

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

# **Shewhart methodology for modelling financial series**

A thesis presented in partial fulfilment of the requirements for the degree

of

**Doctor of Philosophy**

in

**Statistics**

at Massey University, Palmerston North, New Zealand



**Liyana Pathiranaralalage Nadeeka Dilrukshi Premarathna**

2017



## **Abstract**

Quality management techniques are widely used in industrial applications for monitoring observable process variation. Among them, the scientific notion of Shewhart principles is vital for understating variations in any type of process or service. This study extensively investigates and demonstrates Shewhart methodology for financial data.

Extremely heavy tails noted in the empirical distribution of stock returns led to the development of new parametric probability distributions for pricing assets and forecasting market risk. Standard asset pricing models have also extended to account the first four (excess) moments in return distributions. These approaches remain complex, but yet they are inadequate for capturing extreme volatility caused by infrequent market events.

It is well known that the security markets are always subjected to a certain amount of variability caused by noise-traders and other frictional price changes. Unforeseen events which are happening in the world may lead to huge market losses. This research shows that Shewhart methodology for partitioning data into common and special cause variations adds value to modelling stock returns.

Applicability of the proposed method is discussed using several scenarios occurring in an industrial process and a financial market. A set of new propositions based on Shewhart methodology is formed for finer description of the statistical properties in stock returns. Research issues which are related to the first four moments, co-moments and autocorrelation in stock returns are identified. New statistical tools such as difference control

---

charts, odd-even analysis and estimates for co-moments are proposed to investigate the new propositions and research issues. Finally, several risk measures are proposed, and considered with respect to investor's preferences.

The research issues are investigated using partitioned data from S&P 500 stocks and the findings show that in most of the scenarios, contradictory conclusions were made as a result of special cause variations. A modelling approach based on common and special cause variations is therefore expected to lead appropriate asset pricing and portfolio management. New statistical tools proposed in this study can be used to other time series data; a new R-package called QCCTS (Quality Control Charts for Time Series) is developed for this purpose.

## Acknowledgements

This study would have been impossible without the support of many. First and foremost, I would like to express my sincere gratitude to my primary supervisor, Dr. Jonathan Godfrey for his guidance, support and very good understanding throughout this study. You are truly inspirational. Thank you very much for all your guidance, encouragements, smiles and jokes. Your cheerful confidence pushes me outside of my comfort zone ☺. I am also very grateful to my co-supervisor, Dr. K. Govindaraju for his enthusiasm, prompt support and encouragements. Both of you have guided me to be a good researcher. Thank you very much, Raj ☺.

Thank you Dr. Ganes, Prof. Martin Hazelton and Dr. Sarojeni for your guidance and support to begin this journey. I would like to give my special thank to Martin for giving the opportunity to attend conferences with my little boy. To Liz Godfrey, thank you very much for reading my papers and thesis, encouraging me to achieve the finishing line of this study. I also want to thank Peter Lewis, Calitz Hannes for computer technical support; Ann, Fiona, Debbie, Colleen for helping administrative tasks; staff-at the library, the student teaching and learning centre, and the international office at Massey. The statistics research group and friends at post-graduate office, thank you very much for being friendly. We had a nice time together and it is a great pleasure to meet you.

I received PhD grants from the Higher Education for the Twenty-First Century project, Sri Lanka and the National Centre for Advanced Studies in Humanities and Social Sciences,

---

Sri Lanka for the first three years of PhD. The final year of this study is funded by the WC grant, University of Kelaniya and the Institute of Fundamental Sciences, Massey University, New Zealand. The University of Kelaniya, Sri Lanka granted me paid study leave during this time period. I'm very grateful for the financial help from my country and New Zealand. Thank you!

Gesandu and I were truly blessed to have a very supportive Sri Lankan community here in Palmy. Thank you very much for your support and love. We also met lovely teachers and parents at children centre, school and out. Thank you very much for your hospitality and kindness. The moments that we shared together are memorable.

My sincere thanks will also extend to my beloved mother (amma) and my father-in-law. My family and my extended family for their support throughout this journey. My dear father (thaththa), I still miss you but always believe that you keep watching me doing strong. All my teachers, colleagues and friends back home, who supported and encouraged, Thank you very much!

Finally and most importantly, thank you Manoj for your trust, patience, hope and the huge sacrifices that you have been made over the last four years. Gesandu, you are truly amazing at seeing all the up and downs in my life at such a little age. Without your unconditional love, your mum (ammi) would have never achieved the finishing line. We collected valuable and nice memories together that you will cherish one day. I love you both with all my heart. ♡♡

I'm grinning today from ear to ear. 😊😊😊

## **Publications arising from this thesis**

Premarathna, N., Godfrey, A. J. R., and Govindaraju, K. (2016b). Decomposition of stock market trade-offs using Shewhart methodology. *International Journal of Quality & Reliability Management*, 33(9):1311–1331

Premarathna, N., Godfrey, A. J. R., and Govindaraju, K. (2017). Control charts for paired differences:  $\bar{d}$  and  $S_d$  charts. *Quality and Reliability Engineering International*. Early view at <http://dx.doi.org/10.1002/qre.2147>

Premarathna, N., Godfrey, A. J. R., and Govindaraju, K. (2016a). Assessment of induced autocorrelation in stock returns. *Empirical Economics*. Under review





# Contents

|   |             |
|---|-------------|
| <b>List of Figures</b>  | <b>xiii</b> |
| <b>List of Tables</b>   | <b>xvii</b> |
| <b>1 Introduction</b>   | <b>1</b>    |
| 1.1 Motivation . . . . .                                      | 2           |
| 1.2 Shewhart postulates applied to financial data . . . . .   | 7           |
| 1.3 Objectives & thesis structure . . . . .                   | 11          |
| 1.4 Data & software . . . . .                                 | 13          |
| <b>2 Market behaviour and Shewhart's principles</b>           | <b>15</b>   |
| 2.1 Introduction . . . . .                                    | 15          |
| 2.2 Rational and irrational market behaviour . . . . .        | 16          |
| 2.3 Market behaviour and Deming's funnel experiment . . . . . | 18          |
| 2.4 Industrial process vs. stock market . . . . .             | 22          |
| 2.5 Summary . . . . .   | 24          |
| <b>3 Control chart techniques for financial applications</b>  | <b>25</b>   |
| 3.1 Introduction . . . . .                                    | 26          |
| 3.2 Shewhart control charts . . . . .                         | 29          |
| 3.2.1 An example . . . . .                                    | 31          |

## Contents

---

|          |  |           |
|----------|--|-----------|
| 3.3      | Stepwise robust chart procedures . . . . .                                 | 36        |
| 3.3.1    | Robust estimator for $\sigma$ . . . . .                                    | 36        |
| 3.3.2    | Estimation of $\sigma$ for Phase II . . . . .                              | 37        |
| 3.3.3    | Estimation of $\mu$ for Phase II . . . . .                                 | 38        |
| 3.4      | Shewhart control charts for individual observations . . . . .              | 39        |
| 3.5      | R package: QCCTS (Quality Control Charts for Time Series) . . . . .        | 40        |
| 3.6      | Summary . . . . .  | 41        |
| <b>4</b> | <b>Market trade-offs</b> . . . . .   | <b>43</b> |
| 4.1      | Introduction . . . . .   | 44        |
| 4.2      | Descriptive analysis . . . . .   | 46        |
| 4.3      | Trade-offs between the first four moments . . . . .                        | 50        |
| 4.4      | The mean/standard deviation trade-off . . . . .                            | 55        |
| 4.5      | The skewness/kurtosis trade-off . . . . .                                  | 57        |
| 4.6      | Analysis of trading volumes . . . . .                                      | 59        |
| 4.7      | Summary . . . . .  | 60        |
| <b>5</b> | <b>Difference control charts</b> . . . . .                                 | <b>63</b> |
| 5.1      | Introduction . . . . .   | 64        |
| 5.2      | Design of $\bar{d}$ and $S_d$ charts . . . . .                             | 65        |
| 5.3      | Run length properties of the difference charts . . . . .                   | 67        |
| 5.3.1    | Discussion of out-of-control situations when standards are known . . . . . | 69        |
| 5.3.2    | Discussion of RLs properties when standards are unknown . . . . .          | 75        |
| 5.3.3    | Discussion of RLs properties using inverse-Wishart prior . . . . .         | 80        |
| 5.3.4    | Other remarks . . . . .  | 83        |
| 5.4      | A finance application . . . . .  | 85        |
| 5.5      | Summary . . . . .  | 91        |

|          |  |            |
|----------|--|------------|
| <b>6</b> | <b>Bivariate co-moments in stock returns</b>                               | <b>93</b>  |
| 6.1      | Introduction . . . . .   | 94         |
| 6.2      | Theory of co-skewness and co-kurtosis . . . . .                            | 97         |
| 6.2.1    | Sample co-moments . . . . .  | 99         |
| 6.3      | Empirical analysis . . . . .   | 101        |
| 6.4      | Sample estimates from pairwise differences . . . . .                       | 107        |
| 6.5      | Summary . . . . .  | 113        |
| <b>7</b> | <b>Odd-even split</b>  | <b>115</b> |
| 7.1      | Introduction . . . . .   | 117        |
| 7.2      | Partitioning via the Odd-even split . . . . .                              | 118        |
| 7.2.1    | Odd-even split of the AR( $p$ ) model . . . . .                            | 120        |
| 7.2.2    | Odd-even split of the MA( $q$ ) model . . . . .                            | 125        |
| 7.2.3    | Odd-even split of the ARMA( $p, q$ ) model . . . . .                       | 127        |
| 7.3      | Recombined ACFs of the complete series using odd-even split . . . . .      | 130        |
| 7.4      | Summary . . . . .  | 133        |
| <b>8</b> | <b>Autocorrelations in stock returns</b>                                   | <b>135</b> |
| 8.1      | Introduction . . . . .   | 135        |
| 8.2      | Illustration of induced ACFs/PACFs . . . . .                               | 137        |
| 8.2.1    | Linear time series models and induced ACFs . . . . .                       | 137        |
| 8.2.2    | Non-linear time series models and induced ACFs . . . . .                   | 141        |
| 8.3      | The odd-even split and identification of induced autocorrelation . . . . . | 142        |
| 8.4      | Empirical analysis . . . . .   | 146        |
| 8.4.1    | Effect of outliers . . . . .   | 152        |
| 8.4.2    | Other effects . . . . .  | 157        |
| 8.4.3    | Autocorrelation test for market efficiency . . . . .                       | 158        |
| 8.5      | Summary . . . . .  | 162        |

## Contents

---

|  |            |
|--|------------|
| <b>9 Applications</b>  | <b>163</b> |
| 9.1 Applications for investors . . . . .   | 163        |
| 9.2 Comparing performances of stocks using the new measures . . . . .                  | 165        |
| 9.3 What do short term investors look for? . . . . .                                   | 166        |
| 9.4 What do long term investors look for? . . . . .                                    | 168        |
| 9.5 Summary . . . . .  | 173        |
| <b>10 Concluding remarks</b>   | <b>175</b> |
| 10.1 Summary of objectives and achievements . . . . .                                  | 175        |
| 10.2 Future research directions . . . . .  | 180        |
| 10.2.1 Further applications in finance . . . . .                                       | 182        |
| 10.2.2 Applications in other disciplines . . . . .                                     | 184        |
| <b>References</b>  | <b>187</b> |
| <b>Glossary</b>  | <b>210</b> |
| <b>Appendix A Statement of contribution to doctoral thesis containing publications</b> | <b>211</b> |

# List of Figures

|     |  |    |
|-----|--|----|
| 3.1 | S chart for Apple Inc. based on weekly subgroups of daily log returns-Phase I analysis . . . . .   | 32 |
| 4.1 | The first four sample moments for stock returns in S&P 500 list . . . . .  | 47 |
| 4.2 | Average returns in common and special cause periods . . . . .  | 50 |
| 4.3 | Mean/standard deviation trade-off: known distributions . . . . .   | 53 |
| 4.4 | Skewness/kurtosis trade-off: known distributions . . . . .   | 54 |
| 4.5 | Mean/standard deviation trade-off, stocks in the S&P 500 list . . . . .  | 56 |
| 4.6 | Skewness/kurtosis trade-off, stocks in the S&P 500 list . . . . .  | 58 |
| 5.1 | (a) ARLs for shifts in $\mu_Y$ ( $\mu_X = 0, \sigma_X^2 = 1, \sigma_Y^2 = 1$ and $\rho = 0.75$ ) . . . . .   | 70 |
| 5.2 | ARLs for (a) equal shifts in $\mu_X$ and $\mu_Y$ in the opposite direction and (b) unequal shifts in $\mu_X$ and $\mu_Y$ in the same direction ( $\sigma_X^2 = 1, \sigma_Y^2 = 1$ and $\rho = 0.75$ ) . . . . .  | 71 |
| 5.3 | ARLs for shifts in $\sigma_Y^2$ ( $\mu_X = 0, \mu_Y = 0, \sigma_X^2 = 1$ and $\rho = 0.75$ ) . . . . .   | 72 |
| 5.4 | ARLs for (a) shifts in $\sigma_X^2$ and $\sigma_Y^2$ in the same direction and (b) unequal shifts in $\sigma_X^2$ and $\sigma_Y^2$ in opposite directions ( $\mu_X = 0, \mu_Y = 0$ and $\rho = 0.75$ ) . . . . . | 73 |
| 5.5 | ARLs for shifts in $\rho$ ( $\mu_X = 0, \mu_Y = 0, \sigma_X^2 = 1$ and $\sigma_Y^2 = 1$ ) . . . . .  | 74 |
| 5.6 | Distributions of in-control ARLs for $\bar{d}$ chart and Hotelling's $T^2$ chart when the parameters were estimated from $m$ subgroups of size $n = 10$ . . . . .  | 76 |

## List of Figures

---

|      |   |     |
|------|---|-----|
| 5.7  | Distributions of in-control ARLs for $S_d$ chart and $ S $ chart when the parameters were estimated from $m$ subgroups of size $n = 10$ . . . . .   | 77  |
| 5.8  | ARLs for (a) shifts in $\mu_Y$ (b) shifts in $\sigma_Y^2$ when the parameters are estimated from Phase I data . . . . .   | 79  |
| 5.9  | ARLs for shifts in $\rho$ when the parameters are estimated from Phase I data . .   | 79  |
| 5.10 | (a) ARLs for shifts in $\mu_Y$ under a Wishart prior for $\nu = 100$ and (b) ARLs for shifts in $\mu_Y$ under a Wishart prior for $\nu = 500$ (c) ARLs for shifts in $\mu_Y$ under a Wishart prior for $\nu = 1000$ . ( $\mu_X = 0, \sigma_X^2 = 1, \sigma_Y^2 = 1$ and $\rho = 0.75$ ) . . . . .           | 82  |
| 5.11 | (a) ARLs for shifts in $\sigma_Y^2$ under a Wishart prior for $\nu = 100$ and (b) ARLs for shifts in $\sigma_Y^2$ under a Wishart prior for $\nu = 500$ (c) ARLs for shifts in $\sigma_Y^2$ under a Wishart prior for $\nu = 1000$ . ( $\mu_X = 0, \mu_Y = 0, \sigma_X^2 = 1$ and $\rho = 0.75$ ) . . . . . | 82  |
| 5.12 | ARLs for shifts in $\mu_Y$ for $\rho = 0.15, \rho = 0.55$ and $\rho = 0.85$ . . . . .   | 84  |
| 5.13 | ARLs for fixed shifts in $\sigma_Y^2$ for $\rho = 0.15$ and $\rho = 0.55$ . . . . .   | 84  |
| 5.14 | (a) daily price series and (b) log returns . . . . .  | 88  |
| 5.15 | Phase II (a) $\bar{d}$ chart and (b) $S_d$ chart . . . . .  | 89  |
| 5.16 | (a) Hotelling's $T^2$ chart and (b) $ S $ chart . . . . .   | 90  |
| 6.1  | $cs(x, y^2)$ vs $cs(x^2, y)$ for standard bivariate normal data when $\mu_X = 0, \mu_Y = 0, \sigma_X^2 = 1, \sigma_Y^2 = 1, \rho = 0.75$ and $Y$ series with 10% contamination . . . . .  | 100 |
| 6.2  | $ck(x, y^3)$ vs $ck(x^3, y)$ for standard bivariate normal data when $\mu_X = 0, \mu_Y = 0, \sigma_X^2 = 1, \sigma_Y^2 = 1, \rho = 0.75$ and $Y$ series with 10% contamination . . . . .  | 101 |
| 6.3  | Scatter plot of sample correlation coefficient: total vs. common causes . . .   | 105 |
| 6.4  | Co-moments in total and common cause data . . . . .   | 106 |
| 6.5  | Distribution of sample estimates for co-variance, co-skewness and co-kurtosis based on the non-matching pairs of differences for sample size 10 . . . . .   | 109 |
| 6.6  | $s(x, y)_{nmp}, cs(x, y^2)_{nmp}$ and $ck(x, y^3)_{nmp}$ for the total series and a sub-sample size 20 . . . . .  | 110 |

|   |     |
|---|-----|
| 6.7 Newly proposed co-moments measures in total and common cause data using stock returns of Michael Kors Holdings (KORS) . . . . .   | 112 |
| 7.1 Sample ACFs for AR(1) model: $x_t = 0.8x_{t-1} + \epsilon_t$ and its even series $x_{t,e}$ :<br>$x_t = 0.64x_{t-2} + \epsilon_t^*$ . . . . .                                    | 122 |
| 7.2 Sample ACFs for MA(3) model: $x_t = \epsilon_t + 0.4\epsilon_{t-1} + 0.3\epsilon_{t-2} + \epsilon_{t-3}$ and its even series . . . . .  | 127 |
| 7.3 Sample ACFs at even lags of complete series from odd and even series: complete series is equivalent to MA(2): $\theta_1 = 0.3, \theta_2 = 0.6$ , ACF at lag 2 $\approx 4.1$ . . | 132 |
| 8.1 ACFs/PACFs of log returns for a WN price series . . . . .   | 138 |
| 8.2 ACFs/PACFs of MA(1) price series and ACFs/PACFs of log returns, where $-1 < \theta_1 < 0$ . . . . .   | 139 |
| 8.3 ACFs/PACFs of MA(1) price series and ACFs/PACFs of log returns, where $0 < \theta_1 < 1$ . . . . .  | 140 |
| 8.4 ACFs/PACFs of AR(1) price series and ACFs/PACFs of log returns, where AR(1): $-1 < \phi_1 < 0$ . . . . .  | 144 |
| 8.5 ACFs/PACFs of AR(1) price series and ACFs/PACFs of log returns, where AR(1): $0 < \phi_1 < 1$ . . . . .   | 145 |
| 8.6 Comparison of lags 1-4 ACFs of odd vs even series of the original return series   | 153 |
| 8.7 Comparison of ACF values at lag 1 for S&P 500 stocks after removing the effect of outliers . . . . .  | 156 |
| 8.8 JPMorgan Chase & Co. (a) daily price series and (b) stock returns . . . . .   | 159 |
| 8.9 JPMorgan Chase & Co.: sample ACFs of stock returns of (a) complete series (b) odd series and (c) even series . . . . .  | 160 |
| 8.10 (a) sample ACFs of complete series of returns after eliminating outliers (b) odd series (c) even series . . . . .  | 161 |



## List of Figures

---

- 9.1 Measured volatility ( $\sigma$ ) in the common cause periods vs. average stock price 167
- 9.2 Standard deviation (measured volatility) in total, common and special cause periods: Apple stock . . . . . 169
- 9.3 Mean average return vs. measured volatility ( $\sigma$ ) in the special cause periods 170

# List of Tables

|     |  |    |
|-----|--|----|
| 3.1 | Standard deviation $\sigma$ in common cause periods based on the different subgroup sizes: W-week, M-month and Q-quarter . . . . .   | 33 |
| 3.2 | The first four sample moments for the return series of selected stocks . . . .   | 34 |
| 3.3 | Distribution of returns after separation into common and special causes . .  | 35 |
| 4.1 | Correlations between mean/standard deviation and skewness/kurtosis for known theoretical distributions . . . . .   | 51 |
| 4.2 | Correlations between the first four moments for other scenarios in total and partitioned data . . . . .  | 59 |
| 5.1 | Control statistics and control limits used for performance comparison when standards are known and estimated . . . . .   | 68 |
| 5.2 | Performances of $\bar{d}$ chart based on the percentages in-control ARLs compared to Hotelling's $T^2$ chart. The control limits were obtained from estimated parameters from limited Phase I data of subgroups $m$ of size $n = 10$ . | 78 |
| 5.3 | Performances of $S_d$ chart based on the percentages in-control ARLs compared to $ S $ chart. The control limits were obtained from estimated parameters from limited Phase I data of subgroups $m$ of size $n = 10$ . . . . .         | 78 |

## List of Tables

---

|     |   |     |
|-----|---|-----|
| 6.1 | The first four sample moments in the total and common cause data of the return series for S&P 500 index and selected nine stocks: January, 2013 to December, 2015 . . . . .                                       | 103 |
| 6.2 | Sample co-moments in total series . . . . .   | 104 |
| 6.3 | Sample co-moments in common cause periods . . . . .   | 104 |
| 6.4 | False alarm rate comparison of existing co-measures and non-matching pairs based measure for co-moments . . . . .   | 109 |
| 7.1 | Model specifications, notations and definitions for autocorrelation analysis  | 119 |
| 8.1 | Autocorrelation properties of the complete series and log returns. Liner time series models were assumed for the price series (complete series) . . . . .   | 141 |
| 8.2 | Proportions of stocks with significant ACF/PACF in complete, odd and even series in S&P 500 list. The columns under LB test give the proportions of stocks significant ACF/PACF from the Ljung-Box test . . . . . | 148 |
| 8.3 | ACF values of complete, odd and even series for publicly traded companies having the greatest market capitalization, bold font is used to indicate the significance at 5% . . . . .                               | 150 |
| 8.4 | PACF values of complete, odd and even series for publicly traded companies having the greatest market capitalization, bold font is used to indicate the significance at 5% . . . . .                              | 151 |
| 8.5 | Proportions of significant/insignificant ACF/PACF; odd vs. even for lags 1-4  | 152 |
| 8.6 | Proportions of stocks with significant ACF/PACF in $r_{t_{m_3}}$ series and its odd and even series . . . . .   | 155 |
| 9.1 | Standard deviation ( $\sigma$ ) in common cause, special cause, and total data and stock rankings . . . . .   | 172 |