NEW PERSPECTIVES IN THE MODELLING OF ALPHA

EMLYN JAMES FLINT

Master of Commerce - Actuarial Science

Supervised by Dr. Daniel Polakow

School of Actuarial Science

Faculty of Commerce

University of Cape Town

February 2012

Emlyn James Flint: *New Perspectives in the Modelling of Alpha*, Master of Commerce - Actuarial Science, © February 2012

SUPERVISOR:

Dr. Daniel Polakow

Ohana means family.

Family means nobody gets left behind, or forgotten.

— Lilo & Stitch

Dedicated to the loving memory of Granny Flint and Ouma de Jager.

~ 2006 & 2008

ABSTRACT

This dissertation presents an eclectic mix around a central theme of alpha, or value-add. It comprises four essays that are concerned with various theoretical and empirical aspects of alpha. The primary objective is to provide new perspectives in the major areas of modelling alpha; namely, performance measurement, opportunity forecasting and tradability.

Chapter 2 casts a critical eye on an exceedingly popular model utilised in the attribution of fund performance in multi-asset class actively managed funds. This chapter examines the mathematical foundation and assumptions of the Henriksson-Merton (HM) model. The model is plagued by a joint intercept that unwittingly and frequently leads to erroneous conclusions regarding the presence of stock selection abilities in combination with market timing abilities. Utilising a simulation approach, we examine the extent to which this effect impacts on empirical data. We propose that in its current form, the HM piecewise regression incorrectly estimates market timing and stock selection under the presence of restricted (but common) manager skill situations. The model's ability to accurately detect and quantify timing and selection is driven by the absolute difference in up- and down-market alpha values. The HM model cannot accurately detect or quantify timing and selection when the absolute difference in alphas is greater than 80bps per annum. Interestingly, it is also found that the accuracy of the model increases with total fund tracking error. These findings hold under both hypothetical and empirical fund data. We discuss the consequences of these findings.

Cross-sectional volatility (CSV), or return dispersion, is a measure of obvious and increasing relevance to financial mathematicians. That said, the measure itself has been largely neglected from two points of view; namely (i) executing an algebraic decomposition of the measure and (ii) modelling empirical CSV data

via conventional statistical means. To this end, Chapters 3 and 4 investigates the mathematical and empirical properties of CSV respectively.

In Chapter 3, we note that the majority of prior CSV research makes several oversimplifications in order to derive tractable expressions for realised, expected and approximate CSV. However, these simplifications lead to incorrect inferences regarding the underlying mechanisms of CSV. Specifically, CSV has been incorrectly defined as an increasing function of average underlying volatility and a decreasing function of average stock correlation. In Chapter 3, a general theorem for realised CSV is developed, from which a practical and accessible expression for expected CSV is derived, without restrictive market assumptions. Simulating CSV under a broad range of realistic market conditions, it is verified that CSV is actually a monotonically increasing function of the *standard deviation* of the underling volatilities, rather than the average underlying volatility. In addition, it is the *average* CSV that is an inverse function of average stock correlation. Lastly, we note for completeness that there is a material difference between CSV calculated under uniform and market-cap stocks weights, even for relatively uniform markets.

Chapter 4 introduces several novel modelling techniques for empirical CSV. This work initiates addressing the empirical basis for CSV for the first time, specifically (i) it's self-referential properties via auto-regressive modelling, (ii) it's covariation with a suite of broad relevant macro-economic factors, and (iii) the time-varying memory (hazard) characteristics of the series relative to some nominated thresholds. The results from the statistical models reveal that CSV is both extraneously-driven and predictable and possesses an unusual but statistically useful hazard characteristics.

Chapter 5 considers the hidden term-structure of volatility via canonical option valuation. The renowned Black-Scholes option pricing theory implies a constant volatility across term and strike. However, it is known that empirical data violates this assumption. In this chapter, we consider whether the market implied volatility term-structure is justified by that market's historical performance. Stutzer's (1996) nonparametric Canonical Valuation (CV) framework is used in order to develop a tool with which to rank the relative richness of at-the-money options of different

terms over time. We introduce the metric Term-Adjusted-Spread (TAS) and conduct an analysis of the recent volatility term-structure displayed within the South African market. It is shown that the latent South African volatility term-structure is not justified by the market's historical performance, even when considering a number of different historic periods. Potential trading applications using TAS as a measure of direct and relative option mispricing are discussed. Finally, an original, semiparametric option pricing theory is established, which incorporates econometric forecasts of general return distributions within the fundamental CV framework.

If I have seen further it is by

standing on the shoulders of giants.

~ Isaac Newton

ACKNOWLEDGMENTS

This would not have happened without a great supporting - and sometimes leading - cast:

Daniel Polakow - with many fireworks and banners. You deserve a lot more than a line in this thesis.

Jenna, Dylan, Mom, Dad, Jeff & Family - for love, support, stress, frustration, some more support, and some very sound advice.

Laura - for getting me through the rough times and putting up with my shenanigans for the last two years.

The Station, Newlands B and mates - for what is life without friends.

Matt, Kyle, Deepika - friends and office co-habitors. You kept me mostly sane throughout the day.

Michele Laviniere, Russell Investments and OmigSA, for the all important data.

UCT Postgraduate Funding Office and Institute of Applied Statistics, for the all important funding.

People of Finance & Ac Sci: Nashly, Francois, Khader, Ryan, Kanshu, Chun-Sung, Prof Flynn, PvR, Shivani, Dave, Prof MacDonald and John.

As always, all errors in this manuscript are my own.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS

bps basis points

CSV Cross-Sectional Volatility

CAPM Capital Asset Pricing Model

HM Henriksson-Merton

TM Treynor-Mazuy

SML Security Market Line

OLS Ordinary Least Squares

S-P Standard and Poor's

ANOVA Analysis of Variance

ANCOVA Analysis of Covariance

DIA Difference-in-Alphas

DIB Difference-in-Betas

GBM Geometric Brownian Motion

ARIMA Auto-Regressive Integrated Moving Average

MXEF MSCI Emerging Markets Index

BEMETAL Bloomberg Europe 500 Metals and Mining Index

LMEX London Metal Exchange LMEX Metals Index

CRB CMDT Commodity Research Bureau/Reuters US Spot All Commodity

CRY Thomson Reuters/Jefferies CRB Commodities Index

MXEU MSCI Europe Index

IG Inverse Gaussian

CV Canonical Valuation

SAS Strike-Adjusted-Spread

TAS Term-Adjusted-Spread

ALSI All-Share Index

EV Extreme-Value

BSR Best-Subsets Regression

ACF Autocorrelation Function

PACF Partial Autocorrelation Function

GARCH Generalised Auto-Regressive Conditional Heteroscedasticity

AIC Akaike Information Criterion

BIC Bayesian Information Criterion

Part I

DOES ALPHA EXIST?

1

INTRODUCTION

In the financial world, the concept of alpha, or value-add, must rank as one of the most capricious. General financial theory defines alpha as a risk-adjusted measure of performance, or as the excess return after taking into account the risk borne. While academically and strictly true, this definition fails to encompass the contemporary meaning of alpha in its entirety. Perhaps a more intuitive definition is to be found in the words of Tristram Lett, a contemporary Canadian hedge fund manager:

> *"I like to think of alpha as the dark matter of investing. Physicists and astronomers who try to calculate the matter/energy inventory of the universe add up what they know and subtract that from the total and the residual which is very large, they call dark matter. Similarly, financial mathematicians add up what they know (return to betas), subtract it from the observed return and what is left over is called alpha."*[1]

This quote holds several important insights. Firstly, alpha cannot be directly seen nor measured. As with dark matter, it is a construct which, perforce, must exist in order to make sense of the financial universe. Secondly, the fundamental properties of alpha remain unknown. In a physical sense, alpha remains an unknown element, only partially explained by existing theory, which is often of a competing nature. Thirdly, and perhaps a less obvious point, alpha is currently the best explanation that financial theoreticians have. Although one can neither directly observe nor understand dark matter and its properties, the culmination of mankind's physical understanding necessitates that another form of matter must exist within our universe. The concept of alpha plays an equivalent role within the financial

1 Although the physical argument proposed in this quote is not strictly true as there is no mention of dark energy, which accounts for a far larger estimated proportion of the universe than dark matter, the quote still manages to capture the essence of alpha in the more colloquial sense. Quote accessed on 2 February 2012, from: http://www.investmentreview.com/expert-opinion/what-is-alpha-and-does-it-still-exist-4442?Print.

universe. Thus, for all the arguments and debates regarding whether alpha truly exists in practice, the underlying theoretical construct remains a necessity within contemporary finance.

From it's humble beginnings as Jensen's (1968) additional intercept term within the renowned Capital Asset Pricing Model (CAPM), alpha has since become something upon which entire financial institutions have been built, and destroyed. A multitudinous array of financial, mathematical and statistical literature has weighed, measured and deconstructed the mystique surrounding alpha's multi-faceted and ever-changing nature over several decades, and will most probably continue to do so for many more. Hopefully, this manuscript will aid in the fundamental understanding of this mercurial concept.

This dissertation presents an eclectic mix around a central theme of alpha. It comprises four essays that are concerned with various theoretical and empirical aspects of alpha. The primary objective is to provide new perspectives in the major areas of modelling alpha; namely, performance measurement, opportunity forecasting and tradability.

Chapter 2 casts a critical eye on an exceedingly popular model utilised in the attribution of fund performance in multi-asset class actively managed funds. This chapter examines the mathematical foundation and assumptions of the Henriksson-Merton (HM) model. The model is plagued by a joint intercept that unwittingly and frequently leads to erroneous conclusions regarding the presence of stock selection abilities in combination with market timing abilities. Utilising a simulation approach, we examine the extent to which this effect impacts on empirical data. We propose that in its current form, the HM piecewise regression incorrectly estimates market timing and stock selection under the presence of restricted (but common) manager skill situations. The model's ability to accurately detect and quantify timing and selection is driven by the absolute difference in up- and down-market alpha values. The HM model cannot accurately detect or quantify timing and selection when the absolute difference in alphas is greater than 80bps per annum. Interestingly, it is also found that the accuracy of the model increases with total fund tracking error. These findings hold under both

hypothetical and empirical fund data. We discuss the consequences of these findings.

Cross-sectional volatility (CSV), or return dispersion, is a measure of obvious and increasing relevance to financial mathematicians. That said, the measure itself has been largely neglected from two points of view; namely (i) executing an algebraic decomposition of the measure and (ii) modelling empirical CSV data via conventional statistical means. To this end, Chapters 3 and 4 investigates the mathematical and empirical properties of CSV respectively.

In Chapter 3, we note that the majority of prior CSV research makes several oversimplifications in order to derive tractable expressions for realised, expected and approximate CSV. However, these simplifications lead to incorrect inferences regarding the underlying mechanisms of CSV. Specifically, CSV has been incorrectly defined as an increasing function of average underlying volatility and a decreasing function of average stock correlation. In Chapter 3, a general theorem for realised CSV is developed, from which a practical and accessible expression for expected CSV is derived, without restrictive market assumptions. Simulating CSV under a broad range of realistic market conditions, it is verified that CSV is actually a monotonically increasing function of the *standard deviation* of the underling volatilities, rather than the average underlying volatility. In addition, it is the *average* CSV that is an inverse function of average stock correlation. Lastly, we note for completeness that there is a material difference between CSV calculated under uniform and market-cap stocks weights, even for relatively uniform markets.

Chapter 4 introduces several novel modelling techniques for empirical CSV. This work initiates addressing the empirical basis for CSV for the first time, specifically (i) it's self-referential properties via auto-regressive modelling, (ii) it's covariation with a suite of broad relevant macro-economic factors, and (iii) the time-varying memory (hazard) characteristics of the series relative to some nominated thresholds. The results from the statistical models reveal that CSV is both extraneously-driven and predictable and possesses an unusual but statistically useful hazard characteristics.

Chapter 5 considers the hidden term-structure of volatility via canonical option valuation. The renowned Black-Scholes option pricing theory implies a constant volatility across term and strike. However, it is known that empirical data violates this assumption. In this chapter, we consider whether the market implied volatility term-structure is justified by that market's historical performance. Stutzer's (1996) nonparametric Canonical Valuation (CV) framework is used in order to develop a tool with which to rank the relative richness of at-the-money options of different terms over time. We introduce the metric Term-Adjusted-Spread (TAS) and conduct an analysis of the recent volatility term-structure displayed within the South African market. It is shown that the latent South African volatility term-structure is not justified by the market's historical performance, even when considering a number of different historic periods. Potential trading applications using TAS as a measure of direct and relative option mispricing are discussed. Finally, an original, semiparametric option pricing theory is established, which incorporates econometric forecasts of general return distributions within the fundamental CV framework.

Part II

ALPHA PERSPECTIVES

2

UNMASKING THE HENRIKSSON-MERTON MODEL

2.1 INTRODUCTION

Treynor and Mazuy (1966) and Fama (1972) were among the first of many to put forward the idea that fund performance relative to a benchmark could be categorised into two distinct components: (1) forecasts of price movements of selected individual stocks (i.e. "micro-forecasting"); and (2) forecasts of price movements of the general stock market as a whole (i.e. "macro-forecasting"). Since their work, this paradigm has dominated investment performance literature. The two respective components are now more commonly referred to as stock selection and market timing. The stock selector tries to forecast the non-systematic component of the return on individual stocks, while the market timer tries to forecast the differential risk-premia across asset classes.

The first attempt at evaluating stock selection and benchmark sensitivity for an investment portfolio was carried out by Friend et al (1962). Although the study did not specifically introduce model parameters to quantify selection or benchmark sensitivity, it did compare fund performance to that of a comparable market portfolio in a general sense. Following in this research avenue, Sharpe (1964) then proposed the Capital Asset Pricing Model (CAPM). It was the first general linear model to incorporate beta – measuring the sensitivity of excess fund returns to excess market returns – into the fund performance model. Extensions to incorporate alpha – a measure of portfolio returns attributable to the manager's stock selection abilities – into the CAPM framework were subsequently made by Sharpe (1966) and Jensen (1968); both of whom concluded that fund managers deliver negative abnormal returns. However, more recent studies by Berk and

Green (2004) and Nitzsche, Cuthbertson and Sullivan (2006) find evidence of managers achieving positive excess returns.

Over the years, there have been many different methodologies suggested - and studies conducted - to measure market timing and stock selection ability concurrently (Ippolito (1993)).[1] However, the models put forward by Treynor and Mazuy (1966) and Henriksson and Merton (1981) (hereafter referred to as TM and HM respectively) have become the most commonly used in practice. Many contemporary studies continue to focus on these models - or variations thereof - for understanding the portfolio manager's ability to add value through security selection and tactical asset allocation. We include in these recent studies Lhabitant (2001), Bradfield and Swartz (2003), Cuthbertson, Nitzsche and Sullivan (2004), Romacho and Cortez (2006), Chen and Liang (2007), Sehgal and Jhanwar (2008), and Ferruz, Sarto and Vargos (2010) - spanning performance analysis of the Swiss, South African, UK, Portuguese, US, Indian and Spanish investment fund markets respectively.

Despite the algebraic simplicity of the initial HM formulation, and the long-standing usage of the same, we believe the original HM model is critically flawed. This paper examines the nature of the HM model's mathematical construction. Our concern is the regularly disregarded correlation imposed between the model's stock-selection and market-timing coefficients induced by the joint intercept term. It is the intention of this contribution to better understand and quantify the consequences of this forced coupling. Interestingly, because of this flaw, the HM model could potentially fail on two accounts: (i) the estimated coefficients from the model could be incorrect; and (ii) if the empirical data does not conform to the imposed correlation structure intrinsic to the HM model, then while the regression coefficients may be correctly estimated, the interpretation of these coefficients as suitable proxies for timing and selection may be incorrect. Our motivation in this contribution is to understand the consequentiality of these statistical effects and the relevance of the interpretation to critical components of fund manager skill.

1 Ippolito (1993) gives a summary of the findings of the major pre-1990 mutual fund performance attribution studies. Appendix A.1 gives the summary table of his work. More recent studies include Malkiel (1995), Moskowitz (2000), Kosowski (2006), Staal (2006), Fama and French (2010) and Glode (2010).

The paper is organised as follows: Section 2.2 briefly reviews the theoretical literature fund performance measurement before focusing specifically on the HM model and exposing the statistical oversight within the existing model's framework. Section 2.3 provides the data simulation process, details the hypothesis test methodology used and introduces an Analysis of Covariance model as a suitable null case. In Section 2.4, the HM model is tested using simulated fund data in order to see whether it can correctly detect and subsequently correctly quantify a fund manager's market timing and stock selection abilities. This test identifies the accuracy of the model when alpha, beta and fund tracking error variables are varied for both positive and negative market returns. Based on these results, a practical framework is developed stating when the HM model should be used and how the output should be interpreted. The HM model is then applied in Section 2.5 to a sample of American long-only mutual funds in order to test the framework under real market conditions and assess whether the results from the simulation exercise are corroborated. Section 2.6 concludes.

2.2 FUND PERFORMANCE ATTRIBUTION MODELS

For the most part, there have been two methods outlined in the financial literature to quantify a manager's stock selection and market timing abilities. The first method only requires one to assume knowledge of a fund's return distribution in order to use parametric techniques to compare these returns to a benchmark return series. The second approach requires all manager forecasts to be known for the period being studied. With this additional knowledge, one is then able to use nonparametric techniques to quantify the timing and selection ability of the manager. Both methods do require that all returns are taken from stationary distributions. This is due to the fact that changes in distribution parameters could be misinterpreted as timing or selection ability. Although the second method is probably the more accurate of the two, this section will limit itself to looking only at the first approach as manager forecasts generally cannot be observed by investors in practice.

2.2.1 *The Evolution of Timing and Selection Performance Models*

The CAPM framework, independently introduced by Treynor (1961, 1962), Sharpe (1964), Lintner (1965a,1965b) and Mossin (1966), was arguably the first performance attribution model advanced in the financial literature.[2] Built upon numerous market assumptions, CAPM states that the excess portfolio (or security) returns for period t, R_{pt}, are linearly dependent only upon non-diversifiable risk. This is given mathematically by the Security Market Line (SML):

$$R_{pt} = \beta_p R_{mt} + \varepsilon_{pt}, \tag{2.1}$$

where R_{mt} is the excess market return, β_p is the sensitivity of the portfolio to excess market returns and $\varepsilon_{pt} \overset{iid}{\sim} \mathcal{N}(0, \sigma^2)$. Returns are taken in excess to the risk-free return at time t.

Building on this framework, Sharpe (1966) and Jensen (1968) pioneered the work defining the security selection skill of fund managers. This work resulted in the commonly quoted 'market model', which added another parameter to the SML. Thus, the relationship between excess portfolio returns at time t, R_{pt}, and excess market returns at time t, R_{mt}, was now given as:

$$R_{pt} = \alpha_p + \beta_p R_{mt} + \varepsilon_{pt}, \tag{2.2}$$

where α_p - commonly referred to as Jensen's alpha - measures a manager's security selection skill and β_p and ε_{pt} are defined as above. Thus, the market model portfolio returns were now considered to be an affine function, rather than purely a linear function. Intuitively, α_p represents the difference between the managed portfolio's return and those of the portfolio's benchmark. Thus, α_p measures the return attributable to active management.

Fama and French's (1993) three-factor model and Carhart's (1997) four-factor model are the most popular extensions of the market model. The additions were proposed in order to correct for most of the anomalies inherent in the CAPM

2 Jensen (1972b) gives a comprehensive review of this model.

framework. Bollen and Busse (2001, 2005) argue that one should rather make use of Carhart's multi-factor model in order to ensure that managers are not incorrectly rewarded with timing or selection skill for merely exploiting these anomalies. Carhart's modified market model is stated as follows:

$$R_{pt} = \alpha_{pt} + \sum_{k=1}^{4} \beta_k F_{kt} + \varepsilon_{pt}, \tag{2.3}$$

where F_{kt} represents excess market returns, the Fama & French size and book-to-market factors, and Carhart's momentum factor respectively, β_k is the k^{th} factor sensitivity coefficient, and $\varepsilon_{pt} \overset{iid}{\sim} \mathcal{N}(0, \sigma^2)$.

One of the problems noted by Jensen (1972) when using Equation 2.2 to measure fund performance - also applicable to Fama and French (1993) and Carhart's (1997) extensions - was that it failed to take market timing into account, which biases estimates of α_p. It is imperative that both timing and selection be measured concurrently. According to Equation 2.2, fund returns should display a constant beta, irrespective of market situation. However, for a manager skilled at market timing, the portfolio beta for positive or 'up' (negative or 'down') market returns should be higher (lower) than the estimated β_p. In order to capture this proposed non-linearity, Treynor and Mazuy (1966) added a quadratic term to the market model given in equation 2.2. For a perfect market timer, Treynor and Mazuy (1966) reasoned portfolio returns should be related to market returns by a continuous, two-state kinked beta model. However, by assuming that no manager in practice would have perfect forecasting abilities, they argued that larger up- or down-markets should have greater probabilities of a correct forecast attached to them and thus, cause a greater change in asset allocations. Treynor and Mazuy (1966) reasoned that a better-than-average timer should show a continuous, systematic and smooth transition from a portfolio completely invested in cash to one completely invested in the market index. In addition, whether the actual relationship is smooth or kinked, the addition of a quadratic term will always improve a

least-squares statistical fit in comparison to the market model. Mathematically, the TM model is given as

$$R_{pt} = \alpha_p + \beta_p R_{mt} + \gamma_p (R_{mt})^2 + \varepsilon_{pt} \tag{2.4}$$

where $\gamma_p > 0$ indicates positive market timing ability. Note that γ_p is a measure of both the manager's private information and, more importantly, the response to this information. Thus, given the same information set, risk-seeking managers will take larger actions than risk-averse managers. In addition, risk-seeking managers will also act on lower quality information than managers that are risk-averse. For this reason, Lhabitant (2001) cautions against using γ_p to rank market timers.

2.2.2 *The HM Model Framework*

Henriksson and Merton (1981) developed a different approach to that of Treynor and Mazuy (1966). The basis of this method is that fund managers are able to select between discretely different systematic target risk levels, dependent on the forecast made. Henriksson and Merton (1981) extended the market model to include timing by separating excess market returns into up- and down-market variables and fitting separate coefficients to each.[3] While the general HM equation describes a two-state systematic risk model, Henriksson (1984) showed how this technique can be extended to multiple target systematic risk levels. Merton (1981) demonstrated that the total returns for a market timing fund are identical to those found from following a protective put strategy; that is, total fund returns equal the weighted sum of excess market returns and market put options with an exercise price equal to the risk-free rate, R_{ft} (where the weights are specified in Merton's (1981) research). The value of market timing for the fund is that these put options are obtained for free. Following from this finding, the HM model is described as:

$$R_{pt} = \alpha_p^* + \beta_1 X_t + \beta_2 Y_t + \varepsilon_{pt}, \tag{2.5}$$

3 Although Treynor and Mazuy (1966) argue that this situation will only be seen for perfect timers, the rationale behind the HM model is that, if a manager is selecting between between two systematic, target risk levels, timing ability should be independent of the size of the up- or down-market.

where $X_t \equiv R_{mt} - R_{ft}$, $Y_t \equiv max[0, R_{ft} - R_{mt}] = max[0, -X_t]$, and $\varepsilon^*_{pt} \overset{iid}{\sim} \mathcal{N}(0, \sigma^2)$.

Under this specification, Y_t represents the return on the market put options and β_2 represents the number of zero-cost put options provided by the manager's market timing ability. Thus, Henriksson and Merton concluded that a fund displays positive market timing ability if and only if the least square estimate $\beta_2 > 0$. The portfolio manager's security selection is again captured by the intercept coefficient, α^*_p. From equation 2.5, an up-market is defined as $X_t > 0$, while a down-market is defined as $X_t < 0$. In recent years, and especially in industry, the alternate specification of the HM model has gained prevalence due to the more intuitive meaning of the regression coefficients. The alternative HM model is described as:

$$R_{pt} = \alpha^*_p + \beta^*_1 X_{1,t} + \beta^*_2 X_{2,t} + \varepsilon^*_{pt}, \tag{2.6}$$

where $X_{1,t} \equiv min[0, X_t]$, $X_{2,t} \equiv max[0, X_t]$ and once more, $\varepsilon^*_{pt} \overset{iid}{\sim} \mathcal{N}(0, \sigma^2)$. For the remainder of the chapter, the HM model will refer exclusively to Equation 2.6.

The up-market target systematic risk level is given by β^*_1, while β^*_2 estimates the down-market target level. Given that $\beta_2 = \beta^*_2 - \beta^*_1$, market timing ability is now measured by the difference in β^*_2 and β^*_1. When $\beta^*_2 > \beta^*_1$, one says that a manager has positive market timing skill, while a manager that has $\beta^*_1 > \beta^*_2$ has negative market timing skill. Larger differences between the betas equates to greater positive - or negative - timing ability. Figure 1 depicts the general output that would be seen for a manager that displayed both positive market timing and stock selection skills.

Henriksson and Merton (1981) note that the β^*_1 and β^*_2 distributions are not stationary, which causes the standard deviation of the error terms, ε_{pt}, to increase with $|X_t|$. This suggests a correction for heteroscedasticity. However, Henriksson (1984) and Chang and Lewellen (1984) found that there is no material difference when using Ordinary Least Squares (OLS) compared to a weighted least squares method that takes heteroscedasticity into account. Contrastingly Breen, Jagannathan and Ofer (1986) argue that this result is rather due to the specific period

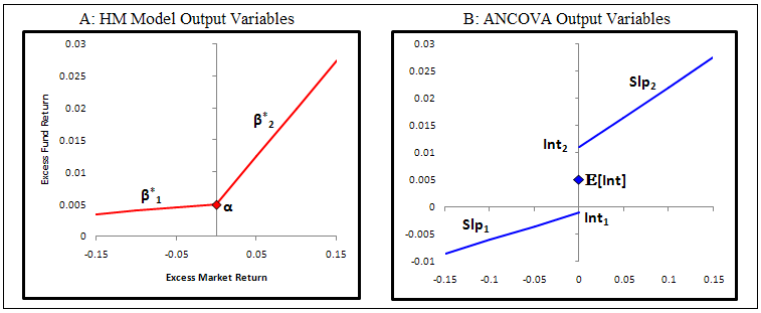


Figure 1: Standard output from the HM and Analysis of Covariance models.

Note. Panel A gives the standard regression output from the HM model for simulated fund data. Note the dependency between the β_i^*'s and the single intercept of the model, α_p^*. Conversely, Panel B displays the standard Analysis of Covariance output for data grouped by positive and negative excess market return on the same fund data. The inclusion of a second intercept ensures independence between all model variables.

analysed and proposed that correcting for heteroscedasticity does significantly alter the coefficient estimates. However, in practice it would seem that most market participants use the ordinary least-squares approach (Bradfield & Swartz (2003)).

The Statistical Flaw

The HM model tests the joint hypothesis of no security selection or market-timing ability attributable to the portfolio manager, with the alternative hypothesis being that the manager does display security selection and/or market timing skill. Using the coefficients in Equation 2.6, this is represented algebraically as:

$$\begin{aligned} H_0 &: (\alpha_p^* = 0) \wedge (\beta_2^* = \beta_1^*) \\ H_1 &: (\alpha_p^* \neq 0) \vee (\beta_2^* \neq \beta_1^*) \end{aligned} \tag{2.7}$$

The regression specification of the HM model assumes that the coefficients - α_p^*, β_1^* and β_2^* - are independent. However, by forcing β_1^* and β_2^* to have the same vertical intercept, α_p^*, the assumption of independence is violated and a distinct correlation structure between the model coefficients is introduced.[4] Interestingly, Henriksson

4 Intuitively, one can think of α_p^* as a movable pivot connecting two lines. As the pivot moves up and down, the gradients of the lines also changes . As a simple example, consider Figure 1. Currently, one infers both positive stock selection and market timing skills from the fitted model coefficients. Together, these coefficients represent the lines of best fit for the down- and up-markets respectively. Now assume that all the excess fund returns in the up-market only (to the right of the y-axis) were shifted upwards by 50bps. This should increase *up-market* α_p^* by 50bps, while *down-market* α_p^*, β_1^* and β_2^* should remain constant. However, what the model actually estimates is that α_p^* increases by approximately half of the 50bps. In addition, because of the single intercept joining both slopes, β_1^* is effectively 'pulled up' slightly and β_2^* is 'pulled down'. The net effect is that one now infers *greater*

(1984) actually noted a substantial negative (positive) correlation between α_p^* and β_2^* (β_1^*), which caused him to question the validity of the specifications used in Equation 2.6. By imputing a forced intercept and thereby inducing the correlation between terms, the model is biased to accepting the premise that managers are either good timers or good selectors but not both. It is thus no coincidence, neither should it be surprising, that this has been the general finding of most prior mutual fund performance studies (Ippolito (1993)).

Grinblatt and Titman (1989) and Glosten and Jagannathan (1994) show that time-variation in true alpha and beta values can lead to biased alpha estimates based on unconditional OLS. In fact, Glosten and Jagannathan (1994) briefly mention a scenario in which totally unskilled managers are able to show a positive selection and negative timing ability by using the HM model. Thus, while the idea that the HM model is somewhat limited and that a manager's stock selection and market timing abilities vary over time is not new, a statistical exposition quantifying the consequences thereof for the HM model has, according to the author's best knowledge, never been undertaken.

2.2.3 *Recent Literature on Econometric Models of State-Dependent Fund Performance*

Although this chapter focuses on the statistical properties of the HM model, a brief introduction to recent time-varying, also referred to as regime-switching, parameter performance attribution models is given for completeness.

Moskowitz (2000), Kosowski (2006) and Staal (2006) consider regime-switching performance models in which performance for market recessionary and expansionary periods are measured separately. Using a series of Markov-switching models, they conclude that actively-managed funds perform poorly during market expansions on a risk-adjusted basis, but contrastingly, perform abnormally well during recessionary periods.[5] Kosowski's (2006) study is extremely comprehensive in the

stock selection but *lower* market timing. Thus stock selection and market timing ability seem to be negatively correlated.

5 Their proposed models are slightly different to the traditional TM and HM conditional timing models in that the market regimes are chosen to be recessionary and expansionary periods as determined by the data.

data period analysed, the number mutual fund research issues addressed and the extent of the literature reviewed. In particular, Kosowski (2006) gives several compelling reasons as to why time variations should be present in a manager's market timing and stock selection abilities. These include time-varying underlying stock-picking abilities, time-varying costs (including transaction, agency and liquidity costs) and time-varying risk measures. Note that the empirical studies of regime-switching performance models given by Moskowitz (2000), Kosowski (2006) and Staal (2006) should not be referred to market timing models. While the above models control for variations in risk exposure by including regime-switching betas, market timing essentially alters a fund's target risk levels over time in the hopes of profiting from different market regimes. Although subtle, this difference is a consequence of the disparities in market and fund manager assumptions underlying each model-type respectively.

Glode (2010) provides another example of a state/regime-dependent model. The model is built upon a partial equilibrium model of the optimal, state-dependent active management policy of a skilled fund manager. Glode (2010) motivates that skilled fund managers will optimally focus on generating positive risk-adjusted returns during poor market states when faced with rational investors. The negative risk-adjusted performance found above is also confirmed theoretically and empirically by Glode (2010).

While regime-switching performance models are slowly gaining prevalence, one must remember that they are different to performance attribution models motivated by a market-timing argument. Although the models highlighted above do address some of the issues affecting the HM model, this is done at the cost of mathematical tractability and uses a very different set of market and manager assumptions. While it would be interesting to compare the performance of the HM timing model - or a more suitable extension thereof - to that of the above regime-switching models, it is, for now, left as a future exercise.

2.3 TESTING THE HM MODEL

We initially generate monthly fund return data (via simulation) with pre-specified timing and selection characteristics. In order to do so, one assumes values for the up- and down-side alphas, α_1 and α_2, the up- and down-side betas, β_1 and β_2, and the standard deviation of the error terms, σ. A fund is then created for the domain $x_i \in [-0.2, 0.2]$ of hypothetical market returns, with sequential fund return data created using the model:

$$y_i = \begin{cases} \alpha_1 + \beta_1 x_i + \epsilon_i, & x_i \in [-0.2, 0] \\ \alpha_2 + \beta_2 x_i + \epsilon_i, & x_i \in [0, 0.2]\,, \end{cases} \tag{2.8}$$

where $\epsilon_i \overset{iid}{\sim} \mathcal{N}(0, \sigma^2)$ are randomly generated for each x_i.[6] The range of monthly market returns for the American (US) Standard & Poor's Composite 500 Index (S&P 500) for the period July 1984 - Nov 2010 is $[-0.2176, 0.1694]$, centred at approximately 0.008 with a standard deviation of 0.0450.[7] These statistics are generally characteristic of most first world markets. The hypothetical market returns range thus reflects empirical data.

2.3.1 *A Benchmark for Testing the HM Model: Analysis of Covariance*

Analysis of Covariance (ANCOVA) is a statistical technique that combines OLS regression and Analysis of Variance (ANOVA) (Rutherford (2001)). While the dependent variables still constitute the data and the model does include the experimental conditions, as in ANOVA, the main difference between the two techniques is that an ANCOVA also includes one or more quantitative predictor variables. Known as covariates, these predictor variables account for sources of

6 Although empirical return data has been shown to violate the assumption of normality, the reader is reminded that one is not testing empirical selection and timing skills, but rather the HM model's statistical efficacy. Thus, use of normally-distributed errors should actually favour the HM model as it complies with strict OLS regression assumptions. We can thus consider the normal distribution as a 'null' case upon which further research can be be built.

7 The domain quoted for the S&P 500 includes outlying returns. When these are removed, one notes an index return range of $[-0.1303, 0.1318]$.

variance of the dependent variable which are not controlled for by the initial experimental conditions. Thus, ANCOVA only determines whether the means from the dependent variable scores are significantly different after controlling for the co-variation between the dependent variable and the covariates.

Given that ANCOVA is a combination of regression and ANOVA, the standard assumptions of independent and identical (i.i.d.) observations within each group are made. In addition, one assumes that each of the underlying population groups share the same variance. Finally, another common assumption is to assume that the data are normally distributed - although this assumption is not critical. ANCOVA thus shares very similar assumptions to the HM model. This ensures that both models will produce results under very similar, assumed market structures and therefore, should provide comparable parameter output. It is proposed that an ANCOVA specification, where the population groups are taken as the up- and down-market data series, should be able to capture any difference in systematic risk levels and intercepts for each respective group.

A benchmark for testing the HM model is thus conveniently given by an ANCOVA model.In a statistical sense, excess market return data is split into two separate groups: up- and down-markets, denoted as subscripts '1' and '2' respectively in Equations 2.7 and 2.8. An ANCOVA is used to test whether there are significant differences between the slopes for each respective group, namely Slp_1 and Slp_2. The slopes estimated represent the target systematic risk levels outlined within the HM model. Independence between coefficients is achieved by fitting separate lines of best fit to each of the groups, as shown in Figure 1B. The ANCOVA also tests whether there is a significant difference between the intercepts of the lines of best fit for the two data groups, given as Int_1 and Int_2[8] Algebraically, for the given market data group definitions, the ANCOVA is testing the hypothesis:

$$\begin{aligned} H_0 &: (\mathrm{Int}_1 = \mathrm{Int}_2) \wedge (\mathrm{Slp}_1 = \mathrm{Slp}_2) \\ H_1 &: (\mathrm{Int}_1 \neq \mathrm{Int}_2) \vee (\mathrm{Slp}_1 \neq \mathrm{Slp}_2) \end{aligned} \tag{2.9}$$

8 Using the notation defined, the expected intercept is given as $\mathbb{E}[\mathrm{Int}] = (\mathrm{Int}_1 + \mathrm{Int}_2)/2$. By mathematical construction of the HM and Ancova models, we find that $\alpha_p^* \approx \mathbb{E}[\mathrm{Int}]$.

Table 1: Fund variable scenarios used to test ANCOVA benchmark validity.

	Fund Variables						Fund Variables				
	α_1	α_2	β_1	β_2	TE		α_1	α_2	β_1	β_2	TE
'Reality'	-0.05	0.05	0.7	1.3	0.065						
Scen 1	-0.05	0.05	0.7	1.3	0.2	**Scen 18**	-0.05	0.05	1	1.1	0.2
Scen 2	-0.05	0.05	0.7	1.3	0	**Scen 19**	-0.05	0.05	1.3	0.7	0.065
Scen 3	0	0	0.7	1.3	0.065	**Scen 20**	-0.05	0.05	1.3	0.7	0.2
Scen 4	0	0	0.7	1.3	0.2	**Scen 21**	-0.05	0.05	0	2	0.065
Scen 5	-0.5	0.5	0.7	1.3	0.065	**Scen 22**	-0.05	0.05	0	2	0.2
Scen 6	-0.5	0.5	0.7	1.3	0.2	**Scen 23**	0	0	1	1	0.065
Scen 7	0.05	-0.05	0.7	1.3	0.065	**Scen 24**	0	0	1	1	0.2
Scen 8	0.05	-0.05	0.7	1.3	0.2	**Scen 25**	-0.5	0.5	0	2	0.065
Scen 9	0.5	-0.5	0.7	1.3	0.065	**Scen 26**	-0.5	0.5	0	2	0.2
Scen 10	0.5	-0.5	0.7	1.3	0.2	**Scen 27**	0.5	-0.5	2	0	0.065
Scen 11	0	0.01	0.7	1.3	0.065	**Scen 28**	0.5	-0.5	0	2	0.2
Scen 12	-0.01	0.01	0.7	1.3	0.2	**Scen 29**	-0.01	0.01	0.9	1.1	0.065
Scen 13	-0.05	0.05	1	1	0.065	**Scen 30**	-0.01	0.01	0.9	1.1	0.2
Scen 14	-0.05	0.05	1	1	0.2	**Scen 31**	0.01	-0.01	1.1	0.9	0.065
Scen 15	-0.05	0.05	0	2	0.065	**Scen 32**	0.01	-0.01	1.1	0.9	0.2
Scen 16	-0.05	0.05	0	2	0.2	**Scen 33**	0.05	-0.05	1.3	0.7	0.065
Scen 17	-0.05	0.05	0.9	1.1	0.065	**Scen 34**	0.05	-0.05	1.3	0.7	0.2

Note. The subscripts "1" and "2" correspond to down- and up-market fund variables respectively. The annualised fund tracking error is given as TE. 'Reality' refers to the case of a good market timer and stock selector in both up- and down-markets.

Testing ANCOVA under Fund Variable Scenarios

In order to test whether an ANCOVA specification can be used as a benchmark for testing the HM model, a wide range of stressed fund variable inputs are considered within a manufactured fund return universe. The reasoning behind this process is that if the ANCOVA can correctly quantify these input parameters over the entire range of scenarios, then the ANCOVA should hold for any subset of the given scenario range. Table 1 gives the scenario list used for this purpose. Numerous funds were then generated from each of these scenarios and an ANCOVA was run on each fund within the scenario to determine whether the variable estimates obtained were statistically indistinguishable from the given variable values. The ANCOVA was specified on two 'groups' of data: up-market (positive x-values) and down-market (negative x-values). Thus, in terms of testing, we are checking

Table 2: Summary of number of incorrect variable estimates given by ANCOVA Specification.

Average % Incorrect	α_1	α_2	β_1	β_2
10 000 Funds per Scenario	5.05	5.07	5.07	5.14

Max/Min Incorrect for 10 000 Fund Test		
5% = 500	Min	Max
α_1	465	566
α_2	476	569
β_1	460	537
β_2	482	542

Note. The subscripts "1" and "2" correspond to down- and up-market fund variables respectively. The table shows the percentage error found when using an ANCOVA specification to classify all variables in the fund space $\{\alpha_1.\alpha_2, \beta_1, \beta_2\}$.

to see whether the two intercept coefficients match the two known alphas and likewise whether the two slope coefficients match the known betas for up- and down-markets. Under a 10 000 fund-per-scenario trial, the ANCOVA specification results, taken at the 95% confidence level, are summarized in Table 2. Given that a 95% confidence level is being used, the ANCOVA should incorrectly estimate the fund parameters 5% of the time. This is confirmed in Table 2. Therefore, the ANCOVA specification can be said to be a suitable benchmark for testing the HM Model.

2.3.2 *Creating the Hypothesis Test Tree*

The hypothesis test of the the HM model is not straightforward. The HM beta estimates need to be compared to the slope estimates from the ANCOVA to test for the HM model's timing efficacy and the HM alpha estimate is compared to the 2 intercept coefficients to test for the model's selection efficacy in both up- and down-markets. Another level of complexity is added by then dividing 'model efficacy' for each effect into the ability to correctly detect the effect and subsequent to correctly quantify that effect. Therefore, In order to capture all the possible permutations, 6 binary variables are created: Timing detection, Timing quantification, Down-

Selection detection, Down-Selection quantification, Up-Selection detection and Up-Selection quantification. Letting "1" represent reliable interpretation by the HM model of the binary variable and "0" the incorrect interpretation, the complete hypothesis test is shown in Table 3. The reason for having both Up- and Down-Selection variables is because many managers market themselves as being able to add different levels of value in positive markets and negative markets. Although the full binary table of 6 variables would imply 64 different permutations within the tree, this is decreased to 27 because of the fact that quantification of any variable depends on the correct initial detection of the variable. Constructing the hypothesis test in such a way not only ensures that the model's mathematical validity is tested but, perhaps more importantly, it gives users a clear and detailed understanding of the limitations of the model's predictive power.

Given that certain hypothesis levels are much more prevalent within the testing results displayed in Sections 2.4 and 2.5, it is appropriate to discuss those particular levels further. Firstly, let us consider the 'null' hypothesis. H_0 states that the HM model is correctly able to detect and quantify timing, up-selection and down-selection. That is, the HM model perfectly captures the manager's complete skill set. Moving downwards in Table 3, the HM model becomes increasingly inaccurate. The H_{5i} levels state that the HM model is correctly able to detect timing and to correctly detect some combination of up- and down-selection, but is unable to correctly quantify these skills. Hypothesis level H_7 goes several steps further by stating that the HM model is only able to correctly detect timing ability and nothing else. Finally, H_9 represents the most extreme 'alternative' hypothesis, stating that the HM model is incorrect in all attempts at correctly detecting and quantifying timing, up-, and down-selection.

2.4 THE HM MODEL UNDER SIMULATED DATA

The fund space is defined by the variable set $\{\alpha_1, \alpha_2, \beta_1, \beta_2, \sigma\}$. We focus our attention on finding the specific ranges for the respective differences in up- and down-side alpha and beta values over which the HM model accurately detects

Table 3: Hypothesis test tree of the HM model in tabular form.

	Timing		Selection				
			Down Market		Up Market		
Hyp. Level	Detect	Quantify	Detect	Quantify	Detect	Quantify	Explanation
H_0	1	1	1	1	1	1	All Correct
H_{1a}	1	1	1	1	1	0	**Timing:** Detect and Quantification
H_{1b}	1	1	1	0	1	1	**Selection:** Detection & Quant.
H_{1c}	1	1	1	1	0	0	Variations
H_{1d}	1	1	0	0	1	1	
H_{2a}	1	1	1	0	1	0	**Timing:** Detection & Quant.
H_{2b}	1	1	1	0	0	0	**Selection:** Detection Variations
H_{2c}	1	1	0	0	1	0	only
H_{3a}	1	0	1	1	1	1	**Timing:** Detection only
H_{3b}	1	0	1	1	1	0	**Selection:** Detection & Quant.
H_{3c}	1	0	1	0	1	1	Variations
H_{3d}	1	0	1	1	0	0	
H_{3e}	1	0	0	0	1	1	
H_4	1	1	0	0	0	0	**Timing:** Detection & Quant. only
H_{5a}	1	0	1	0	1	0	**Timing:** Detection only
H_{5b}	1	0	1	0	0	0	**Selection:** Detection Variations
H_{5c}	1	0	0	0	1	0	only
H_{6a}	0	0	1	1	1	1	**Selection:** Detection & Quant.
H_{6b}	0	0	1	1	1	0	Variations only
H_{6c}	0	0	1	0	1	1	
H_{6d}	0	0	1	1	0	0	
H_{6e}	0	0	0	0	1	1	
H_7	1	0	0	0	0	0	**Timing:** Detection only
H_{8a}	0	0	1	0	1	0	**Selection:** Detection Variations
H_{8b}	0	0	1	0	0	0	only
H_{8c}	0	0	0	0	1	0	
H_9	0	0	0	0	0	0	None Correct

Note. The hypothesis levels are given in descending order according to the number of binary effect variables that the HM model correctly estimates, when tested against the estimates obtained from the ANCOVA. The binary variables are split into Timing, Up-Market Selection and Down-Market Selection, with each category further split into a variable for Detection and Quantification of the respective effect. A "1" represents correct interpretation by the HM model for that binary variable, with "0" representing incorrect interpretation. The numeric values for the hypothesis levels correspond to the level of accuracy by the HM model. For example, H_{5i} represents the levels for which the HM model is only able to accurately detect timing and some combination of up- and down-market selection. Abbreviated explanations for each numerical level are given in the right-hand column.

Table 4: Null case percentage hypothesis levels

Hypothesis Levels*	H_0	H_{2b}	H_{2c}	H_{5c}	H_7	H_{8c}
Percentage per level	**99.3**	0.2	0.2	0.1	0.1	0.1

*Hypothesis levels with no funds are omitted.

and quantifies a manager's timing and selection skills. Once this is known, we can then deduce when the model shows inaccuracies regarding these two manager skills. We also look at the effect of fund tracking error on the accuracy of model timing and selection estimation.

The solution to finding the ranges for Difference-in-Alphas, $|\alpha_1 - \alpha_2|$, and Difference-in-Betas, $|\beta_1 - \beta_2|$ - hereafter referred to as DIA and DIB respectively - over which the HM model is statistically sound starts by looking at a very simple scenario:

$$\begin{aligned} \alpha_1 &= \alpha_2 = 0 \\ \beta_1 &= \beta_2 = 1 \\ \sigma_i &\in [0.014, 0.02]\,, \end{aligned}$$

where σ_i represents the monthly fund error standard deviation, calculated from an annual tracking error range of $[0.05, 0.07]$. The tracking error range is chosen to represent a characteristic long-only active space as closely as possible. This market scenario is taken as the null case. Using 1000 generated funds, the HM model accurately detects and quantifies timing and selection for this simple scenario at a 95 per cent level. The hypothesis level percentages are summarized in Table 4.

2.4.1 *DIA Range*

The total DIA values tested, *ceterus paribus*, range from 0 to 0.4 with sequential increments of 0.002 over the range $[0, 0.08]$ and increments of 0.008 over the range $[0.08, 0.4]$. A total of 161 scenarios span the complete DIA testing range. The

scenarios are split evenly between the cases for which $\alpha_1 < \alpha_2$ and $\alpha_1 > \alpha_2$.[9] Table 5 shows the hypothesis level percentages for extracts of the DIA range, with Figure 2 displaying the equivalent graphically for the entire tested range. There are a number of points to be raised concerning these results:

- The breakdown DIA value using a 95 per cent cutoff level is 0.008. This is very low and emphasises the problem of naively using this model without understanding its limitations.
- The majority shifts swiftly from H_0 to H_7 (correct timing detection only), which shows an initial 95 per cent majority at a very moderate DIA value of 0.04. This trend is what would be expected from increasing DIA values: as the alphas are stretched further apart, first quantification and then detection of stock selection becomes impossible due to the model's single alpha estimate.
- There is a slight secondary effect on the efficacy of the model's timing estimates as the estimated betas are now skewed due to the necessity of the single alpha coefficient. By mathematical construction of the HM regression and ANCOVA models, we find that $\alpha_p^* \approx \mathbb{E}[Int]$, meaning that α_p^* will also approximately equal the average value of the defined up- and down-market alphas. This leads to $\alpha_1 < \alpha_p^* < \alpha_2$ for $\alpha_1 < \alpha_2$, and $\alpha_2 < \alpha_p^* < \alpha_1$ for $\alpha_1 > \alpha_2$. As a result, β_1 and β_2 are both skewed positively or negatively for the respective (α_1, α_2) cases, which renders correct timing detection and subsequent timing quantification increasingly difficult. This is portrayed by the increasing percentage within H_9, which stabilises around 19 per cent, given our suite of input parameters.
- For a certain small range of DIA values, $[0.008, 0.024]$, the error term about the defined up- and down-lines creates a similar amount of noise as to what would be expected from the different alpha values. Thus, the correct DIA values are masked from the HM model by the noise process. This causes the model to be less discerning in both detection and quantification of timing and selection. One could then infer that the greater the fund tracking error,

9 All results displayed correspond to trials with $\alpha_1 < \alpha_2$ only. However, all findings hold equally well for trials with $\alpha_1 > \alpha_2$

Table 5: DIA range percentage hypothesis levels

α_1	α_2	DIA	H_0	H_{2b}	H_{2c}	H_4	H_{5a}	H_{5b}	H_{5c}	H_7	H_{8a}	H_{8b}	H_{8c}	H_9
0	0	0	99.3	0.2	0.2	0	0	0	0.1	0.1	0	0	0.1	0
-0.001	0.001	0.002	98.8	1.3	1.0	0.6	0	0.2	0.1	0.6	0	0	0	0.1
-0.002	0.002	0.004	96.1	1.2	0.8	0.2	0	0.1	0.4	0.5	0	0	0	0.1
-0.003	0.003	0.006	**96.7**	4.9	4.7	1.7	0	1.8	2.1	7.1	0	0.5	0.1	0.4
-0.004	0.004	0.008	**76.7**	4.9	4.2	1.4	0	0.9	1.1	7.1	0	0.5	0.2	0.2
-0.005	0.005	0.01	79.5	8.6	8.0	3.0	0.1	4.1	4.3	15.1	0	0.8	0.8	1.4
-0.006	0.006	0.012	53.8	6.7	6.2	2.8	0.1	4.3	5.6	15.9	0	1.0	0.6	1.2
-0.007	0.007	0.014	55.6	5.2	4.8	2.9	0.3	7.2	8.0	49.7	0	1.5	1.2	3.4
-0.008	0.008	0.016	15.8	6.7	6.2	2.7	0.1	7.4	6.8	47.3	0	1.0	1.1	4.4
-0.009	0.009	0.018	16.3	2.0	1.8	.8	1.3	5.0	6.0	72.2	0.1	1.0	0.8	4.9
-0.01	0.01	0.02	4.1	4.7	5.1	1.6	0.5	7.1	6.7	59.6	0	1.1	1.1	2.9
-0.011	0.011	0.022	9.6	0.1	0.1	0	3.0	1.2	0.5	87.2	0	0.3	0	7.5
-0.012	0.012	0.024	0	0.1	0.1	0	3.0	1.4	1.0	87.7	0.1	0.1	0.2	6.3
-0.013	0.012	0.026	0.6	0.5	0.8	0.2	0.8	32	3.7	82.9	0	0.4	0.5	6.4
-0.014	0.014	0.028	0	0	0	0	3.4	0.9	0.6	88.8	0.2	0	0	6.1
-0.015	0.015	0.03	0.1	0.3	0.2	0.2	2.3	1.0	2.1	87.3	0	0.3	0.3	5.9
-0.016	0.016	0.032	0	0	0	0	3.2	0	0	93.4	0.1	0	0	3.3
-0.017	0.017	0.034	0	0	0	0	2.4	0.1	0.1	91.0	0.1	0.1	0	6.2
-0.018	0.018	0.036	0	0	0	0	3.0	0	0	93.5	0.1	0	0	3.4
-0.019	0.019	0.038	0	0	0	0	3.0	0.3	0	90.8	0.4	0	0	5.5
-0.02	0.02	0.04	0	0	0	0	1.8	0	0	**96.3**	0	0	0	1.9
⋮			⋮					⋮						⋮
-0.18	0.18	0.36	0	0	0	0	0	0	0	80.9	0	0	0	19.1
-0.184	0.184	0.368	0	0	0	0	0	0	0	80.4	0	0	0	19.6
-0.188	0.188	0.376	0	0	0	0	0	0	0	81.7	0	0	0	18.3
-0.192	0.192	0.384	0	0	0	0	0	0	0	82.1	0	0	0	17.9
-0.196	0.196	0.392	0	0	0	0	0	0	0	82.6	0	0	0	17.4
-0.2	0.2	0.4	0	0	0	0	0	0	0	81.0	0	0	0	19.0

Note. The alpha values, α_1 and α_2, represent the down- and up-markets respectively, with the absolute difference in alphas given as $\mathrm{DIA} = |\alpha_1 - \alpha_2|$.The HM model majority percentage shifts from H_0 - that is, breaks down - when $\mathrm{DIA} = 0.008$. The percentage moves swiftly to hypothesis level H_7 and reaches a 95% majority when $\mathrm{DIA} = 0.04$. Hypothesis levels with no funds are omitted.

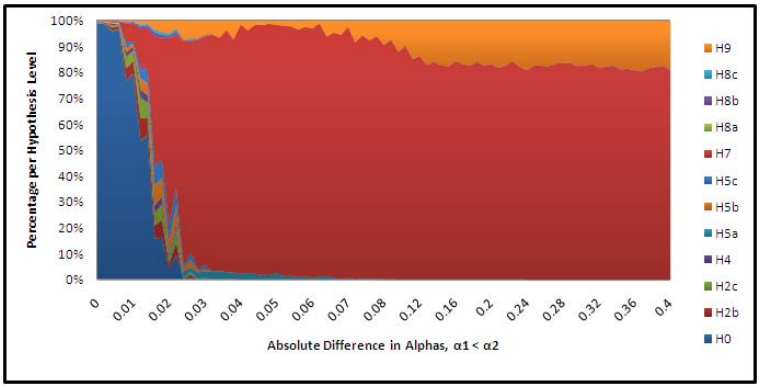


Figure 2: Hypothesis level percentages over DIA range, $\alpha_1 < \alpha_2$.

Note. The graph displays the movement in hypothesis levels for increasing DIA values. The majority percentage shifts from H_0 (given in blue) around DIA values of 0.008 and moves steadily toward H_7 (given in red).

the more accurate the HM model estimates should be. Section 2.4.3 discusses this phenomenon further.

2.4.2 *DIB Range*

The total DIB range tested, ceteris paribus, is $[0, 1.8]$ for both $\beta_1 < \beta_2$ and $\beta_1 > \beta_2$. Figure 3 gives a graphical synopsis for $\beta_1 < \beta_2$. It shows that if $\alpha_1 = \alpha_2$, then irrespective of the DIB value, the HM model is able to correctly detect and quantify both timing and selection on 99 per cent of all occasions. Allowing for random error effects, no matter what the DIB value, the HM model is able to perform accurately. This finding is in stark contrast to those regarding DIA values.

2.4.3 *Tracking Error Effects*

The HM model is tested under the complete DIA range (with 250 simulated funds per DIA value) using a number of different tracking error (TE) ranges. Table 6 gives a summary of the different fund tracking errors applied and their initial

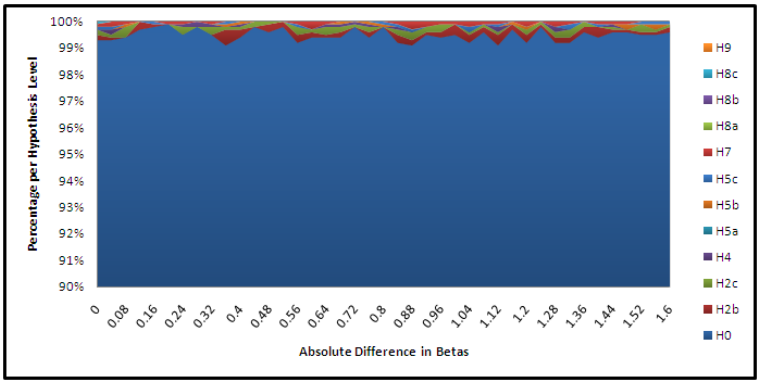


Figure 3: Hypothesis level percentages over DIB range, $\beta_1 < \beta_2$.

Note. The graph displays the movement in hypothesis levels for increasing DIB values. The y-axis is rescaled to $90 - 100\%$. The majority percentage resides within H_0 (given in blue) for the entire DIB range.

Table 6: Summary of the initial DIA break-down value for fund tracking error ranges.

Monthly Error σ-Ranges											
Start	0	**0.014**	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.10
End	0.014	**0.02**	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.10	0.11
Annual Fund TE-Ranges*											
Start	0.00	**0.05**	0.07	0.10	0.14	0.17	0.21	0.24	0.28	0.31	0.35
End	0.05	**0.07**	0.10	0.14	0.17	0.21	0.24	0.28	0.31	0.35	0.38
Break-Down DIA	0.004	**0.008**	0.01	0.014	0.016	0.018	0.022	0.028	0.026	0.026	0.03

*TE = Tracking Error.

Note. Bold values denote the null sigma range used in all previous sections. The first two rows give the start- and end-points for the σ ranges, with the bottom rows giving the corresponding fund tracking error ranges. The final row displays the DIA value at which the HM model breaks down (at a 95% testing level). There is a noticeable upward trend in DIA values over increasing tracking error ranges.

DIA breakdown values which have less than 95 per cent within H_0.[10] From Table 6, it appears that higher fund tracking errors lead to higher HM model DIA breakdown values.[11] Although this finding does not hold for one sigma range, $[0.07, 0.08]$, this can be seen merely as a function of the small number of funds tested per scenario. (For a more detailed discussion, see Appendix A.2). Figure 4 displays the percentage within all hypothesis levels over the complete DIA range when the fund tracking error range is $[0.17, 0.24]$. Comparing this to Figure 2, the

10 Although fund tracking error ranges above 0.10-0.14 are unrealistic in the long-only active space, ranges far above this are seen within hedge funds.

11 Thus, for funds which display very high tracking errors and low DIA values, the HM model is accurate.

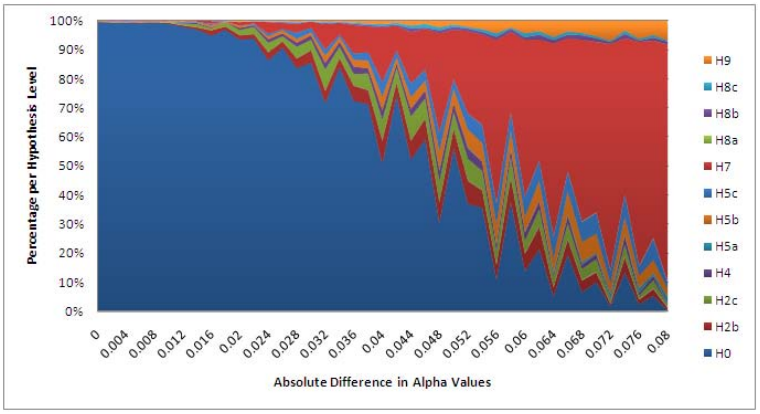


Figure 4: Hypothesis level percentages over DIA range, Tracking Error Range = $[0.17, 0.24]$.

Note. The graph displays the movement in hypothesis levels for increasing DIA values. The majority percentage shifts from H_0 (given in blue) around DIA values of 0.008 and moves steadily toward H_7 (given in red). In contrast to Figure 2, the total hypothesis level pattern is stretched both vertically and horizontally.

same patterns within each hypothesis level are roughly replicated, although they are now stretched both horizontally - over the DIA range - as well as vertically - over hypothesis level percentages. It can thus be inferred that increasing the fund tracking error causes the masking effect to increase across all hypothesis levels fairly equally. Although the HM model essentially becomes less discerning due to the greater noise, the final outcome is that the accuracy of the model increases with fund tracking error.

2.4.4 *Narrow-Range Simulation*

In order to understand the interaction between DIA and DIB, a narrow randomisation over the fund space $\{\alpha_1, \alpha_2, \beta_1, \beta_2, \sigma\}$ is carried out. Hypothetical fund variables are calculated using a Sobol sequence to select random values as inputs for Equation 2.8 from the respective variable ranges given below:

Table 7: Total hypothesis level percentages for narrow randomisation trial

Hyp. Level*	H_0	H_{2a}	H_{2b}	H_{2c}	H_4	H_{5a}	H_{5b}	H_{5c}	H_7	H_{8a}	H_{8b}	H_{8c}	H_9
% per Level	**12.69**	1.73	0.15	0.14	0.01	**55.59**	5.49	5.37	**12.45**	3.97	0.21	0.17	2.04

*Hypothesis levels with no funds are omitted.

$$\begin{aligned} \alpha_1, \alpha_2 &\in [-0.1, 0.1] \\ \beta_1, \beta_2 &\in [0.8, 1.2] \\ \sigma &\in [0.014, 0.02] \end{aligned}$$

A random trial is conducted by using N rows from a scrambled $(N \times 5)$ Sobol matrix. The simulation of up- and down-market alphas and betas is uniformly random to ensure adequate coverage of all reasonable real-world manager skill scenarios. The shape of the parameter distributions is specifically not based on real data as the main purpose of this exercise is to establish whether similar break point conditions as found previously are still evident when parameter interaction is permitted. In fact, the inclusion of real world parameter information could actually mask the statistical limitations of the HM model. Each row within the matrix represents a scenario, with a user-defined number of funds per random scenario. This is done so as to account for any error term effects within each scenario. The summary hypothesis level results from a $(500, 250)$ trial are given in Table 7.[12] Interestingly, the majority percentage now lies in hypothesis level H_{5a} rather than H_7 as before. Level H_{5a} states that the HM model is not only able to detect and quantify timing correctly, but can also correctly detect up- and down-side selection. In order to understand why this is the case, we group the data by DIA, DIB and sigma categories.

From Figure 5 one can see that a similar DIA effect to that noted in Section 2.4.1 is still present and that the H_0 percentages follow a very similar trend to that seen for DIA changes in isolation. The percentage hovers around the 95 per cent level

12 We are using the notation: (# Scenarios, # Funds per Scenario)

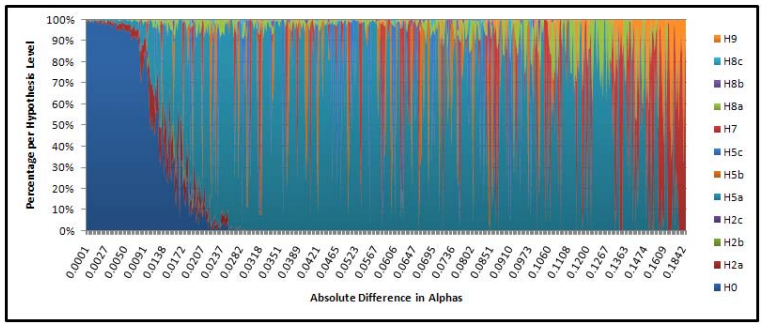


Figure 5: Percentages per hypothesis level within narrow randomisation, categorised by DIA range. ***Note.*** The graph displays the movement in hypothesis levels for increasing DIA values. The majority percentage shifts from H_0 (given in blue) around DIA values of 0.008 and moves steadily toward H_{5a} (given in turquoise) - in contrast to Figure 2 which moves to H_7.

over the DIA range $[0.0061, 0.0076]$, after which the model breaks down and the majority percentage swiftly moves to H_{5a} (as opposed to H_7). This is attributed to the asymmetrical movements now allowed for by α_1 and α_2, which leads to $\alpha_p^* \neq 0$ for most of the funds generated.[13] DIB values, even in conjunction with positive DIA values, do not have a prominent effect on the HM model hypothesis levels. The only time in which DIB values make a minimal difference is in slightly dampening the secondary effect created by positive DIA values; that is, the skew introduced on the model's beta estimates. Similarly, increasing sigma values within the given range show no visible effects on hypothesis level percentages. Thus, for a realistic fund tracking error range, the sigma masking effects can be taken as negligible.

Simulation when DIA is Held at Breakdown Level

We consider model efficacy under different DIB levels when the DIA value is kept constant at the breakdown level of 80bps. In order to achieve this, a narrow randomisation using a fixed DIA value is conducted. The narrow randomisation uses the inputs outlined below:

13 Appendix A.3 gives the algebraic argument behind this inference as well as the reason for the increase in H_7 and H_9 levels for end-of-range DIA values

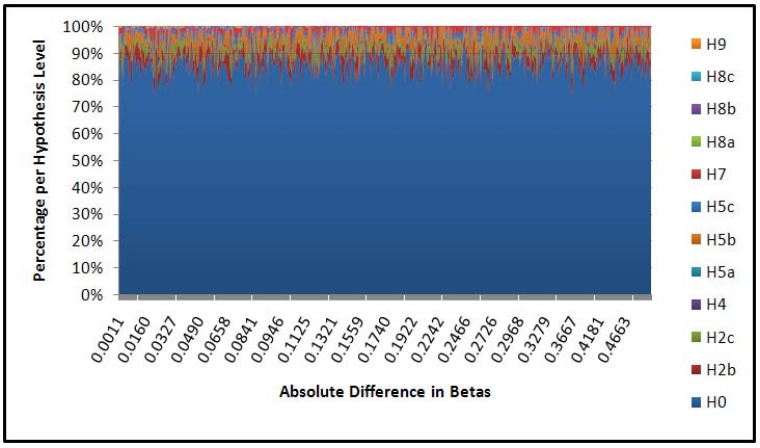


Figure 6: Percentages per hypothesis level when DIA = 0.008, categorised by DIB range.
Note. The HM model maintains a similar degree of accuracy for the entire DIB range tested, shown by the percentage level within H_0 (given in blue).

$$
\begin{aligned}
\alpha_1 &\in [-0.015, 0.015] \\
\alpha_2 &= \begin{cases} \alpha_1 + 0.008 & \text{if } \alpha_1 \leqslant 0 \\ \alpha_1 - 0.008 & \text{if } \alpha_1 > 0 \end{cases} \\
(\beta_1, \beta_2) &\in [0.7, 1.3] \\
\sigma &\in [0.014, 0.02]
\end{aligned}
$$

Figure 6 displays the HM model hypothesis level percentages for the narrow randomisation from a $(500, 250)$ trial, plotted against the DIB range. The HM model displays near identical hypothesis level percentages across the entire DIB range; see Table 8. The average percentage seen within H_0 is approximately 85 per cent, meaning that the HM model still fails to accurately detect timing and selection. This result again emphasises that DIA values are the true driver behind the HM model's efficacy and that the mathematical construction of the model about a single intercept causes the output to be statistically incorrect for $DIA \geqslant 0.008$.

Table 8: Total hypothesis level percentages when DIA = 0.008 for DIB randomisation

Hyp. Level*	H_0	H_{2a}	H_{2b}	H_{2c}	H_4	H_{5a}	H_{5b}	H_{5c}	H_7	H_{8a}	H_{8b}	H_{8c}	H_9
DIB Rand. %	**85.52**	0.17	4.60	2.40	0.44	0.06	3.49	1.48	1.62	0.00	0.12	0.04	0.05

*Hypothesis levels with no funds are omitted.

2.5 THE HM MODEL UNDER EMPIRICAL DATA

Monthly returns from a selection of 148 US mutual funds are used to test the HM model empirically, with the market portfolio proxied by the S&P 500 and the risk-free rate proxied by the US 3-month treasury bill. The mutual funds chosen are those with a "common stock" investment policy and a "capital appreciation", "growth", "growth and income" or "stock and bond" nominated investment objective or style. In addition, mutual funds that are marketed as having a moderate or aggressive asset allocation objective are also included. All funds in the sample have a history of at least five years as of November 2010, with an average (maximum) fund history of roughly 9.5 (18.8) years. All sector, balanced, index and enhanced index funds are excluded from the sample. The reason for this choice of fund data is due to the nature of the sampled funds' target asset allocation strategies. The sampled funds should show some combination of market timing and security selection effects - a useful set of characteristics against which the correctness of the HM model can be benchmarked and tested. In addition, empirical fund return data need not be normally distributed (and, usually, are not). Therefore, empirical testing provides a more stringent test for the HM model, especially given that it fails rather easily for the 'simplest' normal-residual case.[14] The empirical testing of the HM model using this US mutual fund data set corroborates the results found under simulation. Table 9 shows the 'true' fund variables from the ANCOVA specification, while Table 10 provides the HM model's hypothesis level percentages over all funds.

14 A potential problem is that if the data are heteroscedastic across market regimes, then the ANCOVA assumption of homogeneous population group variances is violated. This could lead to incorrect parameter estimates empirically. However, given that the focus of the research here is simulated testing of the HM model with empirical data used only as corroboration, this point is presently left as a caveat. The inclusion of regime-specific volatilities, as well as the choice of market distribution, is left for future research.

Table 9: Summary of true US mutual fund variables.

Fund Variables†	Minimum	Average	Maximum	Std Error
α_1	-0.0095	0.0072	0.0212	0.0067
α_2	-0.0070	0.0059	0.0213	0.0052
DIA	0.0001	0.002	0.0214	0.0047
β_1	-0.1270	0.5213	1.1983	0.2896
β_2	-0.7035	0.1236	0.6864	0.1627
DIB	0.0083	0.4027	0.8574	0.1962
Tracking Error	0.0310	0.0510	0.0851	0.0098

Note. Monthly data from 148 US mutual funds with history greater than 5 years as of Nov 2010 is used to calculate all table values. Funds included in the sample are those with a "common stock" investment policy and a "capital appreciation", "growth", "growth and income" or "stock and bond" investment objective. In addition, mutual funds that are marketed as having a moderate or aggressive asset allocation objective are also included.

Unsurprisingly, it would seem that the HM model is unable to correctly detect and quantify both timing and selection for 111 out of the 148 funds tested. For the majority of funds (62 per cent of the sample) the HM model is only able to correctly detect timing and detect some form of selection in up- and down-markets, shown by the percentage levels in the H_{5i} hypothesis levels. This is in line with earlier findings given that Table 9 shows a total DIA range of $[0.000, 0.021]$ and an average of 0.0062. From the results given in Section 2.4.4, the HM model verges on breakdown over the range $[0.0061, 0.0076]$ for funds within a normal tracking error range of $[0.05, 0.07]$. However, many of the US funds show tracking errors lower than this. As shown in Section 2.4.3, a lower fund tracking error implies a lower DIA breakdown level. For funds with annual tracking errors ranging from $[0, 0.05]$, the DIA breakdown level falls to only 40bps. Using only the 'true' fund variable statistics and the hypothetical fund findings presented previously, one is able to infer when the HM model will be inaccurate. This inference can then be compared to the hypothesis level percentages shown in Table 10.

Of the 148 funds tested, 76 have $DIA > 0.0061$, meaning that the HM model is likely to be inaccurate for those managers' timing and selection skills, given reasonable fund tracking errors.[15] From the remaining funds, 22 have $DIA > 0.004$

15 Using the DIA breakdown levels given in Table 6, those funds with both $DIA > 0.008$ and higher tracking errors display large enough DIA values to infer that the HM model will provide an inaccurate measure of those managers' timing and selection skills.

Table 10: Hypothesis level fund count and percentages for US mutual funds.

Hyp. Level*	H_0	H_{1b}	H_{2a}	H_{2b}	H_{2c}	H_{5a}	H_{5b}	H_{5c}	H_7	H_{8c}
No. Funds per level†	**37**	5	7	1	1	**64**	**14**	**14**	4	1
Percentage per level	**25.00**	3.38	4.73	0.68	0.68	**43.92**	**9.46**	**8.78**	2.70	0.68

*Hypothesis levels with no funds are omitted.
†Total number of funds = 148.

as well as $TE < 0.05$. Using the breakdown DIA values in Table 6 (Section 2.4.3), this again implies that the HM model is likely to be inaccurate with respect to the respective managers' timing and selection abilities. The final 50 funds are those which show fairly low DIA values and relatively normal equity fund tracking errors. However, 8 of these funds are very near to the breakdown levels defined previously, leaving only a total of 42 funds - or 28 per cent of the sample - for which the HM model might be accurate. The fund count given in Table 10 shows that the HM model is only accurate for 37 funds, or 25 per cent of the sample. These results confirm the structural problems intrinsic to the HM model, and further reveal how the previous sections' findings, using simulated data, hold comfortably across empirical settings.

2.6 CONCLUSIONS

In this chapter, we test the HM model's efficacy as a measure of a manager's timing and stock selection skills. The driving force behind percentages per hypothesis level is the DIA value. Given a realistic fund tracking error range and DIB range, the HM model consistently breaks down when DIA values are greater than 80bps per annum. In other words, if fund managers possess a differential ability to stock pick under differing market regimes, to the extent that this ability exceeds 80 bps per annum, the HM model breaks down. As this DIA value increases, the model estimates of timing and selection ability become increasingly erroneous.

In contrast, the model never breaks down when DIB values vary in isolation. In fact, increasing the DIB values concurrently with DIA causes the HM model to be more accurate in detecting timing. This is due to the slight dampening effect

that positive DIB values have on the skewed estimation of the betas, which is introduced by the positive DIA values.

In times of greater fund volatility, the HM model is less discerning and thus actually performs better than for stable periods. This is due to the masking effect that the increased volatility has on the difference in true alpha values, which falsely creates a situation that is closer to the single intercept assumption of the model. This is an interesting finding and one that ought to force researchers to look more carefully at characterising and standardising the volatility backdrop (both longitudinal and cross-sectional) before contrasting different markets, asset classes or managers.

If up- and down-market alpha values move asymmetrically, the HM model shows for the majority of the fund space that it is able to correctly detect both timing and selection. This is in contrast to only correctly detecting timing for symmetric moves.

The US mutual fund sample results are actually slightly worse than those found for the simulated funds. This is mostly likely due to the non-normality of empirical fund return data. Viewed in combination, the results show that the HM model is prone to inaccuracies in detecting and quantifying manager timing and selection skill.

Extensions of this research include creating a variation of the ANCOVA model that takes into account different up- and down-markets risk levels and is tested under a range of different return distributions. Secondly, it could very well be that the Treynor and Mazuy (1966) model suffers from similar problems due to its single alpha estimate and an equivalent analysis of this model may lead to conclusions similar to those found here.

3

UNLOCKING THE ALGEBRA OF CROSS-SECTIONAL VOLATILITY

3.1 INTRODUCTION

Cross-sectional volatility (CSV), also commonly referred to as (return) dispersion, reflects the volatility of market, industry or sector returns over a particular period, rather than over time, as computed by the more conventional longitudinal volatility. Essentially, it measures the opportunity set available to an investor over the specified term. In terms of financial research, CSV is a relatively new concept. Research into this subject only truly started during the late-1990s. Since then, it has become an increasingly active field of literature and application. Both academics and practitioners have realised that quantifying the available opportunity set within a specific basket of assets at any given time can be extremely beneficial in a number of different applications.[1] Letting ${}_{N_t}\chi(r_t)$ represent CSV for time period t, cross-sectional variance - or squared CSV - in an N_t-asset market is defined as:

$$ {}_{N_t}\chi^2(r_t) = \sum_{i=1}^{N_t} w_{it}(r_{it} - r_{mt})^2, \tag{3.1} $$

where w_{it} is the beginning of time period t weight for stock i, r_{it} is the return of stock i over time period t, $r_{mt} \equiv \sum_{i=1}^{N_t} w_{it} r_{it}$ is the weighted market return and N_t is the number of stocks in the market at time t. If one substitutes $w_{it} = 1/N_t$ in Equation 3.1, then this metric is commonly referred to as uniform CSV.[2]

1 To this end, Russell Investments and Parametric Portfolio Associates released a CrossVol™ index in October 2010 that reports the daily and monthly CSV for a wide range of Russell Global indexes.

2 Although CSV strictly refers to cross-sectional *volatility*, many researchers rather provide models and theorems for cross-sectional *variance* due to its much greater mathematical tractability. Thus, when considering the algebraic decomposition of CSV, we will work largely with cross-sectional *variance*, while the simulation analysis, presented later in this chapter, deals directly with cross-sectional *volatility*.

Applications of CSV to date have been in the active management, fund performance attribution and risk management fields, with increasing interest now arising from risk-takers with respect to derivative strategies. In terms of active management, Chadha and Satchell (2008) and Gorman, Sapra and Weigand (2010a, 2010b) argue that CSV is of vital importance in correctly calculating the cross-sectional correlation coefficients proposed in Grinold's (1989) seminal paper on the Fundamental Law of Active Management. In addition, they build on the same framework by incorporating CSV to find the optimal active portfolio weights. In terms of performance attribution, De Silva, Sapra and Thorley (2001) contend that one must always consider the opportunity set within the market when measuring and reporting asset management performance. They propose that both alpha and the information ratio should be standardised according to market CSV level over the reporting period. In the risk management field, there have been several differing streams of research incorporating CSV. Lillo et al (2001) were one of the first to argue for including daily CSV as a supplementary to daily market volatility.[3] A number of researchers build on this work by considering the effects of CSV on the market correlation structure.[4] Another branch of research incorporating CSV within risk management is linking time series and cross-sectional volatility measures. Numerous papers have discussed methods of implementing CSV within time series volatility models.[5]

Despite the volume of interest in the subject, what has not been properly researched is the algebraic properties of CSV itself. In fact, the only paper that obliquely tackles this issue is Hwang and Satchell's (2006) working paper, *Properties of Cross Sectional Volatility*, in which the authors compare the properties of CSV to that of time series volatility. Hwang and Satchel (2006)l also show that CSV can be used as both an explanatory and forecasting variable for time series volatility. Whilst this work stands primarily alone in its efforts to properly understand the

3 Lillo and Mantegna (2000) and Lillo et al (2001) refer to market CSV as *variety*. We find that this name occurs rather infrequently in the literature thereafter.

4 See, for example, Solnik and Roulet (2000), Statman and Scheid (2004), Demirer and Lien (2005), and diBartolomeo and Baig (2006).

5 See, for example, Hwang and Satchell (2005), Ang et al (2006), Connor, Korajczyk and Linton (2006) and Yu and Sharaiha (2007).

dynamics of CSV, its principal focus still remains on using CSV within a time series volatility framework.

Consequently, what has been noted is that much of the existing CSV literature uses several telling simplifications, usually relating to the covariance matrix, the weighting structure and the underlying stock price process. Although these simplifications lead to tractable expressions for realised and expected CSV, this tractability comes at the cost of incorrect inference and, in some cases, actually invalidates the reported expressions entirely. This contribution discusses some of the repercussions of relying upon these oversimplifications and subsequently provides an insight into the fundamental properties of return CSV in a much more general market setting. A comparison is made between the market variables driving CSV under the more relaxed paradigm against those reported when using the common restrictive assumptions. To the author's best knowledge, this type of study has yet to be conducted.

This paper is organised as follows: Section 3.2 summarises the existing measures for realised, expected and approximate CSV under different market frameworks and highlights the most common simplifying assumptions. In Section 3.3, we analyse the algebra of CSV, which leads to a general expression for realised CSV. Assuming fairly general stock price dynamics, we derive an accessible expression for expected CSV. Given a wide range of simulated market conditions, Section 3.4 examines how average asset correlation, average asset volatility and market concentration affect CSV. These results are contrasted to the predictions in previous literature. The effects of the standard deviation of underlying asset volatilities on CSV are also discussed here. Section 3.5 concludes.

3.2 EXISTING EXPRESSIONS FOR REALISED, EXPECTED AND APPROXIMATE CSV

In terms of financial literature, CSV research is still fairly undeveloped. As such, a consensus on the most appropriate measure for CSV has yet to be reached. This section details the most common measures for realised, expected and approximate CSV and some of the key differences between these respective measures. More

importantly though, it also highlights some of the potential drawbacks of certain measures.

Ankrim and Ding (2002) were among the first to provide an expression for expected CSV. Using similar notation as in Equation 3.1, they showed that expected CSV under the probability space $(\Omega_t, \mathcal{F}_t, \mathbb{P})$ could be given as

$$\begin{aligned} \mathbb{E}_{\mathbb{P}}\left[\chi_t^2(r_t)\,|\mathcal{F}_t\right] &= \left[\left(\sum_{i=1}^{N_t} w_{it}\sigma_{it}^2\right) - \sigma_{mt}^2\right] + \left[\sum_{i=1}^{N_t} w_{it}\left(\mu_{it}-\mu_{mt}\right)^2\right] \\ &= \tilde{\sigma}_t^2 - \sigma_{mt}^2 + v_t^2, \end{aligned} \tag{3.2}$$

where μ_{it} and σ_{it} are the mean and volatility of stock i at time t. In this case, $\tilde{\sigma}_t^2 = \sum w_{it}\sigma_{it}^2$ represents the cap-weighted average underlying volatility and $v_t^2 = \sum_{i=1}^{N_t} w_{it}\left(\mu_{it}-\mu_{mt}\right)^2$ the cross-sectional variance for expected stock returns.

Yu and Sharaiha (2007) found a similar expectation when linking time series volatility to its cross-sectional counterpart. Their decomposition is quite robust as it does not depend on any specific asset price process. While Equation 3.2 is theoretically sound, it does not easily allow one to understand the relationship between underlying volatility - or correlation - and CSV. Thus Yu and Sharaiha (2007) proposed an elegant expression for *uniform, instantaneous* CSV based on substituting a simple diffusion process for the time t log-stock price S_{it} into Equation 3.2. That is,

$$\begin{aligned} &dS_{it} = \mu_i dt + \sigma_i dB_i \quad where\, \mathbb{E}\left[dB_i, dB_j\right] = \rho_{ij}dt \;\; \forall i \neq j \\ &\Rightarrow \chi^2_{YS,(t,t+dt)} = \left(\textstyle\sum_{i=1}^{N_t} w_{it}\sigma_{it}^2 - \sigma_{mt}^2\right) dt \end{aligned}, \tag{3.3}$$

where B_i's are Brownian motions. Note that both the drift and variance are time-invariant in the underlying log-price dynamics and also that the log-price process is not the more commonly used Geometric Brownian Motion (GBM). However a time-invariant correlation term is included. Substituting these price dynamics into Equation 3.2 gives the second line in Equation 3.3. Note that the second term from Equation 3.2 falls away with the higher order dt terms and the expectation sign is removed due to the zero-variation in Equation 3.3. Yu and Sharaiha (2007) simplified this expression further by looking at the uniformly-

weighted case and found an elegant approximation for short-term CSV involving only average underlying volatility, $\bar{\sigma}$, and average pairwise correlation, $\bar{\rho}_{ij}$:

$$\begin{aligned} \chi_{(t,t+dt)}(r_{dt}) &\approx \sqrt{\bar{\sigma}^2 - \bar{\sigma}^2_{jk}} \\ &= \bar{\sigma}\sqrt{1-\bar{\rho}}, \end{aligned} \tag{3.4}$$

for N_t sufficiently large and $\bar{\rho} = \bar{\sigma}^2_{jk}/\bar{\sigma}^2$ for $j \neq k$, where $\bar{\sigma}^2_{jk}$ is the cross-sectional, average covariance. What must be remembered is that this expression only holds for instantaneous, uniform CSV under the proposed log-price diffusion process.

Equation 3.4 is prevalent in much of the active management research incorporating CSV. Gorman, Sapra and Weigand (2010a), along with many others, use Equation 3.4 extensively and not only as a short-term CSV measure.[6] By assuming that the asset covariance matrix is a function of only two parameters - average volatility and average correlation - Gorman et al show that Equation 3.4 approximates uniform CSV over any time period. Looking at its mathematical tractability, one can readily understand why this model is heavily favoured. Expressing CSV as a function of two variables fundamental to the active management framework allows one to include CSV intuitively within these results. However, using this approximation may cause the drivers of CSV to be incorrectly identified. Many of the papers using Equation 3.4 infer that CSV increases monotonically with average underlying asset volatility. As Section 3.4 will highlight, this is actually not the case.

An alternative expression for CSV that should also be mentioned is given by Hwang and Satchell (2006). Using a linear multi-factor model to represent stock returns, CSV is expressed as:

$$\sigma^2_{CS,mt} = (N-1)\sigma^2_{m,\varepsilon} + Var^{CS}(\mathbb{E}^{TS}_{t-1}[r_{it}|\mathcal{F}_{t-1}]) + \sum_{k=1}^{K} f^2_{kt} Var^{CS}(\beta_{ikt}), \tag{3.5}$$

6 Similar CSV measures to that given in Equation 3.4 are used by Statman and Scheid (2004) and Solnik and Roulet (2000).

where $\sigma^2_{m,\epsilon} = 1/N \sum \sigma^2_{it}$ is the uniformly-weighted market volatility in a cross-sectional framework, the superscripts X^{CS} and X^{TS} refer to 'cross-sectional' and 'time series' respectively, f_{kt} is the realised value of factor k at time t and β_{ikt} is asset i's coefficient of factor k at time t. Thus, Hwang and Satchell (2006) decompose CSV into three terms: (i) cross-sectional, uniformly-weighted market volatility, (ii) expected return dispersion and (iii) factor coefficient dispersion, weighted by the realised factor values. Equation 3.5 gives one direct insight into the true drivers of CSV by completely capturing its underlying dynamics. However, there are complications with this definition. Equation 3.5 again assumes a uniformly-weighted market and, more importantly, a multi-factor asset return model. Correctly identifying and quantifying these factors in the model is very challenging, thus making practical inference rather difficult. However, by assuming that asset returns are described by the common Capital Asset Pricing Model (CAPM), De Silva, Sapra and Thorley (2001) were able to use Equation 3.5 to find a more tractable form for expected CSV:

$$\mathbb{E}_{\mathbb{P}}\left[\chi^2_t|\mathcal{F}_t\right] = \sigma^2_{\beta t}\left(r_{mt} - r_{ft}\right)^2 + \sigma^2_{\varepsilon t}, \tag{3.6}$$

where $\sigma^2_{\beta t}$ is the cross-sectional variance of stock betas at time t, r_{ft} is the risk-free interest rate at time t and $\sigma^2_{\epsilon t}$ is the variance of the idiosyncratic risk at time t. As with the previous CSV equations, the additional tractability of Equation 3.6 comes at the price of severe assumptions; namely CAPM and a uniformly-weighted market.

Thus, the commonly quoted CSV expressions discussed here all suffer from similar limitations imposed by their underlying simplifications. Although this increases tractability, it does so at the price of realism and, in some cases, correctness.

3.3 UNDERLYING ALGEBRA OF CSV

Let us now consider a more general market setting. Before stating the theorem for squared CSV, one needs to define the concept of *pairwise* CSV. If one considers a very simple two-asset market, the concept of pairwise CSV is fairly intuitive. Using this intuition, one can generalise this concept for the N_t-asset case:

Definition 3.1. For assets i and j and time t, *pairwise squared CSV* is defined as

$$_{i,j}\chi^2_t(r_t) = \left[w_{it}w_{jt}\left(w_{it}+w_{jt}\right)\right]\left(r_{it}-r_{jt}\right)^2. \tag{3.7}$$

Appendix B.1 motivates the choice of definition. Thus, Definition 3.1 shows that pairwise CSV, for any number of stocks within the market, is only dependent on the weights and returns of the two stocks within the measure. Using this result, one can now give a general theorem for realised market CSV:

Theorem 3.1. *Consider a market with N_t stocks. Let r_{it} and w_{it} be the return and weight of stock i respectively for the period t. Squared CSV for the N_t-asset market, $_{N_t}\chi^2_t(r_t)$, is then defined as the sum of pairwise squared CSV and a weighted summation of the product of pairwise asset return differentials. That is,*

$$_{N_t}\chi^2_t(r_t) = \sum_{\substack{i,j=1\\ i\neq j}}^{N_t} {}_{i,j}\chi^2_t(r_t) + 2\sum_{i=1}^{N_t} w_{it}\left[\sum_{\substack{j,k\neq i\\ k>j}}^{N_t} w_{jt}w_{kt}\left(r_{it}-r_{jt}\right)\left(r_{it}-r_{kt}\right)\right]. \tag{3.8}$$

For the special case of uniform CSV, $w_t = 1/N_t$ for all assets.

Proof. See Appendix B.1 □

3.3.1 *Expected CSV*

Section 3.2 outlined the common expectations given for CSV and the inferences drawn from them. However, these expectations were noted to be conditional upon several oversimplifications. This conditioning potentially leads to expectations of limited applicability and, in some cases, correctness. We consider here a less restrictive set of asset price dynamics. Assume that the stock price process, S_{it}, follows a GBM. That is,

$$dS_{it} = S_{it}\left(\mu_{it}dt + \sigma_{it}dB_i\right), \quad \mathbb{E}\left[dB_i, dB_j\right] = \rho_{ij,t}dt \forall\, i \neq j. \tag{3.9}$$

The parameters μ_{it}, σ_{it} and $\rho_{ij,t}$ are either assumed to be constants or deterministic functions that are allowed to change values at the beginning of each period t.[7] This allows for a broad range of price dynamics which are fairly representative of those most commonly assumed in practice. Using Equation 3.9, we obtain the following results.

Theorem 3.2. *Consider a market with N_t assets. Let μ_{it} and σ_{it} represent the drift and volatility of stock i, which follows a GBM process. Under the probability space $(\Omega, \mathcal{F}_t, \mathbb{P})$, the expected squared CSV, conditional on $\mathcal{F}_t$, is written as*

$$\begin{aligned}\mathbb{E}_{\mathbb{P}}\left[{}_{N_t}\chi^2_t\left(r_t\right) \mid \mathcal{F}_t\right] &= \sum_{i=1}^{N_t}\left[\left(\sigma^2_{it} + \mu^2_{it}\right)\left(w_{it} - w^2_{it}\right)\right] \\ &\quad -2\sum_{j=1}^{N_t-1}\sum_{k>j}\left[\left(\rho_{jk,t}\sigma_{jt}\sigma_{kt} + \mu_{jt}\mu_{kt}\right)\left(w_{jt}w_{kt}\right)\right]\end{aligned} \tag{3.10}$$

If $\mu_{it} = a_t$ for all assets, then the $\mathcal{F}_t$-conditional expected squared CSV is given by

$$\mathbb{E}_{\mathbb{P}}\left[{}_{N_t}\chi^2_t\left(r_t\right)|\mathcal{F}_t\right] = \sum_{i=1}^{N_t}\left[\sigma^2_{it}\left(w_{it} - w^2_{it}\right)\right] - 2\sum_{j=1}^{N_t-1}\sum_{k>j}\left[\rho_{jk,t}\sigma_{jt}\sigma_{kt}\left(w_{jt}w_{kt}\right)\right]. \tag{3.11}$$

Proof. See Appendix B.1 □

7 For the drift parameter, a common assumption is that $\mu = r_{ft}$, where r_{ft} is the risk-free return at time t, or that $\mu = 0$. Thus, drift is assumed constant across both time and stocks.

Although not presented here, one finds that, under simulation, the second term in Equation 3.8 is usually quite small in comparison to the first term. A fairly good approximation for N_t-asset CSV is then given entirely by the weighted sum of pairwise CSV. Thus, if one understands the factors that drive pairwise CSV, one immediately has a good understanding of the factors driving N_t-asset CSV. Although it may seem counter-intuitive to deconstruct Equation 3.1 into the more complex Equation 3.8, this decomposition enables one to find an accessible expression for expected CSV without the need for prejudicial simplifications.

Theorem 3.2 states that expected pairwise CSV is a function of several underlying asset variables: namely asset means, volatilities, weights and correlations, as well as the number of underlying assets. There is a large existing body of knowledge on modelling and forecasting these underlying asset parameters. Thus, by using Equation 3.10 in an existing asset modelling framework, one can practically and accurately forecast CSV under a wide range of asset price dynamics.

In comparison to the commonly quoted Equation 3.4, Theorem 3.2 does not give easily discernible relationships between the underlying asset variables and expected CSV. For example, underlying volatility appears in both the positive and negative terms within Equation 3.10. We now explore these relationships further and compare the findings to those given in prior literature.

3.4 VARIABLES DRIVING CSV

In Section 3.3, we showed that expected CSV is a function of several underlying asset variables. Thus one would expect these same variables to drive realised CSV. This is the null hypothesis. In order to test this hypothesis, a broad range of realistic market conditions are simulated under which the resultant CSV is analysed. Table 11 details the chosen market parameters and their possible ranges.

Daily underlying asset returns for the specified T-length period are simulated from independent normal distributions with zero-mean and varying annual stock volatility, σ_i, by using Sobol sequences for the daily underlying return probability

Table 11: List of simulated market variables and allowed ranges

Market Variable	Range
Time Period	$T = 5, 63, 252, 504$ days
Annual Underlying Asset Volatility	$\sigma_i \in [0.05, 0.35]$ - Low-Range
	$\sigma_i \in [0.2, 0.5]$ - Mid-Range
	$\sigma_i \in [0.35, 0.65]$ - High-Range
Average Asset Correlation	$\bar{\rho} = 0.2, 0.3, \ldots, 0.8$
Market Concentration Factor	$C = 0, 1, \ldots, 7$
	$0 =$ uniform weights - $7 \approx$ extreme concentration
Number of Stocks	$N_t = 2, 5, 10, 20, 50, 100$

values.[8] The annual underlying volatility is sampled from a specific uniform range, depending on whether one is considering the 'Low', 'Mid' or 'High' range underlying volatility (defined in Table 11). A positive definite asset correlation matrix is constructed based on a user-defined average pairwise correlation, $\bar{\rho}$, which is allowed to moved in increments of 0.1 and ranges from 0.2 to 0.8. Using the upper triangular matrix obtained from the Cholesky factorisation of the correlation matrix, the independent daily underlying returns are transformed to comply with the defined market correlation structure.

The market weighting structure is defined by the concentration factor, $C = 0, 1, 2, \ldots, 7$, with individual weights being calculated according to an exponential decay function, given in Equation 3.14 (Section 3.4.3). The number of stocks within the market at time t, N_t, is set at 2, 5, 10, 20, 50 or 100. Finally, daily total market returns are calculated from the above parameters and the simulated CSV is measured over four different time periods: namely, 5, 63, 252 and 504 days respectively. The periods are chosen to represent one week, three months, one year and two years respectively.

For each $\bar{\rho}$ level - all else constant - 5000 observations are taken, giving 35000 observations per concentration level. Given that there are eight market concentration levels and that CSV is simulated under four time periods for Low-, Mid- and High-Range stock volatilities respectively, one has a total of 3.36 million

8 Using the quasi-random Sobol sequence, rather than the more common psuedo-random sequence, leads to lower total term return values and thus lower dispersion values than one would commonly expect. While we note this effect for completeness sake, itt does not change the simulation results or inferences in any meaningful way.

observations per stock level analysed. There is a considerable overlap within the resulting CSV findings and thus, for the sake of brevity and clarity, only a subset of results are displayed.

We focus our discussion here on how the mean and standard deviation of asset volatility, $\bar{\sigma}$ and SD (σ), average underlying correlation, $\bar{\rho}$, and market concentration, C, affect CSV. Appendix B.2 provides a detailed discussion of the remaining market variables defined in Table 11.

3.4.1 *Average Underlying Volatility, $\bar{\sigma}$, and its Standard Deviation,* SD (σ_i)

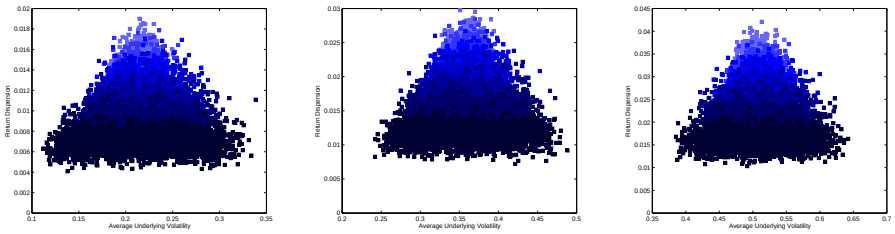
The currently accepted understanding (as per Ankrim and Ding (2002), Hwang and Satchell (2006), Yu and Sharaiha (2007) and Gorman et al (2010a), among others) is that CSV is a monotonically increasing function of the average volatility of the underlying securities. Figure 7 $(a)-(c)$ plots 63-day uniform CSV against average underlying volatility, $\bar{\sigma}$, for the Low-, Mid- and High-Range volatility when $N_t = 100$. Figure 8 again shows CSV against average volatility but now for $N_t = 5, 10, 20, 50$ under Mid-Range input volatility only. Different average correlation values, $\bar{\rho}$, are represented by different colour gradients: $\bar{\rho} = 0.2$ corresponds to the lightest gradient and $\bar{\rho} = 0.8$ is given by the darkest gradient.[9] For all cases, we see a contrary relationship between average volatility and CSV to that predicted within prior literature.

Although Figure 7 reveals that increasing the average underlying volatility increases possible CSV *bounds*, it does not depict a monotonically increasing function. In fact, one finds a very strong concave relationship, even for smaller numbers of underlyings.[10] Thus, it would seem that the relationship between CSV and average underlying volatility has previously only been partially explored.

In order to understand the concavity displayed here, we consider a simple, illustrative example. Assume a uniformly-weighted, two-asset market in which the

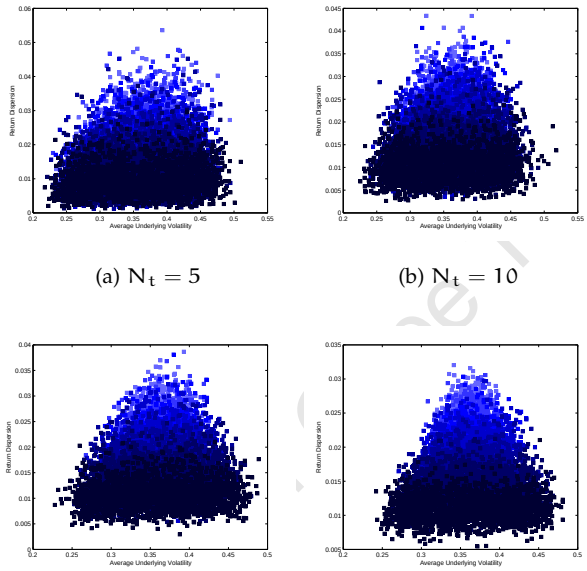
9 For the sake of brevity, all graphs within Section 3.4 are given for $T = 63$. In addition, lower average correlation is always depicted by the lighter colour shades, with the darker shades representing higher average correlation.

10 One would expect more noise from smaller markets due to the greater weight placed upon each asset's return and volatility and thus the higher chance of influential outlying observations.



(a) Low: $\sigma_i \in (0.05, 0.35)$ (b) Mid:, $\sigma_i \in (0.2, 0.5)$ (c) High:$\sigma_i \in (0.35, 0.65)$

Figure 7: Uniform CSV versus Average Volatility for Low-, Mid- and High-Range volatility when $N_t = 100$



(a) $N_t = 5$ (b) $N_t = 10$

(c) $N_t = 20$ (d) $N_t = 50$

Figure 8: Uniform CSV versus Average Volatility for $N_t = 5, 10, 20, 50$ under Mid-Range volatility

asset returns have a zero-mean, Mid-Range volatility and are perfectly correlated. Using Theorem 3.2, expected squared CSV in this case is given by

$$\mathbb{E}_{\mathbb{P}}\left[\chi_t^2(r_t)\right] = \frac{1}{4}(\sigma_1 - \sigma_2)^2, \tag{3.12}$$

and the average asset volatility, $\bar{\sigma}$, is equal to

$$\bar{\sigma} = \frac{1}{2}(\sigma_1 + \sigma_2) \,\forall\, \sigma_1, \sigma_2 \in [0.2, 0.5]. \tag{3.13}$$

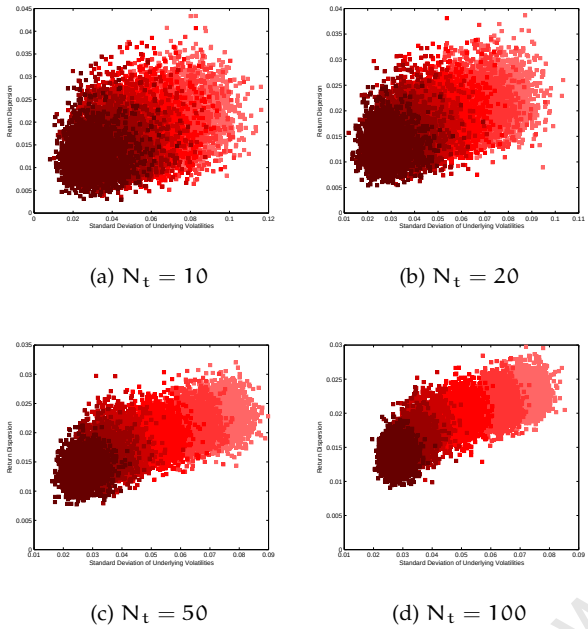
Consider when average volatility equals the mid-point of the input volatility range, that is $\bar{\sigma} = {}^{(0.2+0.5)}/_2 = 0.35$. For this to occur, σ_1 and σ_2 must be equidistant

from $\bar{\sigma}$, with $\sigma_1 \geqslant \sigma_2$ or vice versa. When $\sigma_1 = \sigma_2$, from Equation 3.12 we know expected squared CSV equals zero.[11] However, as the underlying volatilities are moved away from the midpoint, CSV continually increases and finally reaches its maximum when the volatilities are equal to the allowed lower and upper bounds respectively. This implies that there is another volatility-linked variable which monotonically describes CSV: the *standard deviation* of the underlying volatilities, $SD(\sigma_i)$. This is analogous to the volatility of volatility parameter within stochastic volatility models. While this observation is quite obvious in the simple scenario presented above, it is still one which has previously been neglected. This illustrates the difference between using a simplified market to find a general solution in comparison to applying a general solution to a simplified market.

Figure 9 $(a) - (c)$ graphs CSV against the standard deviation of the underlying volatilities, $SD(\sigma_i)$, for Mid-Range underlying volatility and $N_t = 20, 50, 100$. In all cases, one witnesses an increasing relationship with CSV. This is also irrespective of average correlation as there is one still notes a clear increase within each correlation group, especially for $N_t = 100$. Whilst use of a Cholesky factorisation ensures that the simulated correlation matrix is positive definite, the calculated average correlation $\bar{\rho}$ of the simulated matrix is only approximately equivalent to the actual input correlation value. This discrepancy causes a small proportion of $\bar{\sigma}$-values to lie outside the defined input underlying volatility range. This effect is magnified for small numbers of stocks and high underlying volatility inputs. Let us reconsider the simple two-asset scenario and relax the assumption of perfect correlation. When both underlying volatilities are equal to the volatility midpoint, $SD(\sigma_i)$ is at its lowest. This gives near-to-zero CSV values at the average volatility midpoint. However, when the two volatilities are equal to the upper and lower bounds respectively, average volatility is again equal to the midpoint but now $SD(\sigma_i)$ is maximised. Therefore CSV is also maximised. [12]

11 CSV will always be zero when $\sigma_1 = \sigma_2$ in this scenario. If one considers near-perfect correlation when $\sigma_1 = \sigma_2$ over the entire volatility range, then the positive and slightly convex lower CSV bounds in Figure 7 are easily understood.

12 This is slightly higher than the midpoint of the input volatility range as it takes into account the positive correlation between the underlying volatilities. This confounding correlation effect explains the near-to-zero CSV values and the greater noise seen for very small numbers of assets when the underlying volatilities are equal to the volatility midpoint. This effect is especially prevalent when considering Low-Range underlying volatility.



(a) $N_t = 10$ (b) $N_t = 20$

(c) $N_t = 50$ (d) $N_t = 100$

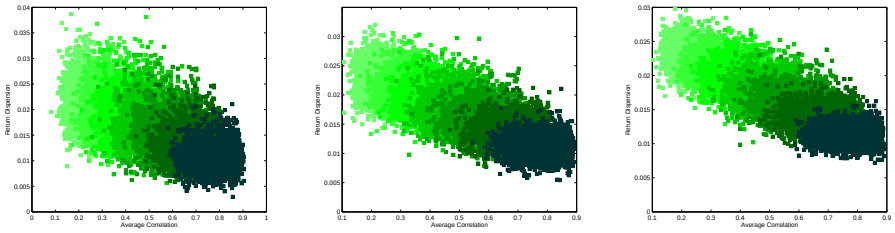
Figure 9: Uniform CSV versus $SD(\sigma_i)$ for Mid-Range volatility when $N_t = 20, 50, 100$

Returning once again to the more general simulation results displayed in Figures 7 and 9, we conclude that CSV is a monotonically increasing function of the *standard deviation* of the underling volatilities, $SD(\sigma_i)$, rather than the average underlying volatility, $\bar{\sigma}$. What is an increasing function of $\bar{\sigma}$ is actually the *range* of possible CSV. In past literature this difference has not been noted, thus leading to important differences regarding CSV drivers.

3.4.2 *Average Underlying Correlation, $\bar{\rho}$*

From past research, one would expect to find an inverse relationship between CSV and average correlation. Figure 10 (a) – (c) graphs CSV against $\bar{\rho}$ using Mid-Range underlying volatility and $N_t = 20, 50, 100$, with the data again grouped by input average correlation value. While one does find the predicted inverse relationship, a small amendment must be made: as average correlation increases, *average* CSV also decreases. Although it may seem a minor discrepancy, this amendment implies that realised CSV within high average correlation markets can be equivalent to CSV within lower average correlation markets. This is because the standard deviation

of the underlying volatilities has a greater effect on CSV values than the average market correlation. Finally, a poignant characteristic seen throughout is that the higher the correlation, the more tightly CSV is grouped around its mean value.

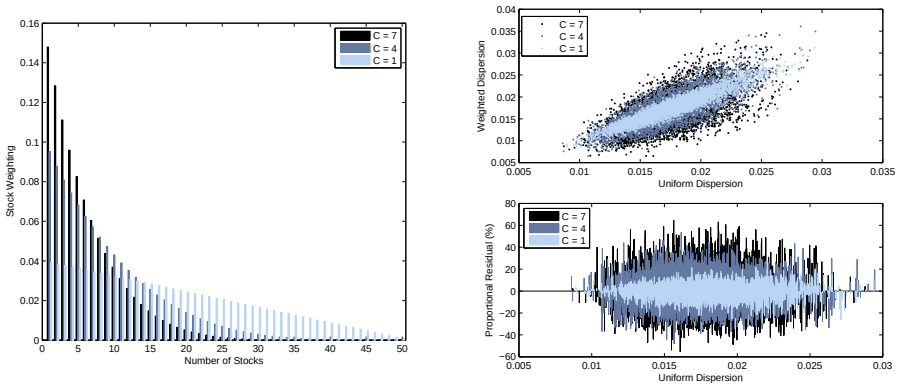


(a) $N_t = 20$ (b) $N_t = 50$ (c) $N_t = 100$

Figure 10: Uniform CSV versus Average Correlation for Mid-Range volatility when $N_t = 100$

3.4.3 *Market Concentration, C: Uniform vs Weighted CSV*

One of the most common assumptions highlighted in Section 3.2 is the choice of a uniformly-weighted market. As we show, this assumption has a severe effect on the estimation of realised CSV. Under the market simulation, stock weights are



(a) Stock weights for $C = 1, 4, 7$ (b) CSV and residuals

Figure 11: Weighted CSV versus uniform CSV under Mid-Range volatility when $\bar{\rho} = 0.5$, $N_t = 50$ and $C = 1, 4, 7$

determined according to the following exponential decay function:

$$W_{i,t} = \frac{[(N_t + 1) - i]^C}{\sum_{i=1}^{N_t} W_{i,t}}, \tag{3.14}$$

where C governs the degree of market concentration (0 = uniform weights - 7 ≈ extreme concentration). [13] As Figure 11 illustrates, market concentration can cause very material differences between uniform and weighted CSV. Furthermore, the more concentrated the market, the greater the difference becomes. For concentrated markets - that is, $C = 4$ and $C = 7$ - we find errors greater than 50 and 60 per cent of uniform CSV respectively and even for a fairly uniform market - $C = 1$ - some of the residuals are still in excess of 25 per cent. Interestingly, these residuals display severe heteroscedasticity. The highest errors are recorded for mid-level CSV and taper towards the upper and lower bounds.

3.4.4 *Intrinsic Market Factors Best-Subsets Regression of Simulated CSV*

In order to consider the effects of higher-order and interactional effects on CSV, we use a best-subsets regression methodology. The independent univariate factors are those given in Table 11. Interaction terms up to the third-order and single factor quadratic terms are also included. This gives one a total of 47 possible factors. Due to the large size of the T and N_t observations relative to those of other factors, both variables are transformed via the natural logarithm. This is done to ensure that the factor variance matrix does not bias the OLS estimates. The maximum number of model factors allowed is eight. Whilst having a greater number of explanatory variables will improve the model's fit, the difference is immaterial in the results obtained here. Thus, the parsimonious approach is taken. Before presenting the results, we note that the intention of this model is merely to highlight the importance of the interaction and higher-order terms on CSV.

Table 12 details the included variables and gives the standardised regression coefficients and their respective standard errors. Standardised regression coefficients display exactly the effects that the model predictors have on the dependent variable. The most important inference from Table 12 is that interaction between the market factors is very important, evidenced by the five interaction factors. While the best-subsets method is a very powerful tool, the best-fitting model is

13 The discussion here only considers the difference between uniform and weighted CSV. For a more general discussion of the effects of market concentration on CSV, see Appendix B.2.

chosen based on a single criterion, namely maximising the Adjusted-R^2. If this model-selection criterion value remains fairly constant across a range of fitted models but the chosen variable subsets differ over those models, then inferences drawn on a specific subset of variables must be questioned. However, if the same variables are seen throughout a range of fitted models which all display fairly consistent selection criterion values, then one is able to make valuable inferences. In this case, the difference in Adjusted-R^2 values within the top 40 models is only 0.01. Thus looking at summary statistics for the chosen variables over this subset of models enables one to see whether there is continuity in factor choice. Table 13 provides select summary statistics for the best model's chosen variables, as well as those regression factors which are not present with the highest Adj-R^2 model but are still chosen in more than 25 per cent of cases. We note from Table 13 that

Table 12: Highest Adj-R^2 Intrinsic Market Factors Regression Model for simulated CSV

Regression Variables	Std. Coefficient*	Std. Error
$SD(\sigma) * C$	-0.463	0.002
$\bar{\sigma} * SD(\sigma) * T$	0.367	0.002
$StdVol * C * N_t$	0.356	0.003
$\bar{\sigma} * T * N_t$	0.411	0.002
$SD(\sigma) * T * N_t$	0.426	0.002
T	-1.234	0.007
T^2	1.098	0.007
N^2	-0.280	0.002
Adjusted R^2	**71.40%**	

Note. The regression model is chosen using the best-subsets method. All estimated coefficients are significant at the 0.1% level

$StdVol * T * N$ is present in 39 of the best 40 regression models and $\bar{\sigma} * SD(\sigma) * T$ is present in all the models. Conversely, we find that $\bar{\sigma} * T * N$ is only present in 37.5% of the best 40 models. Thus it would seem that the standard deviation of the underlying share volatilities, $SD(\sigma)$, and its interactions with other market variables is a crucial factor of market CSV. Interestingly, there is no average correlation term included. The reason for this is that $SD(\sigma)$ would capture most of the effects attributable to average correlation. This is because as average correlation

increased, the underlying volatilities would be closer to each other and thus the standard deviation of these volatilities would be lower.

A total of five variables include the time period variable, T. In addition, it is the only single market factor included within the regression and we note in Table 13 that both it and its quadratic term, T^2, are seen within all of the best 40 models. Thus, the period over which one measures dispersion does indeed make a large difference to the value displayed. This is quite obvious and its effect is probably slightly overemphasised within this simulation due to one simulating stock returns with a zero-mean and non-zero standard deviation. Another factor seen in four of the chosen variables is N_t. This is mostly seen as an interactive term within the chosen variable. Hence the size of the market does impact the opportunity set available to investors.

Table 13: Summary Statistics across 40 Best-Fitting Models for Standardised Regression Coefficients

Best-Fitting Model Variables	**Mean**	**Std. Error**	**Max**	**Min**	**% Chosen**
StdVol * C	-0.387	0.097	-0.569	-0.179	37.5
VBar * StdVol * T	0.492	0.137	0.172	0.727	100
StdVol * C * N	0.325	0.065	0.254	0.486	27.5
VBar * T * N	0.372	0.107	0.168	0.601	37.5
StdVol * T * N	0.407	0.090	0.199	0.548	97.5
T	-1.303	0.091	-1.446	-1.059	100
T^2	1.239	0.099	1.020	1.401	100
N^2	-0.273	0.058	-0.409	-0.196	25
High Proportion Variables					
VBar * C * N	0.325	0.078	0.159	0.482	55
StdVol * C * N	-0.292	0.062	-0.378	-0.177	45
VBar * StdVol * CBar	-0.184	0.031	-0.233	-0.152	25
Adjusted-R^2	70.59%	0.21%	70.39%	71.40%	

We note certain problems within the design of this regression model. The error terms display both heteroscedasticity and non-normality due to excess skewness and excess kurtosis. Given the OLS assumptions, this does bias the regression coefficients. However, this model is not used for forecasting but rather to highlight the importance of the interactions within the intrinsic market variables. Thus, the inferences drawn from the 40 best fitting variables should still be valid. We

also note that due to the size of the data matrix and the number of possible factors within the best-subsets framework, correcting the residuals would be very time-consuming and not provide one with much greater insight.

In conclusion, the best-subsets regression model emphasises the importance of interactions between the intrinsic market factors. It also stands as a caveat against the inclusion of simplifying market assumptions within any dispersion setting.

3.5 CONCLUSIONS

CSV (or dispersion) is a metric of tremendous importance – and an awareness of its increasing usefulness is only likely to increase in current and future market conditions. This work casts a critical eye upon prior CSV research. Because prior research has focused on deriving tractable expressions for realised, expected and approximate CSV, several oversimplifications are latent. These simplifications relate mostly to the asset covariance matrix, the market weighting structure and the underlying stock process. These simplifications have been assumed in order to easily incorporate CSV within the three major fields of application: namely active management, performance attribution and risk management. We discuss how such assumptions lead to incorrect inferences regarding the underlying mechanisms of CSV. Importantly, we reveal how CSV has been incorrectly defined as an increasing function of average underlying volatility and a decreasing function of average stock correlation.

Using a GBM process that allows time- and stock-deterministic drift, volatility and correlation terms to describe stock prices within a weighted market, we develop a general theorem for CSV from which a practical and accessible expression for expected CSV is derived. This derivation shows that CSV is not a simple linear function of average underlying volatility. By analysing CSV under a broad range of realistic market simulations, we verify that CSV is actually a monotonically increasing function of the *standard deviation* of the underling volatilities, rather than the average underlying volatility. Interestingly, we note that what is an increasing function of average volatility is rather the range of simulated CSV values. In

addition, we reveal that it is actually the *average* CSV level that is a linear, inverse function of average stock correlation. For completeness sake, we further note the material difference between CSV calculated under uniform and market-cap stocks weights, even for relatively uniform markets. This difference increases with the level of market concentration.

4

UNDERSTANDING EMPIRICAL CROSS-SECTIONAL VOLATILITY

4.1 INTRODUCTION

Cross-sectional volatility (CSV), or (return) dispersion, refers to the standard deviation of returns for a specific basket of underlying assets over a particular term. It is commonly quoted over shorter terms - monthly or less - and represents the spread of returns within the stated basket from that of the mean stock. CSV thus represents both the opportunity set available for fundamental and statistical-arbitrage stock picking, as well as the basis for dispersion trading. We define CSV, denoted ${}_{N_t}\chi(r_t)$, for an N_t-asset market over term t as:

$$\chi_t(r_t) = \sqrt{\sum_{i=1}^{N_t} w_{it}(r_{it} - r_{mt})^2}, \tag{4.1}$$

where w_{it} is the weight for underlying i at the beginning of term t and $r_{mt} \equiv \sum_{i=1}^{N_t} w_{it} r_{it}$ is the weighted market return over the term.[1]

In comparison to the more common longitudinal volatility, which measures the standard deviation of a series over time, CSV is a fairly new concept within financial research. However, it has already become prominent within several diverse financial fields - in both literature and application.[2] Despite the growing interest in CSV as well as trading CSV, very little work exists in the literature that has attempted to analyse the underlying algebra of CSV or to explore patterns within empirical CSV from a statistical point of view. While we continue to address

1 When substituting $w_{it} = 1/N_t$ for all underlyings in Equation 4.1, the metric is commonly referred to uniform CSV or uniform dispersion.

2 These fields include, but are not limited to: active management, fund performance attribution, risk management and volatility trading. Some of the principal research on incorporating CSV within these financial areas is given in Solnik and Roulet (2000), de Silva, Sapra and Thorley (2001), Ankrim and Ding (2002), Heang and Satchell (2005, 2006), Yu and Sharaiha (2007) and Gorman, Sapra and Weigend (2010a, 2010b).

the former concern in Chapter 3, this contribution is intended to initiate a body of statistical work that will eventually close the gap within the latter field of CSV literature. The challenge, when viewed from a statistical point of view, is how best to approach assessing the predictive capacity latent in the time series that CSV represents.

For this broad purpose, we necessarily need to adopt a heuristic mindset in this paper by approaching the statistical modelling of empirical CSV from three very different angles. First, as is evident, CSV represents a self-referencing series, and hence analysis via ARIMA (Auto-Regressive Integrated Moving Average) methodology is justified. This analysis comprises Section 4.3, and we find that ARIMA models generally disappoint. This finding is noted here for interest and completeness sake. Second, since our series of CSV is empirically grounded, it may make sense to create a model utilising a generalised-linear model framework to assess if macroeconomic factors (and potentially their interactions) usefully predict the evidenced patterns in CSV. A reasonably predictive and robust multi-factor model is presented in Section 4.4 with some interesting conclusions. Lastly, and slightly atypically for the analysis of financial time-series, we adopt a survival-analytic approach to understanding whether CSV possesses a latent memory. Results are discussed in Section 4.5, but we note here that implications of a memory within CSV has piquant implications for the assessment and trading of CSV. Section 4.6 provides a concluding discussion.

We trust that this preliminary exploration of the latent predictability of CSV will prove a useful first-attempt at better understanding what is surely going to evolve as one of the most informative metrics in financial markets in the 21st century.

4.2 EMPIRICAL CSV DATA CHARACTERISTICS

In October 2010, Russell Investments and Parametric Portfolio Associates released the CrossVol^{TM} indexes. These series report daily, 22-day and monthly cap-weighted CSV for a wide range of Russell Global indexes. This contribution looks at monthly CSV values for a subset of these indexes over the period July 1996 until

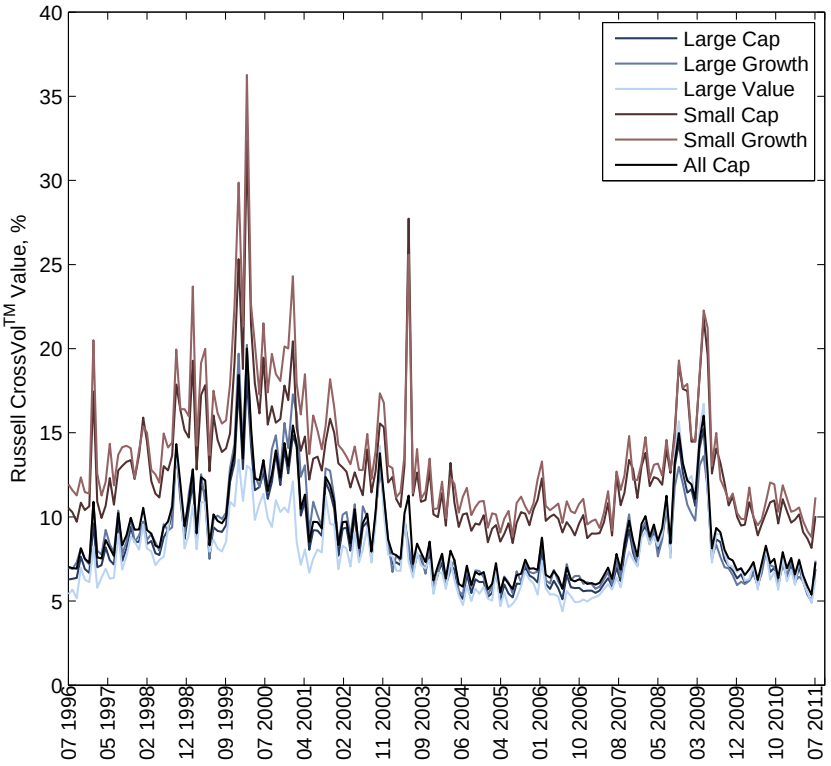


Figure 12: Russell-Parametric cross-sectional volatility series for several Russell Global indexes

July 2011. The CSV data analysed is quoted for the Russell Global All Cap, Large Cap, Large Value, Large Growth, Small Cap and Small Growth indexes, displayed in Figure 12. This dataset showcases several important characteristics. Firstly, the analysed CSV series are all highly correlated. Even the lowest correlated CSV pair - Large Value and Small Growth CSV - display $\rho \approx 0.8$. Secondly, the empirical CSV series are skewed to the right and are quite peaked - this is especially apparent within the Small Cap and Small Growth CSV series. This manifests itself through relatively stable lower CSV values and comparatively more volatile higher CSV values. From this behaviour, it would seem likely that there is useful information hidden within the CSV distribution tails. This observation is considered further in Section 4.5. We now look at several different methods for modelling the empirical CSV.

Table 14: Best-fitting ARIMA models for Russell Global indexes CSV series ranked according to Adjusted-R^2

Russell Global Index	Best-Fitting Model	Adjusted-R^2
All Cap	ARIMA$(5,1,0)$	24.20%
Large Cap	ARIMA$(5,1,0)$	24.87%
Large Value	ARIMA$(5,1,0)$	22.68%
Large Growth	ARIMA$(5,1,0)$	27.25%
Small Cap	ARIMA$(3,1,2)$	22.68%
Small Growth	ARIMA$(5,1,0)$	24.92%

4.3 MODELLING CSV AS AN ARIMA PROCESS

Given that CSV is a time series itself, an obvious starting point is self-referencing, time series models. We consider a wide range of ARIMA(p,d,q) models, where $p < 5$ gives the number auto-regressive terms, $q < 5$ gives the number of moving average terms and $p + q \leqslant 5$ for all models.[3] All original CSV data series are found to be non-stationary under Augmented Dickey Fuller and KPSS tests. Thus the data are transformed via the natural logarithm and subsequently differenced, giving $d = 1$. Although taking logs of the CSV series is not necessary to ensure stationarity, it does lead to markedly better model fits. Unfortunately, the ARIMA$(p,1,q)$ models still provide very poor results both in-sample and when forecasting. Table 14 gives the best-fitting model for each log CSV series ranked according to the model's Adjusted-R^2. In all cases, barring Small Cap CSV, an ARIMA$(5,1,0)$ model is identified as the best-fitting model. The repetition of this model is not particularly surprising given the strong correlations between these empirical CSV series. However, even within the best-fitting models, the maximum Adjusted-R^2 values range between 22.7% and 27.3%. This implies that very little of the variation within the CSV series can be explained by incorporating auto-regressive terms. In addition, the model residuals are both normal and homoscedastic. Thus, the poor results cannot be ascribed to a violation of the underlying model-fitting assumptions.

3 Fitted models which display poor residual behaviour or strong signs of non-stationary and non-invertibility are excluded from consideration.

From this exercise, one must conclude that empirical CSV is not an obviously self-referencing series and hence the general ARIMA methodology does not provide one with an informative analytic tool.

4.4 MODELLING CSV AS A MACROECONOMIC MULTI-FACTOR REGRESSION

We now consider modelling the monthly change in Global CSV - proxied by Russell Global All Cap CrossVol - through a suite of extraneous macroeconomic factors. The monthly factor data covers the period starting at February 2005 until July 2011, giving a total of 75 observations after the exclusion of two confounding outliers. Change in Global CSV (Delta-G-CSV) is modelled in order to ensure normality and stationarity within the dependent variable.[4] Initially, 32 univariate international factors, as well as their interaction and higher-order terms, are considered for inclusion within the model.

In many macroeconomic models, explanatory power is found within the higher-order polynomial and interaction terms and modelling CSV turns out to be no exception. Understandably, when starting from a base of 32 factors, the number of interaction terms becomes computationally prohibitive. Therefore both interaction and polynomial terms are limited to order three and lower. Even so, one is still left with over 6000 possible predictors. Although one could use a direct best-subsets regression method (BSR) in order to find the optimal model, this becomes intractably cumbersome when the number of factors deemed necessary within the model is close to ten. However, by repeatedly running a BSR on a sufficient number of randomly sampled subsets of the initial univariate macroeconomic predictors, a multi-factor model with a surprisingly high predictive capacity is found. This model gives an Adjusted-R^2 of 65%.[5] Table 15 details the included factors and gives the standardised regression coefficients and their respective standard errors. Standardised regression coefficients display exactly the effects that the model predictors have on the dependent variable. Using these coefficients,

4 For the sake of brevity, this discussion is limited to Global All Cap CSV. However, the highly-correlated nature of the empirical CSV series considered would suggest a fair degree of robustness within the displayed results.

5 Appendix C.1 motivates the use of BSR and details the model assumption checking process.

Table 15: Standardised regression coefficients and standard errors for the highest Adjusted-R^2 macroeconomic multi-factor model of Delta-G-CSV

Macroeconomic Factor	Description	Std. Coefficients (Std Error)
MXEF Index	MSCI Emerging Markets Index	0.834 (0.146)
BEMETAL Index^2	Europe 500 Metals & Mining Index	0.633 (0.114)
EURGBP*CRB CMDT	Interaction	0.395 (0.082)
LMEX Index	London Metals Exchange	0.208 (0.112)
CRB CMDT Index*	Commodity Index (in USD)	0.121 (0.113)
EURGBP Currency	EURGBP Exchange Rate	-0.313 (0.077)
MXEU Index	MSCI European Index	-0.632 (0.119)
CRB CMDT*MXEF	Interaction	-1.049 (0.144)
CRY Index^3	Jefferies CRB Commodity Index	-1.152 (0.183)
Adjusted-R^2		65.03%

*The factor *CRB CMDT Index* is included due to the necessity of including the univariate factors within all multiplicative interaction terms.

one is able to identify the underlying economic rationale for the choice of this particular subset of possible predictors.

Within the model, we find that positive effects on the model's forecasts - in order of decreasing magnitude - are attributable to:

- MSCI Emerging and Frontier Markets Index (MXEF),
- European 500 Mining and Metals Index (BEMETAL) - a proxy for world resources markets,
- Euro-Dollar differential term - proxied by the EURGBP-CRB Commodity interaction
- London Metals Exchange (LMEX) - United States (US) Dollar-denominated index only including six metals
- CRB Commodities Index (CRB CMDT) - representing a collection of price actions on a wide range of commodities.

It is a well-established observation that first-world markets are becoming increasingly correlated and that investors are subsequently moving towards emerging and frontier markets in order to participate in independent, high-return (and

consequently high-risk) opportunities.[6] It is thus no surprise that the greatest positive effect on global CSV is attributable to the MXEF factor. In addition, considering that the majority of third-world countries are resource-driven economies, the positive BEMETAL factor also gives one an indicator of investment opportunities within these markets. The Euro-Dollar differential gives an indication of the investor sentiment within the major Western markets. An increase will lead to a decoupling of US and European markets, thus leading to an increase in Global CSV. Although the LMEX index is also a metals-based factor, due to the small set of non-ferrous metals used as its underlyings, it is uncorrelated with the BEMETAL factor ($\rho \approx -0.03$). Thus, taken together, these two factors capture the majority of global mining and metals price action. Finally, the CRB Commodity Index is only included within the model due to the convention of including all univariate factors within all multiplicative interaction terms. Given that global CSV measures the available, investable opportunity set, these predictors are indicative of 'risk-on' market factors.[7]

Conversely, we find negative effects - in order of increasing magnitude - coming from a range of predictors:

- Jefferies CRB Commodities Research Index (CRY) - representing the momentum within world commodity pressure,
- Interaction between US commodities and the MSCI Emerging Index (CMDT-MXEF) - a proxy for the risk differential between first- and third-world economies. This is due to the fact that many third-world economies are based on commodities.
- MSCI Europe Index (MXEU) - a proxy for world equity markets,
- Euro-British Pound differential (EURGBP).

From the discussion on positive predictors, the main drivers of empirical CSV are found to be emerging and frontier markets and the global resources market levels. In addition, most emerging and frontier markets themselves are based

6 Some of the larger emerging markets include the Brazil, Russia, India, China and South Africa; the so-called BRICSA nations.

7 That is, factors which represent those assets that investors are willing to risk money on.

on commodities. Chapter 3 showed that increasing the average correlation of the underlying assets caused the average CSV values to decrease. Thus, the negative effects for both the CRY and CMDT-MXEF predictors, which inherently represent the correlation levels between commodities and between developed and emerging markets respectively, are economically justifiable. In addition, one also notes that negative effects are attached to the MXEU and EURGBP variables, which represent recent factors that capture idiosyncratic risks associated with the Eurozone. Because these four particular predictors negatively effect the model's CSV forecast, one can infer that, in their present setting, they represent 'risk-off' market factors.[8]

Collectively, the chosen model includes a wide array of global 'risk-on/risk-off' predictors that are fairly representative of the current international investment climate, and thus latently, the economic swing between commodity safehavens, an emerging China and Western equity appetite.

4.5 MODELLING CSV THROUGH SURVIVAL ANALYSIS

As alluded to in Section 4.2, one can gain useful information from studying the distribution tails of CSV. Rather than focusing directly on the given extreme CSV observations though, we apply survival analysis to the time between recurring extreme CSV values. Consequently, we are able to determine the latent 'memory' inherent within Global CSV (as proxied by the Russell Global All Cap CSV series).

4.5.1 *The Tools of Survival Analysis within a CSV Framework*

Although survival analysis has largely been used to model 'death' within biological organisms or 'failure' within mechanical systems (where it is commonly termed reliability analysis), the same principles can just as readily be applied to the time between recurring extreme CSV values. Before this can be done, we need to define

8 That is, factors which represents assets that carry the risk of loss. An example of this would be that foreign currencies - here given by the EURGBP currency factor - are normally shorted in a risk-off market.

an extreme CSV value as the *first* value that breaches either the upper threshold, u, or lower threshold, l, respectively. That is,

$$\begin{aligned} U &= \{\chi_t(r_t) | \chi_{t-1} < u \leqslant \chi_t \, \forall t\} \\ L &= \{\chi_t(r_t) | \chi_{t-1} > l \geqslant \chi_t \, \forall t\}. \end{aligned} \tag{4.2}$$

The upper (lower) threshold is always chosen to be greater (less) than the mean CSV value. This threshold is subsequently moved outwards from the mean in 0.5% increments. Following from this, the random variable T_i, representing the time between recurring extreme CSV values, is written as

$$\begin{aligned} T_1 &= \{\min(t-s) | (\chi(s), \chi(t) \in U) \, \forall \, s < t\} \\ T_2 &= \{\min(t-s) | [(\chi(s) \in U, \chi(t) \in L) \vee (\chi(s) \in U, \chi(t) \in L)] \, \forall \, s < t\} \\ T_3 &= \{\min(t-s) | (\chi(s), \chi(t) \in L) \, \forall \, s < t\}. \end{aligned} \tag{4.3}$$

Thus, T_1 measures the time interval between $U(t)$ and $U(t-1)$, T_3 measures the time interval between $L(t)$ and $L(t-1)$ and T_2 measures the time interval between either $U(t)$ and $L(t-1)$ or between $L(t)$ and $U(t-1)$. When displayed graphically, as in Figure 13, these duration variables become much more intuitive. Using points A to D, T_1 is defined as the time between points A and D. T_2 is defined as the times between points A and B and B and D The time interval between points B and C is defined as T_3.

The T_i datasets are dependent on the specified threshold values. In order to avoid overly small datasets, only those threshold values which gives $T_i > 10$ are considered. We focus this discussion on T_2, which measures the minimum time intervals between extreme high CSV values and extreme low values, or vice versa. For information on T_1 and T_2, see Appendix C.2. As previously mentioned, we need to consider the notion of 'memory' within a dataset. In order to do so, we first define the survival function, denoted S. This function classically measures the probability that the time of 'death' for a certain variable is greater than t. In

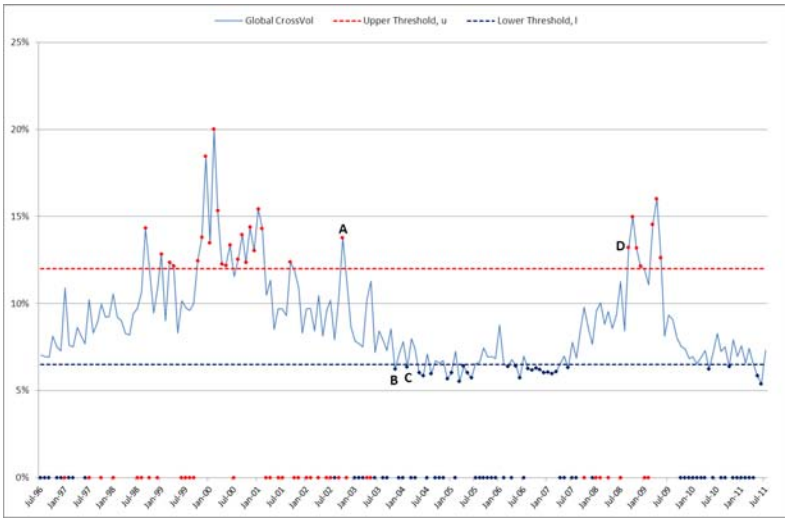


Figure 13: Global CSV for the period July 1996 to July 2011.

Note. The dotted red line represents the upper threshold u, and the dotted navy blue line represents the lower threshold l. All CSV values that are extreme highs, $U(t)$'s, are highlighted with red points, whilst extreme lows, $L(t)$'s, are highlighted with navy blue points.

the case of CSV recurrence times, T_2 denotes the time until the next extreme event occurs, rather than the time until death. Thus we have that

$$S(t) = \mathbb{P}[T_2 > t] = 1 - \mathbb{P}[T_2 \leqslant t] = 1 - F(t), \tag{4.4}$$

where $F(t)$ is the cumulative distribution function of T_2.

Using Equations 4.3 and 4.4, we can now define the concept of 'memory'. The CSV recurrence time series T_2, is defined as *memoryless* if and only if the probability of conditional future survival is unaffected by past survival. That is,

$$\mathbb{P}[T_2 > s + t | T_2 > s] = \mathbb{P}[T_2 > t] \ \forall\, s, t \geqslant 0. \tag{4.5}$$

Although the survival and memory functions are directly applicable in most biological and mechanical studies, they do not directly provide pertinent information about the behaviour of future CSV. Instead, by shifting ones focus onto the probability of CSV moving to an opposite extreme value conditional upon past survival, it is possible to find directly applicable information regarding future CSV

values. The hazard function, defined as $h(t)$, considers exactly this aspect. Letting $f(t)$ denote the probability density function, we write:

$$h(t) = \mathbb{P}[t < T_2 \leqslant t + dt | T_2 \geqslant t] = \frac{f(t)}{S(t)} = \frac{f(t)}{1 - F(t)}.$$

Note that the hazard function is wholly determined by the distribution fitted to the data. If the hazard function is not constant for T_2, then we are able to draw an inference into the future behaviour of market dispersion.

4.5.2 *Fitting an Inverse Gaussian Distribution to* T_2

There are a limited number of distributions used in classical survival analysis. These include the Exponential, Weibull, Gamma, Birnbaum-Saunders, Inverse Gaussian (IG) and Lognormal distributions. In most biological survival studies, past research dictates the distribution that not only gives the best fit but also encompasses the most appropriate properties for the dataset under review. However, given the shortage of CSV modelling research, there is no such precedent. Thus, we first compare how well the candidate distributions fit over all T_i datasets.[9] Using the log-likelihood values to measure the fit, one finds that the IG distribution provides the best fit over all T_2 datasets. Consequently, we focus solely on this distribution. For completeness sake, the IG density function is written as

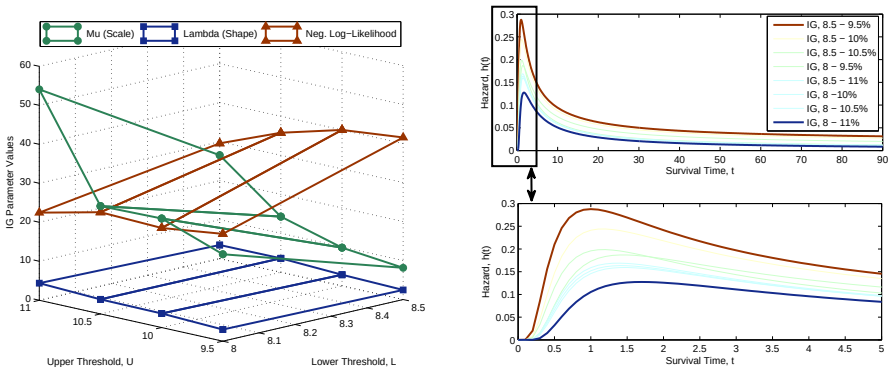
$$f(x; \mu, \lambda) = \left[\frac{\lambda}{2\pi x^3}\right]^{1/2} \exp\left(\frac{-\lambda(x-\mu)^2}{2\mu^2 x}\right), \tag{4.6}$$

for $x > 0$, where $\mu > 0$ is the scale parameter (and mean), and $\lambda > 0$ is the shape parameter.[10] Figure 14 (a) plots μ, λ and negative log-likelihood against the selected threshold values for all T_2 datasets. In addition, Figure 14 (b) displays the hazard function for the fitted IG distributions over the total survival time range t, and for $t \leqslant 5$. The highlighted hazard functions represent the T_2 dataset with the most observations and the correspondingly worst fit - $u = 9.5\%$, $l = 8.5\%$ -

9 Table 17 in Appendix C.2 provides distribution ranks for all T_i.

10 Only those distributional properties directly related to T_2 modelling are discussed here. For a thorough review of the IG distribution and its statistical properties, see Folks and Chhikara (1978).

as well as the T_2 dataset with the least observations and corresponding best fit - $u = 11\%$, $l = 8\%$. Unsurprisingly, the two scenarios correspond to the T_2 datasets with the smallest and largest threshold spreads respectively. We now consider the fitted IG distribution parameters



(a) μ, λ and Neg. Log-Likelihood

(b) Hazard Function, $h(t)$

Figure 14: Fitted IG distribution parameters and hazard functions for T_2

Whilst the shape parameter, λ, remains remarkably constant across all threshold combinations, the scale parameter, μ, increases drastically with upper threshold values. Conversely, the negative log-likelihood values are inverse in movement to the scale parameter. Movements in the underlying IG parameters - in this case, predominantly the scale parameter - translate directly into changes within the shape of the fitted distribution's hazard function.

Rather than the monotonically increasing function commonly seen in many survival analysis applications, the IG hazard function increases until reaching a maximum value to the right of the distribution's mode. The hazard values then asymptotically tends towards a lower bound, given by $h(t) \to \lambda/2\mu^2$ as $t \to \infty$. This explains the lower long-term hazard values seen for T_2 under larger threshold spreads. The greater threshold differential leads to much higher μ values but similar λ, thus giving one the lower asymptotic bound.

The difference in distribution parameters is also responsible for the large difference in maximum hazard values. When looking across all T_2 datasets, the maximum hazard value increases as the threshold differential decreases. In addition, this maximal value also occurs at an earlier time. For the T_2-fitted IG

distributions, the resultant hazard function is maximised for $1 \leqslant t \leqslant 1.7$. Intuitively, this means that for current extreme values, the probability of moving to the opposite extreme is maximised approximately $1 - 1.7$ months into the future. If CSV does not move to the opposite extreme, the probability of a future move slowly decreases and becomes highly unlikely in the long-term.

Appendix C.2 shows that by running similar exercises on T_1 and T_2, we have that the T_1 hazard rate is maximised when $2.8 \leqslant t \leqslant 4.2$ and similarly, that the T_3 hazard rate is maximised for $2.1 \leqslant t \leqslant 4.1$. In addition, while the maximum hazard values for T_1 and T_3 are approximately equal, the values are roughly half that of their T_2 counterpart. Taken in conjunction with the time-range for T_2 maximal hazard values, this means that the most likely path for current extreme CSV is to move initially to the opposite extreme before returning to the current extremes. In addition, this likely passage could be slightly longer if CSV is currently an extreme high than if it were an extreme low.

4.6 CONCLUSIONS

Cross-sectional volatility (or return dispersion) is a metric that has swiftly gained prominence within a large body of financial literature and application. Despite this growing, substantial interest in CSV, there remains very little literature that attempts to explore patterns within empirical CSV from a statistical point of view. This contribution provides a first-attempt at statistically analysing these patterns. In order to assess the latent memory within this time series data, we statistically model empirical CSV under three very different methodologies: namely ARIMA modelling, a generalised-linear multi-factor approach and a survival-analytic method.

The ARIMA modelling provides a very poor fit for all the Russell CSV series considered. Even when allowing for a very general set of model parameters, the maximum Adjusted-R^2 across all models is only 27%.

Conversely, Global CSV (proxied by Russell Global All Cap CSV) is more than adequately modelled by a fairly simple macroeconomic multi-factor model.

Through a combination of random factor sampling and best-subsets regression, an exogenous macroeconomic model is found which displays a reasonable predictive capacity, with an Adjusted-R^2 of 65%. In addition to this capacity, the chosen model factors describe a wide array of global 'risk-on/risk-off' factors that are collectively representative of the current international investment climate, and thus latently, the economic swing between commodity safehavens, an emerging China and Western equity appetite.

Finally survival analysis is used to model the time between recurring extreme Global CSV values. We find that the Inverse Gaussian distribution definitively provides the best data fit for all permitted upper and lower extreme threshold values. We analyse the recurrence times between extreme high and extreme low CSV or vice versa, T_2, and briefly touch on the times between consecutive extreme high CSV values, T_1, and consecutive extreme low CSV values, T_3. The finding of a discernible and statistically significant hazard function for Global CSV is interesting for the following reasons. Firstly, CSV movements have a probabilistic interpretation - the most likely path for extreme observed values of CSV is to the opposite extreme. Secondly, this likely path could potentially be longer if CSV is currently an extreme high. Lastly, in the absence of reversion to opposite means, the hazard rate stabilises predictably after a certain time period. Ideally, one would couple the estimates from the parametrised distribution with the covariance between them, to get an idea of the expected variation in this hazard rate that is naturally occurring. This parametrisation would be done dynamically, so as to study the natural variation in CSV hazard rate over time, and importantly, currently. Lastly, if one could relate expected CSV hazard rates to macroeconomic changes, as we have alluded to in our multi-factor model (Section 4.4) there is a reasonable possibility that with the right statistical construct, short-term CSV becomes predictable. The ability to monetise such information, given the current derivative-based financial instruments poised at this space, is provocative.

5

THE TERM-STRUCTURE OF VOLATILITY

5.1 INTRODUCTION

Since the inception of Black and Scholes (1973) and Merton's (1973) seminal papers, option pricing has been characterised by the celebrated Black-Scholes(-Merton) pricing formula. However, many researchers have since shown that the rather stringent parametric assumptions of the Black-Scholes model disagree with empirical market data. Specifically, the implied volatilities calculated from observed option prices, rather than being constant as assumed by Black-Scholes, have been shown to depend upon both option term and moneyness.[1] In fact, research on understanding and characterising this stylised fact – named the volatility 'skew/smile' – has become very commonplace. In addition, many researchers across a wide range of national and international markets have given considerable evidence to suggest that log-stock and/or log-index returns are not normally distributed.[2]

Due to these violations, the focus of much derivative research has been on testing the implications of these assumptions and the introduction of alternative option pricing theories providing a better empirical fit. One such alternative is the nonparametric Canonical Valuation (CV) method proposed by Stutzer (1996) and further developed by Derman et al (1997) and Stutzer and Chowdhury (1999). This pricing technique uses only historical underlying price data and thus avoids the necessity of specifying underlying return dynamics. Stutzer normalised the historical return distribution via the principle of relative entropy – developed in information theory – in order to find a risk-neutral option price. This method is very robust and can be easily altered to include multiple underlyings as well as empirical option price data. Alcock and Gray (2005) extended Stutzer's ori-

1 See, for example, Rubinstein (1994), Derman and Kani (1994) and Dupire (1994)
2 See for example, Jackwerth and Rubinstein (1996), Egan (2007)

ginal work by developing the theory for a nonparametric, dynamic delta-hedging portfolio, which provided investors and traders with a tractable nonparametric valuation framework for European vanilla and basket options.

A previous criticism of this option pricing theory was that the CV framework could not price American options correctly as the early exercise feature could not be accommodated. This criticism was addressed by Duan (2002) and, more recently, by Alcock and Carmichael (2008), both of whom extended Stutzer's original method in order to allow for the early exercise feature of American options.

Although several other nonparametric pricing approaches have been proposed Alcock and Carmichael (2008) note that the majority of proposed methods rely heavily on existing option prices.[3] In reality, the cited nonparametric methods should be viewed more as numerical interpolation algorithms rather than true nonparametric option valuation theories.

The majority of prior research on nonparametric CV has focused on two particular areas: (i) evaluating the pricing accuracy relative to Black-Scholes under different volatility regimes (Gray and Newman (2005), Alcock and Auerswald (2009)), and (ii) generalising the underlying CV pricing theory (Alcock and Gray (2005), Hayley and Walker (2009), Cadogan (2010)). Another interesting research avenue that has only been partially explored is the relationship between the CV implied volatility surface and the market implied volatility surface.

Zou and Derman (1999) introduced the notion of Strike-Adjusted-Spread (SAS), defined as the spread between the current Black-Scholes implied volatility and the implied volatility obtained via nonparametric CV. It is, in essence, a one-dimensional metric ranking the relative richness of equity options across both strike and term at a fixed date. Thus, SAS is more accurately written as a function of both strike (or moneyness) and term: $SAS(K, T)$. Although Zou and Derman provide several practical applications of SAS, their focus is on evaluating whether the current volatility skew across strike is fair and subsequently providing a

3 These pricing approaches include, but are not limited to, Hutchinson, Lo and Poggio (1994), Rubinstein (1994), Ait-Sahalia and Lo (1998), Ait-Sahalia and Duarte (2003), Hamid and Habib (2005), and Yatchew and Hardle (2006). Refer to Fengler (2005) and Jackworth (2004) for a more complete overview.

measure of relative option value. The SAS implications on term, although alluded to, are not considered.

Another recent study which explores the link between the CV implied volatility surface and the market implied volatility surface is that of de Araujo and Maré (2006). Using the revised method proposed by Duan (2002), de Araujo and Maré conducted a South African-based study on liquid, equity index options which showed that the implied volatility surface obtained via the CV method was very similar to that implied by the market. Following from this insight, de Araujo and Maré motivated for the use of the CV method to generate volatility surfaces for illiquid single stock options. This suggestion is, in essence, an extension of Stutzer (1996) and Zou and Derman's (1999) propositions to use the CV method to price 'unusual', and thus illiquid, basket options. As with previous research though, the direct implications between option term, market implied volatility and CV implied volatility are only obliquely considered.

This contribution revisits the idea of SAS but rather focuses on the term-structure of volatility. In order to do this, we introduce the metric Term-Adjusted-Spread, or $TAS(T, s)$. TAS is again defined as the spread between Black-Scholes market implied volatility and CV implied volatility but is now measured across option terms over time at a fixed strike. By altering the focus in such a way, one is able to compare and contrast the current volatility term-structure implied within the market to the appropriate volatility term-structure implied by the underlying asset's historical distribution. Because of the dependence on historical data, the majority of work done with TAS is empirically led.

Another aspect that naturally arises when considering the volatility term-structure is the statistical differences between the historical return distributions over different terms. What is the relationship between the 3-month and 9-month historical return distributions and how does this manifest itself in the volatility term-structure? By fitting a general distribution to rolling historical data, can one forecast the distribution's future parameters and use this in conjunction with CV? Such questions are of vital importance to both options speculators and replicators.

This contribution provides several key findings. Using TAS it is shown that, within the South African market, the volatility term-structure implied by the market differs markedly to that implied by nonparametric CV. In addition, CV implied volatility varies depending on which historical period is used to construct the future risk-neutral density.[4] This implies a changing historical distribution. Ironically, this leads one to a somewhat more parametric approach, although still motivated by the CV method. Several general statistical distributions are fitted to historical data at a number of specified points. By modelling the calculated distribution parameters at these points via general time-series methods, one can forecast the future asset price distribution. Through the principle of relative entropy, one can subsequently calculate TAS for the fitted distribution families. Although this approach can best be claimed as semiparametric, it is at least empirically led and still makes use of the majority of Stutzer's CV framework.

This chapter is organised as follows. Section 5.2 provides an overview of CV and outlines the motivation for using the information theoretic principle of relative entropy. Section 5.3 considers the distributional properties of historical South African return data and compares both the empirical and best-fitting distributions across respective return horizons. Starting by considering the current market-implied term-structure of volatility within the South African context, Section 5.4 then discusses TAS and its rather interesting findings. The choice of historic asset return period and its effect on TAS is also considered. Section 5.5 then fleshes out the bootstrapping method used to fit several distribution families to the rolling periods of historical data. Following from this, TAS is calculated for the forecast distributions and is compared to that found using the purely empirical return data. Section 5.6 concludes.

5.2 OVERVIEW OF CANONICAL VALUATION

The theory of option pricing is based on the proposition that if no arbitrage opportunities exist within a market, there exists a risk-neutral return distribution

4 Option traders that have particular views on the future asset price distribution can exploit this fact, thus introducing the possibility of informational value-add.

$\mathbb{Q}$, such that the value V, of a contingent claim at time t on an underlying asset priced S_t, is given by the discounted expectation of the payoff. Mathematically, we write

$$V(S,T) = e^{r(T-t)}\mathbb{E}_{\mathbb{Q}}\left[\text{Payoff at } T \mid S_t, t\right] \tag{5.1}$$

where r is the continuous risk-free interest rate and $\mathbb{E}_{\mathbb{Q}}[\ldots \mid \ldots]$ is the conditional expectation under the risk-neutral measure $\mathbb{Q}$. Thus, if one knows $\mathbb{Q}$, one can calculate the value of any European option contract.

In the Black-Scholes framework, the lognormal density function with given volatility is assumed to be the implied risk-neutral continuous distribution $\mathbb{Q}$. Other pricing theories assume different underlying distributions. Stutzer (1996) challenged this assumption by considering the case where one does not want to assume a particular continuous-time process. Based on the fundamental option valuation theory above, Stutzer used a normalisation method commonly used in discrete time models. At each future time T, the price process is discounted by the product of one-period, gross risk-free interest rates $r(t)$ up to T. Denoting the current price of the asset by S_t and dividend payment at time t by D_t, the equivalent martingale probabilities q at time T must satisfy:

$$\begin{aligned} S_t &= \mathbb{E}_{\mathbb{Q}}\left[\frac{S_T + D_T + \sum_{t=1}^{T-1} D_t \prod_{s=t}^{T-1} r(s)}{\prod_{t=1}^{T} r(s)}\right] \\ &= \mathbb{E}_{\mathbb{P}}\left[\frac{S_T + D_T + \sum_{t=1}^{T-1} D_t \prod_{s=t}^{T-1} r(s)}{\prod_{t=1}^{T} r(s)} \frac{dq}{dp}\right], \end{aligned} \tag{5.2}$$

where $\mathbb{P}$ denotes the real-world probability measure and dq/dp denotes the Radon-Nykodym density of the martingale measure with respect to $\mathbb{P}$ at time T. Thus, one must be able to estimate the equivalent martingale measure satisfying the no-arbitrage constraint given in Equation 5.2 in order to calculate the fair value of a European derivative claim from Equation 5.1. In order to do this nonparametrically, Stutzer (1996) proposed a three-part method. Firstly, historical asset returns and risk-free rates are used to estimate the real-world probability distribution $\mathbb{P}$ of the

underlying asset at time T. Secondly, the principle of relative entropy is used to transform the estimated real-world density into an estimated risk-neutral density of the equivalent martingale measure $\mathbb{Q}$ satisfying Equation 5.2. Thirdly, the derivative contract is valued using Equation 5.1 and $\hat{\mathbb{Q}}$, after which the implied CV volatility is calculated.

5.2.1 *Estimating the Future Empirical Distribution*

Consider a European option expiring at time T written on an asset with a current price of S_t. According to Stutzer (1996), the initial step to value this contract is to estimate the real-world probability distribution $\mathbb{P}(S_0, 0; S_T, T)$. Assuming that the historical underlying return distribution is a plausible estimate for the current distribution, one is easily able to construct empirical stock sample price paths. Consider the underlying's historical (rolling) return series of T-year returns, R_i, $i = 1, \ldots, n$, calculated as price relatives. Using these historical returns, we construct a distribution of possible prices for the underlying asset at time T:

$$S_i = S_0 R_i, \; i = 1, \ldots, n \tag{5.3}$$

where S_0 is the current underlying asset price. Each possible future price S_i has an estimated real-world probability $\hat{p} = 1/n$. Using the general assumption that the returns are generated by an unknown ergodic Markov chain, Stutzer (1996) points out that $\hat{\mathbb{P}}$ is an optimal nonparametric estimator of the unknown, invariant real-world distribution $\mathbb{P}$ given that its rate of convergence is the fastest possible among all such consistent estimators.

Using only the historical return distribution has several advantages. From a mathematical standpoint, the smoothness assumptions normally required by kernel-smoothed empirical distribution pricing methods are not required for CV. From a financial viewpoint, use of the historical return distribution to estimate the future real-world return distribution is more likely to ensure that certain stylised market features, such as skewness and leptokurtosis, are captured (Alcock and

Carmichael (2008)).[5] In addition, one is easily able to incorporate stochastic interest rates and dividends as well as options written on multiple underlying assets using this method.

5.2.2 *Estimating the Risk-Neutral Density via Relative Entropy*

Before discussing the estimation the risk-neutral return distribution, we briefly introduce the idea of entropy as a measure of uncertainty.

Information, Entropy and Uncertainty

Information within financial markets plays a large role in shaping an investor's market view. Zou and Derman (1999) show that the information, $I()$, obtained from the occurrence of a random event with assumed probability p can be quantified by a solving a series of differential equations and is given by

$$I(p) = -\ln(p). \tag{5.4}$$

Using Equation 5.4, the entropy of a random variable X, whose i^{th} observation has probability p_i, is defined to be the expected value of the information obtained from an event within the distribution:

$$H(X) = -\sum_{i=1}^{n} p_i \ln(p_i) \tag{5.5}$$

Because all probabilities are less than 1, entropy is always positive. Higher expected values of information imply a greater spread of probabilities and thus, greater uncertainty within the distribution. Therefore, entropy measures the uncertainty of a probability distribution.

Through entropy, one is able to quantify the information gained from changing a distribution. Assume there is a prior probability distribution $\mathbb{P}$ for the random variable X. By incorporating new information, a posterior distribution $\mathbb{Q}$ is formed.

5 See Cont (2001) for more information on stylised asset return features.

By considering the notion of relative entropy, one is able to quantify the reduction in uncertainty. Relative entropy, denoted by the function f, is written as

$$f(\mathbb{P},\mathbb{Q}) = \mathbb{E}_{\mathbb{Q}}[\ln \mathbb{Q} - \ln \mathbb{P}] = \sum_{i=1}^{n} q_i \ln \frac{q_i}{p_i} = -\sum_{i=1}^{n} q_i \ln \frac{p_i}{q_i}. \tag{5.6}$$

From Equation 5.6, we know that $f(\mathbb{P},\mathbb{Q})$ is convex. By using Jensen's equality, Zou and Derman (1999) show that $f(\mathbb{P},\mathbb{Q})$ is strictly non-negative and zero only for $\mathbb{P} \equiv \mathbb{Q}$. Using this fact, they motivate that relative entropy can be considered a 'distance' metric between prior and posterior distributions. Stutzer (1996) intimates that by minimising the relative entropy between prior and posterior distributions – in this case the real-world and risk-neutral density estimates – one preserves maximum uncertainty under the density transformation.

Estimating the Risk-Neutral Density from the Historical Density

This section describes the method used to transform $\hat{\mathbb{P}}$ to the estimated risk-neutral return distribution $\hat{\mathbb{Q}}$. Using the fact that $\hat{p} = 1/n$ and assuming a constant single-period, gross risk-free rate r, the no-arbitrage constraint given in Equation 2 can be simplified as

$$1 = \sum_{i=1}^{n} \left(\frac{R_i}{r^T}\right) \frac{q_i}{\hat{p}_i} \hat{p}_i, \tag{5.7}$$

where q_i is the risk-neutral probability of return R_i. Stutzer (1996) shows that the solution to minimising the relative entropy given in Equation 5.6 subject to the risk-neutral constraint given in Equation 5.7 is given by the following Gibbs canonical distribution:

$$\hat{q}_i = \frac{\exp\left(\gamma^* \frac{R_i}{r^T}\right)}{\sum_{i=1}^{n} \exp\left(\gamma^* \frac{R_i}{r^T}\right)}, \quad i = 1, \ldots, n \tag{5.8}$$

where γ^* is the Lagrange multiplier found by solving the following unconstrained minimisation problem,

$$\gamma^* = \underset{\gamma}{\operatorname{argmin}} \sum_{i=1}^{n} \exp\left[\gamma\left(\frac{R_i}{r^T} - 1\right)\right]. \tag{5.9}$$

Using the risk-neutral probabilities calculated in Equation 5.8, the discounted expected payoff of the derivative contract can be computed.[6] Thus, the price of a European call option with strike price K, expiring at time T is given by

$$C = \sum_{i=1}^{n} \left(\frac{\max(S_0 R_i - K, 0)}{r^T} \right) \hat{q}_i. \tag{5.10}$$

The implied CV volatility can then be solved for from the computed option price C. This imputed volatility is hereafter referred to as 'fair volatility'.

5.3 STATISTICAL PROPERTIES OF THE SOUTH AFRICAN MARKET

An analysis of volatility term-structure through CV is necessarily an empirical exercise. To illustrate applications of CV, we use the South African (SA) All-Share Index (ALSI) Top40 Total Return series dating from 30 June 1995 to 24 October 2011, giving 4085 daily observations.[7] Figure 15 displays the Top40 Total Return series. The risk-free interest rates over the same period are proxied by the 3-month South African Treasury-Bill rates.

Series of rolling, daily gross index returns for a period of N trading days are constructed by calculating

$$R_i = \frac{S_i}{S_{i-N}}. \tag{5.11}$$

Gross returns are calculated for 3-, 6-, 9- and 12-month periods assuming 252 trading days per annum.

A cursory study of Figure 15 reveals three distinct market periods within the given time period. The Top40 index level remained relatively constant from July 1995 until March 2003, after which the index displayed a strong bull market. This bull run was brought to an abrupt halt by the subprime mortgage crisis in August 2008. Given that the chosen historic period can be used as a mechanism allowing

6 This method can be easily extended to include multiple underlying assets and multiple constraints. Appendix D.1 provides a mathematical explanation of these extensions.

7 The total return series is used rather than the index level in order to endogenously include the effect of dividends.

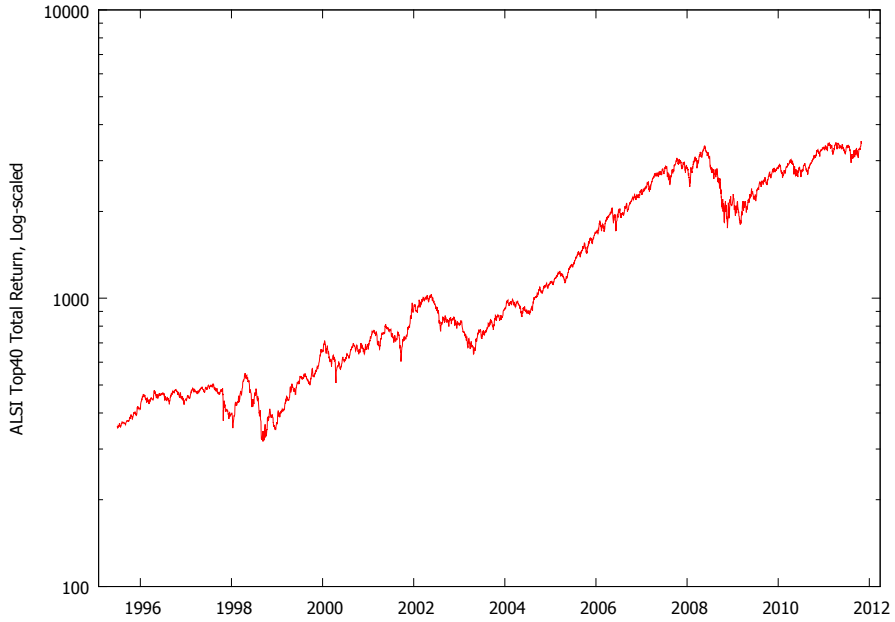


Figure 15: Top40 Total Return Series dating from 31 June 1995 to 31 October 2011

investors to better tailor CV to their respective future market views, and that market behaviour changes rather severely over time, we analyse and implement CV using index level data from three different starting dates.

These starting points are 30 June 1995, 31 March 2003, and 31 July 2008 respectively, motivated by the differing market periods displayed in Figure 15. These periods are referred to as June-95, Mar-03 and Aug-08 respectively. Although the three periods are not independent, the distributional and CV properties still differ widely over each period.

5.3.1 *Empirical and Fitted Historical Distributions*

Although the empirical return distributions generally differ over the historic periods, there are some shared characteristics. Table 16 gives descriptive statistics for the rolling returns for different terms over the three chosen periods. We highlight several observations from Table 16; some rather novel, others expected.

The Jun-95 mean returns are considerably lower than those for Mar-03, which reflects the dampening effect of the initial sideways market on the large bull mar-

Table 16: Descriptive statistics for rolling gross returns

Gross Return Statistics	Time to Expiration (months)			
	3	6	9	12
Mean (annual)	0.1704	0.1731	0.1714	0.1716
	0.2213	0.2296	0.2335	0.2337
	0.1484	0.1997	0.2232	0.2219
Standard Deviation (annual)	0.2258	0.2283	0.2292	0.2363
	0.1937	0.2101	0.2269	0.2400
	0.2001	0.1956	0.1662	0.1367
Skewness	-0.3683	-0.5561	-0.3773	-0.1974
	-1.2033	-1.4269	-1.1095	-0.9181
	-1.0610	-0.8093	-0.2372	0.1366
Excess Kurtosis	0.9885	0.2130	-0.3016	-0.2065
	2.1578	2.3956	1.0525	0.3782
	2.2875	1.0453	0.3032	0.2765

Note. Descriptive statistics are given for rolling, gross returns for different time periods calculated using each of the three index level periods. The first number in each cell corresponds to June-95 returns, the second to Mar-03 returns and the third to Aug-08 returns.

ket period. In contrast, the mean Aug-08 returns are initially very low and increase with term, reaching Mar-03 levels at 12-months. This is directly attributable to the subprime crash dominating shorter-term returns. In terms of standard deviation, we note the opposite trend over term for Aug-08, with a minimal standard deviation of 13.67 per cent for 12-month returns. This is in comparison to the mid-20 per cent values seen within the other return periods for the same term.

Looking at the higher-order moments, Top40 returns display a considerable negative skew in all but one case and are mostly leptokurtic, with Mar-03 returns generally displaying the largest skew and kurtosis. Similar behaviour is documented for the majority of financial markets (Cont (2001)).

Figures 16, 17 and 18 display the empirical return histograms and the best-fitting distributions chosen under maximum likelihood estimation.[8] The non-normality of Top40 return data is verified by the choices of the best-fitting distributions. The Weibull and Extreme Value distributions are prominent across return term and historic period.[9] This is because these distributions are able to capture the negative

8 Appendix D.2 displays the nonparametric distributions of the return data.

9 While one can argue that the Logistic family is also fairly prominent within the chosen best-fitting distributions, it is found that the likelihood values of those competing distributions not chosen are very close to that of the chosen Logistic. Thus, justifying the choice of Logistic distribution based only

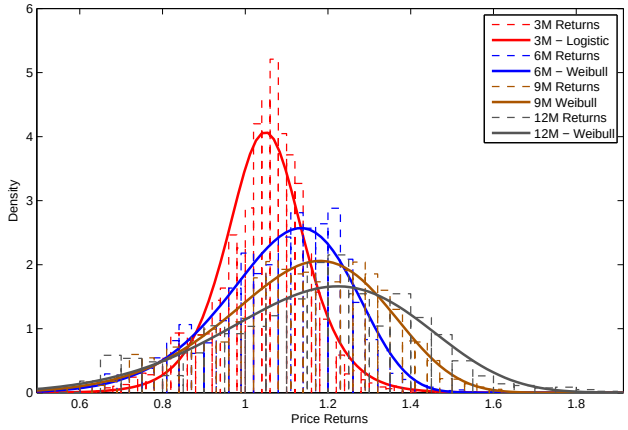


Figure 16: Empirical histograms and best-fitting distributions for June-95 Top40 returns.

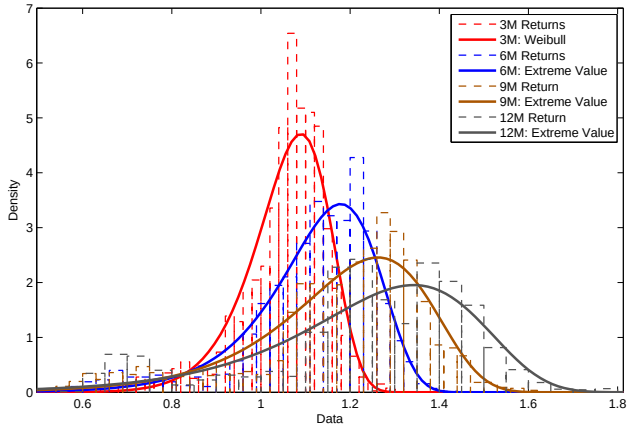


Figure 17: Empirical histograms and best-fitting distributions for Mar-03 Top40 returns.

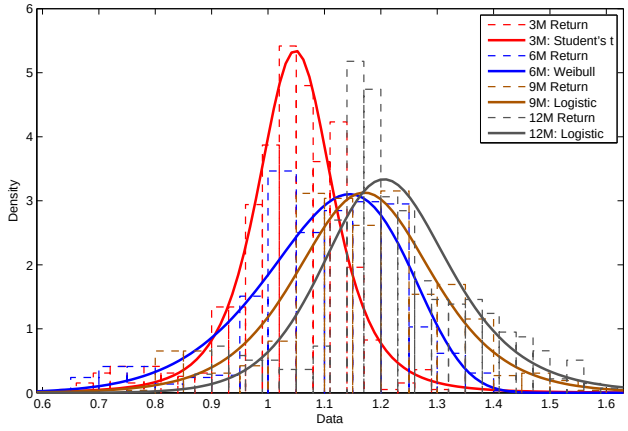


Figure 18: Empirical histograms and best-fitting distributions for Aug-08 Top40 returns.

skew and leptokurtosis inherent within the return series. Under the assumption of normality, one would expect to see a distinctive downwards and rightwards trend across distributions over term. However, this pattern is only really seen for the Mar-03 fitted distributions. The fitted Aug-08 distributions are fairly consistent across higher terms, with the 12-month distribution actually being more peaked than its mid-term counterparts. These departure from normality become even more pronounced when considering the empirical nonparametric distributions directly (see Appendix D.2) and it is exactly these features that CV attempts to capture.

5.4 TERM-ADJUSTED-SPREAD AS A MEASURE OF VOLATILITY TERM-STRUCTURE

Since the advent of implied volatility surface research, equity markets have displayed an irregular volatility term-structure. TAS attempts to answer whether the current volatility term-structure is justified by historical returns. We make use of implied volatility data obtained from at-the-money ALSI Top40 index options with terms of 3-, 6-, 9- and 12-months respectively. The volatility data is measured daily for the period 1 June 2011 to 24 October 2011, covering 102 trading days. Before calculating TAS, the implied volatility term-structure is briefly discussed.

5.4.1 *Market-Imposed Volatility Term-structure*

According to the Black-Scholes framework, volatility is constant over strike and term. However, this has been noted on many occasions to disagree with empirical evidence. While the usual means of presenting an implied volatility surface is against option strikes and terms at a fixed time point, another approach is to hold either strike or term constant and show the evolution of volatility over time. In this case, we consider the evolution of the volatility term-structure over time and

upon maximum log-likelihood becomes much more difficult. This is not the case for the best-fitting Weibull and EV distributions. Thus, we focus our attention on these two distributions and leave further treatment of the Logistic distribution to a later stage.

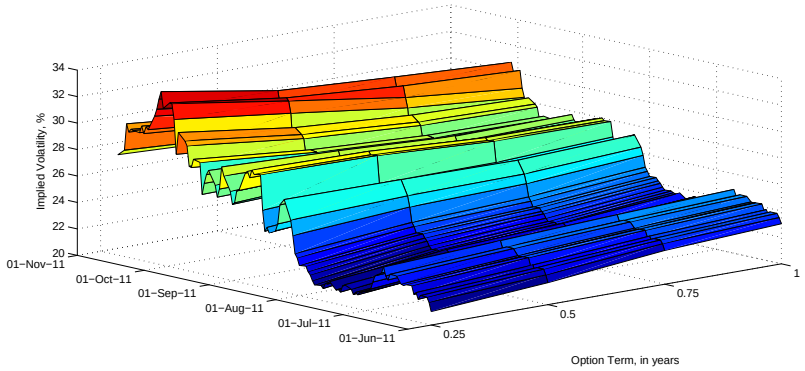


Figure 19: Implied volatilities of ALSI Top40 index options, given for quarterly incremental option terms, measured over the period 1 June 2011 until 24 October 2011.

fix the strike daily to be at-the-money. Figure 19 exhibits the implied volatilities of ALSI Top40 index options plotted against option term and time.[10] Three key observations are apparent: (i) the volatility series of each option term changes rather substantially over time, (ii), the volatility across each term is not constant and this relationship changes with time, and (iii), the volatility series of each option term are highly correlated. While point (ii) is also displayed on the more common strike-term-volatility surface, observations (i) and (iii) remain hidden. It is these two issues that are studied here.

We consider the spread between the volatility series' of each option term. Figure 20 plots this spread – or term-structure – over time for the four chosen option terms. The spread lines are grouped according to the month-difference in option terms. It is clear that the term-structure of volatility over time is far from constant for any of the given spreads. As one would expect, the largest spread is between 12-month and 3-month (9-month term difference) options, ranging approximately 6 per cent. The spread between options with a 6-month difference is slightly smaller, displaying an average range of 4 per cent, while the smallest spread is seen between options with a 3-month term difference, with an average range of 2 per cent. Volatility term-structure is thus very material, in both a theoretical and practical derivative context. A further observation, again to be expected, is the high correlation between all volatility spread series; the lowest correlation is roughly 0.8.

10 ALSI Top40 implied volatility data was kindly provided by Old Mutual Investment Group (SA).

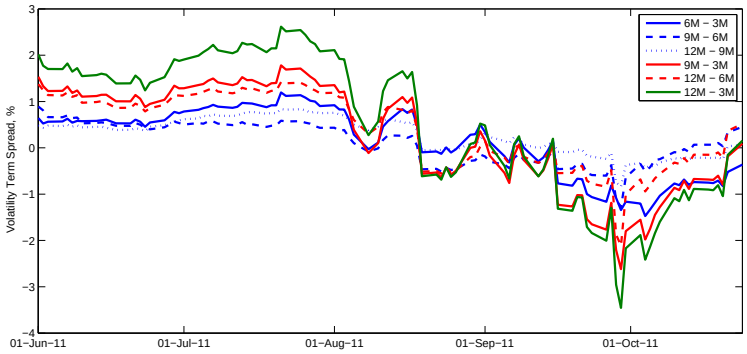


Figure 20: Term-structure of implied volatility for ALSI Top40 options over the period 1 June 2011 until 24 October 2011.

Note. The structure is given as the spread between the volatility series calculated from options with terms 3-, 6-, 9-, and 12-months respectively. The line colour gives the difference in option terms used to calculate the spread. Blue denotes a 3-month difference, red a 6-month difference, and green a 9-month difference.

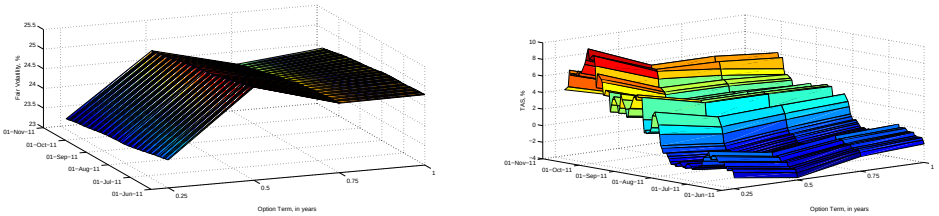
Finally, one notes that large daily movements in the volatility term-structure occur relatively frequently. This indirectly adds credence to the question of whether the volatility term-structure imposed by the current market prices is fair. In order to evaluate this question, we now turn to TAS.

5.4.2 *Assessing Volatility Term-structure through TAS*

The Term-Adjusted-Spread for an option with fixed strike and expiration T at time s is defined as

$$TAS(T,s) = \Sigma(T,s) - \Sigma_{CV}(T,s), \tag{5.12}$$

where $\Sigma(T,s)$ is the Black-Scholes implied volatility of the option, and $\Sigma_{CV}(T,s)$ is the implied volatility calculated from the CV method using some chosen historical period (referred to as 'fair volatility'). Thus $TAS(T,s)$ is a measure of the implied volatility spread as a function of option term T and time s at a fixed strike. Using the June-95 returns and a risk-free rate of 6 per cent, Figures 21a and 21b illustrates



(a) (b)

Figure 21: (a) Fair volatility computed via CV from Jun-95 return series and (b) the resultant TAS, for the period 1 June 2011 until 24 October 2011.

the fair volatility and TAS surfaces respectively, graphed against option term and time.[11]

Interestingly, the fair volatility surface remains very consistent for each term over the time period analysed. This shape is very different to that shown in Figure 19. However, when considering the shape in conjunction with the findings of Section 5.3.1 (and Appendix D.2), which illustrates a distinct relationship between the respective return series distributions, the shape of the fair volatility surface is perhaps better understood. Both the volatility range between option terms and over time is greatly reduced. In addition, the volatility difference is most pronounced between shorter-term options and decreases quite drastically over term. Taken in conjunction, these feature lead to a TAS surface that is mostly driven by the market implied volatilities. Looking back to early June 2011, we note that TAS over all option terms is strongly negative and increases over the period considered, ending at a rather large positive level. The TAS surface thus displays considerable movement.

While it is possible to compare TAS levels over time directly, it is perhaps better to consider the volatility slope across term rather than the actual values. This is due to the stylised fact that the volatility smile is more stable than the volatility levels (Tompkins (2001), Kotzé and Joseph (2009)). Considering the TAS surface from this perspective highlights the fact that 9- and 12-month option contracts seem to be equally mispriced whereas the shorter 3- and 6-month options display

11 While one can use historical risk-free rates, the majority of CV studies normally assume a fixed rate. Appendix D.3 considers the sensitivity of fair volatility, and thus TAS, to this assumption.

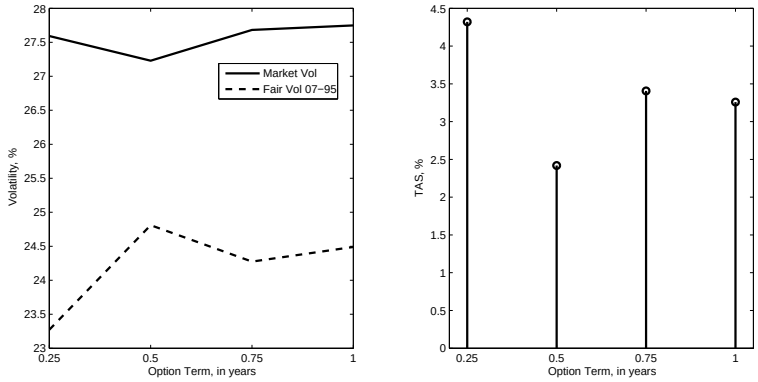


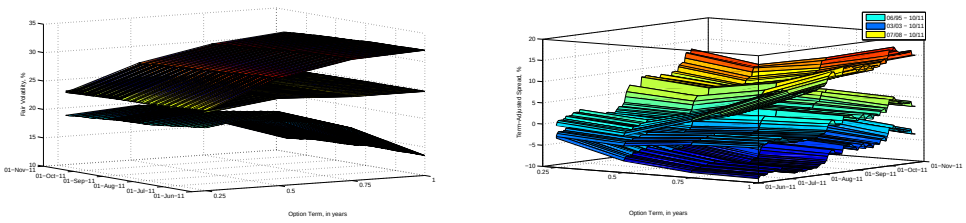
Figure 22: Fair volatility computed via CV from Jun-95 return series and the resultant TAS, for the date 24 October 2011.

different mispricings.[12] This relative value information becomes a practical tool that investors can use to rank options across term. When taken over time as in Figure 21b, this tool can quantitatively measure the market's changing opinion on the volatility term-structure.

One can also consider option richness at a specific date. Figure 22 shows a plot of the fair and market volatilities across term on 24 October 2011, and the corresponding TAS values. All options seem much too rich, with 3-month options being particularly overpriced. Comparatively, 6-month – and the longer-termed – options seem to be better priced. Thus, one could sell short-term options and buy long-term options in an attempt to profit from the TAS spread. Using Figure 21b, one then has a means of tracking the spread and creating some type of buy/sell signal based upon its evolution. TAS can thus provide investors with tradable information regarding the term-structure of volatility, provided that the imputed risk-neutral historical distribution is appropriate. As previously alluded to, Σ_{CV} and thus TAS is dependent on the chosen historic return period, both in terms of underlying asset returns and risk-free interest rates. This dependency imposes an element of investor opinion in the sense of choosing the historic period most relevant to the current market. What occurs when one changes this historic period?[13]

12 This fact follows from the similarities and differences of the June-95 return series distributions respectively.

13 Another important question is: how sensitive is TAS to a constant risk-free rate assumption? It turns out that the effect of the risk-free rate is actually rather trivial. Appendix D.3 provides a detailed discussion on the matter.



(a) (b)

Figure 23: (a) Fair volatility computed via CV using the June-95, Mar-03 and Aug-08 return series and (b) the resultant TAS surfaces, for the period 1 June 2011 until 24 October 2011.

TAS under Different Historical Periods

Using the returns generated from the three historic period defined in Section 5.3.1 – namely, June-95, Mar-03 and Aug-08 – one can construct three distinct estimates of fair volatility and thus TAS. Figures 23a and 23b plot the fair volatility and TAS surfaces computed using each historic period respectively. While the respective fair volatility surfaces all display highly consistent values for each term over time, when compared across the three surfaces, these values and slopes are remarkably divergent, especially for longer-term options.

Once again, it is imperative to consider the fair volatility surfaces computed from the different historic input data in conjunction with the distributional qualities of each period, highlighted in Section 5.3.1. Figure 23a shows that the Mar-03 3-month fair volatility is roughly equal to its June-95 counterpart but is much higher for all other option terms. This is driven by the similarity of the 3-month June-95 and Mar-03 return distributions and the subsequent distributional divergence of the longer-term return series. Aug-08 fair volatility displays a very contrasting picture, being much lower than the June-95 surface. The difference between these two fair volatility surfaces also increases over term. This is again explained by the differences in empirical distributions. Not only are the Aug-08 return distributions flatter than the June-95, they also have much lower standard deviations. This results in the lower fair volatility surface. By definition, the TAS surfaces are switched in comparison to fair volatility so that the Aug-08 surface is now above the June-95 surface, while the Mar-03 surface now appears below. As with fair volatility, the imputed TAS surfaces differ severely over both shape and range of

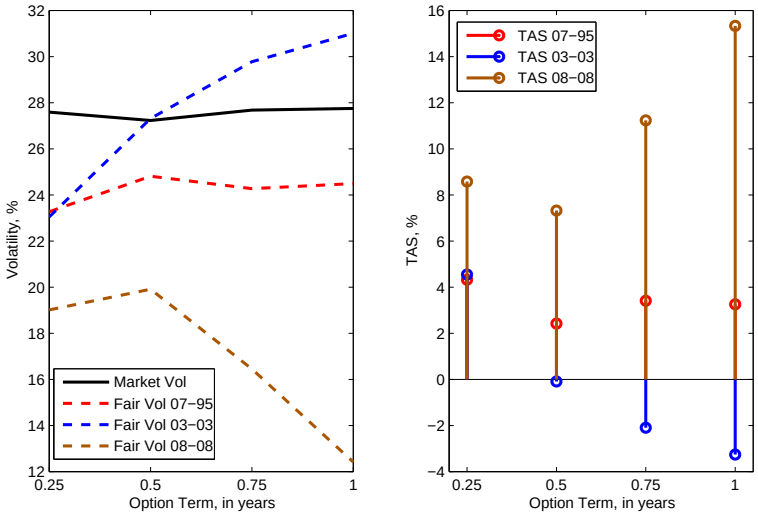


Figure 24: Fair volatility computed via CV from Jun-95 return series and the resultant TAS, for the date 24 October 2011.

values. Let us consider the fair volatility, market volatility and TAS levels of the three return periods as at October 24 2011, given in Figure 24.

Section 5.4.2 described how an investor could potentially use daily TAS values in conjunction with the TAS surface to decide on which options to buy and sell, subsequently evaluate and monitor the strategy over time, and finally decide when to close out the position.

However, all three of these aspects are severely affected by the investor's choice of the historical period relevant to the current market. The severity of this choice is shown in Figure 24b. For example, if the appropriate historical period was actually given by the Aug-08 returns, all options would appear much too expensive with longer-term options showing the highest mispricing. In this case, TAS would lead an investor to consider different strategies to that outlined above in Section 5.4.2. A pure speculator would consider selling long-term options in order to profit on the direct mispricing, while the more risk-averse investor may consider buying short-term options and selling long-term options in order to take advantage of the relative mispricing. A similar analysis of the Mar-03 TAS surface would yield another strategy. Should an investor use TAS calculated under a single historic period, relying on their choice of period to be the most appropriate? Or is it more appropriate to analyse TAS values under a range of different historic periods?

Although a proper treatment of these questions is left for future research, a potential approach to analysing TAS sensibly under multiple historic periods would be to consider a weighted average TAS function with the weighting scheme based upon the investor's opinion. Another potential approach is discussed below.

5.5 RETURN DISTRIBUTION FORECASTS AND TERM-ADJUSTED-SPREAD

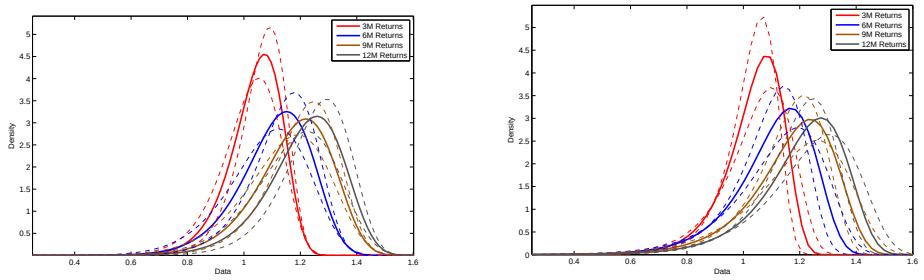
The Weibull and EV families dominate the distributions fitted to the three historic periods under review. Using this insight, we therefore fit the Weibull and Extreme Value distribution families to historic ALSI Top40 total return data at three-month intervals through the following process:

1. Starting from 24 October 2011, gross 3-, 6-, 9- and 12-month returns are bootstrapped from the prior three-year period rolling returns using a sampling effort of 30 per cent.
2. The distributions are fitted to these bootstrapped return series and the parameters are recorded for the chosen date.
3. Moving back through time by three months, Steps 1 and 2 are repeated.

Although one should essentially use independent return periods upon which to fit the distributions, this yields only a handful of parameter estimates. Building a time-series model of the distributional parameters with reasonable forecasting capabilities necessitates rolling time periods. By fitting distributions at three-month intervals to the historic ALSI Top40 data available, one calculates a total of 163 location, scale and shape parameters respectively of the relevant distributions.

The best-fitting ARIMA(r, d, m) models are then chosen for each distribution parameter respectively for each return term. Where necessary, the innovations of the parameters are modelled by a GARCH(p, q) process.[14] The models are limited to $r, m, p, q \leqslant 2$ and $d \leqslant 1$. Appendix D.4 outlines the model selection process and displays the chosen model and its estimated parameters. Using the fitted models, one is able to forecast the future return distributions. Sampling from the forecast

14 GARCH refers to Generalised Auto-regressive Conditional Heteroscedasticy.



(a) Weibull distributions

(b) Extreme Value distributions

Figure 25: Forecast distributions of future returns as at 24 October 2011.

distributions in place of the historical return distributions, one can then apply the CV framework and calculate both the fair volatility and TAS.[15]

As noted in Section 5.1, this approach can only be considered semiparametric. However, it does have several potential advantages to its truly nonparametric counterpart. Firstly, option traders would have a forecast of the future asset return distribution intimated only by the historical data. The trader would be able to both define the period of the historical data, the general distribution family fitted and the time-series models fitted to the distribution parameters. This provides one with greater flexibility to alter the forecast asset return distribution and thus, potentially add value. Secondly, one should be able incorporate early exercise fairly easily and thus price American options. In fact, the idea of modelling time-dependent distribution parameters via time-series models is notionally fairly close to the American option extension proposed by Duan (2002).

Figures 25a and 25b plot the forecast Weibull and Extreme Value return distributions respectively for the different return terms considered. The dashed lines display the distributions calculated using the 95% lower and upper parameter bounds. The Weibull and EV forecast return distributions are remarkably similar across all return terms, with both figures capturing the prevalent negative skew inherent in real-world data. As noted in Section 5.3.1, the assumption of normality dictates rightward-moving and decreasing probability density maxima over return term. One sees that Figures 25a and 25b only partially follow this pattern. The

15 A scrambled Sobol sequence is used to sample 50 0000 random returns from each fitted distribution respectively.

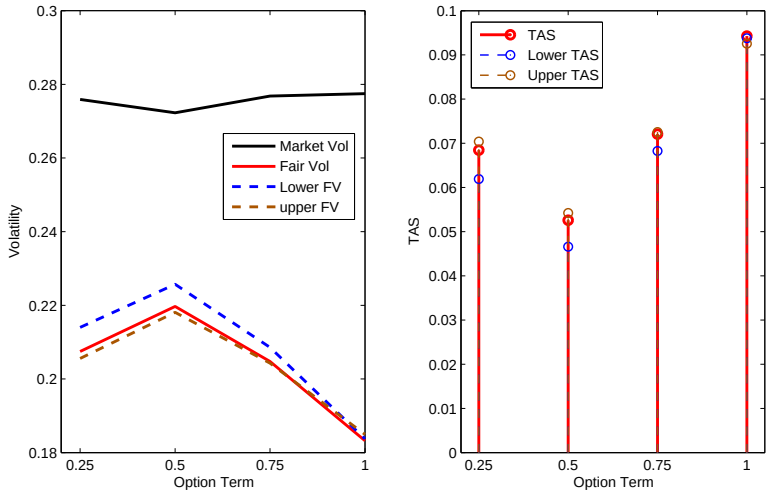


Figure 26: (a) Fair volatility computed via CV from forecast Weibull return distributions and (b) the resultant TAS, for the date 24 October 2011.

density function maxima for 6-, 9- and 12-month returns are approximately equivalent and the horizontal distance between density functions decreases with return term. Comparing these forecast distributions with previous fitted distributions in Section 5.3.1, we find that the general evolution of the distributions over term is most similar to the Aug-08 return distributions. However, the actual values of the forecast density functions are rather different to the respective fitted values. This leads to marked differences in fair volatilities and TAS.

Figures 26 and 27 plot fair volatility and TAS under the forecast Weibull and EV distributions respectively. As before, the dashed lines represent the fair volatility and TAS values calculated using the 95% lower and upper parameter bounds. Interestingly, one now sees a strong divergence between the fair volatilities calculated under the Weibull and EV forecast distributions. Fair volatility under the Weibull forecast returns series is much lower than its respective EV counterpart and evolves very differently over return term. In addition, the Weibull distribution 95% bounds are very tight and show increasing convergence over term, whereas the EV distribution interval remains fairly large and constant across term. As in previous sections, we argue that the TAS shape is perhaps more pertinent than the given values. Under the Weibull forecasts, options become relatively more overpriced across term, with the exception of 3-month options. This TAS evolution

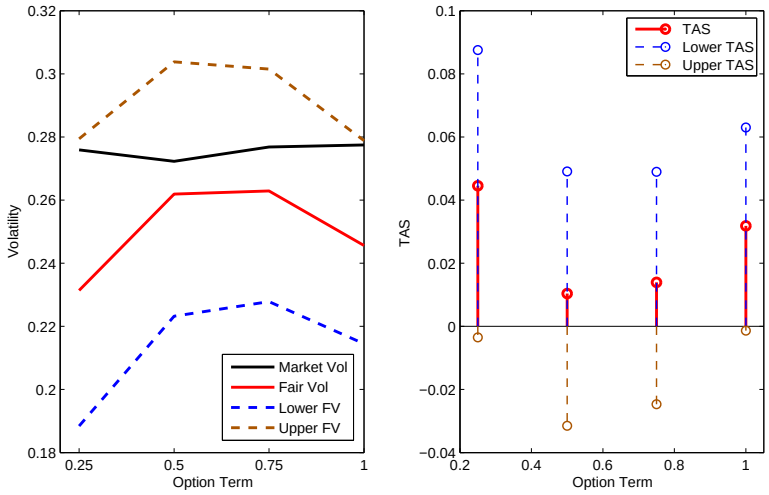


Figure 27: (a) Fair volatility computed via CV from forecast EV return distributions and (b) the resultant TAS, for the date 24 October 2011.

is equivalent to that calculated using Aug-08 returns, although the values are somewhat lower. One should therefore again consider shorting 12-month options and longing 6-month options. Contrastingly, TAS under the EV forecasts shows that relative richness only marginally increases with longer terms and TAS actually is maximised for the 3-month term. Comparing this to the results given in Section 5.3.1, we find some similarity in shape to the Jun-95 TAS function and note that, in this case, the magnitude of TAS values is approximately equal. If one assumed that the future return distribution was more accurately characterised by the EV forecast, then one should consider shorting 3-month options and longing 6-month options. Given the 'closeness' of the Weibull and EV forecast return distributions, the material difference in both TAS values and evolution is rather difficult to understand. We defer a proper mathematical treatment of this characteristic to a later stage and provide here a heuristic explanation. By introducing the risk-neutrality constraint, one is constraining the distribution mean to be equivalent to the risk-free rate. However, the Weibull and EV mean functions describe very different relationships between the respective underlying distribution parameters. Thus, while minimising relative entropy ensures that the minimum amount of uncertainty is lost across the distribution transformation, the underlying mechanics are still dependent upon the mathematics of each distribution. One should

therefore expect the normalisation process to manifest itself differently given different distribution moment generating functions.

5.6 CONCLUSIONS

The Black-Scholes framework implies a constant volatility across term and strike. However, it is known that empirical data violates this assumption. In this chapter, we consider whether the current market implied volatility term-structure is justified by that market's historical performance. Stutzer's (1996) nonparametric CV framework is used in order to develop a tool with which to rank the relative richness of at-the-money options of different terms over time. Stutzer's nonparametric theory suggests using a risk-neutralised historical distribution as the underlyer, which is found by minimising the relative entropy placed on the historical distribution by the constraint of risk-neutrality.

An analysis of the historic and current South African All-Share Index Top40 market is undertaken with respect to both underlying returns and volatility. The distributional properties of returns measured under different terms differ greatly. In addition, the choice of historic period also affects the distributional characteristics. Three historic periods are chosen based on events within the South African market. The return distributions across these periods show large variations. However, prevalent within all distributions is a strong negative skew, leptokurtosis and general asymmetry about the distribution. These stylised facts are evidenced by the best-fitting distributions being dominated by the Weibull and Extreme Value distribution families.

We find that the TOP40 implied volatility is not constant. For each option term, the implied volatility changes on a daily basis and sometimes quite extremely so. In addition, the volatility spread between options with small term-differences is smaller than the spread between options with larger term-differences.

The idea of Term-Adjusted-Spread, an adaptation of Zou and Derman's (1999) Strike-Adjusted-Spread, is introduced. TAS is the difference between an options implied volatility and its fair volatility as estimated by CV. TAS values enable

an investor to relatively value options over different terms, monitor this relation and potentially provide an investor with a closeout signal. This measure cannot be used without some discretion; depending on the investor's selection of the historical period, TAS values and slopes can differ greatly.

In addition to using the nonparametric, risk-neutralised historic distribution, there are also several semi-parametric distributions that one can choose. A subset considered here are time-series forecasts of Weibull and Extreme Value distributions fitted to the historical market data. Although the forecast density functions across each distribution family display equivalence, the imputed fair volatility and TAS values differs markedly. Again, discretion is required when choosing the appropriate distribution. The semi-parametric CV method provides several potential benefits: one has increased flexibility and control over the CV framework, and there is a potential to price American options.

Several extensions of this research include: (i) creating and back-testing practical option strategies based on TAS; (ii) employing variance reduction techniques in order to reduce and/or optimise historic period choice; (iii) studying the effects of additional option constraints; (iv) refining the distribution-forecasting-CV method; and (v) running similar analyses on exotic options.

Part III

APPENDIX

A

HM MODEL APPENDICES

A.1 SUMMARY OF MUTUAL FUND PERFORMANCE STUDIES 1962 - 1991

The seminal fund performance literature on market timing and stock selection ability prior to 1991 is concisely summarised by Ippolito (1993):

Study	Year	Period Covered	No. Funds	Type of Fund	Model	Survivor Bias	Market Index	Average Alpha (b.p.s/yr.)	t-Value	Evidence of Mkt. Timing
Friend, Brown, Herman & Vickers[a]	'62	1953–58	152	All	Index	Yes	S&P Comp.	−20	n/r	?
Treynor & Mazuy	'66	1953–62	57	All	CAPM-NL	Yes	S&P 500??	n/r	no	
Sharpe[b]	'66	1954–63	34	All	CAPM-RV	Yes	Dow-Jones	−34	2.42	n/t
Jensen[c]	'68	1945–64	115	All	CAPM	Yes	S&P 500	−110	5.63	n/t
Friend et al.[d]	'70	1/60–6/68	136	All	CAPM	Yes	VW-NYSE	217	n/r	n/t
					EW-NYSE			22		
Carlson		1948–67	82	Stock	CAPM	Yes	S&P 500	60	n/r	n/t
					CAPM-RV		Dow-Jones	14	11.38	
McDonald	'74	1960–69	123	All	CAPM	Yes	EW-NYSE	62	n/r	n/t
Mains	'77	1955–64	70	All	CAPM	Yes	S&P 500	9	n/r	n/t
Kon & Jen	'79	1960–71	49	All	CAPM-NL	Yes	EW-CRSP	6	n/r	n/t
					CAPM-B			−67		
Shawky[e]	'82	1973–77	255	All	CAPM	Yes	EW-NYSE	−43	1.16	n/t
Alexander & Stover[e]	'80	1966–71	49	All	CAPM-NL	Yes	VW-CRSP	120	1.75	No
Veit & Cheney	'82	1944–78	74	All	CAPM	Yes	S&P 500	103	n/r	No
Kon[f]	'83	1960–6/76	37	All	CAPM-NL	Yes	VW-CRSP	739	2.87	Yes
Chang & Lewellen[e]	'84	1971–79	67	All	CAPM	Yes	VW-CRSP	58	0.75	No
					CAPM-NL			139	2.14	
Henriksson[e]	'84	2/68–6/80	116	All	CAPM	Yes	VW-NYSE	−24	0.80	No
					CAPM-NL			84	1.89	
Berkowitz et al.	'88	76Q1–83Q4	325	All	CAPM	No	S&P 500	68	n/r	n/t
Lee & Rahman	'90	1/77–3/84	93	All	CAPM	Yes	VW-CRSP	−60	n/r	Yes
					CAPM-NL			72		
Lehman & Modest[e]	'87	1968–72	130	All	CAPM-NL	Yes	VW-CRSP	−141	3.68	n/c
		1973–77						−79	1.98	
		1978–82						140	4.01	
		1968–72			APT			−485	14.34	
		1973–77						−545	17.30	
		1978–82						−385	13.32	
Grinblatt & Titman[g]	'89	1975–84	157	Stock	CAPM	No	VW-CRSP	−60	0.76	n/t
							P8 PORT	60	0.61	
Ippolito	'89	1965–84	143	All	CAPM	No	S&P 500	81	4.01	Yes
							VW-NYSE	87	4.20	
Ippolito revised[h]							S&P 500	40	2.19	
							VW-NYSE	51	2.75	

* n/r = not reported; n/t = not tested; n/c = no conclusion; CAPM-NL = nonlinear capital asset pricing model; CAPM-RV = CAPM using a reward-to-volatility ratio; CAPM-B = Black version CAPM; VW = value weighted; EW = equally weighted.

[a] Friend et al. predates CAPM, but it uses a method that effectively adjusts for beta differences. The alpha in the table is the difference between the annual rate of return for the authors' mutual fund sample minus the return on their benchmark portfolio over the period.

[b] Sharpe reports the reward-to-volatility ratio over the period as 0.667 for the Dow Jones index. The average of the ratio for the mutual fund sample is 0.663. I report the difference in terms of basis points. Sharpe did not report the standard error on the mean of the reward-to-variability ratio for his sample of mutual funds. But he listed the ratio for each fund. I calculated the mean of his data to be 0.662, with a standard error of the mean of 0.014. Thus the difference between the ratio for the Dow Jones index (0.667) and the mean of the funds is statistically different at the 95% level of confidence. The t-value in the table is this difference divided by the standard error of the mean.

[c] Jensen had data for 56 funds for the 1945–64 period and 59 funds for the 1955–64 period. He did not report a t-statistic, but listed all his estimated alphas. I calculated the mean and standard error of the mean from these data. The t-statistic is the mean alpha divided by its standard error.

[d] I inferred the alpha in the following way. I recognized that from the market equation, alpha equals (R-M) + (1-Beta)(R-F), where R is the average return for their mutual fund sample, M the return for the market index and F the risk-free return. The value of F over the 1960–68 period (proxied by the three-month T-bill) is 3.65. Friend et al. report that the value of R is 10.7% and that M is 9.9% or 12.4% depending on whether an equally weighted or value-weighted market index is used. They did not report their average beta, but they reported the mean beta within fund categories, together with the number of funds in each category. From this information, the unweighted average beta for the sample as a whole can be calculated. The overall average beta turned out to be 0.78. They reported that the betas were roughly the same regardless of whether they used an equally weighted or a value-weighted NYSE index. The average alpha is solved from the above equation using this value of beta.

[e] t-value equals the average alpha divided by the ratio of the standard deviation of estimated alphas for each fund to the square root of the number of funds.

[f] Kon did not report an average alpha and t-value but did report all 37 alphas for his funds, holding constant market-timing effects. I calculated the average alpha and its standard error from these data.

[g] The strong test was also performed on the data. Using returns *gross* of fees and expenses, and based on the so-called eight-portfolio benchmark, the estimated alpha is 0.0015 per month (t = 2.53). (Using the value-weighted CRSP data, the results are 0.0019 (t = 2.46).) The t-statistics are based on a time-series regression against the benchmark using an equally weighted portfolio for all the funds.

[h] In reproducing the Ippolito results, Elton et al. (1993) reentered the data from the original source and discovered some errors. The revised results reflect the corrected data.

A.2 TRACKING ERROR EFFECTS OVER THE DIA RANGE ON h_0 PERCENTAGE LEVELS

In Section 2.4.3, it is shown that the DIA breakdown level increases with tracking error. What is perhaps more useful though, is to consider the percentages shown

in hypothesis level H_0 over the entire DIA range, categorised by sigma range. This allows one to see how the masking effect of the noise process manifests itself. Figure 28 represents the percentage within hypothesis level H_0 for each DIA value over each sigma range. The colours within the graph represents different H_0 20 percentile ranges respectively. We notice that the overall masking effect from the increasing sigma range does not support a similar linear change on the H_0 percentage level over the DIA range. That is, the vertical space between each section in Figure 28 is not constant. It depends both on the DIA value as well as the sigma range as to how large this effect is. At higher DIA values, larger sigma values are required in order to see an H_0 percentage category shift than for lower levels. Again, this is logical as larger differences in true alpha values would need a much more volatile fund return series in order to be sufficiently masked; and *vice versa* for lower DIA levels.

A.3 EFFECT OF ASYMMETRIC MOVEMENTS IN TRUE ALPHAS

As mentioned previously in Section 2.4.1, $\alpha_p^* \approx \mathbb{E}[\mathrm{Int}]$. While all previous DIA changes were symmetrical movements of α_1 and α_2 about 0, which implies that $\alpha_p^* \approx 0$, asymmetrical movements allow the HM model's alpha coefficient to move away from zero as well. This enhances the HM model's ability to accurately detect (but not quantify) both up- and down-side selection due to the algebraic construction of the selection detection hypothesis, given below:

$$\left\{\left[\left(\alpha_p^* = 0\right) \wedge \left(\mathrm{Int}_i = 0\right)\right] \vee \left[\left(\alpha_p^* \neq 0\right) \wedge \left(\mathrm{Int}_i \neq 0\right)\right]\right\} \wedge \left(\alpha_p^* \neq \mathrm{Int}_i\right) \quad \text{for } i = \{1, 2\}, \tag{A.1}$$

where '=' can be read as 'statistically indistinguishable from', and '≠' as 'significantly different from'.

For H_{5a}, Equation A.1 holds for both $i = 1$ and $i = 2$, while H_{5b} and H_{5c} correspond only to $i = 1$ and $i = 2$ respectively. Thus, because the HM alpha estimate, α_p^*, is not equal to zero for most of the random funds, it is able to detect

selection quite well for the given DIA range. A final point to notice from Figure 5 (Section 2.4.4) is that the percentage levels within H_7 and H_9 increase for the uppermost DIA values. However, this is to be expected due to the fact that for an alpha range symmetrical about zero, very high DIA values will only be seen if $\alpha_1 \approx -\alpha_2$. This in turn means that $\alpha_p^* \approx 0$. Therefore, one would expect for any range tested that the uppermost DIA values would have high percentages within levels H_7 and H_9.

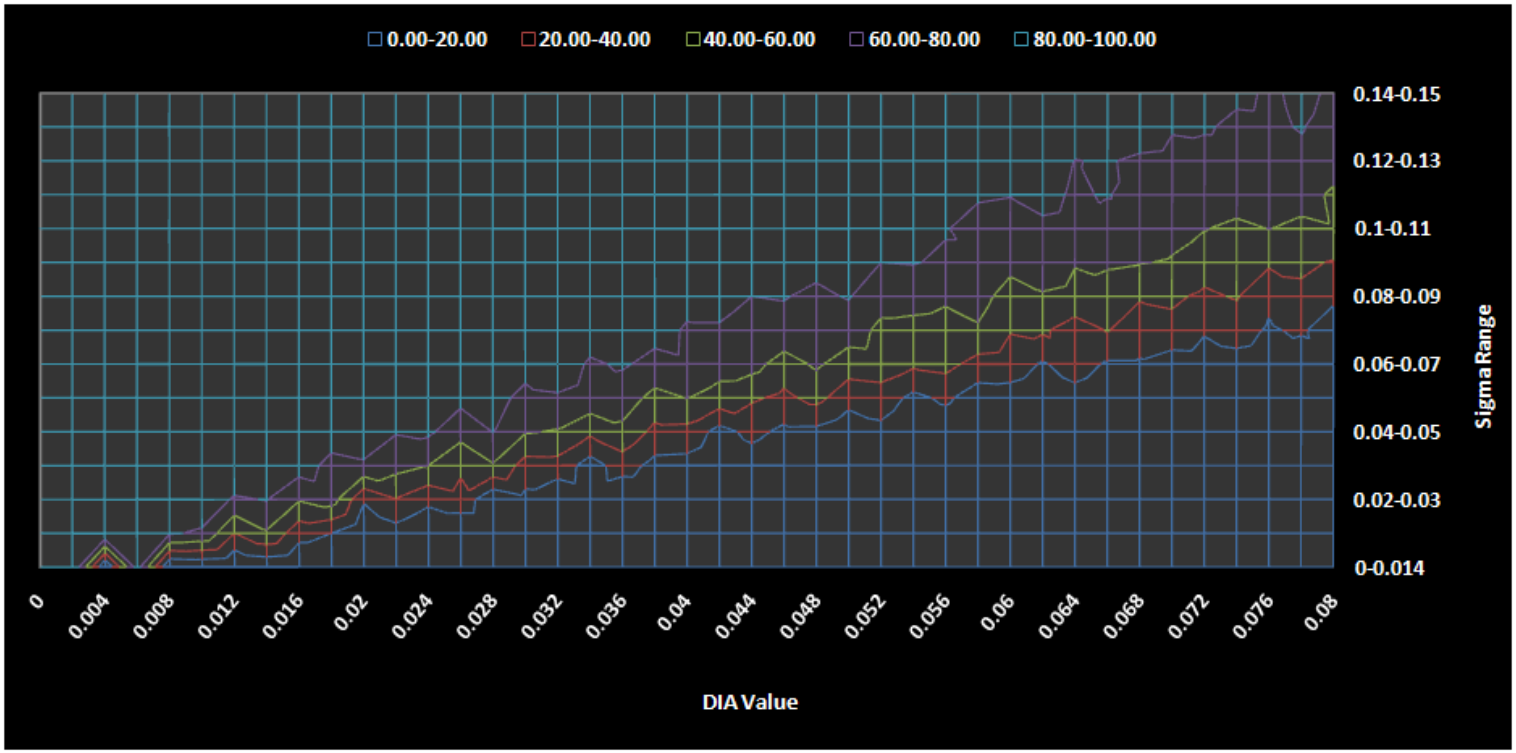


Figure 28: H_0 percentage levels over the sigma range, categorised by DIA value.

Note. Each colour within the graph represents different H_0 20 percentile ranges. For example, if one looks at the uppermost row of squares along the graph, one would say that for a sigma range of $0.14 - 0.15$, the HM model percentage within H_0 is greater than 80 per cent for the DIA range $0 - 0.74$. Alternatively, if one considers the rightmost column, it can be seen that for a 0.08 DIA value, the effect of increasing the tracking error causes the H_0 percentage level to exceed 20 per cent for all sigma values over approximately 0.076. The overall masking effect from the increasing sigma range does not support a similar linear change on the efficacy of the HM model - in terms of H_0 percentage level - over the DIA range. It depends both on the DIA value as well as the sigma range as to how large this effect is. At higher DIA values, larger sigma values are required in order to see an H_0 percentage category shift than for lower levels.

B

CSV ALGEBRA APPENDICES

B.1 PROOFS OF CSV THEOREMS

Motivation of Equation 3.7:

If $N_t = 2$, then we have

$$w_{1t} + w_{2t} = 1,$$

which implies that

$$r_{mt} = w_{1t} r_{1t} + w_{2t} r_{2t}.$$

Using the expressions above, we find that

$$\begin{aligned}
\chi^2_{1,2}(r_t) &= \sum_{i=1}^{2} w_{it} (r_{it} - r_{mt})^2 \\
&= w_{1t} (r_{1t} - r_{mt})^2 + w_{2t} (r_{2t} - r_{mt})^2 \\
&= w_{1t} [(1 - w_{1t}) r_{1t} - w_{2t} r_{2t}]^2 + w_{2t} [(1 - w_{2t}) r_{2t} - w_{1t} r_{1t}]^2 \\
&= w_{1t} [w_{2t} (r_{1t} - r_{2t})]^2 + w_{2t} [w_{1t} (r_{1t} - r_{2t})]^2 \\
&= w_{1t} w_{2t} (w_{1t} + w_{2t}) (r_{1t} - r_{2t})^2 \\
&= w_{1t} w_{2t} (r_{1t} - r_{2t})^2 .
\end{aligned}$$

For $N_t > 2$ and stocks i and j, we have that $w_{it} + w_{jt} \neq 1$. Thus pairwise squared CSV is defined for the general N_t-asset case as

$$\chi^2_{ij}(r_t) = \left[w_{it} w_{jt} \left(w_{it} + w_{jt} \right) \right] \left(r_{it} - r_{jt} \right)^2 . \qquad \square$$

Proof. Equation 3.8:

For the sake of brevity, the documenting of standard algebraic simplifications is abridged. The t-subscripts are discarded for notational ease.

If one considers a N_t-asset market, one has the following:

$$\begin{aligned} r_m &\equiv w_1 r_1 + w_2 r_2 + \ldots + w_N r_N \\ 1 &\equiv w_1 + w_2 + \ldots + w_N. \end{aligned}$$

Using these identities and letting $M < N$, squared CSV is written as

$$\begin{aligned} {}_N\chi^2(r) &= \sum_{i=1}^{N} w_i (r_i - r_m)^2 \\ &= w_1 [w_2 (r_1 - r_2) + w_3 (r_1 - r_3) + \ldots + w_N (r_1 - r_N)]^2 \\ &\quad \ldots \\ &\quad + w_M [w_1 (r_M - r_1) + \ldots w_{M-1} (r_M - r_{M-1}) + w_{M+1} (r_M - r_{M+1}) + \ldots + w_N (r_M - r_N)]^2 \\ &\quad \ldots \\ &\quad + w_N [w_1 (r_N - r_1) + w_2 (r_N - r_2) + \ldots + w_{N-1} (r_N - r_{N-1})]^2. \end{aligned}$$

Applying standard algebraic operations, we simplify the expression above:

$$\begin{aligned} {}_N\chi^2(r) &= (r_1 - r_2)^2 [w_1 w_2 (w_1 + w_2)] + (r_1 - r_3)^2 [w_1 w_3 (w_1 + w_3)] \\ &\quad + \cdots + (r_{N-1} - r_N)^2 [w_{N-1} w_N (w_{N-1} + w_N)] \\ &\quad + 2w_1 [w_2 w_3 (r_1 - r_2)(r_1 - r_3) + w_2 w_4 (r_1 - r_2)(r_1 - r_4) + \ldots + w_{N-1} w_N (r_1 - r_{N-1})(r_1 - r_N)] \\ &\quad \ldots \\ &\quad + 2w_M [w_1 w_2 (r_M - r_1)(r_M - r_2) + \ldots + w_{M-1} w_{M+1} (r_M - r_{M-1})(r_M - r_{M+1}) + \\ &\quad \ldots + w_{N-1} w_N (r_1 - r_{N-1})(r_1 - r_N)] \\ &\quad \ldots \\ &\quad + 2w_N [w_1 w_2 (r_4 - r_1)(r_4 - r_2) + w_1 w_3 (r_4 - r_1)(r_4 - r_3) + \ldots + w_{N-2} w_{N-1} (r_1 - r_{N-2})(r_1 - r_{N-1})]. \end{aligned}$$

Using Definition 3.1, grouping the terms into summations, and re-applying the t-subscripts, we have the result:

$${}_{N_t}\chi^2_t(r_t) = \sum_{\substack{i,j=1 \\ i \neq j}}^{N_t} \chi^2_{i,j}(r_t) + 2\sum_{i=1}^{N_t} w_{it} \left[\sum_{\substack{j,k \neq i \\ k > j}}^{N_t} w_{jt} w_{kt} (r_{it} - r_{jt})(r_{it} - r_{kt}) \right].$$

□

Proof. Equation 3.10:

For the sake of brevity, the documenting of standard algebraic simplifications is abridged. The t-subscripts are discarded for notational ease. Using the squared CSV given in Theorem 3.1 and the pairwise squared CSV expression given in Definition 3.1, we can write the expectation as

$$\begin{aligned}
\mathbb{E}_{\mathbb{P}}\left[{}_{N}\chi^{2}(r)\,|\mathcal{F}_{t}\right] &= \mathbb{E}_{P}\left[\sum_{\substack{i,j=1\\ i\neq j}}^{N}\chi^{2}_{i,j}(r)+2\sum_{i=1}^{N}w_{i}\left(\sum_{\substack{j,k\neq i\\ k>j}}^{N}w_{j}w_{k}(r_{i}-r_{j})(r_{i}-r_{k})\right)\right]\\
&= \sum_{\substack{i,j=1\\ i\neq j}}^{N}\mathbb{E}_{P}\left[\chi^{2}_{i,j}(r)\right]+2\sum_{i=1}^{N}w_{i}\left(\sum_{\substack{j,k\neq i\\ k>j}}^{3}w_{j}w_{k}\mathbb{E}_{P}\left[(r_{i}-r_{j})(r_{i}-r_{k})\right]\right)\\
&= \sum_{\substack{i,j=1\\ i\neq j}}^{N}\left(\left[w_{i}w_{j}(w_{i}+w_{j})\right]\mathbb{E}_{P}\left[(r_{i}-r_{j})^{2}\right]\right)\\
&\quad +2\sum_{i=1}^{3}w_{i}\left(\sum_{\substack{j,k\neq i\\ k>j}}^{3}w_{j}w_{k}\mathbb{E}_{P}\left[(r_{i}-r_{j})(r_{it}-r_{k})\right]\right).
\end{aligned}$$

Thus, one needs to evaluate the expectation for pairwise squared CSV and the return differential product term. For the random variables X and Y, one can write the following identities:

$$\begin{aligned}
\sigma_{X}^{2} &\equiv \mathbb{E}_{P}\left[X^{2}\right]-\mu_{X}^{2}\\
\rho_{XY}\sigma_{X}\sigma_{Y} &\equiv \mathbb{E}_{P}[XY]-\mu_{X}\mu_{Y},
\end{aligned}$$

where μ_X, μ_Y, σ_X^2, σ_Y^2 and ρ_{XY} represent the mean, variance and correlation of the random variables X and Y. Using the asset price dynamics given in Equation 3.9 and letting $M < N$, expected squared CSV can be written as

$$\begin{aligned}
\mathbb{E}_{\mathbb{P}}\left[{}_{N}\chi^2(r)\,|\mathcal{F}\right] &= \left(\sigma_1^2+\mu_1^2\right)\left\{w_1\left[w_2(w_1+w_2)+w_3(w_1+w_3)+\ldots+w_N(w_1+w_N)\right]+2\sum w_i\right\} \\
&\quad+\left(\sigma_M^2+\mu_M^2\right)\left\{w_M\left[\begin{array}{c} w_1(w_1+w_M)+\ldots+w_{M-1}(w_{M-1}+w_M)+ \\ w_{M+1}(w_M+w_{M+1})+\ldots+w_N(w_M+w_N)\end{array}\right]+2\sum w_i\right\} \\
&\quad\ldots+\left(\sigma_N^2+\mu_N^2\right)\times \\
&\quad\left\{w_N\left[w_1(w_1+w_N)+w_2(w_2+w_N)+\ldots+w_{N-1}(w_1+w_{N-1})\right]+2\sum w_i\right\} \\
&\quad-2\left(\rho_{12}\sigma_1\sigma_2+\mu_1\mu_2\right)\left(w_1w_2(w_1+w_2)+\sum w_i\right) \\
&\quad-2\left(\rho_{13}\sigma_1\sigma_3+\mu_1\mu_3\right)\left(w_1w_3(w_1+w_3)+\sum w_i\right) \\
&\quad\ldots-2\left(\rho_{N-1,N}\sigma_{N-1}\sigma_N+\mu_{N-1}\mu_N\right)\left(w_{N-1}w_N(w_{N-1}+w_N)+\sum w_i\right).
\end{aligned}$$

Thus, expected CSV can be written as a weighted sum of the N squared means and variances and the $1/2N(N-1)$weighted differences of the covariances and relevant asset means. By using $\sum w_{it} \equiv 1$, one can simplify the weightings using standard algebraic operations. Collecting the terms and reintroducing the t-subscripts, one has the result

$$\begin{aligned}
\mathbb{E}_{\mathbb{P}}\left[{}_{N_t}\chi_t^2(r_t)\mid\mathcal{F}_t\right] &= \sum_{i=1}^{N_t}\left[\left(\sigma_{it}^2+\mu_{it}^2\right)\left(w_{it}-w_{it}^2\right)\right] \\
&\quad-2\sum_{j=1}^{N_t-1}\sum_{k>j}\left[\left(\rho_{jk,t}\sigma_{jt}\sigma_{kt}+\mu_{jt}\mu_{kt}\right)\left(w_{jt}w_{kt}\right)\right].
\end{aligned}$$

For the special case when $\mu_{it} = a_t$ for all assets, it is trivial to show that

$$\mathbb{E}_{\mathbb{P}}\left[{}_{N_t}\chi_t^2(r_t)\,|\mathcal{F}_t\right] = \sum_{i=1}^{N_t}\left[\sigma_{it}^2\left(w_{it}-w_{it}^2\right)\right]-2\sum_{j=1}^{N_t-1}\sum_{k>j}\left[\rho_{jk,t}\sigma_{jt}\sigma_{kt}\left(w_{jt}w_{kt}\right)\right].$$

□

B.2 NUMBER OF STOCKS, n_t, AND MARKET CONCENTRATION, C

This appendix discusses the effects on CSV of those market variables defined in Table 11 not dealt with in Section 3.4. In addition, a fuller treatment of the effects of market concentration is also provided.

B.2.1 *Number of Stocks, N_t, and Market Concentration, C*

The number of underlying stocks within a market, N_t, has a definite impact on possible CSV values. Figures 9 and 10 show that as N_t increases, possible CSV values change in two specific ways: (i) the lower CSV bound increases and, (ii) the upper CSV bound decreases. Thus the possible CSV range decreases as the number of underlyings increase. If one considers the classical diversification arguments from Modern Portfolio Theory (Markowitz (1952)) this result makes intuitive sense. In short, as the number of stocks in a portfolio increases, the volatility of the portfolio tends towards the market risk lower bound because firm-specific risk is increasingly diversified away. Similarly, maximum possible CSV should also decrease with the number of stocks due to this diversification effect.

The increasing lower bound is also an effect of diversification. For small N_t-values it could quite easily be that stock returns are very closely grouped, leading to very low CSV, which becomes even more pronounced in highly correlated markets. However, as the number of stocks increase, the potential range of returns also increases. Because CSV is merely the weighted sum of return differentials squared, minimum CSV should thus increase with the number of stocks. Therefore it follows that CSV - keeping $SD(\sigma_i)$, $\bar{\sigma}$ and $\bar{\rho}$ constant - converges from both above and below towards some systematic CSV level as the number of stocks increase. Although the authors leave this point as is, it seems reasonable to assume that this insight should affect asset management in terms of the trade-off between the number of stocks held, the total risk of the portfolio and also the possible return/alpha bounds of the portfolio.

The difference between uniform and weighted CSV is discussed in Section 3.4.3, which shows that the level of market concentration materially affects the realised CSV. In addition to this, there is another effect attributable to market concentration. Figure 29 graphs CSV against the standard deviation of Mid-Range underlying volatilities at discrete market concentration levels for $\bar{\rho} = 0.5$ and $N_t = 100$. As the markets become more concentrated (i) the lower CSV bound decreases and, (ii) the upper bound increases. Considering Equation 3.1, this seems fairly intuitive. If

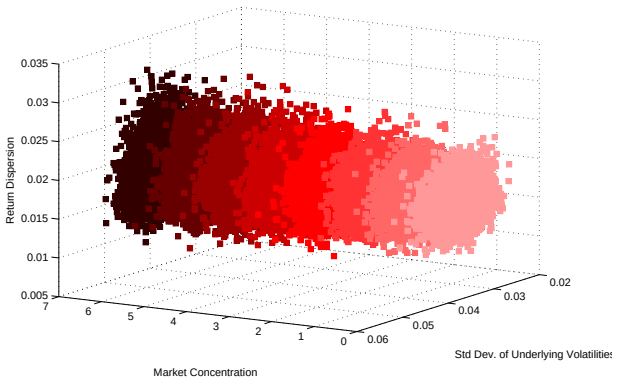


Figure 29: CSV vs. market concentration vs. std. deviation of Mid-Range volatilities for $\bar{\rho} = 0.5$ and $N_t = 100$

those stocks that have returns furthest away from the market return are assigned increasingly large weights, then it follow that the maximum CSV bound will increase. Contrastingly, if those stocks that have returns approximately equal to the market return are increasingly weighted, then the minimum possible CSV bound will decrease. Within the simulation, both weighting scenarios are possible at high concentration levels. This explains the increasing CSV ranges. Thus, market size and market concentration cause opposing movements in CSV.

B.2.2 *Length of Period,* T

Due to the manner in which the simulation is constructed, the underlying volatilities are directly scaled by the square root of T. Thus, for longer time periods, one finds higher underlying volatility values. This in turn increases the standard deviation of the underlying volatilities. Thus CSV grows quadratically with length of time period.

C

CSV MODELLING APPENDICES

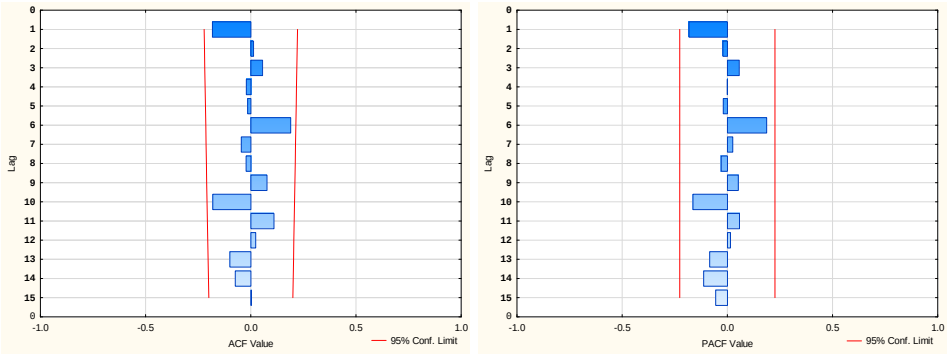
C.1 BEST-SUBSETS REGRESSION AND MODEL CHECKING

Best-subset regression (BSR) uses the Ordinary Least Squares (OLS) method in addition with a factor subset algorithm in order to find a parsimonious model. The user inputs the maximum number of allowed factors within the model. BSR then considers all allowable combinations of interactions and all univariate predictors, and tests the same for statistical significance, usually at an alpha level of 0.05. It preserves those interactions that are significant while dropping those that are not. The process is repeated iteratively until a parsimonious model is derived. It is also not possible to have a multiplicative interaction term present in a model in the absence of its corresponding univariate predictor. The parsimonious model is classically associated with the model that has the highest adjusted-R^2. BSR differs markedly from stepwise-regression procedures, since its final configuration is not path-dependent. This is a critical caveat attached to both the forward and backward stepwise regression methods. The randomisation process outlined in Section 4.4 allows one to deal with the computational intensity of running BSR for very large factor sets.

When fitting a generalised linear multi-factor using OLS, one must ensure that certain assumptions are met. These usually include

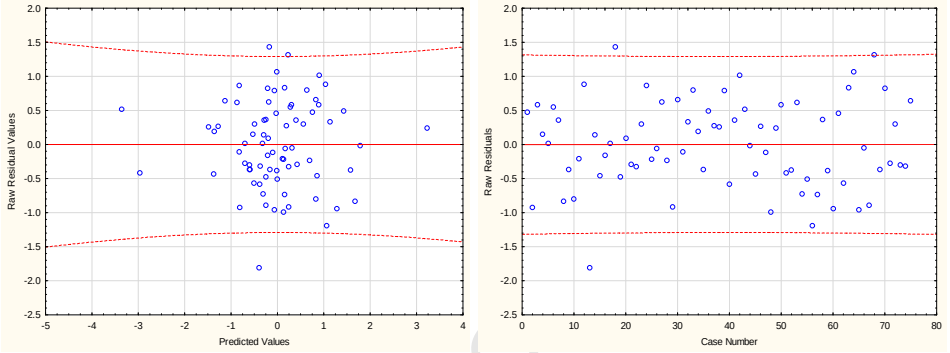
- The model residuals are independent and identically normally distributed, with zero-mean and constant variance. That is, $\varepsilon_t \overset{iid}{\sim} \mathcal{N}(0, \sigma^2)$.
- There is no multicollinearity within the independent variables

The first assumption includes a number of points that need to be validated. In order to do this, consider Figures 30a, 30b, 31a, 31B and 32. Figures 30a and



(a) Autocorrelation function (b) Partial autocorrelation function

Figure 30: ACF and PACF for macroeconomic factor model residuals



(a) Residual-vs-Predicted scatterplot (b) Residual-vs-Case number scatterplot

Figure 31: Scatterplots of model residuals against predicted values and case number respectively

30b plots the autocorrelation and partial autocorrelation respectively - reffered to as ACF and PACF respectively. Because no ACF or PACF lag is statistically significant, one concludes that the residuals display no serial autocorrelation and are thus independent. Figures 31a and 31B show scatterplots of the residuals against the predicted model values and against case number respectively. In this instance, case number refers to time, with zero representing July 1996. Both plots generally display uniformity across domain and range. Thus, one concludes that the residuals display constant variance and are thus identically distributed.[1] Finally, Figure 32 plots the residual histogram with a fitted, zero-mean normal distribution. The fit is relatively good.

1 Although there is a potential outlier highlighted, upon further inspection, this observation does not have materially high leverage and the model results excluding this observation are, for all intents and purposes, equivalent.

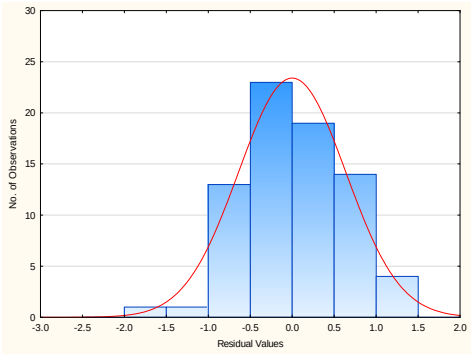


Figure 32: Histogram of model residuals against fitted Normal distribution

Figure 33 displays the histograms of the dependent and univariate independent factors on the main diagonal, and the scatterplots - and thus correlations - between these factors elsewhere. Delta-G-CSV and the univariate macroeconomic factors

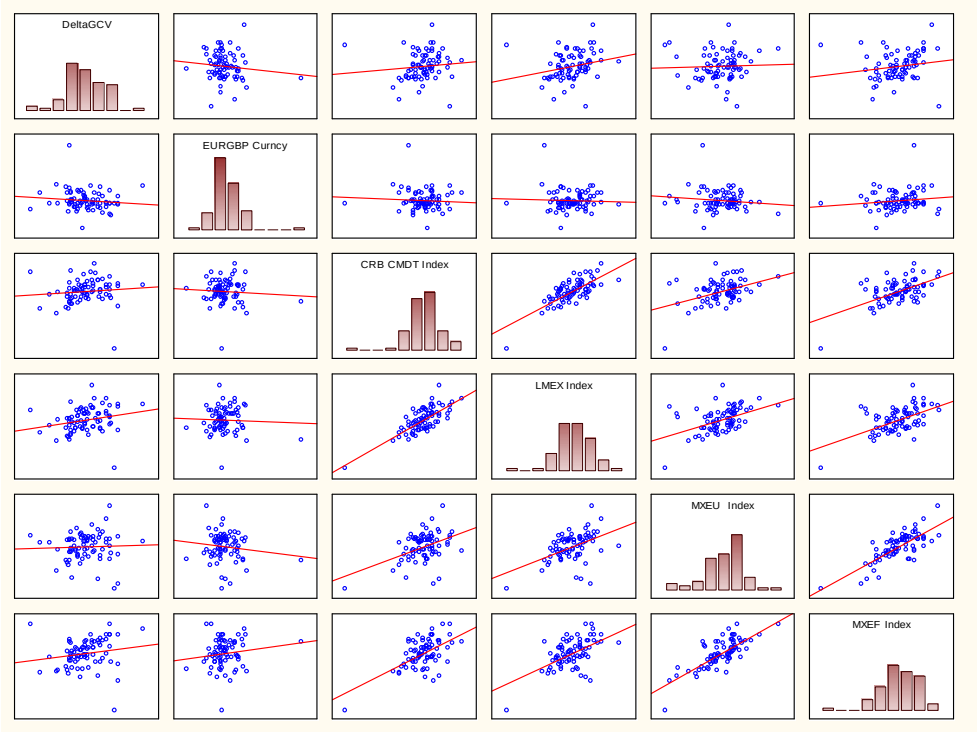


Figure 33: Histograms of the dependent and independent univariate factors and their respective correlations according to factor scatterplots. The gradient of the fitted line gives the correlation between the factors.

are fairly uncorrelated, with the majority showing a slight positive bias and the highest correlation recorded with LMEX. In terms of correlations within the macroeconomic factors, we note substantial positive correlations between the CRB CMDT, LMEX, MXEU and MXEF factors respectively. Remember however, that

the CRB CMDT index is only included by convention of including all univariate factors within interactional factors. CRB CMDT is not statistically significant and displays a standardised coefficient close to zero. Thus the positive correlations attached to this factor are of no concern. What is strange is the high MXEU-MXEF correlation. Whereas the MXEF index factor displays the greatest positive standardised coefficient, the MXEU index displays a fairly large negative coefficient. This is left as an observation to the reader.

C.2 DISTRIBUTION FITTING FOR t_i DATA SETS

Table 17 ranks the fitted distributions by Log-Likelihood values for all T_i CSV recurrence time datasets. Clearly, the IG distribution gives the best fit for all T_i cases. The next-best fitting distribution is the Birnbaum-Saunders, with the Lognormal being a close third. Conversely, the Exponential distribution always gives the worst fit in all cases. This implies that CSV has some form of inherent series memory.

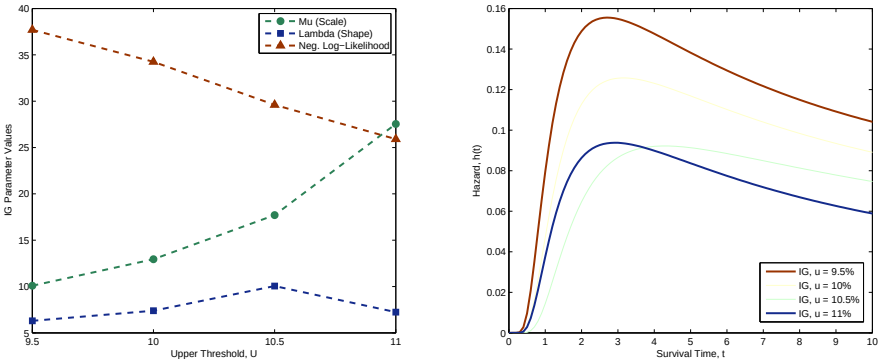
Table 17: Distribution fit to all T_i data, ranked by Log-Likelihood

Rank	T_1		T_2		T_3	
	Dist. (%)*	Ave Log-L†	Dist. (%)	Ave Log-L	Dist. (%)	Ave Log-L
1	IG (100)	31.87	IG (100)	31.07	IG (100)	32.26
2	B-S (50)	45.07	B-S (100)	43.42	B-S (60)	45.51
3	LogN (50)	45.10	LogN (100)	44.42	LogN (60)	45.71
4	Gamma (75)	47.79	Weibull (100)	46.02	Weibull (80)	48.04
5	Weibull (75)	47.97	Gamma (100)	46.45	Gamma (80)	48.33
6	Exp (100)	48.00	Exp (100)	47.31	Exp (100)	48.68

*Dist. (%) give the name of the distribution for rank i and the percentage that the fitted distribution is chosen at that rank.

†The average Log-Likelihood quoted is an absolute value.

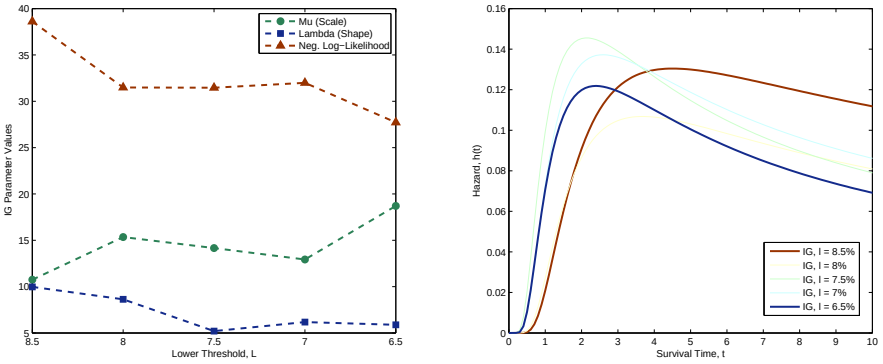
For both T_1 and T_3, we note the same pattern within the fitted IG distribution parameter values as that seen for T_2. As the threshold moves further away from the mean CSV value, the scale parameter increases and the absolute log-likelihood values decrease. Although not as stable as in Figure 14, the shape parameter also remains fairly consistent for both T_1 and T_3.



(a) μ, λ and Neg. Log-Likelihood (b) $h(t)$ for $t \leqslant 10$

Figure 34: Fitted IG distribution parameters and hazard functions for T_1

Turning our attention to the hazard functions, we note that the maximum hazard values are approximately equal for T_1 and T_3 and that this is roughly half that seen for T_2 hazard values. In addition, we see very similar time ranges over which these maximum values occur. For T_1, the maximum hazard values fall within the range $2.8 \leqslant t \leqslant 4.2$ and, similarly for T_3, the maximum hazard values are found for $2.1 \leqslant t \leqslant 4.1$.



(a) μ, λ and Neg. Log-Likelihood (b) $h(t)$ for $t \leqslant 10$

Figure 35: Fitted IG distribution parameters and hazard functions for T_3

D

VOLATILITY TERM-STRUCUTRE APPENDICES

D.1 EXTENSIONS OF CANONICAL VALUATION

It is easy to extend the CV method to incorporate multiple underlying assets. Consider a derivative contract written on M underlying assets. By including $M-1$ additional constraints in the form of Equation 5.7 to the constrained relative entropy minimisation problem, the solution obtained is now given by the multivariate canonical distribution

$$\hat{q}_i = \frac{\exp\left(\sum_{j=1}^{M} \gamma_j^* \frac{R_{i,j}}{r^T}\right)}{\sum_{i=1}^{n} \exp\left(\sum_{j=1}^{M} \gamma_j^* \frac{R_{i,j}}{r^T}\right)}, \quad i = 1, \ldots, n \tag{D.1}$$

where $R_{i,j}$ denotes the i^{th} gross return of asset j for the term T. In this case, the M-component vector γ^* satisfies

$$\gamma^* = \underset{\gamma}{\operatorname{argmin}} \sum_{i=1}^{n} \exp\left[\sum_{j=1}^{M} \gamma_j \left(\frac{R_{i,j}}{r^T} - 1\right)\right]. \tag{D.2}$$

A common example of an additional constraint is to ensure that the at-the-money option price implied by the distribution $\hat{Q}$ is equal to the at-the-money option price quoted in the market. For example, assume that there is a call option with strike level $K \equiv S_0$, expiring at time T and quoted price C. In order to ensure the correct pricing of this option we need include an additional constraint and thus solve for two multipliers, γ_1^* and γ_2^*, which satisfies

$$\gamma^* = \underset{\gamma}{\operatorname{argmin}} \sum_{i=1}^{n} \exp\left[\gamma_1 \left(\frac{R_i}{r^T} - 1\right) + \gamma_2 \left(\frac{\max(S_0 R_i - K, 0)}{r^T} - C\right)\right]. \tag{D.3}$$

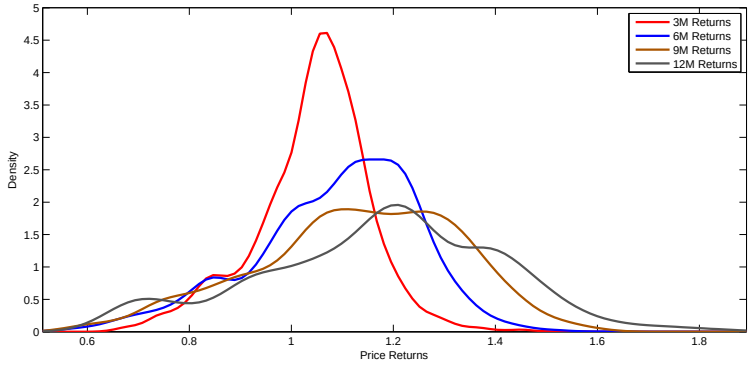


Figure 36: Empirical distributions for June-95 Top40 returns.

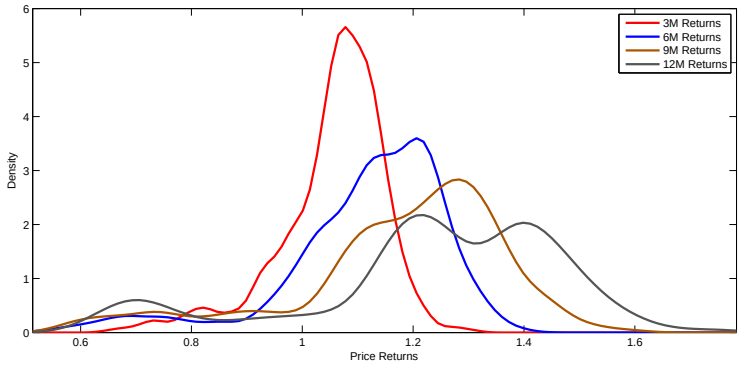


Figure 37: Empirical distributions for Mar-03 Top40 returns.

Substituting the multiplier values into a bivariate canonical distribution, the estimated risk-neutral probabilities are given by

$$\hat{q}_i = \frac{\exp\left[\gamma_1^*\left(\frac{R_i}{r^T}\right) + \gamma_2^*\left(\frac{\max(S_0 R_i - K, 0)}{r^T}\right)\right]}{\sum_{i=1}^{n} \exp\left[\gamma_1^*\left(\frac{R_i}{r^T}\right) + \gamma_2^*\left(\frac{\max(S_0 R_i - K, 0)}{r^T}\right)\right]}. \quad i = 1, \ldots, n$$

D.2 NONPARAMETRIC DISTRIBUTIONS OF TOP40 RETURN DATA

The nonparametric distributions fitted to the Top40 return data are displayed below. Note the non-normality of the data; in particular, the large left tail and general distributional asymmetry.

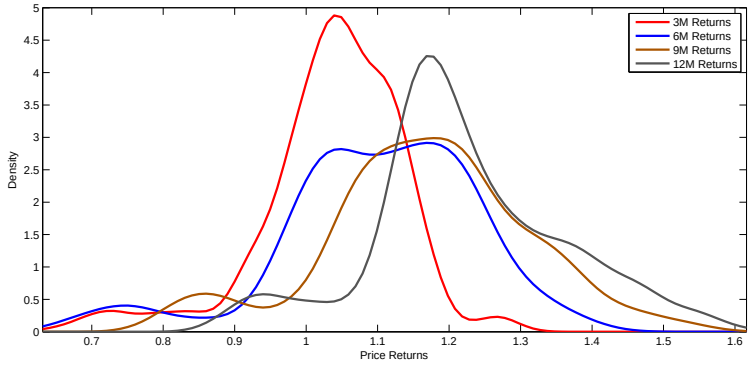


Figure 38: Empirical distributions for Aug-08 Top40 returns.

D.3 CANONICAL VALUATION AND THE RISK-FREE RATE

Although Stutzer's (1996) original paper outlines the use of historic risk-free interest rates, when using empirical option to compare the CV framework to that of the Black-Scholes, Stutzer advocates the use of a constant risk-free rate. By doing so, one removes the confounding effects of a stochastic interest rate on the option pricing comparison. Zou and Derman (1999), also assume a constant risk-free rate for SAS. Whilst their motivation is not specifically outlined, a similar argument to that given by Stutzer can be applied. Because SAS is defined as the spread between the market Black-Scholes implied volatility and the CV implied volatility, stochastic interest rates would again have a confounding effect. Finally, a market-driven reason suggesting the use of a constant rate is that, historically, risk-free interest rates usually move quite slowly over time. Thus, when one is considering a maximum term of one year, one can assume that interest rates will remain fairly stable. This is in contrast to the large, daily movements within underlying asset prices and returns.

The computed fair volatility and TAS levels are still sensitive to the value chosen for the risk-free rate though. However, as shown below, changing the risk-free rate leads to very minor movements in fair volatility. Figures C1, C2 and C3 plot the fair volatility and TAS as at 24 October 2011 under different assumed risk-free rates, $r = 5, 6, 7, 8, 9\%$. Neither the volatility levels nor the shape of the lines changes materially. The largest range occurs for the Mar-03 fair volatility, with a difference

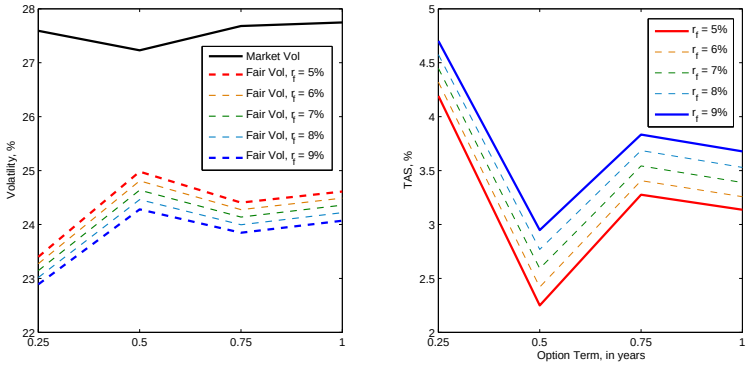


Figure 39: Figure C1: (a) Fair volatility computed via CV using the June-95 return series and different risk-free values, and (b) the resultant TAS lines as at 24 October 2011.

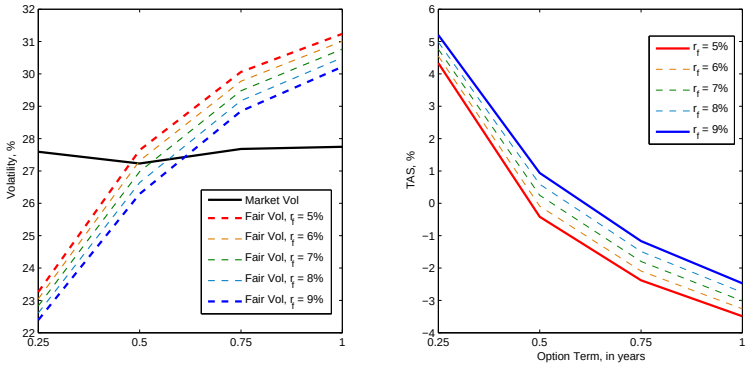


Figure 40: Figure C1: (a) Fair volatility computed via CV using the Mar-03 return series and different risk-free values, and (b) the resultant TAS lines as at 24 October 2011.

of roughly 1 per cent change in fair volatility across a 4 per cent change in risk-free rates. Thus, even an extreme change in interest rates causes only a trivial change in fair volatility levels.

Obviously, the relative effects on TAS are greater due to the much smaller initial values. As for fair volatility though, the TAS shape remains unchanged. Thus, the risk-free rate would only affect whether one considers an option of specific term to be rich or cheap, and even then, only to a small degree. For example, moving from slightly overpriced to slightly underpriced.

Finally, while the relationship between risk-free rate and fair volatility is positive, the relationship between TAS and risk-free rate depends on whether fair volatility is higher or lower than market volatility.

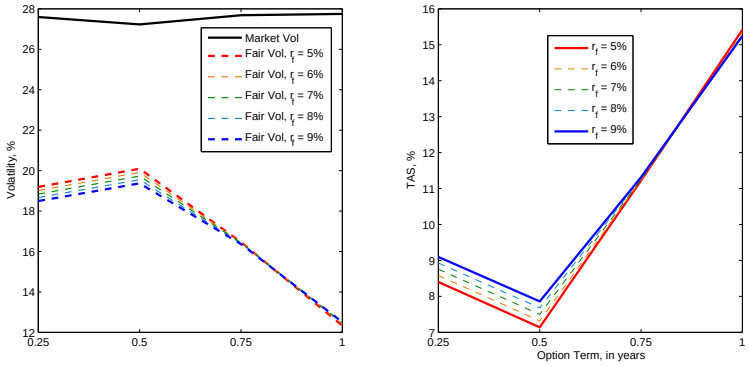
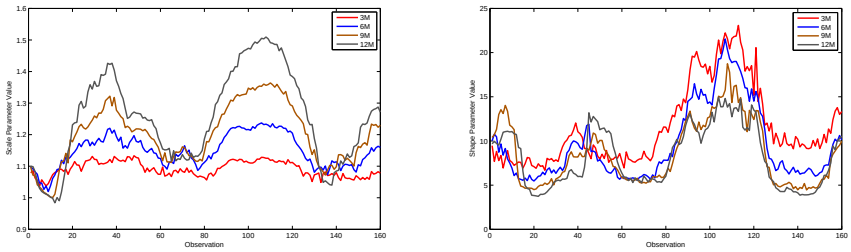


Figure 41: Figure C1: (a) Fair volatility computed via CV using the Aug-08 return series and different risk-free values, and (b) the resultant TAS lines as at 24 October 2011.



(a) Scale, α (b) Shape, b

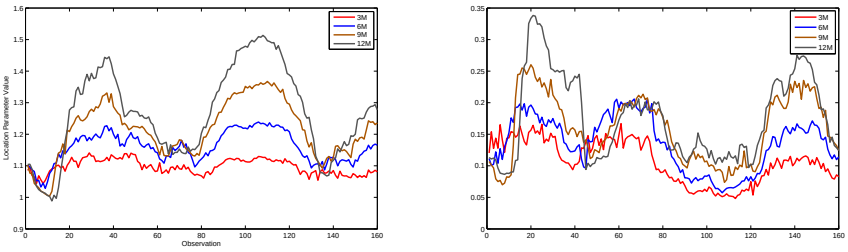
Figure 42: Weibull distribution parameter estimates fitted to Top40 historical data

D.4 MODELLING AND FORECASTING DISTRIBUTION PARAMETERS

This appendix illustrates the ARIMA/GARCH model selection process by reporting the process for a chosen distribution parameter case. In addition, the total range of models chosen and the forecast model parameters is given. Figures 42a and 42b shows the Weibull scale and shape parameters respectively and Figures 43a and 43b do the same for the EV location and scale parameters.

D.4.1 *Selecting an ARIMA/GARCH Model*

As an illustrative example, let us consider the 3-month Weibull scale and shape parameter, α and b. Figures 42a and 42b suggest that neither parameter series is (weakly) stationary. This is confirmed by testing for unit roots within the process via an Augmented Dickey-Fuller test and KPSS test. The results dictate



(a) Location, μ

(b) Scale, σ

Figure 43: Extreme-Value distribution parameter estimates fitted to Top40 historical data

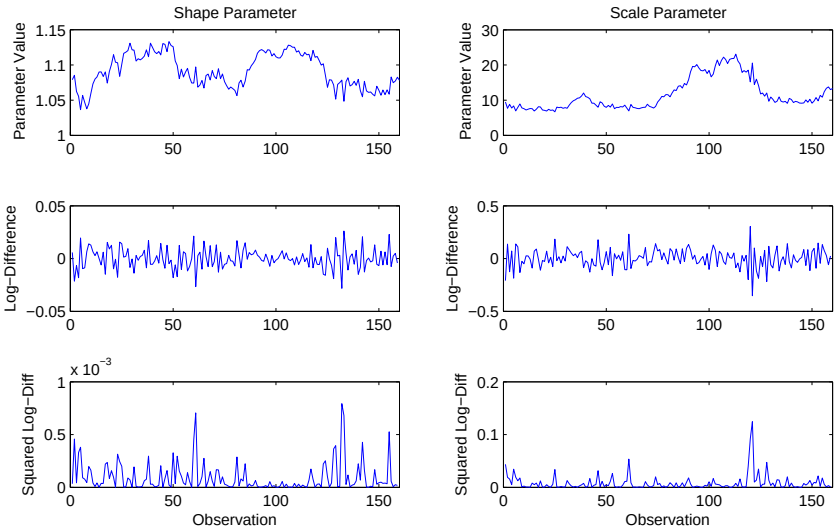
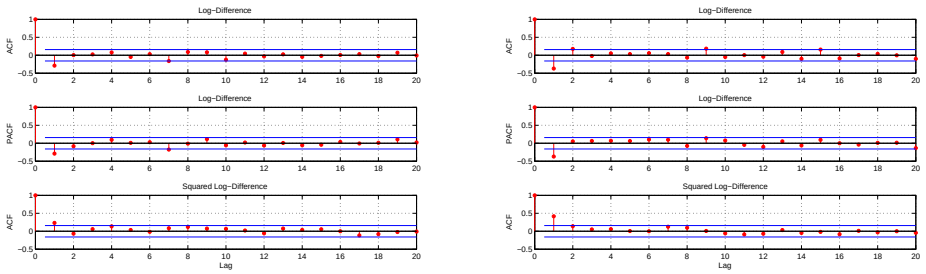


Figure 44: 3-Month Weibull parameter raw values, log-differences and squared log-differences

that one cannot reject the unit root null hypothesis, implying that the series are non-stationary. Figure 44 plots the raw parameter values, the log-differences and the squared log-differences of the parameters. The log differenced scale and shape series show little sign of unit roots. However, looking at the squared log-differences in the lower panels suggest that there is some volatility clustering, or heteroscedasticity. Thus, one should consider a GARCH process for the series residuals, or innovations.

Before modelling the residuals, we first inspect the autocorrelation and partial autocorrelation functions - ACF and PACF - of the respective log-differenced and squared log-differenced series in order to preliminarily identify the ARMA(r, m) and GARCH(p, q) orders. Figures 45a and 45b show the ACF and PACF values



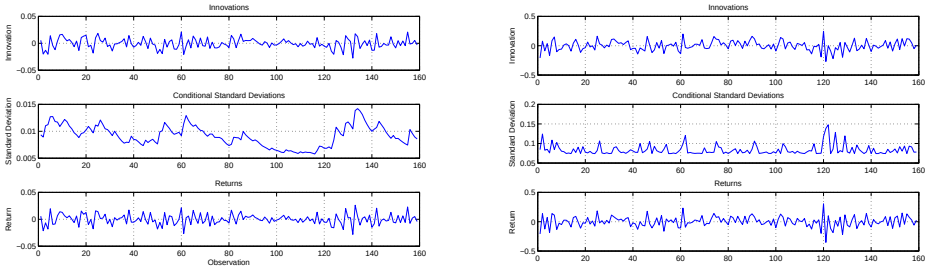
(a) Scale parameter, a (b) Shape parameter, b

Figure 45: Autocorrelation and Partial Autocorrelation functions of the log-differenced and squared log-differenced 3M Weibull parameter values

for a and b with a cutoff value of 5 per cent. The ACF and PACF of both log-differenced series suggest that one should consider an AR(1), MA(1) or ARMA(1,1) model. In addition, the squared, log-differenced series suggest that there are (G)ARCH effects present in both distribution parameter series. The presence of ARCH effects is confirmed by the results obtained from the Ljung-Box-Pierce Q-test and Engle's (1982) ARCH test. Thus the candidate models include a GARCH(p, q) process where $p, q \in \{0, 1\}$.

The candidate models are ranked by Log-likelihood, AIC and BIC criteria values. Maximum Log-likelihood and minimum selection criteria values imply best fit. According to the Box-Jenkins model selection approach, one should always choose the most parsimonious model. Akgiray (1989) states GARCH$(1, 1)$ is extensively used in financial time series modelling, providing a simple representation of the main dynamic characteristics of the return series of a wide range of assets. It is also worth noting that the GARCH$(1, 1)$ model has been proven to have a better forecasting ability when compared to traditional ARCH models. In the majority of cases the BIC value, which imposes the most stringent penalty for the number of model parameters, is used as the selection criterion. In this example, the three-month Weibull log-scale parameter series is best fitted by an ARIMA$(1, 1, 0)$/GARCH$(1, 1)$ process, while the log-shape parameter series is best fitted by an ARIMA$(1, 0, 0)$/GARCH$(0, 1)$ process.

Figures 46a and 46b give the scale and shape model output values.



(a) Scale parameter, a

(b) Shape parameter, b

Figure 46: GARCH output for the 3-month Weibull forecast return distribution, including model innovations, conditional standard deviation and returns.

D.4.2 *Models fitted to Distribution Parameters*

Table 18 details the ARIMA$(r, 1, m)$/GARCH(p, q) models selected for each distribution parameter. Note that the models are always fitted to the log-differenced series and thus $d \equiv 1$ for all models.

Table 19 gives the forecast raw parameter values as well as the 95 per cent confidence interval. The confidence interval values give one lower and upper forecast return distributions, and thus, lower and upper fair volatility and TAS upper and lower limits.

Table 18: Econometric Models selected to model the Weibull and EV distribution parameter series

		Selected Model Order			
Weibull	Return Term	AR(r)	MA(m)	GARCH(p)	ARCH(q)
Scale, a	3M	1	0	1	1
	6M	0	2	0	0
	9M	1	1	0	1
	12M	1	1	1	1
Shape, b	3M	1	0	0	1
	6M	2	1	0	0
	9M	1	1	0	1
	12M	2	0	0	2
Extreme Value		AR(r)	MA(m)	GARCH(p)	ARCH(q)
Scale, a	3M	1	0	1	1
	6M	0	2	0	0
	9M	1	1	0	1
	12M	1	1	1	1
Shape, b	3M	1	0	0	0
	6M	0	0	1	2
	9M	0	0	0	1
	12M	2	0	0	2

Table 19: Weibull and EV parameter forecasts and confidence interval

Weibull	Scale, a			Shape, b		
	Lower CI	Forecast	Upper CI	Lower CI	Forecast	Upper CI
3M	1.063	1.080	1.097	11.549	13.324	15.371
6M	1.138	1.163	1.188	8.808	10.226	11.873
9M	1.207	1.232	1.258	8.905	10.285	11.879
12M	1.240	1.270	1.302	9.381	10.810	12.456

Extreme Value	Location, μ			Scale, σ		
	Lower CI	Forecast	GARCH(p)	Lower CI	Forecast	Upper CI
3M	1.066	1.083	1.100	0.070	0.084	0.100
6M	1.144	1.169	1.194	0.099	0.114	0.131
9M	1.213	1.238	1.263	0.105	0.124	0.146
12M	1.246	1.276	1.306	0.107	0.122	0.139

Note. The confidence interval is calculated at a 95% level

BIBLIOGRAPHY

Yacine Ait-Sahalia and Jefferson Duarte. Nonparametric option pricing under shape restrictions. *Journal of Econometrics*, 116 (1-2):9–47, 2003.

Yacine Ait-Sahalia and Andrew W. Lo. Nonparametric estimation of state-price densities implicit in financial asset prices. *Journal of Finance*, 53 (2):499–547, 1998.

Vedat Akgiray. Conditional heteroscedasticity in time series of stock returns: Evidence and forecasts. *The Journal of Business*, 62 (1):55–80, 1989.

Jamie Alcock and Diana Auerswald. Empirical tests of canonical nonparametrics american option-pricing methods. *Journal of Futures Markets*, 30 (6):509–532, 2009.

Jamie Alcock and Trent Carmichael. Nonparametric american option pricing. *Journal of Futures Markets*, 28 (8):717–748, 2008.

Jamie Alcock and Philip Gray. Dynamic, nonparametric hedging of european style contingent claims using canonical valuation. *Journal of Financial Econometrics*, 2 (1):41–50, 2005.

Andrew Ang, Robert J. Hodrick, Yuhang Xing, and Xiaoyan Zhang. The cross-section of volatility and expected returns. *Journal of Finance*, 61:259–299, 2006.

Ernest M. Ankrim and Zhuanxin Ding. Cross-sectional volatility and return dispersion. *Financial Analysts Journal*, 58 (5):67–73, 2002.

Fisher Black and Myron Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81 (3):639–654, 1973.

Dave Bradfield and Jason Swartz. Binary switches: On or off? *Cadiz Working Paper*, 2003.

Godfrey Cadogan. Canonical option pricing and greeks with implications for market timing. *Working Paper: Available at SSRN: http://ssrn.com/abstract=1625835*, 25 (1):1–19, 2010.

Itinder S. Chadha and Stephen E. Satchell. *Cross-Sectional Volatility and the Mathematics of Managerial Talent in Active Equity Funds*. PhD thesis, Birkbeck, University of London, 2008.

Yong Chen and Bing Liang. Do market timing hedge funds time the market? *Journal of Financial and Quantitative Analysis*, 42 (4):827–856, 2007.

Gregory Connor, Robert A. Korajczyk, and Oliver Linton. The common and specific components of dynamic volatility. *Journal of Econometrics*, 132:231–255, 2006.

Rama Cont. Empirical properties of asset returns: Stylized facts and statistical issues. *Quantitative Finance*, 1 (2):223–236, 2001.

Kieth Cuthbertson, Dirk Nitzsche, and Niall O'Sullivan. UK mutual fund performance: Genuine stock-picking ability or luck. *Conference Paper: available at http://repec.org/mmfc04/55.pdf*, 2004.

Mark de Araujo and Eben Mare. Examining the volatility skew in the south african equity market using risk-neutral historical distributions. *Investment Analysts Journal*, 64:15–20, 2006.

Harindra de Silva, Steven G. Sapra, and Steven Thorley. Return dispersion and active management. *Financial Analysts Journal*, 57 (5):29–42, 2001.

Riza Demirer and Donald Lien. Correlation and return dispersion dynamics in chinese markets. *International Review of Financial Analysis*, 14:477–491, 2005.

Emanuel Derman and Iraj Kani. Riding on the smile. *RISK*, 7:32–39, 1994.

Emanuel Derman, Michael Kaman, Iraj Kani, and Joseph Zou. Is the volatility skew fair? *Goldman, Sachs Quantitative Strategies Research Notes*, pages 1–30, 1997.

Dan DiBartolomeo and Ghazanfer Baig. Time series variation in risk levels and what to do about it. In *Proceedings of the Northfield Seminar*, Newport, RI, 2006.

Jin-Chuang Duan. Nonparametric option pricing by transformation. *Working Paper: Rotman School of Management*, 2002.

Bruno Dupire. Pricing with a smile. *RISK*, 7:18–20, 1994.

William J. Egan. The distribution of standard and poore's 500 index returns. *Working Paper Series: available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=955639*, 2007.

Eugene F. Fama. Components of investment performance. *Journal of Finance*, 27: 551–567, 1972.

Matthias R. Fengler. *Semiparametric Modeling of Implied Volatility, Lecture Note in Finance*. Springer Verlag, Heidelberg, 2005.

Luis Ferruz, Jose L. Sarto, and Maria Vargas. Market timing ability and passive invesment strategies. *Working Paper Series: University of Zaragoza*, pages 1–32, 2010.

J. L. Folks and R. S. Chhikara. The inverse gaussian distribution and its statistical application–a review. *Journal of the Royal Statistical Society. Series B (Methodological)*, 40 (3):263–289, 1978.

Larry R. Gorman, Steven G. Sapra, and Robert A. Weigand. The cross-sectional dispersion of stock returns, alpha and the information ratio. *Journal of Investing*, 19 (3), 2010a.

Larry R. Gorman, Steven G. Sapra, and Robert A. Weigand. The role of cross-sectional dispersion in active portfolio management. *Investment Managament and Financial Innovations*, 7 (3):58–68, 2010b.

Philip Gray and Scott Newman. Canonical valuation of options in the presence of stochastic volatility. *Journal of Futures Markets*, 25 (1):1–19, 2005.

Richard C. Grinold. The fundamental law of active management. *The Journal of Portfolio Management*, 15 (3):30–38, 1989.

M. Ryan Haley and Todd B. Walker. Alternative tilts for nonparametric option pricing. *Journal of Futures Markets*, 30 (10):983–1006, 2009.

Shaikh Hamid and Abraham Habib. Can neural networks learn the black scholes model?: A simplified approach. *Working Paper: Southern New Hampshire University*, 2005.

Roy D. Henriksson. Market timing and mutual fund performance: An empirical investigation. *Journal of Business Finance*, 57:73–96, 1984.

Roy D. Henriksson and Robert C. Merton. On market timing and investment performance of managed portfolios II: Statistical procedures for evaluating forecasting skills. *Journal of Business*, 54:513–533, 1981.

J M. Hutchinson, Andrew W. Lo, and Tomaso Poggio. A nonparametric approach to pricing and hedging derivative securities via learning networks. 1994.

Soosung Hwang and Stephen E. Satchell. GARCH model with cross-sectional volatility: GARCHX models. *Applied Financial Economics*, 15:203–216, 2005.

Soosung Hwang and Stephen E. Satchell. Properties of cross-sectional volatility. Financial Econometric Research Centre Working Paper WP00-4, City University Business School., 2006.

Richard A. Ippolito. On studies of mutual fund performance. *Financial Analysts Journal*, 49:42–50, 1993.

Jens C. Jackwerth. *Option-Implied Risk-Neutral Distributions and Risk Aversion*. Research Foundation of the AIMR, Charlottsville, Virginia, 2004.

Jens C. Jackwerth and Mark Rubinstein. Recovering probability distributions from option prices. *Journal of Finance*, 51 (5):1611–1631, 1996.

Michael C. Jensen. The performance of mutual funds in the period 1945-1964. *Journal of Finance*, 23:389–416, 1968.

Antonie Kotze and Angelo Joseph. Constructing a south african index volatility surface from exchange traded data. *JSE Technical Paper*, pages 1–35, 2010.

Denis Kwiatkowski, Peter C. B. Phillips, Peter Schmidt, and Yongcheol Shin. Testing the null hypothesis of stationarity against the alternative of a unit

root: How sure are we that economic time series have a unit root? *Journal of Econometrics*, 54:159–178, 1992.

Francois-Serge Lhabitant. On swiss timing and selectivity: In the quest for alpha. *Financial Markets and Portfolio Management*, 15:154–172, 2001.

Fabrizio Lillo and Rosario N. Mantegna. Variety and volatility in financial markets. *Physical Review E*, 62 (5):6126–6134, 2000.

Fabrizio Lillo, Rosario N. Mantegna, Jean-Philippe Bouchaud, and Marc Potters. Introducing variety in risk management. Quantitative Finance Papers, 2001.

Harry Markowitz. Portfolio selection. *TheJournal of Finance*, 7 (1):77–91, 1952.

Robert Merton. Theory of rational option pricing. *Bell Journal of Economics and Management Science*, 4 (1):141–183, 1973.

Robert C. Merton. On market timing and investment performance i: An equilibrium theory of value for market forecasts. *Journal of Business*, 51:363–406, 1981.

Joao C. Romacho and Maria. C. Cortez. Timing and selectivity in portuguese mutual fund performance. *Research in International Business and Finance*, 20: 348–368, 2006.

Mark Rubinstein. Implied binomial trees. *The Journal of Finance*, 49:771–818, 1994.

Andrew Rutherford. *Introducing ANOVA and ANCOVA*. Sage Publications Ltd, 6 Bonhill Stree, London, EC2A 4PU, 2001.

Sanjay Sehgal and Manoj Jhanwar. On stock selection skills and market timing abilities of mutual fund managers in india. *International Research Journal of Finance and Economics*, 15:307–317, 2008.

William F. Sharpe. Mutual fund performance. *Journal of Business*, 39:119–138, 1966.

Bruno Solnik and Jacques Roulet. Dispersion as cross-sectional correlation. *Financial Analysts Journal*, 56 (1):54–61, 2000.

Meir Statman and Jonathan Scheid. Dispersion. Santa Clara University Working Paper, 2004.

Michael Stutzer. A simple nonparametric approach to derivative security valuation. *The Journal of Finance*, 51 (5):1633–1652, 1996.

Michael Stutzer and Muinul Chowdhury. A simple non-parametric approach to bond futures option pricing. *The Journal of Fixed Income*, 8 (4):67–76, 1996.

Robert Tompkins. Implied volatility surfaces: Uncovering regularities for options on financial futures. *The European Journal of Finance*, 7 (3):198–230, 2001.

Jack L. Treynor and Kay K. Mazuy. Can mutual funds outguess the market? *Harvard Business Review*, 44:131–136, 1966.

Adonis Yatchew and Wolfgang Hardle. Nonparametric state price density estimation using constrained least squares and the bootstrap. *Journal of Econometrics*, 133 (2):579–599, 2006.

Wallace Yu and Yazid M. Sharaiha. Alpha budgeting - cross-sectional dispersion decomposed. *Journal of Asset Management*, 8:58–72, 2007.

Joseph Zou and Emanuel Derman. Strike-adjusted-spread: A new metric for estimating the value of equity options. *Goldman, Sachs Quantitative Strategies Research Notes: Available at http://ederman.com/new/docs*, 1999.

DECLARATION

I declare that the work presented in this dissertation is, to the best of my knowledge, original and my own work, except as acknowledged in the text and that the material has not been submitted, either in whole or in part, for a degree at this or any other university.

Cape Town, February 2012

Emlyn James Flint