Purely Discontinuous Asset Price Processes

Dilip B. Madan
University of Maryland

April 19, 1999

Abstract

This paper presents the case for modeling asset price processes as purely discontinuous processes of finite variation with an infinite arrival rate of jumps that have arrival rates completely monotone in the jump size. The arguments address both the empirical realities of asset returns and the implications of the economic principle of no arbitrage. Two classes of economic models meeting these conditions are presented and linked. An important example given by the variance gamma process is studied in detail and used to design optimal derivative investment portfolios that are calibrated to actual portfolios to reverse engineer trader preferences and beliefs and infer personalized risk neutral measures termed position measures. Illustrative comparisons of statistical, risk neutral and position measures are also provided.

1 INTRODUCTION

Prices of assets determined in highly liquid financial markets are generally viewed as continuous functions of time. This is true of the Black-Scholes [7], and Merton [33] model of geometric Brownian motion for the dynamics of the price of a stock, and of its many successors that include the stochastic volatility models of Hull and White [23], Heston [22] and the more recent advances into modeling the evolution of the local volatility surface by Derman and Kani [14] , and Dupire [15]. Jumps or discontinuities when considered, have been added on as an additional orthogonal compound Poisson process also impacting the stock as for example in Press[38], Merton [34], Cox and Ross [11], Naik and Lee [37], Bates [6], and Bakshi and Chen [1]. This class of models is broadly referred to as jump-diffusion models and as the name suggests they are mixture models studying the high activity and low activity events by using two orthogonal modeling strategies.

The purpose of this article is to present the case for an alternative approach that stands in sharp contrast to the above mentioned models and synthesizes the study of high and low activity price movements using a class of purely discontinuous price processes. The contrast with the above class of models is that the processes advocated here have no continuous component, as all jump-diffusions must have, and furthermore, the discontinuities are infinite in number with moves of larger sizes coming at a slower rate than moves of smaller sizes. Additionally the jump-diffusion models have what is called infinite variation, in that the sum of absolute price moves is infinity in any interval and one must square these moves before their sum is finite (the property of finite quadratic variation) while the processes we advocate are of finite variation. Unlike jump-diffusions, our processes model price up ticks and down ticks separately and the price process can be decomposed as the difference of two increasing processes representing the increases and decreases of prices. We shall also demonstrate that the finite variation property of the proposed models also enhances their robustness and thereby their relevance for economic modeling.

This paper summarizes the findings of research that I have conducted over the past 15 years in collaboration with a number of coauthors. The research is still on going with a number of new and interesting developments already in place, but we shall focus attention on what has been learned to date. The papers that are summarized here are Madan and Seneta [32] , Madan and Milne [30], Madan, Carr and Chang [29], Carr and Madan [9], [8], Geman, Madan and Yor [18], Bakshi and Madan [5], [4].[1]

The case for purely discontinuous price processes is, as it should be, an argument with many facets. First we summarize the empirical findings on the study of both the statistical and risk neutral processes and observe the empirical need to consider discontinuous processes as relevant candidates. Statistical reality by itself, however, is not a convincing argument. Unsupported by a theoretical understanding of market fundamentals, statistical modeling is at best a spurious coincidence. One must consider the implications of a fundamental economic analysis. We show that economic analysis with the help of some deep structural mathematical results points in the same direction: The use of purely discontinuous price processes. Statistical reality and theoretical conviction are ultimately no match for success. If the wrong model is brilliantly successful in delivering results, while the right one is relatively barren then we have little choice but to work with the incorrect model, bearing in mind its limitations. To address this concern we present some of the successes of modeling with a purely discontinuous price process. We match the success of Brownian motion in option pricing and portfolio management with the success of the purely discontinuous VG process obtained on time changing Brownian motion by a gamma process. The improvement in option pricing is clear, eliminating the implied volatility smile in the strike direction, and we are able to go further in portfolio management and study the optimal management of portfolios of derivative securities, a question that is relatively untouched in the diffusion context. In fact we successfully calibrate observed derivative portfolios as optimal and employ revealed preference methods to infer what we call the position measure but is better known as the personalized state price density. The perspective of purely

[1] The last four of these papers are working papers and can be obtained from my web site: www.rhsmith.umd.edu/finance/dmadan.

discontinuous price processes, we conclude, is not only correct from a statistical and theoretical viewpoint, but is also rich in results and interesting applications.

The statistical findings we summarize confirm from a variety of perspectives that the local motion of the stock price is not Gaussian. This is true of both the time series of moves and the pricing distribution of moves as reflected in option prices. Apart from these standard tests of normality we also consider the behavior of extremal events. Relying on asymptotic laws of maxima and minima of independent sampled observations,(See Embrechts, Kluppelberg and Mikosch [16]) we employ long time series of returns and reject the hypothesis that asset return distributions are locally Gaussian. They lie in the domain of attraction of the Fréchet distribution that includes the log gamma formulation of the VG process. Additionally we investigate empirically the relationship between arrival rates of jumps of different sizes with the jump size. The focus of our attention is on whether arrival rates display a monotonicity with respect to size, decreasing as the size rises, and whether the assumption of an infinite arrival rate is supported by a casual analysis of arrival rates. We conclude in favor of infinite and decreasing arrival rates.

From a theoretical perspective, we concentrate on the implications of no arbitrage, a property that is fundamental to all models for the asset price process. This property is shown to imply that asset prices in continuous time must be modeled by a time changed Brownian motion. The question at issue is then the nature of the time change. We investigate whether the time change could be continuous, with the resultant implication of the continuity of the price process, and show that this is possible only in economies where returns are locally Gaussian and time is locally deterministic and non-random. Given the overwhelming evidence on the lack of a locally Gaussian return distribution we are led to entertain the lack of continuity of the price process. This modeling choice is also consistent with observations on studying the relationship between time changes and economic activity, whereby we learn that time changes are related to some measure of the rate of arrival of orders or trades. As the latter have a random element, and are not locally deterministic, this suggests that such properties are inherited by the time change and hence once again we are led to the class of discontinuous price processes.

Within the class of discontinuous processes we begin our search by focusing attention in the first instance on processes with identical and independently distributed increments: A property shared with Brownian motion, the base model for the underlying uncertainty in the continuous case. This leads naturally via the Lévy-Khintchine theorem for such processes to considering Lévy processes characterized by their Lévy densities whose empirical counterparts are precisely the relationship between arrival rates of jumps of different sizes and the jump size noted earlier in our empirical analysis. When the Lévy density integrates the absolute value of the jump size in the neighborhood of zero, a case we restrict attention to, the process has finite variation and can be decomposed into the difference of two increasing processes that constitute our models for the price up and down ticks. We suggest this model as a partial equilibrium model that clears market buy orders with an up tick price response as the order is cleared through the limit sell book. The converse being the case for market sell orders cleared through the limit buy book at a price down tick.

An alternative and interesting economic model for price responses goes back to traditional dynamic models of price adjustment that represent the rate of adjustment as a function of the level of excess demand in the economy. We term this function relating the rate of change of prices to excess demand, the *force function* of the economy. Modeling excess demand by Brownian motion we may write the price process as the difference between price increases occurring during positive excursions of Brownian motion less the cumulated decreases that occur on negative excursions of Brownian motion. Such a price process is of course open to arbitrage by trades that reverse themselves during a single excursion of Brownian motion. For example, on a single positive excursion, one buys at a price and then sells at a higher price in the same excursion. To avoid such arbitrage, we restrict equilibrium trading to equilibrium times by requiring these to occur at the zero set of Brownian motion. This is organized by evaluating the disequilibrium price process at the inverse local time of Brownian motion. The resulting price process inherits the property of being purely discontinuous from inverse local time, and the process is the difference of two increasing processes that cumulate price responses during positive and negative excursions.

The two models of discontinuous price processes, i) Lévy processes and ii) integrals of force functionals of Brownian motion to inverse local time, are surprisingly related under the hypothesis of complete monotonicity of the Lévy density.[2] Every force function has associated with it a completely monotone Lévy density and for every completely monotone Lévy density there exists an equivalent representation of the price process using a force function. The equivalence is however a consequence of some deep results from number theory and hence the surprise.

We also consider the issue of robustness of the economic model with respect to tolerance of a heterogeneity of views on parameters and observe that the property of bounded variation in the price process is critical for delivering such robustness. Our concern in robustness with respect to views on parameters is that different beliefs should naturally allow for different probabilities, but the probabilities should remain equivalent and not become singular. With infinite variation there are many cases where a change in certain parameters induces singularity of measures.

With the theoretical and statistical foundations in sufficient harmony, and two broad classes of models outlined in sufficient detail we turn our attention to the study of particularly rich examples in this class of models. The basic generalization of geometric Brownian motion we introduce is the VG process that introduces two additional parameters providing control over skewness and kurtosis. The model arises on evaluating Brownian motion with drift at a random time given by a gamma process. The volatility of the gamma process

[2] The Lévy density is completely monotone if each of its two halves on the positive and negative side have the property of sign alternating derivatives or equivalently can be expressed as Laplace transforms of positive functions on the positive half line. Hence, they are essentially mixtures of exponential densities.

provides control over kurtosis while the drift in the Brownian motion before the time change controls skewness. We show that this model is successful in option pricing, eliminating the smile in the strike direction with relative ease.

Fundamental to the world of purely discontinuous price processes is the property of options being market completing assets with a genuine role to play in the economy and a natural demand for these assets by investors. Recognizing these properties we reconsider the problem of optimal derivative investment in continuous time, keeping in place Mertonian [35] objective functions for the investor but expanding the asset space to include all European options on the underlying stock for all strikes and maturities. We find that for $HARA$ utilities and VG statistical and risk neutral measures the derivative investment problem may be solved in closed form and leads in such economies to a healthy demand for at-the-money short maturity options: Precisely the options with the greatest liquidity in financial markets. One may view the Black-Scholes economy as teaching us about stock delta positions in option hedging, while the first lessons of investment in purely discontinuous high activity price processes are about positioning in short maturity at-the-money options.

With some courage we consider replicating actual trader derivative positions as optimal ones, allowing in the process adjustments in the level of risk aversion in power utility and a view on subjective kurtosis that may differ from the statistically observed kurtosis level. Kurtosis is particularly hard to estimate as its variance is of the order of the eighth moment. With this two dimensional flexibility, we are amazingly successful in many instances in calibrating actual spot slides as optimal wealth responses from the perspective of our continuous time optimal derivative investment model.[3] Having inferred risk aversion and the characteristics of subjective probability consistent with replicating observed positions as optimal, we may construct the personalized state price density that values options at a dollar amount yielding a marginal utility that matches the future expected marginal utility from holding the option. We call this state price density the position measure and provide explicit constructions of position measures, contrasting them with the risk neutral and statistical measures. We find generally that position measures are closer to the statistical measure and lie between the statistical and risk neutral measure. This is consistent with the view that traders are aware of relative frequency of occurrence of market moves and their prices and accordingly make markets in option contracts.

The outline for the rest of the chapter is as follows. Section 2 presents a summary of the statistical results. The economic consequences of no arbitrage are described in section 3, while the two equivalent but apparently different economic models of the price process are summarized in section 4. The task of constructing specific examples consistent with the statistical and economic observations of these sections is taken up in section 5. The basic operating model of the VG process is introduced in section 6. Its successes in option pricing are summarized in section 7. Optimal solutions to the asset allocation problem with

[3] The spot slide of a Derivatives book graphs the value of the book as a function of the level of the underlying, typically varying the underlying in the range plus or minus 30% of spot for equity assets.

derivatives are presented in section 8 and employed to infer position measures in section 9. Section 10 concludes.

2 PROPERTIES OF THE PRICE PROCESS

This section summarizes some of the broad properties of the statistical and risk neutral price process. We address issues related to the normality of the motion, the behavior of extreme moves and the shape of the density of arrival rates of price moves. The emphasis in all cases is on the movement over short horizons as we view the macro moves as cumulated short moves.

2.1 Long-tailedness of Historical Returns

We begin by considering some well known results about the long-tailedness of the statistical return distribution and standard chi-square goodness of fit tests of normality of the return distribution. Early results on these issues go back to Fama [19] where both the independence of daily returns and their longtailedness is documented. We now have data at much higher frequencies of observation and report in Table 1 results on S&P 500 futures returns at these frequencies. We focus attention on the level of the observed kurtosis and on χ^2 goodness of fit tests for normality.

TABLE 1
High Frequency Tests of Normality
$S\&P$ 500 Futures Returns Nov.1992-Feb. 1993

	1 Min.	15 Min.	Hourly	Daily
Kurtosis	58.59	13.85	5.97	10.31
χ^2 test statistic	437.12	931.85	98.323	123.84
χ^2 critical value 5%	9.26	5.7	3.57	0.989

Source : Dissertation of Thierry Ané
University of Paris IX Dauphine and ESSEC 1997

We observe from Table 1 that the kurtosis is substantially higher than 3, the kurtosis level of a Normal distribution. The Goodness of fit tests also overwhelmingly reject the hypothesis of normality for returns over short durations. We will note later, in the next section, that this has very significant implications for modeling the dynamics of the price process.

2.2 Long-tailedness in Risk Neutral Distribution

Apart from the statistical return distribution we are also interested in the risk neutral or pricing distribution as implied by option prices. This distribution assesses the futures price of a binary derivative that pays a dollar at a future date if the stock price is in a certain interval, as opposed to the likelihood of the occurrence of this event. The distribution may be recovered from observed option prices with the density being given by the second derivative of the European call option price, of maturity matching the future date, with respect to the option strike as derived in Ross, [41] and Breeden and Litzenberger, [3]. If the distribution describing the current prices of derivatives written on future stock price events is Gaussian then an implication is that the implied volatility obtained from equating the option price to the value given by the Black-Scholes formula, should be constant as one varies the strike for a fixed maturity. On the other hand, if this density is symmetric about a point, then the implied volatilities though no longer necessarily flat with respect to strike, should be symmetric about a point as well. Both these implications are contradicted by what has come to be known as the implied volatility smile.

We present in Table 2 below, the implied volatility smile on S&P 500 Index Options, based on out of the money options using only puts for strikes below, and calls for strikes above, the spot price. These are the more liquid option markets. The time period covered is June 1988 to May 1991 and we focus attention just on the short maturity options. The choice of this focus is motivated by our intention of studying the dynamics of the stock price process, which is but the cumulation of short maturity moves.

TABLE 2
The Smile in Implied Volatilities
at shorter maturities below 60 days.

Moneyness $Spot/Strike$	June 1988- May 1989	June 1989- May 1990	June 1990- May 1991
< 0.94	17.27	16.16	19.70
$0.94 - 0.97$	16.21	15.10	18.23
$0.97 - 1.00$	16.33	15.83	18.65
$1.00 - 1.03$	17.42	17.81	20.87
$1.03 - 1.06$	19.04	20.65	22.27
> 1.06	21.84	25.70	25.57

Source: Bakshi, Cao and Chen, Journal of Finance (1997), page 2015.

We observe from Table 2, reading up the columns, that as the strike level rises, the implied volatility falls sharply followed by a smaller rise as one crosses the level of the spot price. We therefore clearly have a smile shape in the short maturity implied volatility, but the left and right sides are not symmetric. We may conclude from these observations that the left tail of the pricing distribution is fatter than the right tail, and this reflects a negative skewness in the distribution. The existence of the smile itself is evidence of excess kurtosis (relative to the normal distribution) in this density.

2.3 The Behavior of Extreme Moves

Tables 1 and 2 are classical results on the statistical properties of densities associated with price movements in financial markets. They summarize essentially the narrow behavior of the return distribution as may be evidenced by noting that most of the returns considered in the time series analysis are the ones with the smaller magnitudes, and the range of moneyness reported in the implied volatility curves is just within 6 percentage points over an average period of a month. Hence the evidence presented is that of lack of normality in the neighborhood of the zero return and one might wonder whether at least the tail of the distributions is Gaussian. For the risk neutral distribution this has the implication that the implied volatility curve flattens out as one gets into deep out-of-the-money options on both sides, though the level at which the curves flatten out may be different on each side.

To focus attention on the behavior of the tails of the distribution with a view to addressing whether this may be Gaussian, we consider the behavior of extremes. It is shown in Embrechts, Kluppelberg and Mikosch [16], that the asymptotic distribution of the maximum and minimum of independent drawings from a Gaussian distribution is given up to shift and scale by the Gumbel distribution. The other possible asymptotic distributions for these extremal events are, again up to shift and scaling, the Weibull or Fréchet distribution. For distributions that have as support the positive half line, the candidate limiting distributions are just the Gumbell and Fréchet distributions.

The analysis of extreme events requires long time series of data and for this purpose we obtained data on daily returns on the Dow-Jones Industrial average (DJIA) for 100 years from $1897 - 1997$. Partitioning this data into nonoverlapping intervals of 100 days, we constructed a series on the maximum percentage daily rise and the maximum percentage daily drop in the DJIA over the 100 days. We then artificially nested the Gumbel and Fréchet log likelihoods and tested the null hypothesis that the distribution of the extreme event is Gumbell, the limit of the Gaussian tail. Table 3 presents these results.

TABLE 3
Log-Likelihoods of The Distribution of Extremal Price Movements
Maximum Daily Percentage Rise and Fall in the DJIA over
100 Day Nonoverlapping Intervals for 100 years.

	Maximum Daily Drop 100 Days		
	Gumbell	Fréchet	P-Value
1897-1997	768.37	808.58	0.00
1897-1945	380.22	389.98	0.01
1946-1997	409.93	434.74	0.00
	Maximum Daily Rise 100 Days		
	Gumbell	Fréchet	P-Value
1897-1997	811.66	833.77	0.01
1897-1945	395.79	408.92	0.01
1946-1997	358.33	432.95	0.01

Source : Bakshi and Madan (1998),
"What is the Probability of a Stock Market Crash,"
Working Paper, University of Maryland.

Table 3 demonstrates that the normality hypothesis may also be rejected as a model for the tails of the statistical distribution of daily returns. Given the evidence on excess kurtosis, we would conjecture that these tails are heavier than Gaussian and if the property is shared with the risk neutral distribution, as we suspect it is, then implied volatilities must continue to rise as we get deeper out-of-the-money, i.e., the implied volatility curves do not flatten out at either end of the strike range. At this point we do not have documentary evidence on very deep out-of-the-money implied volatilities but observations from current market quotes on S&P 500 index options would suggest that this may well be the case.

2.4 The Structure of the Arrival Rates of Price Moves

The arguments of this paper lead us to considering as models for the dynamics of stock prices, purely discontinuous processes. Such processes, when they have independent and identically distributed increments, are characterized by their Lévy densities that essentially count the rate of arrival of jumps of different sizes. These are a wide class of processes and structural properties if supported by data are beneficial in limiting the class of models that need to be considered. One such structural property is complete monotonicity of the Lévy density, whereby large jumps occur at a smaller rate than small jumps. This is a reasonable property to expect as market participants facing price increases on buy orders and decreases on sell orders have an incentive to minimize these impacts. Another structural property is the aggregate arrival rate of jumps or moves, that could be finite or infinite. We note in this regard that Brownian motion is an infinite activity process as the actual sum of absolute price moves is itself infinite for Brownian motion as it is a process of infinite variation. We note further that jump-diffusions employ a compound-Poisson process for the arrival of jumps that have a finite arrival rate with the magnitude of jumps having, once again, a normal distribution.

The models we propose in this paper, have infinite arrival rates of jumps and in this regard they are closer to Brownian motion, but unlike Brownian motion they are processes of finite variation. This requires that the integral of the Lévy density be infinite, but the density times the jump size should have a finite integral near zero. A typical Lévy density meeting these conditions is of the form $\alpha \exp(-\beta |x|) / |x|^{1+\rho}$ for jump size x with $\rho > 0$. The log arrival rate is in this case linear in the jump size and the log of the jump size, with the coefficient on the log of the jump size being above unity. For $\rho > 1$ we have infinite variation and $\rho = 0$ is the case of the gamma process or in this case the difference of two gamma processes which we will note later is the VG model. On the other hand if the jump sizes are exponentially distributed with a finite arrival rate, as postulated for example in Das and Foresi [12] then the log arrival rates are linear in just the size with the coefficient on log size being 0 or $\rho = -1$. In contrast the log arrival rate of the compound-Poisson process with Gaussian jump sizes (see Cox and Ross, [11]) is linear in the size and the square of the size. Since the exponential of a negative quadratic shifts from being concave near zero to convex near infinity, such a Lévy density is not completely monotone.

A cursory evaluation of these structural properties may be simply made by regressing log arrival rates on the size of jumps, their log and their square. For our 100 year data on daily returns on the DJIA we counted the number of arrivals of jumps in the different size categories and then regressed the log of the empirically observed arrival rate on the size of the jump, its log and its square. For the Cox and Ross [11] model the log arrival rates have a single representation that is not distinguished by the sign of the jump while for the Das and Foresi and VG type models, the parameters vary with sign, so the latter two model estimates allow for this by separating out the positive and negative moves. Table 4 presents the results of these regressions.

TABLE 4
REGRESSION OF LOG ARRIVAL RATES
ON THE SIZES OF JUMPS
Standard Errors are in Parentheses

	LOG ARRIVAL RATES OF DROPS			
	CONSTANT	JUMP SIZE	LOG SIZE	R^2
1897-1997	−9.88 (1.44)	−31.6 (8.36)	−1.92 (0.32)	0.97
1897-1945	−8.51 (1.45)	−33.0 (8.53)	−1.65 (0.32)	0.97
1946-1997	−12.35 (2.22)	−32.0 (17.78)	−2.41 (0.45)	0.95
	LOG ARRIVAL RATES OF RISES			
	CONSTANT	JUMP SIZE	LOG SIZE	R^2
1897-1997	−11.55 (1.71)	−24.5 (9.10)	−2.25 (0.38)	0.96
1897-1945	−10.29 (1.65)	−25.4 (8.97)	−1.99 (0.37)	0.97
1946-1997	−13.66 (3.23)	−25.8 (24.45)	−2.67 (0.65)	0.93
	ARRIVAL RATES FOR JUMP DIFUSSION			
	CONSTANT	JUMP SIZE	SIZE^2	R^2
1897-1997	−3.66 (0.53)	−1.73 (3.86)	−447 (66)	0.70
1897-1945	−3.36 (0.48)	−1.77 (3.66)	−421 (62)	0.71
1946-1997	−3.17 (0.65)	1.54 (8.98)	−928 (191)	0.64

Source : Bakshi and Madan (1998)
"What is the probability of a stock market crash,"
Working Paper, University of Maryland

From Table 4 we observe that the coefficient of log size in the first two regressions is significantly different from zero and may even be close to 2, which definitely argues against a process with a finite arrival rate, as in Das and Foresi [12]. As in a number of cases the coefficient is estimated above two, the process may be one of infinite variation. However, we cannot reject the hypothesis that this coefficient is below 2 and hence we may have a process of finite variation. As will be argued later, there are other reasons for entertaining a finite variation process and in the absence of strong evidence to the contrary we conclude in favor of finite variation processes with infinite arrival rates.

Regarding the comparison with the Cox and Ross [11] process with quadratic log arrival rates, we note that the linear term is in all cases insignificant, suggesting a pure quadratic model, but note further that one explains only up to 70% of the variation in arrival rates compared with up to 97% of the variation using the completely monotone density.

2.5 Summary of Empirical Observations

We note from Tables 1 and 2 that both the statistical and risk neutral distributions are for short intervals, not normal distributions. They have significant levels of excess kurtosis and the risk neutral distribution in particular is also skewed to the left with a heavier left tail than a right tail. This absence of normality continues into the tail of the densities as reflected by an analysis of extremes in Table 3. From Table 4 we infer that a reasonable model could be a pure jump model with an infinite arrival rate - Lévy density integrating to infinity - and a process of finite variation. We also infer from Table 4, some support for a completely monotone Lévy density. Heavy risk neutral tails, if confirmed, imply that implied volatilities are strictly $U-$ shaped and do not flatten out as one moves deep out of the money in both directions.

3 THE IMPLICATIONS OF ECONOMIC THEORY

One of the most far reaching implications of economic theory are now recognized to be the consequences of the no arbitrage hypothesis. From early beginnings with the Ross' [42] theory of arbitrage, and its application to option pricing by Black-Scholes [7] and Merton [33] to the development of the martingale theory of pricing by Harrison and Kreps [20] and Harrison and Pliska [21] this hypothesis has yielded many deep and interesting results. We demonstrate in this section a continuation of these lessons and draw out more exactly the implications of this hypothesis for modeling the dynamics of the asset price.

Before proceeding we note an important proviso with regard to this hypothesis. Financial markets may display arbitrage opportunities and there are many documented "so-called" anomalies that are suggestive of such a possibility, yet it remains true that models of the price process to be employed in developing derivative pricing models, must be free of arbitrage. This is so for the simple reason of preventing traders from arbitraging a firm quoting arbitrageable prices. That models must be arbitrage free goes without question.

3.1 The Stochastic Process Implications of No Arbitrage

Four results, one from mathematical finance and the other three from the theory of stochastic processes form the foundations for the stochastic process implications of the hypothesis of no arbitrage. The first of these results, from mathematical finance, demonstrates that the absence of arbitrage is equivalent to the existence of an equivalent martingale measure. The other results, from the theory of stochastic processes, characterize martingales.

3.1.1 No Arbitrage and Martingales

This result has many proofs or no proof depending on the context and meaning to be attached to the idea of no arbitrage. In discrete time and with finitely many states there is no ambiguity and the result is true with a proof going back to Harrison and Kreps [20]. At the other extreme we have continuous time and states given, at a minimum, by the relatively large set consisting of the paths of the stock price process. Here the existence of martingale measures easily implies the absence of arbitrage, but the implication in the reverse direction is not available, and this is the direction that concerns us here. Essentially the hypothesis of no arbitrage, merely asserting that one cannot combine a portfolio of existing assets to earn a non-negative, non-zero, cash flow at a negative current price is too weak to deduce the existence of a martingale measure. For interesting counterexamples of economies satisfying no arbitrage and yet not satisfying the existence of a martingale measure the reader is referred to Jarrow and Madan [25].

In these richer contexts allowing an infinity of dynamic trading strategies, the hypothesis of no arbitrage must be strengthened to permit deduction of a martingale measure. The strengthening required is topological in nature and requires that one not be able to construct an approximation to an arbitrage opportunity in some limiting sense and then it does follow that there exists an equivalent martingale measure. The first results in this direction are due to Kreps [28]. The difficulty with the result of Kreps [28] is the weak sense in which the limit is taken, as the definition of approximation lacks a sense of uniformity, and what is regarded as an approximation may not be so from the perspective of other economic agents.

The strongest results in this direction are due to Delbaen and Schachermayer [13]. They employ a strong and uniform sense of no arbitrage and show that if there is no random sequence of zero cost trading strategies converging in this strong sense to a non-negative, non-zero cash flow, with the random sequence being uniformly bounded below by a negative constant, then there exists a martingale measure and the converse holds as well. They term this hypothesis, No Free Lunch with Vanishing Risk (NFLVR) and prove that it is equivalent to the existence of an equivalent martingale measure.

3.1.2 Martingales and Semimartingales

The second important result in ascertaining the stochastic process implications of the hypothesis of no arbitrage is Girsanov's theorem. This is pointed out by Delbaen and Schachermayer [13] and amounts to noting that if there exists a change of measure from the true statistical measure P to a martingale measure or risk neutral measure Q such that under Q discounted asset prices are martingales, then it must be that under P the price process was a semimartingale to begin with.

This is a very useful realization as it informs us that models for price processes may safely be restricted to the class of semimartingale processes. Since the class of semimartingales is very wide indeed, one might argue that this is not a very important insight. On the other hand, a lot is known about the structure of semimartingales and for a modeler it is useful to know that the search may be constrained by this structure. Some recent examples of proposals for stock price processes that are not semimartingales include the use of fractional Brownian motion with the arbitrage demonstrated in Rogers [40].

Semimartingales are a difficult concept to communicate in precision, as they go beyond the idea of a simple concept and are in fact a fairly complete and very general theory of random processes, yet given their established importance to the field of mathematical finance today, it is imperative that we communicate some of the flavor of this theory: And do so with brevity. There are at least two approaches, one analytical and the other structural and it is best to consider the structural approach. From this perspective a semimartingale is described by its decomposition into a martingale plus a very general model for the drift of the process. This certainly includes linear drift but also more general models of the drift. One merely requires that this process be of finite and integrable variation, as well as being predictable (i.e. the limit of left continuous functions). Examples include Brownian motion with drift, solutions to stochastic differential equations like the mean reverting Cox, Ingersoll and Ross[10] interest rate process and the VG model [29] with drift to be discussed later in the paper. To appreciate what is not a semimartingale, we consider the discrete time continuous state context studied by Jacod and Shiryaev [26] where they show that the no arbitrage property is lost if zero is not in the relative interior of the support of the multivariate return distribution over the discrete time step and hence the arbitrage. We also learn from this paper, that not all semimartingales are stock price models, as calendar time is a semimartingale with a zero martingale component and has arbitrage if it was a price process. The important property is to get zero into the relative interior of the support, at least in discrete time. Price processes must be semimartingales with a non-zero martingale component.

3.1.3 Semimartingales and Time Changed Brownian Motion

The next result we employ in developing our understanding of the stochastic process implications of no arbitrage is a fundamental characterization of all semimartingales, due to Monroe [36]. This remarkable result shows that every semimartingale can be written as a Brownian motion (possibly defined on some adequately extended probability space) evaluated at a random time. This result is somewhat surprising at first, since Brownian motion, even if evaluated at a random time, is suggestive of a martingale and as noted earlier semimartingales include simple linear drifts like time itself. However, this is only a problem at first glance as the time change need not be independent of the Brownian motion and calendar time t, for example, is Brownian motion $W(t)$ evaluated at the first time $T(t)$ at which this same Brownian motion reaches t.

By this result the study of price processes is reduced to the study of time changes for Brownian motion and one may consider both independent and dependent time changes. One might ask what the time change represents? Ignoring price changes that are the possible result of noise or liquidity trades, changes in the price of an asset occur through trades motivated primarily for reasons of information. The cumulated arrival of relevant information is a reasonable, economically meaningful measure of the time change, that gets translated into buy or sell orders. Geman, Madan and Yor [18] consider many models for the process of buy and sell orders and relate the time change in all these cases to some measure of economic activity. In some cases the measure is just the number of trades while in other cases time is measured by the weighted sum of order arrivals, where the weights vary with the size of the order.

When time is viewed in this economically fundamental manner the question of dependence or independence of the time change becomes an interesting and meaningful question. Certainly, some part of the order process and hence the time change, one would expect, is motivated by observations of the price process. This is the phenomenon of herding or runs on the asset. On the other hand if the market is dominated by independent analysts who view the market price as always providing us with the most efficient and accurate valuation of the asset, i.e. it is a discounted martingale under the right measure then there is no information to be extracted from prices that the market has not already extracted and so no analysts are motivated in their trades by observations of price movements. They are bound to seek independent, and as far as possible, private information, as the motivating basis of their trading decisions. This interpretation of the process suggests an independent time change. We also note that from a mathematical modeling viewpoint, it would be easier to work with independent time changes though it is possible and we shall see cases where both representations are possible for the same process. Generally, the independent time change is the more tractable alternative and so far most of our successes come from processes of this type. The broad consistency of this hypothesis with the efficient markets hypothesis is therefore an attractive feature.

3.1.4 Continuous Time Changes and Semimartingales

We come now to the crux of the issue, the continuity of the price process or otherwise. This brings us to the third and final result from the theory of stochastic processes shedding light on the nature of the price process as a consequence of no arbitrage. We note first that as the price process is a time changed Brownian motion, it will be a continuous process only if the time change is continuous. The implications of supposing such continuity in the time change rely on results characterizing continuous semimartingales (Revuz and Yor [39] page 190).

Let $X(t)$ be a continuous semimartingale, be it the price process or the time change. Let $V(t)$ be the quadratic characteristic of the semimartingale $X(t)$ which exists by virtue of X being a semimartingale. In the terminology of Wall Street the process $V(t)$ is akin to the realized total variance on the process $X(t)$. If the process $X(t)$ has a well defined sense of a variance rate per unit time, or equivalently $V(t)$ is differentiable in t then the quadratic characteristic is absolutely continuous with respect to Lebesgue measure and in this case we may write the process $X(t)$ as a stochastic integral with respect to Brownian motion. Under these conditions there exist processes $a(t)$, $b(t)$ and a standard

Brownian motion $W(t)$ such that

$$X(t) = X(0) + \int_0^t a(s)ds + \int_0^t b(s)dW(s). \tag{1}$$

Consider now the implications of $X(t)$ being a time change and the price process in turn. If $X(t)$ is a time change, then it is an increasing process and so $b(t)$ must be identically zero. This implies that the time change is locally deterministic with no uncertainty in local rate of time change which is then $a(t)$. If we view the time change, as suggested earlier, as a measure of economic activity, proxied by the rate of arrival of information, orders, or size weighted orders then one would expect some local uncertainty in the time change and this argues against the use of a locally deterministic time change and hence by implication, a continuous semimartingale as a model for the price process.

On the other hand if one views $X(t)$ directly as a price process, the representation (1) argues that the local motion of the stock return must be Gaussian. Given the considerable evidence cited against the likelihood of this possibility, we conclude once again that a continuous semimartingale is not an appropriate model for the price process. Now it is possible that there is a continuous martingale component in the price process in addition to a jump component as is the case of jump diffusions but the necessity, of introducing such a diffusion term, onto a functioning purely discontinuous model must be separately argued for. As we will observe, the latter class of models contain many alternatives capable of approximating very closely the structural characteristics of diffusions.

3.1.5 Summary of the consequences of no arbitrage

We showed in this section that no arbitrage implies via the existence of an equivalent martingale measure, that the price process is a semimartingale. We then observed that all semimartingales are time changed Brownian motions, time changed by a random increasing time change. The resulting process could be continuous only if the time change is locally deterministic. Relating time changes to measures of economic activity with some local uncertainty we argued that the price process was not a continuous process. We also observed that such continuity implies that the process is locally Gaussian, for which we have ample evidence to the contrary, and so once again we concluded that the process cannot be continuous. The remaining sections will take up the issue of modeling using purely discontinuous processes and demonstrate their effectiveness. The need to add on an additional continuous process onto a functioning purely discontinuous process must in our view be argued for on theoretical and empirical grounds.

4 ECONOMIC MODELS OF FINITE VARIATION FOR ASSET PRICE PROCESSES

Statistical and Economic analysis suggests that we entertain purely discontinuous price processes with possibly infinite arrival rates, and finite variation. An attractive feature of finite variation processes is that they may be decomposed as the difference of two increasing processes, a property lost in Brownian motion and other processes of infinite variation. This permits, for the first time, a separation of the price process into the process of up ticks and down ticks. Our analysis of optimal contracting in such economies indicates that the major demand for short maturity at-the-money options in such economies arises from a desire on the part of investors to be positioned differently with respect to upward and downward movements in the market, a position not attainable by direct stock investment alone. Hence options, and short maturity at-the-money options in particular, play a fundamental role in such economies: A role that may be consistent with casual observations of high activity in these markets. The next step forward from correctly adjusting ones delta or stock position is the optimal positioning of the up and down deltas via option trades. To effectively answer these questions it is imperative that we focus attention, separately, on the up and down forces of the market. We propose here two classes of models, accomplishing this objective. The models differ in their primitives and are structurally distinct, yet we show in the next section, that under some fairly reasonable conditions, they are in fact equivalent. However, tractability is enhanced by working with both specifications as it can be difficult to find the equivalent formulation from the alternate perspective.

The first class of models takes as primitives two increasing processes that represent cumulated orders to buy and sell at market and models the price responses as these orders are cleared through the limit sell and buy books respectively. Economic activity and the related concepts of economic time reflect cumulated orders of both types in this representation of the price process. We term this class of models the Order Processing Models (OPM).

The second class of models is related to traditional models of dynamic price adjustment with price changes expressed as a function of the level of excess demand in the economy. This response function is termed the force function of the economy as it measures price pressure in its relationship with excess demand. The excess demand itself is modeled by a Brownian motion with the equilibrium points given by the zero set of Brownian motion. Economic time in these models is given by cumulated squared price responses or the realized variance. This class of models we refer to as Dynamic Price Adjustment Models (DPA).

4.1 Prices in the Order Processing Model (OPM)

The primitives in this view of the price process are two increasing processes that represent cumulated market buy orders, $U(t)$, and cumulated market sell orders $V(t)$. We have noted in our discussion of time changes that increasing random processes with local uncertainty are necessarily purely discontinuous. By taking as primitives such increasing random processes, the fundamental uncertainties of the economy are discontinuous and prices modeled as market responses to such inherit this property. Defining the jumps in the processes $U(t)$ at time t by $\Delta U(t) = U(t) - U(t_-)$ where we note that the processes are by construction right continuous with left limits and $U(t) = \lim_{s \downarrow t} U(s)$ while $U(t_-) = \lim_{s \uparrow t} U(s)$ and like wise for $V(t), V(t_-)$ and $\Delta V(t)$. The property of being increasing and purely discontinuous implies that

$$U(t) = \sum_{s \le t} \Delta U(s)$$
$$V(t) = \sum_{s \le t} \Delta V(s)$$

so that the current value of each process is just the sum of all the jumps that have occurred to date.

Price changes are modeled in Geman, Madan and Yor [18] by market responses to these market buy orders. Here we describe the process of price increases. The magnitude $\Delta U(t)$ is viewed as a buy order at the prevailing price of $p(t_-)$ which by construction cannot be accessed. There is a downward sloping demand curve $q^{du}(p(t)/p(t_-), \Delta U(t), t)$ that is $\Delta U(t)$ at $p(t) = p(t_-)$ and an upward sloping supply curve $q^{su}(p(t)/p(t_-), \Delta U(t), t)$ that is zero at $p(t) = p(t_-)$ that must be equated to determine both the quantity transacted $q^u = q^{du} = q^{su}$ and the price response $p(t)$. The solution gives the price response in log form by

$$\ln\left(\frac{p(t)}{p(t_-)}\right) = \Phi^u(\Delta U(t), t).$$

A similar analysis yields the price response to a market sell order

$$\ln\left(\frac{p(t)}{p(t_-)}\right) = \Phi^v(\Delta V(t), t).$$

The price process is obtained as an aggregation of the price responses to market buy and sell orders

$$\ln(p(t)) = \ln(p(0)) + \sum_{s \le t} \Phi^u(\Delta U(s), s) - \sum_{s \le t} \Phi^v(\Delta V(s), s)$$

and is by construction the difference of two increasing processes, and therefore a finite variation process. It is also purely discontinuous in that it is precisely the sum of all its jumps. Geman, Madan and Yor [18] rewrite such processes in many cases as time changed Brownian motion and study the relationship between the time change and the market primitives, showing that the time change is generally a size weighted sum of the market buy and sell order processes. Hence their interpretation as measures of the level of economic activity.

4.2 The Dynamic Adjustment Model (DPA)

This formulation of the price process begins with a traditional price adjustment model of the form

$$\frac{d\ln(p)}{dt} = f(z(t))$$

where $z(t)$ is a measure of excess demand and f represents the force by which prices respond to excess demand in the economy. This function we term the force function of the economy. By construction $f(x) \geq 0$ for $x > 0$ and $f(x) \leq 0$ for $x < 0$.

Excess demand is exogeneously modeled as dominated by new information and is given by a Brownian motion $W(t)$. It follows that

$$\ln(p(t)) = \ln(p(0)) + \int_0^t f(W(s))ds.$$

Equilibrium times are of course given by the zero set of Brownian motion and there are arbitrage opportunities to be made during upward or downward rallies by buying or selling and then reversing the trade before the end of the rally. Such intra rally trades are not available to general market participants whose price access is only at equilibrium times. The restriction to equilibrium times, the zero set of Brownian motion, is accomplished by evaluating the above process at the inverse local time of Brownian motion at zero, $\sigma(t)$. We therefore define

$$\ln(p(t)) = \ln(p(0)) + \int_0^{\sigma(t)} f(W(s))ds \tag{2}$$

This process is once again a purely discontinuous process, inheriting this property from that of inverse local time. It may be decomposed as the difference of two increasing processes

$$\ln(p(t)/p(0)) = \int_0^{\sigma(t)} f^+(W(s))ds - \int_0^{\sigma(t)} f^-(W(s))ds$$

where $f^+(x) = f(x)\mathbf{1}_{(x\geq 0)}; f^-(x) = f(x)\mathbf{1}_{(x\leq 0)}$,and is a process of finite variation under the condition

$$\textstyle\int_{-K}^{K} |f(x)|\, dx < \infty \quad \text{for all } K.$$

It is interesting to enquire into the nature of the force function in the economy. For example, if $f(x) > 0$ for all $x > 0$ and $f(x) < 0$ for $x < 0$ then the price process is one with an infinite arrival rate of jumps. On the other hand there are finitely many jumps in any interval if $f(x) = 0$ in a neighborhood of zero. Another interesting question is whether the force is immediately infinite and decreasing for larger excess demands or whether it rises with the level of excess demand. Geman, Madan and Yor [18] present many explicit solutions that may be employed to answer such questions. They also show that such a process may be written as Brownian motion evaluated at a time change that aggregates the squared price responses and is thereby a measure of realized variance.

5 PRICES AS LÉVY PROCESSES

Finite Variation asset price processes are by construction the difference of two increasing processes and section 4 has described two classes of economic models that give rise to such processes. We now wish to construct specific examples of such processes that may be evaluated empirically in their adequacy as models for the statistical dynamics of the price process, and as models for the pricing densities reflected in option prices. This statistical evaluation is enhanced if one has effective descriptions of the transition densities for use in maximum likelihood estimation and closed form or otherwise fast and accurate computation methods for the prices of European options when the underlying process is in the described class.

Both these objectives are simultaneously met by an analytic closed form for the characteristic function of the log of the stock price at a future date. The density is then easily evaluated by Fourier Inversion and maximum likelihood estimation is feasible, alternatively one may also follow the methods outlined in Madan and Seneta [31] and estimate parameters by maximum likelihood on transformed variates. Option prices are easily obtained from the characteristic function and this is described in Bakshi and Madan [5] and a faster algorithm is provided in Carr and Madan [9]. Carr and Madan show how to analytically write the Fourier transform in log strike, of an exponentially damped call price, in terms of the characteristic function of the log stock price. The damped call price, and call price are then obtained by a single Fourier Inversion that may even invoke the Fast Fourier Transform. The characteristic function of the log stock price is therefore seen as the key to efficient model validation from both a statistical and risk neutral perspective.

5.1 The Characteristic Function of Log Price Relatives

In constructing alternatives to Brownian motion as models of the fundamental uncertainty driving the stock price, that may meet our requirements of being a purely discontinuous process of finite variation with a possibly infinite arrival rate of shocks, we focus in the first instance on keeping all the properties of Brownian motion except those that must be given up. We are well aware that just as more complex models allowing for stochastic volatility and correlations of various sorts can be constructed out of Brownian motions by combining them in various ways the same can be done with any candidate process that replaces Brownian motion.

The first property of Brownian motion that we seek to keep is the analytically rich property of being a process of independent increments, identically distributed over non-overlapping intervals of equal lengths of time. This introduces a homogeneity of the base uncertainty across time, that may be altered through parametric shifts in later developments. In any case, for modeling the local motion, homogeneity should be a reasonable hypothesis from at least the perspective of a local approximation that employs some average density of moves, even if the actual ones are state contingent and time varying.

The second property, which we may or may not keep is that of finite moments of all orders. We are modeling continuously compounded returns and this should in principle be a bounded random variable, even if it is difficult to organize this within a modeling context, and hence the finiteness of moments is really a nonissue. Considerations of analytical tractability may on occasion require us to consider processes with infinite moments, but my prior is to avoid them as far as possible.

The theory of stochastic processes has a lot to teach us about processes meeting these conditions. Such processes are called infinitely divisible and the Lévy-Khintchine theorem (See Feller [17] and Bertion [2]) provides us with a complete characterization of the characteristic function. Specifically, let $X(t) = \log(S(t))$ be the continuous time process for the log of the stock price with mean μt, and further suppose that $X(t)$ is a finite variation process of independent identically distributed increments then there exists a unique measure Π defined on $\mathbb{R} - \{0\}$ such that

$$\phi_{X(t)}(u) \stackrel{def}{=} E\left[\exp(iuX(t))\right] = \exp\left(iu\mu t + t\int_{-\infty}^{\infty}\left(e^{iux} - 1\right)\Pi(dx)\right).$$

The measure Π is called the Lévy measure of the process and $X(t)$ is a Lévy process. When the measure has a density $k(x)$, we may write

$$\phi_{X(t)}(u) = \exp\left(iu\mu t + t\int_{-\infty}^{\infty}\left(e^{iux} - 1\right)k(x)dx\right) \tag{3}$$

and we refer to the function $k(x)$ as the Lévy density.

Heuristically the density $k(x)$ specifies the arrival rate of jumps of size x and the Lévy process $X(t)$ is a compound Poisson process with a finite arrival rate if the integral of the Lévy density is finite. We shall primarily be concerned with Lévy processes with an infinite arrival rate. The Lévy process may always be approximated by a compound Poisson process obtained by truncating the Lévy density in a neighborhood of zero, and using as an arrival rate

$$\lambda = \int_{|x|>\varepsilon} k(x)dx$$

and as a density for the jump magnitude conditional on the arrival, the density

$$g(x) = \frac{k(x)\mathbf{1}_{|x|>\varepsilon}}{\lambda}.$$

The convergence occurs as we let $\varepsilon \to 0$. Geman, Madan and Yor [18] present many examples of candidate Lévy processes that are associated with the two economic models OPM and DPA of section 4.

5.2 Robustness of Finite Variation Lévy Processes

Continuous time processes with continuous sample paths have a certain lack of robustness best illustrated by considering geometric Brownian motion under two different but close volatilities. Two individuals could perhaps hold such different views on volatility but as a consequence their probability measures are no longer equivalent but are in fact singular. The set of paths receiving probability 1 under one measure has probability 0 under the other measure. The measures are not robust, in the sense of equivalence, to different volatility beliefs. This lack of robustness is really a consequence, not of continuity, but of infinite variation. Hence, remaining in the class of finite variation processes enhances robustness of the models to heterogeneity of views on various parameters.

To appreciate this point we note (Jacod and Shiryaev [27] page 159) that two Lévy processes with Lévy densities $k(x)$ and $k'(x)$ are equivalent just if there exists a positive measurable function $Y(x)$ such that

$$k'(x) = Y(x)k(x) \tag{4}$$

and

$$\int_{-\infty}^{\infty} \left|\left(|x| \wedge 1\right)\left(Y(x) - 1\right)\right| k(x)dx < \infty. \tag{5}$$

One may rewrite (5) on employing (4) as

$$\int_{k'<k} \left(|x| \wedge 1\right)\left(k(x) - k'(x)\right) dx + \int_{k'>k} \left(|x| \wedge 1\right)\left(k'(x) - k(x)\right)dx < \infty \tag{6}$$

and observe that on the set $|x| > 1$ the required integrability holds by virtue of the integrability of the Lévy densities on this set. On the set $|x| < 1$ we have the integrability condition

$$\int_{k'<k} |x| \left(k(x) - k'(x)\right)dx + \int_{k'>k} |x| \left(k'(x) - k(x)\right)dx < \infty$$

and this condition essentially requires that the difference between the two Lévy measures be a finite variation process and holds automatically if both Lévy processes are of finite variation. Hence for finite variation processes, equivalence just requires absolutely continuity of the measures with respect to each other or the condition (4) with no integrability conditions. Restrictions on the ability to change parameters like volatility in geometric Brownian motion follow from the integrability conditions for equivalence and apply to processes with infinite variation.

In this regard one may consider the Lévy measure studied in Geman, Madan and Yor [18] of the form

$$k(x) = \frac{e^{-x}}{x^{2+\alpha}} \text{ for } x > 0.$$

For $\alpha > 0$ this process has infinite variation and the parameter generating the infinite variation is α. This parameter cannot be changed if equivalence is to be preserved. Specifically, if

$$k'(x) = \frac{e^{-x}}{x^{2+\beta}}$$

for $\alpha \neq \beta$ and $\alpha, \beta > 0$ the two measures are no longer equivalent and it is the integrability condition (5) that fails.

5.3 Complete Monotonicity (CM)

There are of course many Lévy densities that one may employ in modeling the price process. It is therefore useful if the collection of possible choices can be reduced by invoking some structural properties. One such property is that of complete monotonicity. The idea is to require the arrival rates of large jumps to be less than the arrival rates of small jumps. This suggests that $k(x)$ be decreasing in $|x|$ or that $k'(x) \leq 0$ for $x > 0$ and $k'(x) \geq 0$ for $x < 0$. The first derivative of the Lévy density is therefore of one sign on each side of zero. The property of complete monotonicity requires that all the derivatives, and not just the first, have this property of having the same sign on each side of zero. By a result of Bernstein this property is equivalent to requiring $k(x)$ for $x > 0$ to be the Laplace transform of a positive measure on the positive half line and similarly for $k(x)$ for $x < 0$. Specifically we require that there exist measures G_p and G_n

$$k(x) = \int_0^\infty e^{-ax} G_p(da) \text{ for } x > 0$$
$$k(x) = \int_0^\infty e^{ax} G_n(da) \text{ for } x < 0.$$

The Lévy density is then a mixture of exponential densities. An important result that follows for such Lévy densities is that the two classes of economic models OPM and DPA are equivalent under the CM property.

5.3.1 Equivalence of OPM and DPA under CM

In particular, for every force function defining the price response under DPA, the resulting price process of equation 2 is a Lévy process with a completely monotone Lévy density. Geman, Madan and Yor [18] give numerous examples of force functions and their associated Lévy densities. For example, if the force function is x^m for some integer $m > 0$ then the process is one of independent stable increments with index $\alpha = (1/2 + m)^{-1}$.

Conversely, every Lévy process with such a completely monotone Lévy density can be written as the integral of a functional of Brownian motion up to the inverse local time of the Brownian motion. This equivalence result is an application of analytical results from number theory called Krein's theory and the specification construction of the force function from the Lévy density and vice versa remains a difficult, if not impossible task. Specifically, for the Variance Gamma model that we introduce next, we know the Lévy density quite explicitly but are not aware of what the force function is in this case.

6 THE VARIANCE GAMMA MODEL

Purely discontinuous processes of finite variation with infinite arrival rates contain a particularly tractable and parametrically parsimonious subclass of processes that is constructed from two very well known processes, Brownian motion and the gamma process. This is the "so-called" variance gamma process first studied by Madan and Seneta [32]. The process studied in Madan and Seneta [32], was the symmetric variance gamma process that is obtained on evaluating Brownian motion at gamma time. An asymmetric risk neutral process was developed by Madan and Milne [30] by assuming that a Lucas representative agent with power utility had to hold the risk exposure in a symmetric variance gamma process. It was shown in Madan, Carr and Chang [29] that the resulting risk neutral process was equivalent to evaluating Brownian motion with drift at gamma time. Given the importance of asymmetry or skewness in option pricing, we focus directly on this asymmetric variance gamma process but will refer to it as the variance gamma process. The process is parametrically parsimonious in that only two additional parameters are involved beyond the volatility introduced by Black and Scholes, and these two parameters give us control over skewness and kurtosis, that are precisely the primary concern in modeling and assessing derivative risks.

6.1 The Variance Gamma Process

Let $Y(t;\sigma,\theta)$ be a Brownian motion with drift θ and variance rate σ^2. If $W(t)$ is a standard Brownian motion, we may write the process $Y(t;\sigma,\theta)$ in terms of $W(t)$ as

$$Y(t;\sigma,\theta) = \theta t + \sigma W(t).$$

The variance gamma process is obtained on evaluating the process Y at an independent random time given by a gamma process. For this we define the process $G(t;\nu)$ with independent increments, identically distributed over nonoverlapping intervals of length h, with the increments, $G(t+h;\nu) - G(t;\nu) = g$, having the gamma density

$$p(g,h) = \frac{g^{h/\nu-1}\exp(-g/\nu)}{\nu^{h/\nu}\Gamma(h/\nu)}.$$

The mean of the gamma density is h and the variance is νh. Hence the average random time change in h units of calendar time is h and its variance is proportional to the length of the interval. The gamma density is infinitely divisible with characteristic function

$$E\left[\exp(iug)\right] = \left(\frac{1}{1-iu\nu}\right)^{h/\nu}$$

and the gamma process is an increasing Lévy process with a one sided Lévy density

$$k(x) = \frac{\exp\left(-\frac{x}{\nu}\right)}{\nu x}, \text{ for } x > 0.$$

Both the gamma process and Brownian motion are highly tractable processes about which a lot is known and each process has seen many domains of application. The variance gamma process is the process $X(t;\sigma,\nu,\theta)$ defined by

$$\begin{aligned} X(t;\sigma,\nu,\theta) &= Y(G(t;\nu);\sigma,\theta) \\ &= \theta G(t;\nu) + \sigma W(G(t;\nu)) \end{aligned} \tag{7}$$

or Brownian motion with drift θ and variance rate σ^2 evaluated at the gamma time $G(t;\nu)$. Apart from the variance rate of the Brownian motion σ^2, the two other parameters are θ and ν. We shall observe that it is θ that generates skewness while kurtosis is primarily controlled by ν.

6.1.1 Characteristic Function of the Variance Gamma Process

The characteristic function of the variance gamma process is easily evaluated by conditioning on the gamma process first and then employing the characteristic function of the gamma process itself. It has a simple analytic form of a quadratic raised to a negative power. Specifically,

$$\phi_{X(t)}(u) \stackrel{def}{=} E\left[\exp\left(iuX(t)\right)\right] = \left(\frac{1}{1 - iu\theta\nu + \frac{\sigma^2\nu}{2}u^2}\right)^{\frac{t}{\nu}} \tag{8}$$

The Black-Scholes and Merton model employing Brownian motion is a limiting case of this model since the process converges to Brownian motion with drift as one lets the volatility of the time change ν tend to zero. This may also be observed from the characteristic function on letting t/ν tend to infinity as ν tends to zero and noting that the limit is precisely $\exp(iu\theta t - \sigma^2 u^2 t/2)t$ the characteristic function of Brownian motion with drift.

We also note that if θ is zero, the characteristic function is real valued and the process is therefore symmetric and there is no skewness, hence validating the claim that skewness is generated by $\theta \neq 0$. This observation is even clearer once we have constructed the Lévy measure for the VG process.

6.1.2 Moments of the Variance Gamma Process

The moments of the VG process are easily obtained by exploiting the structure of the process or by differentiating the characteristic function. It is shown in Madan, Carr and Chang [29] that

$$\begin{aligned} E\left[X(t)\right] &= \theta t \\ E\left[\left(X(t) - E\left[X(t)\right]\right)^2\right] &= \left(\theta^2\nu + \sigma^2\right)t \\ E\left[\left(X(t) - E\left[X(t)\right]\right)^3\right] &= \left(2\theta^3\nu^2 + 3\sigma^2\theta\nu\right)t \\ E\left[\left(X(t) - E\left[X(t)\right]\right)^4\right] &= \left(3\sigma^4\nu + 12\sigma^2\theta^2\nu^2 + 6\theta^4\nu^3\right)t \\ &\quad + \left(3\sigma^4 + 6\sigma^2\theta^2\nu + 3\theta^4\nu^2\right)t^2 \end{aligned}$$

We observe again that skewness is zero if $\theta = 0$. Furthermore, in the case of $\theta = 0$ we have that the fourth central moment divided by the square of the second central moment or the kurtosis is $3(1+\nu)$. This leads to the interpretation that the parameter ν controls kurtosis and is in fact (for $\theta = 0$) the percentage excess kurtosis over the kurtosis of the normal distribution, which is 3.

6.1.3 The Variance Gamma Process as a Process of Finite Variation

The Variance Gamma process is a finite variation process and the two increasing processes whose difference is the variance gamma process are both gamma processes. This is observed by considering two independent gamma processes $\gamma_p(t)$ and $\gamma_n(t)$ with mean rates of μ_p, μ_n and variance rates ν_p, ν_n respectively for the positive and negative components. The characteristic functions of the two gamma processes are

$$E\left[\exp(iu\gamma_k(t)\right] = \left(\frac{1}{1 - iu\nu_k/\mu_k}\right)^{\mu_k^2 t/\nu_k} \quad \text{for } k = p, n.$$

Supposing that the two gamma processes have the same coefficients of variation and $\nu_k/\mu_k^2 = \nu$ for $k = p, n$ we may write the characteristic function of the difference of the two gamma processes as

$$E\left[\exp\left(iu(\gamma_p(t) - \gamma_n(t))\right)\right] = \left(\frac{1}{1 - iu\left(\frac{\nu_p}{\mu_p} - \frac{\nu_n}{\mu_n}\right) + u^2\frac{\nu_p}{\mu_p}\frac{\nu_n}{\mu_n}}\right)^{t/\nu}$$

The result follows on comparing this characteristic function with that of the variance gamma process and defining the mean and variance rates of the two gamma processes to be differenced accordingly. Specifically

$$\begin{aligned}
\mu_p &= \frac{1}{2}\sqrt{\theta^2 + \frac{2\sigma^2}{\nu}} + \frac{\theta}{2}, \\
\mu_n &= \frac{1}{2}\sqrt{\theta^2 + \frac{2\sigma^2}{\nu}} - \frac{\theta}{2}, \\
\nu_p &= \mu_p^2\nu, \\
\nu_n &= \mu_n^2\nu.
\end{aligned}$$

6.1.4 The Lévy density for the Variance Gamma Process

The Lévy density for the Variance Gamma process is easily constructed from its representation as the difference of two gamma processes using the well known form for the Lévy density of the gamma process. It follows that the Lévy density of the variance gamma process is

$$k_X(x) = \begin{cases} \frac{1}{\nu}\frac{\exp(-\frac{\mu_n}{\nu_n}|x|)}{|x|} & \text{for } x < 0 \\ \frac{1}{\nu}\frac{\exp(-\frac{\mu_p}{\nu_p}x)}{x} & \text{for } x > 0 \end{cases}$$

The basic form on the Lévy density is that of a negative exponential scaled by the reciprocal of the jump size. Just as in the gamma process, the integral of the Lévy density is infinite and the process is therefore a finite variation process with infinite arrival rates of jumps. It is helpful to write the Lévy density in terms of the original parameters of the process and this leads to the expression

$$k_X(x) = \frac{\exp\left(\frac{\theta x}{\sigma^2}\right)}{\nu\,|x|} \exp\left(-\frac{\sqrt{\frac{2}{\nu} + \frac{\theta^2}{\sigma^2}}}{\sigma} |x| \right) \tag{9}$$

The special case of $\theta = 0$ is a symmetric Lévy measure and hence the absence of skew. Negative values of θ give a fatter left tail and induce negative skewness. We also observe that as ν is increased the rate of exponential decay in the Lévy measure is reduced thus raising the arrival rate of jumps of the larger size. This induces the higher kurtosis related to this parameter. The two additional parameters therefore give direct control of the two moments that data analysis indicates we need to be able to control.

6.1.5 The Return Density for the Variance Gamma Process

The density of $X(t;\sigma,\nu,\theta)$ is available in closed form and is derived in Madan, Carr and Chang [29]. This is a closed form, in that it is expressible in terms of the special functions of mathematics, in particular the modified Bessel function of the second kind. Specifically we have that the density of $X(t) = x$ given $X(0) = 0$, $h(x,t;\sigma,\nu,\theta) = h(x)$ is

$$h(x) = \frac{2\exp\left(\frac{\theta x}{\sigma^2}\right)}{\nu^{t/\nu}\sqrt{2\pi}\sigma\Gamma(\frac{t}{\nu})} \left(\frac{x^2}{\frac{2\sigma^2}{\nu} + \theta^2} \right)^{\frac{t}{2\nu} - \frac{1}{4}} K_{\frac{t}{\nu} - \frac{1}{2}} \left(\frac{1}{\sigma^2} \sqrt{x^2 \left(\frac{2\sigma^2}{\nu} + \theta^2 \right)} \right). \tag{10}$$

There are three terms in the density, an exponential, a real power and the modified Bessel function. This is useful for maximum likelihood estimation of parameters from time series and it is also useful in providing density plots of results. Later we report on closed forms for option prices and this incorporates a closed form for the cumulative distribution function as well, that may be used to determine critical values for extreme points in value at risk calculations.

6.2 The Stock Price Process driven by a VG Process

We replace Brownian motion in the classical formulation of the geometric Brownian motion model by the VG process and define the risk neutral process for the stock price $S(t)$ by

$$S(t) = S(0)\exp\left(rt + X(t;\sigma,\nu,\theta) + \frac{t}{\nu}\ln\left(1 - \theta\nu - \frac{\sigma^2\nu}{2}\right)\right) \tag{11}$$

where r is the constant continuously compounded interest rate. Observe from the characteristic function of the VG process that

$$\begin{aligned}
E\left[\exp(X(t))\right] &= \phi_X(-i) \\
&= \left(\frac{1}{1-\theta\nu-\sigma^2\nu/2}\right)^{\frac{t}{\nu}} \\
&= \exp\left(-\frac{t}{\nu}\ln(1-\theta\nu-\frac{\sigma^2\nu}{2})\right)
\end{aligned}$$

and hence the mean rate of return on the Stock, under the risk neutral process, is the interest rate by construction.

We note further that the limit as ν tends to zero of $\frac{1}{\nu}\ln(1-\theta\nu-\frac{\sigma^2\nu}{2})$ is by L'Hopitals rule $-\theta-\sigma^2/2$ and so for small ν this term is $-\theta t-\sigma^2 t/2$. Noting that $X(t)=\theta G(t)+\sigma W(G(t))$ but for small ν, $G(t)$ is essentially t, we get that

$$\ln S(t) = \ln S(0) + (r-\frac{\sigma^2}{2})t + W(t)$$

or the familiar geometric Brownian motion model for the log of the stock price. Hence we have a generalization of the Black-Scholes and Merton models for the stock price. The generalization has introduced two new parameters ν,θ that we have observed give us control over skewness and kurtosis in the process.

6.2.1 Characteristic function of the log of the stock price

The characteristic function of the $\ln(S(t))$ is easily derived from that of $X(t)$, and is useful in deriving option prices by Fourier methods. Specifically we have that

$$\begin{aligned}
&\phi_{\ln(S(t))}(u) \overset{def}{=} E\left[\exp\left(iu\ln(S(t)\right)\right] \\
&= \exp\left(iu\left(\ln(S(0))+rt+\frac{t}{\nu}\ln(1-\theta\nu-\frac{\sigma^2\nu}{2})\right)\right)\phi_{X(t)}(u) \qquad (12)
\end{aligned}$$

where $\phi_{X(t)}(u)$ is the characteristic function of the VG process given in (8).

6.3 Variance Gamma Option Pricing

When the risk neutral process for the stock is described by the variance gamma process for the log of stock price as in equation (11), European call options on stock of strike K and maturity t have a price, $c(S(0);K,t)$ that is given by evaluating the expected discounted cash flow

$$c(S(0);K,t) = E\left[e^{-rt}\max\left(S(t)-K,0\right)\right]. \qquad (13)$$

This valuation result is an application of the defining property of a risk neutral probability, that traded asset prices, when discounted by the value of the money market account, are martingales under this probability. The valuation result follows on noting that option prices at maturity equal the promised payoff.

The computation of the call price in equation (13) is accomplished in closed form in Madan, Carr and Chang [29]. Other approaches at efficient computation employ Fourier inversion as described in Bakshi and Madan [5] or improvements thereof as explained in Carr and Madan [9]. We present here a brief summary of these results. The reader is referred to the original papers for further details.[4]

6.3.1 The Madan Carr and Chang Closed Form

The method employed by Madan, Carr and Chang [29] to develop a closed form for the VG option price relies on integrating the Black-Scholes formula applied to a random gamma time, with respect to the gamma density for this time. This approach requires the explicit computation of expressions of the form

$$\Psi(a,b,\gamma) = \int_0^\infty N\left(\frac{a}{\sqrt{u}} + b\sqrt{u}\right)\frac{u^{\gamma-1}\exp(-u)}{\Gamma(\gamma)}du, \tag{14}$$

where $N(x)$ is the cumulative distribution function of the standard normal variate. The call option price can be explicitly computed in terms of this Ψ function. Specifically we have that

$$\begin{aligned} c(S(0);K,t) = S(0)\Psi\left(d\sqrt{\frac{1-c_1}{\nu}}, (\alpha+s)\sqrt{\frac{\nu}{1-c_1}}, \gamma\right) - \\ K\exp(-rt)\Psi\left(d\sqrt{\frac{1-c_2}{\nu}}, \alpha\sqrt{\frac{\nu}{1-c_2}}, \gamma\right) \end{aligned}$$

where

$$\begin{aligned} s &= \frac{\sigma}{\sqrt{1+\left(\frac{\theta}{\sigma}\right)^2\frac{\nu}{2}}} \\ \alpha &= -\frac{\theta}{\sigma\sqrt{1+\left(\frac{\theta}{\sigma}\right)^2\frac{\nu}{2}}} \\ \gamma &= \frac{t}{\nu} \\ c_1 &= \frac{\nu(\alpha+s)^2}{2} \\ c_2 &= \frac{\nu\alpha^2}{2} \\ d &= \frac{\ln\left(\frac{S(0)}{K}\right)+rt}{s} + \frac{\gamma}{s}\ln\left(\frac{1-c_1}{1-c_2}\right) \end{aligned}$$

[4] Matlab programs are available for performing these computations in all the three ways described here.

A reduction of the Ψ function (14) to the special functions of mathematics is accomplished in terms of the modified Bessel function of the second kind and the degenerate hypergeometric function of two variables with integral representation (Humbert [24])

$$\Phi(\alpha, \beta, \gamma; x, y) = \frac{\Gamma(\gamma)}{\Gamma(\alpha)\Gamma(\gamma - \alpha)} \int_0^1 u^{\alpha-1}(1-u)^{\gamma-\alpha-1}(1-ux)^{-\beta} e^{uy} du.$$

Explicitly we have that

$$\begin{aligned}
\Psi(a, b, \gamma) = & \frac{c^{\gamma+\frac{1}{2}} \exp\left(sign(a)c\right)(1+u)^{\gamma}}{\sqrt{2\pi}\Gamma(\gamma)\gamma} \\
& K_{\gamma+\frac{1}{2}}(c)\Phi(\gamma, 1-\gamma, 1+\gamma; \frac{1+u}{2}, -sign(a)c(1+u)) \\
& - sign(a)\frac{c^{\gamma+\frac{1}{2}} \exp(sign(a)c)(1+u)^{1+\gamma}}{\sqrt{2\pi}\Gamma(\gamma)(1+\gamma)} \\
& K_{\gamma-\frac{1}{2}}(c)\Phi(1+\gamma, 1-\gamma, 2+\gamma; \frac{1+u}{2}, -sign(a)c(1+u)) \\
& + sign(a)\frac{c^{\gamma+\frac{1}{2}} \exp\left(sign(a)c\right)(1+u)^{\gamma}}{\sqrt{2\pi}\Gamma(\gamma)\gamma} \\
& K_{\gamma-\frac{1}{2}}(c)\Phi(\gamma, 1-\gamma, 1+\gamma; \frac{1+u}{2}, -sign(a)c(1+u))
\end{aligned}$$

where

$$\begin{aligned}
c &= |a|\sqrt{2+b^2} \\
u &= \frac{b}{\sqrt{2+b^2}}.
\end{aligned}$$

Madan, Carr and Chang [29] go on to employ this closed form in a detailed study of the empirical properties of VG option pricing, noting in particular the importance of skewness from the risk neutral viewpoint, and the ability of the VG model to flatten the implied volatility smile in option pricing.

6.3.2 Inversion of Distribution Function Transforms (Bakshi and Madan)

Bakshi and Madan [5] show that very generally one may write a call option price in the form

$$c(S(0); K, t) = S(0)\Pi_1 - K\exp(-rt)\Pi_2$$

where Π_1 and Π_2 are complementary distribution functions obtained on computing the integrals

$$\Pi_1 = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left[\frac{e^{-iuk}\phi_{\ln(S(t))}(u-i)}{iu\phi_{\ln(S(t))}(-i)}\right]du$$
$$\Pi_2 = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left[\frac{e^{-iuk}\phi_{\ln(S(t))}(u)}{iu}\right]du$$

where $k = \ln(K)$ and $\phi_{\ln(S(t))}(u)$ is the characteristic function of the log of the stock price given in this case by (12).

Bakshi and Madan [5] study the general spanning properties of the characteristic functions and their relationship to the spanning properties of options. They also express the general relationships between the two probability elements in option pricing providing a discussion of cases where they are analytically linked in their transforms.

6.3.3 Inversion of the Modified Call Price (Carr and Madan)

Carr and Madan [9] define the Fourier transform of the modified call price by

$$\psi(v) = \int_{-\infty}^{\infty} e^{ivk+\alpha k} c(S(0); e^k, t)dk$$

where $k = \ln(K)$, and the multiplication by $\exp(\alpha k)$ for $\alpha > 0$ and dampens the call price for negative values of log strike. They show generally that

$$\psi(v) = \frac{e^{-rt}\phi_{\ln(S(t))}(v-(\alpha+1)i)}{\alpha^2+\alpha-v^2+i(2\alpha+1)v}.$$

The call option price may then be obtained on a single Fourier inversion of ψ that may also employ the fast Fourier transform to evaluate

$$c(S(0); K, t) = \frac{\exp(-\alpha k)}{\pi}\int_0^\infty e^{-ivk}\psi(v)dv.$$

Carr and Madan [9] also consider other strategies for speeding up the pricing of options using the characteristic function of the log of the stock price and the methods should be useful for a variety of Lévy processes.

6.4 Results on Option Pricing Performance

The variance gamma option pricing model was tested in Madan, Carr and Chang [29] on data for S&P 500 options for the period January 1992 to September 1994. It was noted there that the skew is significant and the three parameter process effectively eliminates the smile in option prices in the direction of moneyness. The pricing errors are generally between 1 and 3 percent for options on the relatively liquid stocks and indices. The maturities we work with get fairly

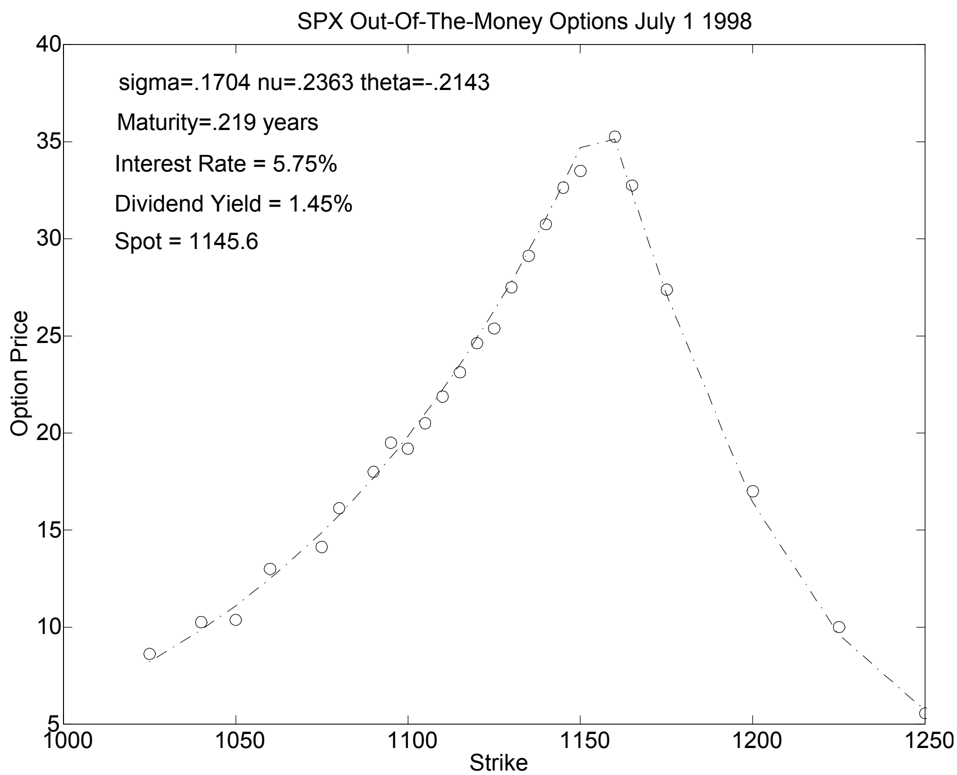


Figure 1: Out-of-the-money option prices on the SPX index and the price curve as fit by the VG model.

small and are as low as a couple of days at times, while the range of strikes are quite wide and may be up to 20 to 30 % out-of-the-money. Yet on this wide range of strikes and low maturities the model provides adequate fits.

Here we provide some illustrations of the results for options on the SPX and Nikkei indices. Figures 1 and 2 provide graphs of the prices of out-of-the-money options on these two indices along with the theoretical price curve as fit by the VG model. For strikes above at-the-money the options are calls while puts are used for the strikes below the spot. The typical V shaped price structure observed in markets is basically consistent with that of the negative exponential in the absolute value of the size of the move, that is the local structure of the VG model. The difficulty for Gaussian based models is precisely the fact that for these models option prices of out-of-the-money options fall of too rapidly, being a negative exponential in the square of the move, compared to market. We observe here that the essential structure of price decay is consistent with the building block of completely monotone Lévy densities, the double negative exponential.

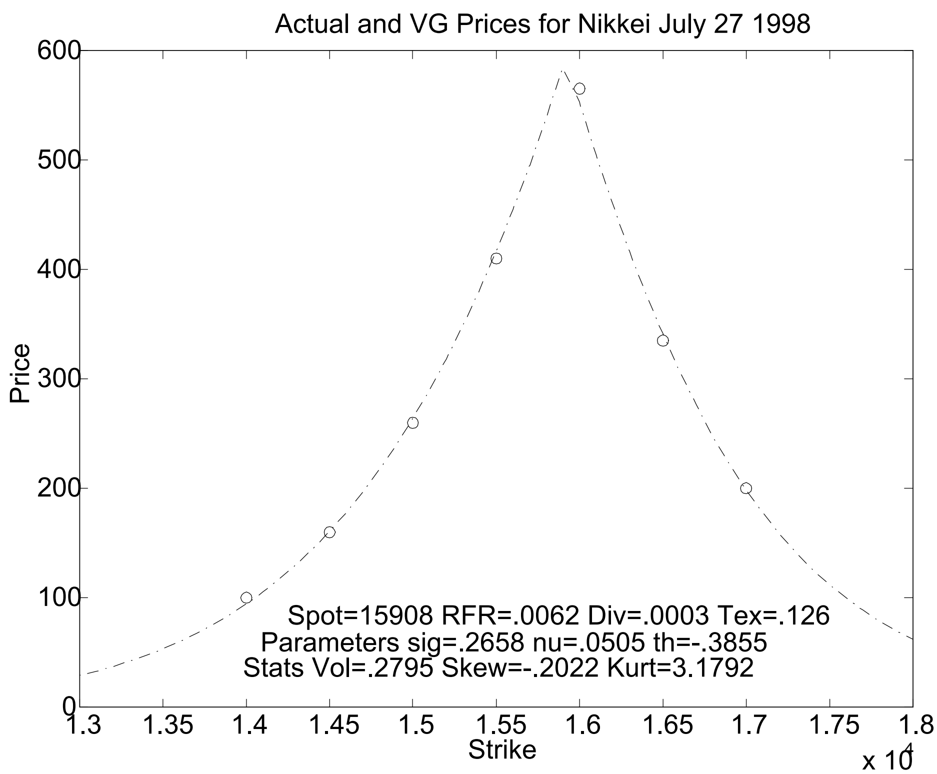


Figure 2: Out-of-the-money option prices on the Nikkei Index and the price curve fit by the VG model.

7 ASSET ALLOCATION IN LÉVY SYSTEMS

Apart from the successes of Lévy processes in option pricing, and the VG model in particular, these processes are associated with financial markets that are incomplete with respect to dynamic trading in the stock and the money market account. In such economies, with stock prices driven by an infinite arrival finite variation Lévy process, European options are market completing assets and one may study the question of the optimal demand for these assets by investors. In contrast, for the traditional economy, where options are redundant assets there is no demand for these assets.

With these observations in mind, Carr and Madan [8] proceed to reformulate the Merton problem for optimal consumption and investment, except now the asset space is genuinely expanded to include all the European options on the stock of all strikes and maturities as well. They study the problem of optimal derivative investment and solve it in closed form for HARA utility when the statistical and risk neutral price processes are in the VG class of processes. They also show that the shape of the optimal financial derivative product is independent of preferences, time horizons and the mean rate of return on the stock, factors that influence the level of investor demand but not the shape. The latter depends primarily on the comparison between the prices of market moves and the relative frequency of their occurrence. Their analysis also suggests that demand would be highest for at-the-money low maturity options in such economies, a fact that is in accord with casual market observations.

7.1 Optimal Derivative Investment

Consider an economy trading a stock with price process $S(t)$ that is a homogeneous Lévy process in the interval $[0, \Upsilon]$ with a Lévy density $k_P(x)$ defined over the real line where x represents the jumps in the log of the stock price. An example is provided by the VG process of equation (11). Also trading in the economy are options on this stock with strikes $K > 0$ and maturities $T < \Upsilon$. The prices of these options are given by the processes $c(S(t); K, T)$ for $t < T$ where these prices are consistent with the absence of arbitrage and are derived in line with martingale pricing methods using the risk neutral measure that is also a homogeneous Lévy process with Lévy density $k_Q(x)$. The subscripts P and Q make the important distinction between the statistical price process and the risk neutral process, with the former assessing the relative frequency of events while the latter assesses their prices.

In such an economy we wish to study the question of optimal derivative investment. At first glance, and in analogy with the solution methods adopted in Merton [35] this is a particularly difficult problem that is not going to be tractable from an analytical perspective. This is because we ask for the optimal positions in a doubly indexed continuum of assets viz. the options of all strikes $K > 0$ and maturities $T > t$ in a context in which many of these options (i.e. those with maturities below t) are expiring on us. Furthermore, the analytical pricing of these options is generally a complex exercise reflecting all the difficulties associated with the kinked option payoff.

For reasons of tractability, we reformulate the problem with the focus on the real uncertainty which is the jump in log price of the stock, x. We view investment, not as a decision on what assets to hold, but in the first instance as a design problem where the investor wishes to design the optimal response of his or her wealth to market moves represented by x. Hence we seek to determine the optimal wealth response function $w(x,u)$ which is the jump in the investor's log wealth if the market were to jump at time u by the amount x in the log price of the stock. The actual investment in options that delivers this optimal wealth response is a secondary problem that may be solved numerically using the spanning properties of options. The structure and solution of this secondary problem is described in further detail in Carr and Madan [29].

From the perspective of the optimal design of wealth responses, the optimal derivative investment problem may be formulated as a Markov control problem. Carr and Madan [8] consider both the infinite time horizon problem with intermediate consumption and the finite horizon problem with no intermediate consumption. Here we present just the former. We denote by $c(t)$ the path of the flow rate of consumption per unit time and suppose the investor has a preference ordering over consumption paths represented by expected utility evaluated as

$$u = E^P \left[\int_0^\infty \exp(-\beta s) U(c(s)) ds \right] \tag{15}$$

where P is the statistical probability measure, β is the pure rate of time preference, and $U(c)$ is the instantaneous utility function. The investor wishes to choose the consumption path $c(\cdot)$ and the wealth response design $w(\cdot)$ with a view to maximizing u.

The investor is constrained by his budget constraint that describes the evolution of his wealth. The wealth, $W(t)$, transition equation is the integral equation

$$\begin{aligned} W(t) &= W(0) + \int_0^t rW(s_{-})ds - \int_0^t c(s)ds \\ &\quad + \int_0^t \int_{-\infty}^{\infty} W(s_{-}) \left(e^{w(x,u)} - 1 \right) \left(m(\omega; dx, ds) - k_Q(x) dx ds \right). \end{aligned} \tag{16}$$

and the budget constraint requires that the wealth process be non-negative, $W(t) \geq 0$ almost surely. The first two terms of the wealth transition are standard and require no explanation, accounting for interest earnings and the financing of the consumption stream. The final term involves integration with respect to two measures, the first is the integer valued random measure $m(\omega; dx, ds)$ that is a Dirac delta measure counting the jumps that occur at various times of various sizes. The second is the pricing Lévy measure $k_Q(x)dxds$. The integration with respect to m accounts for the wealth changes actually experienced by the response design $w(x,u)$. The integration with respect to $k_Q(x)dxds$ accounts for the cost of this wealth response access that must be paid for through time.

The wealth transition equation (16) may be rewritten in a form more directly comparable to Merton's original equation by writing

$$
\begin{aligned}
W(t) = W(0) &+ \int_0^t rW(s_-)ds - \int_0^t c(s)ds \\
&+ \int_0^t \int_{-\infty}^{\infty} W(s_-)\left(e^{w(x,u)} - 1\right)(k_P(x)dxds - k_Q(x)dxds) \\
&+ \int_0^t \int_{-\infty}^{\infty} W(s_-)\left(e^{w(x,u)} - 1\right)(m(\omega; dx, ds) - k_P(x)dxds)
\end{aligned}
\tag{17}
$$

where we have just added and subtracted the integral of the wealth change with respect to the measure $k_P(x)dxds$. In this formulation the final integral in equation (17) is a martingale under the statistical measure P and matches the term representing the martingale component of stock investment in Merton [35]. The first two terms are the same as in Merton [35]. The third term matches the term that evaluates excess returns from stock investment in Merton [35]. Here excess returns are the expected wealth change less the cost or price of this change whereas in Merton we have $\mu - r$.

The investor's optimal derivative investment problem is to choose $c(\cdot), w(\cdot)$, with a view to maximizing the utility u of equation (15) subject to the budget constraint of equation (16).

7.2 Optimal Design of Wealth Responses

Let $J(W)$ be the optimized expected utility when the initial wealth $W(0) = W$. It is shown in Carr and Madan [8] that the optimal wealth response function for the infinite time horizon problem is homogeneous in time and satisfies the equation

$$\frac{J_W(We^{w(x)})}{J_W(W)} = \frac{k_Q(x)}{k_P(x)}. \tag{18}$$

This condition has an intuitive interpretation when it is rewritten as

$$\frac{J_W(We^{w(x)})k_P(x)}{k_Q(x)} = J_W(W)$$

which is that the expected marginal utility per initial dollar spent on cash in each state, x, is equalized across states. If this is not the case then $w(x)$ should be altered to move funds from states with a lower marginal utility to states with a higher marginal utility. Alternatively, the marginal rate of transformation in utility between two states must equal the marginal rate of transformation in markets between the same two states.

The optimal wealth response $w(x)$, is then determined from equation (18), if we know the function $J(W)$ as

$$w(x) = J_W^{-1}\left(J_W(W)\frac{k_Q(x)}{k_P(x)}\right).$$

We learn from this representation that the optimal wealth response design is a possibly smooth function J_W^{-1} applied to the ratio of two finite variation, infinite arrival rate Lévy measures. Such Lévy measures are kinked by construction at zero where the arrival rate goes to infinity. It follows that one would expect to see this property inherited by $w(x)$. This has the implication that at a minimum, optimal wealth response design positions investors with different slopes of their desired wealths with respect to up and down market movements, from at-the-money. Equivalently, there is a demand for short maturity at-the-money options.

7.2.1 HARA VG FINANCIAL PRODUCTS

In the special case when the statistical and risk neutral processes are in the VG class and the utility function $U(c)$ is in the $HARA$ (hyperbolic absolute risk aversion) class of utility functions, the optimal derivative investment problem of section 7.1 is shown in Carr and Madan [8] to have a closed form solution where $J(W)$ is also in the $HARA$ class of utility functions. The kinks in optimal designs discussed generally in section 7.2 can now be explicitly computed for this case.

Specifically, suppose the statistical Lévy measure is symmetric and given by

$$k_P(x) = \frac{1}{\kappa\,|x|}\exp\left(-\sqrt{\frac{2}{\kappa}}\frac{|x|}{s}\right) \tag{19}$$

where κ is the volatility of the statistical gamma time change for a symmetric Brownian motion with volatility s. Further suppose that the risk neutral Lévy measure is as given by (9) and parameters σ, ν, and θ. Let the utility function be

$$U(c) = \frac{\gamma}{1-\gamma}\left(\frac{\alpha}{\gamma}c - A\right)^{1-\gamma}.$$

In this case, defining

$$\begin{aligned}\zeta &= \frac{\theta}{\sigma^2}\\ \lambda &= \frac{1}{s}\sqrt{\frac{2}{\kappa}} - \frac{1}{\sigma}\sqrt{\frac{2}{\nu}+\frac{\theta^2}{\sigma^2}}\end{aligned}$$

and letting R denote the price relative of asset price post jump to its pre jump value then the optimal product takes the form

$$f(R) = \begin{cases} R^{-\frac{\zeta+\lambda}{\gamma}} & R > 1 \\ R^{-\frac{\zeta-\lambda}{\gamma}} & R < 1. \end{cases} \tag{20}$$

and the kink at-the-money is present unless $\lambda = 0$. The shape of this product is independent of the floor of the utility function and depends primarily on the statistical and risk neutral Lévy measures and risk aversion as represented by γ.

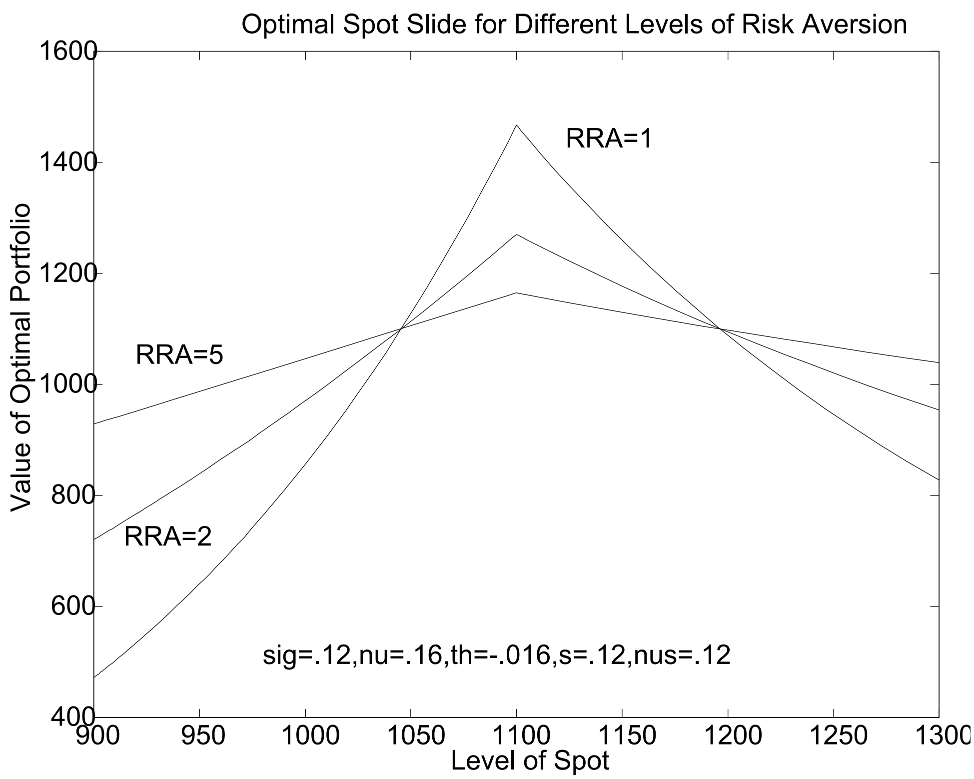


Figure 3: Optimal Spot Slides in the presence of excess risk neutral kurtosis and skew.

We also observe the clear impact of risk aversion on optimal product design. As we raise γ, the effect on this on the optimal wealth response $f(R)$ is to flatten out the movement in the optimal wealth response and to let the payoff approach that of a bond, thereby reflecting a lack of tolerance for movements in wealth.

A variety of possible shapes can arise for the optimal product and these are illustrated in Figures 3 to 6 for a variety of settings on the statistical and risk neutral parameters. Each figure reports three curves, for varying levels of risk aversion (RRA) and the flattening out of the response as we raise risk aversion is apparent in each case. Since these graphs draw optimal portfolio values against the level of the spot asset they are referred to as spot slides.

In Figure 3 the excess risk neutral kurtosis and skew leads to large moves being priced high relative to their likelihood and hence the optimal spot slide shorts these events and we have inverted V shape for the spot slide.

For Figure 4 the skew is strong and the kurtosis is mild. This leads to falls being overpriced while rises are under priced. The optimal slide is basically long the asset, but the positioning with respect to rises, the up delta, and falls, the

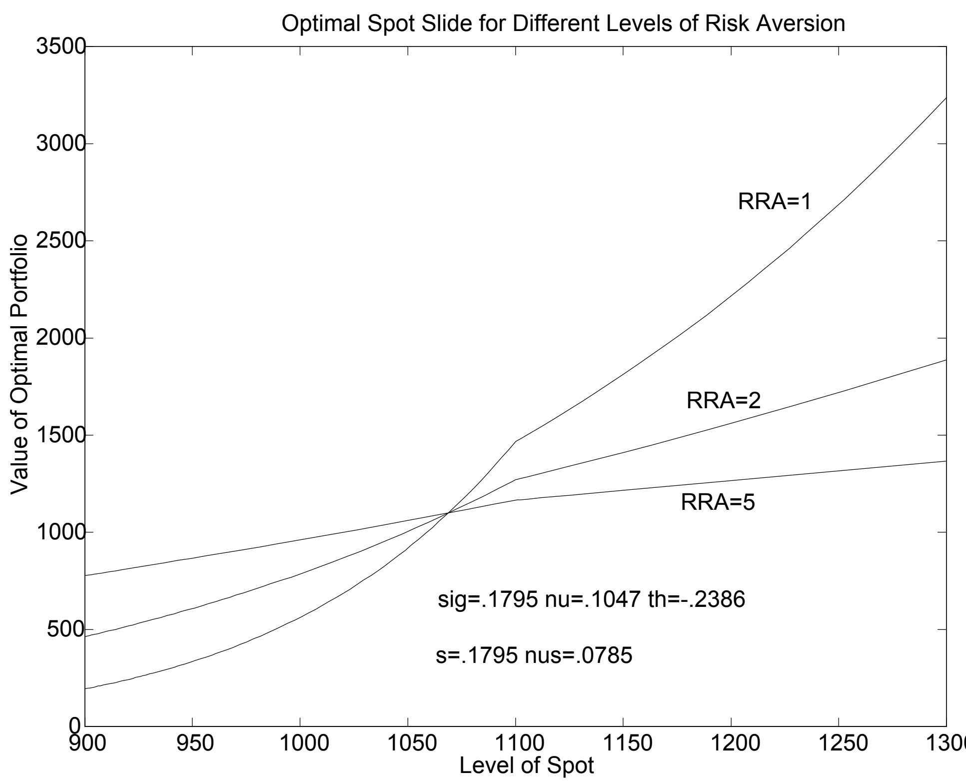


Figure 4: Optimal Spot Slide for a strong Skew and a mild excess kurtosis

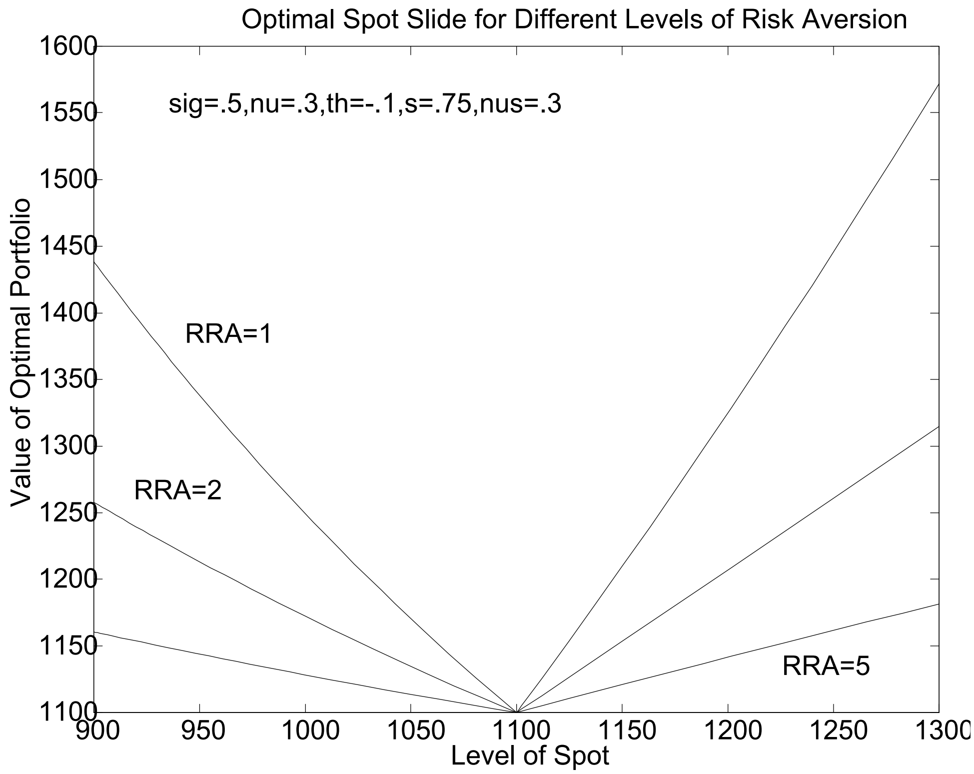


Figure 5: Optimal Spot Slide when statistical volatility dominates risk neutral volatility

down delta, differ.

For Figure 5 we have an excess statistical volatility making large moves relatively cheap securities. This gives rise to the V shaped optimal position.

Figure 6 is a reverse of the situation of Figure 4. The direction of the skew has been reversed and leads to a basically short position, with the kink induced by the behavior of the Lévy densities at the origin.

8 SPOT SLIDE CALIBRATION AND POSITION MEASURES

The inputs for constructing an optimal spot slide are fairly simple and require just the specification of the statistical or time series moments of the return distribution from which one may infer κ and s the statistical Lévy measure parameters. The next step is to obtain data on market option prices, preferably for short maturity options and then to estimate the risk neutral Lévy measure

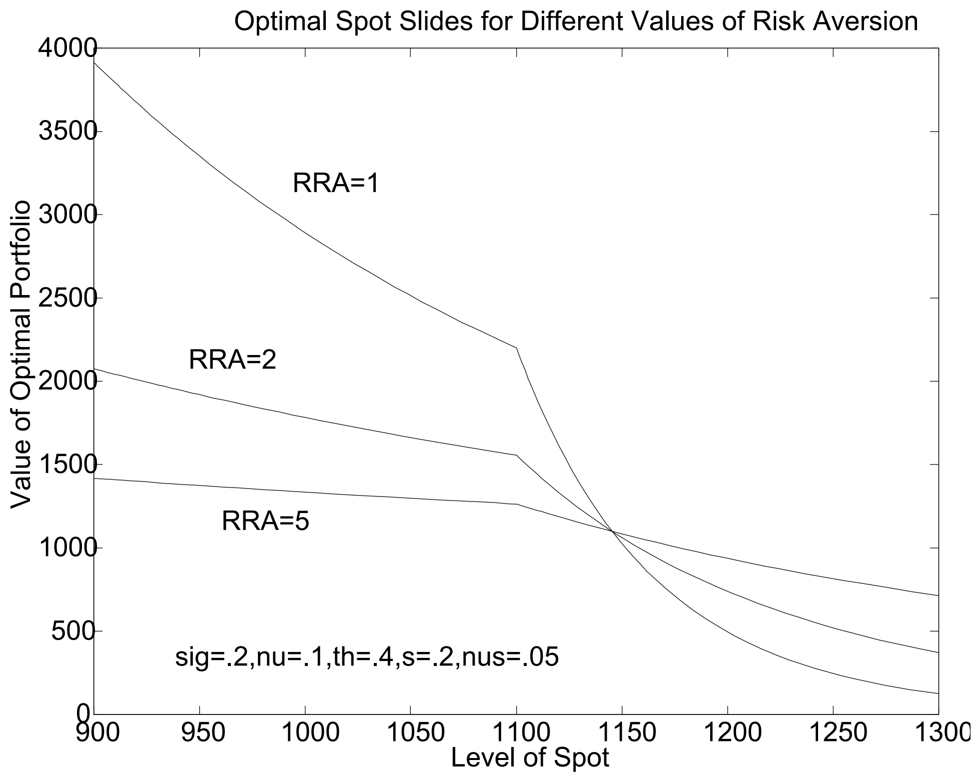


Figure 6: Optimal Spot Slide for a positive Skewness.

and the three parameters σ, ν and θ. Finally, making some assumption on the coefficient of relative risk aversion in a power utility function gives us γ and we are ready to graph the optimal spot slide describing how one should currently be positioned in the derivatives markets.

For a contrast, one may compare with the actual spot slide that aggregates a traders derivatives book and draws the response curve of his book value to market moves. We present here the results of calibrating optimal spot slides to data on actual spot slides. In the calibration we allowed for a reverse engineering of the coefficient of risk aversion γ as there is no other way to estimate this quantity. However, we also observed that the risk neutral excess kurtosis ν is typically an order of magnitude above its statistical counterpart κ and so we allowed this entity to be reverse engineered as well. Such an approach is defensible on noting that the variance of kurtosis estimates are of the order of the eighth moment and as the time series involved are not very long, generally 2 to 4 years, there is some leeway in an appropriate choice of this magnitude. The other parameters, σ, ν, θ, and s are taken at their estimated values.

For a variety of underlying assets and on a number of days, we reverse engineered the values of γ and κ so as to match the optimal spot slide with the actual spot slide observed for that day. Remarkably, we were able in many cases to come close to actual spot slides by just a simple choice on these two parameters (γ, κ). Figure 7 presents an example of an optimal spot slide as calibrated to an actual spot slide on a book of derivatives on a index. The ratio of κ to ν is referred to as β in the graph and describes the relative excess kurtosis of the subjective and risk neutral densities. Though it is often fairly small when calibrated, it is often an order of magnitude above the ratio of the statistical excess kurtosis to the risk neutral excess kurtosis.

Once all these parameters have been estimated and importantly γ and κ have been infered from data on the actual spot slide, one may infer a personalized risk neutral density given by the subjective Lévy measure determined by the parameters s and κ as described by equation (19) that is transformed by the marginal utility process as described in Madan and Milne [30] to obtain the personalized risk neutral Lévy measure, $k_I(x)$ (the subscript I being indicative of an individualized measure)

$$k_I(x) = \exp(-\gamma x) \frac{1}{\kappa |x|} \exp\left(-\sqrt{\frac{2}{\kappa}} \frac{|x|}{s}\right). \tag{21}$$

The Lévy measure (21) is that of a VG process with personalized values for

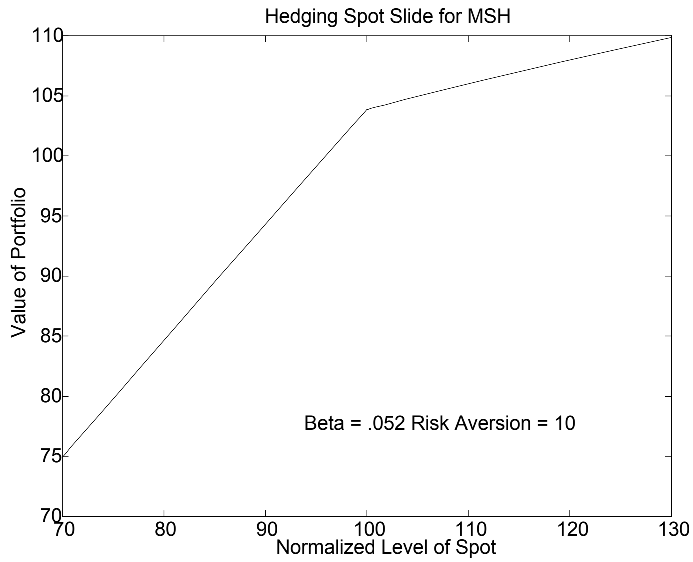


Figure 7: Optimal Spot Slide as calibrated to a book of derivatives on an index.

$\sigma_I, \nu_I, \theta_I$ given by

$$\sigma_I = \frac{s\sqrt{\frac{\kappa}{\nu}}}{\sqrt{1 - \frac{\gamma^2 s^2 \kappa}{2}}} \tag{22}$$

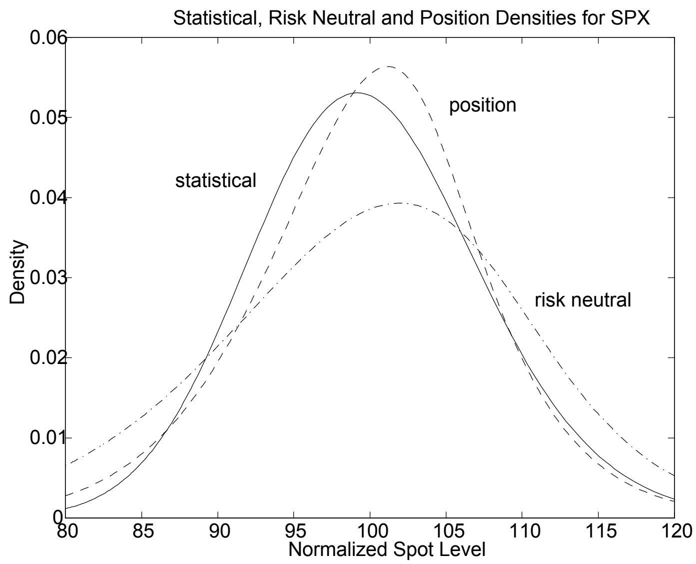
$$\theta_I = -\gamma \frac{\kappa}{\nu} \frac{s^2}{1 - \frac{\gamma^2 s^2 \kappa}{2}}$$

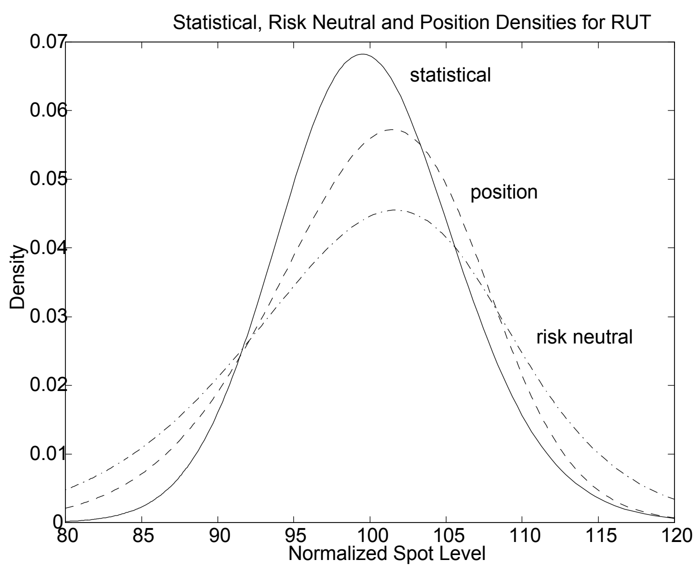
$$\nu_I = \kappa$$

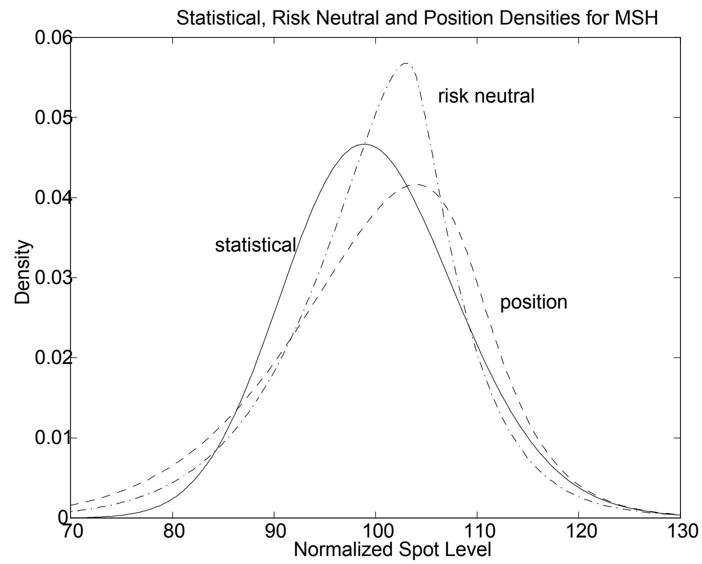
We thus infer a personalized risk neutral process and this may be employed to construct a personalized return density that we term a *position measure* as it is reverse engineered from derivative positions being viewed as optimal and therefore reflects preferences and beliefs that are obtained by a revealed preference exercise. All three densities are in the VG class of processes.

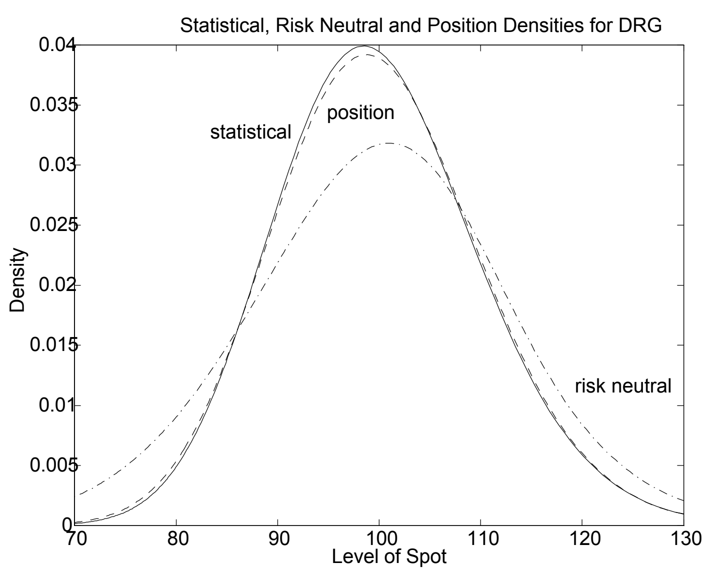
On completing this reverse engineering task we have available a statistical return density estimated from the times series of the return data, a risk neutral density as infered from options data, and a position density as reverse engineered from the actual spot slide of the derivatives book. Figures 9 ,10, 11 and 12 present a range of samples of graphs of these densities on a variety of underlying assets.

We observe a fairly diverse set of shapes of the densities, with varying degrees of skewness and kurtosis as reflected in the size of tails on the left and the right of the distribution. Furthermore, generally the position density is closer to the statistical density than the risk neutral density, reflecting the view that


Statistical, Risk Neutral and Position Densities for SPX
position
statistical
risk neutral
Density
Normalized Spot Level
0.06
0.05
0.04
0.03
0.02
0.01
0
80
85
90
95
100
105
110
115
120


Statistical, Risk Neutral and Position Densities for RUT
statistical
position
risk neutral
Density
Normalized Spot Level
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
80
85
90
95
100
105
110
115
120


Statistical, Risk Neutral and Position Densities for MSH
0.06
0.05
0.04
0.03
0.02
0.01
0
Density
risk neutral
statistical
position
70
80
90
100
110
120
130
Normalized Spot Level


Statistical, Risk Neutral and Position Densities for DRG
0.04
0.035
0.03
0.025
0.02
0.015
0.01
0.005
0
Density
statistical
position
risk neutral
70
80
90
100
110
120
130
Level of Spot

traders respect probability calculation as inferred from time series and position themselves accordingly given the market prices of market moves as reflected in the risk neutral distribution. Occasionally, however, as in the case of Figure 10 the position density may be skewed further to the left than even the risk neutral density and is reflective of greater risk aversion on the part of the trader than is prevalent in the market.

9 CONCLUSION

We argue here that empirical evidence on the statistical and risk neutral price processes for financial assets belong to the class of purely discontinuous processes of finite variation, albeit ones of high activity, as reflected by an infinite arrival rate of jumps. Structurally, the pattern of jump arrival rates is consistent with the hypothesis of complete monotonicity whereby arrival rates at smaller size levels are higher.

Economic considerations of the absence of arbitrage point in the same direction by demonstrating that semimartingales, the candidate no arbitrage price process, is a time changed Brownian motion and the increasing random process of the time change is of necessity purely discontinuous, if it is not locally deterministic. The attribute of finite variation is attractive from two perspectives, one that allows a separation of the up and down tick modeling of the market and we offer two representations of such price processes that are related under complete monotonicity of the Lévy density. The second attractive feature of finite variation is its robustness as reflected in its tolerance of parametric heterogeneity without the resulting measures being singular or disjoint in their sets of almost sure outcomes. This lack of robustness is an inherent property of infinite variation processes and we strongly advocate against the use of these processes as models for the price process unless there is overwhelming evidence in support of such a choice.

The class of stationary processes of independent and identically distributed increments meeting our requirements are characterized as a subclass of Lévy processes. Within this class, an important and analytically rich example is provided by Brownian motion time changed by a gamma process that combines in an interesting way two well studied processes in their own right. We summarize the properties of the resulting process termed the variance gamma process. The process has two additional parameters that enable it combat skew and kurtosis.

Option pricing under the variance gamma process is tractable using a variety of methods and we outline three such methods. The first is a closed form in terms of the modified Bessel function of the second kind and the degenerate hypergeometric function of two variables. The second involves two Fourier inversions for the complementary distribution function and the third employs direct Fourier inversion for the call price using the fast Fourier transform. The results of estimations are illustrated for data on SPX and Nikkei Index options. It is observed that the model eliminates the smile in the strike direction, using effectively for this purpose its two additional parameters.

Infinite arrival rate, finite variation, Lévy processes with completely monotone Lévy densities are processes for the stock price for which options are market completing assets that are part of the primary assets of the economy with a genuine demand for these assets by investors. We study the Merton problem of optimal consumption and investment with the asset space expanded to include out-of-the-money European options as investment vehicles. For $HARA$ utility and VG statistical and risk neutral processes this problem is solved in closed form with optimal portfolios that are kinked at-the-money and display a different slope with respect to upward and downward movements of the market. The positions reflect a role for at-the-money short maturity options, the most liquid end of the options market in practice.

Using our theory of optimal derivative positioning we illustrate how one may reverse engineer the preferences and beliefs of traders from observed spot slides of the derivatives book. This allows us to infer personalized risk neutral densities from observations on positions and we term this density the position density. Illustrations are provided, for comparative purposes of the statistical, risk neutral and position densities. It is observed that position densities are generally closer to the statistical density and lie between the statistical and risk neutral densities. At times however, they may be more skewed than the risk neutral density reflecting risk aversion that dominates market risk aversion.

Acknowledgement 1 *I would like to thank all my co-authors for all the hard work on the various aspects of this project. They are in approximate chronological order, Eugene Seneta, Frank Milne, Eric Chang, Peter Carr, Helyette Geman, Marc Yor and Gurdip Bakshi. This paper is the outcome of a series of talks I have given on the subject matter of the paper and would also like to thank the various institutions and forums that made these talks possible. They are Mathematical Finance Day, Boston, April 1998; Princeton University Seminar, October 1998; Courant Institute Seminar, New York University, October 1998; Midwest Probability Conference on Probability, Northwestern University, October 1998; Seminar in Financial Mathematics, University of Chicago, October 1998; International Press Lectures, Financial Mathematics, University of California at Irvine, November 1998; Quantitative Methods in Finance meeting, Sydney, December 1998; Seminar at ESSEC, January 1999; Seminar, University of Paris IX Dauphine, Paris, January 1999; Université Franche-Comté, Besancon, France, January 1999; the Bachelier Seminar at the Institute Henri Poincaré, Paris, January 1999; the workshop in Financial Mathematics, Oberwolfach, Germany, March 1999 and the Financial Engineering conference, New York, April 1999. The support and encouragement offered by Claudia Albanese, Marco Avellanada, Joseph Cherian, Carl Chiarella, Jaksa Cvitanic, Nicole El Karoui, Hans Föllmer, Robert Jarrow, Yuri Kabanov, Ioannis Karatzas, Vadim Linetsky, Vincent Lacoste, Eckhardt Platen, Marc Pinsky, Stan Pliska, Phillip Protter, Raymond Rishel, Martin Schweizer, Steve Shreve, Meté Soner, and Thaleia Zariphopoulou is also greatly appreciated. Finally I would like to acknowledge the assistance and guidance I have received from my co-workers at Morgan Stanley Dean Witter, they are Doug Bonard, Steven Chung, Georges*

Courtadon, Peter Fraenkel, Santiago Garcia, George George, Kevin Holley, Ajay Khanna, Harry Mendell, and Lisa Polsky. Any remaining errors are solely my responsibility.

References

[1] Bakshi, G. and Z. Chen (1997), "An Alternative Valuation Model for Contingent Claims," *Journal of Financial Economics*, 44, 123-165.

[2] Bertoin, J. (1996), *Lévy Processes*, Cambridge University Press, Cambridge.

[3] Breeden, D. and R. Litzenberger (1978), "Prices of State Contingent Claims Implicit in Option Prices," *Journal of Business*, 51, 621-651.

[4] Bakshi, G. and D.B. Madan (1998), "What is the Probability of a Stock Market Crash," Working Paper, University of Maryland.

[5] Bakshi, G. and D.B. Madan (1998), "Spanning and Derivative Security Valuation," Working Paper, University of Maryland.

[6] Bates, D. (1996), "Jumps and Stochastic Volatility: Exchange Rate Processes implicit in Deutschemark Options," *The Review of Financial Studies*, 9, 69-108.

[7] Black, F. and M. Scholes (1973), "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy*, 81, 637-654.

[8] Carr, P. and D. B. Madan (1998), "Optimal Derivative Investment in Continuous Time," Working Paper, University of Maryland.

[9] Carr, P. and D. B. Madan (1998), "Option Valuation Using the Fast Fourier Transform," Working Paper, University of Maryland.

[10] Cox, J.C., Ingersoll, J.E. and S.A. Ross (1985),"A Theory of the Term Structure of Interest Rates," *Econometrica*, 53, 385-408.

[11] Cox, J. and S.A. Ross (1976), "The Valuation of Options for Alternative Stochastic Processes," *Journal of Financial Economics*, 3, 145-166.

[12] Das, S. and S. Foresi (1996), "Exact Solutions for Bond and Options Prices with Systematic Jump Risk," *Review of Derivatives Research*, 1, 7-24.

[13] Delbaen, F. and W. Schachermayer (1994), "A General Version of the Fundamental Theorem of Asset Pricing," *Mathematische Annalen*, 300, 520-563.

[14] Derman, E. and I. Kani (1994), "Riding on a Smile," *Risk*, 7, 32-39.

[15] Dupire, B. (1994), "Pricing with a Smile," *Risk*, 7, 18-20.

[16] Embrechts, P. Kluppelberg, C. and T. Mikosch (1997), *Modeling Extremal Events,* Springer-Verlag, Berlin.

[17] Feller, W.E. (1971), *An introduction to probability theory and its applications*, 2nd Edition, Wiley, New York.

[18] Geman, H., Madan D.B. and M. Yor (1998), "Asset Prices are Brownian Motion: Only in Business Time," Working Paper, University of Maryland.

[19] Fama, E.F. (1965), "The behavior of stock market prices," *Journal of Business*, 38, 34-105.

[20] Harrison J.M. and D. Kreps (1979), "Martingales and Arbitrage in Multiperiod Securities Markets," *Journal of Economic Theory*, 20, 381-408.

[21] Harrison J.M. and S.R.Pliska (1981), "Martingales and Stochastic Integrals in the Theory of Continuous Trading," *Stochastic Processes and Their Applications*, 11, 215-260.

[22] Heston, S.L. (1993), "A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options," *The Review of Financial Studies*, 6, 327-343.

[23] Hull, John and Alan White (1987), "The Pricing of Options on Assets with Stochastic Volatility," *Journal of Finance*, 42, 281-300.

[24] Humbert, Pierre (1920), "The Confluent Hypergeometric Functions of Two Variables," *Proceedings of the Royal Society of Edinburgh*, 73-85.

[25] Jarrow, R.A. and D. Madan (1998), "Arbitrage, Equilibrium and Rational Bubbles," Working Paper, University of Maryland.

[26] Jacod, J. and A. Shiryaev (1998), "Local Martingales and the Fundamental Asset Pricing Theorems in the Discrete-Time case,"*Finance and Stochastics,* 3, 259-273.

[27] Jacod, J. and A. Shiryaev (1980), *Limit Theorems for Stochastic Processes*, Springer-Verlag, Berlin.

[28] Kreps, D. (1981), "Arbitrage and Equilibrium in Economies with Infinitely Many Commodities," *Journal of Mathematical Economics*, 8, 15-35.

[29] Madan, D.B., Carr P. and E. Chang (1998), "The Variance Gamma Process and Option Pricing," *European Finance Review*, 2, 79-105.

[30] Madan D.B. and F. Milne (1991), "Option Pricing with VG Martingale Components," *Mathematical Finance*, 1, 39-55.

[31] Madan, D.B. and E. Seneta (1989), "Characteristic Function Estimation using maximum likelihood on transformed variables," *Journal of the Royal Statistical Society*, ser. B, 51, 281-285.

[32] Madan, D.B. and E. Seneta (1990), "The Variance Gamma (V.G.) Model for Share Market Returns," *Journal of Business*, 63, 511-524.

[33] Merton, R. C. (1973), "Theory of Rational Option Pricing," *Bell Journal of Economics and Management Science*, 4, 141-183.

[34] Merton, R.C. (1976), "Option Pricing When underlying Stock Returns are Discontinuous," *Journal of Financial Economics*, 3, 125-144.

[35] Merton, R.C. (1971), "Optimum Consumption and Portfolio Rules in a Continuous Time Model," *Journal of Economic Theory*, 3, 373-413.

[36] Monroe, I. (1978), "Processes that can be embedded in Brownian Motion," *The Annals of Probability*, 6, 42-56.

[37] Naik, V. and M. Lee (1990), "General Equilibrium Pricing of Options on the Market Portfolio with Discontinuous Returns," *The Review of Financial Studies*, 3, 493-522.

[38] Press, J.S. (1967), "A Compound Events Model for Security Prices," *Journal of Business*, 40, 317-335.

[39] Revuz, D. and M. Yor (1994), *Continuous Martingales and Brownian Motion*, Springer-Verlag, Berlin.

[40] Rogers, C. (1997), "Arbitrage with Fractional Brownian Motion,"*Mathematical Finance,*7, 95-105

[41] Ross, S.A. (1976), "Options and Efficiency," *Quarterly Journal of Economics*, 90, 75-89.

[42] Ross, S.A. (1976), "Arbitrage Theory of Capital Asset Pricing," *Journal of Economic Theory*, 13, 341-360.