



Research article

An innovative extended Bayesian analysis of the relationship between returns and different risk measures in South Africa

Nitesha Dwarika*

Independent Researcher, 15 Sisal Place, Crossmoor, Chatsworth, Durban, 4092, South Africa

* **Correspondence:** Email: niteshadwarika395@gmail.com.

Abstract: This study investigated the All Share Index (ALSI) returns and six different risk measures of the South African market for the sample period from 17 March 2000 to 17 March 2022. The risk measures analyzed were standard deviation (SD), absolute deviation (AD), lower semi absolute deviation (LSAD), lower semivariance (LSV), realized variance (RV) and the bias-adjusted realized variance (ARV). This study made an innovative contribution on a methodological and practical level, by being the first study to extend from the novel Bayesian approach by Jensen and Maheu (2018) to methods by Karabatsos (2017)—density regression, quantile regression and survival analysis. The extensions provided a full representation of the return distribution in relation to risk, through graphical analysis, producing novel insight into the risk-return topic. The most novel and innovative contribution of this study was the application of survival analysis which analyzed the “life” and “death” of the risk-return relationship. From the density regression, this study found that the chance of investors earning a superior return was substantial and that the probability of excess returns increased over time. From quantile regression, results revealed that returns have a negative relationship with the majority of the risk measures—SD, AD, LSAD and RV. However, a positive risk-return relationship was found by LSV and the ARV, with the latter having the steepest slope. Results were the most pronounced for the ARV, especially for the survival analysis. While ARV earned the highest returns, it had the shortest lifespan, which can be attributed to the volatile nature of the South African market. Thus, investors that seek short-term high-earning returns would examine ARV followed by LSV, whereas the remaining risk measures can be used for other purposes, such as diversification purposes or short selling.

Keywords: risk-return relationship; Bayesian approach; parametric model; asymmetry; density regression; quantile regression; survival analysis

JEL Codes: C22, C58, G12, G32

1. Introduction

The financial markets of the world are vital drivers of economic progression, as they are responsible for the efficient allocation of resources. Over recent years, emerging markets have become increasingly more recognized due to integrating with the global economy. While emerging markets have been subject to positive developments, these markets are still subject to severe market volatility due to not having fully developed to advanced markets (Sehgal and Pandey, 2018). Developed markets are relatively more stable; therefore, trends may be easier to identify (Yildiz et al., 2022). However, such a market does not stimulate risk-seeking investors who want to optimize their chance of earning a superior return.

Emerging markets are generally known to be more profitable than developed markets due to being volatile (Sehgal and Pandey, 2018). South Africa is an example of an emerging market, characterized by various sources of risk that contribute to the volatile nature of the market. According to the National Treasury (2021), such sources include electricity shortages, political and regulatory uncertainty and, most importantly, the fiscal deficit. The lack of fiscal availability has left South Africa susceptible after the external shocks associated with COVID-19. Gross debt is reported to have risen to 80.3% of GDP from 65.6% last year. If left unaddressed, interest costs will become one of the foremost payments of government, with tax payers' money going to bondholders instead of education, health and welfare. More specifically, if the fiscal deficit is left unchecked, South Africa will head into a full-on "debt spiral." Hence, the need to attract local and foreign investment in an emerging market such as South Africa cannot be stressed enough.

By analyzing the market conditions, useful information can be derived, such as the level of risk present and growth prospects (Floros et al., 2020). A traditional method introduced by Markowitz (1952), and used over the years, is the risk-return relationship. The risk-return relationship has led to a number of asset pricing models as well as evaluation methods, contributing to the fields of corporate finance, risk management and portfolio management. For example, asset managers use asset pricing models as a forecast tool to ensure the growth of their client's funds and limit risk exposure. Essentially, the goal of the risk-return relationship is to maximize return and minimize risk.

The relationship between risk and return is assumed to be a positive and linear relationship - "the higher the risk, the higher the return." An accurate prediction leads to the efficient allocation of resources, an increased probability of profitable investment strategies and overall superior returns. Yildiz et al. (2022) stated that the accurate forecast of the risk-return relationship is "vital for asset valuation" and "critical for both the economy and business world." The challenge lies in the choice of the risk variables, as sufficed by the international studies by Stankovic et al. (2020), Fassas and Siriopoulus (2020) and Yildiz et al. (2022).

In South African literature, only one study by Steyn and Theart (2019) is documented to investigate returns in relation to different risk measures, namely, the traditional beta and standard deviation. Investors are often attracted to volatile markets, such as the South African market, in the belief that they can earn superior returns. However, Steyn and Theart (2019) found a negative relationship between risk and return. More importantly, the study found beta to be insignificant, and the standard deviation was significant. This point highlighted that different risk measures have the

ability to impact the significance of the risk-return relationship. Such a finding is in line with international literature that investigated returns with different classes of risk, such as downside risk (Yildiz et al., 2022), realized risk (Floros et al., 2020) and implied risk (Fassas and Siriopoulus, 2020).

Steyn and Theart (2019) is the only known South African study to investigate a different risk measure in the context of the traditional risk-return relationship. In international literature, there is no existing study that investigates all, or the most, different classes of risk in one collective study. Hence, this motivated the investigation of the relationship between returns and different risk measures. Therefore, this is the first South African study to investigate an optimal risk measure in the context of the risk-return relationship.

This study employs the novel Bayesian approach by Jensen and Maheu (2018) and makes an innovative contribution by extending to the methods by Karabatsos (2017) - density regression, quantile regression and survival analysis. The Bayesian approach has gained popularity over the years in the fields of astronomy, psychology and medicine, highlighting its practical efficiency and reliability (Sharma, 2017; Karabatsos, 2017; van Ravenzwaaij et al., 2018). The extended Bayesian approach would be the first to be applied in the field of finance, to the best of the author's knowledge, highlighting the novel insight that the model can provide on the risk-return topic. The Bayesian approach is based on Bayes' (1763) theorem, which refers to the probability estimation of a relationship given prior information. In this case, the relationship is between risk and return, given that the inherent nature of the risk-return variables is taken into account. The prior takes into account the asymmetric, volatile and random properties of the financial data and results in an updated probability estimation known as the posterior.

Jensen and Maheu (2018) and Karabatsos (2017) are the only two known studies, to the best of the author's knowledge, that cover these specific Bayesian methods. The reason for the lack of application of this methodological approach is due to three limitations. First, "the relative lack of familiarity of econometricians and financial economists with Bayesian nonparametric methods and their computational complexity" (Griffin et al., 2018). Second, the prior information can be based on subjective information such as preconceived investor behavior or biases, in the context of this study (Trichilli et al., 2020; Brooks, 2014). Third, the lack of easy-to-use software to implement the model (Sharma, 2017; van Ravenzwaaij et al., 2018). These limitations are overcome as follows. First, by adding to literature on a popular topic of interest—the risk-return relationship. Second, by capturing the risk-return relationship without any predetermined assumptions such as investor behavior. Third, by following the only two known studies that describe these specific Bayesian methods and model implementation.

This study employs a parametric Bayesian approach due to computational efficiency of the software by Karabatsos (2017). A parametric model investigates returns as a linear function of risk in the same way as conventional methods, such as regression analysis and the GARCH approach. However, in contrast to these conventional methods, the Bayesian approach allows for higher moment properties, such as skewness and kurtosis, by clustering analysis, which makes it a more robust choice. The Bayesian approach adjusts to modelling errors and accounts for uncertainty and stochasticity, thus ensuring valid results (Jensen and Maheu, 2018). Employing a more robust model will provide reliable and valuable information to encourage market participants to invest in the South African market.

The layout of this study is as follows. The Introduction provides an overview of the aim of this study, to investigate returns in relation to different risk measures. The Literature Review covers international empirical studies that analyzed groups or different classes of risk and concludes with gaps

found in the literature. The Methodology outlines the data, model specifications of the extended Bayesian approach and the notes for interpretation of the graphical analysis. The Empirical Results are reported with explanations; thereafter, the results are discussed in summary. The Discussion and Conclusion section highlights the key findings of this study and their implications, followed by recommendations made for future research purposes.

2. Literature review

Steyn and Theart (2019) investigated the risk-return relationship for the sample period from July 2004 to September 2018 using basic regression analysis to model the Capital Asset Pricing Model (CAPM). The study analyzed beta and an additional risk measure, standard deviation. It was noted that shares with high systematic risk (beta) resulted in negative returns but were statistically insignificant, whereas standard deviation was significant. This highlighted that different risk measures affected the significance and relationship with returns, especially since the study concluded that the CAPM is an inappropriate model to estimate returns in the South African market.

Sehgal and Pandey (2018) used different risk measures to test the robustness of the risk-return relationship over the sample period June 2005 to June 2016. The study investigated the MSCI World index and MSCI indices of three market categories made up of twenty-one developed, twenty-two emerging and eleven frontier markets. The risk measures used were beta, standard deviation, semi deviation and downside risk. Standard deviation is based on the unrealistic assumption that the return distribution is normal. Downside risk refers to the risk associated with a security's price dropping due to changes in the market conditions. This measure adequately captures risk when the return distribution is asymmetric, according to Yildiz et al. (2022). To elaborate on this point, Hoque and Low (2020) stated that previous literature, from Basher and Sadorsky (2006) to Alles and Murray (2017), dealt with semi-asymmetric properties and did not provide a full representation of the return distribution. This is because the empirical return distribution is subject to asymmetric properties such as kurtosis and skewness. Therefore, other risk measures are considered more practical, such as semi deviation, because it consists of both standard deviation and skewness.

Sehgal and Pandey (2018) estimated the risk-return relationship using a bivariate cross-sectional regression analysis with the Newey-West adjustment to account for autocorrelation and heteroscedasticity in the innovations. It was found that the impact of risk choice had an insignificant impact on the asset pricing analysis. Such a result was due to the use of cross-sectional regression analysis, which fails to account for the time-varying nature of the risk-return variables, according to Yildiz et al. (2022). Yildiz et al. (2022) used panel data analysis to compare the CAPM using beta and downside risk. Panel data analysis effectively accounts for the time-variant and stochastic nature of the risk-return variables. The study analyzed a total of sixteen risk measures from a dataset of twenty developed and twenty-five emerging markets. The study concluded that downside risk is a more optimal risk measure than beta on both a domestic and global level.

Stankovic et al. (2020) tested the mean-variance model by algorithms consisting of quadratic programming, using different measures of downside risk to improve the forecast ability of the risk-return relationship. The study investigated groups of shares from the Belgrade Stock Exchange (BSE), which were made up of the BELEXline and BELEX15 indices, over the sample period from 1 January 2008 to 31 December 2018. The risk measures extended from portfolio variance (PV) and included absolute deviation (AD), lower semivariance (LSV), lower semi absolute deviation (LSAD) and Value

at Risk (VaR). In optimal order of the risk measures, it was found to be AD and LSAD with identical ratios (0.81), PV (0.21), VaR (0.17) and LSV (0.13). Thus, the AD and LSAD risk measures of the BSE had the highest rate of return. Therefore, by analyzing different risk proxies, investors can determine which measure optimizes their return.

In contrast to the positive and linear risk-return relationship found by Stankovic et al. (2020), Fassas and Siriopoulus (2020) found a negative and asymmetric relationship. Fassas and Siriopoulus (2020) investigated the risk-return relationship using sixty-eight implied volatility indices and twenty-five from academic literature. Implied volatility is known as a “forward looking” measure of risk. The authors used quantile regression due to its improved ability to account for outliers and asymmetry. This approach was especially suitable since the daily returns and implied volatility were time-varying and followed an asymmetric distribution. The study found that the implied volatility contained information about future volatility that exceeded historical volatility.

However, Floros et al. (2020) stated that realized risk measures, constructed with high frequency (HF) data, provide more accurate parameter estimates of latent volatility. More information is contained in HF data with respect to intraday movements and microstructure noise in the market. The study analyzed the S&P500 and FTSE100 indices and thirteen realized risk measures over January 2000 to June 2017. The risk-return relationship was estimated by regression analysis, where most of the realized risk measures were found to improve capturing the risk-return relationship. Floros et al. (2020) recommended investigating the forecast accuracy of the risk measure class of implied variance. Moghaddam et al. (2020) conducted a comparative analysis, where the authors found that realized volatility outperformed implied volatility, due to the latter being inconsistent in risk estimation.

In addition, Jensen and Maheu (2018) investigated the risk-return relationship using realized variance (RV) and a bias-adjusted realized variance (ARV). RV is considered an appropriate risk measure in the context of Bayesian modelling. It was found that the ARV had a lower standard deviation because the bias-adjustment optimally captured the risk. More specifically, it accounted for the microstructure noise in the market, in line with Floros et al. (2020).

In conclusion, the South African study by Steyn and Theart (2019) found that different risk measures affected the quantification of the risk-return relationship. Moreover, the authors used only one different risk measure to optimize the risk-return relationship. The first gap identified is that there is no existing study that investigates all, or rather, the greatest number of different classes of risk in one collective study. Standard deviation is considered an appropriate risk measure when the return distribution is normal (Sehgal and Pandey, 2018; Steyn and Theart, 2019). However, it is widely documented in empirical literature that returns follow an asymmetric distribution (Stankovic et al., 2020; Yildiz et al., 2022; Fassas and Siriopoulus, 2020).

The second gap is by Hoque and Low (2020), who stated that previous literature, from Basher and Sadorsky (2006) to Alles and Murray (2017), dealt with semi-asymmetric properties. Therefore, this study effectively accounts for asymmetry by employing the Bayesian approach and investigates different risk measures that are considered more empirically viable in estimating the risk-return relationship. The third gap is that the aforementioned literature did not provide a full representation of the return distribution. The extensions of the parametric Bayesian model employed in this study shall provide a full graphical representation of returns in relation to the different risk measures. This provides a strong motivation for the methodology employed in this study.

3. Methodology

Figure 1 shows the methodological steps to investigate the optimal risk measure for the risk-return relationship.

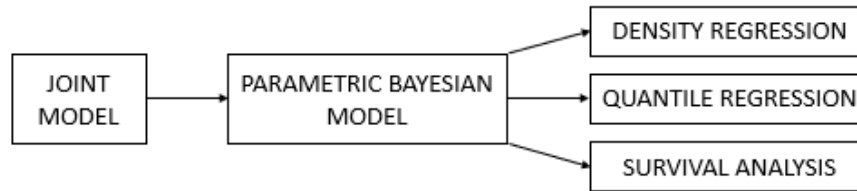


Figure 1. Extended Bayesian approach to find the optimal risk measure for the risk-return relationship.

The Bayesian approach consists of a prior and posterior (Bayes, 1763). The joint distribution between returns and the different risk measures are modelled by a parametric Bayesian model. The prior takes into account the inherent nature of the financial data—the asymmetric, volatile and random properties of the risk-return variables. The posterior is determined by the slice sampling technique by Walker (2007), Kalli et al. (2011) and Papaspiliopoulos (2008) as well as the Gibbs sampler. Thereafter, the extensions are made by density regression, quantile regression and survival analysis.

3.1. Data

Following standard South African literature, this study will obtain the secondary price data from the Integrated Real-time Equity System (IRESS) database and the risk-free rate proxy, the South African Treasury-bill (T-bill), from the South African Reserve Bank (SARB). According to Floros et al. (2020), HF data contains a high degree of information in terms of intraday movements and microstructure noise; therefore, this study will analyze the daily data of the Johannesburg Stock Exchange All Share Index (JSE ALSI). An updated sample period, from 16 March 2000 to 17 March 2022, is used based on data availability and accessibility. The ALSI price data is converted to market returns by $R_t = \ln\left(\frac{P_t}{P_{t-1}}\right)$ where R_t = market returns, P_t = share price for the current day t , and P_{t-1} = share price for the previous day $t - 1$. The annual risk-free rate is converted to a daily value by $\text{Daily } R_f = (1 + \text{yearly } R_f)^{\frac{1}{365}} - 1$ as given by Brooks (2014). Excess returns are computed as the difference between market returns and the risk-free rate.

The risk measures are computed as follows. Let R_i = returns, \bar{R} = average returns and T = number of observations.

Standard deviation is a measure of total risk which comprises systematic and unsystematic risk, following Steyn and Theart (2019) and Sehgal and Pandey (2018):

$$SD = \sqrt{\frac{\sum_{i=1}^T (R_i - \bar{R})^2}{T}} \quad (1)$$

The downside risk measures follow Stankovic et al. (2020) and Sehgal and Pandey (2018). Absolute deviation is given as:

$$AD = \frac{1}{T} \sum_{t=1}^T |R_t - \bar{R}| \quad (2)$$

Lower semivariance:

$$LSV = \frac{1}{T} \sum_{i=1}^T \sum_{t=1}^N (\bar{R} - R_i), \bar{R} > R_i \quad (3)$$

Lower semi absolute deviation:

$$LSAD = \frac{1}{T} \sum_{t=1}^T \left| \sum_{i=1}^N \bar{R} - R_i \right|, \bar{R} > R_i \quad (4)$$

Following Floros et al. (2020) and Jensen and Maheu (2018), realized variance is given as:

$$RV = \sum_{t=1}^{N_{t-j}} R_{t,i} R_{t,i+j} \quad (5)$$

The bias-adjusted realized variance is given by Jensen and Maheu (2018):

$$ARV_t^q = \hat{\gamma}_0 + 2 \sum_{j=1}^q \left(\frac{1-j}{(q+1)} \right) \hat{\gamma}_j, \text{ where } \hat{\gamma}_j = \sum_{t=1}^{N_{t-j}} R_{t,i} R_{t,i+j} \quad (6)$$

where ARV_t^q = bias-adjusted realized variance, q = bias-adjustment and is set to one by Hansen and Lunde (2006).

From the right-hand side, $\hat{\gamma}_0$ = first order of autocorrelation, the second term is the product of $\hat{\gamma}_j$ = bias-adjustment and $\hat{\gamma}_j$ = realized variance by definition. The addition of Bartlett weights is not shown in the equation but is added to ensure a positive bias-adjusted realized variance (Hansen and Lunde, 2006; Jensen and Maheu, 2018).

3.2 Model specifications of the Bayesian Approach

The model specifications are given by Jensen and Maheu (2018), the only known study to the best of the author's knowledge. First, the joint distribution of the return and risk variables is modelled and given as:

$$P(r_t, x | I_{t-1}, \Omega, \Theta) = \sum_{j=1}^{\infty} w_j f(r_t, x | \theta_j, I_{t-1}) \quad (7)$$

where r_t = excess returns, x = risk, I_{t-1} = information set that contains the risk and return variables, $\Omega = \{w_j\}$ set of probability weights, and $\Theta = \{\theta_j\}$ set of parameters, for j = mixture clusters. To simplify the model, Ω and Θ are dropped. The joint distribution is equal to the product of the marginal and conditional distribution by the law of total probability.

$$P(r_t, x | \theta_j, I_{t-1}) = f(r_t | \theta_j, I_{t-1}) * f(x | \theta_j, I_{t-1}) \quad (8)$$

Thereafter, the posterior is determined using the slice sampler by Walker (2007), Kalli et al. (2011) and Papaspiliopoulos (2008) as well as the Gibbs sampling technique.

$$P(r_t, x, u_t | \theta_j, I_{t-1}) = \sum_{j=1}^{\infty} 1(u_t < w_j) * f(r_t, x | \theta_j, I_{t-1}) \quad (9)$$

The slice sampler ensures the elimination of all masses of zero by the auxiliary variable u_t . The Gibbs sampling technique allows the sharing of properties between the initial and updated probability estimation, such as the higher moment properties of skewness and kurtosis.

The extensions are made from the above given parametric Bayesian model. The given Bayesian model is denoted as m . The model specifications of the extensions are given by Karabatsos (2017), the only known study to the best of the author's knowledge.

In the context of this study, let y = excess returns, x = the risk measures, n = number of observations and m = the given Bayesian model. Note that the returns y are conditional on the prior information, which are the risk measures x and given Bayesian model m , as well as the inherent asymmetric nature of the risk-return variables.

The density regression consists of two types of functions, namely, a probability density function (pdf) and cumulative distribution function (cdf). Drawn from Devore (2012), both functions provide a probability estimation or likelihood of the risk-return relationship. The pdf and cdf are given as Equations 10 and 11, respectively:

$$f_n(y|x) = \int_a^b f(y|x, m) \quad (10)$$

The pdf $f_n(y|x)$ estimates the likelihood of the risk-return relationship, which can take on values in the interval from a to b . Figure 2 shows the pdf.

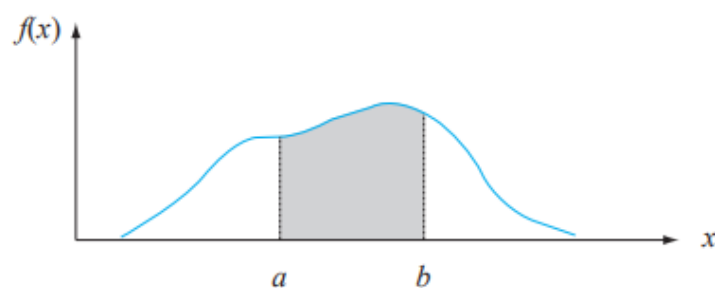


Figure 2. Probability Density Function.

From Figure 2, the pdf is represented as the area under the curve with the interval ranging from a to b .

$$F_n(y|x) = \int_{-\infty}^y F_n(y|x, m) = \int_{-\infty}^y f_n(Y \leq y|x, m) \quad (11)$$

The cdf $F_n(y|x)$ estimates the cumulative probability estimation of the risk-return relationship defined as the random variable y , of which the real value variable Y is less than or equal to y .

Quantile regression (QR) is used as it has an improved ability to capture asymmetric properties and outliers, as noted by Fassas and Siriopoulus (2020).

$$QR = F_n^{-1}(u|x, m) \text{ for some chosen } u \in [0,1] \quad (12)$$

By Karabatsos (2017), QR is equal to the u th quantile, which is an element of the bounded values from zero to one, given the entire sample space of the variables and model denoted by x and m . For example, the median is shown by $QR = F_n^{-1}(0.5|x, m)$.

According to Gao et al. (2019), survival analysis, also known as time-to-event analysis, is a branch of statistics that focuses on how long it takes before a certain event occurs. The application of survival analysis is common in the field of medicine, according to Emmert-Streib and Dehmer (2019). Its application in finance is relatively limited, with early studies mainly focusing on forecasting financial distress and stock market pricing, as sufficed by Gao et al. (2019) and Gepp and Kumar (2015). Moreover, there is no study yet to extend the Bayesian approach by Jensen and Maheu (2018) to survival analysis to investigate the risk-return relationship.

While the survival function focuses on the “life” of the risk-return relationship, the hazard function focuses on its “death.” The survival curve $S_n(y|x)$ is given by Karabatsos (2017):

$$S_n(y|x) = 1 - F_n(y|x, m) \quad (13)$$

The survival function shows the probability estimation of how long the event of interest would survive before a specified period of time (Emmert-Streib and Dehmer, 2019). From Equation 13, it is equivalent to the difference between one and the cdf, by Karabatsos (2017). Related to the survival function is the hazard function, known as the hazard rate or risk function, given as:

$$H_n(y|x) = \frac{f_n(y|x, m)}{S_n(y|x, m)} \quad (14)$$

In this case, the hazard function is shown as a ratio of the pdf and survival function. The hazard function $H_n(y|x)$ gives the probability estimation of how long the event of interest would fail for a duration of time, according to Emmert-Streib and Dehmer (2019). More specifically, the hazard rate estimates an instantaneous failure rate at time t , assuming survival up until t . The cumulative (CML) hazard rate gives the probability estimation of accumulated risk of how long the event of interest would fail for a duration of time. It is given as:

$$H(t) = \int_0^t h(\tau) d(\tau) \quad (15)$$

The CML hazard rate $H(t)$ is given as an integral, which provides the total amount of risk exposure $h(\tau)$ over the interval from 0 to t .

3.3. Notes for the Bayesian approach

The numerical analysis, not shown in this study for the extensions, consists of selecting optimal model specifications and model diagnostics to ensure robust Bayesian test results. The Monte Carlo (MC) mixing values are shown for the preliminary Bayesian model. An approximate value of 0.5 indicates optimal parameter estimates, according to Karabatsos (2017). The original parametric Bayesian model is investigated by normality tests to ensure that the model has adequately captured asymmetry. The

normality tests employed are the Kolmogorov-Smirnov (KS) test and quantile-quantile (QQ) plots. If asymmetry is adequately captured, this would confirm accurate and reliable Bayesian test results.

This study focuses on the graphical analysis to provide conclusive evidence on the risk-return relationship. The graphical analysis consists of 3-D plots of returns with the different risk measures. In the Bayesian approach, the mean value is the probability estimation that is analyzed. Any measure of central tendency can be used—mean, mode or median—but this study used the mean, in line with standard literature. With respect to the graphical analysis, the point estimate can be easily read off the peak of the pdf. The peak of the density represents the measure of central tendency—the mean value. The probability estimation of the cdf is read off as the average of the density. In this case, it is not read off as the peak of the density because the highest point is one for all the densities. That is because a cdf accumulates to one. The quantile regression plot consists of the mean value (on the y-axis), slope and 95% confidence intervals (dotted lines). The direction of the slope indicates the nature of the relationship between risk and return. A downward slope indicates a negative risk-return relationship and an upward slope indicates a positive relationship. The survival analysis functions are interpreted in the same way as the cdf. The probability estimation of the survival functions is read off as the average of the density.

Although not shown in this study, the trellis chart by Karabatsos (2017) can also be used to determine the probability estimations. The trellis chart is a consolidation of multiple charts. In the context of this study, it organizes the 3-D plots into a 2-D diagram, making it an additional convenient method to determine and read off the approximate probability estimations. To determine the probability estimation, trace a line from zero (average midpoint) on the x-axis of RT (excess returns r_t) to the respective curve, then trace a line from that point (of the curve) to the probability on the y-axis. The curves are estimated for the minimum and maximum values of the risk measures. Thus, the point estimate average is taken as the overall value, in line with Bayesian literature.

4. Empirical results and discussion

4.1. Data exploration

The daily closing price data of the JSE ALSI are obtained from IRESS for the sample period from 16 March 2000 to 17 March 2022. The ALSI price data is converted to market returns by taking the natural log form of the difference between the current and previous day's closing prices. The annual T-bill obtained from the SARB is converted to a daily value. Excess returns are calculated as the difference between the ALSI market returns and risk-free rate. The risk measures are calculated, using the specified formulas, following literature. Before analysis, the (excess) ALSI return data and risk measures are confirmed to be stationary to avoid a spurious regression and ensure a valid time series. The statistical properties of the ALSI returns are confirmed to be asymmetric, volatile and stochastic. Thus, justifying the application of the Bayesian approach to take into account the inherent nature of the risk-return variables.

Although not within the scope of this paper, the model specifications for the parametric Bayesian model are as follows. The prior variance of the slope parameters is 1000000 and the prior inverse gamma distribution of the error variance is 0.0001. The model specification values of the Bayesian approach are selected in accordance with the model diagnostics to ensure an optimal model.

The posterior parameter estimates are determined by 20 000 MCMC sampling iterations, a burn-in period of 5000 and a thin number of 5. The MCMC sampling iterations and burn-in period values follow Jensen and Maheu (2018), while the thin number is based on the default value by Karabatsos (2017). The MCMC sampling iterations refer to the repetitive resampling process used to determine the posterior parameter estimates. The burn-in period refers to samples from the initial stages that are discarded due to no longer being able to accurately represent the required distribution. The thin number of 5 means that every fifth sampling iterate of the 20 000 MCMC sampling iterations is collected to determine the posterior estimates.

4.2. Bayesian approach

Table 1 shows the test results of the original Bayesian approach.

Table 1. Posterior parameter estimates of the given Bayesian model.

	Mean	75% Bayesian Interval	97.5% Bayesian Interval	MC Mixing Value
SD	-0.020	(-0.060, 0.020)	(-0.132, 0.097)	0.506
AD	-0.036	(-0.077, 0.008)	(-0.166, 0.090)	0.501
LSAD	-0.000	(-0.000, 0.000)	(-0.000, 0.000)	0.514
LSV	0.000	(-0.000, 0.000)	(-0.000, 0.000)	0.498
RV	-0.000	(-0.001, 0.000)	(-0.002, 0.001)	0.484
ARV	0.309	(-0.330, -0.303)	(-0.354, -0.278)	0.508

From Table 1, in the 75% Bayesian interval, only the ARV is significant, whereas, in the 97.5% interval, all the risk measures are significant, with the exception of ARV. The MC mixing values indicate optimal posterior estimates since the values are all approximately 0.5. Normality tests are employed to the posterior parameter estimates, where asymmetry is adequately captured by the KS test and QQ plots, confirming the robustness of the parametric Bayesian model. The extended Bayesian approach will provide more insight into returns in relation to the different risk measures.

4.3. Extended Bayesian approach

4.3.1. Density regression

Figures 3 to 8 show the pdf of returns with the different risk measures.

From Figures 3 to 8, the probability estimation is a consistent value of 30 for the risk-return relationship in the South African market, which is a relatively low value. This is in contrast to Sehgal and Pandey (2018), who stated that given the volatile nature of emerging markets, it is expected that investors earn superior returns for taking on a high level of risk in their investments. However, this is in line with Steyn and Theart (2019), who found the low-risk anomaly (low risk-return relationship) in South Africa. Figures 9 to 14 show the cdf of returns with the different risk measures.

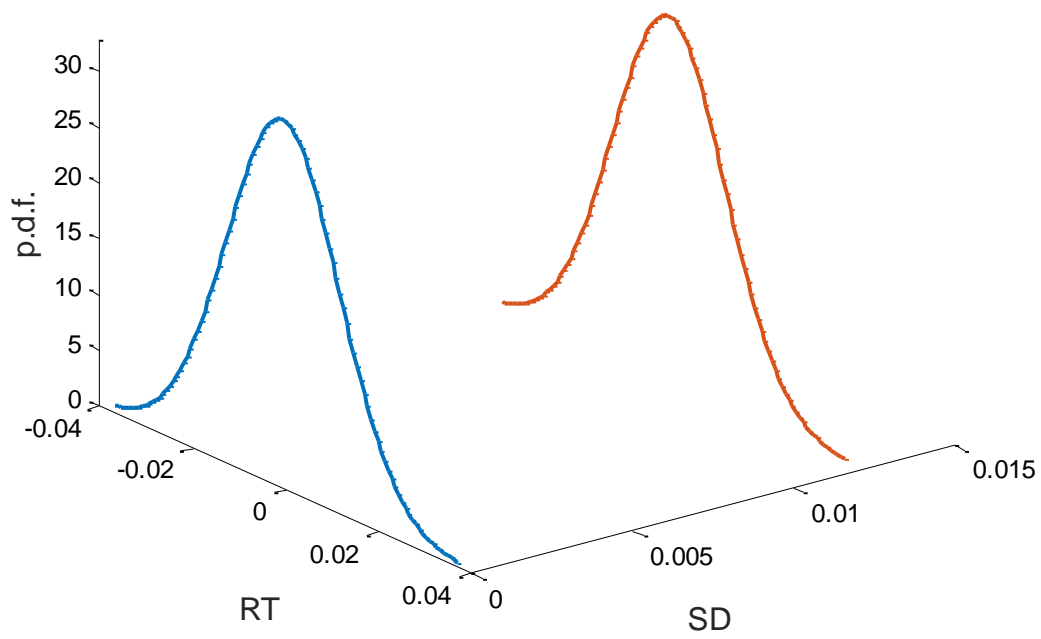


Figure 3. Pdf of returns and SD.

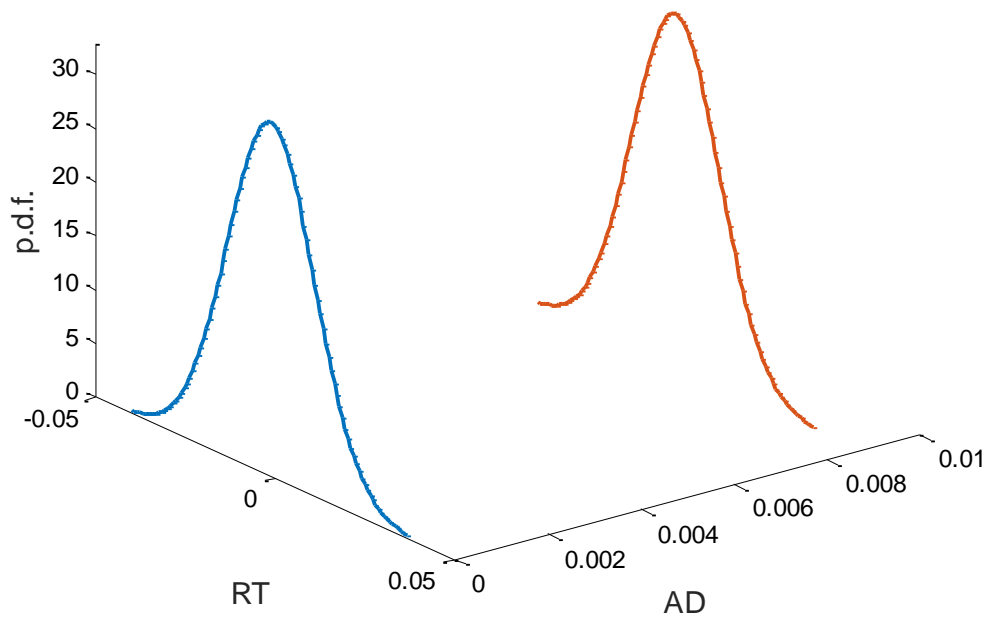


Figure 4. Pdf of returns and AD.

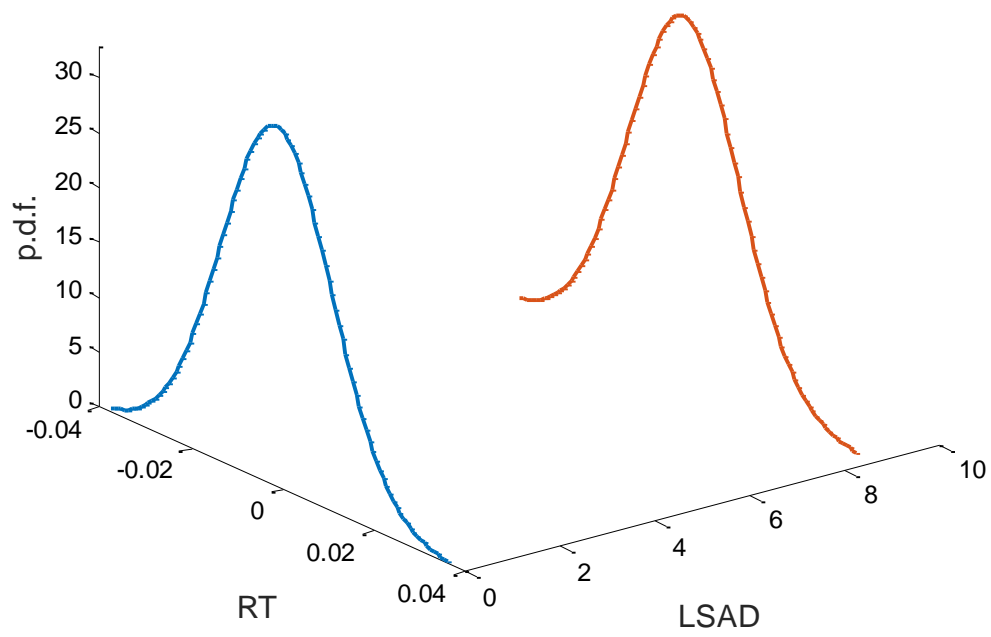


Figure 5. Pdf of returns and LSAD.

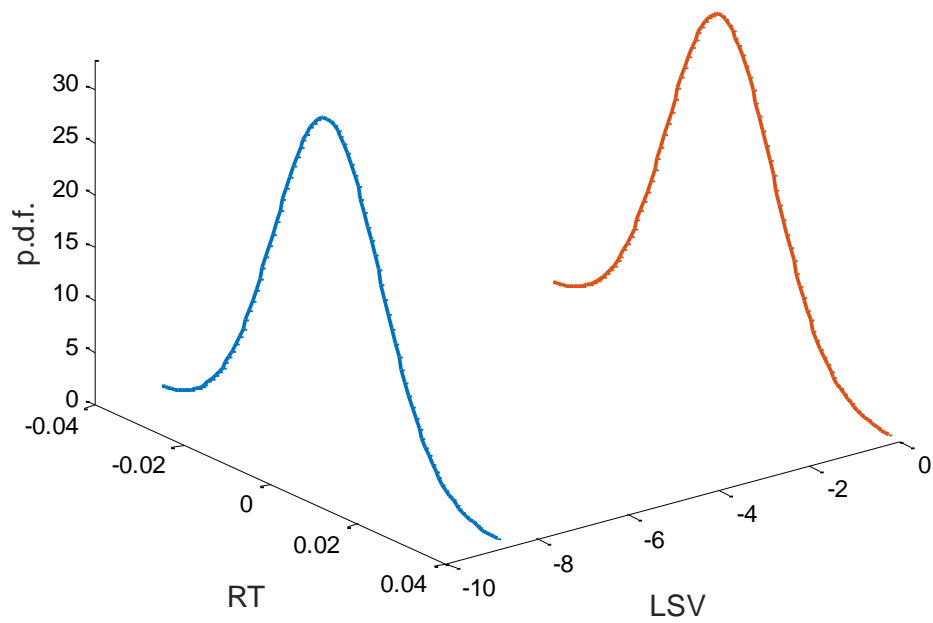


Figure 6. Pdf of returns and LSV.

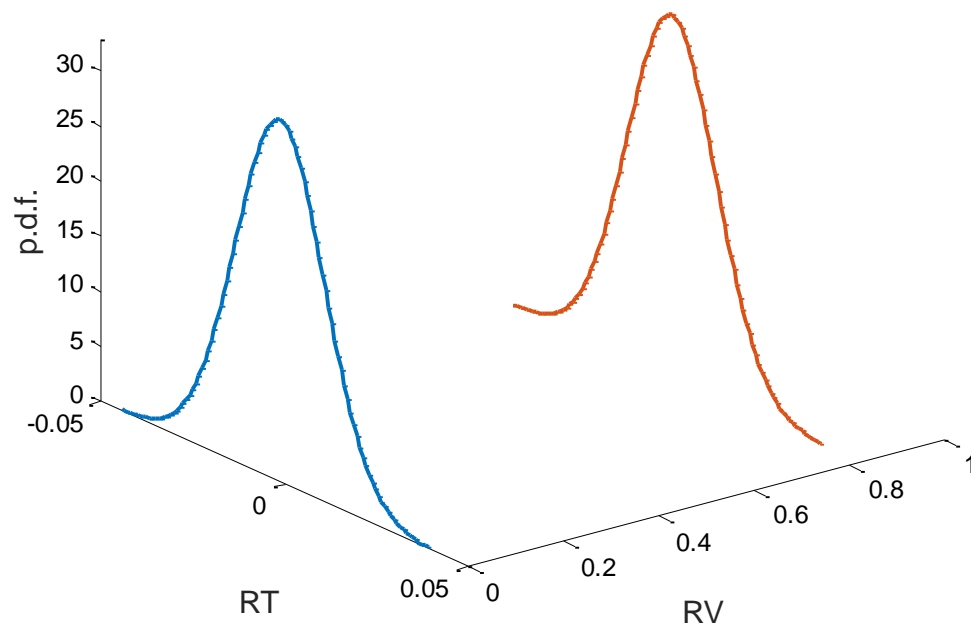


Figure 7. Pdf of returns and RV.

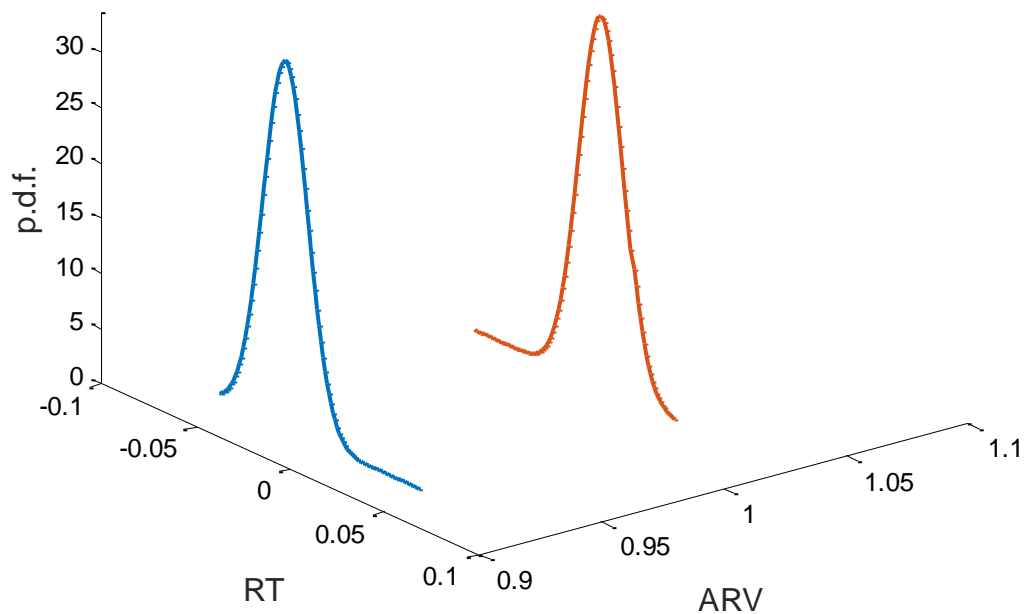


Figure 8. Pdf of returns and ARV.

From Figures 9 to 14, the probability estimation is an overall value of 0.589 for the risk-return relationship in the South African market. This means that the accumulated risk-return relationship is of a substantial value in the market. More specifically, from highest to lowest, the probability estimations of the risk-return relationship are 0.735 (ARV), 0.570 (AD), 0.565 (LSAD), 0.560 (LSV), 0.555 (RV) and 0.550 (SD).

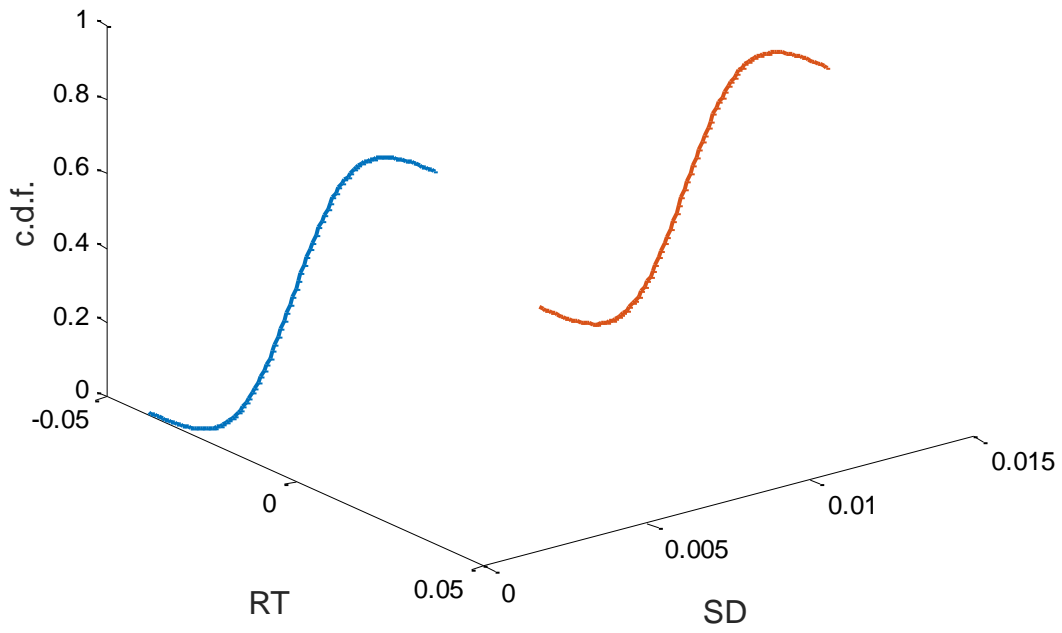


Figure 9. Cdf of returns and SD.

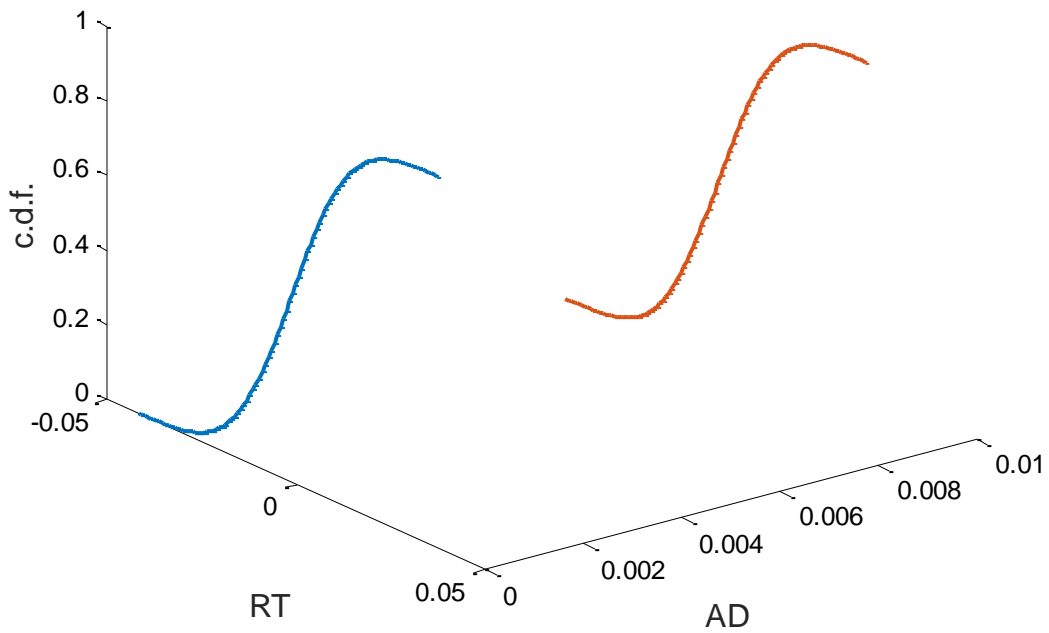


Figure 10. Cdf of returns and AD.

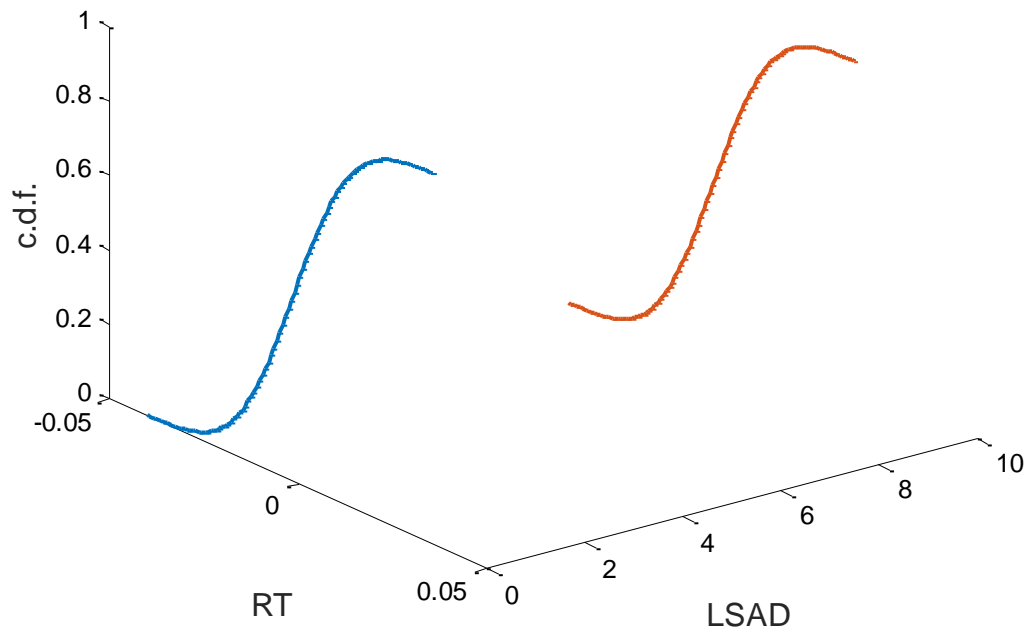


Figure 11. Cdf of returns and LSAD.

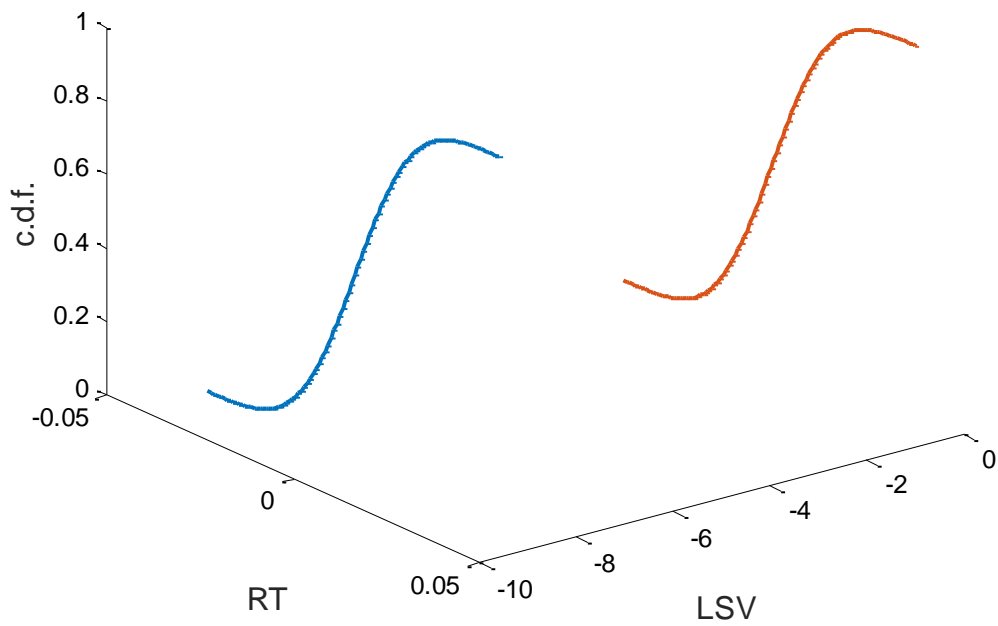


Figure 12. Cdf of returns and LSV.

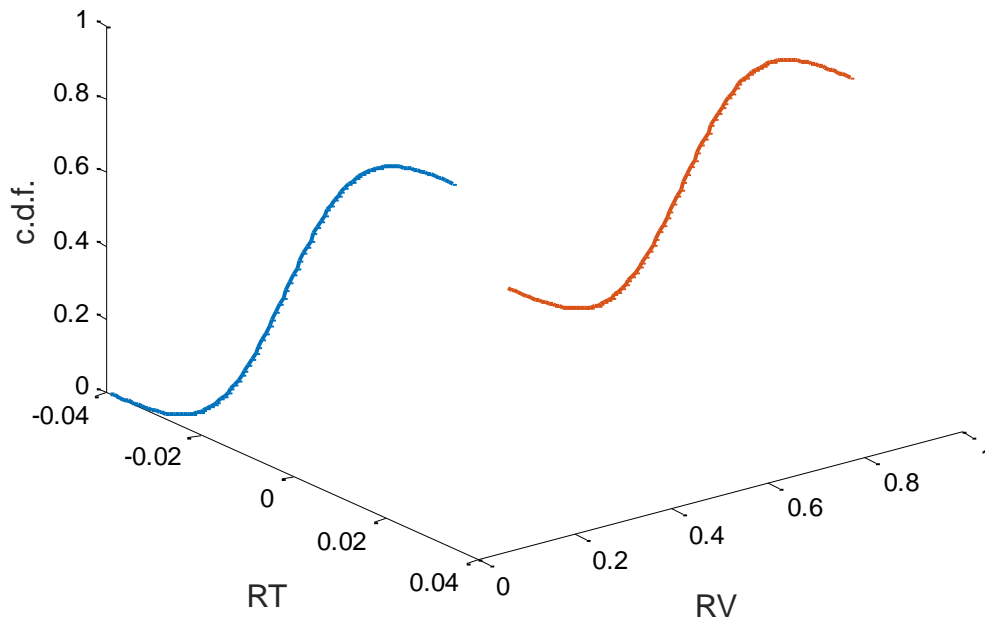


Figure 13. Cdf of returns and RV.

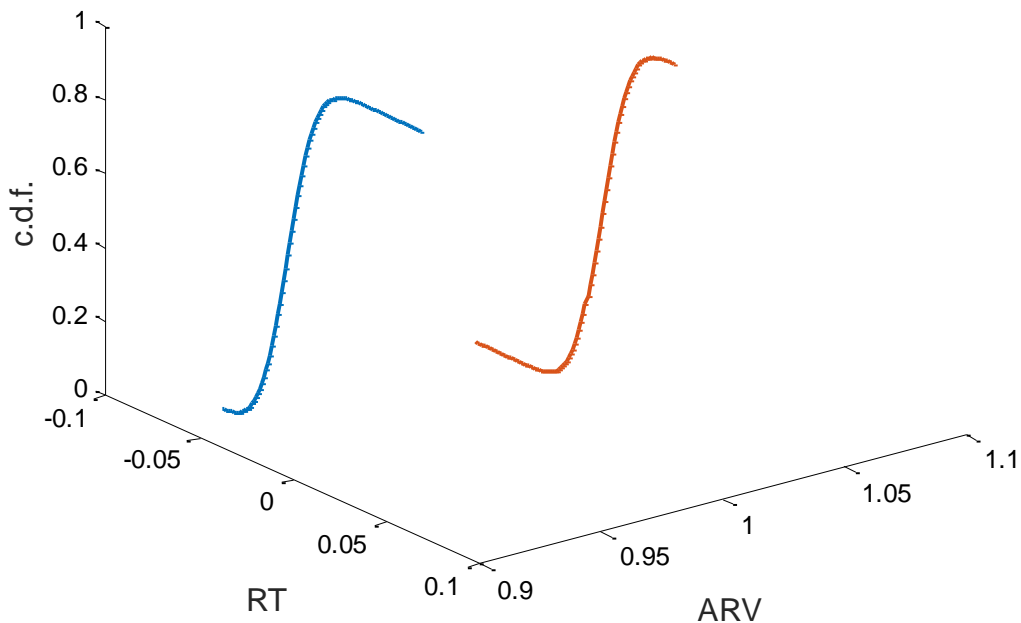


Figure 14. Cdf of returns and ARV.

The key findings are that ARV (0.735) has the highest probability estimation, whereas SD has the lowest (0.55). ARV is considered optimal in capturing asymmetry, latent volatility and microstructure noise, in the context of Bayesian modelling, whereas SD is not since it assumes a normal distribution. The rest of the risk measures are in the range of 0.55 to 0.57, which is substantial. From the density regression, pdf and cdf, it can be concluded that although the chance of investors

earning a superior return is substantial, the probability of abnormal returns does increase over time in the South African market.

4.3.2. Quantile regression

Figures 15 to 20 show the quantile regression plots of returns with the different risk measures.

From Figures 15 to 20, a common observation is that the mean is negative for the risk-return relationship, suggesting a low level of returns in the South African market. This finding is in line with Steyn and Theart (2019). However, it is interesting to note varying shapes for the risk-return relationship, confirming the impact different measures of risk have on returns, in line with Stankovic et al. (2020) and Steyn and Theart (2019). The risk measures SD, AD, LSAD and RV have a negative relationship with returns. In contrast, only LSV and the ARV have a positive risk-return relationship, with the latter showing the steepest slope of all the risk measures. Thus, investors should incline to ARV followed by LSV to optimize their portfolio returns and opt to use the remaining risk measures for diversification purposes or to short sell those stocks.

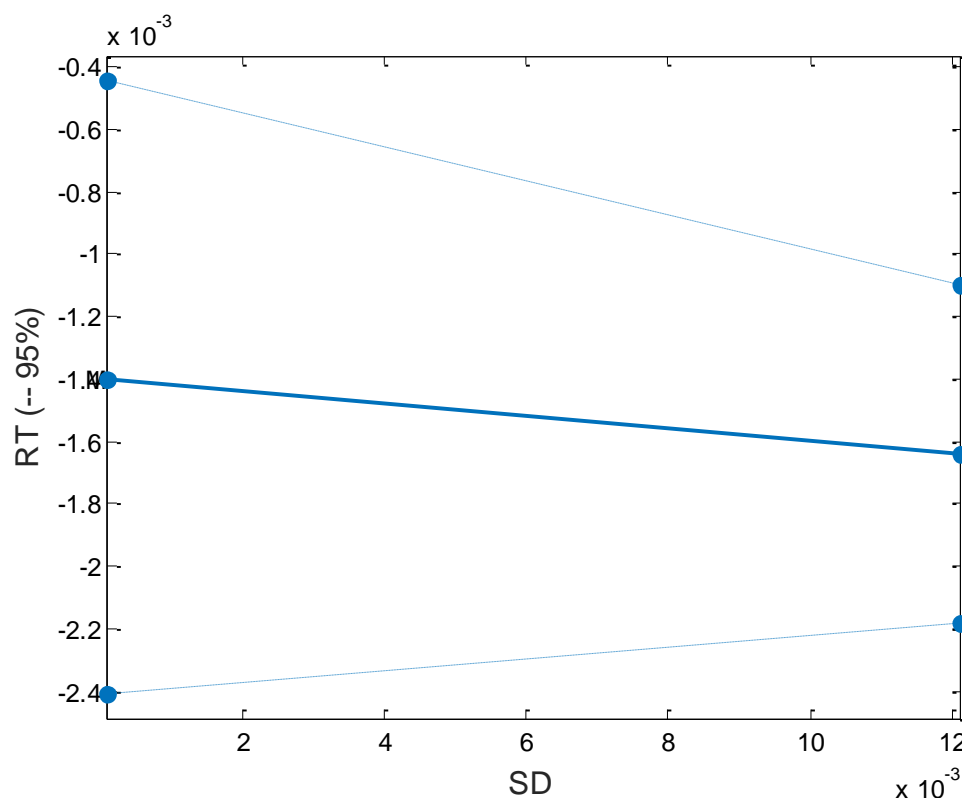


Figure 15. QR of returns and SD.

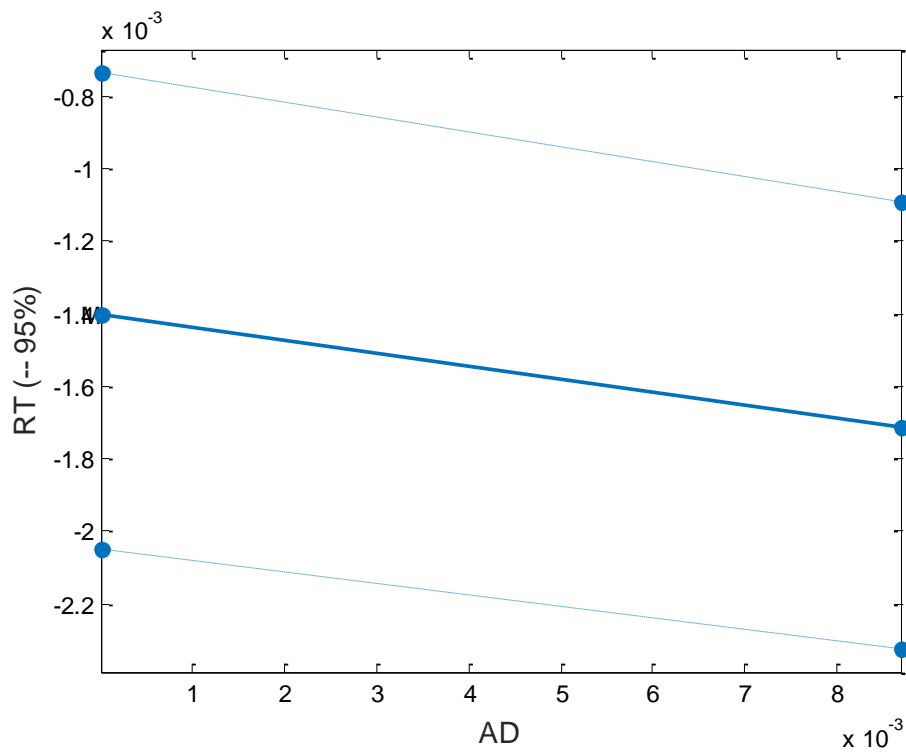


Figure 16. QR of returns and AD.

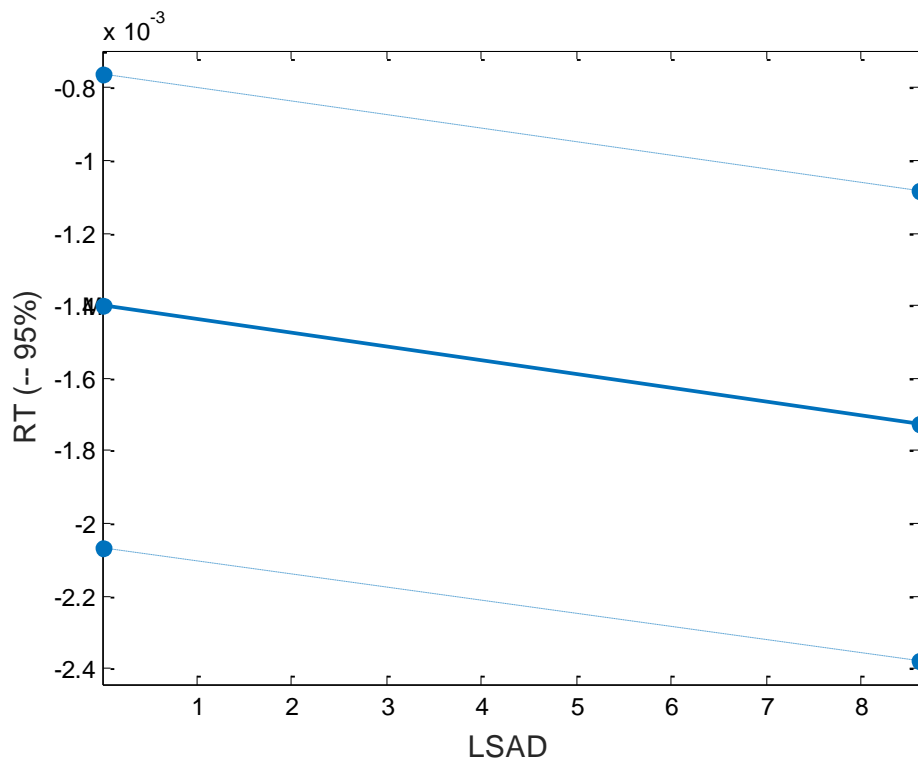


Figure 17. QR of returns and LSAD.

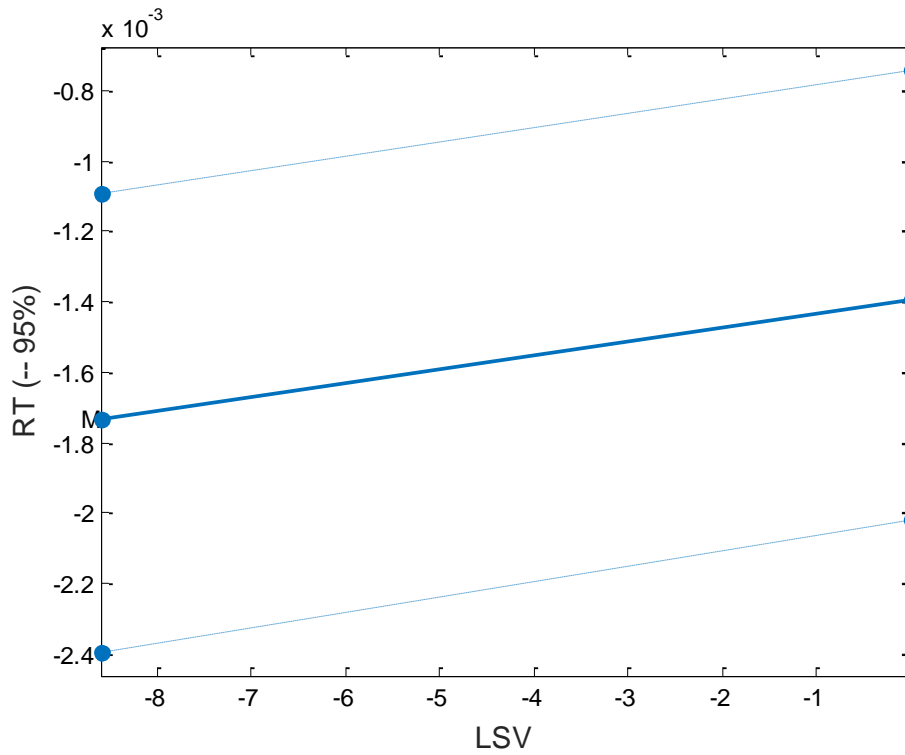


Figure 18. QR of returns and LSV.

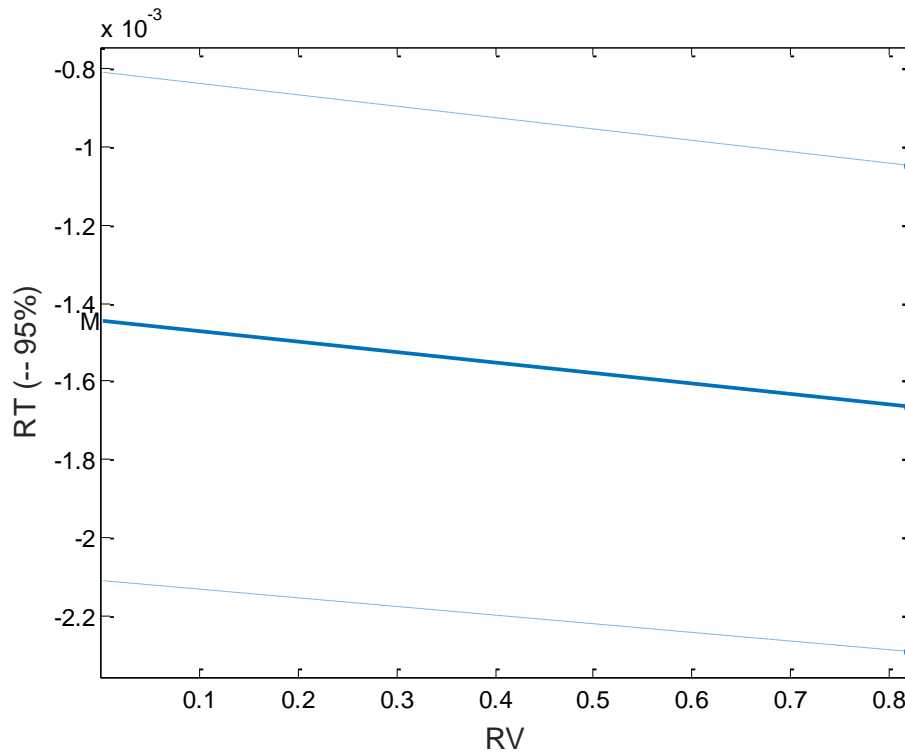


Figure 19. QR of returns and RV.

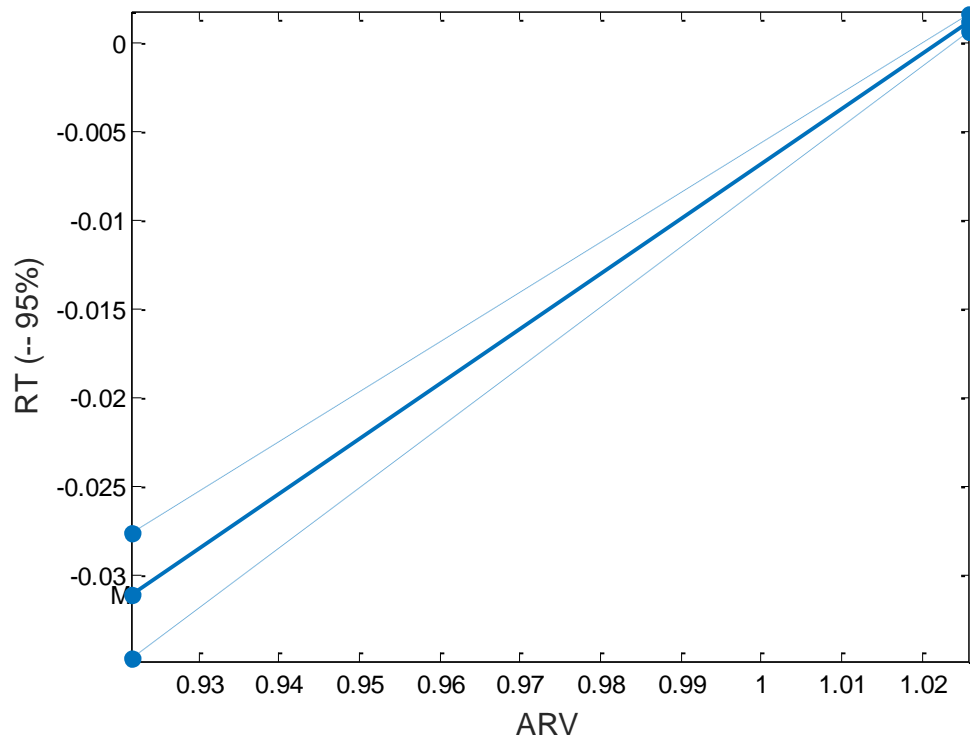


Figure 20. QR of returns and ARV.

4.3.3. Survival analysis

Figures 21 to 26 show the survival function of returns with the different risk measures.

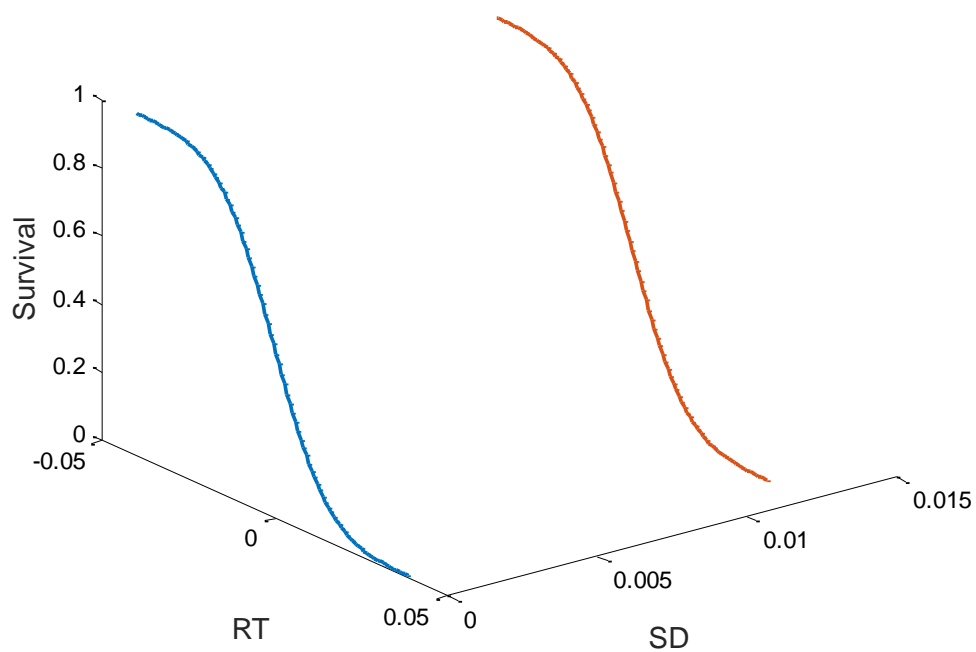


Figure 21. Survival function of returns and SD.

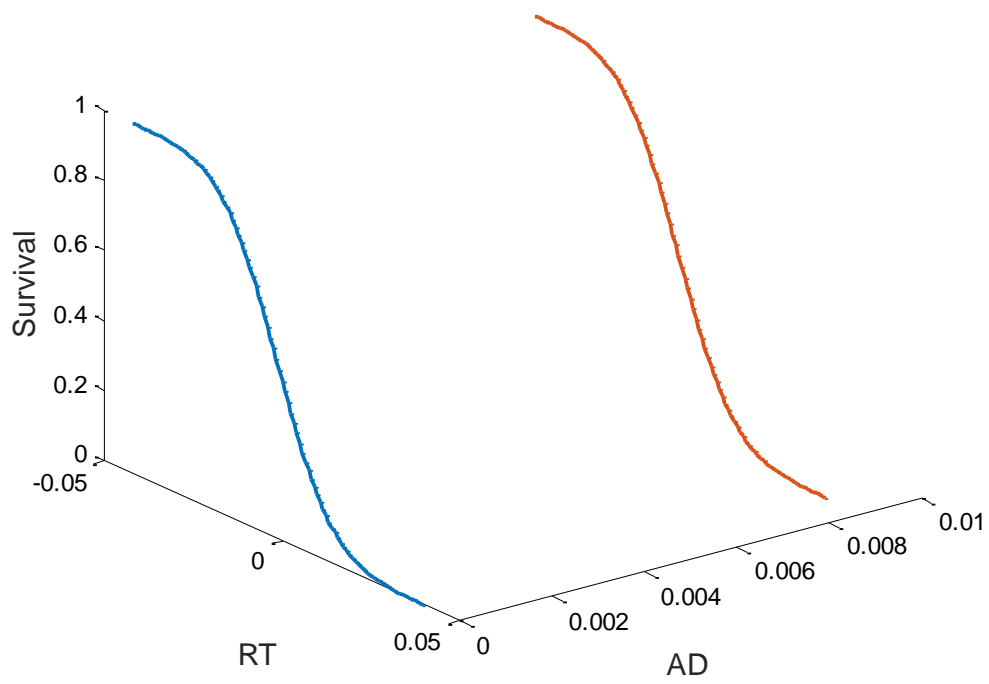


Figure 22. Survival function of returns and AD.

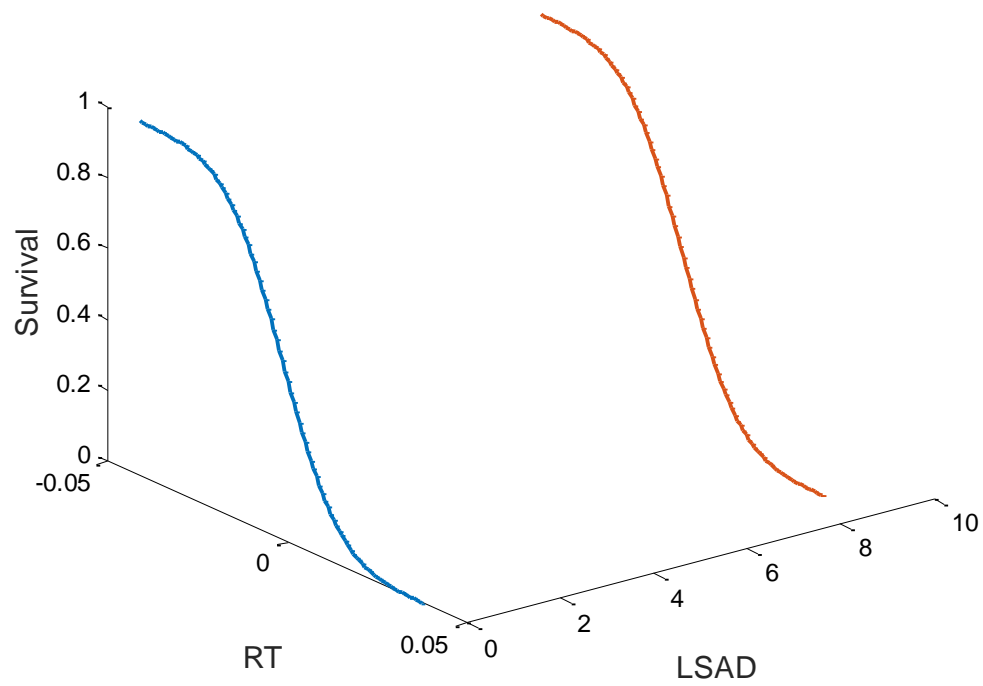


Figure 23. Survival function of returns and LSAD.

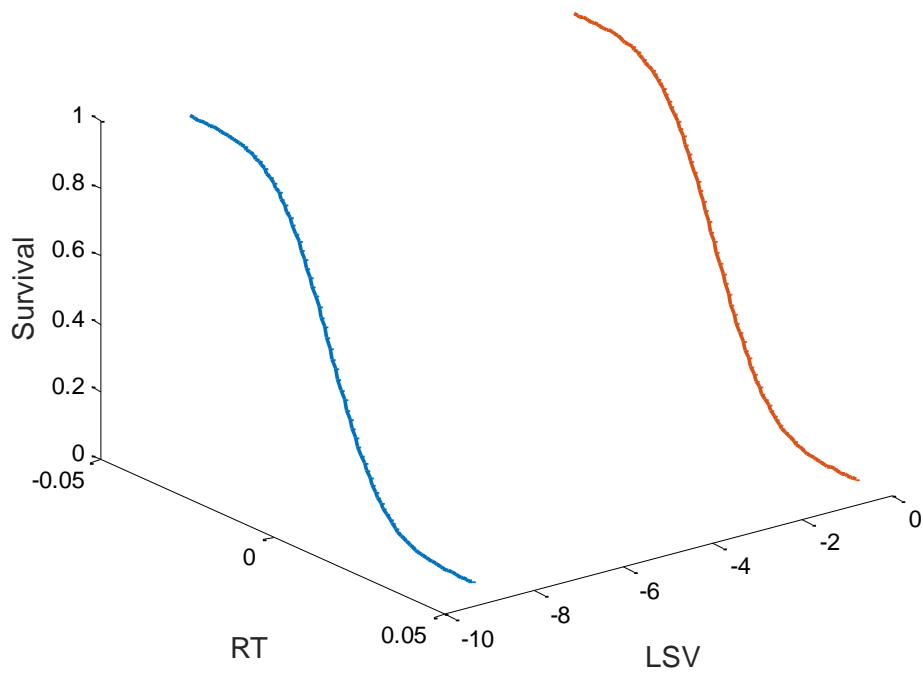


Figure 24. Survival function of returns and LSV.

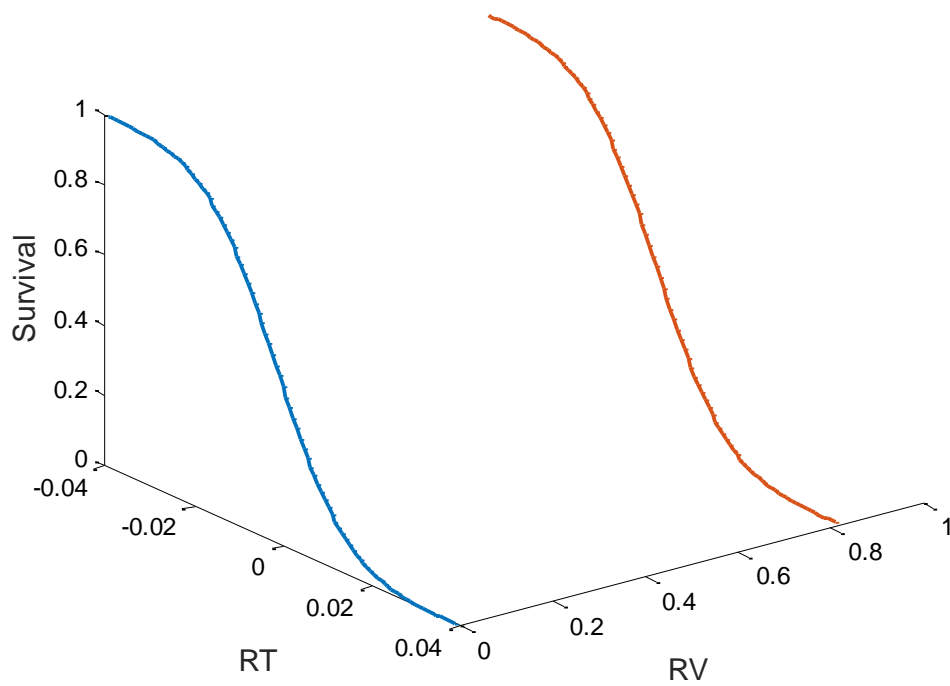


Figure 25. Survival function of returns and RV.

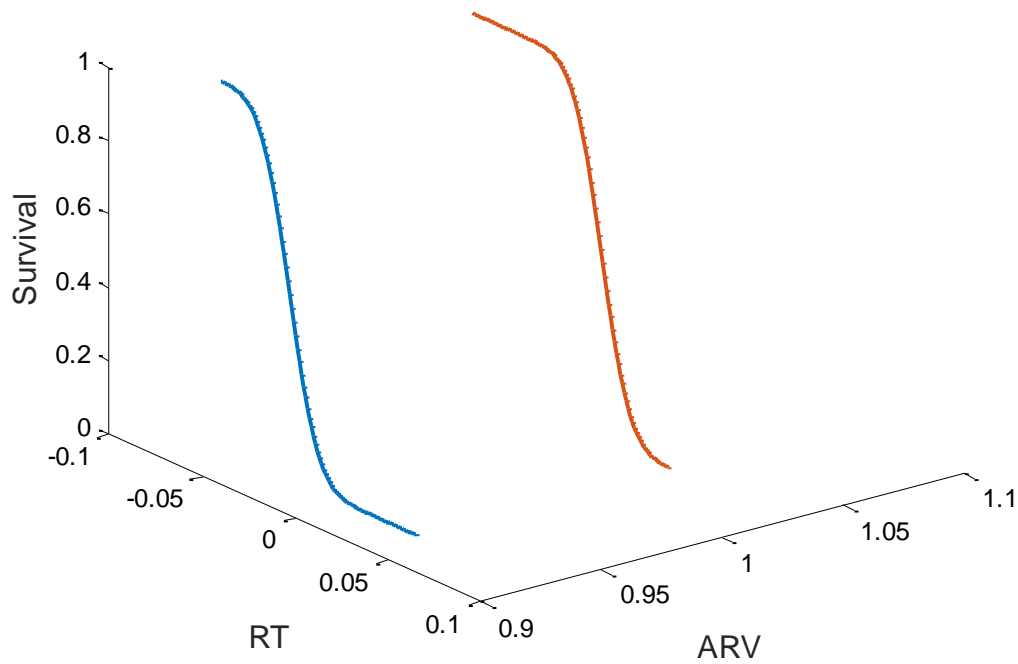


Figure 26. Survival function of returns and ARV.

From Figures 21 to 26, the average probability estimation is 0.420 for the life of the risk-return relationship in the South African market. Given the volatile nature of the market, a consistent and sustained positive risk-return relationship would not be possible, hence the limited timespan. From highest to lowest, the probability estimations of the life of the risk-return relationship are 0.460 (LSAD), 0.550 (SD), 0.450 (LSV), 0.445 (RV), 0.443 (AD) and 0.265 (ARV). The average life of the risk measures ranges from 0.445 to 0.455, with the exception of ARV which has the shortest lifespan.

Figures 27 to 32 show the hazard function of returns with the different risk measures.

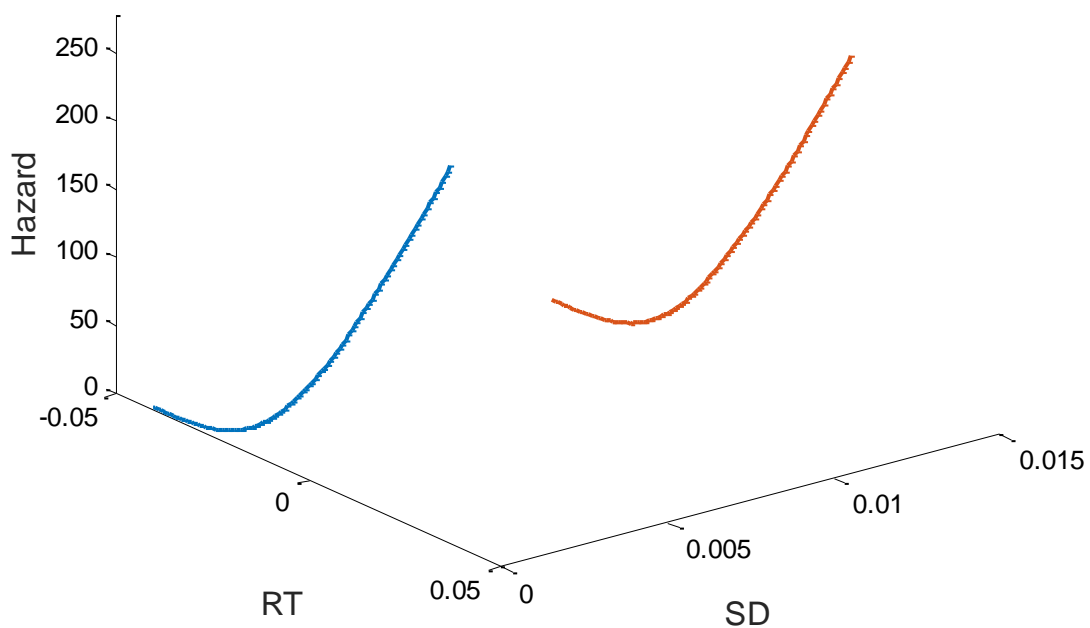


Figure 27. Hazard function of returns and SD.

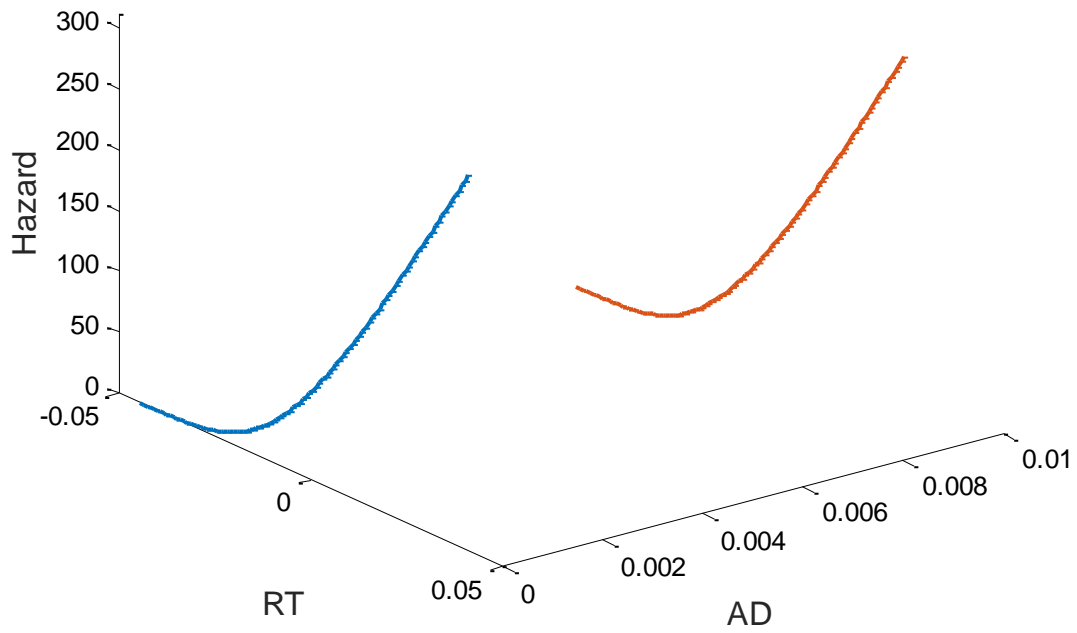


Figure 28. Hazard function of returns and AD.

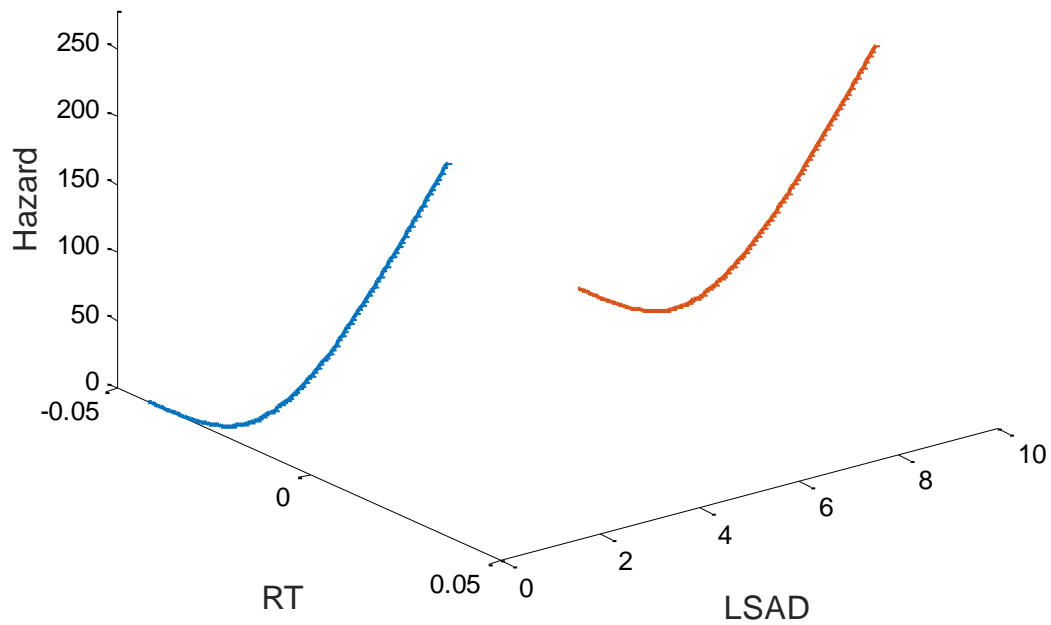


Figure 29. Hazard function of returns and LSAD.

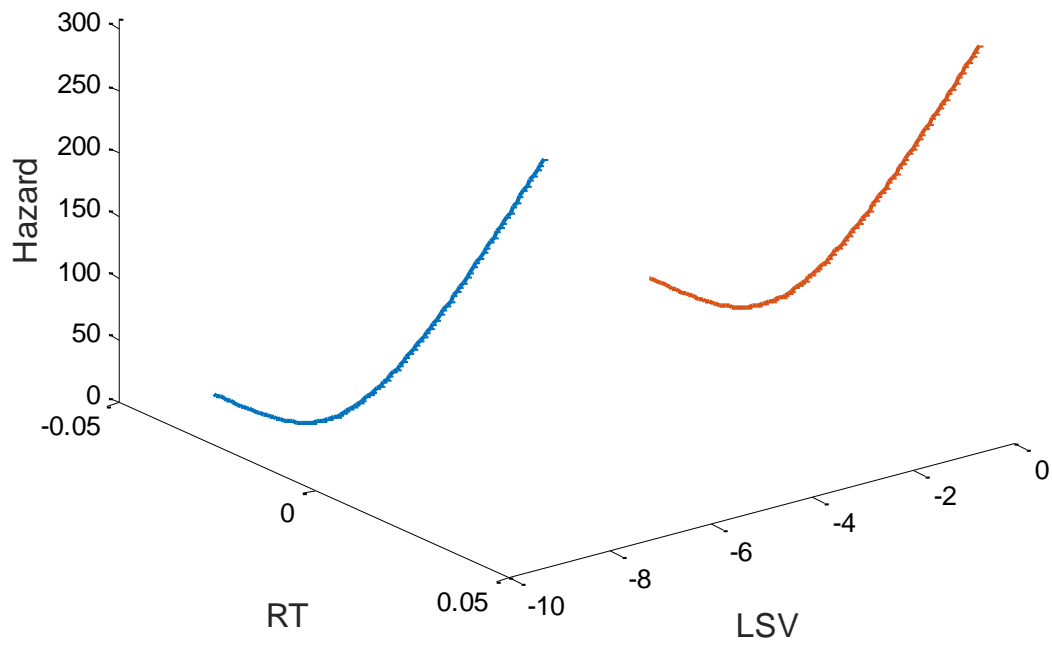


Figure 30. Hazard function of returns and LSV.

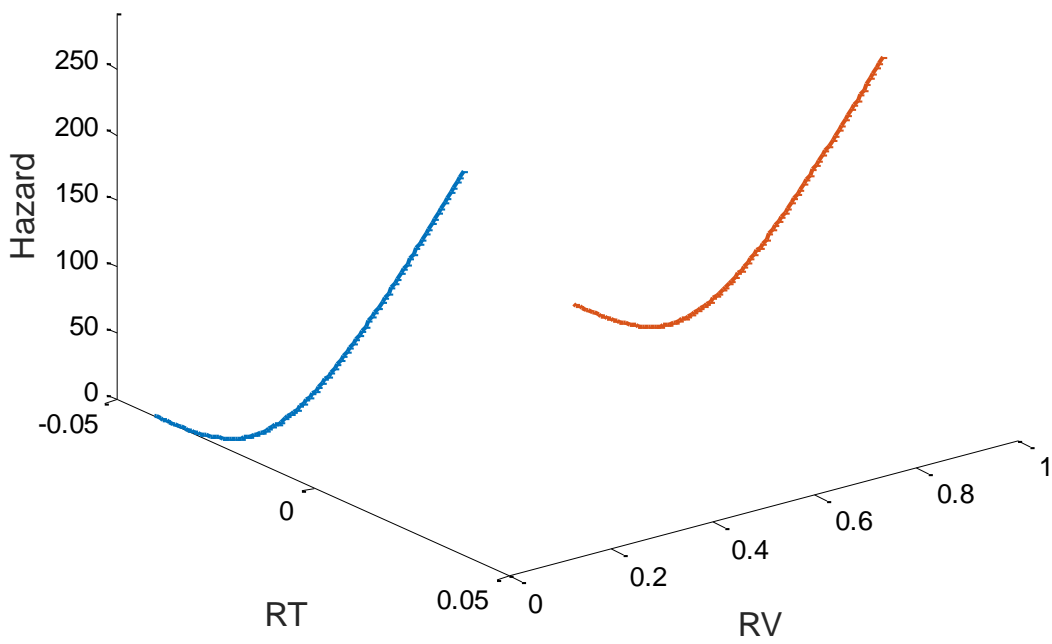


Figure 31. Hazard function of returns and RV.

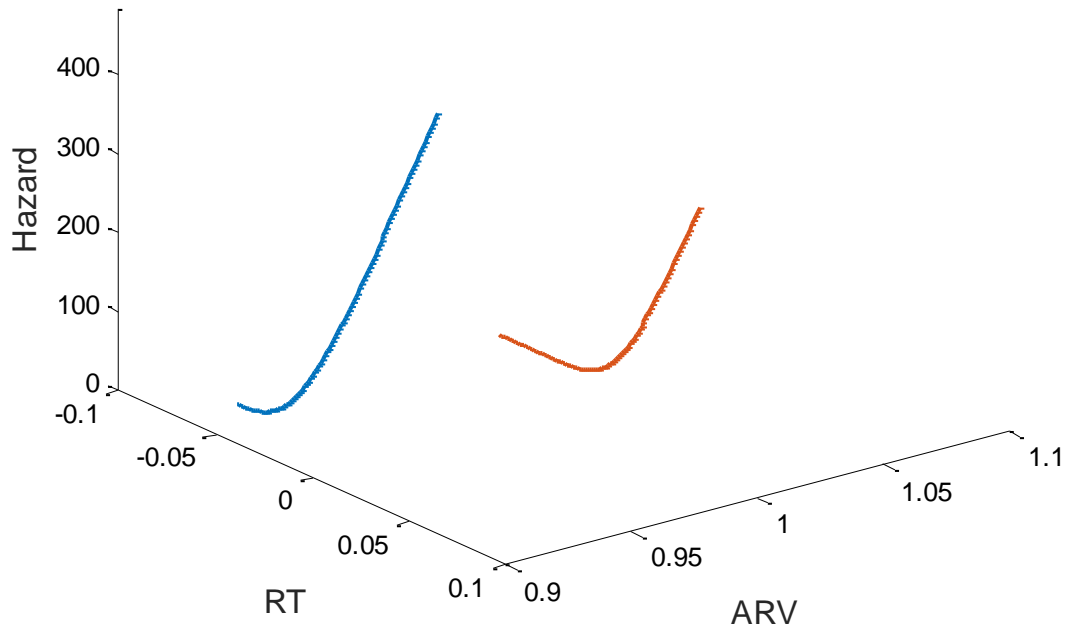


Figure 32. Hazard function of returns and ARV.

From Figures 27 to 32, the average probability estimation is 88.833 for the death of the risk-return relationship in the South African market. More specifically, from highest to lowest, the probability estimations are 161.500 (ARV), 75 (LSV), 74.500 (AD and LSAD), 74.000 (RV) and 73.500 (SD). Similar to the survival function, the death rate of the risk measures has a close range, from 73.500 to 75, whereas ARV has the quickest death rate. Figures 33 to 38 show the CML hazard function of returns with the different risk measures.

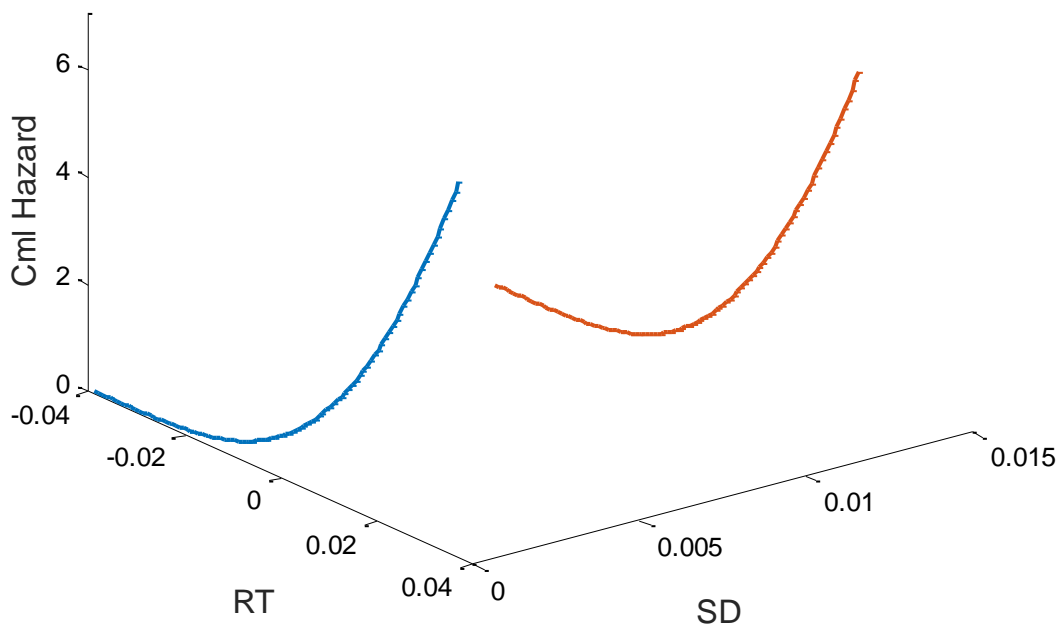


Figure 33. CML hazard function of returns and SD.

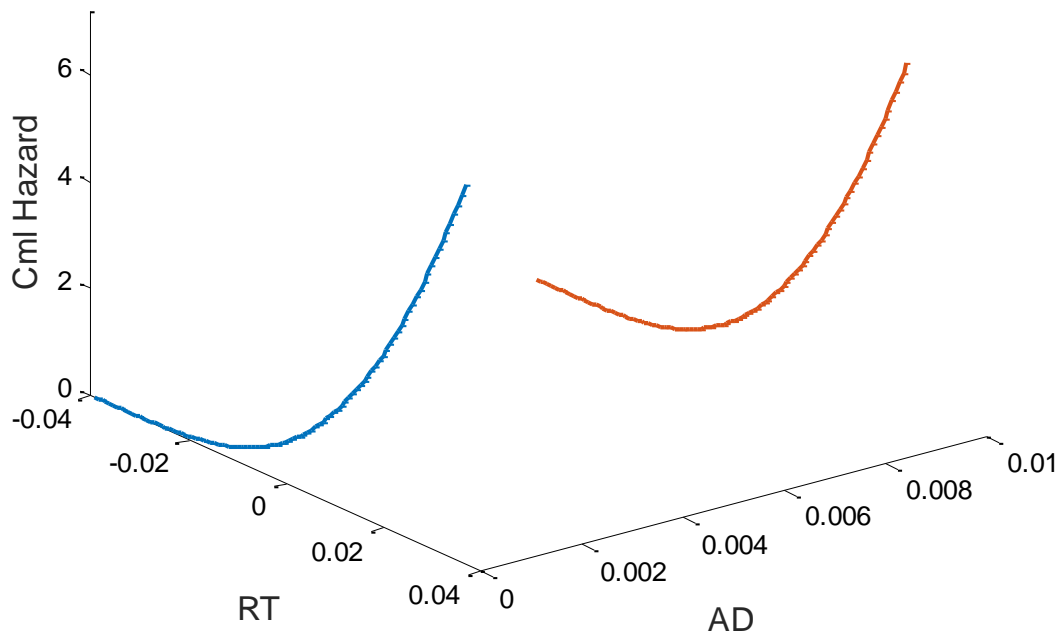


Figure 34. CML hazard function of returns and AD.

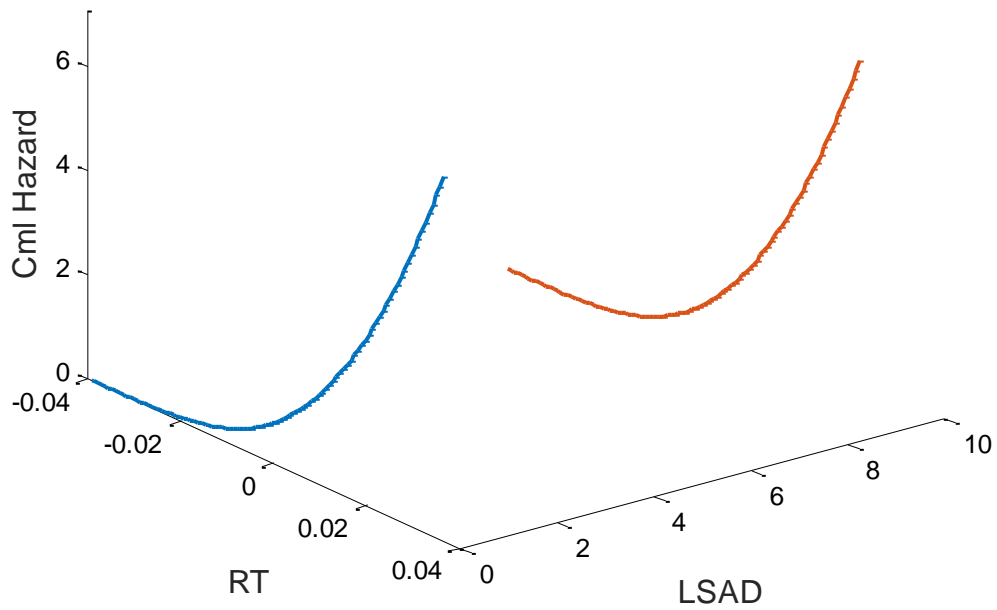


Figure 35. CML hazard function of returns and LSAD.

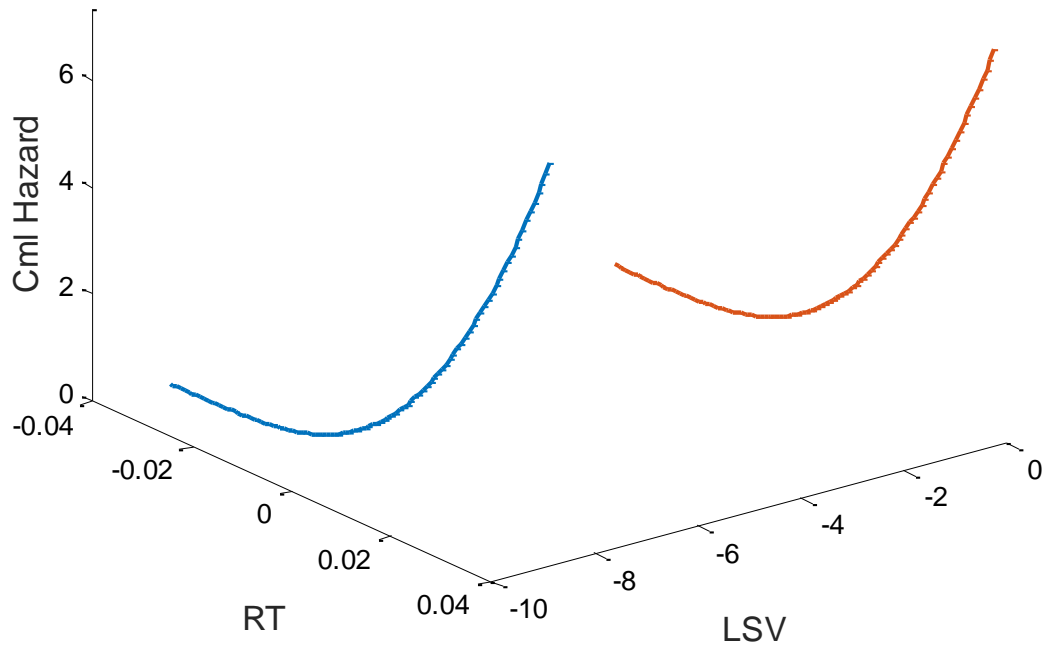


Figure 36. CML hazard function of returns and LSV.

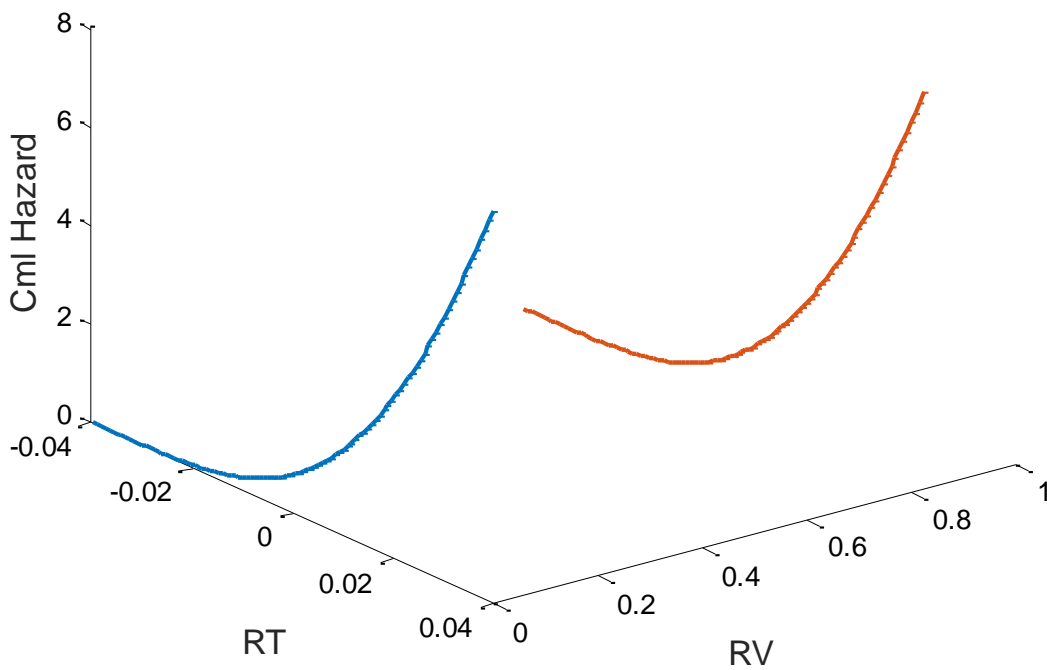


Figure 37. CML hazard function of returns and RV.

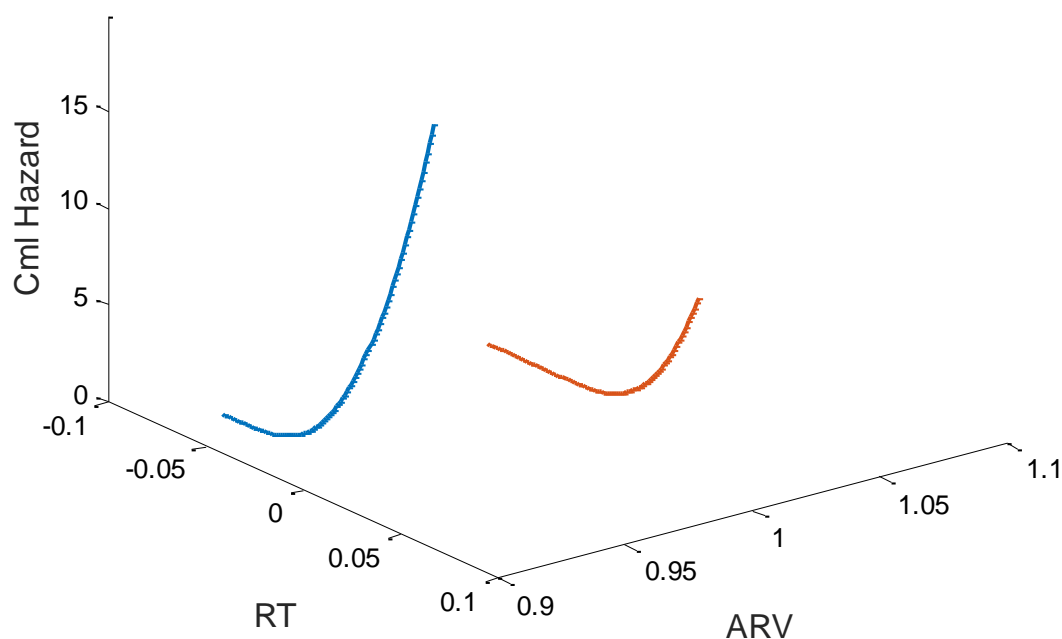


Figure 38. CML hazard function of returns and ARV.

From Figures 33 to 38, the average probability estimation is 1.128 for the risk-return relationship. From highest to lowest, the probability estimations of the accumulated death of the risk-return relationship are 2.850 (ARV), 0.800 (RV) and 73.780 (SD, AD, LSAD and LSV).

The CML hazard functions revealed that ARV has the highest probability of dissipation (2.85), followed by RV (0.8) and the remaining risk measures at 0.78.

From the survival analysis, it can be concluded that ARV earns the highest returns and has the shortest lifespan in probability terms. Investors that seek to earn a quick return would thus incline to this risk measure first and foremost, followed by LSV.

In summary of the empirical results, with respect to the downside risk, AD and LSAD, this study found a negative risk-return relationship. This was in contrast to Stankovic et al. (2020) and Sehgal and Pandey (2018), who found a positive and no relationship, respectively. In line with Stankovic et al. (2020), this study found that LSV had a positive relationship with returns.

The second-lowest and lowest accumulated risk-return relationships were found by the RV and SD risk measures, respectively. The relationship was negative, with an average life and death compared to the other risk measures. The negative relationship between returns and SD is in line with the South African study by Steyn and Theart (2019), who found that this result is in line with the low-risk anomaly in emerging markets.

Given the efficiency of the realized risk measures, as documented by Floros et al. (2020) and Moghaddam et al. (2020), this study expected RV to perform robustly along with ARV. However, only the ARV results stood out with regards to the Bayesian analysis. In contrast to RV, it had the highest accumulated probability estimation of the risk-return relationship. In the survival analysis, the ARV had the highest death rate and accumulated death rate, hence the lowest life span, attributed to the volatile nature of the South African market.

Moreover, LSV and ARV were the only two measures with a positive relationship with returns, where ARV had the higher return due to the steepest slope. The positive relationship between returns

and ARV contrasts with Floros et al. (2020), who found no relationship, and Jensen and Maheu (2018), who found a positive and nonlinear relationship. The implications of these results are discussed and concluded in the next section.

5. Discussion and conclusions

Steyn and Theart (2019) is the only known South African study to investigate a different risk measure in the context of the risk-return relationship. In international literature, there is no existing study that investigates different classes of risk in one collective study. This motivated the investigation of the relationship between returns and different risk measures stemming from international literature. The risk measures analyzed in this study were standard deviation (SD), absolute deviation (AD), lower semi absolute deviation (LSAD), lower semivariance (LSV), realized variance (RV) and the bias-adjusted realized variance (ARV). This study made an innovative contribution on a methodological and practical level, by being the first study to extend from the novel Bayesian approach by Jensen and Maheu (2018) to methods by Karabatsos (2017)—density regression, quantile regression and survival analysis.

The results of this study confirmed that different risk measures impact the significance and relationship with returns, in line with literature, with the exception of Sehgal and Pandey (2018). Sehgal and Pandey (2018) found that different risk measures had an insignificant impact on returns. This was due to the modelling approach employed by the authors, as explained by Yildiz et al. (2022) and Hoque and Low (2020). More specifically, Hoque and Low (2020) stated that previous literature dealt with semi-asymmetric properties and did not provide a full representation of returns. The limitations highlighted by Hoque and Low (2020), provided a gap in literature which this study addressed in three ways.

First, this study took into account different risk measures that were more appropriate in capturing the risk-return relationship. For example, downside risk measures are adequate when the return distribution is asymmetric and the realized variance measures provide accurate measures of latent volatility (Yildiz et al., 2022; Floros et al., 2020, Moghaddam et al., 2020, Jensen and Maheu, 2018). Second, this study employed the original parametric Bayesian approach by Jensen and Maheu (2018), which can effectively account for asymmetric higher moment properties. Third, the extensions by Karabatsos (2017) employed in this study, provided a full graphical representation of returns in relation to the different risk measures.

The most novel and innovative contribution of this study was the application of survival analysis. This is especially relevant given that over the years, there has been increasing popularity in HF trading and data analysis, as sufficed by Floros et al. (2020). Investors are always on the lookout for an arbitrage opportunity to capitalize on and “make a quick profit.” By looking at the “life” and “death” of the relationship between returns and the different risk measures, this provided novel insight for investors to take into account when structuring their investment strategies. More specifically, it can assist investors in determining when to hold their stock, go long or short and ultimately optimize their chance of earning a superior return.

Based on traditional theory, the relationship between risk and return is assumed to be positive and linear (Markowitz, 1952). Following Jensen and Maheu (2018) and Karabatsos (2017), the application of the parametric Bayesian approach is more robust than conventional parametric models such as regression analysis and the GARCH approach. This is because the model has a greater ability to capture higher moment asymmetric properties. As confirmed by normality tests, asymmetry was adequately

captured by the parametric Bayesian model. Hence, the results were considered valid and reliable. The majority of risk measures in this study found a negative relationship with returns, in contrast to theoretical expectations. However, this finding was in line with Steyn and Theart (2019), who found the existence of the low-risk anomaly in the South African market. Hence, investors that structure their investment on the “high risk, high return” strategy would find themselves disappointed as the additional risk would lead to potential losses rather than gains. Such stock would be used for other purposes, such as diversification purposes or short selling.

While this study could build on the low-risk anomaly in South African literature, following Steyn and Theart (2019), this study highlights the remaining two risk measures that found a positive risk-return relationship, namely, LSV and ARV. To recap, Floros et al. (2020) stated that realized risk measures, constructed with HF data, provide more accurate parameter estimates of latent volatility. More information is contained in HF data with respect to intraday movements and microstructure noise in the market. This was confirmed in the study by Jensen and Maheu (2018), who found that the ARV captured risk more adequately than RV. In this study, the results were most pronounced for the ARV risk measure because it is a well-suited metric in Bayesian analysis, as noted by Jensen and Maheu (2018). LSV was found to be the most robust downside risk measure. Therefore, these measures of risk should be analyzed by investors that wish to optimize their returns.

All the risk measures had a high probability of dissipation, so investors should avoid seeking long-term investment strategies in the South African market. The fluctuant nature of the risk-return relationship can be attributed to the volatile nature of the South African market. Thus, South Africa can be considered ideal for short-term high-earning returns by market participants.

Asset managers and portfolio managers are often pressured to demonstrate quick profits to attract and retain clients. Investors are always on the lookout for a “quick profit.” Given that all the risk measures had a high probability of dissipation, the aforementioned market participants can earn potential short-term superior returns in the South African market. However, care should be taken in the development of long-term investment strategies to maintain a superior or positive rate of return. In other words, strategies should be put in place to ensure that a quick return does not come at the expense of shareholders and clients over the long term. Examples include observing corporate control, setting up incentives and having organizational owners.

Supported by the literature of Jensen and Maheu (2018) and Karabatsos (2017), this study concludes as follows. Models that produce a positive and linear relationship between risk and return are generally considered robust. Models that have the ability to capture asymmetric properties effectively are considered optimal. While the parametric Bayesian approach has the ability to capture asymmetric higher moment properties effectively, it is still a parametric model which is subject to a finite sample space. The nonparametric approach overcomes this limitation as it is subject to an infinite sample space and would have a greater capacity to account for asymmetry.

Therefore, a model that has the ability to effectively account for asymmetry is more likely to estimate a nonlinear risk-return relationship, regardless of theoretical expectations. For example, Jensen and Maheu (2018) was the first study to employ the novel nonparametric Bayesian approach and found a positive nonlinear risk-return relationship. Given the magnitude of conflicting results found in literature on the empirical risk-return relationship, a model that has the ability to effectively capture asymmetry, is of greater significance than a model that produces a theoretically expected result.

This study recommends an extended nonparametric Bayesian approach as the nonparametric approach has a greater capacity to account for asymmetry. In line with Jensen and Maheu (2018), this

study found that the ARV risk measure is optimal in the field of Bayesian modelling. Thus, the second recommendation is to research the existence of other traditional robust risk measures in the Bayesian context, or otherwise, and to extend this research to such risk measures.

Data

The secondary price dataset used in this study can be obtained from the IRESS database.

Conflict of interest

The author declares no conflict of interest.

References

- Alles L, Murray L (2017) Asset pricing and downside risk in the Australian share market. *Appl Econ* 49: 4336–4350. <http://dx.doi.org/10.1080/00036846.2017.1282143>
- Basher SA, Sadorsky P (2006) Oil price risk and emerging stock markets. *Glob Financ J* 17: 224–251. <http://dx.doi.org/10.1016/j.gfj.2006.04.001>
- Bayes T (1763) An Essay towards Solving a Problem in the Doctrine of Chances. by the Late Rev. Mr. Bayes, F. R. S. Communicated by Mr. Price, in a Letter to John Canton, A. M. F. R. S. *Phil Trans R Soc* 53: 370–418. <https://doi.org/10.1098/rstl.1763.0053>
- Brooks C (2014) *Introductory Econometrics for Finance*. Cambridge University Press: New York.
- Devore JL (2012) *Probability and Statistics for Engineering and the Sciences*. California Polytechnic State University: San Luis Obispo.
- Emmert-Streib F, Dehmer M (2019) Introduction to Survival Analysis in Practice. *Mach Learn Know Extr* 1: 1013–1038. <https://doi.org/10.3390/make1030058>
- Fassas AP, Siriopoulos C (2020) Implied volatility indices—A review. *Q Rev Econ Financ* 79: 303–329. <https://doi.org/10.1016/j.qref.2020.07.004>
- Floros C, Gkillas K, Konstantatos C, Tsagkanos A (2020) Realized Measures to Explain Volatility Changes over Time. *J Risk Financ Manag* 13: 1–19. <https://doi.org/10.3390/jrfm13060125>
- Gao G, Bu Z, Liu L, et al. (2019) A Survival Analysis Method for Stock Market Prediction. *2015 International Conference on Behavioral, Economic and Socio-cultural Computing (BESC)*, 116–122. <https://doi.org/10.1109/BESC.2015.7365968>
- Gepp A, Kumar, K (2015) Predicting Financial Distress: A Comparison of Survival Analysis and Decision Tree Technique. *Procedia Comput Sci* 54: 396–404. <https://doi.org/10.1016/j.procs.2015.06.046>
- Griffin JE, Kalli M, Steel M (2018) Discussion of “Nonparametric Bayesian Inference in Applications”: Bayesian nonparametric methods in econometrics. *Stat Method Appl* 27: 207–218. <https://doi.org/10.1007/s10260-017-0384-0>
- Hansen PR, Lunde A (2006) Realized variance and market microstructure noise. *J Bus Econ Stat* 24: 127–161. <https://doi.org/10.1198/073500106000000071>
- Hoque ME, Low S (2020) Industry Risk Factors and Stock Returns of Malaysian Oil and Gas Industry: A New Look with Mean Semi-Variance Asset Pricing Framework. *Mathematics* 8: 1–28. <https://doi.org/10.3390/math8101732>

- Jensen MJ, Maheu JM (2018) Risk, Return and Volatility Feedback: A Bayesian Nonparametric Analysis. *J Risk Financ Manag* 11: 1–29. <https://doi.org/10.3390/jrfm11030052>
- Kalli M, Griffin J, Walker S (2011) Slice sampling mixture models. *Stat Comput* 21: 93–105. <http://dx.doi.org/10.1007/s11222-009-9150-y>
- Karabatsos G (2017) A menu-driven software package of Bayesian nonparametric (and parametric) mixed models for regression analysis and density estimation. *Behav Res Methods* 49: 335–362. <https://doi.org/10.3758/s13428-016-0711-7>
- Markowitz H (1952) Portfolio selection. *T J Financ* 7: 77–91. <https://doi.org/10.2307/2975974>
- Moghaddam MD, Liu J, Serota RA (2020) Implied and realized volatility: A study of distributions and the distribution of difference. *Int J Financ Econ* 26: 2581–2594. <https://doi.org/10.1002/ijfe.1922>
- National Treasury (2021) Economic Overview. In: Chapter 2. Budget Review. Republic of South Africa. Available from: <http://www.treasury.gov.za/documents/National%20Budget/2021/review/Chapter%202.pdf>.
- Papaspiliopoulos O (2008) A note on posterior sampling from Dirichlet mixture models. Available from: <http://www2.warwick.ac.uk/fac/sci/statistics/crism/research/2008/08-20wv2.pdf>.
- Sehgal S, Pandey A (2018) Predicting Financial Crisis by Examining Risk-Return Relationship. *Theor Econ Lett* 8: 48–71. <https://doi.org/10.4236/tel.2018.81003>
- Sharma S (2017) Markov Chain Monte Carlo Methods for Bayesian Data Analysis in Astronomy. *Annu Rev Astron Astrophys* 55: 213–259. <https://doi.org/10.48550/arXiv.1706.01629>
- Stankovic JZ, Petrovic E, Dencic-Mihajlov K (2020) Effects of Applying Different Risk Measures on the Optimal Portfolio Selection: The Case of the Belgrade Stock Exchange. *Facta Universitatis Series: Economics and Organization* 17: 17–26. <https://doi.org/10.22190/FUEO191016002S>
- Steyn JP, Theart L (2019) Are South African equity investors rewarded for taking on more risk? *J Econ Financ Sci* 12: 1–10. <https://doi.org/10.4102/jef.v12i1.448>
- Trichilli Y, Abbes MB, Masmoudi A (2020) Islamic and conventional portfolios optimization under investor sentiment states: Bayesian vs Markowitz portfolio analysis. *Res Int Bus Financ* 51: 1–21. <https://doi.org/10.1016/j.ribaf.2019.101071>
- van Ravenzwaaij D, Cassey P, Brown SD (2018) A simple introduction to Markov Chain Monte-Carlo sampling. *Psychon Bull Rev* 25: 143–154. <https://doi.org/10.3758/s13423-016-1015-8>
- Walker SG (2007) Sampling the dirichlet mixture model with slices. *Commun Stat-Simul C* 36: 45–54. <http://dx.doi.org/10.1080/03610910601096262>
- Yildiz ME, Erzurumlu YO, Kurtulus B (2022) Comparative analyses of mean-variance and mean-semivariance approaches on global and local single factor market model for developed and emerging markets. *Int J Emerg Mark* 17: 325–350. <https://doi.org/10.1108/IJOEM-01-2020-0110>



AIMS Press

© 2022 the Author, licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)