Jakša Cvitanić · Vassilis Polimenis · Fernando Zapatero

Optimal portfolio allocation with higher moments

Received: (date)/revised version: (date)

Abstract We model the risky asset as driven by a pure jump process, with non-trivial and tractable higher moments. We compute the optimal portfolio strategy of an investor with CRRA utility and study the sensitivity of the investment in the risky asset to the higher moments, as well as the resulting wealth loss from ignoring higher moments. We find that ignoring higher moments can lead to significant overinvestment in risky securities, especially when volatility is high.

Keywords Pure-jump processes · Optimal allocation · Higher moments

JEL Classification Numbers C61 · G11

The research of J. Cvitanić was supported in part by the National Science Foundation, under grants DMS 04-03575 and DMS 06-31366. Previous versions of this paper have been presented in seminars at USC, Cemfi (Madrid), and the 2006 Winter Meetings of the Econometric Society. We are grateful to Michael Johaness (the discussant), seminar participants, and an anonymous referee, for many comments and suggestions. Remaining errors are our sole responsibility.

Jakša Cvitanić
Caltech, Division of Humanities and Social Sciences, M/C 228-77, 1200 E. California Blvd. Pasadena, CA 91125
E-mail: cvitanic@hss.caltech.edu

Vassilis Polimenis
A. Gary Anderson Graduate School of Business of the University of California at Riverside. Riverside, CA 92521-0203
E-mail: polimenis@ucr.edu

Fernando Zapatero
Marshall School of Business, USC, Los Angeles, CA 90089-1427
E-mail: fzapatero@marshall.usc.edu

1 Introduction

Beginning with Merton (1971), a diffusion has been the standard model of uncertainty, despite empirical evidence that asset returns are not normally distributed. The literature has mostly implicitly assumed that investors are primarily affected in their decisions by the expected return and its variance, and therefore it was acceptable to focus on a distribution characterized by its first two moments.

The development of portfolio allocation theories for non-Gaussian economies has always been challenging, and has generally been met with limited success. Even though some authors have informally argued the contrary, the utility effect of ignoring higher moments may be substantial.[1] The typical approach is to extend early ideas by Rubinstein (1973), and Kraus and Litzenberger (1976), (1983) in developing models that account for higher moments. These models are non-parametric in nature in the sense that no specific distributional assumptions are made. On the other hand, the pricing relations are always approximate. This is because they result from a truncated Taylor expansion either of the underlying distribution, or of the discount factor. Furthermore, these models do not provide an idea of the size of the errors in the approximations, and different choices about what and where to truncate lead to different asset pricing formulae. Among others, Bansal, Hsieh and Viswanathan (1993), Bansal and Viswanathan (1993), and Chapman (1997), approximate a non-linear discount factor. Even though these models have improved empirical performance, it is not clear what equilibrium phenomena they capture.

An economically improved approach is to approximate a utility function by a Taylor series expansion, as in Harvey and Siddique (2000) or Dittmar (2002). Guidolin and Timmermann (2006) combine the above approach with the assumption that the distribution of asset returns is driven by a regime switching process. This approach retains many of the attractive features of the pricing kernels investigated in nonparametric analysis while avoiding many of their limitations. Yet, there are still several criticisms to the use of Taylor series expansions in the asset allocation context, with the most important being that the Taylor series expansion will converge to the true utility only under restrictive conditions.

Jondeau and Rockinger (2005) use a four-term Taylor expansion, but allow for time dependent distributions. An alternative method to study higher moments is "full-scale optimization" (see Adler and Kritzman, 2005, for an explanation of this method). Using this method, Cremers, Kritzman and Page (2004) find that the welfare cost of ignoring higher moments (as in mean-variance optimization) is not substantial.

Our study follows recent continuous time papers by Liu, Longstaff and Pan (2003) and Das and Uppal (2004). These dynamic models improve our ability to study the effect of skewness (and higher moments), since higher moment effects arise naturally due to the jumps in returns rather than being introduced explicitly through a utility function over the moments of the distribution of returns.[2] This approach avoids the truncation problems by providing tractable, closed-form, intertemporal portfolio allocation policies, for an investor with CRRA utility. Liu, Longstaff and Pan (2003) study portfolio choice with the possibility of a large negative idiosyncratic event which would introduce negative skewness in the

[1] For example, see the recent Harvey, Liechty, Liechty and Müller (2004) study.

[2] Independently of this objective, there is solid evidence of the presence of jumps in the time series of asset returns. See, for example, Eraker, Johannes and Polson (2003).

returns. Das and Uppal (2004) consider also a mixed diffusion-Poisson process but focus on systemic jumps. Choulli and Hurd (2001) consider the case of Lévy processes with deterministic jump arrival rates. They find the optimal portfolio for power and exponential utilities for such a case, and they also discuss the corresponding dual problems and shadow prices. Aït-Sahalia, Cacho-Diaz and Hurd (2006) solve the problem for a mixed process with multiple assets and jumps with deterministic arrival rate, for CARA and CRRA investors, focusing on the exposure to jump risk and the impact of jumps on the diversification.

Jumps may be of infinite (finite) activity if their rate of occurrence in any interval of time is certain (uncertain). From an economic standpoint, it is clear that, due to the continuous release of information, stock prices are unlikely to remain constant during any interval; thus we need processes that encapsulate high activity. In the models above, such high activity is generated by the diffusion part of a jump-diffusion, where the (finite activity) jump component is used to generate rare/extreme events.

Our view is that besides helping us understand the effects of skewness and kurtosis on portfolio allocation, the study of jump risks is crucial in understanding the existing plethora of diverse financial instruments when such instruments are redundant, as in complete markets. Lévy processes naturally lead to incomplete markets, where options are not replicable. In such markets, derivative securities are important for asset allocation, as it is demonstrated in Carr, Jin and Madan (2001). Essentially, jumps in the price process force agents to face "large" risks and can also play a role in explaining the need for risk management; our research should thus be relevant to the literature on prescribing capital requirements and on designing insurance contracts covering hedge fund losses. It can be argued that disentangling the pure jump from the diffusive component may be at the core of risk management, since risk managers shouldn't really care about the hedgeable diffusive noise.

We thus complement the literature on portfolio allocation with higher moments by providing the general portfolio effect of jumps that do not necessarily arrive at a slow rate, and are not necessarily large in size. The main analytical contribution of the paper is that we are able to solve the optimal portfolio allocation problem with jumps regulated by a stochastic state variable (which in the literature on Lévy processes has been frequently interpreted as trading activity or "volume"). Furthermore, we also make an observation important for implementation, that for a Lévy process the "moments" of the percentage returns are different from the "moments" of the log-price returns. For example, the volatility of both (percentage and log-price returns) is the same for a diffusion process, but not necessarily for jump processes. This point has led to some confusion in the fast growing financial literature on Lévy processes. For example, a symmetric Lévy process for the log-price does not lead to symmetric percentage returns.

When infinite activity is already generated by the employment of jumps, as in our model, a natural question arises as to whether it is necessary to also employ a diffusion component when modeling asset returns. A related question is whether the jump-diffusion paradigm is more appropriate. It is now becoming increasingly more agreeable that pure jump activity (of the infinite type) exhibits better empirical performance and is more capable of explaining asset pricing phenomena, especially related to option pricing. For example in the recent study by Carr, Geman, Madan and Yor (2002) – henceforth CGMY – a continuous time model that allows for both diffusion risk and jumps, of both finite and infinite activity, is employed to infer econometrically the fine structure of the price processes. Furthermore, CGMY allow the jump component to have either finite or infinite variation, and employ this model to study both the statistical process needed to assess risk and allocate investments and the risk-neutral process used for pricing and hedging derivatives. CGMY find that index returns tend to be pure jump processes of infinite activity and finite variation. Such studies have led a number of authors to argue that the jump components account for the entire activity in index return processes.

Thus, from an empirical point of view, we may dispense with diffusions in describing the fine structure of asset returns, as long as the jump process used is one of infinite activity, since such processes naturally capture both the high activity witnessed in real markets, as well as jumps of varying frequencies and sizes (and rare events). We introduce a time changed diffusion, with time changes that (when conditioned on a state variable[3]) are of a Gamma type – henceforth called the variance-Gamma or VG process. Similar processes have also been used in modeling stochastic volatility (e.g. Carr, Geman, Madan, Yor (2003)). Our process exhibits infinite activity jumps (in both directions), and is motivated by its counterpart without a state variable that has become popular for modelling security prices since it was introduced by Madan and Seneta (1987) and Madan and Milne (1991). Additionally, the VG process seems to fit particularly well some return features, especially with an eye towards option pricing: see Madan, Carr and Chang (1998), Carr, Geman, Madan and Yor (2002) and Carr and Wu (2003, 2004). The VG process is constructed by taking a standard Brownian motion process and sampling it at random times, as given by a Gamma process. In other words, the driving process is $W_{\tau(t)}$ where W_t is a Brownian Motion, and $\tau(t)$ is a random time change of the calendar time. Clark (1973) was the first to propose that we focus on economic activity, rather than calendar time when measuring returns, so that the time change would correspond to the amount of transactions, or trading volume. Geman, Madan and Yor (2001) argue that if the time change is not locally deterministic, then market prices must be purely discontinuous. Geman, Madan and Yor (2002) address the recovery issue, i.e. how much we may learn about trading activity by observing prices. Ané and Geman (2000) show that the rate of the economic activity can be proxied by transaction volume. We assume a stochastic intensity of price jumps (which would be an indicator of volume, in the original Clark 1973 motivation). When the standard Black-Scholes-Merton type models are monitored at random times, (caused, for example, by random trade arrivals) the resulting dynamics are of a general Lévy type.

We find that observed skewness and kurtosis would lead to lower holdings in risky securities than the standard Merton (1971) model would recommend. For the same level of skewness and kurtosis, and correcting for the risk premium that would leave the optimal allocation in a Gaussian setting at the same level, we find that the overinvestment increases with market volatility. We also compute the wealth loss equivalent resulting from overinvestment in the presence of higher moments. Although modest for low volatility settings, it becomes more important as volatility increases. Furthermore, we show that the wealth loss equivalent resulting from overinvesting is always higher than in the Merton (1971) benchmark. The difference is not substantial, but as argued by Brennan and Torous (1999), in a CRRA setting, the wealth loss resulting from overinvesting in the risky asset is small.

[3] This state variable can take various economic meanings depending on the context: stochastic volatility, trading volume, economic or business activity etc. To keep the discussion general, we don't "name" the state variable.

The paper is organized as follows. The first result of section 2 is general, since it applies to the universe of Lévy process driven stocks with the jump arrival intensity conditional on a stochastic state. In section 3 we focus on pure jumps, and we also extend the solution to the many-stocks case, where the jump induced correlation may be either due to the systemic risk (macro effects), as in the context of Das and Uppal (2004), when the jumps exhibit finite activity (low frequency), or due to the micro-structure risk when the jumps exhibit high activity. In section 4, we specialize to the case of stocks where the state captures a "trading volume" type quantity, and the stock compensates the risky states by a proper Sharpe ratio. In this case the results are tractable; the portfolio choice becomes constant. In section 5 we further specialize the model to a particular jump type so that we may derive an exact solution that can be calibrated to asset returns data, and that will allow us to compute optimal portfolios numerically. In section 6 we present and analyze some numerical examples.

2 The general case: An investment model with diffusive and jump risks

In this section we consider a model for the stock price, which is general in the sense that we only assume that the jump arrivals' rates depend on a state variable. This state variable can be interpreted as capturing the "market micro-structure" environment for the stock, for example. We assume that we are given a probability space and a filtration, and all the processes in the paper are adapted to that filtration. There are two securities in our model. There is a risk-free security (bond or bank account) that pays a locally deterministic interest rate r_t, so that the value of the B of this security evolves according to the dynamics

$$dB_t/B_t = r_t dt \tag{1}$$

There is also a risky security, a stock, with stock price process S, and dynamics subject to jump risk. Specifically, the stock's *log* return follows

$$\log S_t - \log S_0 = \int_0^t c_s ds + \int_0^t \sigma_{d,s} dZ_s + X_t \tag{2}$$

where c_t is an adapted process representing a continuous rate of return, and $\sigma_{d,t}$ the diffusive part of the stochastic volatility for the stock. While Z_t is a standard brownian motion, X_t is a pure jump process given by

$$X_t - X_{t-} = \int_{-\infty}^{+\infty} x N(dt, dx) \tag{3}$$

N is a Poisson random counting measure on $R_+ \times R$. We denote by $\Pi(t, dx)$ its compensator measure,

$$E \int_{\mathbb{R}} \phi(t, x) N(dt, dx) = \int_{\mathbb{R}} \phi(t, x) \Pi(t, dx) dt \tag{4}$$

for any measurable, random function $\phi(t, x) = \phi(\omega, t, x)$. For the technical facts regarding stochastic calculus of such jump process, we refer the reader to Jacod and Shiryaev (1987).

In this case, there are three return components, one continuous and locally deterministic, one continuous but stochastic, and another discontinuous, that is,

$$d(\log(S_t)) = c_t dt + \sigma_{d,t} dZ_t + \int_{-\infty}^{\infty} x N(dt, dx) \tag{5}$$

Percentage returns share the continuous risky component, but their jump component differs, and we have (using Itô's rule for jump processes)

$$dS_t/S_t = \left(c_t + \frac{1}{2}\sigma_{d,t}^2\right) dt + \sigma_{d,t} dZ_t + \int_{-\infty}^{\infty} (e^x - 1) N(dt, dx) \tag{6}$$

In purely diffusive dynamics, due to path continuity, the translation from log-returns to percentage returns (the ones the investor really cares about) results in an increase in drift, while keeping volatility the same. Essentially, since diffusive dynamics are locally Gaussian there are no higher cumulants to consider. As we will discuss herein, when the stock dynamics include jumps, one has to carefully account for the different effect such jumps have on the percentage stock return's moments. The first difference is in the drift: while the expected log growth equals

$$c_t + \int_{-\infty}^{+\infty} x \Pi(t, dx)$$

since a log jump of size x implies a percentage return $e^x - 1$, the stock's drift μ_t equals,

$$\mu_t = c_t + \frac{1}{2}\sigma_{d,t}^2 + \int_{-\infty}^{\infty} (e^x - 1) \Pi(t, dx) \tag{7}$$

The structure of (6) resembles that of stock dynamics with jumps, first considered by Merton (1971) and more recently in Liu, Longstaff and Pan (2003). In those papers, stocks follow a jump-diffusion process with Poisson arrivals. Here, instead, the jumps are not constrained to arriving at a slow Poisson rate, but may arrive at extremely fast (potentially infinite) rates.

2.1 The state variable

The intensity of the jump arrivals' rate for all jump sizes is provided by the measure $\Pi(t, dx)$. In order to allow for more realistic stock dynamics, we allow the jump arrival rates $\Pi(t, dx)$ to be stochastic, through a dependence on a state variable v_t, which for some $m_v = m_v(v_t, t)$ and $\sigma_v = \sigma_v(v_t, t)$, follows a positive process v_t given by

$$dv_t = m_v dt + \sigma_v dZ_t^v \tag{8}$$

for a Brownian Motion Z_t^v. Assume that the return rate $c_t = c(v_t, t)$ is a deterministic function of v and t. Since v_t captures the entire information about the jump arrival rates, the stock follows a *conditional* Lévy process; that is, given v_t, the stock return process has independent increments. More specifically

$$\Pi(\omega, t, dx) = \Pi(v_t(\omega), dx) \tag{9}$$

that is, the process X_t has independent increments given the σ-algebra $\mathcal{F}$ generated by v_t. Heuristically, this amounts to assuming that, when the state is such that $v_t = v$, the infinitesimal behavior of the log return process is that of a Lévy process with drift rate $c(v, t)$ and Lévy measure $\Pi(v, \cdot)$.

We further assume that the jumps in stock returns are somewhat "smooth" in the sense that the jump paths exhibit finite variation, that is

$$\int_{-\infty}^{\infty} (|x| \wedge 1)\Pi(v_t, dx) < \infty, \quad a.s., \quad \text{for all } v_t \tag{10}$$

Even though one may think of this state variable as capturing the current level of trading activity or some other relevant micro-structure quantity, here, in order to keep the discussion general, we will not assign a specific economic meaning to v, and will only refer to it as the state variable.

2.2 The Investor

We consider an investor with CRRA power utility, given by

$$U(x) = \frac{1}{1-\gamma} x^{1-\gamma} \tag{11}$$

We assume that the investor maximizes utility of optimal terminal wealth at some future time T, that we denote by $U(W_T)$. The case $\gamma = 1$, corresponds to logarithmic utility. The power utility is considered in Merton (1971), and we will use it as a benchmark in the following sections.

Starting with a positive initial wealth W_0, and given the opportunity to invest in the riskless and risky assets, at each time t, $0 \leq t \leq T$, the investor has to decide what proportion of her wealth, π, to invest in the risky security whose dynamics are given by (6). The rest of the wealth, the proportion $(1 - \pi)$, is invested in the bond (1).[4] The objective of the investor is, then,

$$\max_{\{\pi\}} E \frac{1}{1-\gamma} W_T^{1-\gamma} \tag{12}$$

subject to the budget constraint,

$$dW_t = W_t \left(r_t + \pi_t (c_t + \frac{1}{2}\sigma_{d,t}^2 - r_t) \right) dt + \pi_t W_t \sigma_{d,t} dZ_t + \pi_{t-} W_{t-} \int_{-\infty}^{\infty} (e^x - 1) N(dt, dx) \tag{13}$$

2.3 Optimal Investment Strategy

Given that financial markets are incomplete in the presence of jumps of random size, we determine the optimal investment using standard stochastic dynamic programming rather

[4]All the relevant integrability assumptions needed in this model, in particular for the dynamic-programming HJB equation (used below) to hold, can be found in Øksendal and Sulem (2004); specifically, see their Theorem 3.1. These assumptions are purely technical, and do not contribute to the finance intuition. We therefore adopt the usual approach in the finance literature: We write down the dynamic programming HJB equation, we guess the solution, and we verify that it solves the equation.

than the martingale pricing approach. As usual, we will derive the optimality equation, we will "guess" a solution, and will prove it is indeed a solution. We assume that the drift, diffusive volatility and the interest rate, are deterministic functions of time and state variables:

$$c_t = c(v_t, t),\ \sigma^2_{d,t} = \sigma^2_{d,t}(v_t, t),\ r_t = r(v_t, t). \tag{14}$$

Following Merton (1971), we define the indirect utility function (which will then be a function of (W, v, t)), as

$$J(W, v, t) = \max_{\{\pi_s, t \le s \le T\}} E_t U(W_T) \tag{15}$$

We use the dynamic principle approach to stochastic optimal control, which leads to the following Hamilton-Jacobi-Bellman (HJB) equation for the indirect utility function J,

$$\max_{\pi} \frac{1}{2}\sigma_v^2 J_{vv} + m_v J_v + \left(r + \pi\left(c + \frac{1}{2}\sigma^2_{d,t} - r\right)\right) W J_W + \frac{1}{2}\pi^2 W^2 \sigma^2_{d,t} J_{WW} \tag{16}$$
$$+ \int_{-\infty}^{\infty} \big[J(W(1 + \pi(e^x - 1)), v, t) - J(W, v, t)\big]\Pi(v_t, dx) + J_t = 0$$

where J_W, J_v and J_t denote first partial derivatives of $J(W, v, t)$, and similarly for higher derivatives. We solve this equation by assuming (and then verifying) that the indirect utility function is of a *separable* functional form

$$J(W, v, t) = \frac{1}{1-\gamma} W^{1-\gamma} F(v, t) = U(W) F(v, t) \tag{17}$$

where $F(v, t)$ is a deterministic function capturing the "investment opportunity" that depends on calendar time, and the current state.

Here is the main theoretical result of the paper:

Theorem 1 *Assume that (14) holds and that there is a solution J to (16). Also assume that there is a deterministic function $\pi^*(v, t)$ of (v, t) that solves the following equation:*

$$\gamma \sigma^2_{d,t} \pi^* = \mu - r + \int_{-\infty}^{\infty} \left(R(\pi^*, x)^{-\gamma} - 1\right)(e^x - 1)\Pi(v_t, dx) \tag{18}$$

where

$$R(\pi, x) = 1 + \pi(e^x - 1) \tag{19}$$

is the portfolio return due to a log-jump of size x. Finally, assume that there is a solution $F(v, t)$ to the Partial Differential Equation

$$\frac{1}{2}\sigma_v^2 F_{vv} + m_v F_v + (1-\gamma)\left(r + \pi^*\left(c + \frac{1}{2}\sigma^2_{d,t} - r\right)\right) F - \frac{1}{2}\gamma(1-\gamma)\pi^2\sigma_d^2 F \tag{20}$$
$$+ F \int_{-\infty}^{\infty} \big[R(\pi^*, x)^{1-\gamma} - 1\big]\Pi(v_t, dx) + F_t = 0.$$

with $F(v, T) = 1$. Then, J is the indirect utility function of the form (17), and the optimal investment strategy is given by π^*.

Proof: In Appendix. □

When the investor faces purely diffusive stock dynamics with volatility σ_d, the optimal investment becomes the usual

$$\pi = \frac{\mu - r}{\sigma_d^2 \gamma} \tag{21}$$

Furthermore, observe that here the market incompleteness is manifested by the fact that even though the agent would optimally like to independently choose her wealth exposure (Carr, Jin, Madan (2001)) for every jump x, by being able to only invest in the stock and the risk free asset she can only choose from a single-parametric family of $R(\pi, x)$ functions (19).

Even though prices here determine the optimal allocation, and not the other way around, equation (18) is an interesting *incomplete market* equation, analogous to the traditional asset pricing equations. More specifically, interpret $R(\pi, x)^{-\gamma} - 1$ as the percentage jump in the marginal rate of substitution before and after a jump, while $e^x - 1$ is the asset's percentage return due to the jump. The $\int_{-\infty}^{\infty} (R^{-\gamma} - 1)(e^x - 1)\Pi(dx)$ factor in (18) measures the covariance between jumps in the stock and the investor's marginal utility. Clearly, when the investor is heavily exposed to the stock, positive stock returns will coincide with rich states (low marginal rate of substitution), and the integral above will become very negative, thus limiting the investor's optimal exposure π. For assets with a high return premium $\mu - r$, the investor will thus seek a large exposure until (18) is satisfied.

3 The pure jump case

Since in the calibration exercise we will perform later we will use a pure jump process, it is useful to focus on the pure jump case, and for that purpose we provide a version of Theorem 1 when there is no diffusive risk, and the stock satisfies

$$dS_t/S_t = c_t dt + \int_{-\infty}^{\infty} (e^x - 1) N(dt, dx) \tag{22}$$

Corollary 1 *When there is no diffusive risk, i.e., $\sigma_{d,t} = 0$, the optimal allocation π^* is the deterministic function of (v, t) that satisfies*

$$\mu_t - r_t + \int_{-\infty}^{\infty} \left(R(\pi^*, x)^{-\gamma} - 1\right)(e^x - 1)\Pi(v_t, dx) = 0 \tag{23}$$

and $F(v, t)$ is the solution to the Partial Differential Equation

$$\begin{aligned} &\frac{1}{2}\sigma_v^2 F_{vv} + m_v F_v + (1-\gamma)\left(r + \pi^*(c - r)\right) F \\ &+ F\int_{-\infty}^{\infty} \left[R(\pi^*, x)^{1-\gamma} - 1\right]\Pi(v_t, dx) + F_t = 0. \end{aligned} \tag{24}$$

with $F(v, T) = 1$.

3.1 Optimal portfolios

Even though the focus of this paper is on the skewness and kurtosis effect on investing in a single stock, it is useful to note that the solution in the previous section, and the intuition of Theorem 1, can be extended to the multistock selection problem where there are N stocks and the i^{th} stock, $i = 1...N$, satisfies

$$d(\log(S_t^i)) = c_t^i dt + \int ... \int_{-\infty}^{\infty} x^i N(dt, dx) \tag{25}$$

where $N(dt, dx)$ is the jump counting Poisson measure, and the N-dimensional integral is taken over the entire jump support space. So x^i is the log-jump to the i^{th} stock, when x is the entire jump vector[5]. Observe that jumps may be correlated, and induce an instantaneous covariation between the i^{th} and j^{th} stocks

$$c_t^{i,j} = \int ... \int_{-\infty}^{+\infty} x^i x^j \Pi(v_t, dx)$$

Such jump induced correlation may be either due to the systemic risk (macro effects), as in the context of Das and Uppal (2004), when the jumps exhibit finite activity (low frequency), or due to the micro-structure risk when the jumps exhibit high activity.

In the multi-stock pure-jump case the budget constraint becomes,

$$dW_t = W_t \left(r_t + \sum_i \pi_{i,t}(c_t^i - r_t) \right) dt + W_{t-} \int ... \int_{-\infty}^{\infty} \sum_i \pi_{i,t}(e^{x^i} - 1) N(dt, dx) \tag{26}$$

The dynamic principle approach to stochastic optimal control leads to the following Hamilton-Jacobi-Bellman (HJB) equation for the indirect utility function J,

$$\begin{aligned} \max_{\pi} \frac{1}{2}\sigma_v^2 J_{vv} + m_v J_v + \left(r + \sum_i \pi_i \left(c^i - r\right) \right) W J_W \\ + \int ... \int_{-\infty}^{\infty} \left[J(W(1 + \sum_i \pi_i(e^{x^i} - 1)), v, t) - J(W, v, t) \right] \Pi(v_t, dx) + J_t = 0 \end{aligned} \tag{27}$$

and, similarly to the single stock case, we have the following characterization of the optimal portfolio allocation:

Theorem 2 *Assume that (14) holds for all stocks $i = 1...N$, and that there is a solution J to (27). The optimal portfolio allocation $\pi^*(v, t)$ solves*

$$\mu_t^i - r_t + \int ... \int_{-\infty}^{\infty} \left(R(\pi^*, x)^{-\gamma} - 1 \right) (e^{x^i} - 1) \Pi(v_t, dx) = 0 \tag{28}$$

where

$$R(\pi, x) = 1 + \sum_i \pi_i (e^{x^i} - 1) \tag{29}$$

[5]For simplicity of notation, we restrict the discussion here to the pure jump case.

is the portfolio return due to a log-jump vector x, and J is the indirect utility function of the form (17) with $F(v,t)$ the solution to the Partial Differential Equation

$$\frac{1}{2}\sigma_v^2 F_{vv} + m_v F_v + (1-\gamma)\left(r + \sum_i \pi_i^* \left(c^i - r\right)\right) F \\ + F\int \dots \int_{-\infty}^{\infty} \left[R(\pi^*, x)^{1-\gamma} - 1\right] \Pi(v_t, dx) + F_t = 0. \tag{30}$$

with $F(v,T) = 1$.

4 A Particular case: Jump rate proportional to the state variable

Some of the generality in (9) has to be sacrificed if we are to have (17) satisfied, and find specific solutions. We have to make an assumption as to the dependence of jump arrival rates on the state. Here we are guided by the potential use of such models to represent various micro-structure variables like trading activity, volume, etc. We thus assume that high states will be characterized by a higher rate of jumps. A way to attain such a dependence of jumps on the state, for which we are able to analytically solve for the optimal policy, is to assume that the arrival rate for jumps of any size x is proportional to the state,

$$\Pi(v_t, dx) = v_t \Pi(dx) \tag{31}$$

For example, in an asymmetric information context where each price jump is due to a new order execution,[6] so that the price reflects new information released to the market by the order, the state v_t captures the intensity of trading (or trading volume rate), since the above condition implies that an increase in the trading activity translates into a proportional increase in the instantaneous probability for jumps (i.e. order arrivals) of any size. Alternatively, v_t in (31) can be interpreted as the "rate of business activity" that implies a change in time from calendar time t to total activity time $\tau_t = \int_0^t v_s ds$ (Carr and Wu, 2004).

Our notation will be simplified by using the conditional *cumulant* kernel K_t defined as

$$K_t(s) = \int (e^{sx} - 1)\Pi(v_t, dx) \tag{32}$$

and, analogously, the unconditional

$$K(s) = \int (e^{sx} - 1)\Pi(dx) \tag{33}$$

For example, the return drift (7) can also be written as

$$\mu_t = c_t + K_t(1) \tag{34}$$

Under our assumption (31), we have

$$K_t(s) = v_t K(s) \tag{35}$$

[6]As in the benchmark microstructure models by Kyle (1985), and Glosten and Milgrom (1985).

From (22), the instantaneous variance of percentage returns is given by,

$$\sigma_t^2 = \int (e^x - 1)^2 \, \Pi(v_t, dx) = (K(2) - 2K(1))v_t \tag{36}$$

4.1 Optimal allocation with constant Sharpe ratio

In the model of Merton (1971) with an investor with CRRA utility with risk aversion γ, the investor faces diffusive stock dynamics with constant return rate $\mu = c + \sigma^2/2$ and constant volatility σ, and will invest in the stock an optimal proportion π_M,

$$\pi_M = \frac{\mu - r}{\sigma^2 \gamma} \tag{37}$$

Thus in the Merton's case the investor invests a fraction equal to the ratio η of the stock's Sharpe ratio divided by σ and her risk aversion

$$\pi_M = \frac{\eta}{\gamma} \tag{38}$$

The jump process considered in this paper displays a stochastic instantaneous variance rate σ_t^2. In order to make our model comparable to Merton (1971), we assume that the risk premium $\mu_t - r_t$ adjusts to reflect the changing riskiness of the stock. In particular, we assume that the stock parameters are such that

$$\mu_t - r_t = \eta \sigma_t^2 \tag{39}$$

where η is a constant.

We want to study the effect of higher moments on the optimal strategy determined by the equation (23) and compare our results against the benchmark (37).

The higher moments have to be properly accounted for; a frequent oversight in the jump-prices literature is to use the instantaneous variance of the log-price as the measure of the risk of the stock. This confusion mainly stems from our extensive experience with diffusion processes, where this is appropriate. For a general jump process though, the instantaneous variance of percentage returns is not the same as the one of log returns. To see this in our model, observe that, from (4), (5) and (33), when there is no diffusion ($\sigma_{d,t} = 0$), the instantaneous variance of log returns equals

$$\int x^2 \Pi(v_t, dx) = K''(0) v_t \tag{40}$$

On the other hand, the variance (36) of percentage returns differs from the variance of log-returns (40), unlike the diffusion case.[7]

Under our assumptions, equations (36) and (39) imply

$$\mu_t - r_t = \eta (K(2) - 2K(1)) v_t \tag{41}$$

Furthermore, when the state variable represents trading activity, as in (31), equation (23) leads to the following optimality condition for π^* that is independent of v:

[7] In that case, the kernel $K()$ is quadratic.

Proposition 1 *Under the assumptions of Corollary 1, (31) and (39), the optimal portfolio* π^* *is state invariant, and satisfies*

$$\eta(K(2) - 2K(1)) + M(\pi^*) = 0 \tag{42}$$

where

$$M(\pi) = \int_{-\infty}^{\infty} \left(R(\pi, x)^{-\gamma} - 1\right)(e^x - 1)\Pi(dx) \tag{43}$$

Proof: Straightforward, from (23). □

We see that the investor, when properly compensated by a stochastic risk premium proportional to the rate of variance, does not "time" her strategy with information about the rate of trading activity.

4.2 A model for the state variable

We derived the optimal portfolio strategy (23) by conjecturing a *separable* functional form for the indirect utility

$$J(W, v, t) = \frac{1}{1-\gamma} W^{1-\gamma} F(v, t) = U(W)F(v, t) \tag{44}$$

where $F(v, t)$ is a deterministic discount factor that captures time and state effects. Examples in similar spirit, but without the state variable, can be found in Øksendal and Sulem (2004) and references therein. The explicit functional form of the discount factor F depends on the state dynamics, and we will solve a specific case here. We assume that the state variable v_t follows a square root diffusion

$$dv_t = k(v_o - v_t)dt + \sigma_v v_t^{1/2} dZ_t^v \tag{45}$$

As we show next, when the state follows (45), the utility discount factor attains a log-linear form

$$F(v, t) = e^{A(t)+B(t)v} \tag{46}$$

Theorem 3 *If, in addition to previous assumptions, the state variable follows (45), the indirect utility function is given by*

$$J(W, v, t) := \max_{\{\pi_s, t \le s \le T\}} E_t U(W_T) = \frac{1}{1-\gamma} W^{1-\gamma} e^{A(t)+B(t)v} \tag{47}$$

where A(t) and B(t) are solutions to these Ordinary Differential Equations:

$$r(1-\gamma) + kv_o B + A' = 0 \tag{48}$$

$$\frac{\sigma_v^2}{2} B^2 + \pi^* \Big(\eta\big(K(2) - 2K(1)\big) - K(1)\Big)(1-\gamma) - kB + M_2(\pi^*) + B' = 0, \tag{49}$$

where

$$M_2(\pi) = \int_{-\infty}^{\infty} \left(R(\pi, x)^{1-\gamma} - 1\right)\Pi(dx) \tag{50}$$

is the average jump in utility for the π *policy.*

Proof: In Appendix. □

4.3 The case of constant v

In the case of a constant v, the above ODE's can be solved exactly, and we can get an explicit solution for function J:

Proposition 2 *If, in addition to previous assumptions, v is constant, the optimal expected utility is given by*

$$E_t U(W_T) = U(W_t) e^{a(T-t)+vM_2(\pi)(T-t)} \tag{51}$$

with $a = (1-\gamma)[r + \pi(c-r)]$.

Proof: In Appendix. □

As mentioned below, in Table 1 we present a summary of results.

5 Conditional Variance Gamma model

In order to compare this model to real market dynamics, we need further tractability. The goal here is to study a specific solution calibrated to real returns. Empirically, small jumps are difficult (if not impossible) to discern, but we can work with higher moments as the jumps are responsible for skewness and kurtosis. While with slow Poisson arrivals a diffusion is needed to generate the extreme local activity observed in real securities, when jump arrival rates are infinite, for any time period, no matter how small, there will always be jumps and thus there is no need for a diffusion component anymore. Thus, to keep the following calibration relatively simple,[8] we introduce the *conditional* variance gamma (VG) process that does not include a diffusive component.

The *unconditional* VG process was introduced in Madan and Seneta (1990) and Madan and Milne (1991), and generalized by Madan, Carr and Chang (1998). The VG process is a broadly used, canonical example of a pure jump Lévy process. The infinite, two-sided, pure jump activity of the VG process, can be decomposed into an increasing gamma process that only contains positive jumps, and one only containing negative jumps. In this decomposition, one may think, for example, of the positive component representing the buy orders, while the negative component captures sell orders.

The VG process is a pure jump process, with an infinite arrival rate of small jumps. The small size and infinite arrival rate of jumps generates extreme local activity reminiscent of a diffusion but with right continuous paths of finite variation. Unlike a diffusion that can be approximated by a binomial tree, the infinitesimal change in a VG process can take infinitely many values and is thus fundamentally un-hedgeable, in the sense that trading a finite set of assets does not complete the markets.

Formally, for parameters $\rho > 0$, and θ, the homogeneous Variance Gamma (VG) process is defined as a time-changed Brownian motion. Thus, the resulting stock dynamics are not diffusive, but the result of monitoring the continuous-path Gaussian process W_t^1 at random

[8] A diffusive component could be added easily, as in the earlier section. In reality, it may be difficult to econometrically disentangle and identify the pure jump from the diffusive part. On this issue, see Ait-Sahalia (2004).

times[9] given by a gamma process. That is, instead of the usual return $X_t = \theta t + \rho W^1(t)$, here

$$X_t = \theta \tau_t + \rho W^1(\tau_t) \tag{52}$$

where, for fixed $l > 0$, and $v > 0$, a gamma process, $\tau_t = \gamma_t(l, v)$, with mean rate lv and variance rate $l^2 v$, is used to measure the transformation from real time t to the stopping time τ_t. The gamma process is defined by the density of the increment x over a time interval h, $x = \gamma_{t+h} - \gamma_t$, given by the gamma density function

$$f_h(x) = \frac{e^{-x/l} x^{vh-1}}{l^{vh} \Gamma(vh)} \tag{53}$$

Interestingly, the time changed diffusion X_t is a pure jump process with no diffusive risk. Specifically, it is well known that its Lévy-Khintchine representation is determined by

$$K(s) = v^{-1} t^{-1} \log E e^{s X_t} = -\log(1 - \theta l s - .5 \rho^2 l s^2) \tag{54}$$

and it does not contain a quadratic term, and thus the process has no diffusion component.

It can be shown that the VG process is uniquely decomposed into two gamma processes, one with positive jumps, and the other containing the negative jumps[10]

$$X_t = \gamma_t^u(\lambda_u, v) - \gamma_t^d(\lambda_d, v) \tag{55}$$

5.1 The conditional Lévy process

The process in (52) does not have a stochastic jump structure. It is known that for a constant v the jump measure for a VG process is,

$$\begin{aligned} \Pi(v, dx) &= \frac{v}{x} e^{-x/\lambda_u}, \text{ for } x > 0 \\ &= \frac{v}{|x|} e^{-|x|/\lambda_d}, \text{ for } x < 0 \end{aligned}$$

To arrive at a stochastically varying jump structure we define our underlying process as one which has the above jump measure, but we allow the v parameter to become stochastic. In this case, the jump measure is of the type (31) with

$$\begin{aligned} \Pi(dx) &= \frac{1}{x} e^{-x/\lambda_u}, \text{ for } x > 0 \\ &= \frac{1}{|x|} e^{-|x|/\lambda_d}, \text{ for } x < 0 \end{aligned}$$

[9] Intuitively, these random times can be thought as the arrival times of new market orders.

[10] λ_u and λ_d are the positive solutions to the system $\lambda_u - \lambda_d = \theta l$ and $\lambda_u \lambda_d = .5\rho^2 l$. That is $\lambda_u = \frac{1}{2}\left(\sqrt{\theta^2 l^2 + 2\rho^2 l} + \theta l\right)$ and $\lambda_d = \frac{1}{2}\left(\sqrt{\theta^2 l^2 + 2\rho^2 l} - \theta l\right)$.

5.2 Return moments

In order to compare the investment strategy with conditionally Lévy jumps to the strategy of an investor faced with a diffusion we have to keep in mind that jumps will generally introduce both skewness and excess kurtosis. To enhance intuition we want to initially only compare symmetric but fat-tailed conditionally Lévy returns to Brownian motion. This means that skewness has to be removed. In the literature, it has been wrongly suggested that introducing a symmetric VG process as the log-price process results in zero return skewness. The reason for why this is wrong is similar to the reason we presented in the previous discussion on the instantaneous variance. Diffusive log-prices lead to diffusive returns, only with a different drift, but when symmetric log-price jumps are introduced, the potential for large jumps introduces skewness in infinitesimal returns. This skewness is identified as follows.

From (4) and (22), the conditional third centralized moment of the infinitesimal return is given by

$$\int_{-\infty}^{\infty} (e^x - 1)^3 v_t \Pi(dx) = (K(3) - 3K(2) + 3K(1))\, v_t$$

Thus, from (54), in order to get symmetric returns, the process has to satisfy

$$(1 - 2\theta l - 2\rho^2 l)^3 = (1 - 3\theta l - \frac{9}{2}\rho^2 l)(1 - \theta l - \frac{1}{2}\rho^2 l)^3 \tag{56}$$

Observe that when the log-price returns are symmetric, $\theta = 0$, the percentage returns are not. Based on (56), it can actually be shown that, since net returns are always larger than log-returns, a symmetric log returns VG process always leads to positive percentage returns skewness, and thus makes the stock attractive.

In practice, we could calibrate the stock process by using higher moments, and the observable rate of trading activity, v. As observed earlier, the instantaneous return variance is given by

$$\mathrm{VAR} = (K(2) - 2K(1))v \tag{57}$$

Similarly, instantaneous skewness is given by

$$\mathrm{SKEW} = \frac{K(3) - 3K(2) + 3K(1)}{(K(2) - 2K(1))^{3/2}} v^{-1/2} \tag{58}$$

and instantaneous excess kurtosis by

$$\mathrm{KURT} - 3 = \frac{K(4) - 4K(3) + 6K(2) - 4K(1)}{(K(2) - 2K(1))^2} v^{-1} \tag{59}$$

6 Numerical results

In Tables 1-4 we compute some numerical examples. Our main objective is to study the effect of higher moments in the optimal portfolio allocation of the CRRA investor considered above.

As we explained before, the expected return of the portfolio is adjusted so that $\eta = (\mu_t - r_t)/\sigma_t^2$ is constant. That coefficient, divided by the coefficient of risk aversion γ, explains the proportion of wealth invested in the risky security in the Merton (1971) setting for a stock price that follows a diffusion process. We want then to study the effects of skewness and kurtosis on optimal allocation, with respect to the benchmark Merton (1971) model, and that justifies the previous constraint. We allow for the variance to vary, but with constant η (so that the drift is adjusted accordingly). The variance σ_t^2 is computed as in equation (57). Skewness and excess kurtosis are computed as in equations (58) and (59), respectively. For all computations we take the time horizon $T = 10$.

In Table 1 we consider the effect on the optimal portfolio of the stock price returns moments reported in Campbell, Lo and MacKinlay (1996, page 21). We focus on the case of the value-weighted index. The first row of table 1 is in line with the moments reported by Campbell, Lo and MacKinlay (1996) for daily returns. As we see, the effect of average higher moments for the period considered is moderate, but significant. Average volatility over the time period considered is roughly equivalent to a 12% annual. There are subperiods (like a good part of the 70's) in which volatility was significantly higher. These are also the periods on which accurate portfolio allocation is, arguably, more relevant. In the second line we compute optimal portfolio holdings for a similar level of moments, but with about double the standard deviation which, although not representative of the whole period, is a level of volatility not unusual in financial markets. The impact on optimal portfolio allocation (our benchmark stays constant) is substantial. The third row of that table considers moments computed for monthly returns. Our model considers "instantaneous" returns, therefore statistics of monthly returns do not seem the best choice. However, results are in line with those for daily returns.

In Table 2 we focus on the case in which the skewness is zero, so that we can study the specific effect of excess kurtosis on optimal portfolio allocation. In the Merton (1971) model, for a value of $\eta = 2.55$, a CRRA investor whose opportunity set consisted of a stock satisfying a diffusion process and a risk-free security would hold 85 % and 51 % of wealth in the risky security for degrees of risk aversion of $\gamma = 3$ and $\gamma = 5$, respectively. As observed before, in a setting of lower variance, the effect of kurtosis on optimal allocation is almost negligible. It is modest, but non-trivial, in a setting of higher volatility. We perform the same exercise for a higher value of η. The conclusion is similar, but the impact of kurtosis on optimal allocation is relatively higher than for a lower value of η.

In Table 3 we present several examples of cases in which skewness is strictly negative. As expected, the impact of negative skewness is higher than that of kurtosis (for likely values of both moments). Also as before, higher volatility results in a higher impact of the negative kurtosis.

Table 4 is similar to 5, but skewness is positive and also in line with the values reported in Campbell, Lo and MacKinley (1996).[11] It seems it would take a relatively high level of positive skewness to offset the effect of kurtosis. This level will have to be higher for higher levels of variance.

Overall, we find that higher moments have significant, but not huge, effect on the optimal portfolio allocation. This effect becomes important for high volatility. Our findings are in line with those of Das and Uppal (2004) and Guidolin and Nicodano (2005). However, these papers also compute optimal portfolio allocation for multiple risky assets, with cross-higher moments. They find that the effect of higher moments on portfolio allocation is, in general,

[11] Positive skewness has been rarely documented in financial time series data. The objective of this table is to provide an additional insight on the effect of higher moments.

substantial.

We also study the effect of ignoring higher moments in terms of utility loss. The standard approach is to compute the certainty equivalent or the similar "wealth loss" (as in Liu, Longstaff and Pan 2003). The idea is to compute the percentage of initial wealth the suboptimal allocation (resulting from ignoring higher moments) would amount to losing. That is, the value ϵ that would make the utility of an investor that allocates \$1 suboptimally (ignoring higher moments) equal to an investor that allocates optimally \$(1-ϵ). More explicitly, from equation (51) we can find the expected utility for a given allocation π. We compute the optimal π for the Merton (1971) case (ignoring higher moments). Equation (51) gives us the expected utility for that allocation when higher moments are taken into consideration. We find what is the ϵ such that the utility for \$(1-ϵ) investment with optimal allocation (taking higher moments into consideration) yields the same utility (from equation 51) as \$1 investment according to the optimal allocation following the Merton (1971) rule.

In Table 1 we find the welfare loss for the moments reported in Campbell, Lo and MacKinley (1996). The numbers are relatively low, especially for a context of low volatility (as it was most of the time period covered by the sample used by Campbell, Lo and MacKinley 1996). This is consistent also with Cremers, Kritzman and Page (2004). In Figure 1 we extend that analysis to study the effect of higher volatility. We observe that for context of high (but not impossible) volatility, the wealth loss resulting from ignoring higher moments is significant. Additionally, in Figure 2 we compare the wealth loss resulting from overinvesting in the context of our model with higher moments with a similar overinvesting in a Gaussian setting like in Merton (1971). We observe that the wealth loss in a setting with higher moments is about 30% higher than in the benchmark Merton (1971) setting, although, as pointed by Brennan and Torous (1999), the wealth loss in that setting is not very large.

7 Conclusions

In this paper we study the effect of higher moments on the optimal investment strategy of a risk-averse investor. We analyze our problem for a large class of Lévy processes. For tractability purposes, we consider a particular type of process, the pure-jump Variance Gamma process, which has been widely used in the option pricing literature. We compare optimal asset allocation to that of an investor in a Merton (1971) setting. We find that higher moments affect the optimal allocation of a risk averse investor, although the importance of the deviations will depend strongly on the level of volatility. We characterize the optimal allocation in the presence of multiple risky assets, possibly correlated, but we do not obtain numerical results, since a computational algorithm does not appear obvious. We leave the solution of this numerical problem for future work.

8 Appendix

Proof of Theorem 1: Assuming that J is of the form (17), we take a derivative in the HJB equation (16) with respect to π, and we get (18) as the first-order condition. Substituting back π^* in the HJB equation, we see that the equation is satisfied with J, if F solves (20). The initial condition $F(v, T) = 1$ is self explanatory since at time T there is no more opportunity to invest.

Proof of Theorem 3: When the state follows (45), the Hamilton-Jacobi-Bellman equation (24) becomes

$$\max_{\pi} \frac{1}{2}\sigma_v^2 v J_{vv} + k(v_0 - v)J_v + (r + \pi\,(c - r))\,WJ_W \\ + \int_{-\infty}^{\infty} \big[J(W(1 + \pi(e^x - 1)), v, t) - J(W, v, t)\big]\Pi(v_t, dx) + J_t = 0 \tag{60}$$

We conjecture that the function J satisfies equation (17) in the form

$$J(W, v, t) = U(W)e^{A(t)+B(t)v}$$

where $A(t)$ and $B(t)$ are deterministic functions of time. If that is the case, the optimal investment strategy of this investor is given by the value π^* that solves the equation (42).

We now show that the conjecture is true, by deriving the ordinary differential equations for the time dependent coefficients A and B. It is easy to check that, if the conjecture is true, we have $WJ_W = (1-\gamma)J$, $J_v = JB$, $J_t = J(A' + B'v)$, $W^2 J_{WW} = -\gamma(1-\gamma)J$, $J_{vv} = JB^2$ and $WJ_{Wv} = B(1-\gamma)J$. Substituting the optimality conditions (42) for π back into the HJB equation, we recover an affine relation for v

$$\frac{\sigma_v^2 v}{2}B^2 + \Big(r + \pi\big[\eta(K(2) - 2K(1)) - K(1)\big]v\Big)(1-\gamma) + \\ k(v_o - v)B + vM_2(\pi) + (A' + B'v) = 0$$

where

$$M_2(\pi) = \int_{-\infty}^{\infty} \big(R(\pi, x)^{1-\gamma} - 1\big)\,\Pi(dx) \tag{61}$$

is the average jump in utility for the π policy. For this condition to be satisfied for all v, the constant term and the linear coefficient have to be equal to zero separately, which provides the two ODEs from the statement of Theorem 3 that the A and B functions have to satisfy. If those ODEs are satisfied, then the function J of the conjectured form does, indeed, satisfy the HJB equation.

Proof of Proposition 2: From (13), we see that the wealth at time T when the investor starts with W_t is

$$W_T = W_t e^{r(T-t)+\pi(c-r)(T-t)+\int_t^T \int_{-\infty}^{+\infty} y(x)N(dt,dx)}$$

with $y(x) = \ln R(\pi, x)$ being the wealth exposure of the investor to a jump of size x. The investor's utility becomes

$$E_t U(W_T) = U(W_t)E_t e^{a(T-t)+(1-\gamma)\int_t^T \int_{-\infty}^{+\infty} y(x)N(dt,dx)}$$

where

$$a = (1-\gamma)[r + \pi(c-r)]$$

Denote

$$P_t = e^{(1-\gamma)\int_0^t \int_{-\infty}^{+\infty} y(x)N(dt,dx)}$$

or equivalently

$$P_t = e^{(1-\gamma)Y_t^\pi}$$

with the Lévy process Y^π being defined through

$$Y_t^\pi = \int_0^t \int_{-\infty}^{+\infty} y(x)N(dt,dx)$$

Taking expectations, we get

$$E_t e^{(1-\gamma)\int_t^T \int_{-\infty}^{+\infty} y(x)N(dt,dx)} = \frac{1}{P_t} E_t P_T = e^{(T-t)K_Y(1-\gamma)}$$

with the kernel defined as

$$K_Y(s) = v \int_{-\infty}^{\infty} (R^s - 1)\Pi(dx) \tag{62}$$

Thus

$$\frac{1}{P_t} E_t P_T = e^{(T-t)v\int_{-\infty}^{+\infty}(R^{1-\gamma}-1)\Pi(dx)} \tag{63}$$

This implies the statement of the proposition.

References

1. Adler T., Kritzman, M.: Mean-Variance analysis versus full-scale optimization - out of sample, Revere Street working paper (2005)

2. Aït-Sahalia, Y.: Disentangling diffusion from jumps, Journal of Financial Economics **74**, 487-528 (2004)

3. Aït-Sahalia, Y., Cacho-Diaz, L., Hurd, T.: Portfolio choice with a large number of assets: Jumps and diversification, working paper, Princeton University (2006)

4. Ané, T., Geman,H.: Order flow, transaction clock and normality of asset returns, Journal of Finance **55**, 2259-2284 (2000)

5. Bansal, R., Hsieh, D., Viswanathan, S.: A new approach to international arbitrage pricing, Journal of Finance **48**, 1719-1747 (1993)

6. Bansal, R., Viswanathan, S.: No-arbitrage and arbitrage pricing, Journal of Finance **48**, 1231-1262 (1993)

7. Brennan, M., Torous, W: Individual decision making and investor welfare, Economic Notes **28**, 119-143 (1999)

8. Campbell, J., Lo, A., MacKinlay, C.: The Econometrics of financial markets. Princeton: Princeton University Press 1996

9. Carr, P., Geman, H., Madan, P., Yor, M.: The fine structure of asset returns: An empirical investigation, Journal of Business **75**, 305-332 (2000)

10. Carr, P., Geman, H., Madan, P., Yor, M.: Stochastic Volatility for Lévy Processes , Mathematical Finance **13**, 345-382 (2003)

11. Carr, P., Jin, X., Madan, D.: Optimal investment in derivative securities, Finance and Stochastics **5**, 33-59 (2001)

12. Carr, P., Wu, L.: The finite moment log stable process and option pricing, *Journal of Finance* 58, 753-778 (2003)

13. Carr, P., Wu, L.: Time-changed Lévy processes and option pricing, Journal of Financial Economics **71**, 113-141 (2004)

14. Chapman, D.: Approximating the asset pricing kernel, Journal of Finance **52**, 1383-1410 (1997)

15. Choulli, T., Hurd, T.: The portfolio selection problem via Hellinger processes, working paper, University of Alberta (2001)

16. Clark, P.: A subordinated stochastic process model with finite variance for speculative prices, Econometrica **41**, 135-156 (1973)

17. Cremers, J., Kritzman, M., Page, S.: Optimal hedge fund allocations: Do higher moments matter?, Revere Street Working Paper Series (2004)

18. Das, S., Uppal, R.: Systemic risk and international portfolio choice, Journal of Finance **59**, 2809-2834 (2004)

19. Eraker, B., Johannes, M., Polson, N.: The impact of jumps in returns and volatility, Journal of Finance **58**, 1269-1300 (2003)

20. Geman, H., Madan, D., Yor, M.: Time changes for Lévy processes, Mathematical Finance **11**, 79-96 (2001)

21. Geman, H., Madan, D., Yor, M.: Stochastic volatility, jumps and hidden time changes, Finance and Stochastics **6**, 63-90 (2002)

22. Glosten, L., Milgrom, P.: Bid, ask and transaction prices in a specialist market with heterogeneously informed traders, Journal of Financial Economics **14**, 71-100 (1986)

23. Guidolin, N., Nicodano, G.: Small caps in international equity portfolios: The effects of variance risk, working paper, Federal Reserve Bank of Saint Louis (2005)

24. Guidolin, M., Timmermann, A.: International asset allocation under regime switching, skew and kurtosis preferences, working paper, University of California, San Diego (2006)

25. Harvey, C., Liechty, J., Liechty, Müller, P.: Portfolio selection with higher moments, working paper, Duke University (2004)

26. Harvey, C., Siddique, A.: Conditional skewness in asset pricing tests, Journal of Finance **55**, 1263-1295 (2000)

27. Jacod, J., Shiryaev, A.: Limit Theorems for Stochastic Processes. New York, Heidelberg, Berlin: Springer-Verlag 1987

28. Jondeau, E., Rockinger, M.: Non-normality: How costly is the mean-variance criterion?, working paper, HEC Lausanne (2005)

29. Kraus, A., Litzenberger, R.: Skewness preference and the valuation of risk assets, Journal of Finance **31**, 1085-1100 (1976)

30. Kyle, A.: Continuous auctions and insider trading, Econometrica **53**, 1315-1336 (1985)

31. Kraus, A., Litzenberger, R.: On the distributional conditions for a consumption oriented three moment CAPM, Journal of Finance **38**, 1381-1391 (1983)

32. Liu, J., Longstaff, F., Pan, J.: Dynamic Asset Allocation with Event Risk, Journal of Finance **58**, 231-259 (2003)

33. Madan, D., Carr, P., Chang, E.: The variance gamma process and option pricing, European Finance Review **2**, 79-105 (1998)

34. Madan, D., Milne, F.: 1991, Option pricing with VG martingale components, Mathematical Finance **1**, 39-56 (1991)

35. Madan, D., Seneta, E.: 1987, The variance gamma (VG) model for share market returns, Journal of Business **63**, 511-524 (1987)

36. Merton, R., 1971, Optimum consumption and portfolio rules in a continuous-time model, Journal of Economic Theory **3**, 373-413 (1971)

37. Øksendal, B., Sulem, A.: Applied Stochastic Control of Jump Diffusions. Berlin: Springer 2004

38. Rubinstein, M.: The fundamental theorem of parameter-preference security valuation, Journal of Financial and Quantitative Analysis **8**, 61-69 (1973)

Table 1
Optimal Allocation in a Risky Stock with Campbell, Lo and MacKinlay (1996) Moments

We compute optimal allocation in a risky security that follows the VG process described in the paper and restricted so that parameter $\eta = (\mu - r)/\sigma^2$ is constant. The parameters θ, l and ρ are as in the paper. The column "Var" denotes the variance of the return of the risky security, "Skew" its skewness, and "Kurt-3" its excess kurtosis. Parameters are calibrated so as to get the index moments (that would correpond to the risky security of our model) for daily and monthly returns in Campbell, Lo and MacKinlay (1996). $\hat{\pi}$ represents the proportion of wealth optimally allocated to the risky stock (the balance is allocated to the riskfree security). For all cases, the time horizon is $T = 10$. Optimal allocation to the risky asset security in the benchmark Merton (1971) model when $\eta = 4.5$ and $\gamma = 5$ (as in this table) is 0.9. In parenthesis, underneath the optimal allocation, we record the wealth loss resulting from ignoring higher moments and, therefore, investing 90% in the risky security.

$\eta = 4.5,\ \gamma = 5$

v	θ	l	ρ	Var	Skew	Kurt-3	$\hat{\pi}$
0.1	-0.0095	0.5	0.0360	6.6194 E-05	-1.3672	30.4042	0.8652
							(0.00565%)
0.1	-0.0230	0.5	0.0725	2.6807 E-04	-1.3430	29.6104	0.8204
							(1.3821%)
1.3	-0.0131	0.5	0.0534	1.8772 E-03	-0.2999	2.2941	0.8515
							(0.1518%)
1.3	-0.0269	0.5	0.0895	5.2666 E-03	-0.2990	2.2511	0.8072
							(1.7976%)

Table 2
Optimal Allocation in a Risky Stock with Zero Skewness and Non-zero Kurtosis

We compute optimal allocation in a risky security that follows the VG process described in the paper and restricted so that parameter $\eta = (\mu - r)/\sigma^2$ is constant. Additionally, parameter values are calibrated so that the skewness of stock price returns is zero. The parameters v, θ, l and ρ are as in the paper. The column "Var" denotes the variance of the return of the risky security, and "Kurt-3" its excess kurtosis. $\hat{\pi}$ represents the proportion of wealth optimally allocated to the risky security (the balance is allocated to the riskfree security). We compute optimal portfolio allocation for two different values of η and two degrees of risk aversion, that we denote γ. For all cases, the time horizon is $T = 10$. Optimal allocation to the risky security in the benchmark Merton (1971) model when $\eta = 2.55$ (as in the top panel in this table) is 0.85 for $\gamma = 3$ and 0.51 for $\gamma = 5$. Optimal allocation to the risky security in the benchmark Merton (1971) model when $\eta = 4.2$ (as in the bottom panel in this table) is 0.84 for $\gamma = 5$ and 0.6 for $\gamma = 7$.

$\eta = 2.55$

						$\hat{\pi}$	
v	θ	l	ρ	Var	Kurt-3	$\gamma = 3$	$\gamma = 5$
0.1	-0.00194	0.5	0.036	6.4779 E-05	30.0324	0.8461	0.5082
1/15	-0.0072907	0.2	0.06972	6.4774 E-05	45.0730	0.8441	0.5071
0.1	-0.0194598	0.2	0.1139	2.5913 E-04	30.1302	0.8348	0.5031
1/15	-0.0291862	0.2	0.1395	2.5893 E-04	45.2934	0.8277	0.4998

$\eta = 4.2$

						$\hat{\pi}$	
v	θ	l	ρ	Var	Kurt-3	$\gamma = 5$	$\gamma = 7$
0.1	-0.00194	0.5	0.036	6.4779 E-05	30.0324	0.8321	0.5951
1/15	-0.0072907	0.2	0.06972	6.4774 E-05	45.0730	0.8283	0.5927
0.1	-0.0194598	0.2	0.1139	2.5913 E-04	30.1302	0.8103	0.5814
1/15	-0.0291862	0.2	0.1395	2.5893 E-04	45.2934	0.7973	0.5731

Table 3
Optimal Allocation in a Risky Stock with Negative Skewness and Non-zero Kurtosis

We compute optimal allocation in a risky security that follows the VG process described in the paper and restricted so that parameter $\eta = (\mu - r)/\sigma^2$ is constant. Additionally, parameter values are calibrated so that the skewness of return of the risky security is negative. The parameters v, θ, l and ρ are as in the paper. The column "Var" denotes the variance of the return of the risky security, "Skew" its skewness, and "Kurt-3" its excess kurtosis. $\hat{\pi}$ represents the proportion of wealth optimally allocated to the risky security (the balance is allocated to the riskfree security). We compute optimal portfolio allocation for two different values of η and two degrees of risk aversion, that we denote γ. For all cases, the time horizon is $T = 10$. Optimal allocation to the risky security in the benchmark Merton (1971) model when $\eta = 2.55$ (as in the top panel in this table) is 0.85 for $\gamma = 3$ and 0.51 for $\gamma = 5$. Optimal allocation to the risky security in the benchmark Merton (1971) model when $\eta = 4.2$ (as in the bottom panel in this table) is 0.84 for $\gamma = 5$ and 0.6 for $\gamma = 7$.

$\eta = 2.55$

							$\hat{\pi}$	
v	θ	l	ρ	Var	Skew	Kurt-3	$\gamma = 3$	$\gamma = 5$
0.1	-0.0095	0.5	0.036	6.6194 E-05	-1.3672	30.4042	0.8306	0.4998
0.1	-0.023	0.5	0.0725	2.6807 E-04	-1.3430	29.6104	0.8062	0.4874
1/15	-0.0105	0.5	0.0445	6.6152 E-05	-1.3722	45.0065	0.8288	0.4999
1/15	-0.02652	0.5	0.0892	2.6816 E-04	-1.2992	43.9062	0.8007	0.485
0.1	-0.00344	0.5	0.03518	6.1925 E-05	-0.3009	29.9021	0.8429	0.5065
0.1	-0.01136	0.5	0.07325	2.6819 E-04	-0.3013	29.7997	0.8277	0.4992
1/15	-0.00436	0.5	0.04308	6.1875 E-05	-0.2996	44.8467	0.8411	0.5057
1/15	-0.01541	0.5	0.0898	2.6844 E-04	-0.3012	44.7692	0.8206	0.4960

$\eta = 4.2$

							$\hat{\pi}$	
v	θ	l	ρ	Var	Skew	Kurt-3	$\gamma = 5$	$\gamma = 7$
0.1	-0.0095	0.5	0.036	6.6194 E-05	-1.3672	30.4042	0.8101	0.5801
0.1	-0.023	0.5	0.0725	2.6807 E-04	-1.3430	29.6104	0.7717	0.5548
1/15	-0.0105	0.5	0.0445	6.6152 E-05	-1.3722	45.0065	0.8065	0.5778
1/15	-0.02652	0.5	0.0892	2.6816 E-04	-1.2992	43.9062	0.7616	0.5485
0.1	-0.00344	0.5	0.03518	6.1925 E-05	-0.3009	29.9021	0.8277	0.5920
0.1	-0.01136	0.5	0.07325	2.6819 E-04	-0.3013	29.7997	0.8005	0.5747
1/15	-0.00436	0.5	0.04308	6.1875 E-05	-0.2996	44.8467	0.8241	0.5898
1/15	-0.01541	0.5	0.0898	2.6844 E-04	-0.3012	44.7692	0.7875	0.5665

Table 4
Optimal Allocation in a Risky Stock with Positive Skewness and Non-zero Kurtosis

We compute optimal allocation in a risky security that follows the VG process described in the paper and restricted so that parameter $\eta = (\mu - r)/\sigma^2$ is constant. Additionally, parameter values are calibrated so that the skewness of the return of the risky security is positive. The parameters v, θ, l and ρ are as in the paper. The column "Var" denotes the variance of the return of the risky security, "Skew" its skewness, and "Kurt-3" its excess kurtosis. $\hat{\pi}$ represents the proportion of wealth optimally allocated to the risky security (the balance is allocated to the riskfree security). We compute optimal portfolio allocation for two different values of η and two degrees of risk aversion, that we denote γ. For all cases, the time horizon is $T = 10$. Optimal allocation to the risky security in the benchmark Merton (1971) model when $\eta = 2.55$ (as in the top panel in this table) is 0.85 for $\gamma = 3$ and 0.51 for $\gamma = 5$. Optimal allocation to the risky security in the benchmark Merton (1971) model when $\eta = 4.2$ (as in the bottom panel in this table) is 0.84 for $\gamma = 5$ and 0.6 for $\gamma = 7$.

$\eta = 2.55$

							$\hat{\pi}$	
v	θ	l	ρ	Var	Skew	Kurt-3	$\gamma = 3$	$\gamma = 5$
0.1	-0.00029	0.5	0.0352	6.6194 E-05	0.2993	30.28	0.8496	0.5101
0.1	-0.004282	0.5	0.0704	2.4775 E-04	0.3014	30.5684	0.8419	0.5069
1/15	-0.001215	0.5	0.0431	6.1933 E-05	0.3000	45.4144	0.8477	0.5093
1/15	-0.00799	0.5	0.0861	2.4681 E-04	0.2996	45.9119	0.8348	0.5036
0.1	0.00338	0.5	0.0348	6.1211 E-05	1.0001	31.3288	0.8576	0.5144
0.1	0.00305	0.5	0.07	2.4742 E-04	1.0025	32.0803	0.8575	0.5154
1/15	0.00245	0.5	0.0427	6.1998 E-05	0.9991	46.6805	0.8557	0.5136
1/15	-0.0007	0.5	0.0858	2.4673 E-04	0.9989	47.87	0.8498	0.5118

$\eta = 4.2$

							$\hat{\pi}$	
v	θ	l	ρ	Var	Skew	Kurt-3	$\gamma = 5$	$\gamma = 7$
0.1	-0.00029	0.5	0.0352	6.6194 E-05	0.2993	30.28	0.8373	0.5986
0.1	-0.004282	0.5	0.0704	2.4775 E-04	0.3014	30.5684	0.8205	0.5881
1/15	-0.001215	0.5	0.0431	6.1933 E-05	0.3000	45.4144	0.8336	0.5963
1/15	-0.00799	0.5	0.0861	2.4681 E-04	0.2996	45.9119	0.8074	0.58
0.1	0.00338	0.5	0.0348	6.1211 E-05	1.0001	31.3288	0.849	0.6065
0.1	0.00305	0.5	0.07	2.4742 E-04	1.0025	32.0803	0.8425	0.6034
1/15	0.00245	0.5	0.0427	6.1998 E-05	0.9991	46.6805	0.8451	0.6041
1/15	-0.0007	0.5	0.0858	2.4673 E-04	0.9989	47.87	0.8279	0.5942

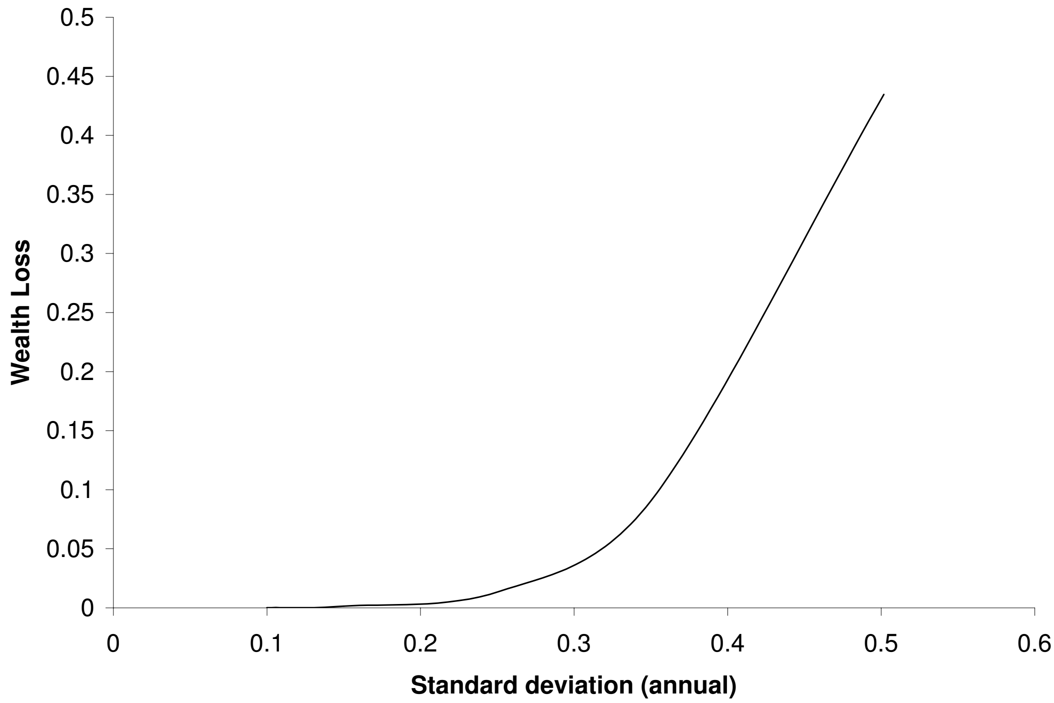

Figure 1: The plot shows the wealth loss resulting of ignoring higher moments to compute optimal allocation for different levels of volatility. Parameter values have been calibrated so as to approximate the level of skewness and excess kurtosis reported by Campbell, Lo and MacKinlay (1996) for monthly returns of a value-weighted index: skewness = -0.29; excess kurtosis = 2.42. We assume a market price of risk such that $\eta = 4.5$, a degree of risk aversion of $\gamma = 5$ and a horizon (in years) of $T = 10$.

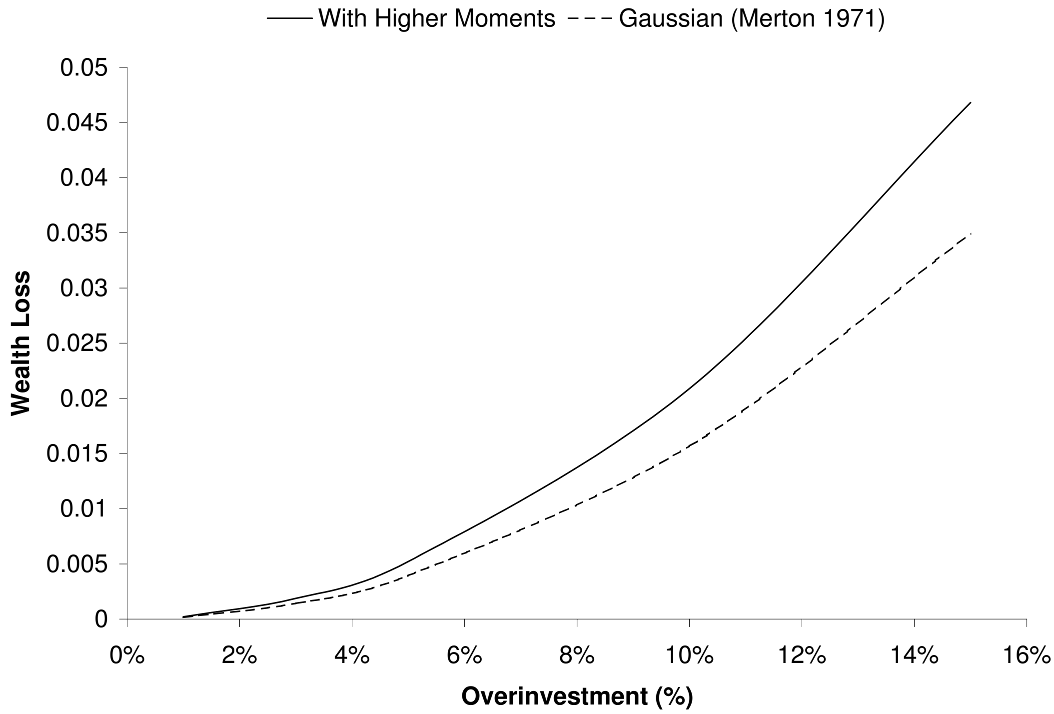

Figure 2: The plot shows the the difference in wealth loss between the model presented in this paper and the Merton (1971) optimal allocation model, resulting from overinvesting. In both cases, parameter values are such that the optimal allocation in the risky security is 80.72%. Overinvestment measures the additional proportion allocated to the risky security. For the model described in this paper, parameter values have been calibrated so as to approximate the level of skewness and excess kurtosis reported by Campbell, Lo and MacKinlay (1996) for monthly returns of a value-weighted index: skewness = -0.29; excess kurtosis = 2.42. Annual volatility is 25.14%. We assume a market price of risk such that $\eta = 4.5$, for the model in this paper and $\eta = 4.036$ for the Merton (1971) model (so that optimal allocation is identical). We assume for both models a degree of risk aversion of $\gamma = 5$ and a horizon (in years) of $T = 10$.