for abnormal conditions, HRP's structure allows it to absorb shocks more effectively than MV. Flexible probabilities in HS-FP simulations can cause certain market events to have disproportionate impacts on portfolio returns, mimicking stressed market environments. HRP's ability to avoid concentration in highly correlated assets ensures that when HS-FP simulations increase the weight of abnormal conditions—such as extreme downturns or spikes in correlations—its allocations remain more stable. It is not overly ex-

posed to any single risk factor or asset that may be affected by these conditions.

Higher-frequency data, such as daily or hourly observations, could improve the framework's ability to identify shifts in market conditions, thereby enhancing the reliability of the forward-looking estimates. Additionally, the case study highlights the potential benefits of introducing a turnover constraint to MV. By curbing excessive portfolio adjustments, turnover constraints could mitigate MV's reactivity to input fluctuations, promoting more consistent performance. Given their widespread use in real-world portfolio management, turnover constraints offer a practical solution for improving MV's robustness under HS-FP.

In contrast, HRP demonstrated greater resilience within the HS-FP framework. Unlike MV, HRP's reliance on the covariance structure, made it less sensitive to estimation errors. Moreover, HRP's diversification-driven approach does not depend on explicit regime identification, allowing it to maintain stable performance even when the historical distribution was adjusted by the FPrs. This robustness underscores HRP's compatibility with stress-testing frameworks like HS-FP, especially in environments where identifying regime shifts is challenging or where non-normal return distributions exacerbate estimation issues.

Overall, these findings emphasise the importance of effective regime identification in forward-looking frameworks like HS-FP. For MV, the failure to accurately capture regime changes not only negates its theoretical advantages but also exacerbates its inherent vulnerabilities to estimation errors. Future research should prioritize methods for enhancing regime detection within HS-FP, including higher-frequency data and advanced probabilistic modelling.

Additionally, adapting MV to incorporate practical constraints, such as turnover limits, could further stabilise its performance and better align it with the demands of dynamic market environments. In contrast, HRP's robustness under HS-FP highlights its suitability for scenarios where regime identification remains imprecise, reaffirming the value of diversification-focused strategies in uncertain conditions.

From comparison 3—comparing the state signal's relative performance within methods—revealed that no single signal consistently outperforms across methods. However, SA Money Supply showed relatively stable performance for both MV and HRP portfolios, suggesting it may warrant further exploration. Notably, the YoY change in CPI as a state signal performed poorly for MV optimisation but worked relatively well for HRP. This disparity appears to stem from timing mismatches in data and market responses, particularly the lag in CPI data relative to repo rate changes. The findings highlight the need for aligning data timing with market events and exploring timely indicators to improve MV's performance.

Chapter 4

Concluding remarks

We began with Vera Rubin’s statement: “*Science progresses best when observations force us to alter our preconceptions.*” MPT, however, is built on assumptions that often fail to reflect the complexities observed in real-world markets (Sec.1.2). This dissertation set out to address the limitations of traditional portfolio optimisation and back-testing methods by reducing key assumptions and introducing a scenario-based, non-parametric approach utilising the HS-FP framework. Conducted within a South African context, the analysis leveraged JSE equity and bond indices to assess the performance of MV optimisation and HRP, alongside relevant benchmarks. The primary objectives were to evaluate the relative strengths of these two portfolio construction methods and to determine whether incorporating forward-looking estimates using the HS-FP framework could enhance their effectiveness or not.

Objective 1: Assessing MV *vs.* HRP

The analysis demonstrated that HRP’s reduced reliance on parameter estimation provides notable benefits, particularly in highly volatile environments like the South African market. While the SR-maximising MV portfolio outperformed HRP in their cumulative performance curves, the HRP excelled on a risk-adjusted basis as reflected by the annualised SR. This suggests that HRP’s ability to achieve correlation-based diversification offers a more robust defence against extreme market fluctuations than the return-focused approach of MV optimisation. These results underscore the importance of prioritising diversification, particularly in markets with significant structural shifts, as illustrated by the extreme volatility observed between 2010 and 2023.

Objective 2: Evaluating the HS-FP Framework

The incorporation of the HS-FP framework provided a mixed but insightful contribution. While forward-looking estimates of risk and return did not consistently enhance overall performance, they highlighted the sensitivity of portfolio outcomes to the assumptions embedded in traditional methods. The regime-aware, state-conditioned kernel smoothing approach offered the most consistent improvements by stabilising portfolio adjustments and increasing responsiveness to market changes. These findings affirm the value of integrating market regime considerations into portfolio stress-testing frameworks, even if immediate performance gains are not always apparent.

Broader Implications

This work stresses the importance of revisiting the assumptions underpinning quantitative portfolio con-

struction methods, particularly the reliance on historical data as a proxy for future performance. By challenging these assumptions, the dissertation advocates for more adaptive frameworks that are better equipped to handle non-stationary, non-ergodic market data. The HS-FP framework presented a compelling alternative to the conventional approaches, but its effectiveness was strongly contingent on accurate regime identification. Its implementation in this study yielded mixed results, but the framework showcased considerable potential for improving portfolio stress-testing under atypical market conditions. Further refinement and development of methods to accurately capture market regimes could provide valuable advancements in aligning portfolio construction processes with the complexities of evolving financial markets.

Through the development of a comprehensive back-testing framework, the research also contributes practical tools and methodologies that can be applied beyond this study. Below, we present a chapter-by-chapter breakdown, provided to outline the progression of the analysis and the key findings that underpin these conclusions.

What was achieved

Chapter 1

I critically examined the foundational assumptions of MPT, particularly its heavy reliance on historical financial time series data. I emphasised that its key assumptions—normality, stationarity, and ergodicity—often fail in real-world scenarios, especially during periods of severe market stress or structural shifts, such as those illustrated in Fig.1.2 (the Dotcom crash, the GFC, and the COVID-19 pandemic). Common portfolio construction methods remain constrained by these assumptions and are also susceptible to the “selection problem.” Using Fig.1.3, I demonstrated that a range of portfolio techniques vary in their effectiveness across different market states, highlighting how some portfolios adapt more successfully to changing conditions than others. Thus, there is a need for approaches that can leverage the information present in these market regime changes that are non-stationary, non-ergodic data, enabling investors to stress-test and evaluate potential portfolio P&L under atypical market circumstances.

For this reason, in Sec.1.3, I introduced Meucci [89]’s non-parametric approach to scenario analysis, “Historical Simulations with Flexible Probabilities,” as an alternative to traditional MPT models. This method provides a more adaptive framework for addressing the complexities and evolving dynamics of financial markets, breaking away from the rigid assumptions of conventional risk measures.

Chapter 2

In Chp.2 I built up the theory needed to run portfolio stress-tests of the MV and HRP optimisations, and dedicated time to building a robust back-testing framework using the conventional approach commonly implemented by investors. This framework inherently assumed that historical data serves as a reliable proxy for future data, effectively treating historical performance as indicative of future outcomes. The analysis focused on South African assets, including the FINDI, INDI, and RESI indices to represent the equity market dichotomy (Fig.2.2), and the ABLI for the bond market. Special attention was given to the treatment of the risk-free asset by removing volatility from the STEFI instrument to better reflect a

realistic cash asset, as shown in Fig.2.3b. Additionally, advanced performance metrics such as the PSR and DSR were introduced to provide a more nuanced evaluation of portfolio performance.

I compared the performance of MV optimisation with HRP to evaluate their respective strengths and weaknesses. To ensure a fair comparison of methods I applied the same constraints to both the MV and HRP portfolios: a 5% equality cash constraint and made use of Winsorised returns. These treatments were decided upon after applying various configurations of constraints (as summarised in Tab.2.3, and figures 2.7 and 2.8), to assess the real-world applicability of these methods.

The simulation results in Fig.2.11 demonstrated that, on a cumulative performance basis, the SR maximising MV portfolio outperformed the HRP. However, Tab.2.5 revealed that HRP outperformed the MV on a risk-adjusted basis. We noted that, given the extreme volatility in the SA market between 2010 and 2023, the results suggested that prioritising correlation diversification, as achieved by HRP, may be more effective than focusing solely on return expectations as the MV portfolio does. This finding highlighted that HRP's success in the conventional back-test was likely driven by its focus on diversification rather than reduced estimation inputs, challenging our initial hypothesis.

Chapter 3

I shifted the focus in this chapter to the non-parametric, scenario-based approach of the HS-FP framework. We conducted three comparisons as well as a case study to gain understanding of the HS-FP simulations and their results. From the comparison in Fig.3.10 we showed that incorporating forward-looking estimates of returns and covariances (HS-FP) into MV and HRP portfolios did not result in clear performance or reliability improvements, though it revealed the portfolios' sensitivity to potential regime shifts. Comparing MV and HRP methods in Fig.3.11, showed that HS-FP adjustments failed to significantly enhance MV's risk-adjusted returns, likely due to challenges in accurately identifying and integrating regime changes, while HRP remained more stable and robust by relying on correlation structures instead of volatile return estimates. Lastly in Fig.3.12, no single state signal consistently outperformed across portfolio methods, underscoring the importance of effective regime detection and highlighting the resilience of diversification-driven approaches like HRP in managing estimation errors and non-normal return distributions.

Potential future work

Sensitivity analysis

In Chp.3 we saw that each HS-FP method requires its own specific combination of parameters and/or state signals. A natural next step would be to conduct a sensitivity analysis where we would examine how varying parameters and combinations of state conditioning signals affect the performance metrics across different HS-FP methods. Ideally for any given configuration, we should be able to run the optimisations and get the performance metrics (for each optimisation type). In principle, we should be able to get the distribution of all the performance metrics. In practice, to achieve this a script could be used to loop over a grid of parameter options, enabling the simulation of multiple portfolios under different parameter combinations. The aim would be to generate a correlation matrix between the parameters and perfor-

mance benchmarks (*e.g.* PSR *etc.*) highlighting how each parameter impacts the metrics. By associating performance metrics with each simulation path, we should be able to gain insights into the sensitivity of the HS-FP methods to parameter variations. A similar approach was applied by Paskaramoorthy et al. [100] to assess IS performance for different hyper-parameter configurations.

A suggestion would be to ring-fence the analysis by assessing the signal-parameter configurations per method rather than attempt to have all three parts (*methods*, *signals* and *parameters*) moving at once when creating possible configurations to feed to back-test simulations. A significant reason as to why this should simplify the problem is that when moving to the multi-market state HS-FP methods, the configuration matrix would need to account for the change from single to multiple signals. By separating the methods it becomes possible to rather customise the configuration matrix for the multi-state case.

Below are proposed ranges/sets for the parameters given in Tab.3.2 that can be used to perform sensitivity analysis:

1. $\lambda \in [3 \text{ mnths}, 36 \text{ mnths}]$ (values should be rounded so that the rolling window falls on the index values. Can also not exceed T).
2. $\tau \in [3 \text{ mnths}, 36 \text{ mnths}]$
3. $z^* \in [Z_1, Z_T]$ for single Z state methods
4. $z^* \in \{\mu(Z); Z_T\}$ for each Z of a multi-state methods
5. $\alpha \in [0, 1]$
6. $h \in [0.05, 1]$ (needs to be > 0 to avoid a solution with NaN probabilities.)
7. $\gamma \in \{1, 2\}$ (1: Exponential kernel, 2: Gaussian kernel)

Creating a dynamic MV-HRP hybrid model

A dynamic hybrid model consisting of the HRP and MV methods may produce promising results when zooming out to view the market from a cyclical aspect. The objective of this would be to find a way to exploit the individual strengths of the two, taking advantage of the stability and resilience provided by the HRP in times of market turmoil, and to use the MV when the market regains stability. A possible approach can be to run two portfolios in parallel, asset allocating across them based on state variables such as CPI and real GDP, *etc.* As the risk in the market increases the exposure of the HRP portfolio will increase, decreasing the exposure of the MV portfolio. Likewise, as the risk begins to dissipate, the exposure of the MV begins to increase, decreasing that of the HRP. A full switch between the portfolios based on single or multiple state variables is another approach to consider, fully reallocating assets to the HRP or MV portfolio exclusively when market states change. This could be determined by fixed thresholds in the state variables, however, a complete switch of strategies could be hugely costly, thus we may want to steer clear of this.

An overlay approach would be far more malleable. We can overlay the HRP over the MV and decide how much of one to take over the other. HRP would be a hedging strategy that would allow for disinvestment from the MV when and if needed during volatility spikes. For this to be effective, we would need to be able to assess the probability of negative event occurring imminently, which can be assessed using hazard rates. This estimation enables TAA between the two portfolios. We want to take the HS-FP framework

and improve the MV to be able to form a hybrid approach with an HRP portfolio. It would need to be a dynamic portfolio that is responsive to changing market environments—being capable of anticipating shocks and consequently adjusting its exposure to a MV benchmark portfolio using HRP as a hedging strategy. The mixture ratio of the portfolios would be a function of the hazard rate. If we have a certain scenario in the HS-FP framework and we see that all state variables are now extreme, then that increases the probability of the hazard rate as the hazard rate equals the probability of being in the tails. For this problem we need Extreme Value Theory (EVT) to find out what can be considered to be extreme in a given scenario, but to know what is extreme you need to define what outliers are.

We could make use of the “Peaks-Over-Threshold” (POT) approach in conjunction with Generalised Pareto Distributions (GPD). If we were to take our feature from our data and plot mean excess curves, we would need to establish at what point the curves become linear. The point at which it becomes linear is the point threshold u for that data set that indicates that whatever point exceeds u it’s extreme [26]. This is based on the Mean-Excess Function: $F_u(x)$. The moment the $F_u(x)$ begins to converge and one can consistently fit a GPD to it, we know that it is at an extreme. The sample of data points that exceed our u will be uniformly close to that of the GPD from a distribution with a scale parameter $\lambda(\hat{u})$ and shape parameter γ if the Gnedenko-Picklands-Balkema-de Haan Theorem holds. With a known distribution, we no longer need to pick a VaR relative to a normal distribution as you know the extreme value distributions will tend towards the GPD. We can now use the VaR as a decision criteria when deciding on the appropriate mix of the blended portfolio [37, 79].

Incorporating a custom bond index

To enhance the robustness and realism of the portfolio optimisation framework, we can extend the methodology by modelling bonds through the creation of a custom bond index using maturity sectors to allow us to better capture the actual instruments we intend to trade. By building our own universe of bonds and categorising them into maturity buckets (*e.g.*, short-end of the yield curve, middle, and long-end of the curve), we can build an index that better reflects the dynamic nature of bonds as they progress toward maturity and transition between sectors. By using this approach we aim to address the limitations of traditional portfolio optimisation methods when integrating bonds.

As seen in the simulations conducted in both Chp.2 and Chp.3, the HRP tend to overweight the ALBI due to its lower volatility compared to equities, leading to optimal portfolios that may under represent the FINDI, INDI, and RESI—potentially miss out on higher returns. By incorporating multiple bond sectors into the HRP framework, we hope to achieve a more balanced allocation. The varying volatility profiles across different bond maturities — higher volatility at the short and long ends and lower in the intermediate maturities — should allow for more nuanced allocations that can reduce the overall bond weighting relative to equities. We would hope that this approach not only mitigates over-concentration in bonds but would also leverage the downside protection that bonds offer without compromising potential equity gains .

To be able to build a bespoke bond index the structure of the index needs to be establish to accurately capture the movement of bonds within the index as well as in and out of the index. Two aspects to keep in mind when conceptualising the bond movements in the index is 1.) *How bonds flow in & out:* if it is a market cap index or based on optimisation objectives; and 2.) *How bonds move within the index:* bonds

shift between maturity bins overtime as their maturity dates approach and pass. Fig.4.1 is a simplified representation of how bond instruments flows through maturity bins within a generic bond index.

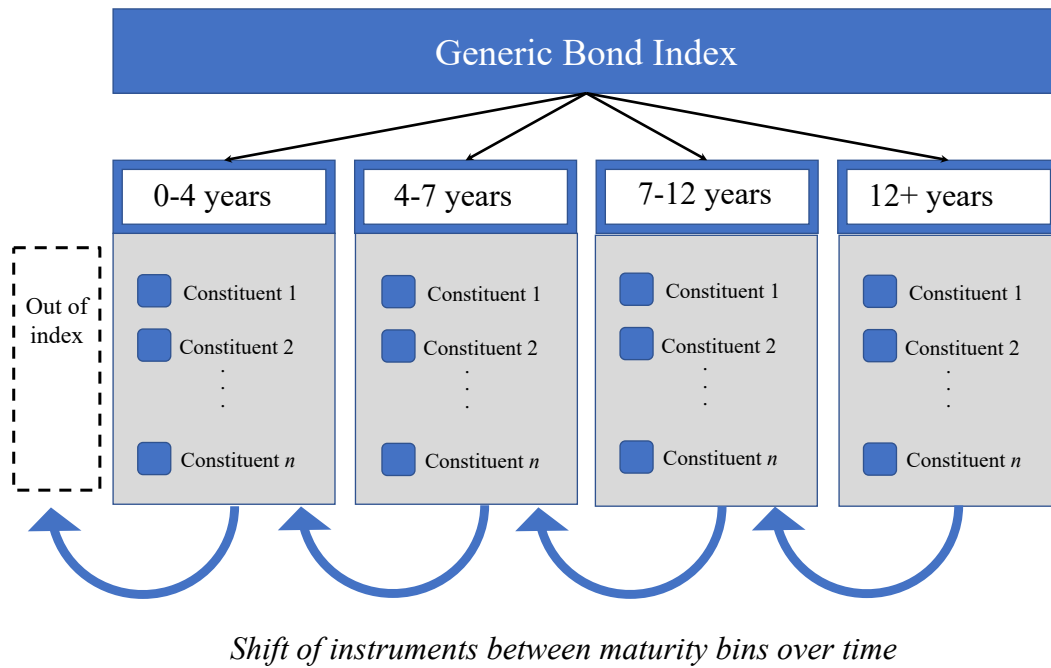


Figure 4.1: A simplified representation of how bond instruments flows through maturity bins within a generic bond index. As maturities approach the bonds move into bins with lesser maturities until the mature and fall out of the index.

The idea would be to use the ALBI to get the build a reference model for the bespoke index. We would want to be able to keep track of the shifting bonds from the different bins in order to recreate the index over time. Once that system work and we can recreate the index reliably then we can use the model for our own choice of bonds.

Appendix

A1. Algorithms

A1.1 Double exponential decay

This algorithm is used in Sec.3.3.2 to transform state variables such as inflation into state signals. Applying double exponential smoothing and standardising a series requires calculating sample moments for standardisation. If we do not want to make distributional assumptions we can use the approach discussed by Meucci [85]. Building on methodology used in BARRA [92] he shows that the process of using traditional exponential smoothing as an estimation technique that accounts for time-changing volatility can be simplified. The need of distributional assumptions and Monte Carlo scenarios can be circumvented by applying flexible probabilities, p_t^{exp} , generated by an exponential decay function to a panel of historical scenarios:

$$p_t(T, \lambda) \equiv p_t^{\text{exp}} \propto e^{-\lambda(T-t)}, \quad (1)$$

where λ is a user-defined decay coefficient, $t = 1$ is the oldest observed value and T is the most recent observation¹. Updating Eq.1 with a time constant allows for a user defined half-life τ (e.g a 3 month half-life) rather than a decay coefficient [89] such that:

$$p_t(T, \tau) \equiv p_t^{\text{exp}} \propto e^{-\frac{\ln 2}{\tau}|t-T|}. \quad (2)$$

This will allow us to specify a level (fast) half-life and a trend (slow) half-life later on as usually used in Double Exponential Moving Average (DEMA).

The generic formula for $\hat{\mu}$ and $\hat{\Sigma}$ under the Meucci framework as discussed in [85] will be to calculate the smoothed series. We will use Meucci's definition of $\hat{\mu}$ as the weighted average defined by:

$$\hat{\mu}_n := \sum_{t=1}^T p_t x_{t,n}, \quad n = 1, \dots, N \quad (3)$$

$$\hat{\Sigma}_{m,n} := \sum_{t=1}^T p_t x_{t,m} x_{t,n} - \hat{\mu}_m \hat{\mu}_n, \quad 1 \leq m \leq n = 1, \dots, N. \quad (4)$$

Here, $\{x_{t,n}\}$ will denote the historical scenarios of the n -th raw state variable time series and p_t as calculated in Eq.2 using the appropriate τ .

¹The notation of $p_t(T, \lambda)$ is taken to mean the flexible probabilities p_t given the specified parameters, in this case T, λ .

If we were to combine Eq.2, 3 and 4 with the notion of applying a double decay approach discussed in [85] we can employ the same idea that the traditional Holt's method makes use of by introducing two smoothing parameters to account for the level and trend. Conventionally α is used for the level (fast) half-life and β for the trend (slow) half-life [27, 48]. α targets the short term volatility component of the series while β targets the longer trend and correlation component.

Given a set of raw state variables $\{S_{n,t} \equiv (S_{1,t}, \dots, S_{N,t})\}$ we will apply an initial exponential decay using $\tau = \alpha$ to calculate the first set of flexible probabilities $p_t^{\text{exp}}(\alpha)$ given our fast half-life α . When $p_t^{\text{exp}}(\alpha)$ is applied to n -th *state variable* $S_{n,t}$ using the weighted average operator in Eq.3 the result is the initial n -th *signal* time series of $\hat{\mu}_n(\alpha)$. We will then calculate a new set of $p_t^{\text{exp}}(\beta)$ based on the slow half-life $\tau = \beta$ and apply it to $\hat{\mu}_n(\alpha)$ resulting in the double decayed signal series $\hat{\mu}_n(\alpha, \beta)$. Lastly, we will standardise $\hat{\mu}_n(\alpha)$ using the moments of $\hat{\mu}_n(\alpha, \beta)$ to evaluating the significance of short-term variations within the context of the long-term trend. Note that when $\propto$ is used it implies a proportionality constant is applied such that the probabilities sum to 1.

Step 1: Decay level component using $\tau = \alpha$

1.1 Calculate $p_t^{\text{exp}}(\alpha)$

$$p_t^{\text{exp}}(\alpha) \propto e^{-\frac{\ln 2}{\alpha}|t-T|}. \quad (5)$$

1.2 Estimate the level smoothed signal $\hat{\mu}_n(\alpha)$ using the weighted average operator in Eq.3:

$$\hat{\mu}_n(\alpha) = \sum_{t=1}^T p_t^{\text{exp}}(\alpha) \cdot S_{n,t} \quad (6)$$

This adjusts the series for the short term volatility. Next we will adjust for the trend using β .

Step 2: Decay trend component using $\tau = \beta$

2.1 Calculate $p_t^{\text{exp}}(\beta)$

$$p_t^{\text{exp}}(\beta) \propto e^{-\frac{\ln 2}{\beta}|t-T|}. \quad (7)$$

2.2 Apply $p_t^{\text{exp}}(\beta)$ to the $\hat{\mu}_n(\alpha)$ for the double decayed series

$$\hat{\mu}_n(\alpha, \beta) = \sum_{t=1}^T p_t^{\text{exp}}(\beta) \cdot \hat{\mu}_n(\alpha) \quad (8)$$

Step 3: Calculate the remaining moment and standardise

3.1 To calculate $\hat{\Sigma}_n(\alpha, \beta)$ needed for the standardising we can rewrite Eq.4 as

$$\hat{\Sigma}_n(\alpha, \beta) = \sum_{t=1}^T p_t^{\text{exp}}(\beta) \cdot (\hat{\mu}_n(\alpha))^2 - (\hat{\mu}_n(\alpha, \beta))^2. \quad (9)$$

3.2 Standardise

$$z_n(\alpha, \beta) = \frac{\hat{\mu}_n(\alpha) - \hat{\mu}_n(\alpha, \beta)}{\hat{\Sigma}_n(\alpha, \beta)} \quad (10)$$

By *smoothing* $\hat{\mu}_n(\alpha)$ using $p_t^{\text{exp}}(\beta)$ we have similarities with the DEMA approach in that a slower half-life is used to smooth a series already smoothed by a fast half-life. By *standardising* $\hat{\mu}_n(\alpha)$ with $\hat{\mu}_n(\alpha, \beta)$ and $\hat{\Sigma}_n(\alpha, \beta)$ we achieve a similar goal to the Holt's method as we end with a series adjusted for both level (α) and trend (β).

A2. Proofs

A2.1 Finite-ensemble average and the expected values

Given a discrete random variable z that can take on values z_j , if z is observed across an ensemble of N independent instances, where n_j represents the number of times the value z_j occurs within this ensemble, then the finite-ensemble average is given by:

$$\mathbb{E}[z(t)]_N = \frac{1}{N} \sum_{i=1}^N z_i = \sum_j \frac{n_j}{N} z_j, \quad (11)$$

where i is the individual observations within the ensemble, j represents each possible value of z , z_i is the i -th observed value, and $\frac{n_j}{N}$ is the fraction of times z_j occurs in the ensemble.

As $N \rightarrow \infty$, $\frac{n_j}{N}$ will converge to the probability p_j of observing z_j :

$$\lim_{N \rightarrow \infty} \frac{n_j}{N} = p_j. \quad (12)$$

This allows us to formulate the large-ensemble limit of the finite-ensemble average as:

$$\lim_{N \rightarrow \infty} \mathbb{E}[z(t)]_N = \lim_{N \rightarrow \infty} \sum_j \frac{n_j}{N} z_j = \sum_j p_j z_j. \quad (13)$$

This shows that:

$$\mathbb{E}[z] = \lim_{N \rightarrow \infty} \mathbb{E}[z(t)]_N = \sum_j p_j z_j. \quad (14)$$

where $\mathbb{E}[z]$ is the expectation value defined as the limit of the finite-ensemble average and $\sum_j p_j z_j$ is the expectation value defined as a weighted sum over all possible values of z . This demonstrates that the two definitions of the expectation value are equivalent as utilised in Def.2 and the ensemble average $\mathbb{E}[z]$ is in fact the large-ensemble limit of the finite-ensemble average $\mathbb{E}[z(t)]_N$.

Note that here we use the terms “ensemble average” and “expectation value” interchangeably for $N \rightarrow \infty$, and specifically refer to $\mathbb{E}[z(t)]_N$ as the “finite-ensemble average” when N is finite.

A2.2 Sharpe Ratio–Maximising Mean–Variance Portfolio with a budget constraint

Theorem 9 (Sharpe Ratio–Maximising Mean–Variance Portfolio with a budget constraint). Let Σ be symmetric and positive definite, and let $E[\mathbf{R}]$ be the vector of expected returns. Define the vector of expected excess returns as $\mathbf{x} = E[\mathbf{R}] - R_f \mathbf{1}$. The portfolio that maximises the SR:

$$\max_{\omega} \left(\frac{\mathbb{E}[R_P] - R_f}{\sigma_P} \right) = \max_{\omega} \left(\frac{\omega' \mathbb{E}[\mathbf{R}] - R_f}{\sqrt{\omega' \Sigma \omega}} \right) \quad (15)$$

is proportional to $\Sigma^{-1} \mathbf{x}$. After imposing the budget constraint $\omega' \mathbf{1} = 1$, the optimal weights are

$$\omega^* = \frac{\Sigma^{-1} \mathbf{x}}{\mathbf{1}' \Sigma^{-1} \mathbf{x}}. \quad (16)$$

The maximum SR achieved is

$$SR_{\max} = \sqrt{\mathbf{x}' \Sigma^{-1} \mathbf{x}}. \quad (17)$$

Proof. Define

$$S(\omega) = \frac{\omega' \mathbf{x}}{\sqrt{\omega' \Sigma \omega}}, \quad (18)$$

where $\mathbf{x} = E[\mathbf{R}] - R_f \mathbf{1}$. Since $S(\omega)$ is homogeneous of degree zero, maximising $S(\omega)$ is equivalent to maximising its square, $\tilde{S}(\omega) = S(\omega)^2$, without any initial constraint on the sum of weights.

We can therefore maximise:

$$\tilde{S}(\omega) = \frac{(\omega' \mathbf{x})^2}{\omega' \Sigma \omega}. \quad (19)$$

To find the optimal ω , we take the gradient with respect to ω and set it to the zero vector.

Step 1: Obtain the gradient of Eq.19 with respect to ω :

To obtain the gradient of Eq.19 with respect to ω we can use the quotient rule for vector calculus, $f(\omega) = \frac{g(\omega)}{h(\omega)}$:

$$\nabla_{\omega} f(\omega) = \frac{h(\omega) \nabla_{\omega} g(\omega) - g(\omega) \nabla_{\omega} h(\omega)}{h(\omega)^2}, \quad (20)$$

and set $g(\omega) = (\omega' \mathbf{x})^2$ and $h(\omega) = \omega' \Sigma \omega$.

To find the partial derivatives of $g(\omega)$ and $h(\omega)$:

1. The partial derivatives of $g(\omega)$ w.r.t ω can be written as $\frac{\partial}{\partial \omega} (\omega' \mathbf{x})^2$. To obtain the partial, let $N = \omega' \mathbf{x}$. Then $g(\omega) = N^2$.

Using the chain rule,

$$\frac{\partial N^2}{\partial \omega} = 2N \frac{\partial N}{\partial \omega}. \quad (21)$$

Since $\frac{\partial(\omega^\top \mathbf{x})}{\partial \omega} = \mathbf{x}$, we have:

$$\frac{\partial g(\omega)}{\partial \omega} = 2(\omega^\top \mathbf{x})\mathbf{x} \quad (22)$$

2. The partial derivatives of $h(\omega)$ w.r.t ω can be written as $\frac{\partial}{\partial \omega}(\omega^\top \Sigma \omega)$.

Since Σ is a symmetric matrix, we know that:

$$\frac{\partial(\omega^\top \Sigma \omega)}{\partial \omega} = 2\Sigma\omega. \quad (23)$$

We can then substitute the above into the gradient of $\tilde{S}(\omega)$, taking the form of Eq.20:

$$\frac{\partial \tilde{S}(\omega)}{\partial \omega} = \frac{(\omega^\top \Sigma \omega) [2(\omega^\top \mathbf{x})\mathbf{x}] - (\omega^\top \mathbf{x})^2 [2\Sigma\omega]}{(\omega^\top \Sigma \omega)^2}. \quad (24)$$

Step 2: Set the gradient to the zero vector:

$$\frac{(\omega^\top \Sigma \omega) [2(\omega^\top \mathbf{x})\mathbf{x}] - (\omega^\top \mathbf{x})^2 [2\Sigma\omega]}{(\omega^\top \Sigma \omega)^2} = \mathbf{0}. \quad (25)$$

Assuming $\omega^\top \mathbf{x} \neq 0$ (i.e., the expected excess return is not zero) and $\omega^\top \Sigma \omega \neq 0$ (i.e., the portfolio variance is not zero), we can multiply both sides by $(\omega^\top \Sigma \omega)^2$ and divide by $2(\omega^\top \mathbf{x})$:

$$(\omega^\top \Sigma \omega)\mathbf{x} - (\omega^\top \mathbf{x})\Sigma\omega = \mathbf{0}. \quad (26)$$

Rearranging, we get:

$$(\omega^\top \Sigma \omega)\mathbf{x} = (\omega^\top \mathbf{x})\Sigma\omega. \quad (27)$$

This implies that $\Sigma\omega$ must be proportional to $\mathbf{x}$:

$$\Sigma\omega \propto \mathbf{x}. \quad (28)$$

Since Σ is positive definite, Σ^{-1} exists. Multiplying by Σ^{-1} from the left:

$$\omega \propto \Sigma^{-1}\mathbf{x}. \quad (29)$$

Let $\omega_{\text{unscaled}} = \Sigma^{-1}\mathbf{x}$ be the optimal **unscaled** portfolio.

It follows, to impose the budget constraint $\omega^\top \mathbf{1} = 1$, we scale ω_{unscaled} such that its components sum to one.

Let $\omega^* = c\omega_{\text{unscaled}}$ for some scalar c .

$$(\omega^*)^\top \mathbf{1} = (c\Sigma^{-1}\mathbf{x})^\top \mathbf{1} = c(\mathbf{1}^\top \Sigma^{-1}\mathbf{x}) = 1. \quad (30)$$

Therefore, the scaling factor is $c = \frac{1}{\mathbf{1}^\top \Sigma^{-1}\mathbf{x}}$.

Thus, the optimal weights satisfying the budget constraint are:

$$\omega^* = \frac{\Sigma^{-1}\mathbf{x}}{\mathbf{1}^\top \Sigma^{-1}\mathbf{x}}. \quad (31)$$

Finally, to find the maximum SR, we substitute $\omega = k\Sigma^{-1}\mathbf{x}$ (where k is any scalar, as the SR is invariant to scaling) into the SR expression. Using $k = 1$ for simplicity:

$$SR_{\max} = \frac{(\Sigma^{-1}\mathbf{x})^\top \mathbf{x}}{\sqrt{(\Sigma^{-1}\mathbf{x})^\top \Sigma (\Sigma^{-1}\mathbf{x})}}. \quad (32)$$

The numerator is $(\Sigma^{-1}\mathbf{x})^\top \mathbf{x} = \mathbf{x}^\top (\Sigma^{-1})^\top \mathbf{x} = \mathbf{x}^\top \Sigma^{-1}\mathbf{x}$ (since Σ^{-1} is symmetric if Σ is). The denominator is $\sqrt{\mathbf{x}^\top (\Sigma^{-1})^\top \Sigma \Sigma^{-1}\mathbf{x}} = \sqrt{\mathbf{x}^\top \Sigma^{-1}\Sigma \Sigma^{-1}\mathbf{x}} = \sqrt{\mathbf{x}^\top I \Sigma^{-1}\mathbf{x}} = \sqrt{\mathbf{x}^\top \Sigma^{-1}\mathbf{x}}$.

Therefore the corresponding maximum SR is:

$$SR_{\max} = \frac{\mathbf{x}^\top \Sigma^{-1}\mathbf{x}}{\sqrt{\mathbf{x}^\top \Sigma^{-1}\mathbf{x}}} = \sqrt{\mathbf{x}^\top \Sigma^{-1}\mathbf{x}}. \quad (33)$$

□

Remark 4. It is important to note that the closed-form solution in Theorem 16 relies on the assumption that short selling is permitted. This is because the inverse covariance matrix Σ^{-1} applied to the excess return vector $\mathbf{x}$ may produce negative weights. When a no-short-selling constraint ($\omega \geq 0$) is imposed, the problem becomes a quadratic program with inequality constraints, for which no general closed-form solution exists. In such cases, numerical optimisation techniques (e.g., quadratic programming or interior-point methods) must be employed to obtain the optimal weights.

A2.3 Inverse-variance allocation

Eq.2.14 for the split factor (weighting factor) used under *Step 3: Recursive Bisection* of the HRP algorithm can be shown to be optimal with a diagonalised covariance matrix. The split factor allows a weight of a parent cluster to be split inversely proportional to a sub-cluster's variance. This is the cases as it stems from the classical portfolio optimisation theory for a standard quadratic optimisation problem of size N :

$$\min_{\omega} \{\omega' \Sigma \omega\} \quad \text{s.t. } \omega' \mathbf{1} = 1, \quad (34)$$

where $\mathbf{1}$ is a vector consisting of 1's with length N . The solution of Eq.34 yields the minimum variance portfolio with the optimal weights:

$$\omega^* = \frac{\Sigma^{-1} \mathbf{1}}{\mathbf{1}' \Sigma^{-1} \mathbf{1}} \quad (35)$$

As we work with the quasi-diagonalised matrix generated in “Step 2: Quasi-diagonalization” the n -th weight can be written as:

$$\omega_n = \frac{\Sigma_{n,n}^{-1}}{\sum_{i=1}^N \Sigma_{i,i}^{-1}}. \quad (36)$$

Considering we are only interested in 2 sets of assets we have $N = 2$, making:

$$\frac{\frac{1}{\Sigma_{1,1}}}{\frac{1}{\Sigma_{1,1}} + \frac{1}{\Sigma_{2,2}}} = 1 - \frac{\Sigma_{1,1}}{\Sigma_{1,1} + \Sigma_{2,2}}. \quad (37)$$

Eq.37 now equates to Eq.2.14, the splitting factor that partitions the weight between the 2 sub-clusters.

2

²This theory was initially written up for use in my Honours thesis at the University of Cape Town [76].

A3. Additional simulation designs

A3.1 Benchmark portfolios for overlapping rolling window simulations

In Chp.1, Sec.1.1.3.2, Fig.1.4 an overlapping rolling window portfolio stress-test was used to compare the performance of various approaches that aim to balance the risk-reward trade-off throughout a SA business cycle. Fig.1.4 was generated as part of an honours-level thesis by myself, alongside Siphesihle Cele, for the degree of Honours in Statistics in the Department of Statistical Sciences at UCT [76]. Tab.1 below contains the benchmark portfolios used to assess the relative performance of an HRP portfolio within the SA market context.

Table 1: The 4 benchmark portfolios used in the rolling (Fig.1.4) and the growing window (Fig.2) dynamic back-tests are seen below. Each benchmark is listed with its base portfolio type indicating the algorithm that underpins the portfolio strategy. The assets that are available for selection by each portfolio is recorded under “Investment Universe”. The table also records the rebalancing strategy that is applied to each actively managed portfolio on a monthly basis, this of course excludes the Buy-and-Hold strategy as it employs passive management.

#	Base Portfolio Type	Investment Universe	Rebalancing strategy
1.	Equally weighted	All industry indices and the ALBI.	CM
2.	SR maximising	All industry indices and the ALBI.	Optimised monthly
3.	SR maximising	All industry indices and the ALBI.	Buy-Hold
4.	Balanced Portfolio	ALSI and ALBI	60 : 40 CM

The first benchmark, Benchmark 1, is a straightforward equally weighted CM portfolio, where all N assets are rebalanced to $1/N$ each month. Benchmarks 2 and 3 are based on the Markowitz [69] MV optimisation to maximise the SR [119] as defined in Eq.2.5. For both Benchmark 2 and 3, the Short Term Fixed Interest (STEFI) index was used as the risk free rate where the mean is calculated over the full period of historical data used in the honours project. Benchmark 2 performs monthly rebalancing through optimisation, while Benchmark 3 uses a passive Buy-and-Hold strategy, holding the SR-maximising portfolio from month 1 without further adjustments throughout subsequent periods. Benchmark 4 is a 60:40 balanced fund portfolio between equity (ALSI) and bonds (ALBI), rebalanced monthly to maintain the 60:40 ratio as a CM.

A3.2 Comparative back-test: growing window simulation

Using the same investment universe, historical data, and benchmarks as in the simulation as seen in Fig.1.4, a growing window back-test can also be applied instead of the rolling window approach. Here, the IS window grows by a month at a time rather than maintaining the window length that shifts over the data. The results of this simulation are shown in Fig.2.³ While the growing window simulation demonstrates improved relative performance of the SR-maximising strategy, it also suggests potential overfitting. This is likely due to the expanding IS data window as the simulation progresses.

³Produced as part of my honours thesis at the University of Cape Town [76].

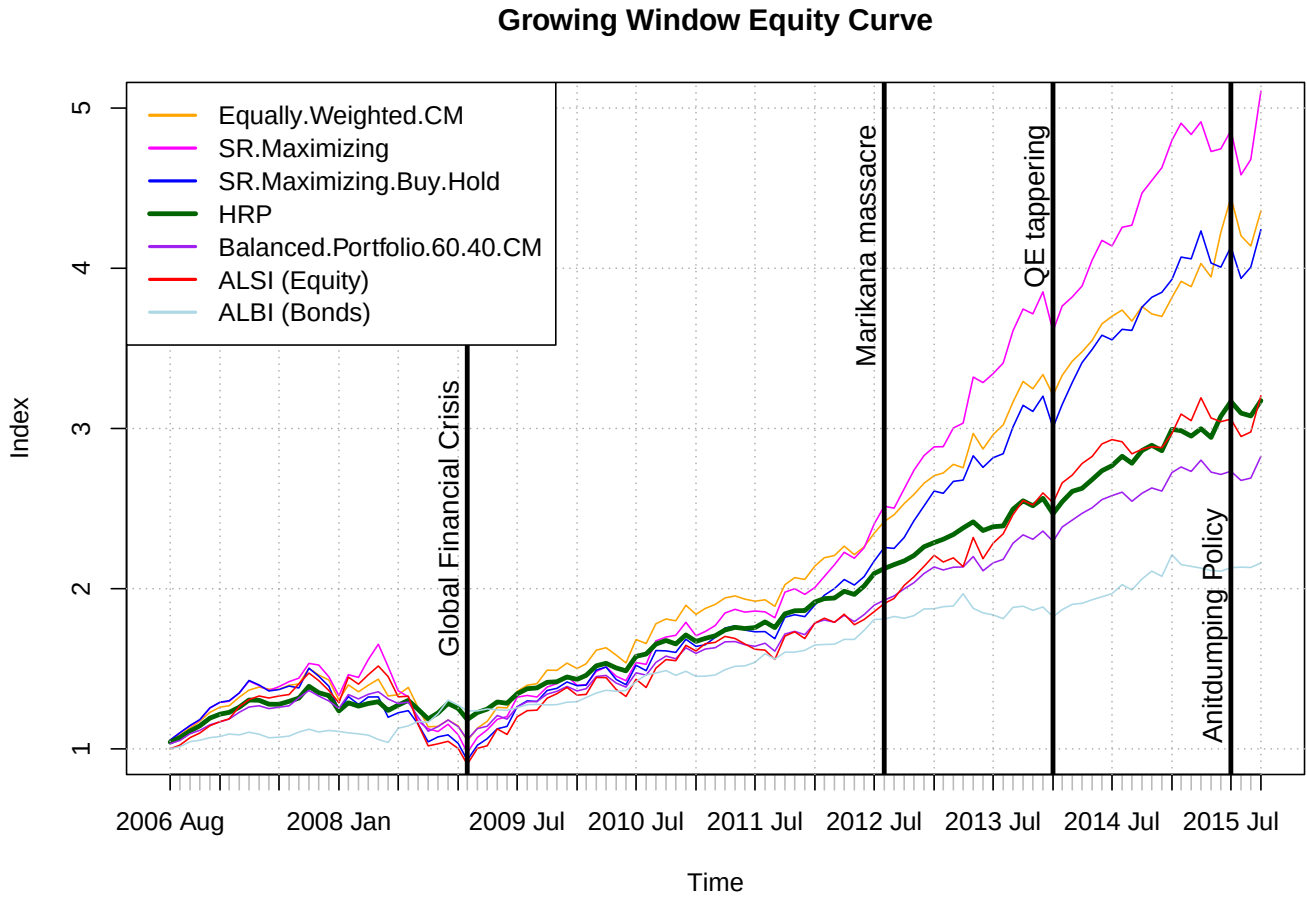
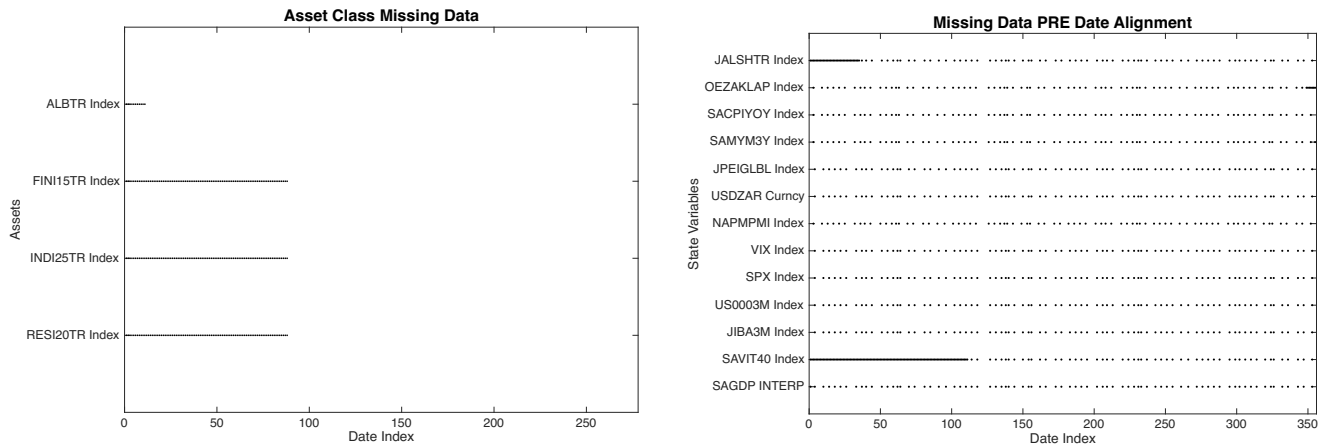


Figure 2: Equity curves over the OOS period spanning Aug'06 to Nov'15 for 5 different portfolio construction strategies along with the ALSI and the ALBI as additionally benchmarks. The curves were generated using a dynamic growing window back-test with the initial window consisting of 3 years of monthly data from Aug'03 to Aug'06. The vertical lines indicate significant economic and political events that had impacts on the local and global market during the time period. Refer to Tab.1 for more detail on the benchmark portfolios used within this simulation and their makeup. See Fig.1.4 for a comparison with the rolling window simulation.

A4. Data engineering

This section captures the EDA that was conducted on the data as well as the pre-processing that took place when preparing the data for use in the simulations. All data was obtained from Bloomberg. Refer to [A5.6 Sourcing data: a step-by-step guide to Bloomberg](#) for details on how to access the data through a Bloomberg terminal.

The state variable data required pre-processing to identify the cause of the inconsistencies and to produce usable data for the simulations. All highlighted state variables listed in Tab.3.1 were obtained with monthly frequencies; however, the SA economic growth rate (*SAGDPANN*) is only reported quarterly, and therefore it needed to be handled slightly differently. The tail of the pre-processed monthly signal data is shown below in Tab.2. The “Date” column in Tab.2 suggests that there is a misalignment in the time series of the variables as the intervals are not strictly monthly after the series were merged into a single timetable object in MATLAB. This confirms the inconsistency of the NA values seen throughout the full dataset in Fig.3b. The plot indicates all data points of the monthly signals that are recorded



(a) Total Returns for the 4 asset class variables the date index points are plotted if they contain missing data in the form of NA values. This is looked at over the time period of 29/02/2000 to 28/02/2023 at a monthly frequency. (b) For the 12 **monthly** state signal variables the date index points are plotted if they contain missing data in the form of NA values. This is over the time period of 29/02/2000 to 28/02/2023 at a monthly frequency.

Figure 3: The initial asset price data as well as state variable data was collected for a period of 23 years, where the start and end dates were kept consistent across data pulls as the 29/02/2000 - 28/02/2023. Upon inspection the asset price data showed limited historical data (as captured by Fig.3a) for the equity sectors being: the financials, industrials and resources sectors. While the state variable data displayed significant inconsistencies with missing data seen throughout the time series data (seen in Fig.3b). Missing values in the state variable data is due to a misalignment of the dates at which the observations are recorded on Bloomberg. Both figures generated by: `data_cleaning_1b.mlx`

as NA values in the State Signal timetable object over the time period 29/02/2000 to 28/02/2023 at a monthly frequency.

Table 2: Tail sample of the monthly state variable data used to assess the cause of the data inconsistency.

Dates	JALSHTR	OEZAKLAP	SACPIYOY	SAMYM3Y	JPEIGLBL	USDZAR Curncy
2022-08-31	11234.66	0.26	7.60	8.12	13029.14	17.12
2022-09-30	10770.73	NA	7.50	8.78	12923.98	18.09
2022-10-31	11297.40	NA	7.60	9.82	13143.78	18.35
2022-11-30	12690.82	NA	7.40	8.76	12988.71	17.20
2022-12-30	12403.84	NA	NA	NA	13090.64	17.04
2022-12-31	NA	NA	7.20	8.66	NA	NA
2023-01-31	13506.61	NA	6.90	9.57	13808.99	17.41
2023-02-28	13210.45	NA	7.00		14202.71	18.36
	NAPMPMI	VIX	SPX	US0003M	JIBA3M	SAVIT40
	52.90	442.75	67687.45	3.10	5.77	25.46
	51.00	571.49	64805.42	3.75	6.47	28.03
	50.00	475.33	71115.49	4.46	6.52	26.99
	49.00	349.25	69241.91	4.78	7.20	22.37
	NA	369.24	65421.24	4.77	7.26	21.68
	48.40	NA	NA	NA	NA	NA
	47.40	338.18	71064.11	4.81	7.47	19.62
	47.70	379.50	72786.35	4.97	7.45	21.08

The inconsistency seen in Tab.2 is due to Bloomberg's varying reporting conventions. Some tickers report only on business days (*e.g.*, JALSHTRI), while others include weekends (*e.g.*, OEZAKLAP, SACPIYOY). Therefore when we initially attempted to synchronise the data in MATLAB, the `synchronize()` function assigned NA values where data points were missing. For instance, if month-end falls on a Sunday, JALSHTRI reported Friday's data, while OEZAKLAP and SACPIYOY showed Sunday's values. Synchronizing these series introduces NA values for mismatched dates, as shown in Fig.4.

Dates	JALSHTR Index	OEZAKLAP Index	SACPIYOY Index	SAMYM3Y Index	JPEIGLBL Index	USDZAR Curncy	NAPMPMI Index	VIX Index	SPX Index	US0003M Index	JIBA3M Index	SAVIT40 Index
Wednesday, 30 September 2009	2 791,42	2,11	6,10	4,01	3 451,22	7,51	54,90	192,31	7 937,92	0,29	7,02	25,31
Friday, 30 October 2009	2 959,37	NA	NA	NA	3 615,00	7,82	NA	241,04	8 138,28	0,28	7,24	22,48
Saturday, 31 October 2009	NA	3,02	5,90	2,68	NA	NA	57,60	NA	NA	NA	NA	NA
Monday, 30 November 2009	3 022,43	3,86	5,80	0,59	3 451,63	7,40	55,40	181,90	8 131,22	0,26	7,23	23,58
Thursday, 31 December 2009	3 110,79	4,59	6,30	1,79	3 443,40	7,40	55,80	160,00	8 229,46	0,25	7,23	22,93
Friday, 29 January 2010	3 001,79	NA	NA	NA	3 543,26	7,63	NA	186,28	8 125,11	0,25	7,19	25,15
Sunday, 31 January 2010	NA	5,16	6,20	0,60	NA	NA	56,30	NA	NA	NA	NA	NA
Friday, 26 February 2010	3 012,99	NA	NA	NA	3 642,99	7,72	NA	149,58	8 472,12	0,25	7,19	26,01
Sunday, 28 February 2010	NA	5,55	5,70	0,12	NA	NA	55,50	NA	NA	NA	NA	NA
Wednesday, 31 March 2010	3 250,11	5,73	5,10	1,55	3 530,69	7,29	58,80	127,79	8 495,57	0,29	6,67	21,15
Friday, 30 April 2010	3 248,17	5,73	4,80	1,67	3 618,24	7,39	58,10	162,92	8 767,85	0,35	6,60	22,95
Monday, 31 May 2010	3 082,13	5,57	4,60	1,35	3 667,15	7,67	57,40	244,00	8 288,75	0,54	6,61	31,42
Wednesday, 30 June 2010	2 984,48	5,33	4,10	2,47	3 763,25	7,67	56,50	264,38	7 889,36	0,53	6,62	30,45
Friday, 30 July 2010	3 224,60	NA	NA	NA	3 736,20	7,30	NA	171,52	8 040,34	0,45	6,57	26,00
Saturday, 31 July 2010	NA	5,07	3,70	3,77	NA	NA	56,10	NA	NA	NA	NA	NA
Tuesday, 31 August 2010	3 109,27	4,82	3,50	4,39	3 867,51	7,37	56,40	192,34	7 747,70	0,30	6,38	27,33
Thursday, 30 September 2010	3 381,14	4,60	3,20	5,09	3 718,54	6,96	55,30	165,48	7 967,98	0,29	6,03	23,56
Friday, 29 October 2010	3 501,18	NA	NA	NA	3 779,00	7,00	NA	147,88	8 254,05	0,29	5,92	22,64
Sunday, 31 October 2010	NA	4,44	3,40	6,36	NA	NA	56,90	NA	NA	NA	NA	NA
Tuesday, 30 November 2010	3 485,37	4,30	3,60	7,23	3 719,60	7,10	57,30	166,94	8 372,25	0,30	5,55	23,54

Index Description	ALSI	Domestic Leading Indicator	SA CPI	SA Money Supply	JPM EMB Index	US/SA Exchange Rate	SA PMI	US Volatility Index	S&P 500	US 3Month LIBOR	SA 3Month JIBAR	SA Volatility Index
-------------------	------	----------------------------	--------	-----------------	---------------	---------------------	--------	---------------------	---------	-----------------	-----------------	---------------------

Figure 4: Sample of monthly state variable data showing the misalignment in NA values. The indices in blue are reported only on the last business day of each month while the indices in green are also reported over weekends. Therefore the green indices are always lagged by a day if the last day of the month falls over a weekend.

To fix the inconsistent dates in the timetable objects we shift the `timedate` array of the individual state signal variables before synchronizing:

```

1 for signal = 1:width(SIG_data)
2     SIG_data{1,signal}.Dates = dateshift(SIG_data{1,signal}.Dates, 'end',
3     , 'month');
4 end
5 SIG_aligned_data = SIG_data;
6 % Synchronize timetable objects
7 SIG_aligned_dataTable = synchronize(SIG_aligned_data{:});

```

After adjusting the array of dates prior to synchronising the series, we observe in Fig.5 that the final state variable time series shows a significant reduction in the number of NA values. This adjustment effectively aligns the reporting dates across the different signals.

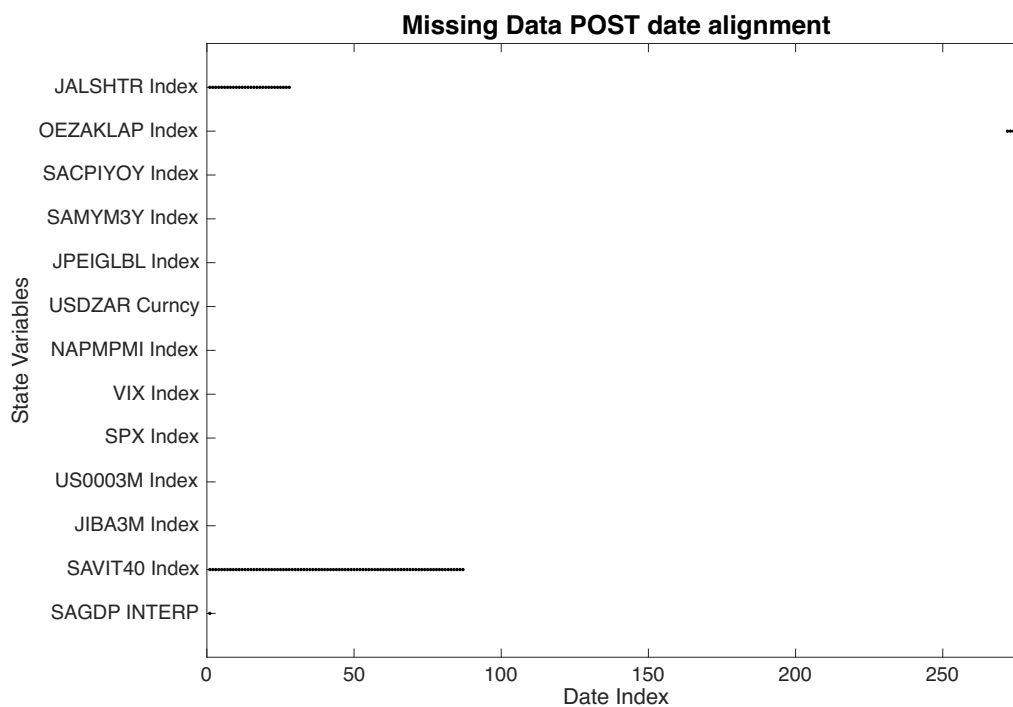


Figure 5: State variables with dates adjusted to ensure that all data points align with the last calendar day of each month, regardless of original reporting frequency received from Bloomberg. Generated by: `data_cleaning_1b.mlx`

A5. Reproducible research

This section provides a comprehensive guide to reproducing the results presented in this thesis. We outline the system requirements, detail the data management practices used, and provide an overview of the codebase developed for this project. It includes a user example of the code toolbox, the pre-processing steps, and descriptions of the methodologies and algorithms utilised. By following the step-by-step guide to accessing the data and utilizing the user example on how to implement the project's codebase, you

should be able to replicate the simulations generated in Chp.2 and Chp.3. This transparency ensures that others can verify this work and build upon it for future research.

A5.1 Systems specification

1. Hardware:

- (a) An Apple Silicon Mac with an M1 chip will be used, having 8GB of RAM with an 8-core CPU on MacOS version 14.1 (23B74).

2. Software:

- (a) Data is be collected from Bloomberg using the Excel Add-In. See [A5.6 Sourcing data: a step-by-step guide to Bloomberg](#) for data access instructions.
- (b) MATLAB (version R2023b, Update 3) was the primary programming language used for data manipulation and algorithm development. On Apple devices with M series chips, MATLAB R2023b runs natively on Apple silicon.

A5.2 Data management

We suggest incorporating a one-to-one mapping between the file path and its corresponding file name with the objective of enhancing user-friendliness. The data file naming convention as well as an illustrative example is shown below:

Naming convention: DATA-SOURCE-TYPE-DATA_LABEL-FRQY-START-END.extension,

where the **START** and **END** dates follow the ISO 8601 convention of YYYYMMDD.

Example: To store the data collected from Bloomberg, including signal data such as the USD_ZAR exchange rate between February 11, 2000, and March 10, 2023, with a recorded weekly frequency, the file is named using the nomenclature:

DATA-BB-SIG-USDZAR-CRNY-W-20000211-20230310.xlsx.

This standardised naming convention serves the dual purpose of facilitating swift user identification and aims to mitigate the risk of data misuse.

A5.3 Data structures

To perform data imports and create the necessary data structures:

1. Load the data from Excel as a MATLAB data structure, then convert it to a MATLAB `timetable` object.
2. Build the `timetable` object incrementally, variable by variable, to avoid loading all the data into memory at once (as this would be computationally expensive). We suggest extracting only the required data from each Excel file and then only concatenate it into the pre-defined `timetable` object.

For the `timetable` objects, we utilised MATLAB's built-in functionality to ensure that each dataset is well-documented by setting relevant properties. Data descriptions included:

1. Data type
2. Data frequency (*e.g.*, monthly)
3. Latest update date
4. Original source (*e.g.*, Bloomberg)
5. Processed status (yes/no)
6. Intended export file (*e.g.*, `DATA-PROCESSED-SIG-PX-LAST-M-TTABLE-20000229-20230228.mat`)

The `VariableDescriptions` field was used to elaborate on each variable in turn such that the descriptions somewhat correspond to those given in tables such as Tab.3.1.

A5.4 Toolbox construction

The code files fit together in such a way that they are able to support flow of raw data from a Bloomberg source through to the end simulations. The data and work flows are captured in Fig.6 below.

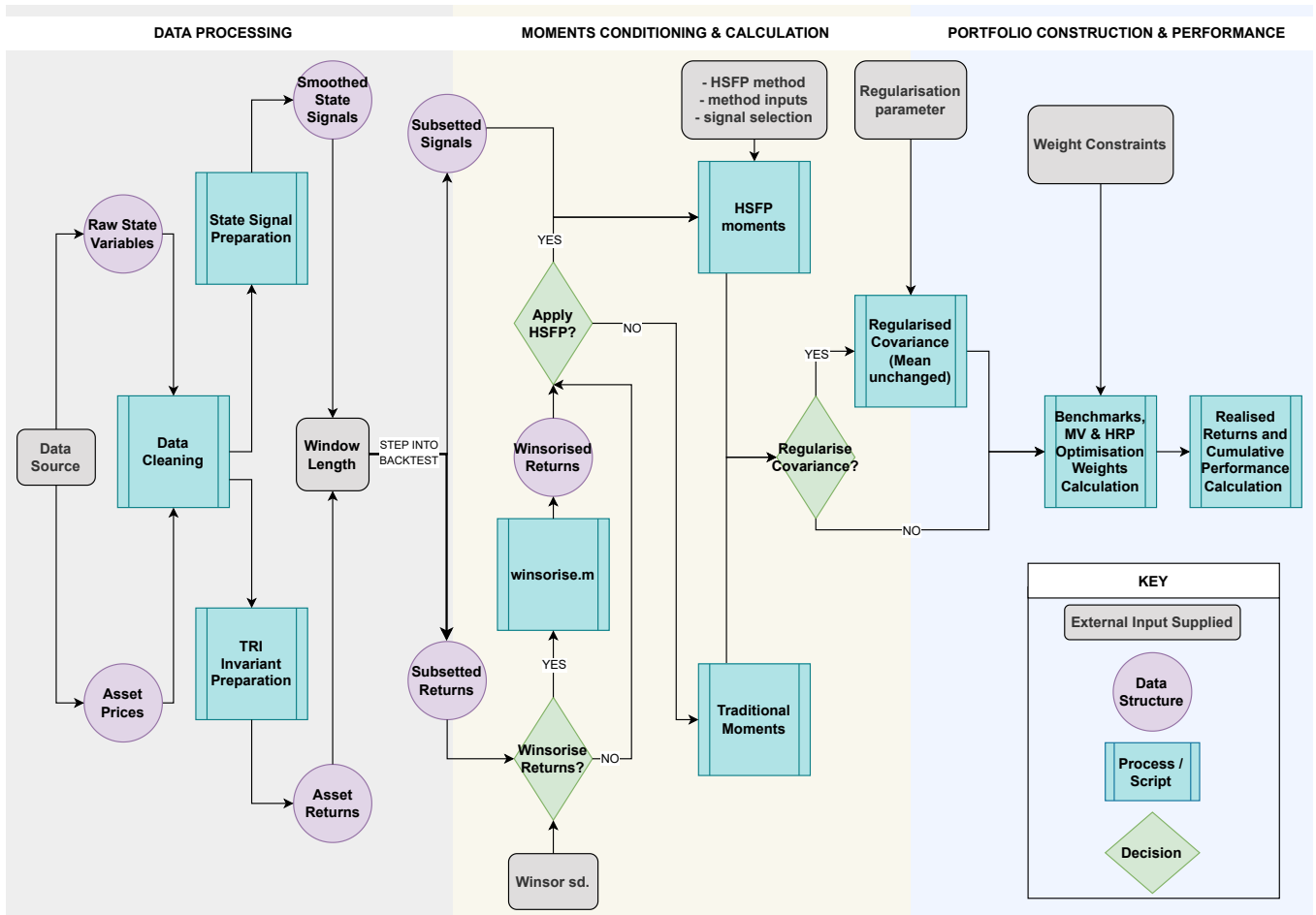


Figure 6: This diagram shows the flow of raw data from the data source such as Bloomberg to the final back-test output being the performance measure of the various portfolios and benchmarks. The progression through the full project workflow is shown including the signal processing and application of the HS-FP methods. To apply only the conventional approach used in Sec.2 to back-testing, exclusively follow the rightward (downward) arrows out of each node.

A5.4.1 Class construction

In this research, only two user-defined classes were created. The primary class utilised for the back-testing procedure is denoted as *backtestedPortfolios*, and an additional class, named *HSFPparameters*, was defined through the use of class definition files.

The following design consideration were taken into account when building the classes:

1. Method and/or property validations are defined to restrict possible values as well as define property attributes to control access and visibility.
2. A hierarchy of base classes and derived classes were created to separate specific definitions of multiple derived classes allowing for inherited as well as unique independently derived class properties. This allows for the use of organised modular code in the modelling of different variations of similar objects [120]. An example could be making a “benchmark” class a derived class of the back-testing class.
3. Different types of classes can also be considered. *E.g.* MATLAB Handle classes rather than value classes.
4. Enumeration makes the code more user-friendly, providing the ability to define an ordering for the values of a class, tab completion, as well as case insensitivity.

A5.4.2 Class properties

Below is the class definition for the *backtestedPortfolios* object class. This class definition needs to be used in conjunction with the *HSFPparameters* class.

backtestedPortfolios class definition:

```

1  classdef backtestedPortfolios
2      properties (SetAccess = private) % Can only change property through
        a method
3          PortfoliosList      cell          % cell array
4          AssetList           cell          % cell array
5          Attributions        table         % structure
6          RollingPerformance  timetable     % Timetable
7          HSFP_Prs            = []          % Empty for storage
8          NumPortfolios       (1,1) double  % scalar
9          DataType            (1,1) string = 'backtestedPortfolios'
10         ExcessReturns        double
11         GeometricSR          double
12         Benchmark_PRet      double
13         MaxDrawDowns        table         = []
14         OptPortWts           cell          % Cell array of TTs with dif
            dimensions
15         Realised_tsPRet_TT   timetable
16     end
17
18     properties % Able to be set by user when creating a class object

```

```

19
20 WindowLength      (1,1) double = 36
21 Returns           timetable
22 Invariants        timetable
23 Signals           timetable           % Timetable
24 RegLambda         (1,1) double = 0    % Default (no
    regularisation)
25 MaxDD_cal         (1,1) string {mustBeMember(MaxDD_cal,{'on',
    ...
    'off'}})} = 'off'    % Default (no max Drawdown calc)
26 % if MaxDD_cal set to 'on': calculate max drawdown (might fail
27 for
28 % certain cases sense can be turned off)
29 WinsorStd         double {mustBeNonnegative} = []    % if empty
    don't winsorise
30 MVWts_lb          (1,:) double         % vector of lower bounds
31 MVWts_ub          (1,:) double         % Vector of upper bounds
32 CashConstraint    (1,1) double {mustBeNonnegative} = 1 % Default
    no constraint
33 Method            (1,1) string {mustBeMember(Method,{'none',
    ...
    'rolling_w','exp_decay','crisp', 'kernel', ...
    'e_pooling', 'ew_ensemble', 'cb_ensemble'}})} = 'none'
34 HSFParameters     HSFParameters %structure/class
35 PSRBenchmark      (1,1) string {mustBeMember(PSRBenchmark,{'
    zeroSkill',...
    'strategyWise'}})} = 'zeroSkill'
36
37 end

```

HSFPparameter class definition:

```

1 % HSFPparameters class definition
2 classdef HSFPparameters
3     properties
4         RollWindow      (1,1) double {mustBePositive} = 12
5         Tau              (1,1) double {mustBePositive} = 12
6         Alpha            (1,1) double {mustBeGreaterThanOrEqual(Alpha
            ,0)} = 0.1
7         Z_target         {mustBeA(Z_target,['numeric', "
            string","char"])} = 'latest'
8         Gamma            (1,1) double {mustBePositive} = 2
9         h                (1,1) double {mustBePositive} = 0.5
10        Tau_prior        (1,1) double {mustBePositive} = 12
11        Ensemble_wt_method (1,1) string {mustBeMember(

```

```

12         Ensemble_wt_method,...
13         {'simple','log-linear'}}} = 'log-linear'
14
15     end
16
17     methods
18     %%%% CONSTRUCTOR FUNCTION %%%%%%%%%
19     function HSFPparameters = HSFPparameters(RollWindow,Tau,Alpha,
20         Z_target,Gamma,h,Tau_prior,Ensemble_wt_method)
21         if (nargin == 8)
22             HSFPparameters.RollWindow = RollWindow;
23             HSFPparameters.Tau = Tau;
24             HSFPparameters.Alpha = Alpha;
25             HSFPparameters.Z_target = Z_target;
26             HSFPparameters.Gamma = Gamma;
27             HSFPparameters.h = h;
28             HSFPparameters.Tau_prior = Tau_prior;
29             HSFPparameters.Ensemble_wt_method = Ensemble_wt_method;
30         elseif (nargin > 0)
31             error("You need to provide all 8 required inputs (or
32                 leave empty for default values): " + ...
33                 "RollWindow, Tau, Alpha, Z_target, Gamma, h,
34                 Tau_prior, Ensemble_wt_method ")
35         end
36     end
37 end

```

Note: For the Method property of a `backtestedPortfolio` object the following parameters are required based on the HS-FP method of choice:

1. `none` (default)
2. `rolling_w` (window)
3. `exp_decay` (tau)
4. `crisp` (alpha, z_target)
5. `crisp` (alpha, z_target) Note: `z_target` options 'mean' or 'latest'
6. `kernel` (h, gamma, z_target)
7. `e_pooling` (alpha, tau_prior, z_target)
8. `ew_ensemble` (alpha, tau_prior, z_target)
9. `cb_ensemble` (alpha, tau_prior, z_target, ensemble_wt_method)

A5.4.3 Class user example

To run a back-test you need to initialise a `backtestedPortfolio` object and a `HSFPparameters` object. Follow the instructions below and refer to the example code that follows. In the example below we are setting up the simulation to run a time-conditioned HS-FP back-test using a rolling window approach, with no cash constraint or additional treatments *e.g.* Winsorisation:

```

1 % 1. Define inputs:
2 CC = 1; % no constraint if == 1
3 W_sd = []; % exclude if no winsorising is wanted
4 % 2. Initialise HSFPparameters object with model specific parameters:
5 HSFPparameters_test = HSFPparameters;
6 HSFPparameters_test.Tau_prior = 36;
7 % 3. Intialise backtestedPortfolios object:
8 test_class1 = backtestedPortfolios;
9 test_class1.HSFPparameters = HSFPparameters_test;
10 % 4. Provide return & signal series and specfiy model choice:
11 test_class1>Returns = M_Ret_TT;
12 test_class1.Method = 'rolling_w';
13 test_class1.Signals = timetable(SIG_SMOOTHED_TT.Time, SIG_SMOOTHED_TT.
    SAVIT40_Index);
14 % 5. Specifiy any additional treatments to be applied
15 test_class1.WinsorStd = W_sd;
16 test_class1.CashConstraint = CC;
17 test_class1.RegLambda = 0.00;

```

Once both the `backtestedPortfolio` object and the `HSFPparameters` object have been defined we can use the `00Pbacktest_analysis` function from the `backtestedPortfolios` object to run the simulation:

```

1 test_backtestObject = 00Pbacktest_analysis(test_class1);
2 test_backtestObject.Attributions

```

Upon completion of the simulation, the user has several options for accessing results and visualising data:

1. Call the `test_backtestObject.Attributions` property to retrieve simulation results.
2. Use the `PerformancePlot(backtestedPortfolios)` function to generate performance curves:
 - Example: `backtest_plot_gcf = PerformancePlot(backtestedPortfolios)`
3. Use the `OptControlPlot(backtestedPortfolios)` function to visualise the optimal weight surface for each portfolio:
 - Example: `RollW.Wt_Surfaces_Plot = OptControlPlot(backtestedPortfolios)`

A5.5 Test case documentation

Implementing López de Prado and Lewis [67] unsupervised learning methods for clustering multiple trials is extremely computationally intensive. If one were to want to replicate the results given in Sec.2.5.1 and feel the need to speed up your code there are basic steps to follow. First ensure that you remove all the unnecessary code, vectorise and tweak the algorithms, and make sure that nothing heavy runs more often than it needs to. If that still does not product desirable speed improvements, one can implementing computationally intensive sections of your application using a compiled language, incorporating high-performance C, C++, or Fortran code.⁴ Another alternative is using parellisation which involves dividing the computational tasks into smaller sub-tasks that execute simultaneously on multiple processors or cores. This approach should speeds up processing times and improve performance by making efficient use of available computing resources.

Lastly, after tackling the software aspect, the hardware can examined to improve efficiency. This can be achieved by memory optimisations, distributing the workload across all CPU cores, as well as attempting to make use of accelerators such as GPUs. Unfortunately for macOS users GPU acceleration with Parallel Computing Toolbox is not available as of yet (when using [MATLAB 2023b](#)). Optimising data retrieval by reducing the amount of data fetched from memory, adopting a sequential memory access pattern, or enhancing data locality (by accessing data already in the cache) can result in notable performance improvements. Also note that to maximise the efficiency of certain computationally intensive code such as clustering algorithms, it is prudent to avoid running the graph-generating code. This ensures that the JIT (Just-In-Time) compiler dedicates its resources to optimising the clustering code. The JIT compilation does introduce initial overhead during code execution, however, its primary advantage is optimising frequently executed sections of the code. The JIT compiler selectively compiles ‘hotspots’ based on runtime profiling information, adapting the code to the runtime environment and enhancing subsequent executions. Refer to the [MATLAB Performance](#) page for additional resources.

A5.6 Sourcing data: a step-by-step guide to Bloomberg

As UCT has an educational license with Bloomberg and Bloomberg is one of the leading providers of real-time financial data, it naturally became my primary source of market data. In this section and subsequent sections, I provide the steps one can follow to access the data used for analysis in this dissertation. Bloomberg can be accessed through one of the four UCT terminals in the Research Wing on Jagger Level 5 of the Chancellor Oppenheimer Library on the upper campus. To use the terminals users can visit the Business Corner Bloomberg Terminals [booking page](#) to reserve a timeslot. Note that a student or staff card is required to book a terminal as well as enter the library. If a visitor would like to access the terminals, please contact the library Commerce Team for assistance (one can only enquire, however, access is never guaranteed).

Gaining access to Bloomberg once at a UCT Bloomberg terminal:

1. Based on your booking, select the appropriate computer in the Business Corner of the Research Wing.

⁴In the case of this research it was seen to be a time-consuming alternative therefore parellisation was the solution chosen.

2. The computer will prompt you to log into the network using your student/staff number and your network password.
3. You will then be prompted to provide login details for the PC itself. Here, the library usually requires the user name to be “user” and to leave the password field blank.
4. Once logged in, select the Bloomberg icon on the desktop or find it in the search bar.
5. The login details for the specific terminal are provided by the library staff and should be listed on a notice document alongside the computer. Enter these details to gain access.

There is an abundance of instruments that can be accessed through Bloomberg with thousands of fields across them. Below I provide an example of how to access and download closing price data from an equity index. These same steps may not necessarily work for all instruments, but they provide some basics to begin with.

Index price data: FTSE/JSE ALSI example

1. Follow the access steps listed above to open up the Bloomberg terminal.
2. Use the search (SRCH) function entering “All Share Index”, and select “FTSE/JSE All Share Index Total Return Value” from the drop-down menu.
3. Use the “Historical Price Table” function (abbreviation: HP) to access the historical price data.
4. To specify a specific date window change the “Date Range” field. For this dissertation, the date range specified was 2002/01/06 00:00:00 to 2023/03/12 00:00:00.
5. Select the data frequency in the “Period” field (here I chose Weekly).
6. Ensure that the “Market” field is selected to be “Last Price (PX.LAST)”.
7. For comparability, it is best to have all prices denominated in a single currency. To do this update the “Currency” field. (ZAR in my case).
8. Select the export to Excel option - this should open a new Excel workbook with the desired data, which you can go ahead and save as you wish.

A5.7 Sourcing bond data: additional steps needed

There is so much information available on any single bond, that it can be overwhelming and/or a hassle to obtain. In this section, I outline steps to access granular information on the SA bond market. As explained in Sec.4, when I began this project, one of the initial goals was to concentrate on fixed-income instruments and to build a self-constructed dynamic bond index, as well as a yield curve that we would use in the HS-FP framework.

Unfortunately, due to restrictions on accessing data, I was not able to achieve what I initially set out to do. Sec. [A5.7.1 Accessing additional bond data through Bloomberg](#) outlines the necessary steps to compile a detailed dataset on the SA bond market. To construct a historical price path for bonds with different maturities, it was essential to obtain historical prices and not only spot prices. However, Bloomberg could only supply spot prices. This led to Sec. [A5.7.2 Accessing additional bonds data: contacting the JSE](#) where I detail the efforts made to address this limitation.

By providing details of how I tried to navigate the data restrictions, I hope to offer insights for those facing similar challenges.

A5.7.1 Accessing additional bond data through Bloomberg

Below I have separated the accessing of different types of data for all listed SA fixed income instruments into three parts: 1.) Accessing general information such as maturity date, issuer, *etc.* 2.) accessing price data, and 3.) accessing further descriptive information such as call dates *etc.*

1. General info: (Issuer, ticker, coupon amount, maturity date, currency, amount outstanding *etc.*)

- (a) Use the SRCH function for “South Africa Bonds” and select “Fixed Income Search for ‘South Africa Bonds’ ”.
- (b) Under the “Select Universe” section click on “11) Asset Classes” then under “Additional Options” select “Consolidate Duplicate Bonds (REGS, 144A and STRIPs)”.
- (c) Bloomberg will then cut down the search to the appropriate number of bonds. In this case, as of the 12th of March ’23, it cut it from 8,590,891 securities to 2,144 SA-specific matches (2,130 (Bonds) + 14 (Preferreds) + 0 (Municipals) = 2,144 securities).
- (d) To view the list of bonds click on “1) Results”.
- (e) Changed the “Currency” field to ZAR to ensure consistency of results.
- (f) You can customise what data you would like to extract by clicking on the “Add Column” or “Edit Columns” tabs.
- (g) Exported data using Excel export option as detailed in point 8. of the ALBI section.

2. Prices (PX_LAST):

- (a) Return to the SRCH function for the SA bonds.
- (b) Under the “Additional Analysis Options” section select “3) Evaluate Price — FIW” to generate a Fixed Income Worksheet (FIW).
- (c) Here you have the option to narrow your search if you wish by using the “Reset Facets” menu.
- (d) Select the tab on the top that is called “List” then select the “Pricing” tab under that.
- (e) Here you can select and reorganise the column list by right-clicking on the columns.
- (f) For this research the bonds were grouped by “Country/Region of Risk” and then by “Maturity”.
- (g) I then reordered based on “Maturity” to make it easy to partition later on.
- (h) Exported data using Excel export option as detailed in point 8. of the ALBI section.

3. SA bond descriptive data: Coupon type, maturity type, next call date, next call price *etc.*

- (a) Once under the FIW you can switch from the “Pricing” tab (under the “List” tab) to the “Descriptive” tab.
- (b) This should result in descriptive fields such as “Coupon Type”, “Maturity type” *etc.* Once again you can customise the columns by right-clicking on any one of them.
- (c) Exported data using Excel export option as detailed in point 8. of the ALBI section.

A note of the JIBA listed on Bloomberg:

“To model the risk-free rate, we use the three-month Johannesburg Interbank Average Rate (JIBAR), which is calculated by taking the median value of the daily fixed rates taken from the five contributing banks as a proxy.” [68]

A5.7.2 Accessing additional bonds data: contacting the JSE

There were two practical ways to obtain historical price data: either I needed to find an alternative source to Bloomberg, or I would need to aggregate the data myself, recording the Bloomberg spot prices on a month-by-month basis.

My first port of call was to contact the primary source of data, the JSE. Under the ‘JSE Market Data Policies’ document ([Version 4.0](#)) - Section 18.4, one can request *ad-hoc* access to historical data for educational purposes by completing a ‘JSE Historical Data Order Form’ and emailing it to MDclientsjse.co.za along with proof of registration at a tertiary institute, and potentially a signed ‘Student Dispensation Letter’. Under this policy, you are required to provide additional information under different subsections. Below I have provided the section numbers (according to [Version 4.0](#)) of each question (*verbatim*) along with an example of how to answer them [62]. To determine what data you would request you can refer to the Bonds Market Data Products Specification ([V7.0](#)) document which breaks down all available data type on bonds.

1. The data required:

- (a) **Date Range:** 01.01.2000 - 01.01.2023 (20 years worth)
- (b) **Frequency:** Weekly (if possible)
- (c) **Data specifics:** Based on the [Bonds Market Data Products Specification \(V7.0\)](#) the following data is requested:
 - i. Zero Curve (see Section 5.1 of Bonds Market Data Products Specification)
 - ii. Mark To Market (MTM) Detailed (Section 5.3)
 - iii. Inflation-linked bond indices: Composite Inflation-Linked Index (CILI) (Section 5.9)
 - iv. Constituents (Section 5.10)
 - v. DAILY TRI (ATTRIBUTION REPORT) (Section 5.11)
 - vi. TRI (TOTAL RETURN INDEX) (Section 5.12)

2. The intended use of the data:

- (a) The data will be used for my Masters in Mathematical Statistics research dissertation which has an element focused on the South African Bond market. The data is for analysis purposes only and will not be updated at any point.

3. How often it will receive the data:

- (a) The data will be a once-off request and will not need to be updated, therefore the data will only be requested and received once.

4. How the data is intended to be made available (e.g. through a terminal device at the educational institution, or through web-based access):

- (a) Web-based access would be ideal if possible as I work remotely.

Further breakdown on requested data:

1. Zero Curve (see Section 5.1)[\[53\]](#) Contains:
 - (a) Bond curve data
 - (b) Swap Curve
 - (c) Real Bond curve
2. Mark To Market (MTM) Detailed (Section 5.3) - available at 17:30 daily. Contains:
 - (a) MTM (use as Cloing Price)
 - (b) BEASSA TRI (Total Return Index)
 - (c) COBI TRI (Credit Indices)
 - (d) Yield Curve
3. Inflation linked bond indices: Composite Inflation-Linked Index (CILI) (Section 5.9) Contains:
 - (a) Government (IGOV)
 - (b) State Owned Enterprises (ISOE)
 - (c) Corporates (ICORP)

Please note: "The CILI is split into four sub-indices based on term to maturity forming the 1-3 year, 3-7 year, 7-12 year and 12+ year sub-indices."

4. Constituents (Section 5.10) Contains:
 - (a) A complete list of all constituents of the BEASSA total return indices reweighting published monthly and recons are quarterly.
 - (b) Monthly re-weighting of the All Bond Index (ALBI) - split into Sector yrs
 - (c) "A complete list of all constituents of the Composite Inflation –Linked Indices (CILI) Reweighting is published monthly by the JSE. The Recon is published quarterly"
5. DAILY TRI (ATTRIBUTION REPORT) (Section 5.11) Contains:
 - (a) This report shows all intermediate calculations needed to calculate the daily index level as well as the individual bond performance figures.
6. TRI (TOTAL RETURN INDEX) (Section 5.12) Contains:
 - (a) The clean and total return index values
 - (b) Relevant Greeks (Modified Duration, Convexity).
 - (c) Periodic return numbers (Month to date and Year to date)

A6. Additional Software Documentation

A6.1 Multiple simulation test cases

When we built the algorithms for the unsupervised clustering algorithm [67] we experienced severe computational limitations on my personal PC (see machine specs as listed in [A5.1 Systems specification](#)). We went through a process of elimination to attempt to reduce the time taken to run the algorithm and to try and establish where the algorithm was spending most of its time. We ran all test cases (TC) on a clear device with no other applications running and all connectivity cut to ensure no other applications were diverting computing resources. Below in Fig.7 is an example of how we designed our TCs. The purpose is to build up from the most basic test case *e.g.* “TC-01-01” and iteratively add complexity to the test cases to systematically identify inefficiencies or defects in the software design. In the example in Fig.7 aim is 1.) to assess if parallelisation of the simulation runs decreases total runtime, and 2) to isolate the component that is the cause of an extremely long run time in the algorithm used to calculate the DSR.

DATA PROCESSING: WINSORING RETURNS, REGULARISATION OF COVARIANCE MATRIX, IMPLEMENTING CASH CONSTRAINT								
Test Case ID	Test Case Objective	Assumptions/Pre-conditions	Step	Test Procedure	Source Data	Test Inputs	Expected Results	Remarks
Multiple Simulation Build for DSR and hyper-parameter Selection								
TC-01-01	Verify that the multiple portfolio simulation works without parallelisation as well as establish runtimes.	User has defined a matrix of multiple portfolio configurations of chosen hyper-parameters as well as a base backtestedPortfolio object contain all static properties wanted for the simulation.	1	Specify multiple hyper-parameter configurations to backtest. Provide base backtestedPortfolio object with desired static properties.	Asset Returns: DATA-PROCESSED-ASSETCL_INVARIANT-RETURNS-M-TTABLE-20070731-20230228.mat	> 5 evenly spaced values within a selected range for each varying hyper-parameter. E.g. Lambda_reg_range = linspace(0,1,5);	N simulated backtestedPortfolio objects each with unique configurations of selected hyper-parameters.	Runtime: ~43.20 sec Profiler Results: Highest runtime component is the sgp for MV portfolio.
			2	Loop through configuration matrix updating varying object properties linked to hyper-parameters.	DATA-PROCESSED-SIG-ADJUSTED-SMTH-3M3Y_NORM-M-TTABLE-20070531-20230228.mat	NOTE: Actually simulation will use 15 evenly spaced values.		
TC-01-02	Verify that the multiple portfolio simulation works with parallelisation as well as establish runtimes.	Same user requirements as above. Additionally: Start a parallel pool for processes in Matlab to ensure runtime reflect only the algo runtime.	1	Specify multiple hyper-parameter configurations to backtest. Provide base backtestedPortfolio object with desired static properties.	Asset Returns: DATA-PROCESSED-ASSETCL_INVARIANT-RETURNS-M-TTABLE-20070731-20230228.mat	> 5 evenly spaced values within a selected range for each varying hyper-parameter. E.g. Lambda_reg_range = linspace(0,1,5);	N simulated backtestedPortfolio objects each with unique configurations of selected hyper-parameters.	Runtime: ~17.30 sec Profiler Results: Highest runtime component is the sgp for MV portfolio.
			2	Loop through configuration matrix updating varying object properties linked to hyper-parameters.	DATA-PROCESSED-SIG-ADJUSTED-SMTH-3M3Y_NORM-M-TTABLE-20070531-20230228.mat	NOTE: Actually simulation will use 15 evenly spaced values.		
Multiple Simulation Clustering to Estimate E[X]								
Varying the number of simulated paths								
TC-02-01	Test the base kmean clustering algorithm (clusterKMeansBase) of multiple portfolio simulation works without parallelisation as well as establish runtimes.		1	Extract all realised return series from the test case simulations for both the MV and HRP portfolios.	Test Case 125 Simulations: DATA-SIM_TC125_BKTST-PERFORMANCE-M-CELL_ARR.mat	> 125 simulated portfolio paths created for test case. > max number of clusters = 6 > number of iterations = 6		General: Runtimes of components to varying to establish the 'issue' component that is slowing the process. Runtime: ~0.16 sec Profiler Results: Highest runtime component varies at each run.
			2	Generate the covariance matrix for MV realised return paths as well as for the HRP paths.				
			3	Input the correlation matrix and specify the max k and number of iterations required.				
TC-02-02	Test the base kmean clustering algorithm (clusterKMeansBase) of multiple portfolio simulation works with parallelisation as well as establish runtimes.		1	Define options to be specified in the kmean algorithm to invoke parallelisation	Test Case 125 Simulations: DATA-SIM_TC1000_BKTST-PERFORMANCE-M-CELL_ARR.mat	> 125 simulated portfolio paths created for test case. > max number of clusters = 6 > number of iterations = 6		General: Same runtime issue as TC-02-01. Takes longer than the non-parallelised case. Runtime: ~0.50 sec Profiler Results: Highest runtime component varies.
			2	Repeat TC-02-01 with the kmeans options to invoke parallel computing.				
TC-02-03	Test whether increasing the number of simulations in the TC produces more consistent runtime results. (Without parallelisation)		1	Repeat TC-02-01 with increased number of simulation paths.	Test Case 1000 Simulations: DATA-SIM_TC1000_BKTST-PERFORMANCE-M-CELL_ARR.mat	> 1000 simulated portfolio paths created for test case. > max number of clusters = 6 > number of iterations = 6		General: More distinct and consistent runtime profile. Runtime: ~5.50 sec Profiler Results: Highest runtime component is the silhouette score calculation.

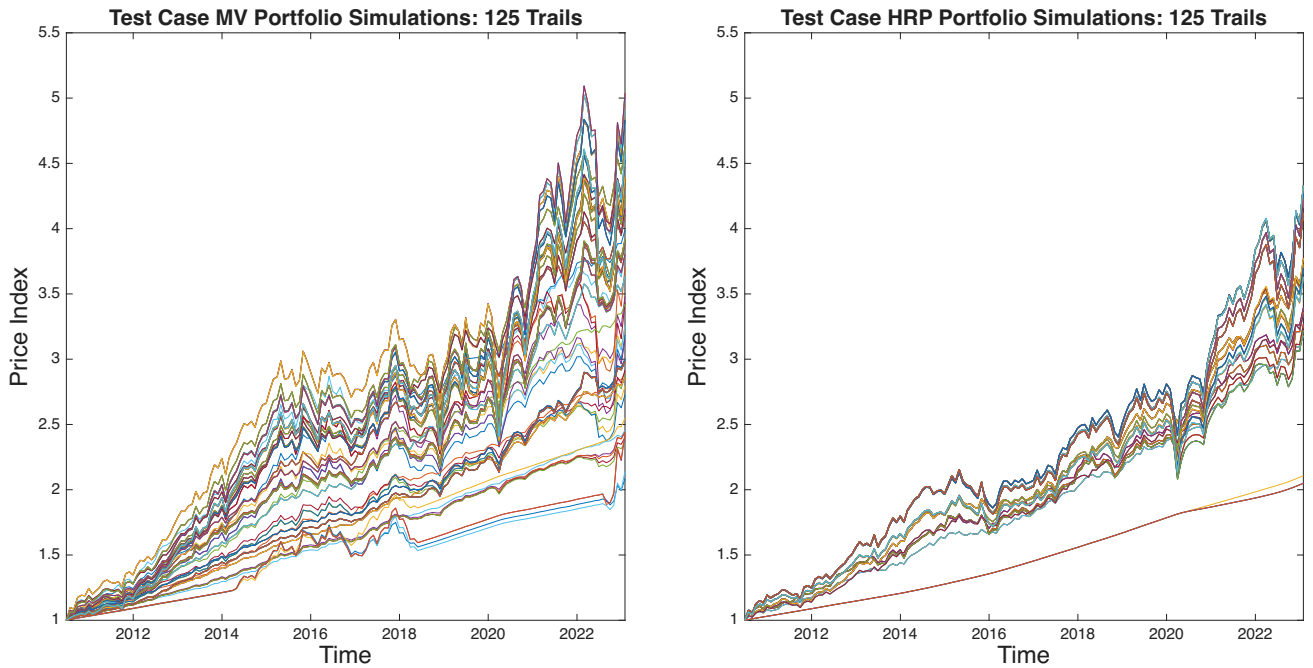
Figure 7: Example of how to set up a simulation and clustering test case documentation processes. The purpose is to build up from the most basic test case *e.g.* “TC-01-01” above and iteratively add complexity to the test cases to systematically identify inefficiencies or defects in the software design. Here the aim is to isolate the component that is the cause of an extremely long run time in the algorithm used to calculate the DSR.

In the first sequence of TCs the use of parallelisation in “TC-01-02” is tested to determine if parallelising the build of the simulations decreases runtime, which it does from ~43.20 seconds to ~17.30 seconds⁵. In the second set of test cases the goal is to identify what part of the clustering algorithm takes the longest time. We first test whether parallelisation benefits the K-Mean clustering algorithm, in which case there were negligible changes in run time between “TC-02-01” and “TC-02-02”.

When assessing which components took up the most time we saw that in “TC-02-01” and “TC-02-02,”, using a small sample of 125 simulations produced consistent clustering results when the seed was set.

⁵To assess run times and to obtain a code execution summary of its components you can use the [MATLAB Code Profiler](#).

However, the run times of the algorithm’s components (*e.g.* running the MV and HRP optimisations, running the K-Means algorithm *etc.*) fluctuated significantly due to the overall extremely short duration of the small simulation (less than 0.3 seconds on average). This variability made it difficult to pinpoint the component consistently slowing down the process. To address this, we needed to increase the number of samples to create a more distinct separation between the run times of different components. By increasing the number of simulations to 1000 as seen in “TC-02-03”, it became evident that the silhouette function that is used to score the cluster quality consistently had the highest runtime. The silhouette function is a standard MATLAB function from the Stats package and unfortunately we could not step into the function to assess the potential reason for the long run times.



(a) Multiple trial simulation of a Mean-Variance portfolio with varying hyper-parameters. These simulations consist of varying the values of each parameter as seen in Sec.2.5.1. Generated by: `testcase_dsr_4.mlx`

(b) Multiple trial simulation of a Hierarchical Risk Parity portfolio with varying hyper-parameters. These simulations consist of varying the values of each parameter as seen in Sec.2.5.1. Generated by: `testcase_dsr_4.mlx`.

Figure 8: The simulations above were generating in the process of building test cases for the full multiple simulations used in the DSR and cluster analysis conducted in Sec.2.5.1. Across 125 simulated trials of an MV portfolio, a large degree of variability is visible in the cumulative performance curves. This is suggestive of a higher probability of obtaining a false positive considering the number of trials conducted. The HRP portfolio shows a narrower spread of the performance curves. Two path trajectories are distinct, the lower group of curves displaying a slight degree of variation and the upper group exhibiting larger variation. Upon investigation, the lower curves are those generated when only the standard deviation hyper-parameter is varied within the simulation. Evidently, winsorisation does not play an important role in the improvement of the HRP method. This is discussed in greater depth in Sec.2.4.2.3. Make ref to Fig.2.9

A6.2 Code base file structure and descriptions

The code base was built in such a way that it can later be wrapped in an Object-oriented manner into custom classes. The files are hierarchically organised in a Toolboxes folder which contains specialised Toolbox sub-folders—each with their purpose in the overall code base. These can later be wrapped into methods in a class object. Live scripts with test code can be used to create an instantiation of the object going into the constructor and the `Contents.m` files would then function as help files for the class object [74, 121]. All live test scripts are numbered in order of the running sequence that needs to be followed. *E.g* if one wants to start from scratch and process the raw Bloomberg data files, you would run `data_cleaning_1b.mlx` → `state_signal_prep_2.mlx` → `tri_invariants_3.mlx`. The same principle should be applied to all toolbox test_code folders.

A6.2.1 Class definition files

Below is the file structure accompanied with each file’s description.

OOP

```

|-----backtestedPortfolio.m
|-----HSFPparameters.m
|-----testClass.mlx

```

File descriptions:

Table 3: Object-orientated Programming Code. MATLAB code files can be found at [OOP](#). An Object Orientated Programming (OOP) Approach was used to construct this code. The code defines the MATLAB class `backtestedPortfolios` that performs the back-testing of all portfolios by calculating performance metrics (like annualised returns, variances, SR, *etc.*) and generates the figures for portfolio performance and weight surfaces. An additional subclass is defined, `HSFPparameters`. It ensures that each parameter for the HS-FP methods is properly validated and initialised either with user-provided values or with predefined default settings. See [A4.6 Toolbox construction](#) for details.

File Names	Description
1. backtestedPortfolios.m	Main class definition. This class is called within the execution files of the “backtesting” toolbox in order to initialise the back-test simulations given user defined or default inputs.
2. HSFPparameters.m	Subclass defined to be used in <code>backtestedPortfolios.m</code> that initialises an <code>HSFPparameters</code> object which stores the parameter definitions for the HS-FP methods.
3. testClass.mlx	Used to test all classes and subclasses in various unit tests to ensure it returns constant results.

A6.2.2 Toolbox definition files

Below is the file structure accompanied with each file's description.

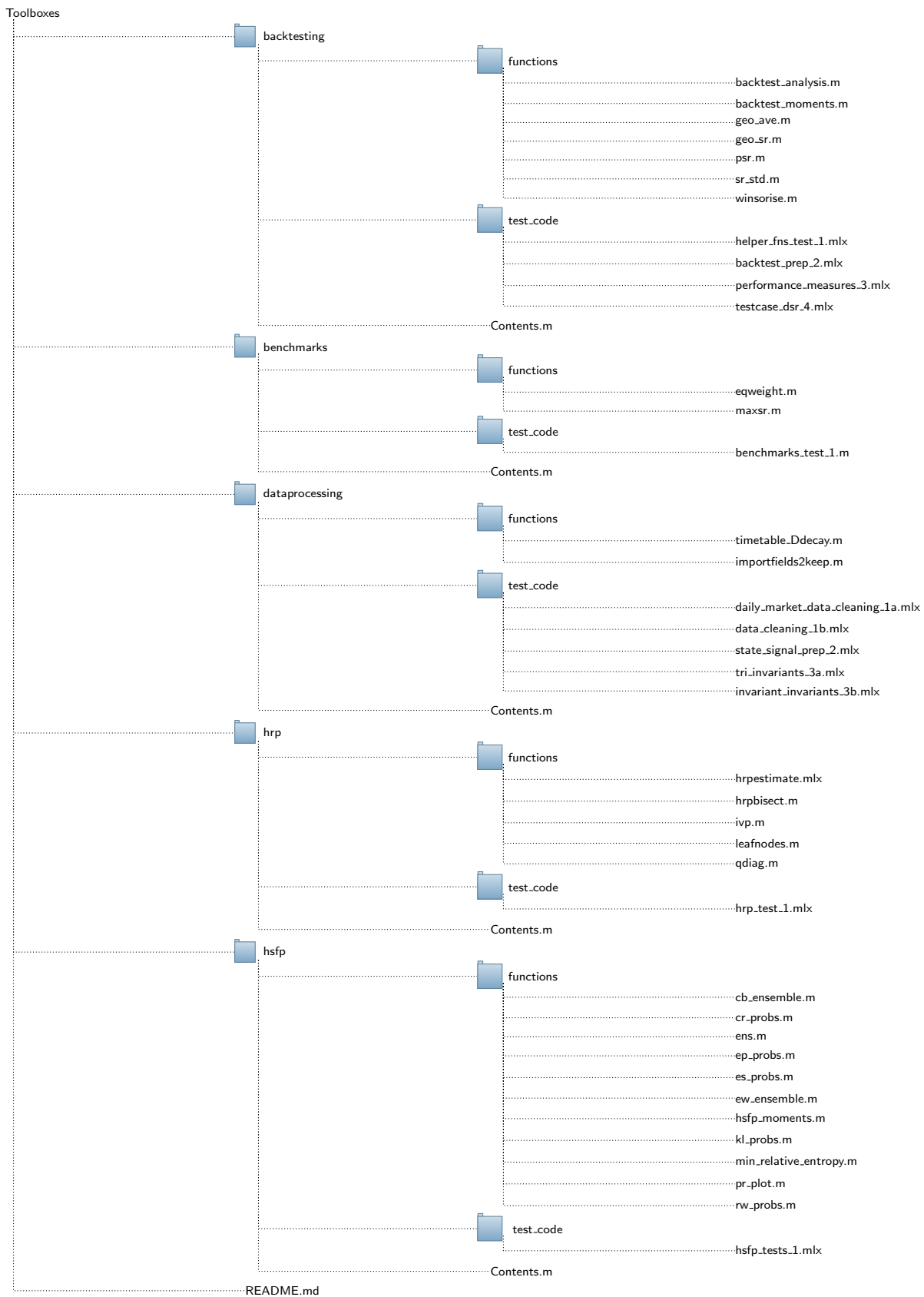


Table 4: **Back-testing Code.** MATLAB code files can be found at [Toolboxes/backtesting](#). These files consist of the test code as well as the functions use when building the back-test simulations.

File Names	Description
1. helper_fns_test_1.m	Test files for the helper functions that are called within the back-test: Winsorisation in Eq.2.38, the geometric average functions in Eq.1.23 and the geometric Sharpe Ratio function in Eq.1.32.
2. backtest_prep_2.mlx	Used to test different components of the back-test. The framework needs to accommodate several aspect to increase the functionality and usability. It demonstrates the building blocks that went into ‘OOPbacktest_analysis’.
3. performance_measures_3.mlx	Test code to compute performance metrics: SR, PSR, DSR and Portfolio attributes: returns, annualised returns, vol, annualised vol. See Sec.1.2 and 2.2 Note: The DSR function is called in this file but the test code for building the DSR can be found in “testcase_dsr_4.mlx”
4. testcase_dsr_4.mlx	First test area used to build and calculate the DSR (Sec.2.2.1.2). This file is a test case after which the full case with 3375 simulations is called in dataprocessing_dsr_5.mlx.
5. dataprocessing_dsr_5.mlx	In testcase_dsr_4.mlx the building of the DSR performance metric was created. In this file we implement the full case seen in Sec.2.5.1 which involves simulating 3375 paths by varying 3 hyperparameters with 15 data points per hyper-parameter.
6. models_results_6.mlx	Tests the reliability and implement the HSFP framework within a backtest. This file is used to generate the results in Sec.3.6.
7. backtest_analysis.m	Base function that is eventually used in the OOP class “backtestPortfolio.m”.
8. backtest_moments.m	Takes in a backtest object and allows a user to specify the HS-FP method they want to use to calculating Flexible Probabilities as seen in Sec.3.4.
9. clusterIVP.m	Computes the IVP when given a covariance matrix to be used by in the unsupervised approach to calculating the DSR by López de Prado and Lewis [67].
10. clusterKMeansBase.m	This function performs a first-pass estimate of $\mathbb{E}[K]$ for an unsupervised learning approach of López de Prado and Lewis [67]. Estimating $\mathbb{E}[K]$ is to establish the number of independent trials K that exist within M simulated trials.
11. clusterKMeansTop.m	Used to estimate the $\mathbb{E}[K]$ by applying a second-pass K-means clustering on the correlation matrix of backtest results, refining the clusters based on their silhouette scores and t-statistics.
12. geo_ave.m	Calculate the geometric average of input data, allows for annualisation or adjusting for other time periods using the number of periods within a yr e.g $f = 12$ to annualise from monthly data. See Eq. 1.23.
13. geo_sr.m	Calculates the SR based off geometric averages of the risky asset and a risk-free asset given realised returns of the risky and risk-less. Based on Eq. 1.32 as <i>ex post</i> SR.
14. psr.m	Calculate the Probabilistic Sharpe Ratio (PSR) - Bailey & de Prado (2012) [7]. $\Pr(SR^*)$ = probability that $\hat{SR} > SR^*$ where, $\hat{SR}$ = estimated Sharpe ratio, and SR^* is a benchmark comparative strategy. See Sec.2.2.1.1.
15. sr_std.m	Calculate the standard deviation of the Sharpe ratio estimator distribution given non-normal returns as in Eq.2.20. However, the formula generalises to Eq. 2.19 for normal returns is $\mu_3 = 0$ and $\mu_4 = 3$.
16. winsorise.m	Used to winsorise the returns to control for outliers. Brings in the outliers to a point that is a function for the standard deviation e.g within 2 standard deviations. See Sec.2.4.2.3

Table 5: **Benchmark Models Code.** MATLAB code files can be found at [Toolboxes/benchmarks](#). This code runs a selection of benchmark portfolios needed for comparison within the back-test simulations used to assess portfolio out-of-sample performance.

File Names	Description
1. benchmark_test_1.mlx	Main live script file used to test that all benchmark portfolios can be generated correctly when called by the back-test framework.
2. eqweight.m	Equal-weighted portfolio allocation based on number of assets in returns. $Prs = 1/N$.
3. maxsr.m	Mean-Variance Optimisation with a Maximum Sharpe Ratio objective function for portfolio weights. Portfolio constraints use are that the portfolio is fully-invested long-only portfolios <i>i.e.</i> (non-negative weights that sum to 1).

Table 6: **Data Import and Processing Code.** MATLAB code files can be found at [Toolboxes/dataprocessing](#). Initial data processing needs to be performed using this code before any other code can be run. The files need to be run in numerical order to ensure all processing runs successfully.

File Names	Description
1. daily_market_data_cleaning_1a.mlx	Generate epochs plots for the market fluctuation figure in Sec.1.1.2.
2. data_cleaning_1b.mlx	Main live script file used to import the signal and market data. Deals with missing data and frequency mismatches. Lastly synchronises individual timetable objects.
3. state_signal_prep_2.mlx	Converts raw state variables into usable state signals as well as creates composite signals such as equity risk premium <i>etc.</i> and introduces lags where seen beneficial. Uses <code>timetable_Ddecay.m</code> to smooth and standardise the variables.
4. tri_invariants_3a.mlx	Transforms asset price data to simple returns and generates the cash proxy as a forward rate of the JIBA3M. The distribution of the returns are checked for normality and scatter plots are used to check for IID behaviour. This data is destined for the standard and conventional backtests in Chp.2.
5. tri_invariants_3a.mlx	Used to transform equity prices into log-difference returns and forward rates for cash assets into invariants, verifying normality with QQ plots and IID properties with lagged scatter plots. This ensures data is approximately <i>i.i.d</i> and approximately normally distributed for Meucci's HS-FP models in Chp.3.
6. timetable_Ddecay.m	Uses Meucci's double exponential decay method (2010) (based on Holt's linear method) to smooth variables in a timetable object to reduce noise. Follows on to standardising the data for comparative purposes.
7. importfields2keep.m	Loops through sheets of excel doc, where each sheet stores individual ticker data to import selected fields from all sheets.

Table 7: **Hierarchical Risk Parity Optimisation Code.** MATLAB code files can be found at [Toolboxes/hrp](#). This code runs a HRP optimisation on the return data provided to it (raw, processed or post HS-FP returns) and generates the optimal HRP weightings for the assets.

File Names	Description
1. hrpeestimate.mlx	Main live script file that houses the main call function (hrpeestimate) that generates the Hierarchical Risk Parity weightings given a set of asset returns by means of calling subsequent sub-functions.
2. qdiag.m	Produces a sorted asset index list for quasidiagonalized covariance matrix
3. leafnodes.m	leaf nodes for a given group node id
4. hrpbisect.m	recursively bisects assets and allocates weights using vip
5. ivp.m	performs simple inverse-variance allocation

Table 8: **Historical Simulation with Flexible Probabilities Code.** MATLAB code files can be found at [Toolboxes/hsfp](#). These file contain the functions (as well as supporting functions) needed to generate FPrs using Meucci [89] methods.

File Names	Description
1. hsfp_test.1.mlx	Main live script file used to test time/state/time—single state/ multiple state conditioning of probabilities.
2. cb_ensemble.m	Generates the Time & Multiple States Conditioned Bayesian Ensemble Posterior flexible probabilities given multiple states : $\bar{q}$ entropy pooled probabilities for $\bar{q}$ signals. Calculates Bhattacharyya coefficient, Hellinger distance, Diversity Indicators as well as ω_q for each signal. Outputs f_P as the Bayesian ensemble posterior stemming for the conditioning via $\bar{q}$ signals.
3. cr_probs.m	Single State variable Market conditioning using Crisp Probabilities for the HS-FP framework. Optimal Prs are set to 1 if z_t falls within the bandwidth range around the selected target value.
4. ens.m	Calculates the Effective Number of Scenarios $\hat{T}$ using Meucci's more general expression: $\hat{T} \equiv \left(\sum_{t=1}^T p_t^\gamma \right)^{-1/(\gamma-1)}$. The exponential of the entropy can also be specific, in which case: $\hat{T} \equiv e^{-\sum_{t=1}^T p_t \ln p_t}$.
5. ep_probs.m	Calculates the updated distribution (posterior) of Flexible Probabilities using entropy pooling . Calls on min_relative_entropy.m Fn to produce the posterior given an exponential decay prior and the moments of Crisp probability to form the constraints.
6. es_probs.m	Facilitates the calculation of exponentially smoothed time-conditioned probabilities used in the HS-FP framework developed by Meucci (2010).
7. ew_ensemble.m	Calculates the equally weighted posterior ensemble flexible probabilities via entropy pooling given $\bar{q}$ conditioning variables (signals).
8. hsfp_moments.m	Calculates the Mean and Covariance of asset returns using FPrs .
9. kl_probs.m	Single State variable Market conditioning using Kernel-based Probabilities for the HS-FP framework. Optimal Prs are proportional to the distance that z_t is away from the target z^* .
10. min_relative_entropy.m	Contains nested functions: The Lagrangian dual functions & the Hessian function.
11. pr_plot.m	Generates filled staircase plot of probabilities over time period of T .
12. rw_probs.m	Facilitates the calculation of the rolling window time-conditioned probabilities used in the HS-FP framework developed by Meucci (2010).

References

- [1] Erdinç Akyıldırım and Halil Mete Soner. A brief history of mathematics in finance. *Borsa Istanbul Review*, 14(1): 57–63, 2014. ISSN 2214-8450. doi: <https://doi.org/10.1016/j.bir.2014.01.002>. URL <https://www.sciencedirect.com/science/article/pii/S2214845014000039>.
- [2] Robert Almgren and Neil Chriss. Optimal portfolios from ordering information. In *Forecasting Expected Returns in the Financial Markets*, pages 55–100. Elsevier, 2007.
- [3] Robert D Arnott, Campbell R Harvey, and Harry Markowitz. A backtesting protocol in the era of machine learning. *Available at SSRN 3275654*, pages 3–6, 2018. doi: 10.2139/ssrn.3275654.
- [4] Anthony B Atkinson. The timescale of economic models: how long is the long run? *The Review of Economic Studies*, 36(2):137–152, 1969.
- [5] Carl R Bacon. *Practical Portfolio Performance Measurement and Attribution*. John Wiley & Sons, 2023. ISBN 9780470059289. doi: 10.1002/9781119206309.
- [6] David H Bailey and Marcos Lopez De Prado. The deflated sharpe ratio: correcting for selection bias, backtest overfitting, and non-normality. *The Journal of Portfolio Management*, 40(5):94–107, 2014. doi: 10.3905/jpm.2014.40.5.094.
- [7] David H Bailey and Marcos Lopez de Prado. The sharpe ratio efficient frontier. *Journal of Risk*, 15(2):13, 2012. doi: 10.2139/ssrn.1821643.
- [8] David H Bailey, Jonathan Borwein, Marcos Lopez de Prado, and Qiji Jim Zhu. Pseudo-mathematics and financial charlatanism: The effects of backtest overfitting on out-of-sample performance. *Notices of the American Mathematical Society*, 61(5):458–471, 2014. doi: 10.2139/ssrn.2308659.
- [9] David H Bailey, Jonathan Borwein, Marcos Lopez de Prado, and Qiji Jim Zhu. The probability of backtest overfitting. *Journal of Computational Finance*, forthcoming, 2015. doi: 10.2139/ssrn.2326253.
- [10] Discovery Bank. Savings accounts rates and fees, 2023. URL <https://www.discovery.co.za/bank/discovery-savings-fees>.
- [11] South African Reserve Bank. *A review of the rate-setting process of the Johannesburg Interbank Agreed Rate (Jibar) as an interest rate benchmark*. Financial Markets Department, Nov 2012.
- [12] Patrizia Baudino, Roland Goetschmann, Jérôme Henry, Ken Taniguchi, and Weisha Zhu. *Stress-testing banks: a comparative analysis*. Bank for International Settlements, Financial Stability Institute, 2018.
- [13] Jakob Bernoulli. Ars conjectandi: Usus & applicationem praecedentis doctrinae in civilibus. *Moralibus & Oeconomicis*, 4(1713):1713, 1713.
- [14] Michael J Best and Robert R Grauer. On the sensitivity of mean-variance-efficient portfolios to changes in asset means: some analytical and computational results. *The review of financial studies*, 4(2):315–342, 1991.
- [15] A. Bhattacharyya. On a measure of divergence between two multinomial populations. *Sankhyā: The Indian Journal of Statistics (1933-1960)*, 7(4):401–406, 1946. ISSN 00364452. URL <http://www.jstor.org/stable/25047882>.

- [16] George D. Birkhoff. Proof of the ergodic theorem. *Proceedings of the National Academy of Sciences*, 17(12):656–660, 1931. doi: 10.1073/pnas.17.2.656. URL <https://www.pnas.org/doi/abs/10.1073/pnas.17.2.656>.
- [17] Fischer Black and Robert Litterman. Asset allocation: combining investor views with market equilibrium. *Goldman Sachs Fixed Income Research*, 115(1):7–18, 1990. doi: 10.3905/jfi.1991.408013.
- [18] Fischer Black and Robert Litterman. Asset allocation: combining investor views with market equilibrium. *Goldman Sachs Fixed Income Research*, 115, 1990.
- [19] Z. Bodie, A. Kane, and A.J. Marcus. *Investments*. McGraw Hill series in finance, insurance, and real estate. McGraw Hill LLC, 2023. ISBN 9781266837227. URL https://books.google.co.za/books?id=W7B_zwEACAAJ.
- [20] J.Y. Campbell and L.M. Viceira. *Strategic Asset Allocation: Portfolio Choice for Long-term Investors*. Clarendon lectures in economics. Oxford University Press, 2002. ISBN 9780198296942. URL <https://books.google.co.za/books?id=Me7K1UVJ4vAC>.
- [21] J.Y. Campbell, A.W. Lo, and A.C. MacKinlay. *The Econometrics of Financial Markets*. Princeton University Press, 2012. ISBN 9781400830213. URL <https://books.google.co.za/books?id=7Gkri6HWWkgC>.
- [22] Denis Chaves, Jason Hsu, Feifei Li, and Omid Shakernia. Risk parity portfolio vs. other asset allocation heuristic portfolios. *Journal of Investing*, 20(1):108, 2011. doi: 10.3905/joi.2011.20.1.108.
- [23] Vijay K Chopra and William T Ziemba. The effect of errors in means, variances, and covariances on optimal portfolio choice. In *Handbook of the fundamentals of financial decision making: Part I*, pages 365–373. World Scientific, 2013.
- [24] Steve Christie. Is the sharpe ratio useful in asset allocation? *Macquarie Applied Finance Centre Research Paper*, 2005.
- [25] Roger G Clarke, Harindra De Silva, and Brett H Wander. Risk allocation versus asset allocation. *The Journal of Portfolio Management*, 29(1):9–30, 2002. doi: 10.3905/jpm.2002.319860.
- [26] Stuart Coles, Joanna Bawa, Lesley Trenner, and Pat Dorazio. *An introduction to statistical modeling of extreme values*, volume 208. Springer, 2001. ISBN 978-1-85233-459-8. doi: 0.1007/978-1-4471-3675-0.
- [27] Eberly College of Science. T.2.5.2 - exponential smoothing: Stat 501 - regression methods, 2023. URL <https://online.stat.psu.edu/stat501/lesson/t/t.2/t.2.5/t.2.5.2-exponential-smoothing>.
- [28] Lara Dalmeyer and Tim Gebbie. Geometric insights into robust portfolio construction. *International Journal of Theoretical and Applied Finance*, 27(07n08):2540024, 2024. doi: 10.1142/S0219024924500249. URL <https://doi.org/10.1142/S0219024924500249>.
- [29] Marcos Lopez De Prado. Building diversified portfolios that outperform out-of-sample. *The Journal of Portfolio Management*, 42(4):59–69, 2016. doi: 10.3905/jpm.2016.42.4.059.
- [30] Marcos López De Prado. The 10 reasons most machine learning funds fail. *The Journal of Portfolio Management*, 44(6):120–133, 2018. doi: 10.2139/ssrn.3104816.
- [31] Victor DeMiguel, Lorenzo Garlappi, and Raman Uppal. Optimal Versus Naive Diversification: How Inefficient is the 1/N Portfolio Strategy? *The Review of Financial Studies*, 22(5):1915–1953, 12 2007. ISSN 0893-9454. doi: 10.1093/rfs/hhm075. URL <https://doi.org/10.1093/rfs/hhm075>.
- [32] Nadia Drake. Vera rubin, pioneering astronomer, dies at 88. *National Geographic*, Dec 2016. URL <https://www.nationalgeographic.com/science/article/vera-rubin-dark-matter-galaxy-rotation-nobel-science>.
- [33] EB1009436 Dynkin. Kolmogorov and the theory of markov processes. *The Annals of Probability*, 17(3):822–832, 1989.
- [34] W. Feller. *An Introduction to Probability Theory and Its Applications, Volume 2*. Wiley Series in Probability and Statistics. Wiley, 1991. ISBN 9780471257097. URL <https://books.google.co.za/books?id=rxadEAAQBAJ>.
- [35] International Monetary Fund. *South Africa: Special Data Dissemination Standard - Interest Rates*. IMF Bulletin Board, Nov 2012. URL <https://dsbb.imf.org/sdds/dqaf-base/country/ZAF/category/INR00>.

-
- [36] Owen Futter, Blanka Horvath, and Magnus Wiese. Signature trading: A path-dependent extension of the mean-variance framework with exogenous signals. *arXiv preprint arXiv:2308.15135*, 2023. doi: 10.2139/ssrn.4541830.
 - [37] Tim Gebbie. Sta3045f: Advanced stochastic processes part ii: Modelling random events, 2022.
 - [38] Ramazan Gençay, Michel Dacorogna, Ulrich A Muller, Olivier Pictet, and Richard Olsen. *An introduction to high-frequency finance*. Elsevier, 2001. doi: 10.1016/B978-0-12-279671-5.X5000-X.
 - [39] Maxim Golts and Gregory C Jones. A sharper angle on optimization. *Available at SSRN 1483412*, 2009.
 - [40] R.M. Gray. *Probability, Random Processes, and Ergodic Properties*. Springer New York, 2009. ISBN 9781475720242. doi: 10.1007/978-1-4419-1090-5. URL <https://books.google.co.za/books?id=2Qr3BwAAQBAJ>.
 - [41] Massimo Guidolin and Allan Timmermann. Asset allocation under multivariate regime switching. *Journal of Economic Dynamics and Control*, 31(11):3503–3544, 2007. doi: 10.2139/ssrn.940652.
 - [42] Thomas Guillerme and Natalie Cooper. Effects of missing data on topological inference using a total evidence approach. *Molecular phylogenetics and evolution*, 94, 08 2015. doi: 10.1016/j.ympev.2015.08.023.
 - [43] J.D. Hamilton. *Time Series Analysis*. Number v. 10 in Book collections on Project MUSE. Princeton University Press, 1994. ISBN 9780691042893. URL https://books.google.co.za/books?id=B8_1UBmqVUoC.
 - [44] Cecil Hastings Jr, Frederick Mosteller, John W Tukey, and Charles P Winsor. Low moments for small samples: a comparative study of order statistics. *The Annals of Mathematical Statistics*, 18(3):413–426, 1947. doi: 10.1214/aoms/1177730388.
 - [45] Nicholas J. Higham. Computing the nearest correlation matrix—a problem from finance. *IMA Journal of Numerical Analysis*, 22(3):329–343, 07 2002. ISSN 0272-4979. doi: 10.1093/imanum/22.3.329. URL <https://doi.org/10.1093/imanum/22.3.329>.
 - [46] Michael A Hitt, Jean-Luc Arregle, and R Michael Holmes Jr. Strategic management theory in a post-pandemic and non-ergodic world. *Journal of management studies*, 58(1):259, 2021.
 - [47] Arthur E Hoerl and Robert W Kennard. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 12(1):55–67, 1970. doi: 10.2307/1271436.
 - [48] Charles C Holt. Forecasting seasonals and trends by exponentially weighted moving averages. *International journal of forecasting*, 20(1):5–10, 2004. doi: 10.1016/j.ijforecast.2003.09.015.
 - [49] Arthur E Horel. Application of ridge analysis to regression problems. *Chemical Engineering Progress*, 58:54–59, 1962.
 - [50] John Hull. *Risk Management and Financial Institutions, 6th Edition*, volume 733. John Wiley & Sons, 2023. ISBN 978-1-119-93248-2.
 - [51] Hilary Joffe. Replacing jibar with zaronia, May 2023. URL <https://www.rmb.co.za/news/replacing-jibar-with-zaronia#:~:text=And%20it%20reflects%20the%20rate,transactions%20and%20calculated%20in%20arrears>.
 - [52] Robert Jones, Terence Lim, and Peter J Zangari. The black-litterman model for structured equity portfolios. *Journal of Portfolio Management*, 33(2):24, 2007.
 - [53] Limited JSE. Bonds market data products specifications version 7.0, August 2023. URL <https://clientportal.jse.co.za/Content/Market%20Data%20Interest%20Rate%20Derivatives/Bonds%20Market%20Data%20Products%20Specifications.pdf>.
 - [54] Alexander Kempf and Christoph Memmel. Estimating the global minimum variance portfolio. *Schmalenbach Business Review*, 58:332–348, 2006. doi: 10.1007/BF03396737.
 - [55] Judy L Klein. *Statistical visions in time: a history of time series analysis, 1662-1938*. Cambridge University Press, 1997.

- [56] Petter N Kolm, Reha Tütüncü, and Frank J Fabozzi. 60 years of portfolio optimization: Practical challenges and current trends. *European Journal of Operational Research*, 234(2):356–371, 2014. doi: 10.1016/j.ejor.2013.10.060.
- [57] S. Kullback and R. A. Leibler. On information and sufficiency. *The Annals of Mathematical Statistics*, 22(1):79–86, 1951. ISSN 00034851. URL <http://www.jstor.org/stable/2236703>.
- [58] D.H Leavens. Diversification of investments. *Trusts and Estates*, 80:469–473, 1945.
- [59] Olivier Ledoit and Michael NMI Wolf. Honey, i shrunk the sample covariance matrix. *SSRN Electronic Journal*, 2003. doi: 10.2139/ssrn.433840.
- [60] Wai Lee. *Theory and Methodology of Tactical Asset Allocation*. Frank J. Fabozzi Series. Wiley, 2000. ISBN 9781883249724. URL <https://books.google.co.za/books?id=e2V1TvPITqAC>.
- [61] ABSA Limited. Savings and investment accounts, 2023. URL <https://www.absa.co.za/personal/save-invest/products/>.
- [62] JSE Limited. Market data policies - jse version 4.0, 2014. URL https://www.jse.co.za/sites/default/files/jse_document_manager/RW/Internal/JSE%20Data%20Agreements/Equities%2C%20Derivatives%20and%20Interest%20Rates/01.%20JSE%20Market%20Data%20Policies.pdf.
- [63] JSE Limited. Albi index details, 2023. URL <https://www.jse.co.za/albi#:~:text=Index%20Reconstitution-,End%20of%20every%20quarter%20in%20Feb%2C%20May%2C%20Aug%2C%20Nov,first%20Thursday%20of%20reconstitution%20month.&text=Effective%20the%20Friday%20following%20the%20first%20Thursday%20of%20reweighting%20month.,-Dissemination&text=The%20ALBI%207%2D12%20Year,maturity%20of%207%2D12%20years>.
- [64] JSE Limited. Savi top 40, 2023. URL <https://www.jse.co.za/trade/derivitive-market/equity-derivatives/savi-top-40>.
- [65] Nedbank Limited. Save and invest for the life you want, 2023. URL <https://personal.nedbank.co.za/save-and-invest/accounts/on-demand/money-market.html>.
- [66] Andrew W. Lo. The statistics of sharpe ratios. *Financial Analysts Journal*, 58(4):36–52, 2002. ISSN 0015198X. URL <http://www.jstor.org/stable/4480405>.
- [67] Marcos López de Prado and Michael J Lewis. Detection of false investment strategies using unsupervised learning methods. *Quantitative Finance*, 19(9):1555–1565, 2019. doi: 10.2139/ssrn.3167017.
- [68] Bloomberg Finance L.P. Bloomberg terminal version 112.3.80, March 2023. URL <https://www.bloomberg.com/professional/support/release-notes/>.
- [69] Harry Markowitz. Portfolio selection. *The Journal of Finance*, 7(1):77–91, 1952. ISSN 00221082, 15406261. doi: 10.2307/2975974. URL <http://www.jstor.org/stable/2975974>.
- [70] Harry M Markowitz. The sveriges riksbank prize in economic sciences in memory of alfred nobel 1990, 1990. URL <https://www.nobelprize.org/prizes/economic-sciences/1990/markowitz/biographical/>.
- [71] Harry M. Markowitz. The early history of portfolio theory: 1600–1960. *Financial Analysts Journal*, 55(4):5–16, 1999. doi: 10.2469/faj.v55.n4.2281.
- [72] Jacob Marschak. Money and the theory of assets. *Econometrica, Journal of the Econometric Society*, pages 311–325, 1938. doi: 10.2307/1905409.
- [73] Maria Soledad Martinez Peria, Giovanni Majnoni, Matthew T. Jones, and Winfrid Blaschke. Stress testing of financial systems: An overview of issues, methodologies, and fsap experiences. *IMF Working Papers*, 01(88):1, 2001. doi: 10.5089/9781451851168.001.
- [74] Help Center MATLAB. Create help summary files — contents.m, 2022. URL https://uk.mathworks.com/help/matlab/matlab_prog/create-a-help-summary-contents-m.html.

-
- [75] Nina Matthews. Time- and market-conditioned flexible probabilities: in the context of the south african market, 2023. URL <https://github.com/nematthews/Meucci-HS-FP-framework-SA-context>.
 - [76] Nina Matthews and Siphesihle Cele. Comparing the hierarchical risk parity algorithm and mean-variance portfolio selection, 2021.
 - [77] Nina Matthews and Siphesihle Cele. Honours thesis: cluster analysis for portfolio selection: Code and data needed for replication., 2021. URL <https://github.com/nematthews/Honours-Thesis-Cluster-Analysis-for-Portfolio-Selection>.
 - [78] David G McMillan. Stock return predictability: Using the cyclical component of the price ratio. *Research in International Business and Finance*, 48:228–242, 2019.
 - [79] Alexander J. McNeil and Rüdiger Frey. Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach. *Journal of Empirical Finance*, 7(3):271–300, 2000. ISSN 0927-5398. doi: 10.1016/S0927-5398(00)00012-8. URL <https://www.sciencedirect.com/science/article/pii/S0927539800000128>. Special issue on Risk Management.
 - [80] Antonio Mele. Asymmetric stock market volatility and the cyclical behavior of expected returns. *Journal of financial economics*, 86(2):446–478, 2007.
 - [81] Elmar Mertens. Comments on variance of the iid estimator in lo (2002). Technical report, Technical report, Working Paper University of Basel, 2002.
 - [82] Attilio Meucci. Beyond black-litterman: Views on non-normal markets. *Available at SSRN 848407*, 2005. doi: 10.2139/ssrn.848407.
 - [83] Attilio Meucci. *Risk and Asset Allocation*, volume 1. Springer, 2005. ISBN 978-3-642-00964-8. doi: 10.1007/978-3-540-27904-4.
 - [84] Attilio Meucci. Fully flexible views: Theory and practice. *Fully Flexible Views: Theory and Practice, Risk*, 21(10): 97–102, 2008. URL <https://ssrn.com/abstract=1213325>.
 - [85] Attilio Meucci. Historical scenarios with fully flexible probabilities. *GARP Risk Professional*, pages 47–51, 2010. doi: 10.2139/ssrn.1696802.
 - [86] Attilio Meucci. ‘the prayer’ ten-step checklist for advanced risk and portfolio management. *Social Science Research Network*, 2011.
 - [87] Attilio Meucci. Mixing probabilities, priors and kernels via entropy pooling. *GARP Risk Professional*, pages 32–36, 2011. URL <https://ssrn.com/abstract=1944303>.
 - [88] Attilio Meucci. Effective number of scenarios in fully flexible probabilities. *GARP risk professional*, pages 32–35, 2012. URL <https://ssrn.com/abstract=1971808>.
 - [89] Attilio Meucci. Estimation and stress-testing via time-and market-conditional flexible probabilities. *Available at SSRN 2312126*, 2013. doi: 10.2139/ssrn.2312126.
 - [90] Richard O Michaud and Robert O Michaud. *Efficient Asset Management: Practical Guide to Stock Portfolio Optimization*. Oxford University Press, March 2008. ISBN 9780195331912.
 - [91] Kristine Monteith, James L. Carroll, Kevin Seppi, and Tony Martinez. Turning bayesian model averaging into bayesian model combination. In *The 2011 International Joint Conference on Neural Networks*, pages 2657–2663, 2011. doi: 10.1109/IJCNN.2011.6033566.
 - [92] MSCI. *Barra Risk Model Handbook*. MSCI Inc., 2004. URL <https://www.msci.com/www/research-report/barra-risk-model-handbook/015929568>.
 - [93] Nasdaq. The most rewarding portfolio construction techniques: An unbiased evaluation, Sep 2013. URL <https://www.nasdaq.com/articles/most-rewarding-portfolio-construction-techniques-unbiased-evaluation-2013-09-24>.

-
- [94] Office of the United States Trade Representative. African growth and opportunity act (agoa), 2023. URL <https://ustr.gov/issue-areas/trade-development/preference-programs/african-growth-and-opportunity-act-agoa>.
 - [95] Esa Ollila, Daniel P. Palomar, and Frédéric Pascal. Shrinking the eigenvalues of m-estimators of covariance matrix. *IEEE Transactions on Signal Processing*, 69:256–269, 2021. doi: 10.1109/TSP.2020.3043952.
 - [96] John Douglas Opdyke. Comparing sharpe ratios: so where are the p-values? *Journal of Asset Management*, 8:308–336, 2007.
 - [97] Antonio Páez and Geneviève Boisjoly. *Exploratory Data Analysis*, pages 25–64. Springer International Publishing, Cham, 2022. ISBN 978-3-031-20719-8. doi: 10.1007/978-3-031-20719-8_2. URL https://doi.org/10.1007/978-3-031-20719-8_2.
 - [98] Jochen Papenbrock. *Asset clusters and asset networks in financial risk management and portfolio optimization*. PhD thesis, Karlsruher Institut für Technologie (KIT), 2011.
 - [99] Andrew Paskaramoorthy, Tim Gebbie, and Terence L. Van Zyl. The efficient frontiers of mean-variance portfolio rules under distribution misspecification. In *2021 IEEE 24th International Conference on Information Fusion (FUSION)*. IEEE, November 2021. doi: 10.23919/fusion49465.2021.9626916. URL <https://ieeexplore.ieee.org/document/9626916>.
 - [100] Andrew B Paskaramoorthy, Tim J Gebbie, and Terence L van Zyl. A framework for online investment decisions. *Investment Analysts Journal*, 49(3):215–231, 2020. doi: 10.1080/10293523.2020.1806460.
 - [101] Andre F. Perold. Large-scale portfolio optimization. *Management Science*, 30(10):1143–1160, 1984. ISSN 00251909, 15265501. URL <http://www.jstor.org/stable/2631383>.
 - [102] Ole Peters. Optimal leverage from non-ergodicity. *Quantitative Finance*, 11(11):1593–1602, 2011. doi: 10.1080/14697688.2010.513338. URL <https://doi.org/10.1080/14697688.2010.513338>.
 - [103] Ole Peters. Author correction: The ergodicity problem in economics. *Nature Physics*, 16(1):111–111, 2019.
 - [104] Ole Peters and Alexander Adamou. Ergodicity economics. In *Ergodicity Economics*, 2017. URL <https://api.semanticscholar.org/CorpusID:202546236>.
 - [105] K.E. Petersen. *Ergodic Theory*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 1989. ISBN 9781316583203. URL https://books.google.co.za/books?id=is_LCgAAQBAJ.
 - [106] Jacques Pezier et al. Global portfolio optimization revisited: A least discrimination alternative to black-litterman. Technical report, Henley Business School, University of Reading, 2007.
 - [107] Johann Pfitzinger and Nico Katzke. Perspectives on optimal hierarchical portfolio selection. *Perspectives on Optimal Hierarchical Portfolio Selection (April 4, 2023)*, 2023. doi: 10.2139/ssrn.4409173. URL https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4409173.
 - [108] Eugene Podkaminer, Wylie Tollette, and Laurence Siegel. Preparing a multi-asset class portfolio for shocks to economic growth. *The Journal of Portfolio Management*, 45(2):106–116, 2018.
 - [109] Edward Qian and Stephen Gorman. Conditional distribution in portfolio theory. *Financial Analysts Journal*, 57(2): 44–51, 2001.
 - [110] Riccardo Rebonato. *Coherent Stress Testing: A Bayesian Approach to the Analysis of Financial Stress*. John Wiley & Sons, 2010. ISBN 978-0-470-97148-2.
 - [111] S.I. Resnick. *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling*. Number v. 10 in Heavy-tail phenomena: probabilistic and statistical modeling. Springer, 2007. ISBN 9780387242729. URL <https://books.google.co.za/books?id=p8uq2QFw9PUC>.
 - [112] Peter J Rousseeuw. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of computational and applied mathematics*, 20:53–65, 1987. doi: 10.1016/0377-0427(87)90125-7.

-
- [113] Andrew Donald Roy. Safety first and the holding of assets. *Econometrica: : Journal of The Econometric Society*, 20(3):431–449, 1952. ISSN 00129682, 14680262. doi: 10.2307/1907413. URL <http://www.jstor.org/stable/1907413>.
 - [114] Johnny Ryan. *A History of the Internet and the Digital Future*. Reaktion Books, 2010. ISBN 9781861898357. URL <https://books.google.co.za/books?id=100YhHefumoC>.
 - [115] D.W. Scott. *Multivariate Density Estimation: Theory, Practice, and Visualization*. Wiley Series in Probability and Statistics. Wiley, 2015. ISBN 9781118575482. URL <https://books.google.co.za/books?id=TZg3BwAAQBAJ>.
 - [116] Antoinette G. Sebastian and Tim J. Gebbie. Systematic asset allocation using flexible views for south african markets. *arXiv: Portfolio Management*, 2019. doi: 10.48550/ARXIV.1910.05555. URL <https://api.semanticscholar.org/CorpusID:204509607>.
 - [117] William F. Sharpe. Mutual fund performance. *The Journal of Business*, 39(1):119–138, 1966. ISSN 00219398, 15375374. URL <http://www.jstor.org/stable/2351741>.
 - [118] William F Sharpe. Adjusting for risk in portfolio performance measurement. *The Journal of Portfolio Management*, 1(2):29–34, 1975. doi: 10.3905/jpm.1975.408513.
 - [119] William F. Sharpe. *The Sharpe Ratio (Fall 1994)*, pages 169–178. Princeton University Press, Princeton, 1998. ISBN 9781400829408. doi: 10.1515/9781400829408-022. URL <https://doi.org/10.1515/9781400829408-022>.
 - [120] Documentation The MathWorks Inc. Hierarchies of classes — concepts, 2023. URL https://uk.mathworks.com/help/matlab/matlab_oop/hierarchies-of-classes-concepts.html.
 - [121] Help Center The MathWorks Inc. User-defined classes, 2023. URL https://uk.mathworks.com/help/matlab/matlab_oop/user-defined-classes.html.
 - [122] Erik N. Torgersen, Søren Johansen, Bo Lindqvist, J. Hilden, and O. Barndorff-Nielsen. Comparison of statistical experiments [with discussion and reply]. *Scandinavian Journal of Statistics*, 3(4):186–208, 1976. ISSN 03036898, 14679469. URL <http://www.jstor.org/stable/4615635>.
 - [123] R.S. Tsay. *Analysis of Financial Time Series*. Wiley Series in Probability and Statistics. Wiley, 2005. ISBN 9780471746188. URL https://books.google.co.za/books?id=ddL4tTLb_08C.
 - [124] P van Rensburg. Market segmentation on the johannesburg stock exchange ii. *Studies in Economics and Econometrics*, 26(1):83–100, 2002. doi: 10.1080/10800379.2002.12106327.
 - [125] Aditya Vyas. The hierarchical risk parity algorithm: An introduction, Jan 2021. URL <https://hudsonthames.org/an-introduction-to-the-hierarchical-risk-parity-algorithm/>.
 - [126] Diane Wilcox and Tim Gebbie. An analysis of cross-correlations in an emerging market. *Physica A: Statistical Mechanics and its Applications*, 375(2):584–598, 2007. ISSN 0378-4371. doi: <https://doi.org/10.1016/j.physa.2006.10.030>. URL <https://www.sciencedirect.com/science/article/pii/S0378437106010351>.
 - [127] John Burr Williams. *The Theory of Investment Value*. Harvard Univ. Pr., 1938.
 - [128] Stavros A Zenios. *Financial Optimization*. Cambridge University Press, 1993. ISBN 9780521577779. doi: 10.1017/CBO9780511522130. URL <https://books.google.co.za/books?id=-FJdgjZOqSAC>.