Myopic Portfolio Choice with Higher Cumulants

Eduardo Dávila[*]

Harvard University

`edavila@fas.harvard.edu`

January 2011

Abstract

This paper explores analytically the implications of higher order moments in the distribution of returns for the myopic optimal portfolio of an expected utility maximizer. The use of cumulant generating functions and entropy is crucial to find these solutions. With constant absolute risk aversion (CARA) utility, I find in closed form the optimal amount of risky asset for many distributions. My focus is in the problem with one risky asset but the results can be extended to multivariate returns. When possible, I set up a simple equilibrium model and solve for the equilibrium price, analyzing the effect of higher cumulants on the risk premium. My approach also provides new intuition about the effect of higher order events with constant relative risk aversion (CRRA). In that case, with a traditional budget constraint, the model stills need to be evaluated numerically, but I introduce a new formulation to obtain analytical results. The main practical conclusion of this paper is that tail events can matter for portfolio choice when departures from normality are large.

1 Introduction

This paper explores analytically how higher order moments in the distribution of returns affect the optimal portfolio choice of an expected utility maximizer. In order to do so, I make extensive use of cumulant generating functions (CGF) and entropy, importing the approach used by Martin (2008) and Backus, Chernov and Martin (2009) for consumption based asset pricing into classical portfolio choice.

Should the anticipation of tail events change the ex ante behavior of market participants? Following periods of crisis, this question often emerges as a subject of heated discussion. The approach developed here allows me to shed some light on this issue, as it provides a very clear interpretation of how jumps (rare disasters) or nongaussian tails can modify the optimal portfolio of assets.

Since Markowitz (1952), the portfolio selection problem has been a basic pillar of financial economics. Markowitz's assumption of mean variance preferences is usually rationalized with quadratic utility, normal distribution for returns or local approximations to utility, although the work by Chamberlain (1983) formally establishes that it is the class of elliptical distributions that allows for mean variance analysis. This and other results concerning higher order moments have been discussed and analyzed in many different contexts by Samuelson (1970); Jean (1971); Kraus and Litzenberger (1976); Scott and Horvath (1980); Kane (1982); Owen and Rabinovitch (1983); Ingersoll (1987); Kogan and Uppal (2001); Ortobelli et al. (2003); Harvey et al. (2004); Brandt et al. (2005) or Ait-Sahalia, Cacho-Diaz and Hurd (2009)[1], among many others. While this vast literature has demonstrated the relevance of higher order moments for portfolio decisions, most of the work relies on approximations and simulations. My paper adds some intuition to many of these results by working out simple analytical versions.

The basic approach followed in this paper is straightforward only under constant absolute risk aversion (CARA) utility. However, I also provide new intuition with a constant relative risk aversion (CRRA) formulation, primarily under the assumption of a new kind of budget constraint. Additionally, when possible, I close the model in order to understand equilibrium pricing effects.

The main contribution of this paper is the introduction of a tractable analytic framework to understand the implication of higher order terms in the portfolio choice problem. As discussed by Campbell and Viceira (2002) and Brandt (2009), many other considerations are crucial in more realistic portfolio choice environments. These include long term investment horizons, life cycle considerations, uninsurable labor income risk, learning or model uncertainty; I do not address such issues here.

Many of the results in this literature tilt optimal portfolios towards risky assets in a way

---

[1]The introduction of this last paper provides a detailed set of references in both the Markowitz (1952) and Merton (1969) tradition.

that does not match actual data. I tractably explore whether departures from the assumption of normal returns can rationalize those behaviors. The main conclusion is that, unless we assume important unrealized shocks (rare disasters), standard parametric choices for excess return, even accounting for skewness or kurtosis, are not enough to change portfolio prescriptions.

I calibrate the model according to an actual portfolio choice problem in order to illustrate the magnitudes of the new effects. My results also seem particularly promising in environments with segmented markets, as in the market microstructure and behavioral finance literatures, as argued in section 6.

Section 2 of the paper analyzes CARA utility and section 3 shows the results with CRRA utility. I discuss the results along with possible extensions of this work in section 4 and I conclude in section 5. The appendix shows some properties of CGF's and entropy, summarizes some properties of the excess return distributions and proves the multivariate generalization for the CARA case.

2 CARA utility and the cumulant generating function

2.1 The portfolio choice problem

I assume that time is discrete and I focus my attention in the decision between periods t and $t+1$. There is a single agent who at period t maximizes expected utility $U(\cdot)$ defined over final levels of terminal wealth W_{t+1}. In this section a assume constant absolute risk aversion (CARA) utility: $U(W_{t+1}) = -e^{-AW_{t+1}}$.

The agent initially has an amount of wealth $\tilde{W}_t$ that can be invested in a riskless asset, with gross return $1+r$, or in a risky asset, with gross return given by $1+r+\tilde{\theta}$. Therefore, $\tilde{\theta}$ is the stochastic excess return generated by investing in the risky asset. The control variable for the agent is the amount (in dollars) invested in the risky asset, denoted by x. The problem can be formulated as

$$\max_x \mathbb{E}_t \left[U(W_{t+1}) \right] \tag{1}$$

Subject to

$$W_{t+1} = \left(\hat{W}_t - x \right)(1+r) + x \left(1+r+\tilde{\theta} \right)$$
$$= W_t + x\tilde{\theta} \tag{2}$$

Where I have redefined $W_t = \tilde{W}_t(1+r)$. In order to use the definition of the CGF, I rely on the following lemma.

Lemma 1. *Without loss of generality, the objective function* (1) *can be expressed as* [2]

$$\begin{cases} \max_x \log \mathbb{E}_t \left[U\left(W_{t+1}\right) \right] & \text{if } \mathbb{E}_t \left[U\left(W_{t+1}\right) \right] \geq 0 \\ \min_x \log -\mathbb{E}_t \left[U\left(W_{t+1}\right) \right] & \text{if } \mathbb{E}_t \left[U\left(W_{t+1}\right) \right] < 0 \end{cases}$$

When discussing the CARA utility case with normal and Laplace excess return, I solve a multivariate version of the problem. In that case, instead of a single risky asset, I assume a multivariate distribution for the whole set of possible risky assets. The rest of the problem remains the same. The budget constraint then becomes

$$\begin{aligned} W_{t+1} &= \left(\hat{W}_t - \sum_{i=1}^{k} x_i \right) (1+r) + x_1 \left(1 + r + \tilde{\theta}_1 \right) + \ldots + x_k \left(1 + r + \tilde{\theta}_k \right) \\ &= \tilde{W}_t (1+r) + x_1 \tilde{\theta}_1 + \cdots + x_k \tilde{\theta}_k \\ &= W_t + x \cdot \tilde{\theta} \end{aligned}$$

with $\cdot$ denoting dot product.

2.2 Introducing the CGF

I denote the cumulant generating function of a random variable Y as $CGF_Y(t) = \log \mathbb{E}\left[e^{ty} \right]$. For reference, Billingsley (1995) describes CGF's with detail, and Martin (2008) and Backus, Chernov and Martin (2009) demonstrate many of its modeling virtues in practice. I describe some of the properties of CGF's in appendix A.

I now state the main result with CARA utility of the paper that is applied subsequently to many different distributions for excess return.

Proposition 1. *The objective function of an agent with CARA utility that solves* (1) *with* (2) *can be written as*

$$\min_x CGF_{\tilde{\theta}} (-Ax) \tag{3}$$

Proof. The problem to solve is $\max_x \mathbb{E}_t \left[U\left(W_t + x\tilde{\theta} \right) \right]$. Given Lemma 1, the proposition follows after a short manipulation

---

[2]Simple calculus is enough to see this result. The agent solves $\max_x f(x)$, with FOC $f'(x) = 0$ and subject to SOC $f''(x^*) < 0$. If he solves $\max_x log\left(f(x) \right)$, the new FOC is $\frac{f'(x)}{f(x)} = 0$ and the SOC is given by $f''(x) f(x) - f'(x) \cdot f'(x)$, but this shows the result, since the $f'(x)$ terms are zero in the critical points and the function must be positive to be well defined.

$$\max_x \mathbb{E}_t\left[-e^{-AW_{t+1}}\right] \iff \min_x \log \mathbb{E}_t\left[e^{-A\left(W_t + x\tilde{\theta}\right)}\right]$$

$$\iff \min_x \log \mathbb{E}_t\left[e^{-Ax\tilde{\theta}}\right]$$

$$\iff \min_x CGF_{\tilde{\theta}}\left(-Ax\right)$$

$\square$

The appendix shows the multivariate extension, whose result is identical to this proposition if x is reinterpreted as a vector and the CGF is understood as the multivariate cumulant generating function.

As shown by Billingsley (1995), CGF's are convex. From that fact, it follows that if $CGF\left(x\right)$ is convex, $CGF(-Ax)$ has also to be convex [3]. We can consequently be sure that the problem is well defined and that a unique minimum solution exists.

To gain some intuition, consider expanding $CGF_Y\left(t\right)$ as

$$CGF_Y\left(t\right) = \mu t + \sigma^2 \frac{t^2}{2} + \text{skewness} \cdot \frac{\sigma^3 t^3}{3!} + \text{kurtosis} \cdot \frac{\sigma^4 t^4}{4!} + \dots \tag{4}$$

Given proposition 1, (4) can be written as

$$\min_x CGF_{\tilde{\theta}}\left(-Ax\right) \iff \min_x -A\mu x + \frac{\sigma^2}{2}A^2 x^2 - \frac{\kappa_3}{6}A^3 x^3 + \frac{\kappa_4}{24}A^4 x^4 + \dots$$

And denoting x^* as the optimal portfolio choice

$$x^* = \arg\min_x \mu x - \frac{\sigma^2}{2}Ax^2 + \frac{\kappa_3}{6}A^2 x^3 - \frac{\kappa_4}{24}A^3 x^4 + \dots$$

Therefore, with CARA utility, the optimal portfolio can be understood as the minimizer of a polynomial expression with coefficients given by the cumulants of the excess return. We can see that, given a choice of risky asset x, a CARA agent likes positive odd moments, as mean and skewness, and dislikes even moments as variance and kurtosis.

Taking the FOC that would give the optimal portfolio choice and dividing by A

$$\mu + -\sigma^2 Ax^* + \frac{\kappa_3}{2}A^2 \left(x^*\right)^2 - \frac{\kappa_4}{6}A^3 \left(x^*\right)^3 + \dots = 0$$

In the general form, the solution to the problem would be

---

[3]Proof: $f \equiv CGF\left(-Ax\right)$, then $f' = -A \cdot CGF'\left(-Ax\right)$ and $f'' = A^2 \cdot CGF''\left(-Ax\right)$, but we know that $CGF'' > 0$, therefore it is true that $CGF\left(-Ax\right)$ must be convex.

$$\sum_{n=1}^{\infty} \frac{(-1)^n \kappa_n A^{n-1}}{(n-1)!} (x^*)^{n-1} = 0 \tag{5}$$

We can clearly see that all the higher order moments matter for the portfolio choice distribution in a non negligible way. One possible way to proceed would be to truncate the expansion of the CGF and then try to understand the results. Martin (2008) shows that this approach can have important drawbacks if used in consumption based asset pricing when there are important tail events. From (5) it is straightforward to see that the same concern also applies here.

In order to overcome these problems, I derive multiple analytical solutions for particular distributions of the excess return. The distributions discussed in the paper simply illustrate the potential effects of higher order moments and they are by no means the only cases that can be solved. I am confident that further research will be able to identify more tractable and flexible distributions.

I divide the distributions for excess return in two categories:

1. Distributions with only two free parameters: once we have fixed mean and variance for this type of distributions, all the higher order cumulants are automatically determined. This category includes the normal, Laplace, Gumbel, uniform, logistic and skew-logistic.

2. Distributions with more than two free parameters: holding constant mean and variance, there are other parameters that determine variation in higher order cumulants. This category include the truncated normal, skew-normal, normal with Bernoulli shocks and normal with poisson-normal shocks.

In the cases of the normal and Laplace I am able to solve a multivariate generalization. With normal, Laplace and Gumbel distributions I can solve for the equilibrium price in a simple market model, since these are the only cases with a true closed form solution. After deriving the analytical results, I calibrate the different distributions to match historical data for excess return. Appendix B provides detailed information about every distribution considered throughout the paper.

2.3 Benchmark case: normal excess return

This is the classical case used in many papers. The excess return for the risky asset follows a normal distribution $\tilde{\theta} \sim N\left(\mu, \sigma^2\right)$.

A normal $N\left(\mu, \sigma^2\right)$ has the following CGF

$$CGF_{N(\mu,\sigma^2)}(t) = t\mu + \frac{1}{2}t^2\sigma^2$$

According to proposition 1, the agent solves

$$\min_x CGF_{\tilde{\theta}}(-Ax) \iff \min_x -Ax\mu + \frac{1}{2}A^2x^2\sigma^2$$

Which yields the traditional result for the optimal amount of the risky asset in the portfolio

$$x^* = \frac{\mu}{A\sigma^2} \tag{6}$$

This result conveys the intuition that a risk averse agent will tilt his portfolio towards riskier assets the higher the excess return and the lower his risk aversion and the variance of the excess return. It is also noteworthy that the optimal amount of risky asset is invariant to the wealth of the agent, a direct consequence of CARA utility.

2.3.1 Multivariate normal returns

The multivariate generalization is straightforward. Assume that the vector $\tilde{\theta}$ is distributed jointly normal with mean μ, and covariance matrix Σ. A multivariate normal $N(\mu, \Sigma)$ has the following CGF

$$CGF_{N(\mu,\Sigma)}(t) = t'\mu + \frac{1}{2}t'\Sigma t$$

According to the generalization of proposition 1, shown in the appendix, the agent solves

$$\min_x CGF_{\tilde{\theta}}(-Ax) \iff \min_x -Ax'\mu + \frac{1}{2}(-Ax)'\Sigma(-Ax)$$

By using standard matrix calculus, the solution is

$$x^* = \frac{\Sigma^{-1}\mu}{A} \tag{7}$$

2.4 Laplace/Double exponential distribution

This is the first new analytical result provided by the framework proposed in this paper[4]. Assume that $\tilde{\theta}$ follows a Laplace distribution $LA(\mu, b)$. This distribution is symmetric and leptokurtic[5], with mean μ and variance $2b^2$. It is also referred to as the double exponential distribution, since each half of the distribution can be seen as an exponential distribution scaled by $1/2$. It

---

[4]It is a fact that Ingersoll (1987) has an equivalent derivation when discussing elliptical distributions, although he doesn't use CGF's as a building block.

[5]Remember that a distribution is called leptokurtic when its kurtosis is greater than 3 (the kurtosis of the normal) and platykurtic when its kurtosis is less than 3. Given a value for the variance, a leptokurtic distribution has a higher peak than normal and fatter tails, while a platykurtic distribution has a wider peak and less mass in the tails.

is a member of the family of the generalized normal distributions and of the class of elliptical distributions.

A Laplace distribution $LA(\mu, b)$ has the following CGF

$$CGF_{LA(\mu,b)}(t) = \log\left[\frac{\exp(\mu t)}{1 - b^2 t^2}\right]$$

According to proposition 1, the agent solves

$$\min_x CGF_{\tilde{\theta}}(-Ax) \iff \min_x \log\left[\frac{\exp(-Ax\mu)}{1 - b^2 A^2 x^2}\right]$$

After optimizing, the solution becomes

$$x^* = \frac{\mu}{A2b^2}\left(1 - b^2 A^2 (x^*)^2\right) \tag{8}$$

This last expression looks similar to the traditional case with CARA utility and normal excess return, but with an attenuation or dampening effect generated by the excess kurtosis of the distribution. The quadratic can be solved explicitly and the only valid solution is[6]

$$x^* = \frac{\sqrt{1 + \frac{\mu^2}{b^2}} - 1}{A\mu} \tag{9}$$

Since $\sqrt{1 + x} < 1 + \frac{x}{2}$, for a given mean and variance, the optimal portfolio choice with Laplace returns is always smaller than with normal returns. The comparative statics, $\frac{\partial x^*}{\partial \mu} > 0$ and $\frac{\partial x^*}{\partial b^2} < 0$ are also as expected.

2.4.1 Multivariate Laplace returns

As with the normal case, I can solve for the multivariate version of the Laplace. The CGF of the multivariate Laplace distribution $L(\mu, \lambda, \Omega)$ is defined by[7]

$$CGF_{L(\mu,\lambda,\Omega)}(t) = \mu' t - \log\left(1 - \lambda\frac{t'\Omega t}{2}\right)$$

According to the generalization of proposition 1, the agent solves

$$\min_x CGF_{\tilde{\theta}}(-Ax) \iff \min_x -\mu' x A - \log\left(1 - \frac{\lambda}{2}A^2 x'\Omega x\right)$$

By using standard matrix calculus, the solution x^* becomes

---

[6]The other root implies a negative optimal amount with $\mu > 0$ and vice versa with $\mu < 0$.

[7]λ is a variance parameter and Ω is a matrix of covariance parameters.

$$x^* = \frac{\Omega^{-1}\mu}{\frac{\lambda}{2}A}\left[1 - \frac{\lambda}{2}A\left(x^*\right)'\Omega x^*\right] \tag{10}$$

It is obvious that (10) is simply a generalization of (8). Unfortunately, there is no closed form solution for multivariate quadratic equations, but the intuition that optimal portfolio decisions are more conservative with fat tails than with normal tails continues to hold.

2.5 Gumbel distribution

The Gumbel distribution is a two parameter distribution (μ, β) that shows negative skewness and excess kurtosis[8]. It is a particular case of the generalized extreme value distributions widely used in extreme value theory.

A Gumbel distribution $G(\mu, \beta)$ has the following CGF [9]

$$CGF_{G(\mu,\beta)}(t) = t\mu + \log\left(\Gamma\left(1 + \beta t\right)\right)$$

According to proposition 1, the agent solves

$$\min_x CGF_{\tilde{\theta}}(-Ax) \iff \min_x -Ax\mu + \log\Gamma\left(1 - \beta Ax\right)$$

After optimizing, the solution becomes[10]

$$\Psi\left(1 - \beta Ax^*\right) = -\frac{\mu}{\beta}$$

$$x^* = \frac{1 - \Psi^{-1}\left(-\frac{\mu}{\beta}\right)}{\beta A} \tag{11}$$

We must assume that $1 > \beta Ax$ in order for the CGF to be well defined. The digamma function in that range is everywhere increasing and negative, so a unique solution is guaranteed. Even though there is no closed form solution in this case, the result can be aggregated to find an expression for the asset price, as I demonstrate in the next section. It is straightforward to derive comparative statics implications: higher μ and lower β or A reduce the optimal allocation to the risky asset.

---

[8]Some sources define a Gumbel distribution as the additive inverse of the way I am doing it. In that case the function would show positive skewness and all its cumulants would be positive. Check the appendix for details about the moments of my formulation and the shape of the distribution.

[9]Γ denotes the gamma function here and throughout the rest of the paper.

[10]Ψ denotes the digamma function. The digamma function is defined as the logarithmic derivative of the gamma function: $\Psi(v) = \frac{d}{dv}\log\Gamma(v) = \frac{\Gamma'(v)}{\Gamma(v)}$.

2.6 Uniform distribution

Assume now that the excess return has a uniform distribution between a and b. A uniform $U(a, b)$ has the following CGF

$$CGF_{U(a,b)}(t) = \log\left(e^{tb} - e^{ta}\right) - \log\left(t\left(b - a\right)\right)$$

According to proposition 1, the agent solves

$$\min_x CGF_{\tilde{\theta}}(-Ax) \iff \min_x \log\left(e^{-Axa} - e^{-Axb}\right) - \log Ax\left(b - a\right)$$

After optimizing, the solution becomes

$$x^* = \frac{1}{A} \frac{e^{-Aax^*} - e^{-Abx^*}}{-ae^{-Aax^*} + be^{-Abx^*}}$$

Although it is extremely simple to compute the result analytically, there is not a true closed form solution[11]. Once again the comparative statics $\frac{\partial x^*}{\partial b} > 0$ and $\frac{\partial x^*}{\partial a} < 0$ are as expected.

2.7 Logistic distribution

The logistic distribution is a symmetric two parameter distribution (μ, s). It resembles the normal distribution but it has fatter tails (its excess kurtosis is $\frac{6}{5}$). Its mean is μ and its variance is $\frac{\pi^2}{3}s^2$.

A logistic $LG(\mu, s)$ has the following CGF[12]

$$CGF_{LG}(t) = \mu t + \log\left[B\left(1 - st, 1 + st\right)\right]$$

According to proposition 1, the agent solves

$$\min_x CGF_{\tilde{\theta}}(-Ax) \iff \min_x -Ax\mu + \log\left[B\left(1 + sAx, 1 - sAx\right)\right]$$

Using the properties of the beta function, this can be rewritten as

---

[11] If we made the additional assumption that the distribution is symmetric around zero, such that $a = -b$, we would find that $x^* = \frac{1}{Ab}\tanh(Abx^*)$. However, as we can expect, the only solution to this problem is to set $x^* = 0$, since the mean of the excess return is zero and there is positive variance, but perhaps this relation with hyperbolic tangents can be interesting to study in detail.

[12] $B(\cdot, \cdot)$ denotes the (Euler) beta function. It can be written as $B(v, w) = \frac{\Gamma(v)\Gamma(w)}{\Gamma(v+w)}$, with $\Gamma(\cdot)$ denoting the gamma function.

$$\min_{x} -Ax\mu + \log \Gamma \left(1 + sAx\right) + \log \Gamma \left(1 - sAx\right)$$

After optimizing, the solution becomes

$$\frac{\mu}{s} = \Psi \left(1 + sAx^*\right) - \Psi \left(1 - sAx^*\right)$$

Once again, the comparative statics give the expected result that higher μ and lower A and s increase the optimal x^*.

2.8 Skew-logistic distribution

Johnson, Kotz and Balakrishnan (2002) describe many generalized logistic models. The so called Type IV generalized logistic is the most general case and has a *CGF*

$$CGF_{SL}(t) = \frac{\Gamma(\beta - t)\Gamma(\alpha + t)}{\Gamma(\alpha)\Gamma(\beta)}$$

With the restriction that $-\alpha < t < \beta$. All the cumulants depend non trivially on the values of α and β, as described in appendix B.

According to proposition 1, the agent solves

$$\min_{x} CGF_{\tilde{\theta}} \left(-Ax\right) \iff \min_{x} \frac{\Gamma(\beta + Ax)\Gamma(\alpha - Ax)}{\Gamma(\alpha)\Gamma(\beta)}$$

After optimizing, the solution becomes

$$\psi(\beta + Ax^*) = \psi(\alpha - Ax^*)$$

Given that the digamma function is strictly increasing, the optimal choice is

$$x^* = \frac{\alpha - \beta}{2A} \tag{12}$$

The solution (12) is one of the simplest that can be found in the paper. When $\alpha = \beta$, the distribution collapses to the the type III generalized logistic distribution, and the optimal position is to hold 0 of the risky asset. If $\alpha > \beta$ there is positive skewness and if $\alpha < \beta$ negative skewness.

2.9 Calibration of two-parameter distributions

With the previous analytical results, it is helpful to compare the effect of different higher order terms in the optimal portfolio choice problem. I proceed to calibrate each distribution by matching its first two moments to the historical series of excess returns. I use annual data for excess returns from the Kenneth French webpage, from 1926 until 2009. The estimated mean is $\kappa_1 = 0.08$ and the standard deviation $\kappa_2^{1/2} = 0.2$. Using different values for mean and variance only changes the results slightly and the basic intuitions are unchanged.

Figure 1 represents on the vertical axis the ratio $\frac{x^*}{x^*_{Normal}}$ for different values of relative risk aversion[13].
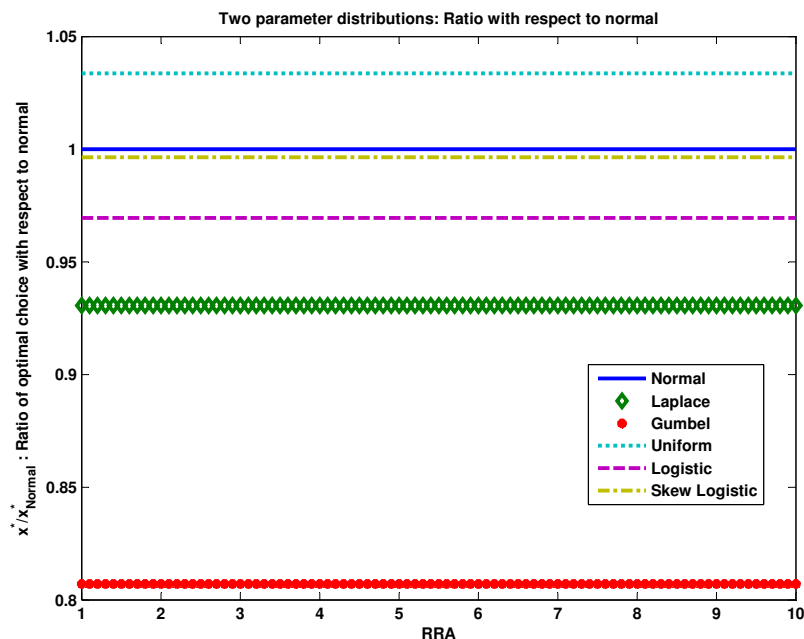

Figure 1: Two parameter distributions $\kappa_1 = 0.08$, $\kappa_2^{1/2} = 0.2$

It may seem surprising that the ratios are flat for all wealth levels; this simply reflects a two-fund separation result, as discussed by Ingersoll (1987).

As we expected, holding the mean and the variance constant, a uniform distribution that features less mass in the tails makes the CARA agent to opt for slightly more aggressive positions. The other distributions feature either fatter tails or negative skewness in different combinations (see appendix with the plot of the cdf's). This is reflected in the plot. The most relevant conclusion from this analysis is that the optimal portfolio for these parameterizations only

---

[13]I convert relative risk aversion values into absolute risk aversion before using the formulas derived in the paper.

departs from the normal prescription in less than 20 percent.

2.10   Truncated normal

I now move to distributions with free parameters in addition to mean and variance. The truncated normal is distributed as a normal confined into the interval $[a, b]$, with $-\infty \leq a < b \leq \infty$. This distribution for excess return is not subject to criticisms linked to limited liability. A truncated normal $TN\left(\mu, \sigma^2, a, b\right)$ has the following CGF

$$CGF_{TN(\mu,\sigma^2,a,b)}(t) = \mu t + \frac{1}{2}\sigma^2 t^2 + \log \frac{\Phi\left(-\sigma t + \frac{b-\mu}{\sigma}\right) - \Phi\left(-\sigma t + \frac{a-\mu}{\sigma}\right)}{\Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)}$$

According to proposition 1, the agent solves

$$\min_x CGF_{\hat{\theta}}(-Ax) \iff \min_x -Ax\mu + \frac{1}{2}\sigma^2 A^2 x^2 + \log \frac{\Phi\left(Ax\sigma + \frac{b-\mu}{\sigma}\right) - \Phi\left(Ax\sigma + \frac{a-\mu}{\sigma}\right)}{\Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)}$$

After optimizing, the solution becomes

$$x^* = \frac{\mu - \sigma \frac{\phi\left(Ax^*\sigma + \frac{b-\mu}{\sigma}\right) - \phi\left(Ax^*\sigma + \frac{a-\mu}{\sigma}\right)}{\Phi\left(Ax^*\sigma + \frac{b-\mu}{\sigma}\right) - \Phi\left(Ax^*\sigma + \frac{a-\mu}{\sigma}\right)}}{A\sigma^2}$$

Two relevant cases are $a \to -\infty$ and $b \to \infty$. [14]

$$\lim_{a \to -\infty} x^* = \frac{\mu - \sigma \frac{\phi\left(Ax^*\sigma + \frac{b-\mu}{\sigma}\right)}{\Phi\left(Ax^*\sigma + \frac{b-\mu}{\sigma}\right)}}{A\sigma^2}$$

$$\lim_{b \to \infty} x^* = \frac{\mu + \sigma \frac{\phi\left(Ax^*\sigma + \frac{a-\mu}{\sigma}\right)}{1 - \Phi\left(Ax^*\sigma + \frac{a-\mu}{\sigma}\right)}}{A\sigma^2} = \frac{\mu + \sigma\lambda\left(Ax^*\sigma + \frac{a-\mu}{\sigma}\right)}{A\sigma^2}$$

The last limiting result shows analytically that optimal portfolio decisions with limited liability are tilted towards more aggressive positions.

2.10.1   Calibration

I assume $b = \infty$ for this calibration, what can be interpreted as limiting liability or truncating losses at the value a. Figure 2 shows the ratio between the optimal choice with truncation

---

[14] I have denoted the hazard rate of the normal by $\lambda\left(\cdot\right) = \frac{\phi(\cdot)}{1 - \Phi(\cdot)}$.

and the optimal choice with normal returns for different values of a. The vertical lines denote the values of one and two negative standard deviations. As in the previous case, this ratio is independent of risk aversion.
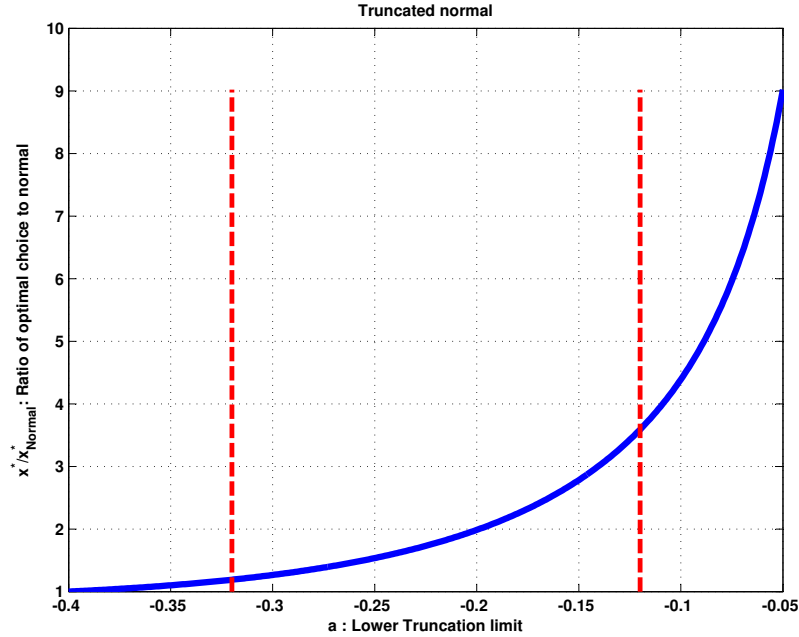

Figure 2: Truncated normal distribution $\kappa_1 = 0.08$, $\kappa_2^{1/2} = 0.2$, $b = \infty$ $a \in [-0.4, -0.05]$

2.11 Skew-normal distribution

Excess return $\tilde{\theta}$ is now distributed as a skew-normal $\tilde{\theta} \sim SW(\xi, \omega, \alpha)$. This distribution can be thought of as a generalized version of the normal that allows for skewness. This is a 3-parameter distribution, where ξ can be understood as a location parameter, ω as a scale parameter and α as a skewness parameter. When $\alpha = 0$, the distribution collapses into a normal, where ξ is exactly the mean and ω^2 is the variance. This distribution has positive skewness for $\alpha > 0$ and it has negative skewness for $\alpha < 0$. We should note that independently of parameters, the excess kurtosis is always negative, i.e., the kurtosis is less than the kurtosis for a normal. In appendix B I provide some details about this distribution; see also Azzalini (1985) for more details.

I use a reparameterization of α to ease notation: $\delta = \frac{\alpha}{\sqrt{1+\alpha^2}}$ [15]

A skew-normal $SW(\xi, \omega, \alpha)$ has the following CGF[16]

---

[15]Note that δ keeps the same sign as α, and collapses to 0 in the normal case. It is also the case that $\delta \in [-1, 1]$, mapping the values of $\alpha \in \mathbb{R}$ to $[-1, 1]$.

[16]$\Phi(\cdot)$ and $\phi(\cdot)$ denote respectively the cdf and pdf of the standard normal.

$$CGF_{SW(\xi,\omega,\alpha)}(t) = \log\left[2\exp\left(\xi t + \frac{\omega^2 t^2}{2}\right)\Phi(\omega\delta t)\right]$$

According to proposition 1, the agent solves

$$\min_x CGF_{\tilde{\theta}}(-Ax) \Longleftrightarrow \min_x -Ax\xi + \frac{\omega^2 A^2 x^2}{2} + \log\left[\Phi(-\omega\delta Ax)\right]$$

After optimizing, the solution becomes[17]

$$x^* = \frac{\xi}{A\omega^2} + \frac{\delta}{A\omega}\lambda(\omega\delta Ax^*) \tag{13}$$

As we expected, in the case where $\alpha = \delta = 0$, the problem is the standard CARA with normal excess return, and the optimal solution is exactly as in (6). When there is positive or negative skewness, an additional term adjusts for that.

2.11.1 Calibration

I adjust mean and variance to their empirical counterparts and allow for different values of skewness δ. As before, the two-fund separation result applies, so the only variable of interest is $\frac{x^*}{x^*_{Normal}}$, independently of risk aversion.

---

[17]$\lambda(\cdot) = \frac{\phi(\cdot)}{1-\Phi(\cdot)}$ denotes again the hazard rate of the normal. I use the fact that $\phi(-s) = \phi(s)$ and $\Phi(-s) = 1 - \Phi(s)$.
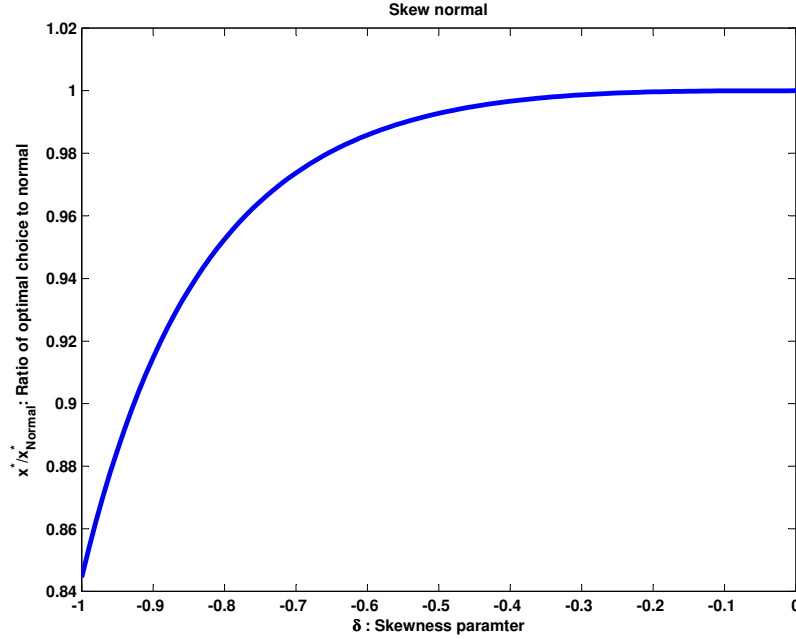
Figure 3: Skew-normal distribution $\kappa_1 = 0.08$, $\kappa_2^{1/2} = 0.2$, $\delta \in [-1, 0]$

Figure 3 demonstrates that for small values of skewness the effects on the optimal choice are quantitatively small.

2.12 Normal with Bernoulli Shock

In this case, I explicitly deal with tail events for the first time. Following the disasters literature, e.g., Backus, Chernov and Martin (2009), I assume that the excess return is given by the sum of a normal distribution and a Bernoulli shock, i.e., $ER = X_1 + X_2$, with $X_1 \sim N(\mu, \sigma)$ and $X_2 \sim B(a, b, p)$, where X_1 and X_2 are independent. The Bernoulli distribution $B(a, b, p)$ is defined by

$$B(a, b, p) = \begin{cases} a, & \text{prob} = p \\ b, & \text{prob} = 1 - p \end{cases}$$

The CGF for $B(a, b, p)$ is given by

$$CGF_{B(a,b,p)}(t) = \log\left(pe^{ta} + (1 - p)e^{tb}\right)$$

The use of CARA utility allows us to combine uncorrelated sources of risk, so simple addition of CGF's is sufficient. Hence the CGF of the excess return is $CGF_{ER} = CGF_{N(\mu,\sigma)} + CGF_{B(a,b,p)}$

$$CGF_{ER}(t) = t\mu + \frac{1}{2}t^2\sigma^2 + \log\left(pe^{ta} + (1-p)\,e^{tb}\right)$$

According to proposition 1, the agent solves

$$\min_x CGF_{\tilde{\theta}}(-Ax) \iff \min_x -Ax\mu + \frac{1}{2}A^2x^2\sigma^2 + \log\left(pe^{-Aax} + (1-p)\,e^{-Abx}\right)$$

After optimizing, the solution becomes

$$x^* = \frac{\mu}{A\sigma^2} + \frac{1}{A\sigma^2}\frac{ape^{-Aax^*} + b(1-p)\,e^{-Abx^*}}{pe^{-Aax^*} + (1-p)\,e^{-Abx^*}}$$

This condition has a straightforward interpretation: the first term corresponds to the classic normal solution and the second term provides the correction for the disaster effect in the excess return. When p approaches 0 or 1, the Bernoulli shock simply enters in the solution as a mean shifting correction, but when $p \in (0,1)$, risk aversion and the amount of risk held interact to determine the optimal solution.

To see this more clearly, suppose $b = 0$ and $a < 0$, a period with exceptionally low risk premium that occurs with probability p. The optimal amount invested in the risky asset is then

$$x^* = \frac{\mu}{A\sigma^2} + \frac{a}{A\sigma^2}\frac{1}{1 + \frac{1-p}{p}e^{Aax^*}}$$

As argued above, when $p \to 1$, there is a shift of mean and the optimal portfolio becomes $x^* = \frac{\mu+a}{A\sigma^2}$. When $p \to 0$, we approach to the normal benchmark. For intermediate values of p, given that $0 \leq \frac{1}{1 + \frac{1-p}{p}e^{Aax}} \leq 1$, the optimal choice lies in between. Note that when $A \to \infty$ and $a < 0$, the effect of the Bernoulli shock in x^* is at a maximum.

2.12.1 Calibration

One way to present the results is to control for overall mean and variance and then allow the Bernoulli parameters to vary. Unfortunately this approach imposes restrictions in higher moments that are not very transparent to variation in parameters. To avoid this problem, I calibrate the underlying normal distribution to the historical returns and allow both a and p, the size and probability of disaster, to vary. My procedure can be understood as a thought experiment in which there is an unobserved shock in addition to the conventionally estimated historical returns. The ratio plotted in figure 4 is the ratio of the optimal portfolio with both Bernoulli and normal excess return to the case with only normal excess return.
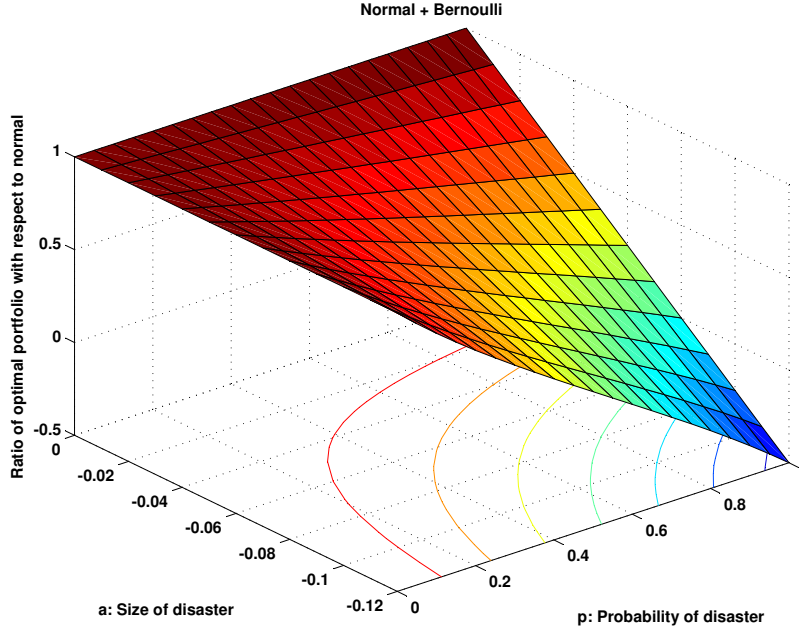
Figure 4: Normal with Bernoulli shocks $\mu = 0.08$, $\sigma = 0.20$, $a \in [-0.12, 0]$, $b = 0$, $p \in [0, 1]$

From the calibration we can conclude that a small probability of a large disaster can cause the optimal portfolio to be reduced by a substantial amount.

2.13   Normal with poisson-normal shock

Paralleling the analysis of Barro (2006) and Martin (2008) with consumption risk, I show how the solutions for the optimal portfolio look when the excess return is given by $ER = X_1 + X_2$, with $X_1 \sim N(\mu, \sigma)$ and $X_2 \sim PN(\lambda, \theta, \delta)$, a poisson-normal mixture with poisson parameter λ, and normal parameters θ and δ^2, where X_1 and X_2 are independent. A poisson-normal mixture implies that conditional on j (the realization of a poisson distribution with parameter λ), X_2 is given by a normal with mean $j\theta$ and variance $j\delta^2$.

Again, CARA utility allows us to combine uncorrelated sources of risk, so the *CGF* of the excess return premium is $CGF_{ER} = CGF_{N(\mu,\sigma)} + CGF_{PN(\lambda,\theta,\delta)}$. Since $CGF_{PN(\lambda,\theta,\delta)}(t) = \lambda \left(e^{t\theta + \frac{(t\delta)^2}{2}} - 1 \right)$

$$CGF_{ER}(t) = t\mu + \frac{1}{2}t^2\sigma^2 + \lambda \left(e^{t\theta + \frac{(t\delta)^2}{2}} - 1 \right)$$

According to proposition 1, the agent solves

$$\min_x CGF_{\tilde{\theta}}\left(-Ax\right) \iff \min_x -Ax\mu + \frac{1}{2}A^2x^2\sigma^2 + \lambda\left(e^{-Ax\theta + \frac{x^2A^2\delta^2}{2}} - 1\right)$$

After optimizing, the solution becomes

$$x^* = \frac{\mu + \lambda\theta + CGF\left(-Ax^*\right)\theta}{A\sigma^2 + \lambda A\delta^2 + A\delta^2 CGF\left(-Ax^*\right)}$$

$$= \frac{\mu\left[1 + \frac{\theta}{\mu}\left[\lambda\left(e^{-Ax^*\theta + \frac{(x^*)^2A^2\delta^2}{2}}\right)\right]\right]}{A\sigma^2\left[1 + \frac{\delta^2}{\sigma^2}\left[\lambda\left(e^{-Ax^*\theta + \frac{(x^*)^2A^2\delta^2}{2}}\right)\right]\right]}$$

As anticipated, a negative value of θ or a large variance δ^2 reduces the optimal amount of the risky asset. Note also that when $\mu = \theta$ and $\delta^2 = \sigma^2$, i.e., the underlying normal excess return shares the same parameters as the poisson-normal shocks, the solution is identical to the normal case. How can this happen? Conditional on the poisson jump j, the excess return scales both mean and variance as $j\mu$ and $j\sigma^2$. Therefore, for any given jump j, the effects on the mean to variance ratio cancel out. More generally we can conclude that $x^* \gtreqless x^*_{Normal}$ if and only if $\frac{\theta}{\delta^2} \gtreqless \frac{\mu}{\sigma^2}$.

2.13.1 Calibration
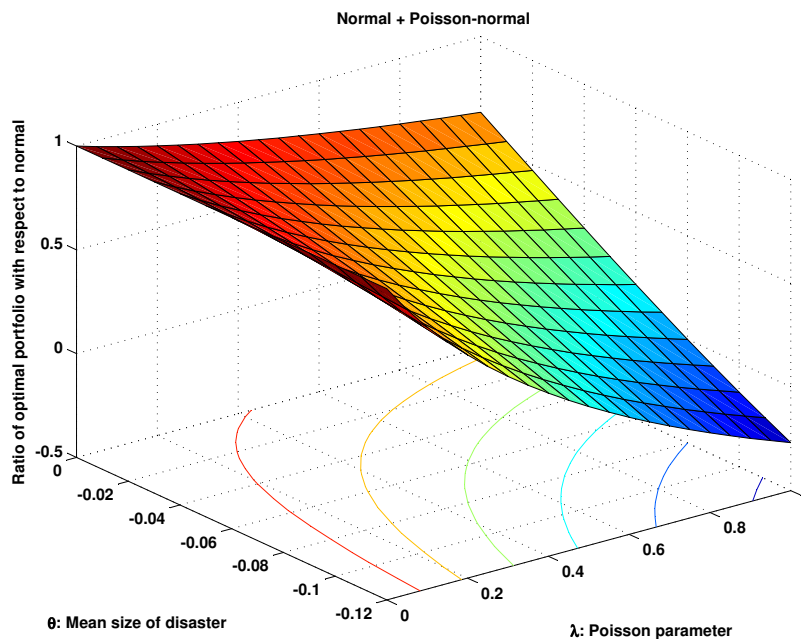
Figure 5: Normal with poisson-normal shocks $\mu = 0.08$, $\sigma = 0.2$, $\lambda \in [0, 1]$, $\theta \in [-0.12, 0]$, $\delta = 0.15$

From figure 5 we see that depending on the size and magnitude of disaster, the optimal portfolio choice can be heavily distorted. Note that, by adding different shocks it would possible to replicate any desired number of moments and then solve for the optimal portfolio as a simple non linear equation.

2.14 Comparison of results

The results derived so far are summarized in the following table:

Distribution	Parameters	Solution
Distributions with two parameters		
Normal	$N\left(\mu,\sigma^2\right)$	$x^* = \frac{\mu}{A\sigma^2}$
Laplace	$L\left(\mu,b^2\right)$	$x^* = \frac{\sqrt{1+\frac{\mu^2}{b^2}}-1}{A\mu}$
Gumbel	$G\left(\mu,\beta^2\right)$	$x^* = \frac{1-\Psi^{-1}\left(\frac{\mu}{\beta}\right)}{\beta A}$
Uniform	$U\left(a,b\right)$	$x^* = \frac{1}{A}\frac{e^{-Aax^*}-e^{-Abx^*}}{-ae^{-Aax^*}+be^{-Abx^*}}$
Logistic	$LG\left(\mu,s^2\right)$	$\frac{\mu}{s} = \Psi\left(1+sAx^*\right)-\Psi\left(1-sAx^*\right)$
Skew-logistic	$SL\left(\alpha,\beta\right)$	$x^* = \frac{\alpha-\beta}{2A}$
Distributions with more than two parameters		
Truncated normal	$TN\left(\mu,\sigma^2,a,b\right)$	$x^* = \frac{\mu-\sigma\frac{\phi\left(Ax^*\sigma+\frac{b-\mu}{\sigma}\right)-\phi\left(Ax^*\sigma+\frac{a-\mu}{\sigma}\right)}{\Phi\left(Ax^*\sigma+\frac{b-\mu}{\sigma}\right)-\Phi\left(Ax^*\sigma+\frac{a-\mu}{\sigma}\right)}}{A\sigma^2}$
Skew-normal	$SW\left(\xi,\omega^2,\delta\right)$	$x^* = \frac{\xi}{A\omega^2}+\frac{\delta}{A\omega}\lambda\left(\omega\delta Ax^*\right)$
N.+Bernoulli	$N\left(\mu,\sigma^2\right)+B\left(a,b,p\right)$	$x^* = \frac{\mu}{A\sigma^2}+\frac{1}{A\sigma^2}\frac{ape^{-Aax^*}+b(1-p)e^{-Abx^*}}{pe^{-Aax^*}+(1-p)e^{-Abx^*}}$
N.+Poisson-N.	$N\left(\mu,\sigma^2\right)+PN\left(\lambda,\theta,\delta^2\right)$	$x^* = \frac{\mu\left[1+\frac{\theta}{\mu}\left[\lambda\left(e^{-Ax^*\theta+\frac{(x^*)^2A^2\delta^2}{2}}\right)\right]\right]}{A\sigma^2\left[1+\frac{\delta^2}{\sigma^2}\left[\lambda\left(e^{-Ax^*\theta+\frac{(x^*)^2A^2\delta^2}{2}}\right)\right]\right]}$

Table 1: Comparison of optimal solutions across distributions

2.15   Some equilibrium implications

The combination of CARA utility with normal returns has become the workhorse model in the market microstructure and behavioral finance literature; e.g., see O'Hara (1998), Shleifer (2000) and Vives (2008). Given the new set of solutions that I just found, it is also interesting to explore how important the effects of higher order moments are when an equilibrium model is imposed.

I recast the problem of section 2 in an environment that uses prices rather than returns. The objective function remains unchanged

$$\max_x \mathbb{E}\left[U\left(W_{t+1}\right)\right] \tag{14}$$

But now the budget constraint takes the equivalent form

$$\begin{aligned} W_{t+1} &= \left(\hat{W}_0 - xP_t\right)(1+r) + P_{t+1}x \\ &= W_0 + x\left[P_{t+1} - P_t(1+r)\right] \end{aligned} \tag{15}$$

In this formulation, P_t is known when making the portfolio decision and all the uncertainty comes from the value of P_{t+1}.

The reformulated problem is identical to the one solved in proposition 1.

$$\max_x \mathbb{E}\left[-e^{-AW_{t+1}}\right] \iff \min_x \log \mathbb{E}\left[e^{-A(W_0+x[P_{t+1}-P_t(1+r)])}\right]$$
$$\iff \min_x \log \mathbb{E}\left[e^{-Ax\tilde{\theta}}\right]$$
$$\iff \min_x CGF_{\tilde{\theta}}(-Ax)$$

Where I have defined $\tilde{\theta} \sim P_{t+1} - P_t(1+r)$. In this environment, the relevant variable is the expected future price of the asset once the current price and the riskfree rate are controlled for. Since $P_t(1+r)$ is not random at t, it represents a mean shift in the return that should not affect the shape of the distribution of P_{t+1}.

In the rest of this section I assume $r = 0$ for simplicity. I also assume the existence of a representative investor that holds the market portfolio. We could equivalently posit many identical agents, and simply divide the risk aversion by N, the number of investors[18]. In the classic CARA utility with normal prices case we find the well known result

$$S = \sum_i x_i = \frac{V - P_t}{A\sigma^2}$$

Where V denotes the expectation of P_{t+1} and S the total amount available of the asset in the economy. The implied equilibrium price is

$$P_t = V - SA\sigma^2 \tag{16}$$

Now suppose that the excess return follows a distribution with fatter tails as the Laplace. The optimal portfolio choice is then given by

$$x = \frac{\sqrt{1 + \frac{[V-P_t]^2}{b^2}} - 1}{A[V - P_t]}$$

Assume that there is a fixed supply S of the asset and that there is a representative investor. The equilibrium price becomes

$$P_t = V - \frac{SA2b^2}{1 - b^2A^2S^2} \tag{17}$$

---

[18]Heterogeneous risk aversion can also be easily handled in this context with minor modifications.

As expected, an increase either the quantity of risk, S or b, or the price of risk, A, increases the risk premium in the economy. It is interesting to note how A and S enter symmetrically as in the normal case.

In order to keep the CGF well defined, we have to impose that $1 - b^2 A^2 S^2 > 0$. This assumption implies that, for a given mean and variance, if the returns of an asset follow a Laplace distribution instead of a normal, the risk premium in that particular market must be larger in equilibrium. This result can be understood as a dual version of the rare disasters literature: if an economy experiences thick tailed stochastic variation in returns, the required risk premium must be larger.

Finally, I solve the same problem with Gumbel returns. From equation (11), the optimal portfolio for every agent is $x^* = \frac{1 - \Psi^{-1}\left(-\frac{\mu}{\beta}\right)}{\beta A}$. Assuming a representative agent and a total amount of asset S, we can write $\mu = -\beta \psi (1 - S\beta A)$. Since $\mathbb{E}\left[\tilde{\theta}\right] = V - P_t = \mu - \beta \tilde{A}$[19], the final expression for P_t becomes

$$P_t = V + \beta \tilde{A} + \beta \psi (1 - S\beta A)$$

The term $\beta \tilde{A}$ is a correction for the fact that β also affects the mean of the return. If β increases, the mean excess return also increases according to $\beta \tilde{A}$. Given an expected $\mathbb{E}_t [P_{t+1}]$, it is necessary that the current price also increases, but this effect is not due to a risk premium. Having accounted for $\beta \tilde{A}$, the actual risk premium is given by $\beta \psi (1 - S\beta A)$. Since we have assumed that $1 > S\beta A$, it can be shown that $\beta \psi (1 - S\beta A) < 0$, as expected.

The comparative statics are standard once again. Higher variance $\frac{\pi^2}{6}\beta^2$, higher amount of risk S and higher risk aversion A induce a higher risk premium.

The following table summarizes the analytical results of this section:

Distribution		Price P_t	Risk Premium
Normal	$N(\mu, \sigma^2)$	$P_t = V - SA\sigma^2$	$SA\sigma^2$
Laplace	$LA(\mu, b)$	$P_t = V - \frac{SA2b^2}{1 - b^2 A^2 S^2}$	$\frac{SA\sigma^2}{1 - b^2 A^2 S^2}$
Gumbel	$G(\mu, \beta)$	$P_t = V + \beta \tilde{A} + \beta \psi (1 - S\beta A)$	$-\beta \psi (1 - S\beta A)$

Table 2: Results for the equilibrium model

In order to evaluate the relevance of higher order cumulants in the risk premium, I plot the ratio of each risk premium with respect to the normal case. Note that the ratio is greater than the unity and that it is not monotonic in the amount of risk tolerance in the economy.

---

[19]$\tilde{A}$ denotes the Euler-Mascheroni constant; $\tilde{A} \approx 0.57721....$. The mean of a Gumbel $G(\mu, \beta)$ is $\mu + \tilde{A}\beta$.
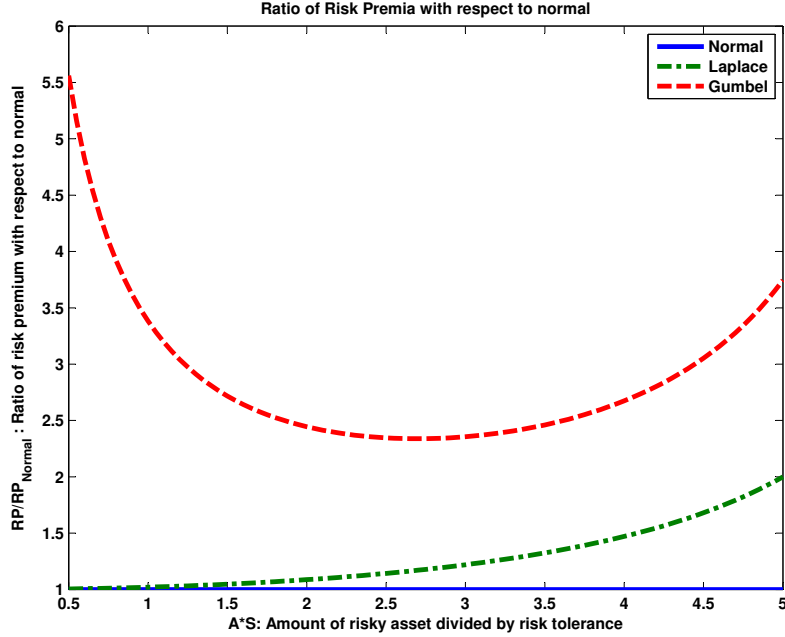
Figure 6: Equilibrium Risk Premium for $\sigma = 0.2$

3 CRRA utility, cumulants and entropy

3.1 General results

After characterizing the explicit solution to the CARA utility problem, I derive new implications of higher order returns in constant relative risk aversion (CRRA) environments. The results derived in this subsection work for any budget constraint specification that respects $W_{t+1} = (1 + R_{p,t+1}) W_t$ and are valid for any CRRA coefficient in $(0, \infty)$. My approach allows me to generalize many of the equations derived in the second chapter of Campbell and Viceira (2002).

The utility function that implies CRRA is $U(W_{t+1}) = \frac{W_{t+1}^{1-\gamma}}{1-\gamma}$, a specification that has become the workhorse in macroeconomics. In this section, γ denotes relative risk aversion.

Since I do not use a specific budget constraint in this subsection, I am not forced to specify the control variable x. The literature usually optimizes utility in the share of wealth, rather than the dollar amount. My derivations in 4.1 and 4.2 are consistent with this approach. In subsections 4.3 and 4.4, I introduce a different budget constraint.

Under CRRA, the objective function is

$$\max_x \mathbb{E}_t \left[\frac{W_{t+1}^{1-\gamma}}{1-\gamma} \right] \tag{18}$$

24

Taking logs of this expression[20] and applying the definition of entropy[21], extensively used by Alvarez and Jermann (2005) and originally introduced in Theil (1967), the problem becomes

$$\max_x \log \mathbb{E}_t \left[\frac{W_{t+1}^{1-\gamma}}{1-\gamma} \right] \iff \max_x (1-\gamma) \mathbb{E}_t [w_{t+1}] + \mathbb{L}\left(W_{t+1}^{1-\gamma}\right) \tag{19}$$

This expression[22] is the generalization of equation 2.16 in Campbell and Viceira (2002). In the case of lognormal wealth, it can be shown that $\mathbb{L}\left(W_{t+1}^{1-\gamma}\right) = \frac{1}{2}(1-\gamma)^2 \sigma_{wt}^2$.

We can gain more intuition about the problem by working with the budget constraint. Since $W_{t+1} = (1 + R_{p,t+1}) W_t$, then $w_{t+1} = r_{p,t+1} + w_t$.

Substituting in (19), the problem becomes[23]

$$\max_x \mathbb{E}_t [r_{p,t+1}] + \frac{\mathbb{L}\left((1 + R_{p,t+1})^{1-\gamma}\right)}{1-\gamma} \tag{20}$$

When an agent has CRRA utility, he maximizes the expected log return minus an additional term that accounts for the entropy of the $1-\gamma$ power of portfolio returns. Entropy is linked to all higher order cumulants starting with the second cumulant. Using this link between entropy and cumulants we can rewrite (20) as

$$\max_x \mathbb{E}_t [r_{p,t+1}] + \sum_{j=2}^{\infty} \frac{(1-\gamma)^{j-1} \kappa_j [r_{p,t+1}]}{j!} \tag{21}$$

where $\kappa_j [r_{p,t+1}]$ denotes the cumulants of $r_{p,t+1}$, the log return of the overall portfolio.

This last expression generalizes equation 2.18 in Campbell and Viceira (2002). My formulation allows us to prove generally that log utility agents ($\gamma = 1$) only care about the portfolio that maximizes expected log returns, leaving aside higher order cumulants of log returns. By using a central limit theorem, this fact relates to the "growth optimal portfolio" literature of the 60's and 70's, amusingly reflected in the monosyllabic Samuelson (1979).

The expression in (21) may induce us to believe that when $\gamma < 1$ an agent likes large even moments. We can show that this counterintuitive result is misleading. Applying again the entropy decomposition on $\mathbb{E}_t [r_{p,t+1}]$ and using its properties, we find

---

[20]Variables that appear as lower case denote log version of its capital counterparts; for example $w_{t+1} \equiv \log W_{t+1}$. When applied to gross return, they denote log grosss return; for example $r_{p,t+1} = \log (1 + R_{p,t+1})$.

[21]Entropy of random variable X is defined as $\mathbb{L}(X) = \log \mathbb{E}[X] - \mathbb{E}[\log X]$I discuss many of its properties in appendix B.

[22]I'll discuss for simplicity only the case with $\gamma < 1$. When $\gamma > 1$ a result equivalent to lemma 1 applies and all the conclusions remain the same.

[23]Where I have used the property that $L(aX) = L(X)$ for a given $a \in \mathbb{R}$.

$$\mathbb{E}_t\left[r_{p,t+1}\right] = \log\left[\mathbb{E}_t\left(1+R_{p,t+1}\right)\right] - \mathbb{L}\left(1+R_{p,t+1}\right)$$

$$= \log\left[\mathbb{E}_t\left(1+R_{p,t+1}\right)\right] - \sum_{j=2}^{\infty}\frac{\kappa_j\left[r_{p,t+1}\right]}{j!} \tag{22}$$

Using (22), we can reexpress (21) as

$$\max_x \log\left[\mathbb{E}_t\left(1+R_{p,t+1}\right)\right] - \sum_{j=2}^{\infty}\frac{\kappa_j\left[r_{p,t+1}\right]}{j!} + \sum_{j=2}^{\infty}\frac{(1-\gamma)^{j-1}\kappa_j\left[r_{p,t+1}\right]}{j!}$$

$$\max_x \log\left[\mathbb{E}_t\left(1+R_{p,t+1}\right)\right] + \sum_{j=2}^{\infty}\frac{\left((1-\gamma)^{j-1}-1\right)\kappa_j\left[r_{p,t+1}\right]}{j!}$$

This is exactly a generalization of equation 2.20 in Campbell and Viceira (2002). Using the binomial theorem $(1-\gamma)^{j-1} = \sum_{k=0}^{j-1}\binom{j-1}{k}(-\gamma)^k$, we can show neatly the influence of γ

$$\max_x \log\left[\mathbb{E}_t\left(1+R_{p,t+1}\right)\right] + \sum_{j=2}^{\infty}\frac{\left(\sum_{k=0}^{j-1}\binom{j-1}{k}(-\gamma)^k - 1\right)\kappa_j\left[r_{p,t+1}\right]}{j!}$$

$$\max_x \log\left[\mathbb{E}_t\left(1+R_{p,t+1}\right)\right] + \sum_{j=2}^{\infty}\frac{\left(\sum_{k=1}^{j-1}\binom{j-1}{k}(-\gamma)^k\right)\kappa_j\left[r_{p,t+1}\right]}{j!} \tag{23}$$

From this last expression we observe that no matter whether γ is greater or smaller than 1, agents with larger γ penalize high variances and, in general, even cumulants. As is usual in finance, the distinction between arithmetic and geometric returns plays an important role in this result.

Note that the choice of portfolio x determines the expected log return and the cumulants $\kappa_j\left[r_{p,t+1}\right]$. We also observe that even though higher cumulants can be relevant, the $j!$ factorial term in the denominator eventually kills any effect. Nevertheless, as discussed in section 2, it is certainly possible that when truncating the exact expansion in (23), many higher order terms are necessary for an accurate approximation.

3.2 Using parametric entropy

Before discussing how the budget constraint affects the results, it is useful to discuss three parametric examples of entropy. The classic case discussed in Campbell and Viceira (2002) assumes lognormal returns $r_{p,t+1} \sim N\left(\mu, \sigma^2\right)$. The objective function can be written following (21) as

$$\max_{x} \mu + \frac{1}{2}\left(1 - \gamma\right)\sigma^2$$

The agents face a tradeoff between arithmetic return $\mu + \frac{\sigma^2}{2}$ and risk aversion correction $-\gamma\frac{\sigma^2}{2}$.

If we assume that the portfolio log return $r_{p,t+1}$ is distributed as the sum of a normal and a disaster type Bernoulli shock $r_{p,t+1} \sim N\left(\mu, \sigma^2\right) + B\left(a, b = 0, p\right)$, then the objective function using (21) can be written as

$$\max_{x} \mu + \frac{1}{2}\left(1 - \gamma\right)\sigma^2 + \frac{\log\left(1 - p\left(1 - e^{(1-\gamma)a}\right)\right)}{1 - \gamma}$$

To gain some intuition, take a linear Taylor expansion of the additional term. This adds $\frac{\log\left(1 - p\left(1 - e^{(1-\gamma)a}\right)\right)}{1-\gamma} \approx pa$, the mean of the Bernoulli component log return, but it is now obvious how nonlinearities affect the optimal choice of x.

A third parametric model assumes that the portfolio log return $r_{p,t+1}$ is distributed as the sum of a normal and a poisson-normal shock $r_{p,t+1} \sim N\left(\mu, \sigma^2\right) + PN\left(\lambda, \theta, \delta\right)$. In this case, the problem is

$$\max_{x} \mu + \frac{1}{2}\left(1 - \gamma\right)\sigma^2 + \frac{\lambda\left(e^{(1-\gamma)\theta + \frac{[(1-\gamma)\delta]^2}{2}} - 1\right)}{1 - \gamma}$$

Another linear Taylor expansion to the additional term implies that $\frac{\lambda\left(e^{(1-\gamma)\theta + \frac{[(1-\gamma)\delta]^2}{2}} - 1\right)}{1-\gamma} \approx \lambda\left[\theta + \frac{1}{2}\left(1 - \gamma\right)\delta^2\right]$. With this linearization, the individual is assumed to prefer the arithmetic mean of the normal shock and dislike the variance in a similar fashion as with the traditional normal log return. As in the Bernoulli case, the nonlinearities affect the optimal choice of x.

Distribution of log portfolio return $(r_{p,t+1})$		Objective function
Normal	$N\left(\mu, \sigma^2\right)$	$\mu + \frac{1}{2}\left(1 - \gamma\right)\sigma^2$
N.+Bernoulli	$N\left(\mu, \sigma^2\right) + B\left(a, b, p\right)$	$\mu + \frac{1}{2}\left(1 - \gamma\right)\sigma^2 + \frac{\log\left(1 - p\left(1 - e^{(1-\gamma)a}\right)\right)}{1-\gamma}$
N.+Poisson-N.	$N\left(\mu, \sigma^2\right) + PN\left(\lambda, \theta, \delta^2\right)$	$\lambda\theta\mu + \frac{1}{2}\left(1 - \gamma\right)\sigma^2 + \frac{\lambda\left(e^{(1-\gamma)\theta + \frac{[(1-\gamma)\delta]^2}{2}} - 1\right)}{1-\gamma}$

Table 3: Objective functions implied by different $r_{p,t+1}$ distributions

These parametric expressions help us understand how the objective function can be decomposed once we know the overall return of the overall portfolio log return. Unfortunately, as discussed in detail below, once a particular budget constraint is introduced, it is hard for an

agent to select portfolio weights in such a way that the overall portfolio log return follows one of these particular specifications.

3.3 Budget constraints and optimal exposure

The results derived so far under CRRA do not rely on any particular budget constraint specification. The most natural assumption is

$$W_{t+1} = (1 + R_{p,t+1}) W_t = [x (1 + R_{t+1}) + (1 - x) (1 + R_{f,t+1})] W_t$$

where $1 + R_{f,t+1}$ denotes the gross risk free rate, x the optimal share of wealth invested in the risky asset and $1 + R_{t+1}$ the return in the risky asset. This formulation implies that $r_{p,t+1} = \log (x (1 + R_{t+1}) + (1 - x) (1 + R_{f,t+1}))$, and rules out any analytical solution to the problem because it is impossible to derive the distribution and moments of the log portfolio given individual return distributions. The deep concern here is how to find the distribution of the log of a sum, but getting around that problem analytically is simply not feasible. [24]

My approach in this and the next subsections is to specify a new budget constraint and use optimal exposure as the choice variable x. I assume that the expected utility maximizer has the option to choose his market exposure χ, given the following budget constraint

$$W_{t+1} = (1 + R_{p,t+1}) W_t = (1 + R_{m,t+1})^{\chi} W_t \tag{24}$$

where $1 + R_{m,t+1}$ denotes the return on a given reference asset, such as the market portfolio. This formulation is nonstandard, but it captures the idea of an investor choosing among different index exchange traded funds (ETF's). In this setting, $\chi > 1$ is a leveraged ETF while $\chi < 1$ is an inverse ETF. The new choice variable χ can be also understood as a χ percentage change in the portfolio return given a one percent change in the underlying market index

$$\chi = \frac{dr_{p,t+1}}{dr_{m,t+1}}$$

This formulation can be expressed in logs as $r_{p,t+1} = \chi \cdot r_{m,t+1}$. As is required for CRRA utility, this formulation rules out default in all states; when $\chi \to -\infty$, $1 + R_{p,t+1} \to 0$ and when $\chi \to \infty$, $1 + R_{p,t+1} \to \infty$.

We can substitute the new budget constraint in the definition for log portfolio return in (21) to obtain

---

[24]The only way to proceed analytically is by using Campbell and Viceira (2002) approximation, that becomes exact in continuous time. That approximation is driven by the gaussianity of diffusions and Ito's lemma and it wouldn't make any sense in the framework discussed here.

$$\max_{\chi} \mathbb{E}_t \left[\chi \cdot r_{m,t+1} \right] + \sum_{j=2}^{\infty} \frac{(1-\gamma)^{j-1} \kappa_j \left[\chi \cdot r_{m,t+1} \right]}{j!} \tag{25}$$

Using the properties of cumulants in (25)

$$\max_{\chi} \chi \mathbb{E}_t \left[r_{m,t+1} \right] + \sum_{j=2}^{\infty} \frac{(1-\gamma)^{j-1} \chi^j \kappa_j \left[r_{m,t+1} \right]}{j!}$$

It is evident that as long as risk aversion is smaller than 1, the optimal decision implies $\chi = \infty$[25]. I delay discussion of this peculiar result and continue here with my derivations assuming $\gamma > 1$.

Denoting χ^* as the optimal exposure, we get

$$\mathbb{E}_t \left[r_{m,t+1} \right] + \sum_{j=2}^{\infty} \frac{\left[\chi^* (1-\gamma) \right]^{j-1} \kappa_j \left[r_{m,t+1} \right]}{(j-1)!} = 0 \tag{26}$$

Given the new budget constraint introduced in (24), the value of χ^* that solves (26) gives the optimal exposure. It is remarkable to note that, although they represent substantially different problems, equations (5) and (26) are quite similar.

To summarize, the solution to the problem of maximizing (18) subject to (24).

$$\chi^* = \begin{cases} +\infty & \text{if } \gamma \leq 1 \\ \chi^* s.t. \sum_{j=1}^{\infty} \frac{\left[\chi^* (1-\gamma) \right]^{j-1} \kappa_j \left[r_{m,t+1} \right]}{(j-1)!} = 0 & \text{if } \gamma > 1 \end{cases}$$

To gain intuition, we can solve for the optimal exposure with lognormal market returns. For $\gamma > 1$, the optimal exposure is

$$\chi^* = \frac{\mathbb{E}_t \left[r_{m,t+1} \right]}{(\gamma - 1) \kappa_2 \left(r_{m,t+1} \right)}$$

. The optimal exposure increases with the mean of the log market returns and decreases with its variance. As expected, when $\gamma \to \infty$, $\chi^* \to 0$.

Returning to the puzzling fact that a risk averse agent with $\gamma \leq 1$ opts for an infinite exposure to the market, in order to understand better this fact, I analyze the objective function for an agent with lognormal market returns.

$$\max_{x} \chi \mathbb{E}_t \left[r_{m,t+1} \right] + \frac{1}{2} (1-\gamma) \chi^2 \kappa_2 \left[r_{m,t+1} \right]$$

Decomposing this expression

---

[25]From the original objective function $\max_{\chi} \mathbb{E}_t \left[(1 + R_{m,t+1})^{\chi(1-\gamma)} \right]$, it is also easy to show that, as long as $1 \leq \gamma$, the optimal choice of χ, given that $1 + R_{m,t+1}$ is strictly positive, is $\chi = \infty$.

$$\max_x \chi \left[\underbrace{\mathbb{E}_t \left[r_{m,t+1} \right] + \frac{\chi \kappa_2 \left[r_{m,t+1} \right]}{2}}_{\text{Arithmetic Return}} - \underbrace{\frac{\gamma \chi \kappa_2 \left[r_{m,t+1} \right]}{2}}_{\text{Risk Correction}} \right]$$

For a fixed geometric average return, an increase in its exposure χ, increases the arithmetic return (the one the agent cares about) by an amount $\frac{\chi \kappa_2 [r_{m,t+1}]}{2}$, but reduces his utility due to risk aversion by $\frac{\gamma \chi \kappa_2 [r_{m,t+1}]}{2}$. When $\gamma > 1$, this tradeoff is always negative and guarantees a bounded solution. When $\gamma \leq 1$, the agent doesn't dislike risk as much, so the increase in arithmetic return always dominates, pushing the optimal solution out to $\chi^* = \infty$.

When we deviate from the lognormal case, a similar intuition applies. This derivation can be considered a piece of evidence against using a CRRA coefficient lower than or equal to 1 if we believe that these results are not reasonable.

3.4 Going back to CGF's

Given the budget constraint in (24), an approach that mimics the one used with CARA utility becomes valid.

Proposition 2. *The objective function of an agent with CRRA utility that solves* $\max_x \mathbb{E}_t \left[\frac{W_{t+1}^{1-\gamma}}{1-\gamma} \right]$, *assuming* $\gamma > 1$, *with* $W_{t+1} = (1 + R_{m,t+1})^\chi W_t$ *can be written as*

$$\min_\chi CGF_{r_{m,t+1}} \left((1-\gamma) \chi \right)$$

Proof. Given the assumption that $\gamma > 1$

$$
\begin{aligned}
\max_\chi \mathbb{E} \left[\frac{W_{t+1}^{1-\gamma}}{1-\gamma} \right] &\iff \log \min_\chi \mathbb{E} \left[e^{(1-\gamma) \ln(W_{t+1})} \right] \\
&\iff \log \min_\chi \mathbb{E} \left[e^{(1-\gamma) \ln[(1+R_{m,t+1})^\chi W_t]} \right] \\
&\iff \log \min_\chi \mathbb{E} \left[e^{(1-\gamma)\chi r_{m,t+1}} \right] \\
&\iff \min_\chi CGF_{r_{m,t+1}} \left((1-\gamma) \chi \right)
\end{aligned}
$$

$\square$

This result is extremely interesting, since all the analytic results derived with CARA utility will go through. We should replace the distributions for excess return for the distributions of log market returns and we should use $1 - \gamma$ were previously we had coefficients of absolute risk aversion. These expressions, equivalent to those in table 1, are summarized here in the following table.

$r_{m,t+1}$ **Distrib.**	**Parameters**	**Optimal Exposure for $\gamma > 1$**
Distributions with two parameters		
Normal	$N\left(\mu,\sigma^2\right)$	$\chi^* = \frac{\mu}{(\gamma-1)\sigma^2}$
Laplace	$L\left(\mu,b^2\right)$	$\chi^* = \frac{\sqrt{1+\frac{\mu^2}{b^2}}-1}{(\gamma-1)\mu}$
Gumbel	$G\left(\mu,\beta^2\right)$	$\chi^* = \frac{1-\Psi^{-1}\left(\frac{\mu}{\beta}\right)}{\beta(\gamma-1)}$
Uniform	$U\left(a,b\right)$	$\chi^* = \frac{1}{(\gamma-1)}\frac{e^{-(\gamma-1)ax^*}-e^{-(\gamma-1)bx^*}}{-ae^{-(\gamma-1)ax^*}+be^{-(\gamma-1)bx^*}}$
Logistic	$LG\left(\mu,s^2\right)$	$\frac{\mu}{s} = \Psi\left(1+s\left(\gamma-1\right)\chi^*\right) - \Psi\left(1-s\left(\gamma-1\right)\chi^*\right)$
Skew-logistic	$SL\left(\alpha,\beta\right)$	$\chi^* = \frac{\alpha-\beta}{2(\gamma-1)}$
Distributions with more than two parameters		
Truncated normal	$TN\left(\mu,\sigma^2,a,b\right)$	$\chi^* = \frac{\mu-\sigma\frac{\phi\left((\gamma-1)x^*\sigma+\frac{b-\mu}{\sigma}\right)-\phi\left((\gamma-1)x^*\sigma+\frac{a-\mu}{\sigma}\right)}{\Phi\left((\gamma-1)x^*\sigma+\frac{b-\mu}{\sigma}\right)-\Phi\left((\gamma-1)x^*\sigma+\frac{a-\mu}{\sigma}\right)}}{(\gamma-1)\sigma^2}$
Skew-normal	$SW\left(\xi,\omega^2,\delta\right)$	$\chi^* = \frac{\xi}{(\gamma-1)\omega^2} + \frac{\delta}{(\gamma-1)\omega}\lambda\left(\omega\delta\left(\gamma-1\right)x^*\right)$
N.+Bernoulli	$N\left(\mu,\sigma^2\right)+B\left(a,b,p\right)$	$\chi^* = \frac{\mu}{(\gamma-1)\sigma^2} + \frac{1}{(\gamma-1)\sigma^2}\frac{ape^{-(\gamma-1)ax^*}+b(1-p)e^{-(\gamma-1)bx^*}}{pe^{-(\gamma-1)ax^*}+(1-p)e^{-(\gamma-1)bx^*}}$
N.+Poisson -N.	$N\left(\mu,\sigma^2\right)+PN\left(\lambda,\theta,\delta^2\right)$	$\chi^* = \frac{\mu\left[1+\frac{\theta}{\mu}\left[\lambda\left(e^{-(\gamma-1)x^*\theta+\frac{(x^*)^2(\gamma-1)^2\delta^2}{2}}\right)\right]\right]}{(\gamma-1)\sigma^2\left[1+\frac{\delta^2}{\sigma^2}\left[\lambda\left(e^{-(\gamma-1)x^*\theta+\frac{(x^*)^2(\gamma-1)^2\delta^2}{2}}\right)\right]\right]}$

Table 4: Comparison of optimal solutions across distributions

In order to illustrate these results I plot in figure 7 the optimal exposure χ^* for the normal log return. The historical mean for the log market return is 0.573 and its variance is around 0.04. The optimal exposure for other distributions can be inferred from the ratios shown in figure 1.
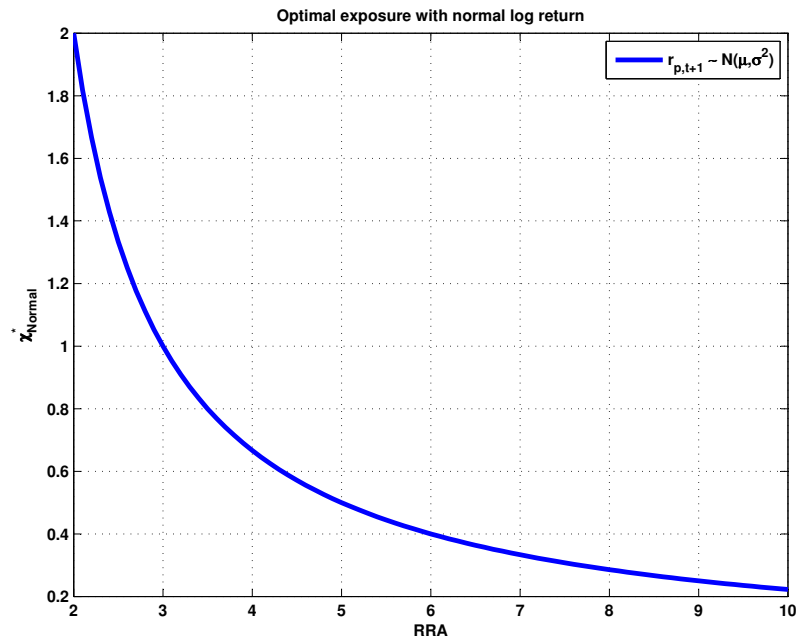
Figure 7: Optimal exposure when log returns are normal

4 Discussion of the results and possible extensions

This paper has adapted different techniques to understand, from an analytical perspective, how higher order cumulants in the distribution of returns can modify the optimal portfolio choice of expected utility maximizers. The paper has focused on comparing these new solutions to the baseline normal case and concludes that non-normality can potentially account for a nonnegligible reduction of risky positions, from values around 20% under Gumbel excess return with CARA utility, to almost any feasible value when we introduce possibly unobserved shocks in the same CARA environment.

The analytical approach in the CRRA case unfortunately faces the insurmountable problem of approximating a sum of logarithms when dealing with the standard budget constraint. I hope that the use of a new budget constraint in section 5 and the results derived there can be understood as a different way of partially get around that problem in a tractable way. That approach seems to rule out on theoretical grounds the use of CRRA coefficients smaller or equal than 1.

Although the work in this paper is framed as an application to classical portfolio choice analysis, it can be easily extended in many directions. Since the CARA formulation is crucial in both market microstructure and behavioral finance, e.g., O'Hara (1998) and Shleifer (2000), there may exist different extensions that would fit in those environments, making the model

32

potentially testable. Extensions to asymmetric information could be interesting as well, and allow modelers to leave the pervasive assumption of normality in those environments behind, e.g., Vives (2008).

Additionally, while I have chosen to explore a particular set of distributions throughout the paper, I hope that research efforts will unveil new distributions whose CGF's and entropy specifications fit nicely with the approach developed here.

Another way to extend this paper is to work through the results by Demange and Laroque (1995) or Athanasoulis and Shiller (2000) and search for optimal market structures given the existence of asymmetric and fat tailed shocks to returns. More generally, different parametric exercises in the tradition of Lucas (1987) or even Rotemberg and Woodford (1997) could be performed inside this framework.

5 Conclusion

The ultimate goal of this paper has been to incorporate many of the recent advances in the rare disasters literature into an analytically tractable normative portfolio choice framework. Despite the simplicity of my formulation, it helps to shed light on the effect of higher order moments on the distribution of returns can determine portfolio choice decisions.

CARA utility makes the problem extremely tractable when combined with many different distributions, allowing for a multivariate extension and equilibrium analysis. CRRA utility also generates new insights to the portfolio problem, mainly when a new budget constraint is introduced. With a traditional budget constraint, numerical simulations are still necessary.

The main practical conclusion of this work is that non gaussian higher moments in the distribution of returns should be taken into consideration by expected utility maximizers. The degree of relevance of these effects depends crucially on the parametric assumption adopted, but reductions of the 20% or larger in risky positions with respect to the normal benchmark are easy to obtain for some excess return distributions featuring excess kurtosis, negative skewness or disaster shocks.

A   Cumulants, CGF and Entropy

A.1   Cumulants and CGF

The cumulants κ_j of a given distribution are related to higher order moments. A detailed reference is chapter 4 in Stuart and Ord (1994) and also Billingsley (1995).

Cumulants represent

$$\kappa_1 \equiv \text{mean}$$
$$\kappa_2 \equiv \text{variance}$$
$$\kappa_3 \equiv \text{skewness} \cdot \sigma^3$$
$$\kappa_4 \equiv \text{excess kurtosis} \cdot \sigma^4$$
$$\kappa_5 \equiv \text{mix of moments...}$$

And some facts about $CGF's$ are

- CGF is convex

- $CGF(0) = 0$, $CGF'(0) = \mathbb{E}[X]$, $CGF''(0) = Var[X]$

- We can write $CGF(t) = \sum_{n=1}^{\infty} \frac{\kappa_n t^n}{n!}$

- $\kappa_j(aX) = a^j \kappa_j(X)$

- If $Y = a + bX$, then $CGF_Y(t) = at + CGF_X(bt)$

- The normal distribution has $\kappa_3 = \kappa_4 = \ldots = 0$

- The only distribution with $\kappa_m = \kappa_{m+1} = \ldots = 0$ for $m \geq 3$ is the normal. This result implies that we don't have so much freedom when choosing higher order cumulants.

A.2   Entropy

Entropy can be written as

$$\begin{aligned}
\mathbb{L}(X) &= \log \mathbb{E}[X] - \mathbb{E}[\log X] \\
&= \log \mathbb{E}\left(e^{\log X}\right) - \mathbb{E}[\log X] \\
&= CGF_{\log X}(1) - \kappa_1[\log X] \\
&= \sum_{j=2}^{\infty} \frac{\kappa_j[\log X]}{j!}
\end{aligned}$$

where κ_j are cumulants of the $\log(X)$.

- $\mathbb{L}(aX) = \mathbb{L}(X)$, with a constant

- $\mathbb{L}(XY) = \mathbb{L}(X) + \mathbb{L}(Y)$ if X and Y are independent

If we assume that $\log X = \alpha + \beta \log Y$ or $X = e^{\alpha + \beta \log Y} = e^{\alpha} e^{\beta \log Y}$, entropy becomes

$$\mathbb{L}(X) = \mathbb{L}\left(e^{\alpha} e^{\beta \log Y}\right) = \mathbb{L}\left(e^{\beta \log Y}\right)$$

But we know that in this case

$$\kappa_j(\log X) = \kappa_j(\log Y) \cdot \beta^j$$

If we specify $\log Y = G + Z$, with G and Z independent, then $\log X = \alpha + \beta(G + Z)$

$$\mathbb{L}(X) = \mathbb{L}\left(e^{\beta G}\right) + \mathbb{L}\left(e^{\beta Z}\right)$$

and

$$\kappa_j(\log X) = \beta^j \kappa_j(G) + \beta^j \kappa_j(Z)$$

Assuming that $G \sim N(\mu, \sigma^2)$ and that Z follows a Bernoulli distribution

$$\mathbb{L}(X) = \beta^2 \frac{\sigma^2}{2} + \log\left(1 - \omega + \omega e^{\beta\theta}\right) - \beta\omega\theta$$

And if Z follows a poisson-normal distribution **CHECK CONCORDANCE WITH TEXT IN NOTATION**

$$\mathbb{L}(X) = \beta^2 \frac{\sigma^2}{2} + \omega\left(e^{\beta\theta + \frac{(\beta\delta)^2}{2}} - 1\right) - \beta\omega\theta$$

If X is lognormal $\log X \sim N(\kappa_1, \kappa_2)$, then its entropy is $\mathbb{L}(X) = \frac{\kappa_2}{2}$.

B   Distributions

B.1   Summary of the moments

This is a table with the first four moments of the distributions used in this paper

Distribution		Mean	Variance	Skewness	Excess Kurtosis
		κ_1	κ_2	κ_3/σ^3	κ_4/κ_2^2
Distributions with two parameters					
Normal	$N\left(\mu,\sigma^2\right)$	μ	σ^2	0	0
Laplace	$L\left(\mu,b\right)$	μ	$2b^2$	0	3
Gumbel	$G\left(\mu,\beta\right)$	$\mu-\beta\tilde{\gamma}$	$\frac{\pi^2}{6}\beta^2$	$-\frac{12\sqrt{6}}{\pi^3}\zeta\left(3\right)$	$\frac{12}{5}$
Uniform	$U\left(a,b\right)$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	0	$-\frac{6}{5}$
Logistic	$LG\left(\mu,s\right)$	μ	$\frac{\pi^2}{3}s^2$	0	$\frac{6}{5}$
Skew-logistic	$SL\left(\alpha,\beta\right)$	$\psi\left(\alpha\right)-\psi\left(\beta\right)$	$\psi'\left(\alpha\right)+\psi'\left(\beta\right)$	$\frac{\psi''(\alpha)-\psi''(\beta)}{[\psi'(\alpha)+\psi'(\beta)]^{3/2}}$	$\frac{\psi'''\left(p\right)+\psi'''\left(q\right)}{[\psi'(\alpha)+\psi'(\beta)]^2}$
Distributions with more than two parameters					
Truncated normal	$TN\left(\mu,\sigma^2,a,b\right)$	See below			<0
Skew-normal	$SW\left(\xi,\omega^2,\delta\right)$	$\xi+\omega\delta\sqrt{\frac{2}{\pi}}$	$\omega^2\left(1-\frac{2\delta^2}{\pi}\right)$	$\frac{4-\pi}{2}\frac{\left(\delta\sqrt{\frac{2}{\pi}}\right)^3}{\left(1-\frac{2\delta^2}{\pi}\right)^{3/2}}$	$2\left(\pi-3\right)\frac{\left(\delta\sqrt{\frac{2}{\pi}}\right)^4}{\left(1-\frac{2\delta^2}{\pi}\right)^2}$
Bernoulli	$B\left(a,b,p\right)$	$pa+(1-p)b$	$(1-p)p\left(a-b\right)^2$	See below	
Poisson-N.	$PN\left(\lambda,\theta,\delta^2\right)$	$\lambda\theta$	$\lambda\left(\theta^2+\delta^2\right)$	See below	

Table 5: Properties of the different distributions

$\tilde{\gamma}$ is the Euler Mascheroni constant $\approx 0.57721....$

$\zeta\left(\cdot\right)$ denotes the Riemann Zeta function.

- Mean of truncated normal: $\mu+\frac{\phi(\frac{a-\mu}{\sigma})-\phi(\frac{b-\mu}{\sigma})}{\Phi(\frac{b-\mu}{\sigma})-\Phi(\frac{a-\mu}{\sigma})}\sigma$

- Variance of truncated normal: $\sigma^2\left[1+\frac{\frac{a-\mu}{\sigma}\phi(\frac{a-\mu}{\sigma})-\frac{b-\mu}{\sigma}\phi(\frac{b-\mu}{\sigma})}{\Phi(\frac{b-\mu}{\sigma})-\Phi(\frac{a-\mu}{\sigma})}-\left(\frac{\phi(\frac{a-\mu}{\sigma})-\phi(\frac{b-\mu}{\sigma})}{\Phi(\frac{b-\mu}{\sigma})-\Phi(\frac{a-\mu}{\sigma})}\right)^2\right]$

- Skewness of truncated normal: positive if $|a|\leq|b|$ and $a\leq\mu\leq b$. See Johnson, Kotz and Balakrishnan (2002).

- Skewness of Bernoulli: $\frac{(1-p)p(a-b)^3(1-2p)}{\left[(1-p)p(a-b)^2\right]^{3/2}}=\frac{(1-2p)}{[(1-p)p]^{1/2}}$

- Kurtosis of Bernoulli: $\frac{(1-p)p(a-b)^4\left(6p^2-6p+1\right)}{\left[(1-p)p(a-b)^2\right]^2}=\frac{\left(6p^2-6p+1\right)}{(1-p)p}$

- Skewness of poisson-normal: $\frac{\lambda\theta\left(\theta^2+3\delta^2\right)}{[\lambda(\theta^2+\delta^2)]^{3/2}}$

- Kurtosis of poisson-normal: $\frac{\lambda\left(\theta^4+6\delta^2\theta^2+3\delta^4\right)}{[\lambda(\theta^2+\delta^2)]^2}$

B.2 Pdf's of the distributions

This graph plots the different distributions used for calibration. All of them have mean of 0.08 and standard deviation of 0.2. The skew normal is plotted with maximal positive skewness.
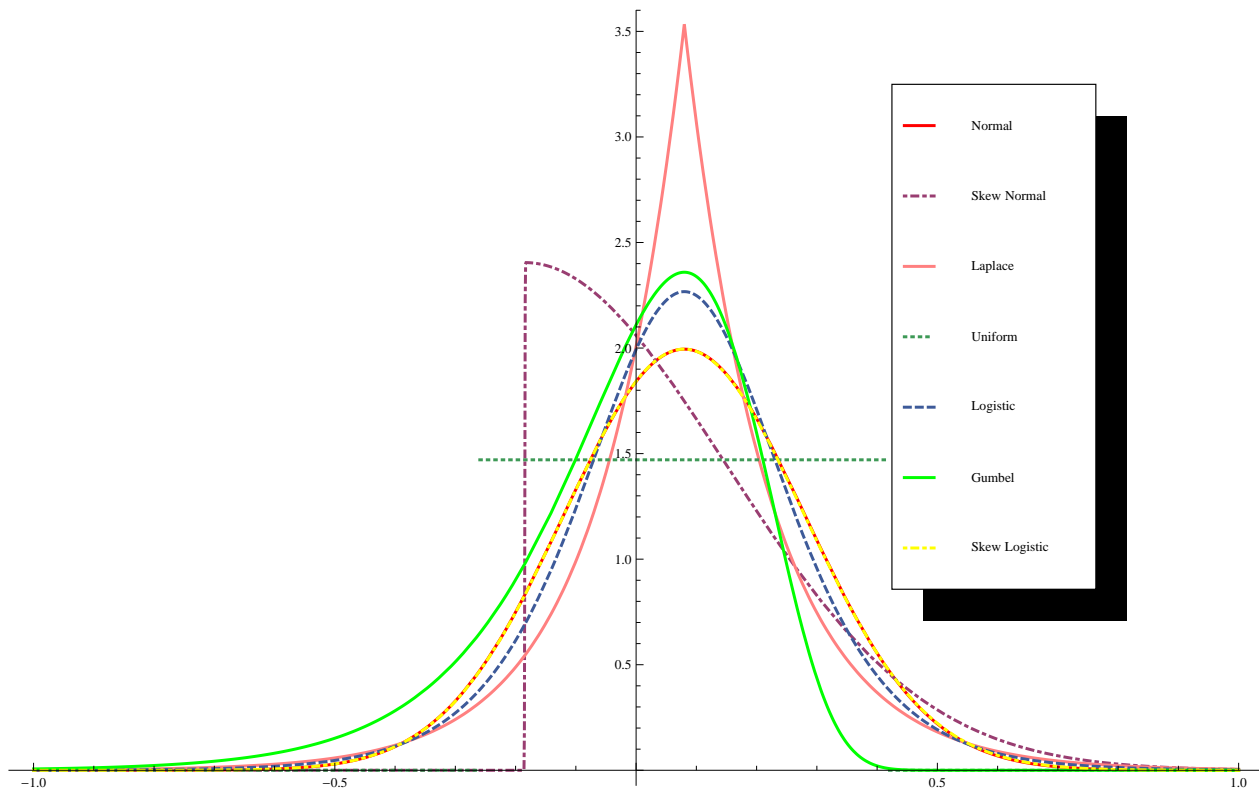

Figure 8: Comparing distributions

C Multivariate CARA case

The bivariate CGF of two random variables X and Y is defined by

$$CGF_{X,Y}(t) = CGF_{X,Y}(t_1, t_2) \equiv \log \mathbb{E} e^{t_1 X + t_2 Y}$$

And the bivariate cumulants are defined by

$$CGF_{X,Y}(t_1, t_2) = \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \kappa_{jk} \frac{\theta_1^j}{j!} \frac{\theta_2^k}{k!}$$

In an analogous way we can define a multivariate CGF

$$CGF_Z(t) \equiv \log \mathbb{E}e^{t \cdot Z}$$

with its respective multivariate cumulants (see for example Stuart and Ord (1994)).

I generalize here proposition 1 to the multivariate case. The problem to solve becomes

$$\max_x \mathbb{E}_t [U(W_{t+1})] \tag{27}$$

Subject to

$$W_{t+1} = W_t + x \cdot \tilde{\theta} \tag{28}$$

Proposition 3. *The objective function of an agent with CARA utility that solves* (27) *with* (28) *can be written as*

$$\min_x CGF_{\tilde{\theta}_1, \ldots, \tilde{\theta}_k} (-Ax_1, \ldots, -Ax_k)$$

Proof. The problem to solve is $\max_x \mathbb{E}_t \left[U\left(W_t + x \cdot \tilde{\theta}\right)\right]$. Given Lemma 1, let's proceed in the following way

$$
\begin{aligned}
\max_x \mathbb{E}_t \left[-e^{-AW_{t+1}}\right] &\iff \min_x \log \mathbb{E}_t \left[e^{-AW_{t+1}}\right] \\
&\iff \min_x \log \mathbb{E}_t \left[e^{-A\left(W_t + x \cdot \tilde{\theta}\right)}\right] \\
&\iff \min_x \log \mathbb{E}_t \left[e^{-Ax \cdot \tilde{\theta}}\right] \\
&\iff \min_x \log \mathbb{E}_t \left[e^{-Ax_1 \tilde{\theta}_1, \ldots, -Ax_k \tilde{\theta}_k}\right] \\
&\iff \min_x CGF_{\tilde{\theta}_1, \ldots, \tilde{\theta}_k} (-Ax_1, \ldots, -Ax_k) \\
&\iff \min_x CGF_{\tilde{\theta}} (-Ax)
\end{aligned}
$$

$\square$

References

Ait-Sahalia, Y., J. Cacho-Diaz, and TR Hurd. 2009. "Portfolio choice with jumps: A closed form solution." *The Annals of Applied Probability*, 19(2): 556–584.

Alvarez, F., and U.J. Jermann. 2005. "Using asset prices to measure the persistence of the marginal utility of wealth." *Econometrica*, 73(6): 1977–2016.

Athanasoulis, S.G., and R.J. Shiller. 2000. "The significance of the market portfolio." *Review of Financial Studies*, 13(2): 301.

Azzalini, A. 1985. "A class of distributions which includes the normal ones." *Scandinavian journal of statistics*, 12(2): 171–178.

Backus, D.K., M. Chernov, and I. Martin. 2009. "Disasters implied by equity index options." *NBER Working Paper.*

Barro, R.J. 2006. "Rare Disasters and Asset Markets in the Twentieth Century*." *The Quarterly Journal of Economics*, 121(3): 823–866.

Billingsley, P. 1995. *Probability and Measure, Wiley.* New York.

Brandt, M.W. 2009. "Portfolio choice problems." *Handbook of financial econometrics*, 1(1).

Brandt, M.W., A. Goyal, P. Santa-Clara, and J.R. Stroud. 2005. "A simulation approach to dynamic portfolio choice with an application to learning about return predictability." *Review of Financial Studies*, 18(3): 831.

Campbell, J.Y., and L.M. Viceira. 2002. *Strategic asset allocation: portfolio choice for long-term investors.* Oxford University Press, USA.

Chamberlain, G. 1983. "A characterization of the distributions that imply mean-variance utility functions." *Journal of Economic Theory*, 29(1): 185–201.

Demange, G., and G. Laroque. 1995. "Optimality of incomplete markets." *Journal of Economic Theory*, 65(1): 218–232.

Harvey, C.R., J. Liechty, M.W. Liechty, and P. Muller. 2004. "Portfolio selection with higher moments." *Working Paper, Duke University.*

Ingersoll, J.E. 1987. *Theory of financial decision making.* Rowman & Littlefield Pub Inc.

Jean, W.H. 1971. "The Extension of Portfolio Analysis to Three or More Parameters." *Journal of Financial and Quantitative Analysis*, 6(1): 505–515.

Johnson, N.L., S. Kotz, and N. Balakrishnan. 2002. *Continuous Multivariate Distributions, volume 1, Models and Applications.* New York: John Wiley & Sons,.

Kane, M. 1982. "A. Kane, Skewness preference and portfolio choice." *Journal of Financial and Quantitative Analysis*, 17(1).

Kogan, L., and R. Uppal. 2001. "Risk aversion and optimal portfolio policies in partial and general equilibrium economies." *NBER working paper.*

Kraus, A., and R.H. Litzenberger. 1976. "Skewness preference and the valuation of risk assets." *Journal of finance*, 1085–1100.

Lucas, R.E. 1987. *Models of business cycles.* Basil Blackwell New York.

Markowitz, H. 1952. "Portfolio Selection." *Journal of Finance*, 7(1): 77–91.

Martin, I. 2008. "Consumption-based asset pricing with higher cumulants." *manuscript.*

Merton, R.C. 1969. "Lifetime portfolio selection under uncertainty: The continuous-time case." *The review of Economics and Statistics*, 247–257.

O'Hara, M. 1998. *Market microstructure theory.* Wiley.

Ortobelli, S., I. Huber, ST Rachev, E. Schwartz, et al. 2003. "Portfolio choice theory with non-Gaussian distributed returns." *Handbook of Heavy Tailed Distributions in Finance*, 547–594.

Owen, J., and R. Rabinovitch. 1983. "On the class of elliptical distributions and their applications to the theory of portfolio choice." *Journal of Finance*, 38(3): 745–752.

Rotemberg, J.J., and M. Woodford. 1997. "An optimization-based econometric framework for the evaluation of monetary policy." *NBER macroeconomics annual*, 12: 297–346.

Samuelson, P.A. 1970. "The fundamental approximation theorem of portfolio analysis in terms of means, variances and higher moments." *The Review of Economic Studies*, 37(4): 537–542.

Samuelson, P.A. 1979. "Why we should not make mean log of wealth big though years to act are long." *Journal of Banking and Finance*, 3(2/3): 4.

Scott, R., and P. Horvath. 1980. "On the Direction of Preference for Moments of Higher Order than Variance." *Journal of Finance*, 35: 915–919.

Shleifer, A. 2000. *Inefficient markets: An introduction to behavioral finance.* Oxford University Press, USA.

Stuart, A., and J.K. Ord. 1994. "Kendall's advanced theory of statistics. Vol. 1: Distribution theory."

Theil, H. 1967. "Economics and information theory." *Amsterdam: North.*

Vives, X. 2008. *Information and learning in markets: the impact of market microstructure.* Princeton Univ Pr.