

Aggregational Gaussianity In The South African Equity Market

Daniel Polakow, University of Cape Town, South Africa
David Taylor, University of Cape Town, South Africa
Obeid Mahomed, University of Cape Town, South Africa

ABSTRACT

While stylized facts of South African asset returns have been studied extensively, Aggregational Gaussianity has largely been overlooked. The aggregational aspect arises from the n -day log-return being the sum of n one-day log-returns and empirical asset returns tending to normality as the term increases. This fact is commonly corroborated graphically using overlapping return series depicted in $Q-Q$ plots. Using a resampling-based statistical methodology to test for Aggregational Gaussianity while catering for overlapping data, an alternate picture emerges. Here the authors describe evidence from the South African market for a discernible absence of Aggregational Gaussianity and briefly discuss the implications thereof.

Keywords: Stylized Facts; Aggregational Gaussianity; Normality; Risk Management

1. INTRODUCTION

Financial time series are well-known to be characterized by certain stylized facts. A stylized fact in finance refers to an empirical finding that is sufficiently ubiquitous across all instruments, markets and time periods as to be accepted as truth. Those stylized facts common to a wide set of financial assets include: a) an absence of autocorrelations, b) heavy tails, c) gain/loss asymmetry, d) intermittency, e) volatility clustering, f) conditional heavy tails, g) a slow decay of autocorrelation in absolute return values, h) a leverage effect, i) volume/volatility correlation, and j) asymmetry in time scales (see Cont, 2001; Sewell, 2011; Li et al., 2010 for an African study). Another well-known stylized fact of financial time series is the property of Aggregational Gaussianity. Aggregational Gaussianity (hereafter AG) is the phenomenon in which the empirical distribution of log-returns tends to normality (or Gaussianity) as the frequency of observations decreases (or the time scale Δt over which the returns are calculated increases); see Eberlein and Keller (1995) and Rydberg (2000). In a financial time-series context, log-returns calculated over shorter time periods are known to be leptokurtotic (heavy-tailed) and often skewed. As the time-interval over which the returns are calculated is increased, the distribution of returns better approximate normality. Hence, the shape of the distribution is different at different time scales, or terms.

High-frequency data display power-law decay – for reasons involving self-similarity and scaling laws (Schmitt, Schertzer, & Lovejoy, 1999). Extreme-value theory is particularly useful in terms of modelling the tails in this context (Embrechts, Kluppelberg, & Mikosch, 1997). Intermediate returns (for example, daily returns) often display log-linear properties best modelled by hyperbolic or normal Inverse Gaussian distributions (Kulikova & Taylor, 2010). Log-returns over longer periods of time are, in principle, the sum of returns over shorter periods – hence the central limit theorem should give a tendency towards normality. An explanation for leptokurtosis in returns was advanced by Clark (1973) and Blattberg and Gonedes (1974). These studies proposed the idea that transactions are not spread uniformly across time, which implies that the underlying distribution of price changes is a normal-mixture. This fact is well known and observed in many financial time series.

Mathematically, the absence of AG suggests that stable distributions are unsuitable models for log-returns (Eberlein & Keller, 1995). Bingham and Kiesel (2004) note the general rule of thumb is that terms in excess of 16 days typically conform to normality. This is evidenced in Herlemont (2003) where AG is documented from three

months on the CAC-40, and Boavida (2011) where AG is documented for six months, but not for twelve months in the US markets. The authors note that AG has been documented on the US from one month by Campbell, Lo, and MacKinlay (1997), although statistical methodologies between this and the Boavida (2011) studies differ. The authors also note that Kim, Morley, and Nelson (2005) failed to characterize monthly returns on NYSE portfolios as normal. AG has been documented in many markets, on many underlying instruments and asset classes from around the world (see Blattberg & Gonedes, 1974; Diebold, 1986; Diebold, 1988; Madan & Seneta, 1990; Eberlein & Keller, 1995; Campbell, Lo, & MacKinlay, 1997; Herlemont, 2003; Boavida, 2011). Some papers simply assume AG at terms of one month to justify their statistical methodology (Hossain et al., 2009).

The relevance of being able to assume normality in the underlying log-returns is a cornerstone on which much of the theory and applied financial models are built. This is true for both risk and derivative pricing. Dependency on normality pervades all of our classical models - from generic measures of risk, such as VaR, the standard mean-variance portfolio selection models and the Sharpe-Litner-Mossin capital asset pricing models, through to the Black-Scholes-Merton option pricing formula and the use of Itô calculus – all assuming that the marginal distribution of underlying log-price (returns) is normal. In this context, the knowledge that it does not take long before log-returns can safely be assumed to be normal is commonly assumed.

Empirical documentation of AG was evident as early as the days when Maurice Kendall examined the statistical properties of economic time series (Kendall, 1953) and again by Eugene Fama in 1965 (Fama, 1965). AG is currently deemed a consistent feature of returns in the financial literature.

Sewell (2011) remarks:

Given that market log returns are additive, due to the central limit theorem (above), one might expect market log returns above anything but the highest frequency to be approximately normally distributed. This is only the case over the longest of time periods, such as annual returns.

As a simple illustrative example of how AG manifests and is documented, examine the log-returns on the CAC-40 Index on the Paris Bourse for the period 1990 – 2003 inclusive (Herlemont, 2003). Here, Herlemont uses a standard Q-Q plot to assess departures from normality at terms from 10 seconds to 10 minutes (Figure 1) and daily to 6-monthly (Figure 2). A Q-Q plot is a qualitative and graphical method for comparing two probability distributions by plotting their expected versus observed quantiles simultaneously. The goodness-of-fit is visually assessed by the extent to which the observed quantiles differ from expected quantiles from an imputed probability distribution, such as normal.

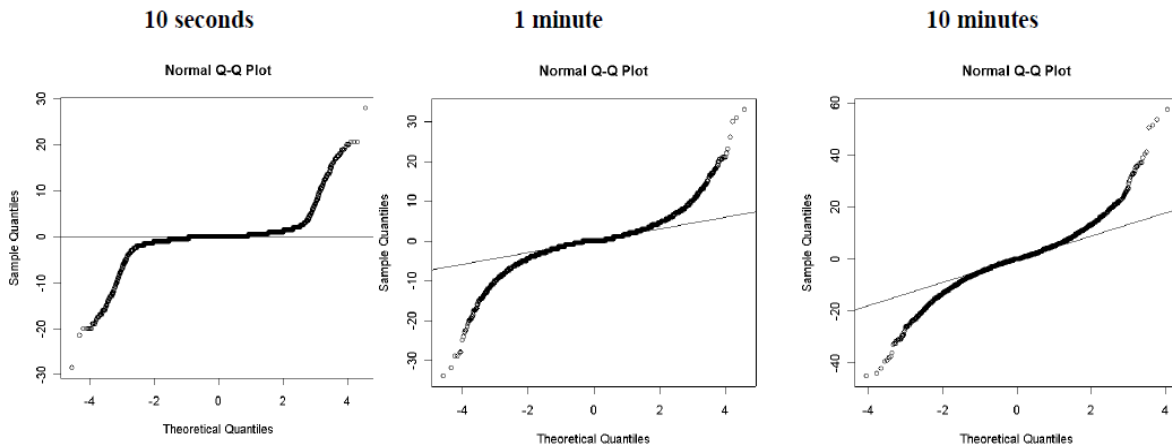


Figure 1: Q-Q- Plots Taken From a Study by Herlemont (2003) on the CAC-40 (Period 1990-2003) for a Variety of Terms - 10 Seconds, 1 Minute and 10 Minutes

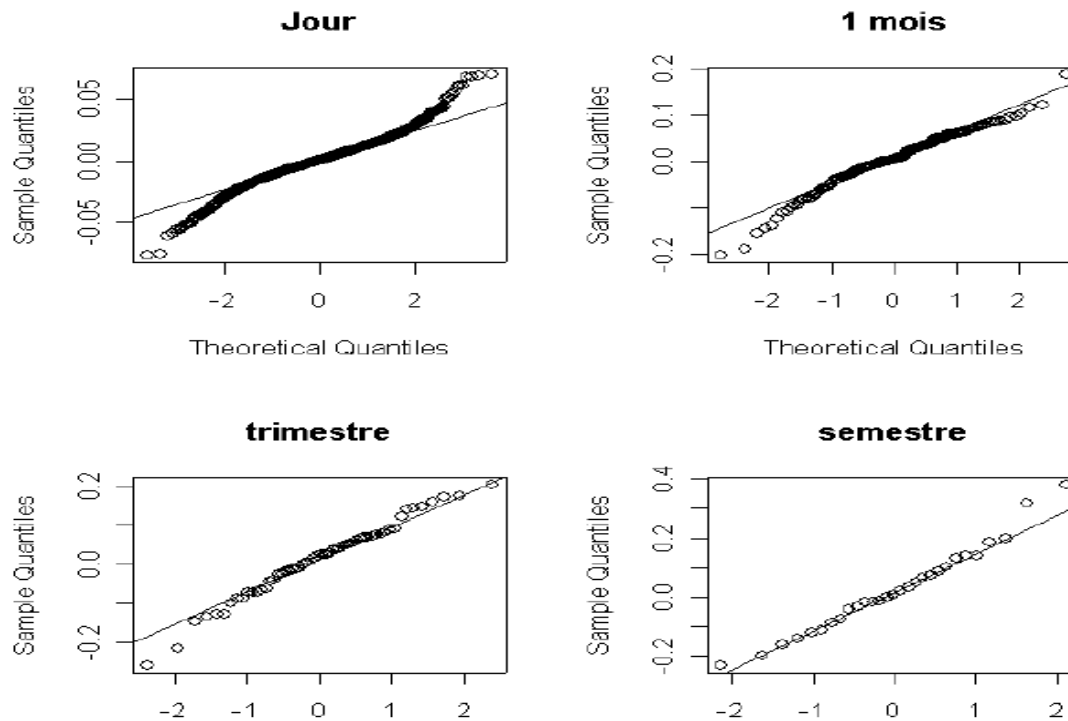


Figure 2: Q-Q- Plots Taken From a Study by Herlemont (2003) on the CAC-40 (Period 1990-2003) for a Variety of Terms - Jour (Daily), Mois (Monthly), Trimestre (Quarterly) and Semestre (6 Monthly)

What is clear from Herlemont's work is that the shift in distributional structure from kurtotic to Gaussian is gradual and as anticipated by AG, with Gaussianity assumed progressively to almost seamlessly at terms of three months and above. Hence, the presence of AG in the CAC-40 is simply proven. It is worthwhile to note that while the use of a Q-Q plot is qualitative rather than inferential, when the underlying assumption is simply being corroborated, no further inferential statistics is really required. When results are more ambiguous, the authors would need to cast the same investigation into an inferential framework and one would require a more complex statistical methodology. The use of inferential statistics in such studies is not uncommon (see Boavida, 2011, for example).

2. METHODS AND RESULTS

In this study, the authors investigate the time-varying distributional characteristics of the Johannesburg Stock Exchange (JSE) All-Share equity index. The All-Share index comprises a market-cap weighted index of 166 constituents. The index is anomalous from the perspective that it is extremely concentrated (five stocks currently make up 39.25% of the index and 10 stocks 57.33% of the index), being dominated by large-cap stocks, the majority of which are dual listed, and resource biased (currently, 33.24% of the top 20 stocks in the All-Share Index are resource companies, 50.28% industrial, and 16.48% financial). These sectorial and concentration anomalies have been the subject of previous scrutiny (see Mbululu, 2009; Flint, 2012).

Daily returns data have been obtained from Bloomberg for the period January 1996 through June 2012. A total of 4,271 daily observations are examined, comprising just over 16 years of data. From the daily returns, we compute the compounded weekly, monthly, quarterly, and annual rolling returns (each calculation of accumulated return for terms greater than a day is termed the 'terminal value'). The various frequencies of returns comprises five distinct 'terms' - daily, weekly, monthly, quarterly, and annually. These observations for the total 16-year period are displayed in Q-Q plots (Figure 3).

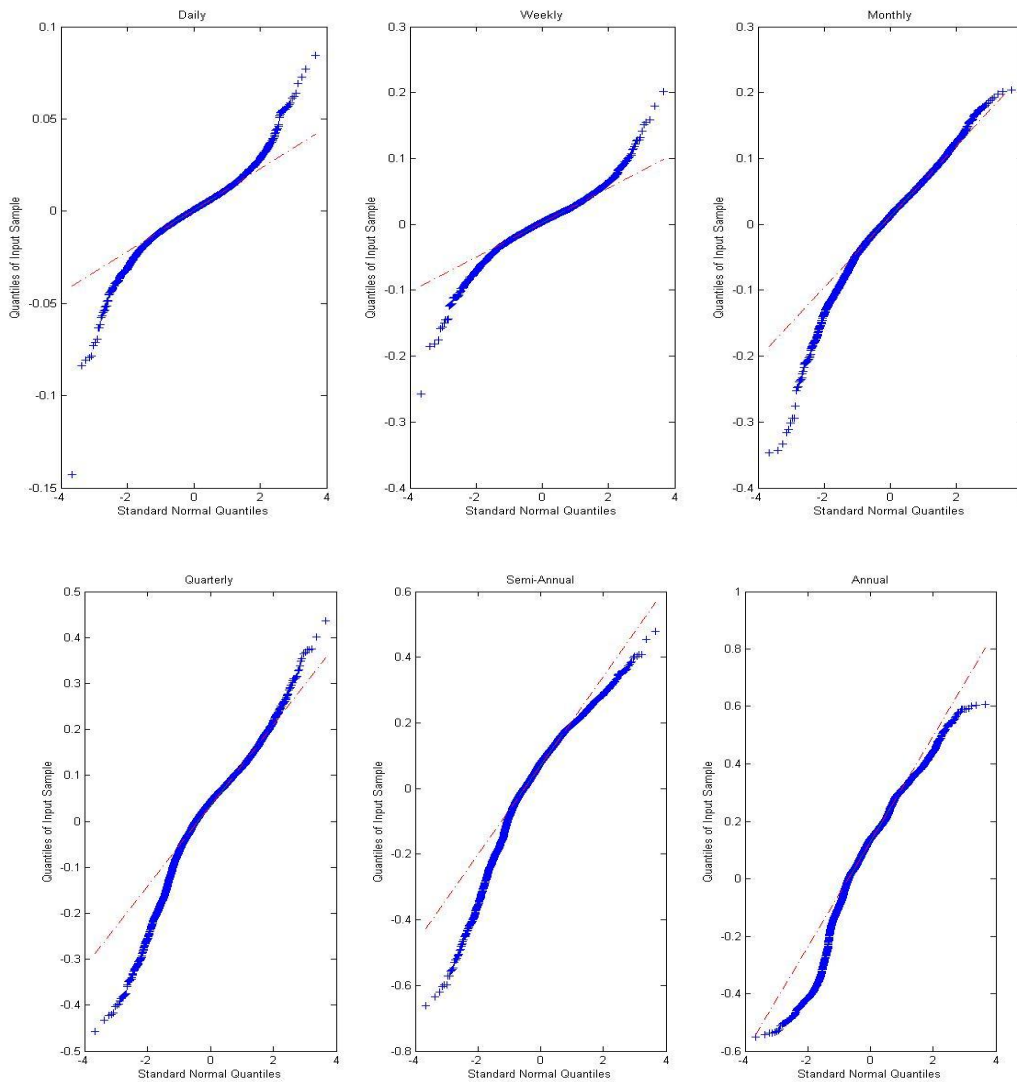


Figure 3: Q-Q Plots From the JSE All-Share Index for the Period 1996-May 2012 (Inclusive) for a Variety of Terms From Daily to Annually (Return Data is Overlapping)

From Figure 3, it appears as though AG seems to hold on the JSE All-Share Index, with the probability distribution of returns becoming progressively more Gaussian as term increases. However, there is still appreciable errant tail behaviour in quarterly and annual returns to be concerned about. Recent work on the Top 40 Index and its constituents (Kulikova & Taylor, 2010) suggests that daily returns data are best described by student-t, variance-gamma, and normal inverse Gaussian distributions. Subsequent investigations suggest that the best fit distribution choice is slowly time-varying and switches between these three. Since two of these distributions do not aggregate to the normal law and are not closed under convolution, this is problematic.

The question now arises regarding the plausibility of these findings within the qualitative Q-Q plot framework. Unfortunately, this qualitative analysis does not provide the necessary statistical rigour required to support or refute the existence of AG. For this objective, one needs to move beyond Q-Q plots and into an inferential framework.

Historical inferential approaches to assessing departures from normality usually rest on a barrage of goodness-of-fit tests; for recent examples, see Ruppert (2004), Boavida (2011), and Khan and Huq (2012). The

authors note with concern that not all tests are equally robust. For example, the commonly used Jarque-Bera test is known to suffer from misdiagnosis in distributions with short tails (Thadewald & Büning, 2007). More concerning is the observation that the admission of specific endpoints or overlapping returns into an inferential framework introduces two palpable sources of statistical bias:

- In any analysis of financial returns over variable terms, there is a range of terminal-values that may be deemed suitable for admission into a statistical analysis. Where a study limits these values to specific term-endpoints (for example, month-end or year-end; see, for example, Boavida, 2011), the statistical analysis is implicitly assuming that the self-same conclusions would be drawn if other endpoints were arbitrarily selected. This is called ‘end-point’ bias. In a world where close-out, calendar, intra-month, holiday, and weekend effects are well documented (see Sewell, 2011, for a summary), this assumption is limiting and potentially calamitous to generalizations.
- The derivation of return intervals over rolling periods introduces serial-dependency into the analysis - something one should be aware of if they are then going to sample the terminal values in an attempt to impute the probability distribution of the same (Rydberg, 2000). Samples from highly-correlated return values will not be independent (as is required for inferential diagnostics) and have a vastly deflated variance. Any inferential conclusions drawn from an analysis using overlapping data is necessarily unconvincing (Polakow, 2010).

Hence, overlapping returns requires special statistical treatment. In order to obviate the effect of end-point bias and auto-correlation amongst the terminal notes, one needs to ensure that samples are spaced as far from each other to be as independent as possible. A single sample of quasi-uniformly distributed terminal return values would incur ‘end-point bias’, with the consequence that the inferences made from a single sample are not generalisable. Hence, the authors adopt a quasi-random resampling regime in which only a percentage of the terminal return values are admitted into the distributional assay based on the availability of unique non-overlapping samples of such terminal values for any specific term. For this purpose, the well-known Sobol' quasi-random uniform sequencing algorithm was used (Sobol', 1967) in conjunction with a resampling (without replacement) technique. This sampling regime and sampling effort adequately prevents overlapping returns due to sampling-error (refer to Appendix A for a discussion on the resampling technique.)

A single quasi-uniform sample of terminal values comprises a single trial. A total of 1,000 trials is run for each term. Return data are sampled from the period 1996-2012, inclusive. After each trial, two goodness-of-fit tests are conducted for normality - the Shapiro-Wilk (SW) test and Anderson-Darling (AD) test. There are many well-known univariate statistical tests for assessing departures from normality. Recent research (Razali & Wah, 2011) concludes that, for a range of simulated symmetric and asymmetric distributions, the Shapiro-Wilk test has the greatest power for large sample sizes, given a stated probability, followed closely by the Anderson-Darling test. The conventional Kolmogorov-Smirnov test ranks lowest of all the well-known goodness-of-fit tests in such circumstances and hence is not deployed in this study. This test is known to be most useful in detecting departures from normality in small sample sizes. Both Anderson-Darling and Shapiro-Wilk tests are two-sided and an alpha level of significance of 5% is adopted for both. Interestingly, if the Shapiro-Wilk and Anderson-Darling tests are run on the total sample, with the high degree of auto-correlation, all six terms fail both AD and SW tests at an alpha of 5%. The authors report on the percentage of trials that are rejected by the null-hypothesis (a null-hypothesis of no difference to normality) for each sampling term. Their a priori expectations for the presence of AG are two-fold and jointly:

- the percentage of rejections should decrease monotonically as sampling term increases if AG is a clear feature of returns on the JSE and
- as mirrored in the international literature (see, for example, Herlemont, 2003), terms from two- to four-weeks and longer should be overtly Gaussian.

The first expectation follows due to the Berry-Esseen theorem which states that if there is a sequence of independent random variables with finite mean, variance and third absolute moment, then convergence of the distribution of the sample mean of these variates to the normal distribution (as motivated by the central limit

theorem) occurs uniformly with the speed of convergence being at least of the order of $\frac{1}{\sqrt{n}}$. This result was discovered independently by Berry (1941) and Esseen (1942). Conformity to both observations would provide strong evidence of AG. Conformity to either of the observations (but not both) signals something other than AG. Results for the AD and SW tests are presented in Table 1.

Each trial comprised a sample of the terminal return values being drawn from the population with the aid of the algorithm described in Appendix A. One thousand trials are conducted.

Table 1: Percentage of Trials That Fail the Goodness-of-Fit Tests for Normality (Via AD and SW Tests)

Term	AD	SW
Daily	100.00%	100.00%
Weekly	100.00%	100.00%
Monthly	75.70%	77.60%
Quarterly	63.90%	59.60%
Semi-Annual	32.70%	30.40%
Annual	9.90%	8.50%

The results evident in Table 1 raise several key observations that do not support the phenomenon of AG on the JSE All-Share Index, which are:

- First, the proportion of observations conforming to normality generally decreases as term is increased, but at different rates. There is constancy up to terms of a week and an inflection point at around a quarterly term.
- Second, terms from daily through to semi-annually (inclusive) have an exceptionally high incidence of the underlying return data being generated by non-normal processes. At annual terms, the proportion of samples failing normality is still in the region of 9%. These two points are depicted graphically in Figure 4.

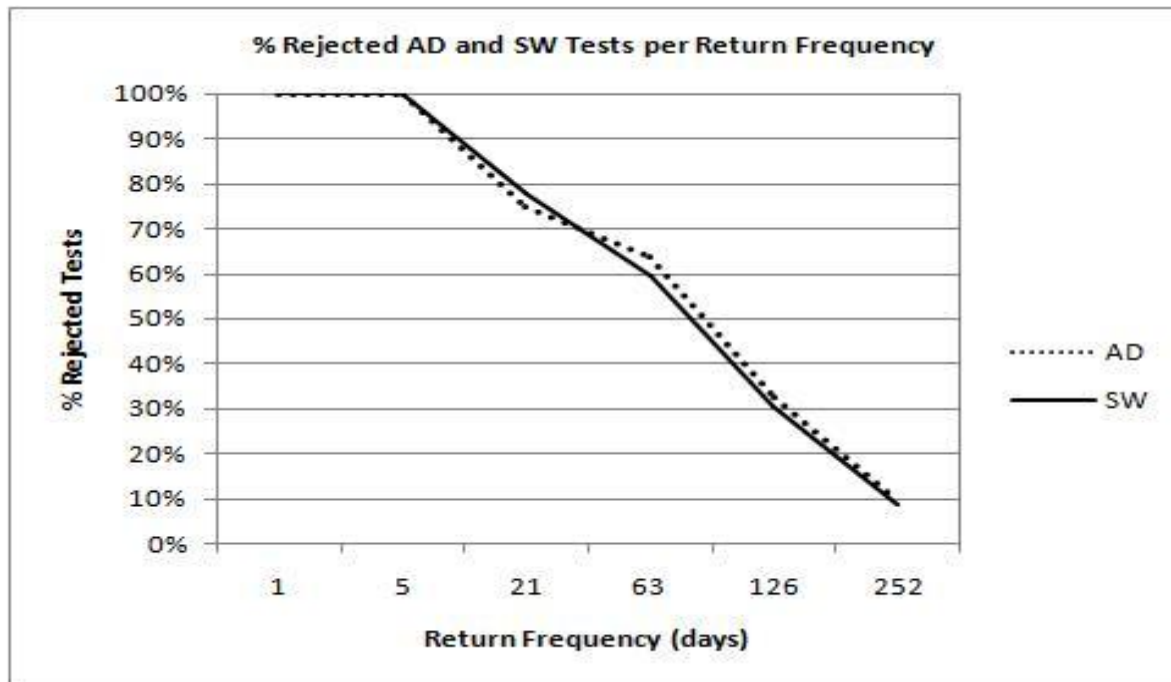


Figure 4: Graphical Illustration of the Proportion of Trials Failing Normality Tests (Via AD and SW) as Term is Increased

Table 1 provides convincing evidence that AG is not a clear feature of the JSE All-Share equity index. The study also documents (the authors believe for the first time) that AG is not a feature even out to terms of twelve

months. AG is not, therefore, a feature of the South African equity market and long-run normality is an assumption that is unfounded.

The natural question at this point would be to understand if the absence of AG is a feature of the South African equity market (an important component of the emerging world markets) or whether these findings were characteristic of the developed world as well. For this reason, the authors cast the same methodology on the S&P500 for interest sake and replicate the same study within another market. The results are contained in Appendix B, which shows that the absence of the AG property is more pronounced in the US equity market, which is much more diversified and less concentrated than the JSE.

Lastly, the longitudinal properties of AG are also interesting. The authors further break up the 16 years into two periods of differing economic relevance (pre-crisis: 1996-2006, and credit-crisis to present: 2007-2012). The methodology follows in the same way as noted above (and motivated in Appendix A.). The results depicted in Figure 5 indicate that in neither of these economically distinctive backdrops was AG a clear feature of returns on the JSE.

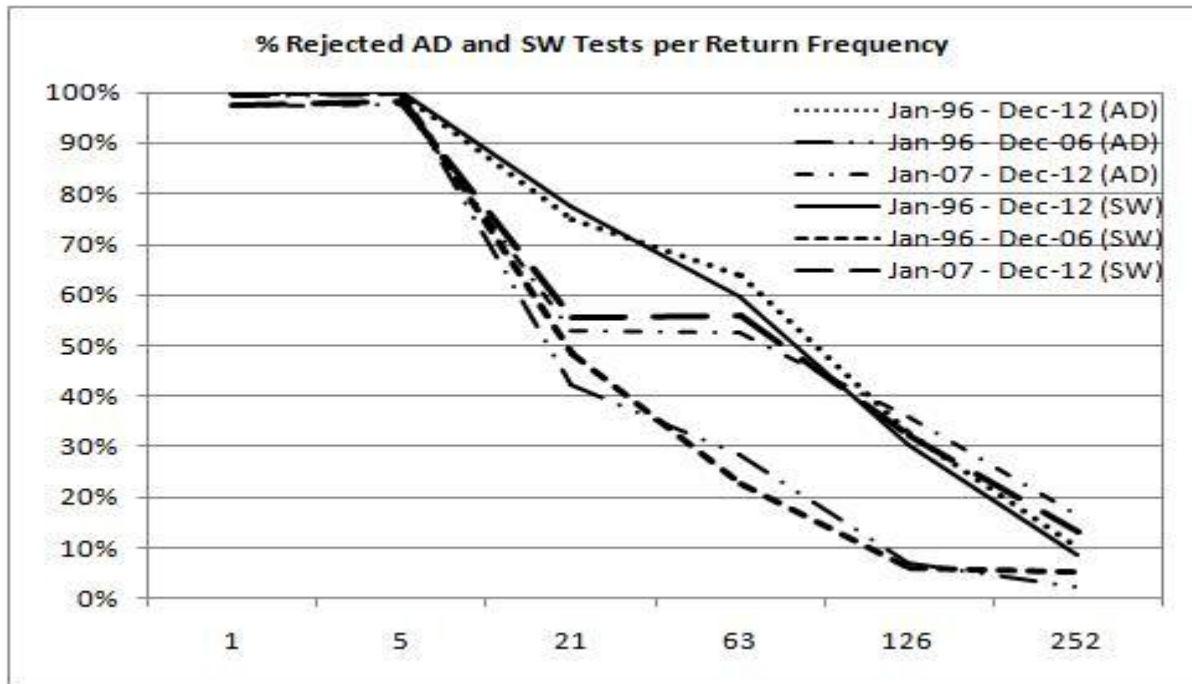


Figure 5: Graphical Illustration of the Proportion of Trials Failing Normality Tests (Via SW And AD) as Term is Increased (Three Periods Identified: 1996-2012 (The Total Period); Then Two Separate Periods: 1996-2006 and 2007-2012)

3. CONSEQUENCES AND CONCLUSIONS

Our inferential investigation reveals that Aggregational Gaussianity (AG) is not a stable property of the South African equity market. Although the distributions are converging to Gaussian, they are converging slower than is generally assumed, and even at a term of one year, they cannot be regarded as normal. The consequences for this may be manifold and are certainly complex.

The results presented here should be of interest to risk managers, in general, and to anyone using quantitative models that invoke normality as an invariant property of term. A discussion of the impact on risk-measures and models assuming normality would require a lengthy exposition, which are omit since the authors believe that many of the implications should be self-evident.

One aspect is noted, however. It is apparent that most previous studies documenting AG have been statistically inadequate, either by virtue of their non-inferential (i.e., graphical) framework or by virtue of their lack of regard for the impact of auto-correlated returns data. The data used in this study exhibits auto-correlation, which has a consequential impact on the non-inferential evidence. It is believed that this study will result in a revision of the understanding of AG in world capital markets. When more robust and reliable statistical techniques are deployed, the results do not support AG as a stylized fact. To this end, the authors trust that the methodology documented here presents an advancement.

Pricing derivatives using the Black-Scholes-Merton model assumes an underlying Brownian motion driving the noise process for stock returns. Returns are then univariate normal and conditional prices are marginally log-normal. It is generally accepted that any pricing approach that relies on such assumptions will underestimate the tail-risk (Bingham et al., 2003). If the true process driving the returns is not Brownian, but is aggregationally Gaussian, then the tail-risk for European derivatives will decrease over long time periods and risk managers and option traders can benefit from this. This manifests as the 'square-root of time' rule when scaling volatility. Option implied volatility surfaces, which often flatten out at longer maturities, seem to conform this as a general market view (see, for example, Gatheral, 2006). Volatility surface structure is evidence of the presence of risk-aversion with respect to the underlying asset process – the more skew and more variable the returns distribution, the higher the aversion and the steeper the 'smile' (Bakshi et al., 2003). However, in the absence of AG, the observed term-structure and flattening of the volatility surface, and the consequential mark-to-market process, represents a potentially serious risk, despite the high uncertainty in estimating realized volatilities.

Of allied concern is the relationship between the real-world (often historical) measure of risk managers and the risk-neutral measure of derivatives traders. One of the attractive features of geometric Brownian motion, with its underlying stochastic driver, is a straightforward understanding of the change of measure function - the Radon-Nikodym derivative (see, for example, Shreve, 2005). This measure change may be viewed as a tilting of either distribution via an unobserved risk-aversion parameter (Bakshi et al., 2003), which appears in most utility functions in various forms. Its effect in the world of Brownian motion is simple. The tilting induces a change in the drift term and leaves the second moment of the marginal distribution - the variance - invariant. This is powerful because it implies that volatility (the standard proxy for risk) has an equivalent meaning under each measure. As a result, historical (therefore statistical) volatility may be compared with implied volatility. In the absence of short, medium, and long-term AG, the comparability between what is regarded as risk-neutral variance and real-world variance becomes blurred. Since implied volatility surfaces are anything but flat, this effect is clearly already known to options markets.

Although the Black-Scholes solution for option prices is relatively robust in the face of price processes that are away from geometric Brownian motion, it is heavily dependent on an accurate estimation of the future (realized) volatility over its lifetime. It is well-known that estimates of variance, skewness, and kurtosis that are carried out on high frequency data (to gain statistical significance) do not translate easily to longer terms and are often biased. For instance, it is generally accepted that the accuracy of forecasts of volatility (variance) improves as the sampling frequency of data increases relative to the forecast horizon for short horizons; however, for longer horizons, model forecasts based on higher frequency data tend to deteriorate largely due to mean reversion considerations (see Poon & Granger, 2003, for an in-depth discussion).

In the wake of the latest market crisis, risk management clearly requires more thought and greater modelling prowess. The interplay between the historical distribution(s) of returns and their characteristics, and that of the pricing kernel methodology of derivative valuation needs urgent attention. The time-varying properties of these distributions and the clear lack of AG shown in the results above would be a perfect point of departure.

AUTHOR INFORMATION

Daniel Polakow is an Associate Professor and the director and founding member of the Institute of Applied Statistics at the University of Cape Town, South Africa. His research interests include alternative investments, financial engineering, derivatives, risk, and statistical modelling. E-mail: daniel.polakow@uct.ac.za

David Taylor is an Associate Professor, the convenor of the MPhil in Mathematical Finance, and the director and founding member of the African Collaboration for Quantitative Finance and Risk Research (ACQuFRR) at the University of Cape Town, South Africa. His research interests include continuous-time asset pricing, interest rate modelling, discrete-time stochastic volatility modelling, and financial time-series analysis. E-mail: david.taylor@uct.ac.za (Corresponding author)

Obeid Mahomed is a research administrator within the Actuarial Science department, a postgraduate student specialising in mathematical finance as well as a research associate of the African Collaboration for Quantitative Finance and Risk Research (ACQuFRR) at the University of Cape Town, South Africa. His research interests include asset pricing, risk management, counterparty credit risk, and fund management strategies. E-mail: obeid.mahomed@uct.ac.za

REFERENCES

1. Bakshi, G., Kapadia, N., & Madan, D. (2003). Stock return characteristics, skew laws, and the differential pricing of individual equity options. *The Review of Financial Studies*, 16(1), 101-43.
2. Berry, A. (1941). The accuracy of the Gaussian approximation to the sum of independent variates. *Transactions of the American Mathematical Society*, 49(1), 122-36.
3. Bingham, N., & Kiesel, R. (2004). *Risk-neutral valuation: Pricing and hedging of financial derivatives*. Springer-Verlag, London.
4. Bingham, N., Kiesel, R., & Schmidt, R. (2003). A semi-parametric approach to risk management. *Quantitative Finance*, 3(6), 426-41.
5. Blattberg, R., & Gonedes, N. (1974). A comparison of the stable and student distributions as statistical models for stock prices. *Journal of Business*, 47, 244-80.
6. Boavida, J. (2011). *Excess returns and normality*. (Master's thesis). Technical University of Lisbon.
7. Bratley, P., & Fox, B. (1988). Implementing Sobol's Quasirandom Sequence Generator. *ACM Transactions on Mathematical Software*, 14, 88-100.
8. Bratley, P., Fox, B., & Niederreiter, H. (1992). Implementation and tests of low discrepancy sequences. *ACM Transactions on Mathematical Software*, 2, 195-213.
9. Campbell, J., Lo, A., & MacKinlay, A. (1997). *The econometrics of financial markets*. Princeton University Press, Princeton, NJ.
10. Clark, P. (1973). A subordinate stochastic process model with finite variance for speculative prices. *Econometrica*, 41, 135-55.
11. Cont, R. (2001). Empirical properties of asset returns: Stylized facts and statistical issues. *Quantitative Finance*, 1, 223-36.
12. Diebold, F. (1986). *Temporal aggregation of ARCH processes and the distribution of asset returns*. Special studies paper, Federal Reserve Board, Division of Research and Statistics.
13. Diebold, F. (1988). *Lecture notes in economics and mathematical systems*. Springer-Verlag, London.
14. Eberlein, E., & Keller, U. (1995). Hyperbolic distributions in finance. *Bernoulli*, 1, 281-99.
15. Embrechts, P., Kluppelberg, C., & Mikosch, T. (1997). *Modelling extremal events*. Springer-Verlag, London.
16. Esseen, C-G. (1942). On the Liapunoff limit of error in the theory of probability. *Arkiv för matematik, astronomi och fysik*, A28, 1-19.
17. Fama, E. (1965). The behaviour of stock market prices. *Journal of Business*, 38, 34-105.
18. Flint, E. (2012). *New perspectives in the modelling of alpha*. (Master's thesis). University of Cape Town, South Africa.
19. Gatheral, J. (2006). *The volatility surface*. Wiley: Finance Series, USA.
20. Herlemont, D. (2003). *Futures markets of stock indices: Stylized facts & automated trading*. (Working Paper). Yats Consulting, USA. Retrieved from <http://www.yats.com/doc/intraday-facts-article-en-ppt.pdf>
21. Hossain, A., Zaman, F., Nasser, M., & Islam, M. (2009). *Pattern recognition and machine intelligence*. Springer-Verlag, Berlin, lecture Notes in Computer Science.
22. Kendall, M. (1953). The analysis of economic series - Part I: Prices. *Journal of the Royal Statistical Society Series A (General)*, 116, 11-25.

23. Khan, M., & Huq, U. (2012). Distribution of risk and return. A statistical test of normality on Dhaka Stock Exchange. *Research Journal of Finance and Accounting*, 3(3), 28-38.
24. Kim, C-J., Morley, J., & Nelson, C. (2005). The structural break in equity premium. *Journal of Business and Economics*, 23, 181-91.
25. Kulikova, M., & Taylor, D. (2010). A Conditionally heteroskedastic time series model for certain South African stock price returns. *Investment Analysts Journal*, 72, 43-52.
26. Li, Y., Hamill, P., & Opong, K. (2010). Do benchmark African equity indices exhibit the stylized facts? *Global Finance Journal*, 21, 71-97.
27. Madan, D., & Seneta, E. (1990). The variance gamma (VG) model for share market returns. *Journal of Business*, 63, 511-24.
28. Mbululu, D. (2009). *Day-of-the-week effect: Evidence from nine sectors of the South African Stock Market*. (Master's Thesis). Rhodes University, Grahamstown, South Africa.
29. Polakow, D. (2010). If a portfolio manager who cannot count finds a four-leaf clover, is he still lucky? *Investment Analysts Journal*, 72, 53-71.
30. Poon, S-H., & Granger, C. (2003). Forecasting volatility in financial markets: A review. *Journal of Economic Literature*, 41(2), 478-539.
31. Razali, N., & Wah, Y. (2011). Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling Tests. *Journal of Statistical Modelling and Analytics*, 2(1), 21-33.
32. Ruppert, D. (2004). *Statistics and data analysis for financial engineering*. Springer-Verlag, London.
33. Rydberg, T. (2000). Realistic statistical modelling of financial data. *International Statistical Review*, 68(3), 233-58.
34. Schmitt, F., Schertzer, D., & Lovejoy, S. (1999). Multifractal analysis of foreign exchange data. *Applied Stochastic Models & Data Analysis*, 15, 29-53.
35. Sewell, M. (2011). *Characterization of financial time series*. Research note rn/11/01, UCL Department of Computer Science.
36. Shreve, S. (2005). *Stochastic calculus for finance I*. Springer-Verlag, London.
37. Sobol', I. (1967). Distribution of points in a cube and approximate evaluation of integrals. *USSR Computational Mathematics and Mathematical Physics*, 7, 86-112.
38. Thadewald, T., & Büning, H. (2007). Jarque-Bera test and its competitors for testing normality - a power comparison. *Journal of Applied Statistics*, 34(1), 87-105.

APPENDIX A – QUASI-RANDOM RESAMPLING METHOD

Sobol' sequences are an example of a quasi-random low discrepancy sequence known to have one of the smallest discrepancy measures, see Bratley and Fox (1988) and Bratley et al. (1992). The aim of such a sequence in one-dimension is to fill the unit interval $[0,1]$ optimally by forming successively finer partitions thereof. Since it is well known that the elements of the unit interval are uncountably infinite, the Sobol' algorithm allows one to progressively sample unique elements of the interval (provided one has the capability of high numerical precision). For finite sampling problems though, where we adapt such an algorithm to sample from a finite number of elements, say, $\{1,2,\dots,N\}$ - the countably finite nature of this set implies that an application of the Sobol' sequencing algorithm (to this finite set) does not preclude the possibility of overlapping or non-distinct samples. We therefore employ a sampling regime based on the Sobol' sequencing algorithm, but ensure that we sample without replacement.

Samples of overlapping returns for lower frequencies (viz. weekly, monthly, quarterly, semi-annual, and annual) are constructed from daily returns, chronologically. When considering such a sample construction for statistical experiments, one should be cognisant of the following observations. Let X_1, X_2, \dots, X_M denote a sample of daily log returns and let $Y_1, Y_2, \dots, Y_{M-f+1}$ denote the corresponding set of f -day overlapping returns (for example, $f = 5$ for weekly returns) constructed from the set of daily returns; i.e., $Y_i = (X_i + X_{i+1} + \dots + X_{i+f-1})$ where $i = 1, 2, \dots, M-f+1$. We may then infer the following:

- the maximal size of a non-overlapping sample of f -day returns that can be drawn from the total population of daily returns, of size M , is given by $\lfloor M / f \rfloor$, where $\lfloor y \rfloor$ denotes the integer part of y ;
- expressed as a percentage of the total overlapping population size, $M-f+1$, the maximal size of a non-overlapping sample of f -day returns is given by $\lfloor M / f \rfloor / (M-f+1) = \lfloor M - (M \text{ modulo } f) \rfloor / \lfloor f (M-f+1) \rfloor$, which provides additional insight into our choice of sample size for a given return frequency; and
- there are only ${}^n C_k$ unique samples of maximal size $\lfloor M / f \rfloor$, given the overlapping population construction, where $n = \lfloor M / f \rfloor + (M \text{ modulo } f)$ and $k = \lfloor M / f \rfloor$.

Since we adopt a quasi-random sampling method as opposed to a deterministic combinatorics- or permutations-based sampling method, we avoid the complexity of having to compute all possible unique combinations of non-overlapping return samples, for a given sample size and return frequency. The observations made above, with regard to the maximal sample size of non-overlapping data for a given return frequency, may therefore appear superfluous. However, those remarks are still useful as they provide one with some a priori idea of the bounds of this sampling problem, particularly from a sample size perspective.

Assume that we are drawing a sample of size N from the sample of f -day returns, $Y_1, Y_2, \dots, Y_{M-f+1}$, as defined above. Further, we shall endeavour to ensure that the sample is non-overlapping; i.e., if Y_i is selected, then any of $Y_{i-1}, Y_{i-2}, \dots, Y_{i-f+1}$ or $Y_{i+1}, Y_{i+2}, \dots, Y_{i+f-1}$ may not be selected, where negative subscripts are of course non-existent for $i < f$. Algorithmically, our quasi-random sampling method may be summarised as follows:

- Generate an element of the one-dimensional Sobol' sequence, say $u \in [0,1]$.
- Convert the Sobol' element, u , into a natural number in the set $\{1, 2, \dots, M-f+1\}$ by setting $x = \lfloor u (M-f+1) \rfloor$, where $\lfloor y \rfloor$ denotes the integer nearest to y .
- Select the f -day return Y_x , which now forms part of the sample.
- Remove $Y_{x-f+1}, Y_{x-f+2}, \dots, Y_{x-1}, Y_x, Y_{x+1}, \dots, Y_{x+f-1}$ from the population of f -day returns.
- Repeat steps 1 to 4, N times.

Of course there is an upper bound on the sample size N , given that we are sampling without replacement. In order to demonstrate the effect of overlapping data (or autocorrelation) on the results of the statistical hypothesis tests for normality, we assume that the population is replenished once the upper sample size bound has been breached. We then compute a statistic which measures the overall proportion or percentage of the sample that contains overlapping data, in order to corroborate the violation of the upper sample size bound.

The results for the JSE All Share equity index for the period ranging from January 1996 to December 2012 are now presented. The first task is to justify the appropriate sample sizes for the respective statistical hypothesis tests for each return frequency. The maximal size of a non-overlapping sample for an f-day return is given by the quantity $\frac{M}{f} - (M-f+1)$ which is used as a guide for the sample size for each of the respective return frequencies under consideration, viz. daily (f = 1), weekly (f = 5), monthly (f = 21), quarterly (f = 63), semi-annually (f = 126), and annually (f = 252). The problem of overlapping sample data for daily returns is redundant - without any loss of generality, sample sizes for daily returns are based on those utilized for corresponding weekly returns. Drawing 1,000 samples was considered following the algorithm presented above for each of the aforementioned return frequencies for varying sample sizes (sampling efforts) expressed as a percentage of the respective maximal non-overlapping sample size.

The results, which have been depicted in Figures 6, 7, and 8, clearly reflect the effect of sample size on the normality tests. A higher sample size, which also increases the chance of overlapping sample data, results in a greater proportion of normality hypothesis tests being rejected and vice versa. Figure 6 shows that a sample size (or sampling effort) of 70% of the maximal non-overlapping sample size provides the largest possible sample size across all return frequencies with zero or minimal overlapping data - as one can observe, a sampling effort of 80% and beyond results in increasing percentages of overlapping data. We therefore conjecture that a sampling effort of 70% is an appropriate level for such a quasi-random inferential analysis.

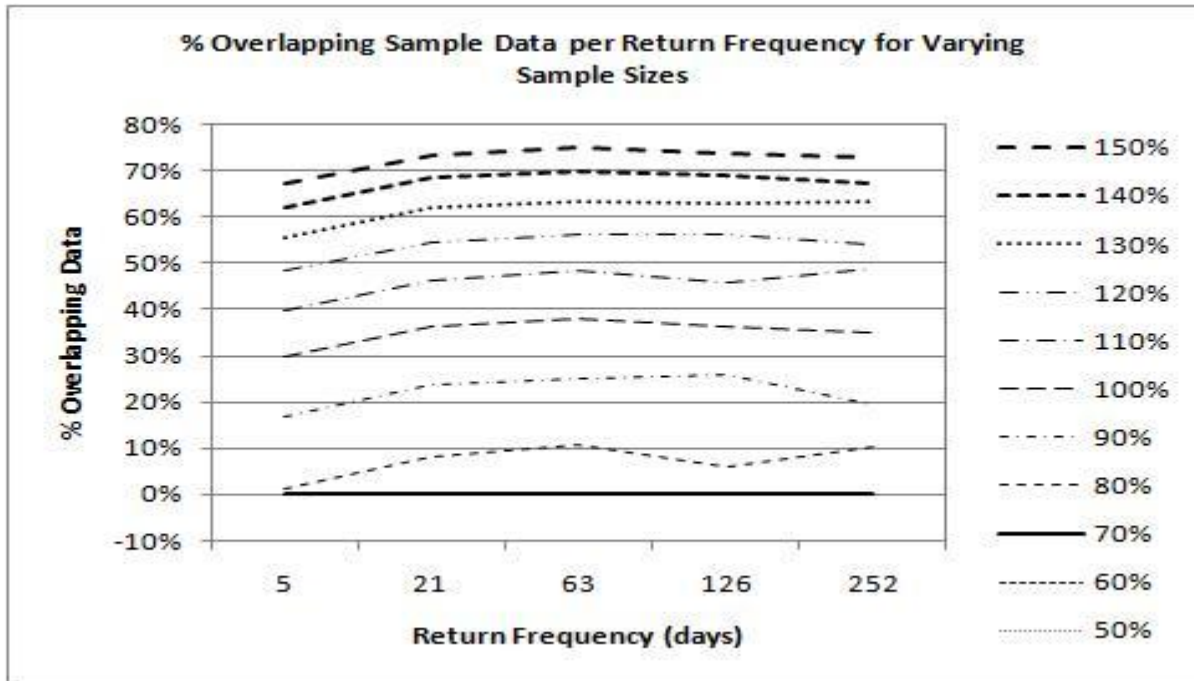


Figure 6: Graphical Presentation of the Percentages of Overlapping Data in the 1,000 Samples Drawn for Each Return Frequency for Varying Sample Sizes Ranging From 50% to 150% (Expressed as a Percentage of the Maximal Size of a Non-Overlapping Sample)

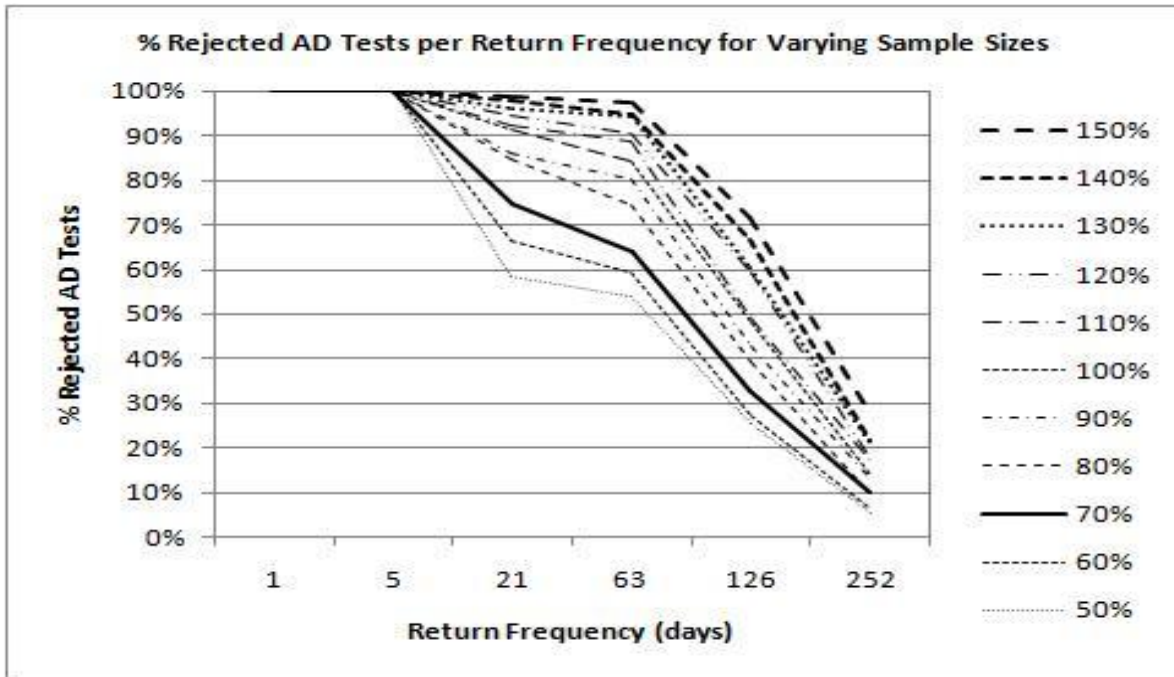


Figure 7: Graphical Illustration of the Proportion of the 1,000 Trials That Failed the AD Normality Test (5% Level Of Significance) for Each Return Frequency for Varying Sample Sizes Ranging From 50% to 150% (Expressed as a Percentage of the Maximal Size of a Non-Overlapping Sample)

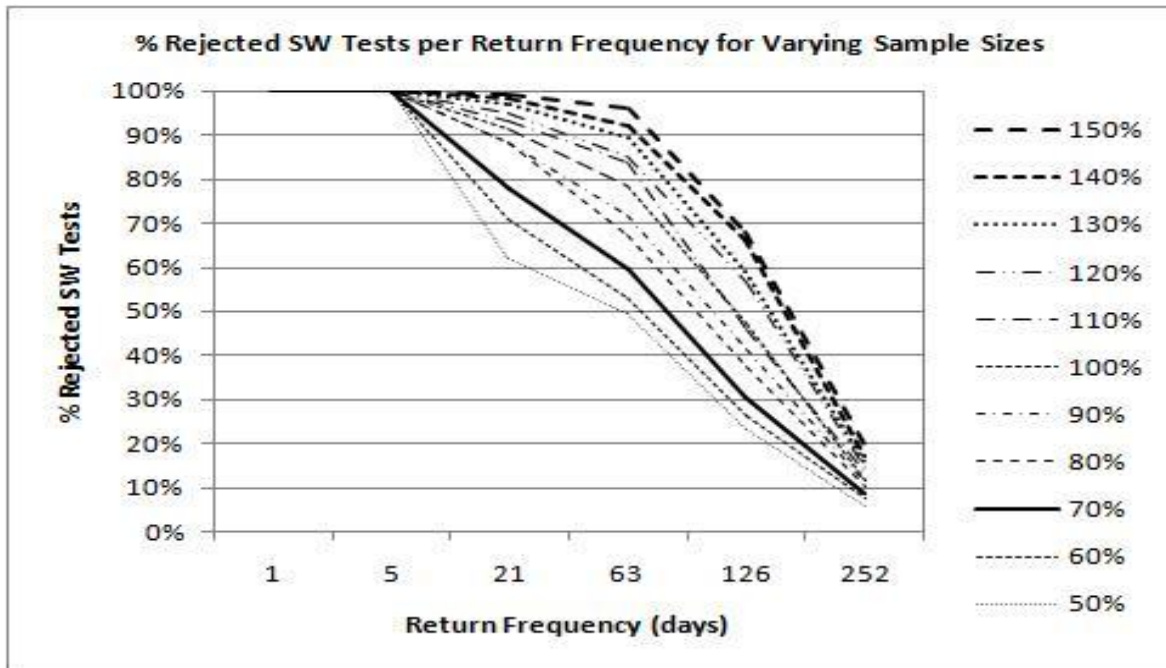


Figure 8: Graphical Illustration of the Proportion of the 1,000 Trials That Failed the SW Normality Test (5% Level Of Significance) for Each Return Frequency for Varying Sample Sizes Ranging From 50% to 150% (Expressed as a Percentage of the Maximal Size of a Non-Overlapping Sample)

Assuming that a sampling effort of 70% is appropriate, the next question that one may ask is whether 1000 samples (trials) is sufficient for satisfactory convergence of results? Figures 9 and 10 are testament to the satisfactory convergence characteristics of the quasi-random inferential framework.

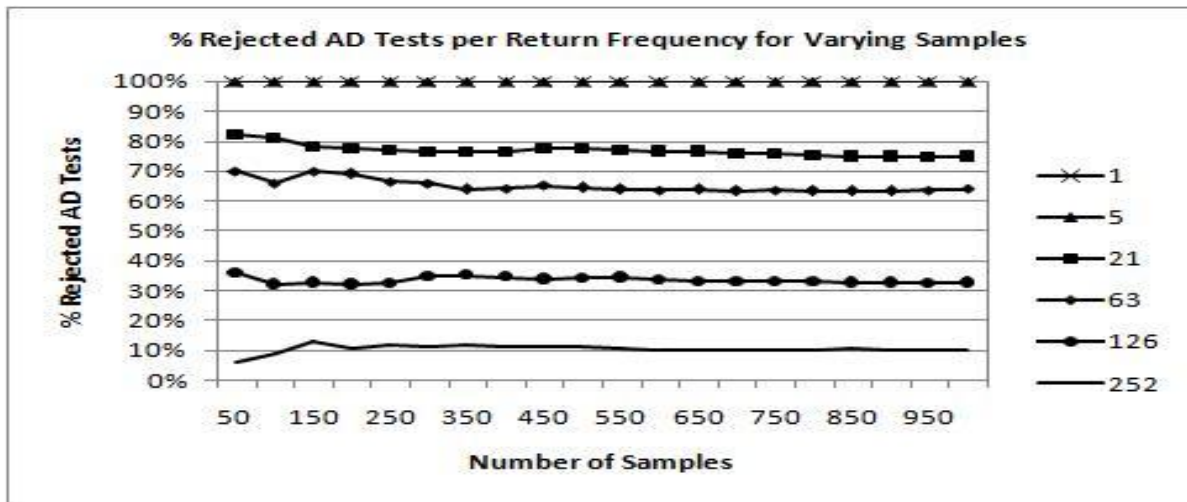


Figure 9: Graphical Depiction of the Proportion of Failed AD Normality Tests (5% Level of Significance) For Each Return Frequency for Varying Numbers of Samples (Trials) Ranging From 50 to 1,000

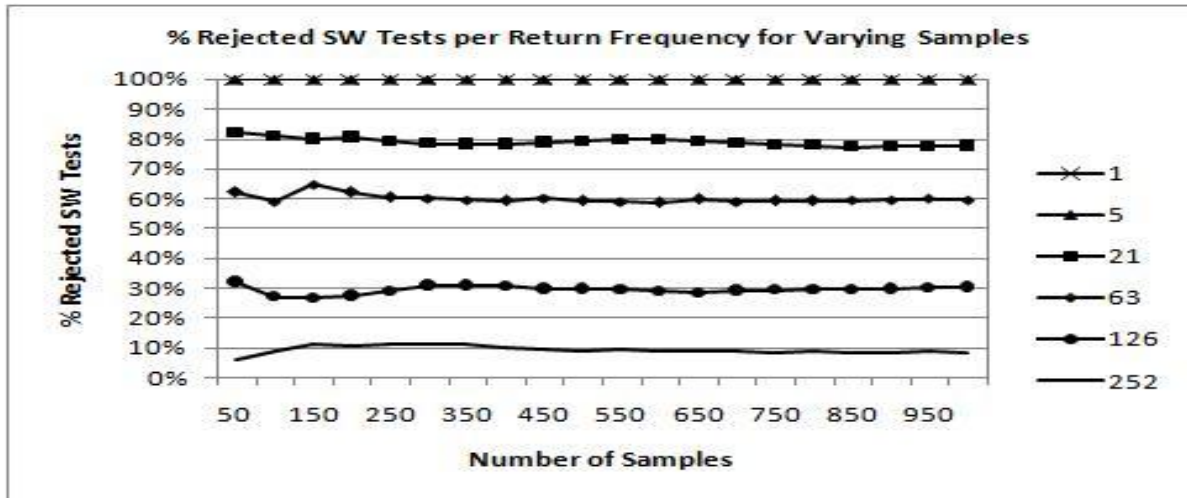


Figure 10: Graphical Depiction of the Proportion of Failed SW Normality Tests (5% Level of Significance) for Each Return Frequency for Varying Numbers of Samples (Trials) Ranging From 50 to 1,000

APPENDIX B – RESULTS FOR THE S&P 500

Here we consider the time-varying distributional characteristics of the Standard & Poor's 500 stock market index, which is a representation of the 500 leading publicly traded companies in the United States stock market. We consider the same time period that was considered for the South African market, viz. January 1996 to December 2012. The procedure described in Appendix A was applied, and the results are summarised below. The absence of AG is even more pronounced in the United States stock market, as the incidence of the underlying return data being generated by non-normal processes dominates the corresponding statistics for the South African market for all respective return frequencies. Results for the AD and SW tests are presented in Table 2 and depicted in Figures 11 and 12, respectively. Note that Figure 12 is a replication of the longitudinal analysis done for the South African market, depicted in Figure 5.

Each trial comprised a sample of the terminal return values being drawn from the population with the aid of the algorithm described in Appendix A. 1000 trials are conducted.

Table 2: Percentage of Trials That Fail the Goodness-of-Fit Tests for Normality (Via AD and SW Tests)

Term	AD	SW
Daily	100.00%	100.00%
Weekly	100.00%	100.00%
Monthly	94.10%	92.30%
Quarterly	57.70%	62.10%
Semi-Annual	44.50%	50.80%
Annual	30.30%	21.90%

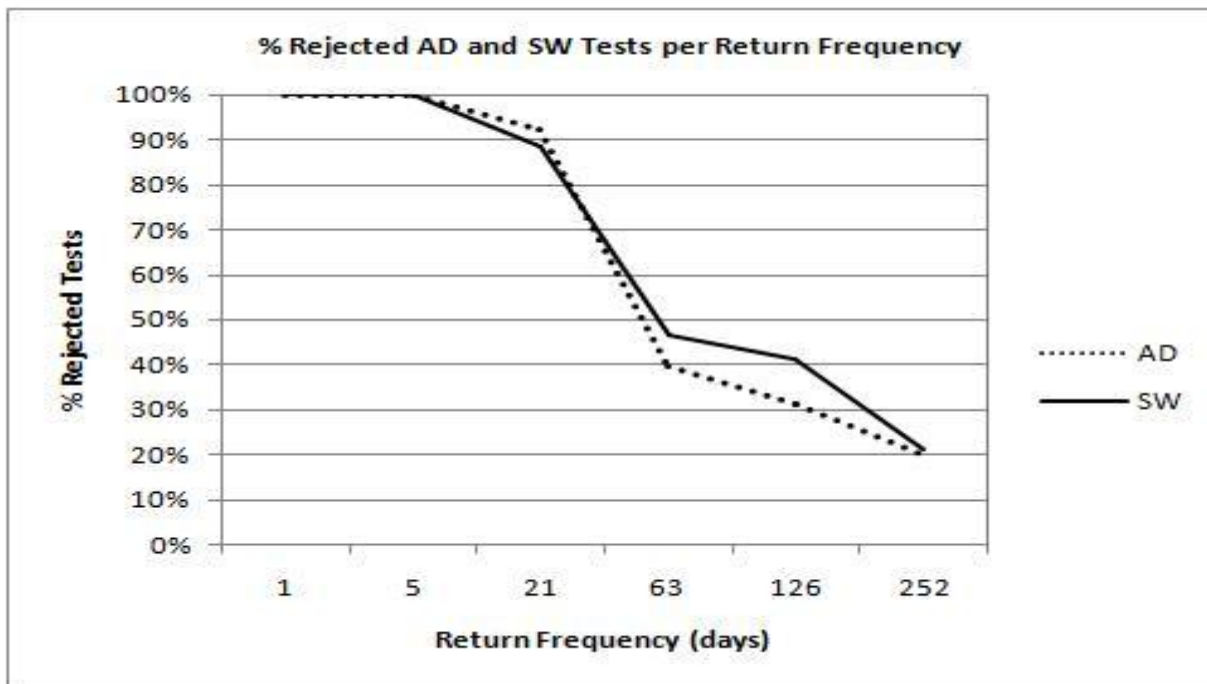


Figure 11: Graphical Illustration of the Proportion of Trials Failing Normality Tests (Via SW and AD) as Term is Increased

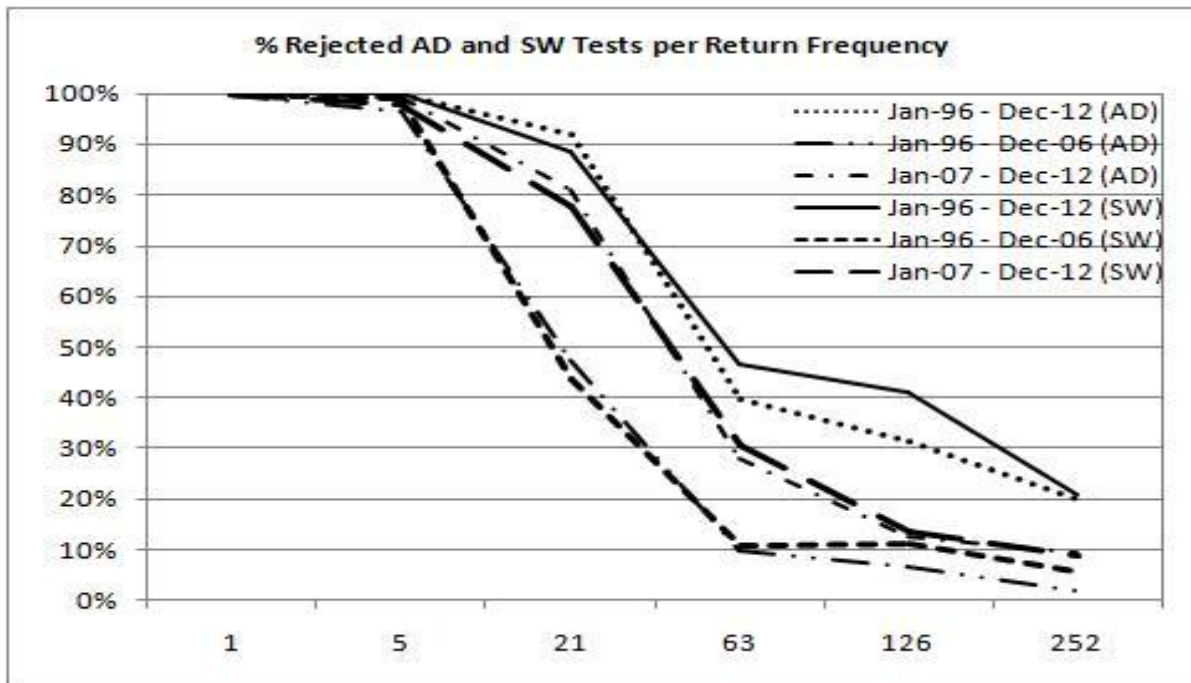


Figure 12: Graphical Illustration of the Proportion of Trials Failing Normality Tests (Via SW and AD) as Term is Increased (Three Periods Identified: 1996-2012 (The Total Period); Then Two Separate Periods: 1996-2006 and 2007-2012)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.