

# Index Option Returns and Generalized Entropy Bounds

Yan Liu\*

July 14, 2012

## Abstract

I develop a continuum of new nonparametric bounds. They stem from the solution of an optimization problem that is dual to the Hansen and Jagannathan (1991) approach, and are shown to complete the nonparametric bound universe that the literature has so far discovered. Through the lens of these bounds, I estimate rare event distributions from option market returns. Standard disaster models and their perturbations are shown not able to meet the bounds implied by simple static option trading strategies. Nonetheless, their abilities in magnifying pricing kernel dispersions through tail distortions do seem promising. My results suggest more sophisticated modeling of disaster models to reconcile with the index option data.

## 1 Introduction

Asset markets generate risk and return characteristics that continuously challenge our thinking. To rationalize market abnormalities, economists create models within a few generally accepted economic principles. These models are constantly scrutinized and possibly rejected with the advent of new empirical facts, and new models are again proposed to incorporate new findings. In this process, a few important diagnostic tools have been developed by the literature to restrict the behavior of a plausible model. For instance, under the basic no-arbitrage condition, Hansen and Jagannathan (HJ, 1991) construct bounds on the second moment of the stochastic discount factor for a given

---

\*I thank Ravi Bansal for his advice and encouragement. I appreciate comments from Campbell Harvey, Lukas Schmid, Hengjie Ai, John Graham, and other participants at the Duke Finance Brown Bag seminar.

asset menu. This nonparametric bound provides a simple way to summarize asset market information and helps screen candidate discount factors. Snow (1991) extends their work by showing how to bound higher moments of the pricing kernel. Bansal and Lehmann (BL, 1997) and Alvarez and Jermann (AJ, 2005) derive restrictions on the entropy, a separate metric on dispersion, based on the equity risk premium. These nonparametric bounds build on the fundamental no-arbitrage condition and provide unique lens through which we can diagnose asset pricing models, characterize asset market data in an efficient and economically interesting way, and potentially design new models that can meet the most stringent requirements set by market returns.

I contribute to this literature by providing a unifying theory on non-parametric bounds. Starting from the no-arbitrage condition alone, I show the existence of a continuum of bounds that restrict the  $\delta$ -th norm of the pricing kernel, with  $\delta \in (-\infty, 0) \cup (0, 1)$ . Next, I show these bounds can be naturally interpreted as the restrictions placed by an optimizing investor with a power utility function. In particular, I define an augmented return space with respect to a pricing kernel, and show that an agent's single period portfolio choice problem in this augmented return space puts a constraint on the corresponding moment of the pricing kernel. In a strict duality sense, I show my approach is complementary to the Hansen and Jagannathan approach in restraining pricing kernel moments. Then, motivated by the similarities between my bounds and the BL/AJ entropy bound, I complete my bound spectrum by showing that the entropy bound is a limiting case of my bounds. Finally, I depict the entire nonparametric bound universe and discuss its exhaustiveness.

To investigate bound informativeness and ease empirical implementations, I suggest using a new quantity which I call the *generalized entropy function*. It is a natural generalization of the entropy concept (see Stutzer (1996), Backus, Chernov and Zin (2011), Hansen and Sargent (2008)), and encodes all the information of a pricing kernel. More importantly, the system of nonparametric bounds can be conveniently brought in to restrict the generalized entropy function evaluated at different values. Through a cumulant expansion, similar to Martin (2008) and Backus, Chernov and Martin (BCM, 2011), I show how various moments of the pricing kernel are distilled into the generalized entropy function and more importantly, how weighted asset return moments provide information on it. An example featuring a finite state complete market economy is given to gain insights on the workings of bounds, and to highlight the powers of the newly developed bounds in teasing out tail information of the pricing kernel.

For the empirical applications, I attack the pseudo problem in a well-known class of models. The rare disaster models, as pioneered by Rietz (1988) and recently rejuvenated by a sequence of papers by Robert Barro and his coauthors (Barro (2006), Barro and Ursua (2008) and Barro et al. (2009)), use tail information to explain market abnormalities, in particular the equity risk premium. The empirical difficulty in this literature is the measurement issue of rare event distributions. Similar

to BCM, who gauge disaster models' implications against index option data along several metrics, I also use option data to infer tail information in the pricing kernel. Unlike BCM, I consider static portfolio strategies involving option returns and utilize nonparametric bounds to restrict the tail behavior. My approach is distinctively different and arguably advantageous in several aspects. First, no specific assumption is made on the linkage between consumption and the market index. As a result, my results are robust against model misspecifications. Second, instead of fitting a parametric model trying to summarize the option cross-section, I take the realized option returns as given and study their portfolio implications. This again alleviates the goodness-of-fit concern of empirical option pricing models. Finally, a formal statistical testing framework is developed to accommodate multiple asset classes. This not only enhances the scope of assets that we can simultaneously consider, but also generates statistical significance for different model parameterizations.

Turning to the empirical findings, I first document the superior performance of simple trading strategies involving deep out-of-the-money (OTM) put options from a nonparametric bound perspective. Bounds implied by OTM puts universally dominates bounds generated from the market index or risk-neutral straddles. This highlights the significance of the pricing of jump risks, and is exactly the kind of information one needs to impinge on models with tail risks. Next, I use market implied bounds to confront a standard rare disaster model. In particular, I mark up the permissible parameter region where mean estimates of all nonparametric bounds are simultaneously satisfied. Through this process, the discriminatory power of the newly developed bounds stands out and more importantly, we are able to better distinguish alternative tail specifications with asset market data alone, offering a way to partially circumvent the pseudo problem. Lastly, to take statistical uncertainty into account, I develop a formal testing framework which in principle can accommodate multiple assets and different types of bounds. Under this framework, I reject the benchmark disaster model and a few alternatives. Nonetheless, the model's ability in meeting asset market bounds does look impressive and I believe the idea of magnifying pricing kernel dispersion through tail distortions is promising. Taken as a whole, my results suggest more sophisticated specifications of disaster models, possibly through the time-dependency of disaster probabilities along the lines of Barro and Ursua (2008) and Watcher (2008).

The rest of the paper is organized as follows. Section 2 develops a unifying theory on nonparametric bounds. Section 3 studies bound informativeness by introducing a generalized entropy concept. Section 4 estimates nonparametric bounds constructed from option returns and uses them to test standard disaster models. Section 5 concludes.

## 2 A unifying theory on non-parametric bounds

Hansen and Jaganathan (1991), Snow (1991), Bansal and Lehmann (1997) and Alvarez and Jermann (2005) derive non-parametric bounds under the basic no-arbitrage condition. Depending on the forms of the non-linear transformations on the pricing kernel, strong no-arbitrage condition may be required to generate meaningful bounds. For instance, the entropy bound employs a logarithmic transformation of the pricing kernel. As a result, it will generate meaningful bounds only if the pricing kernel is strictly positive with probability one. To the contrary, the variance bound by Hansen and Jaganathan (1991) in general has no sign restrictions<sup>1</sup> since the pricing kernel is raised to the second power. To be specific about the asset pricing environment and facilitate discussion, I briefly introduce some notations that will be used throughout the paper.

Let  $\aleph$  be the collection of gross returns. Conceptually, it includes returns of all the tradable fundamental assets and portfolios of them. Under the assumption of no-arbitrage, there exists a pricing kernel  $M$  that prices all returns in  $\aleph$ , i.e.,

$$E[MR] = 1, \forall R \in \aleph. \quad (1)$$

Hansen and Jaganathan (1991, 1994) define  $Q^{++}$  and  $Q^+$  to be the set of strictly positive and nonnegative pricing kernels, respectively. Similarly, let  $\aleph^{++} = \{R : R \in \aleph \text{ and } R > 0 \text{ with probability one}\}$  and  $\aleph^+ = \{R : R \in \aleph \text{ and } R \geq 0 \text{ with probability one}\}$  be the set of strictly positive and nonnegative returns, respectively. For the same reason as in the entropy bound, we generally require  $M \in Q^{++}$  and  $R \in \aleph^{++}$  to produce informative bounds. Therefore, except for some discussions on weaker conditions towards the end, I impose these two constraints for the rest of this section. Notice that  $M \in Q^{++}$  is an implication of the strong no-arbitrage condition and  $R \in \aleph^{++}$  is also a weak condition for gross returns of primitive assets due to limited liability. However, a portfolio of assets with excessive short positions may generate negative returns with positive probability. The theories I develop later will not apply to these aggressive portfolio strategies.

### 2.1 A continuum of new bounds

Let  $M \in Q^{++}$  and  $R \in \aleph^{++}$  be the stochastic discount factor and an arbitrary return, respectively. Under the no-arbitrage condition, we have the following proposition:

**Proposition 1.** *:  $E(M^{\frac{1}{p}}) \leq [E(R^{-\frac{q}{p}})]^{\frac{1}{q}}$ , for any  $p > 1, q > 1, \frac{1}{p} + \frac{1}{q} = 1$ .*

---

<sup>1</sup>Hansen and Jaganathan (1991, 1994) consider generalizations for which the pricing kernel is restricted to be nonnegative or strictly positive.

*Proof.* The proof involves simple manipulations of the pricing equation and Hölder's inequality.

$$\begin{aligned}
E(M^{\frac{1}{p}}) &= E[(MR)^{\frac{1}{p}} R^{-\frac{1}{p}}] \\
&\leq [E([(MR)^{\frac{1}{p}}]^p)]^{\frac{1}{p}} \cdot [E(R^{-\frac{1}{p}})^q]^{\frac{1}{q}} \\
&= [E(MR)]^{\frac{1}{p}} [E(R^{-\frac{q}{p}})]^{\frac{1}{q}} \\
&= [E(R^{-\frac{q}{p}})]^{\frac{1}{q}}
\end{aligned}$$

The second line applies Hölder's inequality to  $(MR)^{\frac{1}{p}}$  and  $R^{-\frac{1}{p}}$ , and the last line uses the pricing condition  $E(MR) = 1$ .  $\square$

The above proof is distinctly different from the proof of HJ variance bound or more generally Snow's high-moment bounds. These bounds seek for restrictions on the average  $p$ -th power of the pricing kernel, with  $p$  greater than one. As a result, direct applications of Chachy-Schwarz inequality or Hölder's inequality on the no-arbitrage condition suffice<sup>2</sup>. For fractional powers on the pricing kernel, the trick is to create a power-transformed gross return to go with the pricing kernel and annihilate them both through the no-arbitrage condition by applying Hölder's inequality in the middle.

Proposition 1 gives a bound on the  $1/p$ -th moment of  $M$  by the  $-q/p$ -th moment of a return. As  $p$  runs from one to  $+\infty$ ,  $1/p$  covers every value in  $(0, 1)$ . At the same time,  $-q/p = 1/(1-p)$  goes from  $-\infty$  to zero. Therefore, we are exhausting negative moments of the return on the right hand side. However, due to the symmetric roles of  $M$  and  $R$  in the no-arbitrage condition, we can obtain a continuum of bounds on negative moments of the pricing kernel by reversing the roles of  $M$  and  $R$ .

**Corollary 1.** :  $E(M^\delta) \geq [E(R^{\frac{-\delta}{1-\delta}})]^{1-\delta}, \forall \delta \in (-\infty, 0)$ .

We can rewrite the bounds in Proposition 1 to make them conformable with the notations in Corollary 1:

$$E(M^\delta) \leq [E(R^{\frac{-\delta}{1-\delta}})]^{1-\delta}, \forall \delta \in (0, 1) \quad (2)$$

Combining Corollary 1 and equation (2), we find lower bounds for  $E(M^\delta)$  when  $\delta < 0$  and upper bounds when  $\delta \in (0, 1)$ . The change of direction at zero seems cumbersome and inevitable, but I will show later that it is simply a matter of scaling. Under appropriate transformations, the system of bounds will be smoothly connected at zero.

---

<sup>2</sup>Hölder's inequality works by bounding an expected product with the product of two power-transformed moments, with both powers strictly greater than one. See Casella and Berger (2001), Chap.4.

The bounds developed above apply for any  $R \in \aleph^{++}$ . To provide the tightest restrictions on the pricing kernel, we can search for the optimal return  $R$  corresponding to each power  $\delta$ , similar to Snow (1991) and Bansal and Lehmann (1997). Define  $\rho(\delta)$  as:

$$\rho(\delta) = \begin{cases} \sup_{R \in \aleph^{++}} [E(R^{\frac{-\delta}{1-\delta}})]^{1-\delta} & \text{if } \delta \in (-\infty, 0), \\ \inf_{R \in \aleph^{++}} [E(R^{\frac{-\delta}{1-\delta}})]^{1-\delta} & \text{if } \delta \in (0, 1) \end{cases} \quad (3)$$

then  $\rho(\delta)$  gives the sharpest lower (upper) bound on  $E(M^\delta)$  when  $\delta \in (-\infty, 0)$  ( $\delta \in (0, 1)$ ).

## 2.2 Interpreting bounds

What are the economic stories behind these bounds? In particular, what do the two sides of these inequalities measure? Do equality conditions reveal something fundamental about the economy? With these questions in mind, I offer to provide a utility-based interpretation.

Let us first introduce a risk-aversion index  $\gamma(\delta)$  defined as

$$\gamma(\delta) \equiv \frac{1}{1-\delta}, \text{ for } \delta \in (-\infty, 0) \cup (0, 1).$$

Note that  $\gamma(\delta)$  has a well-defined support as a risk-aversion coefficient:  $\gamma(\delta) \in (0, 1)$  if  $\delta \in (-\infty, 0)$  and  $\gamma(\delta) \in (1, +\infty)$  if  $\delta \in (0, 1)$ . Next, define the augmented return space as

$$\aleph^{**} = \{R : E(MR) = 1 \text{ and } R > 0 \text{ with probability one}\}$$

It is crucial to see the difference between  $\aleph^{++}$  and  $\aleph^{**}$ : the former restricts returns to be from tradable market assets while the latter contains all positive returns that satisfy the no-arbitrage condition. In other words,  $\aleph^{++}$  passively takes whatever the market has while the potentially much larger  $\aleph^{**}$  include what the market could have. The difference between  $\aleph^{++}$  and  $\aleph^{**}$  measures the degree of market completeness.

Endowed with this augmented return space  $\aleph^{**}$ , we seek to find the optimal portfolio choice for an agent with unit endowment and risk-aversion  $\gamma(\delta)$ . This optimization problem can be written as:

$$U_\delta(M) = \sup_{R \in \aleph^{**}} E\left[\frac{R^{1-\gamma(\delta)}}{1-\gamma(\delta)}\right] \quad (4)$$

The maximized utility  $U_\delta(M)$  depends on the discount factor  $M$ , whose information is implicitly given in  $\aleph^{**}$ . The following proposition gives the solution to this maximization problem.

**Proposition 2.** *The solution to the maximization problem in (4) is given by:*

$$U_\delta(M) = \frac{[E(M^{\frac{\gamma(\delta)-1}{\gamma(\delta)})]^{\gamma(\delta)}}{1 - \gamma(\delta)} = \frac{[E(M^\delta)]^{\frac{1}{1-\delta}}}{1 - \gamma(\delta)} \quad (5)$$

$$\tilde{R}_\delta(M) = M^{-\frac{1}{\gamma(\delta)}} / E(M^{\frac{\gamma(\delta)-1}{\gamma(\delta)}}) \quad (6)$$

*Proof.* The appendix contains a detailed proof. The inequalities in Proposition 1 and Corollary 1 establish the finiteness of the objective function, making the optimization problem well-defined. The proof then proceeds in two steps. First, the optimal portfolio return  $\tilde{R}_\delta$  is solved as a function of the Lagrange multiplier associated with  $E(MR) = 1$ . Second, the Lagrange multiplier itself is solved through the no-arbitrage equation.  $\square$

Now the economic meanings of the bounds we discussed before stand out. To ease interpretation, we can transform the bounds shown in Corollary 1 and equation (2) to

$$\frac{[E(M^\delta)]^{\frac{1}{1-\delta}}}{1 - \gamma(\delta)} \geq \frac{E[R^{1-\gamma(\delta)}]}{1 - \gamma(\delta)}, \forall R \in \mathfrak{N}^{++}. \quad (7)$$

By Proposition 2, the quantity on the left hand side is the maximized utility over the augmented return space  $\mathfrak{N}^{**}$  for an agent with a risk-aversion coefficient of  $\gamma(\delta)$ . It is the highest achievable utility if the market is complete in the sense that  $\mathfrak{N}^{**} = \mathfrak{N}^{++}$  or, as a weaker requirement, the optimal choice  $\tilde{R}_\delta(M)$  is actually tradable. For a given risk aversion  $\gamma(\delta)$ , the log of  $\tilde{R}_\delta(M)$  loads negatively on the log pricing kernel, i.e.

$$\log \tilde{R}_\delta(M) = A_\delta(M) - \frac{1}{\gamma(\delta)} \log M \quad (8)$$

where  $A_\delta(M)$  is a scaling constant. Given the stylized fact that market returns are generally counter-cyclical, these bounds are easier to satisfy and thus potentially more informative than HJ or Snow's high-moment bounds.

From a global perspective, although the marginal investor determines the marginal rate of substitution, we see that investors with different levels of risk aversion all have a say in the behavior of the pricing kernel. Their optimal portfolio choices automatically place a sequence of thresholds that the discount factor has to overcome. As the power  $\delta$  goes through its admissible region, we are essentially running through the support of the risk-aversion coefficient of power-utility agents. This interpretation is in spirit similar to Bansal and Lehmann (1997)'s interpretation of their growth-optimal portfolio, albeit we are able to significantly generalize their argument.

From a methodological perspective, it is also interesting to compare the approach I have taken to interpret my bounds and the Hansen-Jaganathan type of projection methods in constructing their mean-variance bound(See Hansen and Jaganathan

(1991), Gallant, Hansen and Tauchen (1990) and Bekaert and Liu (2003)). HJ bounds are constructed by projecting the pricing kernel onto the space of available asset pay-offs. The  $L_2$ -norm of the projected pricing kernel has the minimal standard deviation across all valid pricing kernels. I start from a candidate pricing kernel and ask what an optimizing agent will do in an ideal world where all admissible returns are tradable. Then, by limiting the asset space to available returns, the agent's real-world objective function dictates a lower bar that the starting candidate pricing kernel has to satisfy. As a matter of fact, in a strict duality sense, these two approaches are complementary and thus compatible with each others. I rigorously define the duality concept and prove it in the appendix. On a practical level, HJ's approach is more transparent when certain moment of the pricing kernel (variance, sharpe ratio, etc.) is the focus and my approach will be more intuitive if a return moment (e.g., a CRRA investor's objective function) is the interest.

### 2.3 Characterizing the non-parametric bound universe

Thus far, I have established bounds for moments of various orders of the pricing kernel, with zero as the only undefined case. Additionally, I extend the log-utility interpretation of the entropy bound to general power utilities. These results prompt one to wonder if the entropy bound is the one that fills the hole of my bound spectrum. Indeed, it is. The following proposition formally establishes this.

**Proposition 3.** *The bounds given in Proposition 1 and Corollary 1 both imply the entropy bound:  $E(\log(M)) \leq -E(\log(R))$ .*

*Proof.* The proof applies the same intuition as how power utility converges to log-utility. I will only show how bounds in Proposition 1 imply the entropy bound. Essentially the same proof can be done for bounds from Corollary 1. I start by scaling the bounds in equation (2):

$$\frac{E(M^\delta) - 1}{\delta} \leq \frac{[E(R^{\frac{-\delta}{1-\delta}})]^{1-\delta} - 1}{\delta}$$

This is true because  $\delta > 0$ . Taking limits as  $\delta \downarrow 0$  and under regularity conditions<sup>3</sup>, the left hand side is readily seen to converge to  $E(\log(M))$  using L'Hôspital's rule. An careful application of the rule to the right hand side will deliver  $-E(\log(R))$  as the limit.  $\square$

We are now ready to summarize the non-parametric bound universe that the literature has discovered. Figure 1 shows a diagram of this bound universe. At power equals

---

<sup>3</sup>We need conditions on moments of  $M$  and  $R$  to be able to exchange limits and expectation. Dominated convergence will suffice. See Davidson (1994) Part IV for specific conditions.



one, the expected marginal rate of substitution is bounded within  $[\frac{1}{\min R}, \frac{1}{\max R}]$  for a generic  $R \in \mathfrak{N}^{++}$ . This seemingly informative bound becomes redundant in the presence of a risk-free rate  $R_f$ , since then  $E(M) = 1/R_f$ . Starting from one and going right, one will encounter the spectrum of Snow's high-moment bounds and HJ bound is sitting right at  $s = 2$ . Going left, one sees the continuum of bounds I just developed and BL/AJ entropy bound fills the hole at  $s = 0$ . It is intriguing to see the symmetric structure of these bounds around  $s = 1$  and the order of discovery of these bounds by the literature.

Lastly, we ask if the bound system is complete. This is more than a technical question because we do not want to leave out information on the pricing kernel that can be provided by the return space. In particular, given the existence of these one-sided inequalities for essentially any moment of the pricing kernel, one may wonder if other bounds, possibly with opposite directions of inequalities, can further enrich the bound universe. After all, an interval bound certainly looks more appealing than a one-sided bound since both sides are well restricted. The following proposition eliminates such possibilities and indirectly shows the exhaustiveness of the above bound universe.

**Proposition 4.** *For a given power  $s$  and the corresponding upper (lower) bound on  $E(M^s)$ , the lower (upper) side of  $E(M^s)$  is generally unbounded. Hence, the non-parametric bound system is exhaustive.*

*Proof.* The idea is to construct a sequence of pricing kernels that can simultaneously price a return but has unbounded limit for a given moment. I leave this proof to the appendix.  $\square$

## 2.4 Discussions and extensions

The new continuum of bounds can be extended along several dimensions. First, they can be adjusted to incorporate payoffs that are zero with positive probability. A hold-to-maturity option return is a good example. The idea is to consider the truncated return that is strictly positive. For a return  $R$ , let the truncated return be

$$R^+ = \begin{cases} R & \text{if } R > 0, \\ 0 & \text{if } R = 0. \end{cases}$$

Notice that the no-arbitrage condition holds for the truncated return, i.e.,  $E(MR^+) = E(MR) = 1$ . Consequently, bounds developed before can be directly applied to this truncated return  $R^+$  instead.

Second, it can be adapted to study the dynamic behavior in the marginal rate of substitution. Let  $M_{t,t+n} = M_{t+1}M_{t+2}\dots M_{t+n}$  be the time aggregated pricing

kernel and  $R_{t,t+n} = R_{t+1}R_{t+2}\dots R_{t+n}$  be a generic multi-period return. Then long-horizon asset market returns provide bounds on the unconditional moments of the time aggregated pricing kernel. These unconditional moments of the multi-period pricing kernel reveal the dynamic dependency of the single period pricing kernels. Different power transforms shed lights on different forms of dynamic dependency. For instance, one natural way to scale an  $n$ -period pricing kernel is to take a fractional power of  $1/m$  on the time aggregated pricing kernel. A bound on thus scaled kernel is given by

$$E(M_{t,t+n}^{\frac{1}{m}}) \leq [E(R_{t,t+n}^{\frac{-1}{m-1}})]^{\frac{m}{m-1}}, m \geq 2$$

When  $m \rightarrow \infty$ , properly scaled version of the above bound converges to the multi-period entropy bound. Backus, Chernov and Zin (2011) use the multi-period entropy bound to study time-dependency in discount rates. Setting  $m$  at  $n$ , the left hand side is  $E[\exp(\frac{1}{n} \sum_{j=1}^n \log M_{t+j})]$  so the average of the log pricing kernel is revealed. By decomposing the pricing kernel into a permanent and a transitory component (see Alvarez and Jermann (2005)), such a scaling leaves the expectation of the permanent component intact while allows the transitory component to die out at a rate of  $n$ . Bounds give us information on how fast they have to die out.

Lastly, conditioning information can be incorporated to sharpen bounds on unconditional moments (See Bekaert and Liu (2004), Gallant, Hansen and Tauchen (1990) and Ferson and Siegel (2003)). Notice that simply adding instruments in the conditional Euler equation (See Hansen and Jagannathan (1990)) is different from the use of a return series that is generated from a dynamic trading strategy. The utility-based interpretation of the new bounds naturally favors the use of the latter approach. As shown in Ferson and Siegel (2003), the use of returns from dynamic strategies significantly sharpens HJ bounds. We would expect it to be important for the new bounds as well, especially considering some of the well-documented predictive evidence for the first and second moment of returns by the literature.

### 3 Bound informativeness

Having established an exhaustive non-parametric bound system, one naturally wonders what unique insight a bound on a certain moment of the SDF can provide us. After all, an exploration of a continuum of powers is physically impossible and we would like to select a few bounds that are both representative and informative. With this goal in mind, I perform a dissection of the bound universe. Analogous to the situation of a doctor performing a surgery, I need a “surgery knife” to decompose the bound system. Therefore, I first develop a useful tool in studying bounds. I define a quantity that is a natural generalization of the entropy concept popularized by Bansal and Lehmann (1997), Alvarez and Jermann (2005) and Backus, Chernov

and Zin (2011) in the recent asset pricing literature. I show that it is both economically meaningful and analytically convenient. With this tool in hand, I employ the cumulant-expansion technique (See Backus, Chernov and Martin (2011) and Martin (2008)) to inspect what both sides of a bound tell us. Lastly, a concrete example is given to gain deeper understanding on the workings of a bound.

### 3.1 A useful quantity

The recent asset pricing literature proposes a convenient quantity to study the link between a pricing kernel and asset returns (See Alvarez and Jermann (2005), Bansal and Lehmann (1997), Backus, Chernov and Zin (2011) and Martin (2011))<sup>4</sup>. The entropy of the marginal rate of substitution is used to measure pricing kernel dispersion and it is bounded below by the expected differences in returns

$$L(M) \equiv \log E(M) - E(\log M) \geq E(\log R) - \log(R_f) \quad (9)$$

where  $R \in \mathfrak{N}^{++}$  is an arbitrary return and  $R_f = 1/E(M)$  is the gross risk-free rate if we assume one exists. Researchers rely on the entropy bound to gauge the amount of dispersion that an asset pricing model must generate. However, in lieu of the continuum of bounds that are relatives of the entropy bound, we shall expect to gain additional insights by utilizing other forms of bounds. I propose a quantity that is a natural generalization of the original entropy concept. The continuum of bounds developed in the previous section can then be brought in to provide further constraints on the pricing kernel. In essence, I am normalizing the system of bounds in reference to the entropy bound.

The Generalized Entropy Function (GEF) of a positive pricing kernel is defined as

$$GEF(s; M) \equiv \log E(M) - \frac{1}{s} \log E(M^s) \quad (10)$$

for any real-valued  $s$ . It is an extension of the original entropy because its limit at zero is exactly the entropy, i.e.,

$$\lim_{s \rightarrow 0} GEF(s; M) = L(M)$$

Given the finiteness of all moments,  $GEF(s; M)$  is generally an everywhere continuous function on the real line. Moreover, many nice properties of entropy are maintained for the GEF. For instance, GEF will be zero at any power  $s$  if and only if  $M$  is a constant. Similar to entropy, it is scale-invariant, i.e.,  $GEF(ws; M) = GEF(s; M)$  for any constant  $w$ . Hence, the GEF transformation leaves the pricing kernel numeraire

---

<sup>4</sup>Notably, entropy is gaining popularity in many fields of economics and finance. See Stutzer (1996), Hansen and Sargent (2008), Ghosh, Julliard and Taylor (2011) and Van Nieuwerburgh and Veldkamp (2010).

invariant. This is an empirically appealing property because we do not need to worry about the adjustment between a nominal and a real pricing kernel. Additionally, GEF is pivotal around  $(1, 0)$  in the sense that every GEF has to pass  $(1, 0)$  on the two dimensional plane. This gives us a fixation point to think about the GEF's corresponding to different pricing kernels. Finally, in the familiar lognormal case,  $GEF(s; M) = (1 - s)\sigma_M^2/2$  where  $\sigma_M^2$  is the variance of the log pricing kernel.

The bound universe in the previous section can be brought in to restrict  $GEF(s; M)$ . The implied restrictions from my bounds can be shown as

$$GEF(s; M) \geq \frac{s-1}{s} \log E(R^{\frac{s}{s-1}}) - \log(R_f), \forall s \in (-\infty, 1) \quad (11)$$

Notice how the two types of bounds in Corollary 1 and equation (2) nicely line up with each other in terms of the direction of inequalities. The undesirable flip in direction at zero disappears once we introduce the generalized entropy function. In this sense, GEF is a more appropriate apparatus in studying  $E(M^s)$  when  $s < 1$ . When  $s > 1$ , Snow's continuum of high-moment bounds imply:

$$GEF(s; M) \leq \frac{s-1}{s} \log E(R^{\frac{s}{s-1}}) - \log(R_f), \forall s \in (1, +\infty) \quad (12)$$

Figure 2 displays the plot a generic GEF with asset market bounds.

Of course, sharper restrictions can be found by explicitly searching for the optimal bounds, as in equation (3). Considering the practical convenience offered by the generalized entropy function, for the rest of the paper I will focus on bounds given in the forms of (11) and (12) unless otherwise specified. I will refer to the system of bounds given in (11) as generalized entropy bounds or, with a slight abuse of terminology, simply entropy bounds. The bounds in (12) will be termed high-moment bounds.

### 3.2 A CGF expansion

The Cumulant-generating Function (CGF) is another recently developed tool in studying higher moments in the pricing kernel (See Backus, Chernov and Martin (2011) and Martin (2008)). By Taylor expanding the log expected pricing kernel into power series, it becomes transparent to see how various higher moments contribute to the overall entropy. Backus, Chernov and Martin (2011) use entropy as the sole dispersion measure and study how a standard disaster model and an empirical option pricing model imply different moment statistics of the pricing kernel. Assuming representative agent and i.i.d. consumption growth rates, Martin (2008) link fundamentals to moments of the consumption growth and perform a calibration exercise. I contribute to this literature by showing that asset market returns provide valuable

information on the entire CGF and not just at zero, which corresponds the original entropy. This significantly strengthens the link between asset pricing models and market returns, and can potentially help us better distinguish candidate stochastic discount factors.

I start by performing a Taylor expansion of our newly defined  $GEF(s; M)$ . This amounts to Taylor expanding  $E(M^s) = E(e^{s \log M})$  around  $s = 0$ .

$$\begin{aligned}
GEF(s; M) &= \sum_{i=1}^{\infty} \frac{\kappa_i(\log M)}{i!} - \frac{1}{s} \sum_{i=1}^{\infty} \frac{\kappa_i(\log M)}{i!} s^i \\
&= \sum_{i=2}^{\infty} \frac{\kappa_i(\log M)}{i!} (1 - s^{i-1}) \\
&= \frac{\kappa_2(\log M)}{2!} (1 - s) + \frac{\kappa_3(\log M)}{3!} (1 - s^2) \\
&\quad + \frac{\kappa_4(\log M)}{4!} (1 - s^3) + \frac{\kappa_5(\log M)}{5!} (1 - s^4) \dots \tag{13}
\end{aligned}$$

The first two lines Taylor expand the two components in  $GEF(s; M)$  and group the relevant terms. The last line explicitly writes out the first few terms in this expansion. Here  $\kappa_i(\log M)$  denotes the  $i$ -th ‘‘cumulant’’ of the log discount factor and is defined as the  $i$ -th derivative of  $\log E(e^{s \log M})$  at  $s = 0$ . Cumulants are closely related to moments:  $\kappa_1(\log M)$  and  $\kappa_2(\log M)$  are the mean and variance of  $\log M$ , respectively and  $\kappa_3(\log M)$  and  $\kappa_4(\log M)$  are related to the usual skewness ( $\nu_1$ ) and excess kurtosis ( $\nu_2$ ) through:  $\nu_1 = \kappa_3(\log M)/[\kappa_2(\log M)]^{\frac{3}{2}}$ ,  $\nu_2 = \kappa_4(\log M)/[\kappa_2(\log M)]^2$  (See Backus, Chernov and Martin (2011) for detailed information on individual cumulants).

The expansion reveals how cumulants are weighted by polynomials of  $s$  in  $GEF(s; M)$ . In particular, the  $i$ -th scaled cumulant  $\kappa_i(\log M)/i!$  is multiplied by  $(1 - s^{i-1})$ . By varying the value of the argument  $s$ ,  $GEF(s; M)$  puts different weights on various moments. In this way,  $GEF(s; M)$  conveys information about all the moments of the pricing kernel. In particular, when evaluated at  $s = 0$ ,  $GEF(s; M)$  equals the original entropy, which is an overall sum of  $\kappa_i(\log M)/i!$ . I argue that  $GEF(s; M)$  is especially useful in teasing out the tail information in the pricing kernel. Take, for example, a standard disaster model along the lines of Barro (2006, 2009). In such a model, large drops in consumption in disastrous states generate a huge amount of negativity in skewness and all the other odd moments. Consequently, the marginal rate of substitution, which loads negatively on consumption growth, will display excess positivity for all odd moments. On the other hand, even moments will mechanically increase as well in the presence of outliers in state prices. This creates an identification problem for the ultimate source of dispersion. Backus, Chernov and Marin (2011) argue that odd cumulants in the original entropy expansion reflect the inherent asymmetry in jumps. However, given that all moments are equally weighted at  $s = 0$ , it would be very difficult to see the differential effect of negative jumps on odd and even mo-

ments. I suggest taking large negative  $s$  values to inflate this wedge. Large negative  $s$  makes the multipliers associated with even moments positive and the ones with odd moments negative. Thus, a “net” jump effect is singled out by taking the difference of odd and even moments. In fact,  $s < -1$  is the only region for the net jump effect to appear. This highlights the importance of the extensions on powers for the bound system I developed in the previous section.

Similar expansions can be applied to returns on the right hand side of the bounds. This step is important in that it gives us some ex-ante guidance on the selection of the most informative asset market returns. I cumulant-expand the right hand side of equation (11) as

$$\begin{aligned} GEF(s; M) &\geq \frac{s-1}{s} \sum_{i=1}^{\infty} \frac{\kappa_i(\log R)}{i!} \left(\frac{s}{s-1}\right)^i - \log R_f \\ &= [E(\log R) - \log R_f] + \sum_{i=2}^{\infty} \frac{\kappa_i(\log R)}{i!} \left(\frac{s}{s-1}\right)^{i-1} \end{aligned} \quad (14)$$

where  $\kappa_i(\log R)$ 's are the return cumulants. We see that in addition to the continuously-compounded equity risk premium, an extra term  $\sum_{i=2}^{\infty} \frac{\kappa_i(\log R)}{i!} \left(\frac{s}{s-1}\right)^i$  comes out of the expansion. For large negative  $s$  values,  $\frac{s}{s-1}$  is close to one so the first few higher-order cumulants will enter significantly into the right hand side. This means that to search for the tightest lower bounds from returns, characteristics such as excess (positive) skewness and kurtosis are probably as important as high risk compensation. Option strategy returns naturally fit these descriptions (See Coval and Shumway (2001) and Broadie, Chernov and Johannes (2008)). In the empirical section of the paper, I follow this intuition to explore the restrictions that option strategy returns put on the discount factor.

### 3.3 A complete market economy example

To better illustrate the workings of entropy bounds, I consider a complete market economy with finite states. Under this setup, limiting distributions for the discount factor are easily derived and the asset market has a transparent structure. Through this process, our goal is to gain intuition on: 1. How the generalized entropy measure reduces to a tail measure when  $s$  is sufficiently negative; 2. How exactly do security returns bound the pricing kernel?

Consider a one-period economy with only two states of nature. The state prices are given by

$$M = \begin{cases} M_1, & \text{with probability } p_1 \\ M_2, & \text{with probability } p_2 = 1 - p_1 \end{cases} \quad (15)$$

Imagine that state one is a bad state and state two is a normal state so that  $M_1 > M_2$ . In thinking of a disaster model, we will have  $M_1 \gg M_2$ . In this economy,  $GEF(s; M)$  has the following limiting behavior:

$$\begin{aligned} GEF(s; M) &= \log(EM) - \frac{1}{s} \log(EM^s) \\ &\rightarrow \log\left(\underbrace{p_1 \frac{M_1}{M_2} + p_2}_{\bar{W}}\right) \end{aligned}$$

when  $s \rightarrow -\infty$ .

What is  $\bar{W}$ ? In the presence of a risk-free rate  $R_f$ ,  $\bar{W}$  can be expressed as  $\frac{1}{R_f} \cdot \frac{1}{M_2}$ . This is the expected wealth for a risk-neutral optimizing agent with an endowment of one unit of riskless bond. To see this, notice that the expected returns for the two Arrow-Debreu securities are  $1/M_1$  and  $1/M_2$ , respectively. A risk-neutral agent only cares about expected returns. Hence, she will shy away from the relatively expensive insurance asset and have a concentrated position on the second Arrow-Debreu security. Her expected end-of-period wealth is thus equal to the beginning-of-period wealth ( $1/R_f$ ) times the expected return of the second elementary security ( $1/M_2$ ). This interpretation of  $\bar{W}$  also reminds us of the utility-based interpretation of entropy bounds from the previous section.

Intuitively, a risk-neutral investor's portfolio choice contains superior information about tail events because her incentive in "selling" the insurance asset is the strongest. To see how the tail event probability is reflected in  $\bar{W}$ , note that

$$\bar{W} = \log\left[1 + p_1\left(\frac{M_1}{M_2} - 1\right)\right] \approx p_1\left(\frac{M_1}{M_2} - 1\right)$$

under mild conditions<sup>5</sup>. Therefore, the rare event probability is amplified by the multiplier  $(\frac{M_1}{M_2} - 1)$ . The more frequently tail events happen, or the larger the state price is for a disastrous state, the more profit that a risk-neutral agent can exploit. In this sense, the tail event information is partially identified through the risk neutral agent's optimization behavior.

On the other hand, the right hand side of equation (11) reduces to  $\log(\frac{1}{R_f} \cdot E(R))$  when  $s \rightarrow -\infty$ . Hence, the generalized entropy bounds simply say that in expectation no return can exceed the return of the second Arrow-Debreu security. This is trivially true given the structure of the economy. What is not so trivial, even in this two-state economy, is how entropy bounds internally extract information from a discount rate and provide economically sensible links to market securities.

---

<sup>5</sup>In a standard disaster model, where the disaster state probability is in the range of 1%-2% and the state price ratio is within 10, the approximation is good.

The above example can be easily extended to an  $N$ -state case, although the marginal gain in intuition is limited. In essence, the generalized entropy bounds feed on the idea that a power transformation of the discount rate can be viewed as the objective function for a certain type of investors. By taking large negative powers, the rare event information stands out and this information is calibrated against the maximized utility of a nearly risk-neutral agent.

## 4 Option market bounds and rare disaster models

Tail information, long recognized as a potential source to generate economic risk premiums (See Rietz (1988)), has recently been elevated to quantitatively explain asset market abnormalities. Barro (2006), Barro and Ursua (2008) and Barro, Nakamura, Steinsson and Ursua (2009) extrapolate the tail distribution for the US consumption growth process by looking at international macroeconomic data. Based on the exchange economy and representative agent framework, they argue that thus calibrated rare event distribution can explain key moments of the US asset returns, in particular the equity risk premium. Gabaix (2009), Wachter (2009) and Gourio (2008) extend the basic disaster model intuition to account for other salient features of the asset markets.

At the heart of the rare disaster literature is the so-called pseudo problem: given the rare occurrence of disasters, one cannot measure their distributions accurately from a relatively short univariate time series. As a remedy, researchers pool data from other sources to avoid the inherent small sample problem. An alternative approach is to use asset market returns to infer the tail information in the pricing kernel. Since option prices are informative about the investors' ex-ante valuation of extreme event risks, they can be a useful source of information. Backus, Chernov and Martin (2011) use equity index options to deduce the distribution of consumption growth. I also consider index option data but employ a different approach to study its implications on the pricing kernel. I use the nonparametric bounds developed above to restrict the behavior of a pricing kernel. By doing this, tail information of the kernel is distilled into various moment inequalities. My approach has several advantages. First, it is based on the basic no-arbitrage condition alone and thus free from any misspecification on the linkage between the pricing kernel and asset returns. Second, no parametric model is needed to fit the cross-section of option prices. Instead, individual option trading strategies are estimated and fed into the nonparametric bounds. Lastly, a formal statistical testing framework is developed. It in principle can accommodate multiple testing asset classes with different power transforms. I show the discriminatory power of the generalized entropy bounds at negative power values.



## 4.1 Data description

For the empirical analysis, I use monthly data on S&P 500 index, the associated index options and the risk-free rate. The riskfree rate is from Kenneth French's on-line data library. The longer sample for index returns run from July 1926 to December 2011. The shorter sample, which coincides with the span of the option data available from OptionMetrics, is from January 1996 to December 2011, giving us 192 months of data. All nominal returns are converted to ex post real returns using the inflation rate based on CPI.

I collect European style S&P 500 index options for the period from January 1996 to December 2011 from the OptionMetrics database. The data set contains daily settlement prices for options with various strike prices and maturities, as well as liquidity measures such as open interests and trading volumes. It also includes dividend yield for the market index and interpolated zero coupon yields. They will be necessary for us to construct option strategies later on. To mitigate microstructure issues, I drop option data with average bid-ask prices less than one eighth of a dollar, with open interests less than 100 contracts or with zero trading volumes. Finally, I use put-call parity relationships to filter out data that obviously violate the no-arbitrage condition.

To construct equally-spaced monthly returns from option prices, I follow a procedure similar to Coval and Shumway (2001), Buraschi and Jackwerth (2001) and Driessen and Maenhout (2005). First, options with strike-to-spot ratio closest to 92%, 96% and 100% are targeted on the first trading day of each month. Next, these option contracts are followed and identified until the beginning of the subsequent month and the monthly holding period returns are calculated. In this process, option contracts that expire on the third week of the same month have to be excluded. Due to liquidity concerns, I focus on short-maturity options with around seven weeks to maturity at the buying date and about two weeks to maturity at the selling date. These options have large trading volumes and are arguably less affected by liquidity problems (See Bondarenko (2003)).

The equally-spaced return series are convenient to handle empirically, especially when we confront asset pricing models with these returns. This is because most discrete time models specify a fixed sample frequency. Additionally, as argued by Driessen and Maenhout (2005), thus constructed returns are more sensitive to changes in jump or volatility risks than hold-to-maturity returns. Such return feature is important for this paper since risk-sensitive returns can potentially provide the most informative bounds on the pricing kernel. Finally, from a more technical point of view, our bound theory requires returns to have a positive support. However, hold-to-maturity returns are at times extremely high and this creates trouble in constructing a short strategy that always generate positive returns.

Instead of using raw option returns data, I focus on returns from a few well-known derivative strategies. There are two main reasons for this: 1. These economically meaningful strategies offer clear interpretations of the risks (jump risks or volatility risks or both) that are being traded. This in turn can ease the way we interpret the bounds on the pricing kernel; 2. The recent literature on option pricing anomalies have major focus on these trading strategies (See Bondarenko (2003), Coval and Shuway (2001) and Driessen and Maenhout (2005)). To be consistent and comparable with the existing literature, I choose to adopt this tradition.

In particular, I take the following two strategies as our benchmark strategies:

- An out-of-the-money (OTM) put option with 96% moneyness;
- An at-the-money(ATM) straddle.

A deep OTM put is a hedge against market crashes and much less so against volatility movements. Its price is therefore only sensitive to market jump risks. On the other hand, an ATM market-neutral straddle makes profits when either the market volatility is high or when market crashes, so it is exposed to both volatility and jump risks. These two option strategies are among the most commonly traded strategies by market practitioners and have been extensively studied by the recent option pricing literature (See Coval and Shumway (2001), Jackwerth (2000), Bondarenko (2003)). These facts, together with the clean economic interpretation of their risk exposures, make them our ideal choices of benchmark strategies.

In addition to the benchmark strategies, I further consider two types of “crash-neutral” variants of them. They are simply the two original options mixed with an offsetting short position in the 92%-OTM put option<sup>6</sup>. By doing this, large returns when market crashes are capped off for long positions in the benchmark strategies and symmetrically, short positions are protected against large downward movements in market returns<sup>7</sup>. It would be interesting to see if these alternative strategies can provide any additional information beyond what the benchmark strategies can.

Table 1 shows the summary statistics for the returns from various derivative strategies, as well as from the market index. Consistent with the literature, these option strategies generate large negative average returns and Sharpe Ratios. The flip side of this would be the potential profits generated by short positions in these strategies.

---

<sup>6</sup>For details on the construction of the crash-neutral strategies, see Coval and Shumway (2001) and Jackwerth (2000)

<sup>7</sup>This is only approximately so because the beginning-of-period 92%-OTM put and 96%-OTM put may have different maturities. In fact, the deeper 92%-OTM put may have a higher price than the 96%-OTM put simply because the former’s maturity is significantly longer than the latter. This creates trouble in the usual interpretation of crash-neutral strategies: for instance, a short leg in 92%-OTM put becomes a long leg. Therefore, I also create robust-crash-neutral put and straddles which essentially delete these abnormal dates. See the description under Table 1 for details.

Moreover, the return distribution are highly non-normally distributed, as reflected by the large magnitude of skewness and kurtosis. These higher-moment characteristics on index option strategy returns will be helpful later for us to form informative bounds on the pricing kernel. Nonetheless, in spite of the magnitude, the mean returns for my sample are notably smaller and about half of the size of those reported by Coval and Shumway (2001), Broadie, Chernov and Johannes (2007) and Bondarenko (2003). This is mainly driven by the unstableness in mean returns for these derivative strategies. The above mentioned papers mainly focus on the returns before 2005 and many include the 1987 crash episode, whereas mine starts at 1996 and extends all the way to the most recent period. It is worthwhile to mention that the recent six years (2006-2011) see significant increases in returns for these strategies. For instance, the average 92%-OTM and 96%-OTM put returns are -12.8% and -13.6% per month, much larger than the sample averages for early years. A full investigation on the change in these return characteristics are beyond the scope of this paper. To the extent that my sample is under-representative of the true option returns population and involves overestimated mean returns, the bounds I constructed below can be regarded as conservative lower bounds on the generalized entropy function of the pricing kernel.

## 4.2 Bounds implied by option strategies

With these return data from the asset market and relying on the analytical tools developed in the previous section, I explore their implications on the behavior of a pricing kernel. Ideally, to provide the sharpest bounds, we need to search for the optimal dynamic portfolio strategy that maximize a certain unconditional moment of the return. However, parameter estimates for even simple static portfolio choice problems are usually very unstable, partly due to the volatile nature of market returns (See Brandt (1999), Sahalia and Brandt (2001)). Here since we considering portfolios that involve highly non-normal and mechanically correlated option returns, the estimation inaccuracy issue can only get worse. Eventually, these estimation uncertainty will translate into bounds uncertainty and may significantly lower our confidence in the bound informativeness. To alleviate this estimation issue, I choose to consider simple static option strategy that has the following form

$$R_P = R_f + \alpha_S(R_S - R_f) \tag{16}$$

where  $\alpha_S$  denotes the fraction of wealth allocated to a generic return  $R_S$ . Hence, only the tradeoff between a safe asset and a return is considered<sup>8</sup>. By focusing on this type

---

<sup>8</sup>By doing this, I also ignore the possible utility gain from a combination of the market index and a derivative strategy. For a CRRA investor with a risk aversion coefficient more than one, Driessen and Maenhout (2005) show the allocation to the market index is rarely significant. This indirectly shows the limited utility gains by adding in the market index.

of strategy, we are able to depict bounds provided by clean economically meaningful individual option strategies.

Figure 3 plots the bound frontiers given in inequality (11) and (12) when the power  $s$  equals to 2, 0.5, 0, -1, -3 and -8, and Table 2 reports the optimal portfolio weights at a riskfree rate of zero per month<sup>9</sup>. Notice that when the power is equal to two, the admissible region for the pricing kernel is below the depicted curves, whereas at other powers it is above the depicted curves. I intentionally leave this “inconvenient” feature in the picture to emphasize the flip in the direction of bounds around  $s = 1$ . Four different types of portfolios are shown on this plot: two types involving the two benchmark derivative strategies and for comparison purposes, the other two types involving the market index.

Several patterns emerge from Figure 3 and Table 2. First, strategies involving the put option clearly dominate the other strategies across all powers and a wide range of riskfree rates. At  $s = 2$ , which is at the HJ bound, Sharpe Ratio essentially determines the strength of restrictions that a given security return puts on the pricing kernel. As a result, in accordance with the rankings of absolute Sharpe Ratio given in Table 1, 96%-OTM put returns implies the sharpest constraint, which is followed by the index and lastly the straddle returns. At other powers, the optimal bound belongs to the spectrum of generalized entropy bounds and hence has a utility-based interpretation as discussed before. In particular, a bound at a given power  $s$  corresponds to the transformed optimized utility of a power utility agent with a risk-aversion of  $\frac{1}{1-s}$ . In a closely related empirical paper, Driessen and Maenhout (2005) study the asset allocation problem of an investor who has access to index options. Their empirical results lend support to my results, especially at  $s = 0.5$  and  $s = 0$ . At  $s = 0.5$ , both their paper and my results show that an agent with a risk-aversion of two will have a significant short position in the 96%-OTM put option: their paper, allowing the market index to enter into the asset menu as well, has an estimate of about  $-10\%$  for  $\alpha_S$  while I have an estimate of approximately  $-20\%$  by excluding the index. At  $s = 0$ , which corresponds to the logarithmic utility case as in the original entropy bound, their estimate is around  $-15\%$  and mine again roughly doubles their estimate. Putting the difference in asset menus and sample periods aside, both studies establish the economic benefits by allowing investors to trade deep OTM put options. The two plots for  $s = 0.5$  and  $s = 0$  in Figure 3 vividly illustrate these benefits by highlighting the differences in utility gains among the four candidate strategies. As  $s$  goes negative, the related risk-aversion coefficient becomes even smaller so the relative gain in expected returns by shorting put options further outweighs the loss from variance and other higher-order moments. Consequently, the optimal bound requires even larger absolute position in the OTM put option. Notably, the ATM

---

<sup>9</sup>The mean and the standard deviation of the riskfree rate for the short sample is 4bp per month and 40bp, respectively. Therefore, I center it at zero and extend to  $\pm 3$  standard deviations away from the center in Figure 3.

straddle returns for my sample yield an unimpressive mean (either from a longing or a shorting perspective) relative to 96%-OTM put options and a very high variance relative to the market index. Accordingly, strategies involving the straddle returns appear to imply an inferior bound compared to either the put option strategy or the index strategy.

To offer a deeper interpretation of the above empirical findings, it is worthwhile to repeat the insights in bound interpretations given in Section 2.2. Although the marginal or representative investor determines the market prices of jump and volatility risks<sup>10</sup>, investors with different risk-attitudes all reveal valuable information about these prices through their trading behavior. For example, all else being equal, higher prices of jump risks imply more expensive deep OTM put options. With a fixed physical jump distribution, this implies a lower ex-ante and ex-post average return for buying puts. However, for a less risk-averse agent who does not value the put option's hedging ability as much as the average consumer, she treats the increase in put prices as a lucrative trading opportunity. By shorting more, she increases her expected utility. Following this logic, the above empirical findings highlight the more important roles of priced jump risks as opposed to volatility risks in option prices. Notice that this is not saying that volatility risks should have little pricing impact, since we are only looking at this through the lens of static power utility agents' optimization problems. To have priced volatility risks emerge as an effective bound, we may have to investigate more complicated dynamic strategies given the strong predictability in volatilities. This is left as a future research topic.

Besides the two benchmark derivative strategies, alternative strategies may appear attractive for certain CRRA investors and thus provide tighter bounds on the pricing kernel. Figure 4 and 5 show bounds implied by two alternative put option strategies and two crash-neutral strategies, respectively, and Table 3 shows the corresponding weights. Figure 4 shows that a strategy that shorts the 96%-OTM put option turns out to be the dominating one across all powers. This is somewhat surprising since we would expect its performance to lie in between the ones involving the 92%-OTM put and the ATM put. Closer looks at the admissible portfolio weights reveal that strategies shorting the 92%-OTM put have strong weight restrictions: given that the maximum net return observation is close to 4 (See Table 1), the maximum proportion that one can short on a 92%-OTM put is  $1/(1 - (4 + 1)) = -0.25$  at a riskfree rate of zero. In absolute value, this is significantly smaller than the allowable 0.45 short position on the 96%-OTM put, as shown in Table 2. Consequently, despite the fact that the 92%-OTM put has a more negative average return than the 96%-OTM put, weight constraints prevent investors from exploiting it any further. This phenomenon,

---

<sup>10</sup>The presence of jump and stochastic volatility in the index price is well established in the option pricing literature. Buraschi and Jackwerth (2001), Coval and Shumway (2001) and Bakshi and Kapadia (2003) show the presence of volatility premium. Bates (2002), Pan (2002) and Ait-Sahalia, Wang and Yared (2001) show the presence of jump risk premium.

although statistically irrelevant for the portfolio dominance results (strategies involving the 96%-OTM put are more attractive partly because they allow more aggressive short positions), does point to the instability issue of the in-sample portfolio choice problems<sup>11</sup>. I will shortly come back to have a full discussion of this issue.

Figure 5 shows that bounds from the benchmark put option strategy dominate crash-neutral strategies as well at  $s = 2$ , and fare well for the other three powers. In particular, the deep OTM put strategy weakly dominates the crash-neutral put strategy at  $s = 0.5$ , and is on par with the latter at  $s = 0$  and a little less informative at  $s = -4$ . The key observation is that difference between bounds based on these two strategies is significantly smaller than the difference between them and the other two strategies. Taken as a whole, Figure 4 and 5 suggest the superior role of deep OTM put strategies (especially strategies with 96%-OTM put) in effectively shaping the admissible space of candidate pricing kernels. They reveal information on the pricing of jump risks in the economy and empirically provide the sharpest bounds that any valid discount rate has to satisfy.

The above optimized bounds can be directly used to confront candidate pricing kernels. However, they may appear too stringent or cumbersome from several practical concerns. First, in-sample portfolio choice generate portfolio weights that are too volatile (See Brandt (2000), Driessen and Maenhout (2005)). What is even more detrimental in our setup is the boundary-dependence of the optimal weights. As seen from Table 2 and 3, weights for several strategies are close to their boundary values for large negative powers. This exacerbates the in-sample instability issue since extreme observations are more sample-dependent than sample moments. Second, transaction costs and margin requirements for real-world option trading strategies may limit the amount that we can short. Although transaction costs are small for the index option markets (Bakshi, Cao and Chen (1997)) and our long positions in the riskfree asset can serve as margin, these microstructure effects may become non-negligible if we have excessive short positions on index options. Third, weights for these optimal bounds depend on the prevailing riskfree rate and are thus time-varying. This is cumbersome for most applications when we do not observe the bond rate, since a different bound has to be calculated for a different riskfree rate. Casual inspections of Figure 4 and 5 reveal that option implied bounds are stable across a wide range of bond rates. This prompts us to think about option strategies that are invariant to the bond rates.

Driven by these concerns, I propose a simple way to create conservative and interest rate independent portfolio strategies. For the most informative 96%-OTM

---

<sup>11</sup>Notice that the in-sample portfolio choice problems are well-defined both theoretically and numerically within our context. In particular, although sometimes the solutions are close to the boundary (See Table 2 when  $s = -1, -3$  or  $-8$ ), the infinitely large marginal utility at the boundary for an CRRA investor will restrict the optimal choice to be well within the boundary. However, the boundary-dependence of the optimal solution is exacerbated in our context because the CRRA investor's risk-aversion is sometimes close to zero.

put strategies, I set a lower threshold  $\alpha_L = -35\%$ <sup>12</sup> on the short position to avoid excessive shorting and fix the put weight at the optimal zero-interest put weight given in Table 2 if it does not exceed the threshold, or at  $\alpha_L$  otherwise. By doing this, we end up having the following three types of OTM put strategies: a longing 50% strategy at  $s = 2$ ; a shorting  $-20\%$  strategy at  $s = 0.5$  and a shorting  $-35\%$  strategy at other powers. Figure 6 display the bounds for these conservative strategies together with the bounds from the two benchmark strategies. Not surprisingly, they agree well with the optimal strategies for  $s = 2, 0.5$  and  $0$ . At  $s = -4$ , the discrepancy is about 1% and much smaller than the difference of around 8% between the straddle strategy and the optimal OTM put strategy. For the rest of the paper, I use these simple yet efficient bounds to study restrictions on the pricing kernel. Again, had we missed important information by utilizing these sub-optimal bounds, they still provide valid, albeit conservative restrictions on the pricing kernel.

### 4.3 Rare disaster models and option return bounds

I consider a representative-agent exchange economy model with infrequent large declines in consumption growth (See Barro (2006, 2009)). More precisely, I focus on models with an iid environment. This is a first step in understanding the distribution of tail events in consumption growth. Moreover, since we restrict ourselves to simple static portfolio strategies in the construction of bounds, it is only fair to consider a pricing kernel with iid shocks. I also adopt a Poisson-normal distribution for the jump component in consumption growth<sup>13</sup>. This parametric setup has two appealing features: 1. It is flexible enough to match some of the empirical regularities on rare event distributions as documented by Barro (2006); 2. It is infinitely divisible and thus allows us to “zoom” into an arbitrarily small frequency. I state the model at the annual frequency but will use its monthly counterpart to match asset market bounds constructed from monthly returns data.

The time-additive utility representation for the representative agent is given by

$$E_0\left(\sum_{t=0}^{\infty} \beta^t \frac{C_t^{1-\gamma}}{1-\gamma}\right)$$

where  $\gamma$  governs investor risk-aversion. The pricing kernel is known to be

$$\log M_{t+1} = \log \beta - \gamma \log G_{t+1} \tag{17}$$

---

<sup>12</sup>This number is chosen to approximately equate the optimal weight at  $s = 0$  and a riskfree rate of zero. It is also about 10% away from the boundary weight, which helps alleviate the boundary-dependent problem.

<sup>13</sup>For its applications in the Macro-finance literature, see Naik and Lee (1990), Martin (2007) and Backus, Chernov and Zin (2011)

where  $G_{t+1} \equiv C_{t+1}/C_t$  is the consumption growth rate. Log consumption growth is assumed to be driven by two independent shocks,

$$\log G_{t+1} = \epsilon_{t+1} + \eta_{t+1} \quad (18)$$

where  $\epsilon_{t+1} \sim \mathcal{N}(\mu, \sigma^2)$  is the normally-distributed component and the distribution of the jump component  $\eta_{t+1}$  is given by

$$\eta_{t+1}|(J = j) \sim \mathcal{N}(j\theta, j\nu^2), J \sim \text{Poisson}(\omega) \quad (19)$$

To derive the entropy-related quantities for this kernel, we start from calculating the moment-generating functions(MGF) of the two shocks. The MGF for the normal shock is  $E(e^{s\epsilon_{t+1}}) = \exp(\mu s + \sigma^2 s^2/2)$  and the MGF for the Poisson-normal part, shown in Backus, Chernov and Martin (2011), is

$$E(e^{s\eta_{t+1}}) = \exp(\omega[e^{s\theta+(s\nu)^2/2} - 1]) \quad (20)$$

Then, by independence of the two shocks, the cumulant-generating function (CGF) can be shown to be

$$CGF(s) \equiv \log E(e^{s \log M_{t+1}}) = s \log \beta - \gamma \mu s + \frac{1}{2} \gamma^2 \sigma^2 s^2 + \omega[e^{-\gamma s \theta + (\gamma s \nu)^2/2} - 1] \quad (21)$$

By setting  $s = 1$ , the continuously compounded one-period riskfree rate  $r_f \equiv -\log E(M_{t+1})$  is given by

$$r_f = -(\log \beta - \gamma \mu + \frac{1}{2} \gamma^2 \sigma^2 + \omega[e^{-\gamma \theta + (\gamma \nu)^2/2} - 1]) \quad (22)$$

Finally, combining the above two pieces, the generalized entropy function (GEF) can be shown as

$$GEF(s) = \frac{1}{2} \gamma^2 \sigma^2 (1 - s) + \omega[e^{-\gamma \theta + (\gamma \nu)^2/2} - \frac{1}{s} e^{-\gamma s \theta + (\gamma s \nu)^2/2} - \frac{s - 1}{s}] \quad (23)$$

When  $s \rightarrow 0$ ,  $GEF(s)$  reduces to the original entropy

$$L(M) = \frac{1}{2} \gamma^2 \sigma^2 + \omega[e^{-\gamma \theta + (\gamma \nu)^2/2} + \gamma \theta - 1] \quad (24)$$

Using  $GEF(s)$  as a model diagnosing tool, we are now ready to explore the effects of rare disasters on the pricing kernel. The full stacked parameter vector for a disaster model is  $(\beta, \gamma, \mu, \sigma^2, \omega, \theta, \nu^2)'$ , which has seven components. To better concentrate on the economically interesting disaster intensity  $\omega$  and size  $\theta$ , I perform a ‘‘partial derivative’’ exercise. First, as shown in Panel A of Table 4, I fix the two preference parameters, the riskfree rate and two variance statistics related to the consumption growth process. The total variance in consumption growth  $\sigma^2 + \omega(\theta^2 + \nu^2)$  is fixed at the sample estimate based on the US real consumption data (Backus, Chernov and



Martin (2011)) and the variance for the normal component of the Poisson-normal shock is estimated from realized disasters based on international data (Barro (2006)). These second-order moments can be estimated with much more precision than first-order moments ( $\mu$  and  $\theta$ ) and my choices agree well with the disaster literature. Next, I fix the expected loss in a disaster state at  $\omega\theta = -0.006$  and consider three  $\omega$  and  $\theta$  combos that represent a light, mild and severe disaster type, respectively, as shown in Panel B of Table 4<sup>14</sup>. These three types of disaster distributions imply an increasing magnitude in disaster size and roughly agree with the historical consumption data of the US, an average country in Barro's sample (Barro (2006)) and a few European countries which experience large drops in per capita GDP during World War II, respectively. Note that these alternative tail specifications are difficult to differentiate empirically, since by design they imply exactly the same mean consumption growth, total volatility in growth and interest rate. This goes to the heart of the so-called pseudo problem in disaster models. It is therefore interesting to see if *GEF* can better distinguish these models, and additionally, if asset market returns provide support for some of them.

Figure 7 shows the corresponding *GEF* plots for the three disaster models. I focus on the region from around -3.5 to 1. As  $s$  goes beyond 1 or below -3.5, either the severe disaster case or the light disaster case will display explosive behavior. In addition, I will show later that bounds at these powers are less informative anyway. Starting from  $s = 1$  where all three *GEF*'s equal zero and going left, the *GEF* of the severe disaster type rises more steeply compared to the other two and reaches its peak around  $s = 0.8$ . At its peak, the *GEF* more than triples the one for either the mild or light disaster type. As going further left, it remains the dominating one until  $s$  reaches -3, at which the *GEF* from the light disaster type catches up. The mild disaster type follows a similar pattern, rising more than the light disaster case initially and meeting it at around -2. Eventually, all of them will start rising for large negative powers, with the light disaster case being the earliest one in turning heads.

To see how various weighting schemes on cumulants generate the patterns in the *GEF*, Table 8 displays the contributions from the second to sixth weighted cumulants to the overall (generalized) entropy<sup>15</sup>. Since the severe disaster type involves explosive behavior at the sixth moment, I choose to focus on the first two types of disaster distributions. We can view the case at  $s = 0$  as the benchmark, since all individual cumulants are weighted equally. At  $s = 0$ , both disaster types imply the same second cumulant (variance) by design. However, the mild disaster case implies higher third

---

<sup>14</sup>The mean  $\mu$  of the normal shock component is used to match the interest rate.

<sup>15</sup>Strictly speaking, as in equation 13, the  $j$ -th individual cumulant of the log pricing kernel is  $\kappa_j(\log M)$ . However, with a slight abuse of terminology, I sometimes refer to the factorial adjusted cumulant  $\frac{\kappa_j(\log M)}{j!}$  as cumulant since factorial provides a natural scaling of the raw cumulant. The  $j$ -th weighted cumulant is defined as  $\frac{\kappa_2(\log M)}{j!}(1 - s^{j-1})$ . See Backus, Chernov and Martin (2011) on the derivation of the individual cumulants for Poisson-normal shocks

to sixth cumulants. As a result, its overall entropy is higher<sup>16</sup>. At  $s = 2$ , which corresponds to the Hansen-Jaganathan bound, the signs for cumulants are reversed and more importantly higher-order cumulants are magnified compared to the case at  $s = 0$ . This is because the polynomial multipliers attach more weights to higher-order terms at  $s = 2$ . This also explains the relatively large magnitude in entropy when the power goes above one, as shown in Figure 7. At  $s = -1$ , interestingly, all odd cumulants vanish and the entropy is a sum of even moments only. Given the vast literature on skewness preferences<sup>17</sup>, it is interesting to see if even moments alone can stand some of the empirical regularities from option returns. In other words, stronger restrictions on the pricing kernel might be found by focusing on  $s = -1$  since skewness or in general odd cumulants are no longer present to help boost up entropy. More dramatically, at  $s = -3$ , odd cumulants show up as negative and hence are canceling out the effects of even cumulants. The empirical inquiry at  $s = -1$  applies to this case as well. Moreover, these bites coming from negative odd cumulants explain why the light disaster GEF dominates the mild disaster GEF at  $s = -3$ : relatively larger disaster size  $\theta$  for the mild disaster case generates disproportionately high absolute odd cumulants and they in turn reduce the overall entropy. In sum, the generalized entropy function generates interesting combinations of cumulants and it remains an empirical question to see if asset returns can bind it in a meaningful way.

I now confront the rare disaster models with option implied bounds. Analogous to the usual calibration approach in the Macro-finance literature, the plan is to mark up the admissible parameter space corresponding to a set of asset market bounds. Similar approaches have been taken by Hansen and Jaganathan (1991) to depict the efficient mean-variance frontier from market Sharpe ratios and by Bansal and Lehmann (1994) to restrict the representative agent's risk-aversion based on equity premium. Again, to sharpen our focus on economically interesting quantities, I choose to consider the triple  $(\omega, \theta, \gamma)'$ . To avoid negative volatility, this time I choose to fix the variance  $\sigma^2$  of the normal shock component<sup>18</sup>. Also, risk aversion is released as a free parameter to match return bounds. Other than these changes, the rest are the same as in Panel A of Table 4.

Bounds based on option strategies essentially delineate a specific domain in the three-dimensional  $(\omega, \theta, \gamma)$  space. To ease visual inspections, I plot the contours

---

<sup>16</sup>Of course, other higher-order cumulants matter, so we need to extrapolate from the patterns seen from the second to sixth cumulant.

<sup>17</sup>See Kraus and Litzenberger (1976), Rubinstein (1973), Harvey and Siddique (2000)

<sup>18</sup>We could do this for the previous calibration exercise. In fact, there will be little change in model implications if we adopt this, because only local variants of the baseline disaster model ( $|\omega\theta|$  is small) are considered. I choose to fix the total variance before so it has a partial derivative flavor. This time, however, since we are searching over the entire  $(\omega, \theta, \gamma)$  space, a fixed total variance may sometimes imply a negative  $\sigma^2$ . Therefore, I set  $\sigma^2$  at  $0.035^2$  instead. In fact, given the rare occurrence of disasters (especially for the US), I think it makes more sense to equalize  $\sigma^2$  with the estimate based on historical data.

on the two-dimensional  $(\omega, \gamma)$  plane. Figure 9 shows these contours for four values on disaster size  $\theta$ . Many interesting patterns emerge from this figure. Firstly, as expected, larger magnitude in  $\theta$  require smaller risk aversion. At  $\theta = -0.10$ , a 4% annual equity risk premium asks for a risk aversion of around 5.5 if a disaster occurs every 60 years. When  $\theta$  drops to -0.50, a risk aversion of 2.3 would suffice. Fixing the disaster frequency at  $1/60$ , as  $\theta$  changes from -0.10 to -0.50, the required risk aversion by the most stringent entropy bound drops from slightly more than 10 to around 6. In particular, at Barro's calibration ( $\omega = 1/60, \theta = -0.38$ ), I calculate that a risk aversion of 7.2 is needed to satisfy the entropy bound at  $s = -1$ . Such a risk aversion may be regarded as too high to reconcile with many aspects of an individual's risk taking behavior. Secondly, option strategy returns require much larger risk aversion coefficient than the equity risk premium. In particular, at  $1/\omega = 100$ , the HJ bound implied by longing 50% in the 96%-OTM put typically requires an extra 0.5 units in risk aversion across different disaster size specifications. On top of that, the entropy bound at  $s = 0$  (by shorting 40% on 96%-OTM put) asks for an additional 2 units in risk aversion. Finally, the incremental requirement imposed by the most stringent entropy bound at  $s = -1$  is around 0.5 to 1 in units of risk aversion. Changes in risk aversion may be hard to quantify economically; a better way to read the economic significance off the figure is to reverse the roles of the x- and y-axis. For instance, when  $\theta = -0.50$  and by fixing the risk aversion at 6, an entropy bound at  $s = -1$  implies that disasters need to happen on average at least once every 50 years, whereas a period of around 100 years would suffice for the entropy bound at  $s = 0$ . A 50% drop in consumption that happens twice every century is certainly less preferably than the case with only one drop every century. The difference in bounds' implications is hence economically significant. Thirdly, in unreported results, I consider alternative interest rates and entropy bounds at even more negative powers. An annual riskfree rate in the range of  $(1.00, 1.06)$  results in little change in the contour plots. This is mainly because of the insensitivity of option implied bounds to interest rates, as we discussed before. Entropy bounds at more negative powers do not substantially improve on the parameter frontiers depicted by the entropy bound at  $s = -1$ .

To relate my findings to the literature, it is important to emphasize the methodological difference. In particular, Backus, Chernov and Martin (2011) try to infer rare event information from option data. However, two strong assumptions are made in their paper to make the task feasible: 1. Dividend is a levered claim on consumption; 2. A Merton type option pricing model is chosen to summarize the cross-section option data. Both of them are based more on convenience than on reality, and it is hard to measure how small deviations from them could affect the inference on tail information. My approach, on the other hand, is model free by nature. Based on the fundamental no-arbitrage condition, it asks how much (generalized) dispersion a pricing kernel has to generate in order to rationalize the profits from trading jump risks. Although it cannot deliver a definitive point estimate, informative bounds can still tell a lot on the robust features of a pricing kernel. In terms of the empirical

findings, they conclude that option prices imply lower probabilities of extreme adverse events than what international macroeconomic data imply. My results, to the contrary, show that more frequent and/or severe disasters are probably needed to reconcile with the historical option return series<sup>19</sup>

#### 4.4 Testing rare disaster models with option market bounds

In the previous section, we show how to derive bounds on the pricing kernel from various moments of the historical option returns. However, to fully appreciate the strength of these bounds in restricting the pricing kernel, we need to study its statistical significance. This is the task of this section. I start by laying down a testing framework. Suppose the set of parameters governing a specific model is  $\Pi$ . In the case of a disaster model,  $\Pi = (\beta, \gamma, \mu, \sigma^2, \omega, \theta, \nu^2)'$ . The transformed moment vector of the pricing kernel is defined as

$$\Omega_M(\Pi; S) = \begin{bmatrix} [EM^{s_1}]^{\frac{1}{1-s_1}} \cdot I_{s_1 \in [0,1)} \\ [EM^{s_2}]^{\frac{1}{1-s_2}} \cdot I_{s_2 \in [0,1)} \\ \dots \\ [EM^{s_N}]^{\frac{1}{1-s_N}} \cdot I_{s_N \in [0,1)} \end{bmatrix} \quad (25)$$

where  $S = (s_1, s_2, \dots, s_N)'$  denotes the collection of powers we are interested in and  $I_{s_j \in [0,1)}$  equals  $-1$  if  $s_j \in [0, 1)$  and  $1$  otherwise. At  $s = 0$ ,  $I_{s \in [0,1)} = -1$  and the corresponding transformed moment  $E(M^s)^{\frac{1}{1-s}}$  should be understood as  $E(\log M)$ . These sign functions  $I_{s_j \in [0,1)}$ 's adjust the directions of bounds so that the left hand side (moments of the pricing kernel) should always dominate the right hand side (moments of market returns). Similarly, the transformed return moment vector is given by

$$\Omega_R(R; S) = \begin{bmatrix} E(R_1^{\frac{s_1}{s_1-1}}) \cdot I_{s_1 \in [0,1)} \\ E(R_2^{\frac{s_1}{s_1-1}}) \cdot I_{s_2 \in [0,1)} \\ \dots \\ E(R_N^{\frac{s_1}{s_1-1}}) \cdot I_{s_N \in [0,1)} \end{bmatrix} \quad (26)$$

where  $R_j$  denotes a specific type of return from the market. At  $s = 0$ , the return moment  $E(R^{\frac{s}{s-1}})$  should be treated as  $-E(\log R)$ . For a valid parameterization  $\Pi$  of the pricing kernel, the difference between  $\Omega_M(\Pi; S)$  and  $\Omega_R(R; S)$  should be positive, i.e.,

$$\Theta(S) \equiv \Omega_M(\Pi; S) - \Omega_R(R; S) > 0. \quad (27)$$

---

<sup>19</sup>My results partially agree with the option return literature (See Jackwerth (2000) and Bondarenko (2003)), which find that more frequent crashes need to be added to explain away put returns.

The element-wise positivity of  $\Theta(S) = (\theta(s_1), \theta(s_2), \dots, \theta(s_N))'$  constitutes our basic testable assumptions. In actual estimation, the population moments of returns can be replaced by their sample counterparts for a given sample size  $T$ :

$$\hat{\Theta}(S) = \Omega_M(\Pi; S) - \begin{bmatrix} \sum_{t=1}^T R_{t1}^{\frac{s_1}{s_1-1}} \cdot I_{s_1 \in [0,1)} \\ \sum_{t=1}^T R_{t2}^{\frac{s_2}{s_2-1}} \cdot I_{s_2 \in [0,1)} \\ \dots \\ \sum_{t=1}^T R_{tN}^{\frac{s_N}{s_N-1}} \cdot I_{s_N \in [0,1)} \end{bmatrix} \quad (28)$$

By applying the Generalized Method of Moments of Hansen and Singleton (1982), we can easily find the asymptotic distribution of  $\hat{\Theta}(S)$  for this just identified system. However, our problem is a non-standard multivariate inequality test and the usual Wald or Likelihood-ratio tests will not apply. The difficulty lies in the specification of a null hypothesis that can generate easy-to-calculate critical values. Following the multivariate testing literature by Gourieroux, Holly and Monfort (1982), Wolak (1987) and especially a recent paper by Patton and Timmermann (2010), I set the null at  $\Theta(S) = 0$ , which is least favorable to the alternative  $\Theta(S) > 0$ . To find p-values, I simulate a large number of draws from the empirical limiting distribution of  $\hat{\Theta}(S)$  and calculate the fraction of draws that an element-wise positive  $\theta$  vector is observed. This is similar to the simulation exercise in Patton and Timmerman (2010).

The above testing procedure ignores the estimation uncertainty in the GMM asymptotic variance-covariance matrix. This uncertainty could be large given the presence of highly skewed and fat-tailed option returns. As an alternative, I bootstrap the historical return data to provide robust p-values. In particular, I re-sample the historical return series with replacement for a large number of times. For each sample, I calculate the in-sample  $\Omega_R(R; S)$  vector and compare it to  $\Omega_M(\Pi; S)$ . Finally, I count the number of times that  $\Omega_M(\Pi; S)$  does not lie above  $\Omega_R(R; S)$  in an element-wise sense.

Ideally, we would like to jointly consider generalized entropy bounds at various powers with multiple assets. However, a few statistical issues restrict the way that we can form these tests. First, testing moment bounds at different powers using the same asset is problematic. This is because the disturbance terms are perfectly related in a nonlinear way, which violates the basic ergodicity assumption for most asymptotic theories to work. Second, the limiting distribution becomes increasingly unstable as we increase the dimension of the testing statistics. Given a few hundred monthly observations, this puts a practical limit on the size of the panel of bounds we can think of. Facing these issues, I choose to consider two types of simple bound test. One is the univariate entropy bound test using the market return only. This also serves as the benchmark test since many recent papers studying equity risk premium impose such a bound (See Backus, Chernov and Martin (2011, 2012), Martin (2008) and Alrevaz and Jermann (2005)). The other type is the joint test of the entropy

bound of the index and a generalized entropy bound of an option trading strategy. As in the previous section, bounds at powers of 2, 0.5, 0, -1 and -2 are considered and the corresponding optimal option trading strategies are given in Section 4.2.

To focus on important quantities and to isolate the effects of the riskfree rate, I use the following procedure to generate a null hypothesis for the parameter vector  $\Pi$ . First, similar to before, I treat  $(\beta, \sigma^2, \nu^2)$  as nuisance parameters and set them at  $(0.99, 0.035^2, 0.2^2)$ . Second, I set the possible real annualized riskfree rate at 0.95, 1.00 and 1.05, which roughly correspond to the lower bound, mean and upper bound of a sensible estimate based on the US history. Lastly, given a point of interest for the triple  $(\omega, \theta, \gamma)'$ , I choose  $\mu$  to match the required interest rate. By doing this, I give  $(\omega, \theta, \gamma)'$  a lot of freedom in meeting bounds from option data and the burden in hitting a target interest rate is transferred onto  $\mu$ . Of course, a large deviation of  $\mu$  from the historical consumption growth indicates a failure of the hypothesized  $\Pi$ . I choose to report the model implied mean consumption growth  $\mu + \omega\theta$  instead<sup>20</sup>.

Table 5 reports the testing results for the baseline disaster model with  $\omega = 0.02, \theta = -0.35$ . This roughly corresponds to the  $\omega = 0.017, \theta = -0.38$  estimate based on the empirical distribution of international disasters by Barro (2006), and close to the baseline parameter choice of the rare disaster literature (See Martin (2009), Backus, Chernov and Zin (2011)). At  $\gamma = 5$ , both the entropy bound for the index and the HJ bound for the optimal option strategy are satisfied at the historical mean, as demonstrated by the positivity of the corresponding  $M_{diff}$  statistics<sup>21</sup>. Notably, at  $s = 0$ , the modeled implied entropy is in excess of the risk premium by at least 5% annually. Hence, not surprisingly, the MKT test have p-values well above 0.90 across all interest rate specifications, suggesting no evidence in the rejection of the entropy bound from the index returns. This is also the case when HJ bound is added, indicating no discriminating power from HJ bound either. When we include more demanding option strategy returns at the original entropy bound ( $s = 0$ ), the p-values drop to 10-18%. The reduction in p-value is impressive, considering the close to one p-value when the index return is the only testing asset. Nonetheless, the model survives at conventional significance levels. When  $s$  goes to  $-1$  and  $-2$ , p-values drop to well below 5%, suggesting a strong rejection of the baseline disaster model across all target interest rates.

To remedy this, one must enhance the amount of dispersion in the generalized entropy function. As we learned from the previous section (see Figure 9), one good way to achieve this is to increase the marginal investor's risk aversion. Indeed, when  $\gamma$  is raised to 6, the model is on the borderline of acceptance at the 5% significance

<sup>20</sup>This is because  $E_c = \mu + \omega\theta$  is the strict model counterpart to the mean historical consumption growth. For reasonable  $\omega$  and  $\theta$  pairs, the difference between  $E_c$  and  $\mu$  is small.

<sup>21</sup>At  $s = 0$ , the  $M_{diff}$  statistic can be directly interpreted as the dispersion in the pricing kernel in excess of the risk premium. At other  $s$  values,  $M_{diff}$  measures generalized excess dispersion and its sign indicates if a bound is satisfied at the historical mean return moment.

level, and at  $\gamma = 7$  the p-values show no sign of rejection at all. However, as risk aversion gets larger, the implied mean growth rate becomes implausibly high. For instance, the implied mean growth rate  $E_c$  is about 5% when  $\gamma = 6$  and  $R_f = 1.00$ ; when  $\gamma$  rises to seven,  $E_c$  is in the range of 7-9%. These numbers are in contradiction with the historical consumption data for the US. Clearly, a tension exists between risk aversion and riskfree rate, and it is stronger than what thin-tailed consumption growth models imply. An inspection of equation (22) tells us why this is the case. Compared to standard models with normal shocks, disaster models carry an extra exponential term  $\omega[e^{-\gamma\theta+(\gamma\nu)^2/2} - 1]$ . Since disasters rarely happen ( $\omega$  is small), this term is small for low  $\gamma$  values. However, as risk aversion rises, it grows exponentially and quickly dominates the intertemporal substitution effect  $\gamma\mu$  and the precautionary savings effect  $\frac{1}{2}\gamma^2\sigma^2$ . In particular, at  $\gamma = 7$  and assuming a mean consumption growth rate of 2%, this term is four times the value of the substitution effect and 20 times the value of the precautionary savings effect. Clearly, under a disaster model framework, the agent's hedging demand for rare event risk is strong for even moderately high risk aversion levels. Given a mean consumption growth at around 2%, this strong hedging motive requires low risk aversion to reconcile with the relatively high interest rate. On the other hand, asset market returns ask for higher risk aversion to square with generalized entropy bounds. These two forces make the determination of the representative agent's risk attitude no longer a one-sided exercise. Taken as a whole, it seems that the baseline disaster model is unable to explain the bond market and the option market at the same time.

If the extrapolated disaster distributions cannot explain the US market, what can? I next test disaster models with alternative distributional assumptions. Table 6 reports the testing results at  $\omega = 0.02$ ,  $\theta = -0.10$ , which is arguably what the US has experienced<sup>22</sup>. Table 7 and 8 show testing results based on perturbations around the baseline case. In Table 6, as expected, the rejections from violations of the entropy bounds are generally stronger than those in the baseline case. In fact, risk aversion has to go all the way up to 9 to pass the bound test at  $s = -1$  at 5% level. At the same time, the hedging demand at  $\theta = -0.10$  is substantially lower than that at  $\theta = -0.35$ . As a result, the US type disaster model still implies a sensible mean consumption growth even at  $\gamma = 9$ . If one is willing to accept such a high risk aversion level, then the US type disaster specification can fit the asset markets. For the severe type model in Table 7, a risk aversion of 5.5 suffices to satisfy option return bounds, but a 8% mean growth rate in consumption seems implausibly high. Turning to the less severe but more frequent type shown in Table 8, a somewhat high mean growth rate around 3.5% and a somewhat high risk aversion of  $\gamma = 7$  are simultaneously needed to pass all the tests.

---

<sup>22</sup>For the US, a consumption decline in the magnitude of 10% only happened once: in 1931, the per capita consumption dropped by 9.9%. Hence, strictly speaking, my assumption on the disaster frequency doubles what the US actually experienced. A lower assumed  $\omega$  value makes the rejections in Table 6 even more stronger.

Finally, one may wonder if the demanding entropy bounds at negative powers come from the excessive short position that we allow investors to hold. In other words, given the statistical uncertainty around the optimal allocation rule, maybe the selected representative option trading strategies imply in-sample moment characteristics that are too harsh for any representative agent type of model to satisfy. To address this robustness concern, I reduce  $\alpha_L$  by half and redo all the tests at  $\gamma = 5$ . Table 9 show the results. We see that although all the p-values associated with  $s = 0, -1$  and  $-2$  are somewhat larger than their counterparts in the previous tables, the rejections are still strong. In particular, the baseline model is rejected at below 5% significance level for  $s = -1$  and  $s = -2$  across all target interest rates. For the US type and the mild type disaster models, none of the specifications can pass the generalized entropy bound tests at  $s = 0, -1$  or  $-2$ . For the severe type, the p-values are now well above 5% but it still does not qualify as a successful model since the implied mean consumption growth rate is too high.

As we reduce the magnitude of  $\alpha_L$ , we do see that the bound mean statistics  $M_{diff}$  gets better (closer to zero) at negative powers. For instance, for the baseline model, the bound inequality at  $s = 0$  is improved by around 8%-10%. This is equivalent to say that we reduce the mean (log) return of the representative option trading strategy by 8%-10%. With such a significant mean drop, how come the p-values do not increase much? This stems from the small variation of the transformed option returns. To see this, note that at negative power  $s$ , we are essentially taking fractional powers of the returns, e.g., at  $s = -1$ ,  $E(R^{\frac{s}{s-1}}) = E(R^{\frac{1}{2}})$ . Consequently, the highly volatile option strategy returns are flattened out by these power transformations. This reduced variation in the transformed sample, together with a negative mean bound estimate, makes the rejection of bounds highly significant. This feature again highlights the discriminatory power of generalized entropy bounds at negative powers.

In summarizing the above empirical findings, I show how standard disaster models under several parameterizations fail to meet the nonparametric bounds based on robust option trading strategies. The discriminatory power of bounds with negative powers are highlighted. However, I consider these findings as suggestive as opposed to conclusive. First, although I find evidence against streamline disaster models, their abilities in magnifying (generalized) entropy through tail distortions are impressive. Within a risk aversion of ten, even the US type specification can meet all the entropy bounds with a reasonable amount of mean consumption growth. This leads one to conjecture that more sophisticated variants of disaster models may meet the challenge (See Barro and Ursua (2008) and Watcher (2008)). Second, even for the current version of disaster model, I have not exhausted all plausible parameter choices. For instance, the variance  $\nu^2$  for the normally distributed individual jump is shown to have important effect on the riskfree rate for at a high risk aversion level. Yet I only set it at 0.2 for simplicity. A more extensive exploration of the seven-component parameter vector  $\Pi$  may yield a winner.



Despite these caveats, there are a few important takeaways from the above exercise. First and foremost, confronting a pricing kernel with the equity risk premium alone is not enough, especially when the tail behavior is considered. In fact, except at  $\gamma = 2$  and for a riskfree rate below 1, the equity risk premium constraint is satisfied across all specifications, most of the time with a p-value close to one. This reveals its lack of power in discriminating tail behaviors of the pricing kernel. By subjecting a discount rate to a spectrum of option trading strategies at different powers, we gain a better sense of its all-around performance. Secondly, it is crucial to consider generalized entropy bounds at negative powers, not only because of its informativeness by mean moment restrictions as demonstrated by Figure 9, but also because of the statistical powers they afford through fractional power transformations of returns. This analytical feature, combined with the moment characteristics of option returns, can potentially provide the most exacting and thus informative moment constraints on the pricing kernel. Of course, return robustness should be checked to avoid overly restrictive constraints.

## 5 Conclusion

Under the fundamental no-arbitrage condition, this paper develops a spectrum of new nonparametric bounds that significantly enrich the nonparametric bound universe. These bounds essentially describe the discrepancy between what an optimizing agent could achieve if all admissible assets were tradable and what she actually achieves in the real-world market, thus providing an economically interesting way to restrict candidate pricing kernels. Motivated by these new bounds, I propose to use the generalized entropy function, a natural extension of the original entropy, to systematically study asset market implied bounds. Through cumulant-expansions on both sides of the generalized entropy bounds, I show how the new bounds provide distinct information on the pricing kernel. Their abilities in teasing out tail information are highlighted.

Equipped with these analytical tools, I study index option returns, since their unique moment characteristics can potentially provide the sharpest restrictions on the pricing kernel. Empirically, I find that strategies with short positions in deep OTM put options dominate both the market index and other standard derivative trading strategies in shaping moments of the pricing kernel. This fact stresses the pricing impact of jump risks in the index, and is expected to be helpful in inferring rare event distribution in the pricing kernel. I then postulate a pricing kernel in the form of a standard disaster model, and use option return bounds to differentiate between alternative parameterizations. Both point estimates and formal testing results indicate the deficiency of standard disaster models in reconciling with option data.

Both tail distortion and time-dependency might be needed to attain bounds implied by option returns.

A study of the joint behavior of time-varying disaster distribution and option returns is an obvious avenue to pursue. This not only helps achieve option return bounds, but can also generate insights on the time series properties of option returns. The newly developed bound system, especially an extended version that can cope with conditioning information, is expected to be instrumental as well. On the other hand, it remains interesting to see other applications of the generalized entropy bounds, possibly on welfare analysis, model diagnosis, and the creation of new models.

## References

- [1] Alvarez, Fernando and Urban Jermann, 2005, Using asset prices to measure the persistence of the marginal utility of wealth, *Econometrica* 73, 1977-2016.
- [2] Ait-Sahalia, Y., Wang, Y. and Yared, F., 2001, Do options markets correctly price the probabilities of movements of the underlying asset?, *Journal of Econometrics* 102, 67-110.
- [3] Backus, David, Mikhail Chernov and Ian Martin, 2011, Disasters implied by equity index options, *Journal of Finance* 66, 1969-2012.
- [4] Backus, David, Mikhail Chernov and Stanley Zin, 2011, Sources of entropy in representative agent models, Working paper, August.
- [5] Bakshi, G., C. Cao and Z. Chen, 1997, Empirical performance of alternative option pricing models, *Journal of Finance* 52, 2003-2049.
- [6] Bakshi, G. and Kapadia, N., 2003, Delta-hedged gains and the negative market volatility risk premium, *Review of Financial Studies* 16, 527-566.
- [7] Bansal, Ravi and Bruce N. Lehmann, 1997, Growth-optimal portfolio restrictions on asset pricing models, *Macroeconomic Dynamics* 1, 333-354.
- [8] Barro, Robert J., 2006, Rare disasters and asset markets in the twentieth century, *Quarterly Journal of Economics* 121, 823-867.
- [9] Barro, Robert J., Emi Nakamura, Jon Steinsson and Jose F. Ursua, 2009, Crises and recoveries in an empirical model of consumption disasters, manuscript, June.
- [10] Barro, Robert J. and Jose F. Ursua, 2008, Macroeconomic crises since 1870, *Brookings Papers on Economic Activity* no. 1, 255-350. .
- [11] Bates, D., 2002, The market for crash risk, unpublished working paper, University of Iowa.
- [12] Bekaert, Geert and Jun Liu, 2004, Conditioning information and variance bounds on pricing kernels, *Review of financial studies* 17, 339-378.
- [13] Brandt, M., 1999, Estimating portfolio and consumption choice: A conditional euler equations approach, *Journal of Finance* 54, 1609-1646.
- [14] Broadie, Mark, Mikhail Chernov and Michael Johannes, 2009, Understanding index option returns, *Review of Financial Studies*, 22, 4493-4529.
- [15] Bondarenko, O., 2003, Why are Puts So Expensive?, working paper, University of Illinois, Chicago.

- [16] Buraschi, A. and Jackwerth, J., 2001, The price of a smile: Hedging and spanning in option markets, *Review of Financial Studies* 14, 495-527.
- [17] Casella, George and Roger L. Berger, 2001, Statistical Inference, Duxbury Advanced Series.
- [18] Coval, J. and T. Shumway, 2001, Expected Option returns, *Journal of Finance* 56, 983-1009.
- [19] Davidson, James, 1994, Stochastic limit theory, Oxford University Press.
- [20] Driessen, Joost and Maenhout, Pascal, 2007, An empirical portfolio perspective on option pricing anomalies, *Review of Finance*, 11, 561-603.
- [21] Ferson, Wayne and Andrew Siegel, 2003, Stochastic discount factor bounds with conditioning information, *Review of Financial Studies*, 16, 567-595.
- [22] Gabaix, X., 2009, Variable rare disasters: an exactly solved framework for ten puzzles in macro-finance. Unpublished Working Paper, New York University.
- [23] Gallant, R., L.P. Hansen and G. Tauchen, 1990, Using conditional moments of asset payoffs to infer the volatility of intertemporal marginal rates of substitution, *Journal of Econometrics* 45, 141-179.
- [24] Gourio F., 2008, Time-series predictability in the disaster model, *Finance Research Letters*, 5 (2008), pp. 191203.
- [25] Gourieroux, C. , Holly, A. and Monfort, A., 1982, Likelihood ratio test, Wald test, and Kuhn-Tucker test n linear models with inequality constraints on the regression parameters, *Econometrica* 50, 63-80.
- [26] Hansen, L.P., and K. Singleton, 1983, Stochastic consumption, risk aversion, and the temporal behavior of asset returns, *Journal of Political Economy*, 91, 249-265.
- [27] Hansen, L.P. and R. Jaganathan, 1991, Implications of security market data for models of dynamic economies, *Journal of Political Economy* 99, 225-262.
- [28] Hansen, L.P. and R. Jaganathan, 1994, Assessing specification errors in stochastic discount factor models, Staff report 167, Federal Reserve Bank of Minneapolis.
- [29] Harvey, Campbell R. and Akhtar Siddque, 2000, Conditional skewness in asset pricing tests, *Journal of Finance* 55, 1263-1295.
- [30] Jackwerth, J., 2000, Recovering risk aversion from option prices and realized returns, *Review of Financial Studies* 13, 433-451.
- [31] Kraus, Alan and Robert H. Litzenberger, 1976, Skewness preference and the valuation of risk assets, *Journal of Finance* 31, 1085-1100.

- [32] Longstaff, Francis A., and Monika Piazzesi, 2004, Corporate earnings and the equity premium, *Journal of Financial Economics* 74, 401-421.
- [33] Martin, Ian, 2007, Disasters and asset pricing: evidence from option markets, manuscript, February.
- [34] Martin, Ian, 2008, Disasters and the welfare cost of uncertainty, 2008, *The American Economic Review* 98, 74-78.
- [35] Martin, Ian, 2009, Consumption-based asset pricing with higher cumulants, manuscript, January.
- [36] Naik, Vasanttilak and Moon Lee, 1990, General equilibrium pricing of options on the market portfolio with discontinuous returns, *Review of Financial Studies* 3, 493-521.
- [37] Pan, J., 2002, The jump-risk premia implicit in options: evidence from an integrated time-series study, *Journal of Financial Economics* 63, 3-50.
- [38] Patton, Andrew J. and Timmermann, Allan, 2010, Monotonicity in asset returns: new tests with applications to the term structure, the CAPM, and portfolio sorts, *Journal of Financial Economics* 98, 605-625.
- [39] Rietz, Thomas A., 1988, The equity premium: a solution, *Journal of Monetary Economics* 22, 117-131.
- [40] Rubinstein, Mark E., 1973, The fundamental theorem of parameter-preference security valuation, *Journal of Financial and Quantitative Analysis*, Vol. 8, No. 1, 61-69.
- [41] Snow, K.N., 1991, Diagnosing asset pricing models using the distribution of asset returns, *Journal of Finance* 46, 955-983.
- [42] Wachter, Jessica A., 2008, Can time-varying risk of rare disasters explain aggregate stock market volatility?, Working paper, April.
- [43] Wolak, F.A., 1987, An exact test for multiple inequality and equality constraints in the lineal regression model. *Journal of the American Statistical Association* 82, 782-793.

## A Tables and Figures

	Mean	Std.	Sharp	Skew.	Kurt.	Max	Min	Corr.Index
Index,long	0.007	0.055	0.113	0.225	10.323	0.384	-0.285	NA
Index,short	0.005	0.049	0.087	-0.605	3.666	0.118	-0.177	1.000
0.92-OTM put	-0.236	0.728	-0.325	2.756	13.049	3.944	-0.877	-0.305
0.96-OTM put	-0.201	0.611	-0.330	1.519	4.839	2.261	-0.849	-0.277
1.00-OTM put	-0.082	0.627	-0.132	1.897	7.990	3.256	-0.849	-0.323
ATM straddle	-0.017	0.340	-0.052	1.702	6.809	1.501	-0.778	0.001
C-neutral put	-0.310	0.996	-0.312	-2.553	18.805	2.135	-7.438	-0.096
C-neutral straddle	-0.015	0.428	-0.035	-0.528	13.847	1.665	-2.823	0.077
R-C-neutral put	-0.212	0.724	-0.293	-0.075	7.469	3.135	-2.960	-0.119
R-C-neutral straddle	-0.025	0.422	-0.059	-0.660	14.602	2.665	-1.823	0.059

Table 1: Summary statistics. This table reports the summary statistics of the returns for the S&P 500 index and several derivative strategies. The long index series is from July 1926 to December 2011 and the short index series is from January 1996 to December 2011, which is also the time span for all the option strategy returns. The first column displays the strategy name and the last column reports the correlation of the strategy returns with the short-sample market index. “C-neutral put” and “C-neutral straddle” denote crash-neutral put and straddle returns, for which the original 96%-OTM put and ATM straddle are mixed with a short leg on 92%-OTM put option, respectively. See Coval and Shumway (2001) for the construction of the crash-neutral put and Jackwerth (2000) for the construction of the crash-neutral straddle. “R-C-neutral put” and “R-C-neutral straddle” denote robust crash-neutral put and straddle returns, respectively. They are the original crash neutral series excluding the date when the 92%-OTM put maturity date is more than three trading weeks longer than the 96%-OTM put maturity date at the moment of buying. By doing this, data for six months are deleted from the 192 months of data, including two months for which the 92%-OTM put has higher price than the 96%-OTM put. Skewness and kurtosis are the standardized central third and fourth moments, respectively. The riskfree rate is 60bp annualized for the long sample and 54bp for the short sample. These rates are the inputs in calculating the Sharpe Ratios for the corresponding samples.

Power	Market	96%-OTM put	ATM straddle
Panel A			
$s = 2$	-1.954	0.489	0.149
$s = 0.5$	0.941	-0.199	-0.070
$s = 0$	1.839	-0.341	-0.137
$s = -1$	3.450	-0.437	-0.260
$s = -3$	5.428	-0.442	-0.461
$s = -8$	5.667	-0.442	-0.663
Panel B			
$\alpha_{min}$	-9.589	-0.450	-0.680
$\alpha_{max}$	5.369	1.175	1.281

Table 2: **Optimal portfolio weights for benchmark strategies and the index.** Panel A shows the optimal portfolio weights for the optimization problem described in figure 3 at a fixed riskless rate of zero per month. Panel B shows the range of the admissible portfolio weights that guarantees a positive portfolio return series. For a return series  $\{R_t\}_{t=1}^T$ , the range is given by  $[\alpha_{min}, \alpha_{max}] = [1/(1 - \max[\{R_t\}_{t=1}^T]), 1/(1 - \min[\{R_t\}_{t=1}^T])]$ .

Power	92%-OTM put	ATM put	R-C-neutral put	R-C-neutral straddle
Panel A				
$s = 2$	0.406	0.207	0.253	0.138
$s = 0.5$	-0.129	-0.088	-0.184	-0.070
$s = 0$	-0.202	-0.162	-0.331	-0.138
$s = -4$	-0.253	-0.306	-0.467	-0.545
Panel B				
$\alpha_{min}$	-0.254	-0.307	-0.468	-0.601
$\alpha_{max}$	1.141	1.178	0.253	0.354

Table 3: **Optimal portfolio weights for alternative put and crash-neutral strategies.** Panel A shows the optimal portfolio weights for the optimization problem described in figure 4 and 5 at a fixed riskless rate of zero per month. Strategies involving the 92%-put option, ATM put option, crash-neutral put option and crash-neutral straddle are shown. Panel B shows the range of the admissible portfolio weights that guarantees a positive portfolio return series. For a return series  $\{R_t\}_{t=1}^T$ , the range is given by  $[\alpha_{min}, \alpha_{max}] = [1/(1 - \max[\{R_t\}_{t=1}^T]), 1/(1 - \min[\{R_t\}_{t=1}^T])]$ .



Parameter	Value
Panel A	
$\beta$	0.99
$\gamma$	5
$r_f$	0.02
$\sigma^2 + \omega(\theta^2 + \nu^2)$	$0.035^2$
$\nu^2$	$0.2^2$
Panel B	
$\omega_L$	0.04
$\theta_L$	-0.15
$\omega_M$	0.02
$\theta_M$	-0.30
$\omega_S$	0.01
$\theta_S$	-0.60

Table 4: **Parameter specifications for disaster models** This table shows the parameter specifications for disaster models. Panel A shows the fixed parameters. The total variance in consumption growth is given by  $\sigma^2 + \omega(\theta^2 + \nu^2)$ . Panel B shows the disaster intensity and size combos that represent three types of disaster distributions: Light disaster type  $(\omega_L, \theta_L)$ , Mild disaster type  $(\omega_M, \theta_M)$  and Severe disaster type  $(\omega_S, \theta_S)$ .

			MKT		MKT + OPT			
			$s = 0$	$s = 2$	$s = 0.5$	$s = 0$	$s = -1$	$s = -2$
$\gamma = 2$	$R_f = 0.95$ ( $E_c = -0.025$ )	$M_{diff}$	-0.081	-1.018	-0.315	-0.500	-0.387	-0.566
		$P^a$ -value	0.029	0.004	0.001	0.003	0.000	0.000
		$P^b$ -value	0.031	0.005	0.001	0.005	0.000	0.000
	$R_f = 1.00$ ( $E_c = 0.001$ )	$M_{diff}$	-0.030	-1.120	-0.263	-0.449	-0.361	-0.532
		$P^a$ -value	0.252	0.017	0.009	0.025	0.001	0.000
		$P^b$ -value	0.234	0.017	0.011	0.027	0.001	0.000
	$R_f = 1.05$ ( $E_c = 0.025$ )	$M_{diff}$	0.019	-1.218	-0.215	-0.400	-0.337	-0.500
		$P^a$ -value	0.666	0.023	0.036	0.063	0.002	0.000
		$P^b$ -value	0.664	0.022	0.041	0.066	0.003	0.001
$\gamma = 5$	$R_f = 0.95$ ( $E_c = 0.018$ )	$M_{diff}$	0.057	2.698	-0.213	-0.362	-0.303	-0.444
		$P^a$ -value	0.906	0.904	0.046	0.096	0.003	0.002
		$P^b$ -value	0.901	0.908	0.052	0.100	0.006	0.002
	$R_f = 1.00$ ( $E_c = 0.028$ )	$M_{diff}$	0.108	2.628	-0.162	-0.311	-0.278	-0.409
		$P^a$ -value	0.995	0.994	0.100	0.139	0.010	0.003
		$P^b$ -value	0.995	0.995	0.104	0.135	0.012	0.004
	$R_f = 1.05$ ( $E_c = 0.044$ )	$M_{diff}$	0.157	2.561	-0.114	-0.262	-0.253	-0.376
		$P^a$ -value	1.000	1.000	0.179	0.173	0.017	0.005
		$P^b$ -value	1.000	1.000	0.182	0.175	0.020	0.007
$\gamma = 6$	$R_f = 0.95$ ( $E_c = 0.039$ )	$M_{diff}$	0.202	9.103	-0.093	-0.217	-0.223	-0.331
		$P^a$ -value	1.000	1.000	0.227	0.225	0.030	0.014
		$P^b$ -value	1.000	1.000	0.225	0.216	0.034	0.016
	$R_f = 1.00$ ( $E_c = 0.048$ )	$M_{diff}$	0.253	9.088	-0.043	-0.165	-0.197	-0.296
		$P^a$ -value	1.000	1.000	0.364	0.288	0.046	0.024
		$P^b$ -value	1.000	1.000	0.350	0.269	0.047	0.032
	$R_f = 1.05$ ( $E_c = 0.056$ )	$M_{diff}$	0.302	9.074	0.005	-0.116	-0.173	-0.263
		$P^a$ -value	1.000	1.000	0.519	0.336	0.071	0.042
		$P^b$ -value	1.000	1.000	0.506	0.322	0.069	0.045
$\gamma = 7$	$R_f = 0.95$ ( $E_c = 0.074$ )	$M_{diff}$	0.485	10.831	0.149	0.067	-0.072	-0.121
		$P^a$ -value	1.000	1.000	0.885	0.590	0.275	0.202
		$P^b$ -value	1.000	1.000	0.886	0.587	0.257	0.200
	$R_f = 1.00$ ( $E_c = 0.081$ )	$M_{diff}$	0.536	10.831	0.199	0.118	-0.046	-0.086
		$P^a$ -value	1.000	1.000	0.944	0.659	0.348	0.292
		$P^b$ -value	1.000	1.000	0.948	0.644	0.338	0.286
	$R_f = 1.05$ ( $E_c = 0.088$ )	$M_{diff}$	0.585	10.831	0.246	0.167	-0.021	-0.052
		$P^a$ -value	1.000	1.000	0.976	0.727	0.435	0.361
		$P^b$ -value	1.000	1.000	0.979	0.715	0.416	0.365

Table 5: **Baseline disaster model testing results.** This table reports the testing results for the baseline disaster model with  $\omega = 0.02, \theta = -0.35$ .  $R_f$  is the annual riskfree rate and  $E_c = \mu + \omega\theta$  is the implied mean consumption growth. MKT denotes the test of the entropy bound with the market return alone, and MKT+OPT denotes the joint test of the entropy bound for the market return and generalized entropy bound at power  $s$  for the corresponding option strategy return.  $M_{diff}$  is the mean bound as in equation (28). It measures the mean bound for the index under MKT and the mean bound for the corresponding option strategy at  $s$  under MKT+OPT.  $P^a$ -value and  $P^b$ -value are the P-values generated from the theoretical limiting distribution and a bootstrap re-sampling procedure, respectively.

		MKT		MKT + OPT				
			$s = 0$	$s = 2$	$s = 0.5$	$s = 0$	$s = -1$	$s = -2$
$\gamma = 2$	$R_f = 0.95$ ( $E_c = -0.028$ )	$M_{diff}$	-0.089	-1.054	-0.319	-0.507	-0.392	-0.574
		$P^a$ -value	0.020	0.002	0.001	0.003	0.000	0.000
		$P^b$ -value	0.023	0.004	0.001	0.004	0.000	0.000
	$R_f = 1.00$ ( $E_c = -0.003$ )	$M_{diff}$	-0.037	-1.156	-0.268	-0.456	-0.366	-0.540
		$P^a$ -value	0.196	0.012	0.008	0.019	0.001	0.000
		$P^b$ -value	0.191	0.013	0.010	0.023	0.001	0.000
	$R_f = 1.05$ ( $E_c = 0.022$ )	$M_{diff}$	0.012	-1.254	-0.219	-0.407	-0.342	-0.507
		$P^a$ -value	0.601	0.016	0.036	0.060	0.001	0.000
		$P^b$ -value	0.610	0.017	0.037	0.062	0.003	0.000
$\gamma = 5$	$R_f = 0.95$ ( $E_c = -0.004$ )	$M_{diff}$	-0.054	-0.731	-0.298	-0.472	-0.364	-0.518
		$P^a$ -value	0.107	0.023	0.003	0.010	0.000	0.000
		$P^b$ -value	0.110	0.027	0.006	0.013	0.001	0.001
	$R_f = 1.00$ ( $E_c = 0.006$ )	$M_{diff}$	-0.003	-0.831	-0.247	-0.421	-0.338	-0.483
		$P^a$ -value	0.468	0.060	0.019	0.047	0.001	0.000
		$P^b$ -value	0.473	0.062	0.020	0.048	0.002	0.001
	$R_f = 1.05$ ( $E_c = 0.016$ )	$M_{diff}$	0.046	-0.926	-0.198	-0.372	-0.314	-0.451
		$P^a$ -value	0.860	0.060	0.052	0.087	0.005	0.002
		$P^b$ -value	0.862	0.064	0.060	0.090	0.005	0.003
$\gamma = 7$	$R_f = 0.95$ ( $E_c = 0.006$ )	$M_{diff}$	0.010	2.272	-0.252	-0.409	-0.320	-0.387
		$P^a$ -value	0.595	0.590	0.019	0.054	0.002	0.004
		$P^b$ -value	0.588	0.584	0.024	0.060	0.005	0.005
	$R_f = 1.00$ ( $E_c = 0.013$ )	$M_{diff}$	0.061	2.199	-0.201	-0.357	-0.294	-0.352
		$P^a$ -value	0.925	0.921	0.053	0.094	0.006	0.009
		$P^b$ -value	0.927	0.928	0.058	0.099	0.009	0.012
	$R_f = 1.05$ ( $E_c = 0.020$ )	$M_{diff}$	0.110	2.128	-0.152	-0.309	-0.26	-0.319
		$P^a$ -value	0.995	0.994	0.111	0.135	0.012	0.015
		$P^b$ -value	0.995	0.994	0.116	0.135	0.013	0.019
$\gamma = 9$	$R_f = 0.95$ ( $E_c = 0.022$ )	$M_{diff}$	0.167	10.814	-0.123	-0.252	-0.221	0.418
		$P^a$ -value	1.000	1.000	0.162	0.186	0.031	0.998
		$P^b$ -value	1.000	1.000	0.160	0.185	0.032	0.998
	$R_f = 1.00$ ( $E_c = 0.028$ )	$M_{diff}$	0.218	10.813	-0.073	-0.201	-0.195	0.455
		$P^a$ -value	1.000	1.000	0.289	0.239	0.044	0.999
		$P^b$ -value	1.000	1.000	0.274	0.232	0.050	1.000
	$R_f = 1.05$ ( $E_c = 0.033$ )	$M_{diff}$	0.267	10.813	-0.025	-0.152	-0.171	0.490
		$P^a$ -value	1.000	1.000	0.417	0.297	0.070	0.999
		$P^b$ -value	1.000	1.000	0.409	0.286	0.072	1.000

Table 6: **US type disaster model testing results.** This table reports the testing results for the US type disaster model with  $\omega = 0.02, \theta = -0.10$ .  $R_f$  is the annual riskfree rate and  $E_c = \mu + \omega\theta$  is the implied mean consumption growth. MKT denotes the test of the entropy bound with the market return alone, and MKT+OPT denotes the joint test of the entropy bound for the market return and generalized entropy bound at power  $s$  for the corresponding option strategy return.  $M_{diff}$  is the mean bound as in equation (28). It measures the mean bound for the index under MKT and the mean bound for the corresponding option strategy at  $s$  under MKT+OPT.  $P^a$ -value and  $P^b$ -value are the P-values generated from the theoretical limiting distribution and a bootstrap re-sampling procedure, respectively.

			MKT		MKT + OPT			
			$s = 0$	$s = 2$	$s = 0.5$	$s = 0$	$s = -1$	$s = -2$
$\gamma = 2$	$R_f = 0.95$ ( $E_c = -0.023$ )	$M_{diff}$	-0.077	-0.974	-0.312	-0.496	-0.385	-0.563
		$P^a$ -value	0.037	0.007	0.001	0.005	0.000	0.000
		$P^b$ -value	0.034	0.006	0.001	0.004	0.000	0.000
	$R_f = 1.00$ ( $E_c = 0.003$ )	$M_{diff}$	-0.026	-1.075	-0.260	-0.444	-0.359	-0.529
		$P^a$ -value	0.281	0.018	0.009	0.027	0.000	0.000
		$P^b$ -value	0.278	0.022	0.011	0.034	0.001	0.000
	$R_f = 1.05$ ( $E_c = 0.028$ )	$M_{diff}$	0.023	-1.172	-0.212	-0.396	-0.334	-0.497
		$P^a$ -value	0.703	0.020	0.037	0.067	0.002	0.001
		$P^b$ -value	0.703	0.029	0.046	0.072	0.003	0.001
$\gamma = 5$	$R_f = 0.95$ ( $E_c = 0.049$ )	$M_{diff}$	0.213	9.786	-0.076	-0.206	-0.224	-0.339
		$P^a$ -value	1.000	1.000	0.279	0.230	0.028	0.012
		$P^b$ -value	1.000	1.000	0.262	0.232	0.031	0.014
	$R_f = 1.00$ ( $E_c = 0.059$ )	$M_{diff}$	0.264	9.777	-0.026	-0.154	-0.198	-0.305
		$P^a$ -value	1.000	1.000	0.417	0.286	0.049	0.019
		$P^b$ -value	1.000	1.000	0.413	0.291	0.053	0.027
	$R_f = 1.05$ ( $E_c = 0.069$ )	$M_{diff}$	0.313	9.768	0.022	-0.106	-0.174	-0.271
		$P^a$ -value	1.000	1.000	0.566	0.349	0.066	0.033
		$P^b$ -value	1.000	1.000	0.554	0.334	0.075	0.037
$\gamma = 5.5$	$R_f = 0.95$ ( $E_c = 0.075$ )	$M_{diff}$	0.379	10.818	0.068	-0.040	-0.137	-0.220
		$P^a$ -value	1.000	1.000	0.707	0.438	0.123	0.072
		$P^b$ -value	1.000	1.000	0.699	0.427	0.125	0.076
	$R_f = 1.00$ ( $E_c = 0.084$ )	$M_{diff}$	0.430	10.817	0.118	0.011	-0.111	-0.185
		$P^a$ -value	1.000	1.000	0.821	0.512	0.172	0.110
		$P^b$ -value	1.000	1.000	0.828	0.499	0.175	0.102
	$R_f = 1.05$ ( $E_c = 0.093$ )	$M_{diff}$	0.479	10.817	0.165	0.060	-0.086	-0.151
		$P^a$ -value	1.000	1.000	0.906	0.587	0.228	0.158
		$P^b$ -value	1.000	1.000	0.904	0.574	0.220	0.156

Table 7: **Severe type disaster model testing results.** This table reports the testing results for the severe type disaster model with  $\omega = 0.01, \theta = -0.60$ .  $R_f$  is the annual riskfree rate and  $E_c = \mu + \omega\theta$  is the implied mean consumption growth. MKT denotes the test of the entropy bound with the market return alone, and MKT+OPT denotes the joint test of the entropy bound for the market return and generalized entropy bound at power  $s$  for the corresponding option strategy return.  $M_{diff}$  is the mean bound as in equation (28). It measures the mean bound for the index under MKT and the mean bound for the corresponding option strategy at  $s$  under MKT+OPT.  $P^a$ -value and  $P^b$ -value are the P-values generated from the theoretical limiting distribution and a bootstrap re-sampling procedure, respectively.

		MKT		MKT + OPT				
			$s = 0$	$s = 2$	$s = 0.5$	$s = 0$	$s = -1$	$s = -2$
$\gamma = 2$	$R_f = 0.95$ ( $E_c = -0.026$ )	$M_{diff}$	-0.085	-1.039	-0.317	-0.503	-0.389	-0.569
		$P^a$ -value	0.025	0.003	0.001	0.003	0.000	0.000
		$P^b$ -value	0.028	0.004	0.002	0.003	0.000	0.000
	$R_f = 1.00$ ( $E_c = -0.001$ )	$M_{diff}$	-0.033	-1.141	-0.266	-0.452	-0.363	-0.535
		$P^a$ -value	0.222	0.013	0.007	0.022	0.000	0.000
		$P^b$ -value	0.215	0.014	0.010	0.026	0.001	0.000
	$R_f = 1.05$ ( $E_c = 0.024$ )	$M_{diff}$	0.016	-1.239	-0.217	-0.403	-0.339	-0.502
		$P^a$ -value	0.633	0.018	0.036	0.060	0.001	0.000
		$P^b$ -value	0.634	0.021	0.038	0.066	0.003	0.001
$\gamma = 5$	$R_f = 0.95$ ( $E_c = 0.005$ )	$M_{diff}$	-0.009	-0.011	-0.266	-0.427	-0.336	-0.477
		$P^a$ -value	0.426	0.249	0.010	0.043	0.001	0.001
		$P^b$ -value	0.415	0.250	0.014	0.045	0.002	0.002
	$R_f = 1.00$ ( $E_c = 0.015$ )	$M_{diff}$	0.043	-0.104	-0.215	-0.376	-0.310	-0.442
		$P^a$ -value	0.841	0.396	0.040	0.080	0.004	0.001
		$P^b$ -value	0.840	0.383	0.048	0.088	0.006	0.003
	$R_f = 1.05$ ( $E_c = 0.025$ )	$M_{diff}$	0.092	-0.193	-0.167	-0.327	-0.286	-0.409
		$P^a$ -value	0.983	0.376	0.091	0.117	0.007	0.002
		$P^b$ -value	0.984	0.361	0.096	0.122	0.008	0.003
$\gamma = 7$	$R_f = 0.95$ ( $E_c = 0.027$ )	$M_{diff}$	0.159	7.682	-0.136	-0.260	-0.235	-0.274
		$P^a$ -value	1.000	1.000	0.139	0.185	0.027	0.032
		$P^b$ -value	1.000	1.000	0.138	0.179	0.026	0.039
	$R_f = 1.00$ ( $E_c = 0.035$ )	$M_{diff}$	0.210	7.655	-0.085	-0.208	-0.209	-0.239
		$P^a$ -value	1.000	1.000	0.253	0.231	0.035	0.053
		$P^b$ -value	1.000	1.000	0.244	0.222	0.041	0.056
	$R_f = 1.05$ ( $E_c = 0.042$ )	$M_{diff}$	0.259	7.629	-0.037	-0.159	-0.184	-0.206
		$P^a$ -value	1.000	1.000	0.392	0.296	0.060	0.086
		$P^b$ -value	1.000	1.000	0.376	0.274	0.058	0.088
$\gamma = 8$	$R_f = 0.95$ ( $E_c = 0.046$ )	$M_{diff}$	0.335	10.803	0.012	-0.083	-0.135	-0.012
		$P^a$ -value	1.000	1.000	0.533	0.381	0.122	0.459
		$P^b$ -value	1.000	1.000	0.517	0.369	0.125	0.468
	$R_f = 1.00$ ( $E_c = 0.052$ )	$M_{diff}$	0.387	10.803	0.062	-0.032	-0.108	0.024
		$P^a$ -value	1.000	1.000	0.683	0.454	0.178	0.560
		$P^b$ -value	1.000	1.000	0.689	0.443	0.178	0.555
	$R_f = 1.05$ ( $E_c = 0.058$ )	$M_{diff}$	0.435	10.803	0.109	0.017	-0.084	0.058
		$P^a$ -value	1.000	1.000	0.811	0.525	0.241	0.649
		$P^b$ -value	1.000	1.000	0.806	0.508	0.230	0.640

Table 8: **Mild type disaster model testing results.** This table reports the testing results for the mild type disaster model with  $\omega = 0.04, \theta = -0.15$ .  $R_f$  is the annual riskfree rate and  $E_c = \mu + \omega\theta$  is the implied mean consumption growth. MKT denotes the test of the entropy bound with the market return alone, and MKT+OPT denotes the joint test of the entropy bound for the market return and generalized entropy bound at power  $s$  for the corresponding option strategy return.  $M_{diff}$  is the mean bound as in equation (28). It measures the mean bound for the index under MKT and the mean bound for the corresponding option strategy at  $s$  under MKT+OPT.  $P^a$ -value and  $P^b$ -value are the P-values generated from the theoretical limiting distribution and a bootstrap re-sampling procedure, respectively.

		MKT		MKT + OPT				
		$s = 0$	$s = 2$	$s = 0.5$	$s = 0$	$s = -1$	$s = -2$	
$(\omega = 0.02, \theta = -0.35)$	$R_f = 0.95$ ( $E_c = 0.018$ )	$M_{diff}$	0.057	2.698	-0.213	-0.286	-0.151	-0.198
		$P^a$ -value	0.901	0.907	0.045	0.006	0.003	0.002
		$P^b$ -value	0.906	0.908	0.045	0.008	0.004	0.003
	$R_f = 1.00$ ( $E_c = 0.028$ )	$M_{diff}$	0.108	2.628	-0.162	-0.234	-0.125	-0.164
		$P^a$ -value	0.993	0.994	0.097	0.019	0.011	0.015
		$P^b$ -value	0.995	0.995	0.102	0.022	0.013	0.013
	$R_f = 1.05$ ( $E_c = 0.038$ )	$M_{diff}$	0.157	2.561	-0.114	-0.186	-0.101	-0.131
		$P^a$ -value	1.000	1.000	0.186	0.047	0.032	0.032
		$P^b$ -value	1.000	1.000	0.180	0.055	0.033	0.040
$\omega = 0.02, \theta = -0.10$	$R_f = 0.95$ ( $E_c = -0.004$ )	$M_{diff}$	-0.054	-0.731	-0.298	-0.396	-0.212	-0.272
		$P^a$ -value	0.106	0.025	0.003	0.000	0.000	0.000
		$P^b$ -value	0.112	0.026	0.004	0.000	0.000	0.000
	$R_f = 1.00$ ( $E_c = 0.006$ )	$M_{diff}$	-0.003	-0.831	-0.247	-0.345	-0.186	-0.238
		$P^a$ -value	0.474	0.061	0.016	0.001	0.000	0.000
		$P^b$ -value	0.478	0.059	0.019	0.002	0.001	0.001
	$R_f = 1.05$ ( $E_c = 0.016$ )	$M_{diff}$	0.046	-0.926	-0.198	-0.296	-0.162	-0.205
		$P^a$ -value	0.863	0.063	0.054	0.004	0.001	0.002
		$P^b$ -value	0.857	0.058	0.062	0.007	0.003	0.004
$\omega = 0.04, \theta = -0.15$	$R_f = 0.95$ ( $E_c = 0.005$ )	$M_{diff}$	-0.009	-0.011	-0.266	-0.351	-0.184	-0.232
		$P^a$ -value	0.427	0.256	0.013	0.001	0.000	0.001
		$P^b$ -value	0.419	0.246	0.014	0.001	0.001	0.001
	$R_f = 1.00$ ( $E_c = 0.015$ )	$M_{diff}$	0.043	-0.104	-0.215	-0.300	-0.158	-0.197
		$P^a$ -value	0.842	0.383	0.041	0.004	0.002	0.004
		$P^b$ -value	0.841	0.374	0.049	0.006	0.003	0.003
	$R_f = 1.05$ ( $E_c = 0.025$ )	$M_{diff}$	0.092	-0.193	-0.167	-0.251	-0.133	-0.164
		$P^a$ -value	0.984	0.375	0.097	0.013	0.006	0.009
		$P^b$ -value	0.984	0.365	0.095	0.018	0.008	0.014
$\omega = 0.01, \theta = -0.60$	$R_f = 0.95$ ( $E_c = 0.049$ )	$M_{diff}$	0.213	9.786	-0.076	-0.129	-0.072	-0.094
		$P^a$ -value	1.000	1.000	0.273	0.123	0.091	0.095
		$P^b$ -value	1.000	1.000	0.269	0.129	0.094	0.094
	$R_f = 1.00$ ( $E_c = 0.059$ )	$M_{diff}$	0.264	9.777	-0.026	-0.078	-0.046	-0.059
		$P^a$ -value	1.000	1.000	0.418	0.243	0.193	0.209
		$P^b$ -value	1.000	1.000	0.408	0.240	0.204	0.199
	$R_f = 1.05$ ( $E_c = 0.069$ )	$M_{diff}$	0.313	9.768	0.022	-0.029	-0.022	-0.026
		$P^a$ -value	1.000	1.000	0.569	0.402	0.341	0.361
		$P^b$ -value	1.000	1.000	0.560	0.396	0.338	0.347

Table 9: **Robust option strategies testing results.** This table reports the testing results for various disaster models with robust option strategies. In particular, the short position in 96%-OTM put is halved to 20% for option strategies at  $s = 0, -1$  and  $-2$ . In the table,  $R_f$  is the annual riskfree rate and  $E_c = \mu + \omega\theta$  is the implied mean consumption growth. MKT denotes the test of the entropy bound with the market return alone, and MKT+OPT denotes the joint test of the entropy bound for the market return and generalized entropy bound at power  $s$  for the corresponding option strategy return.  $M_{diff}$  is the mean bound as in equation (28). It measures the mean bound for the index under MKT and the mean bound for the corresponding option strategy at  $s$  under MKT+OPT.  $P^a$ -value and  $P^b$ -value are the P-values generated from the theoretical limiting distribution and a bootstrap re-sampling procedure, respectively.

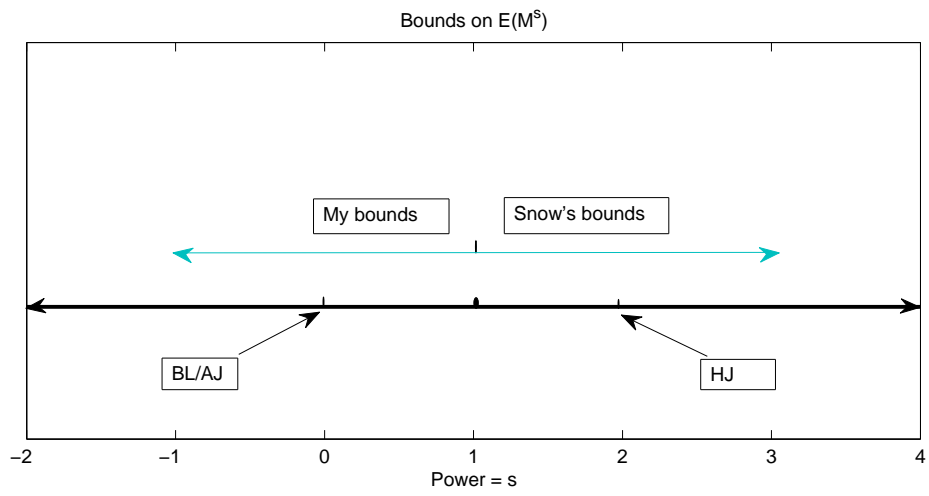


Figure 1: **The non-parametric bound universe**

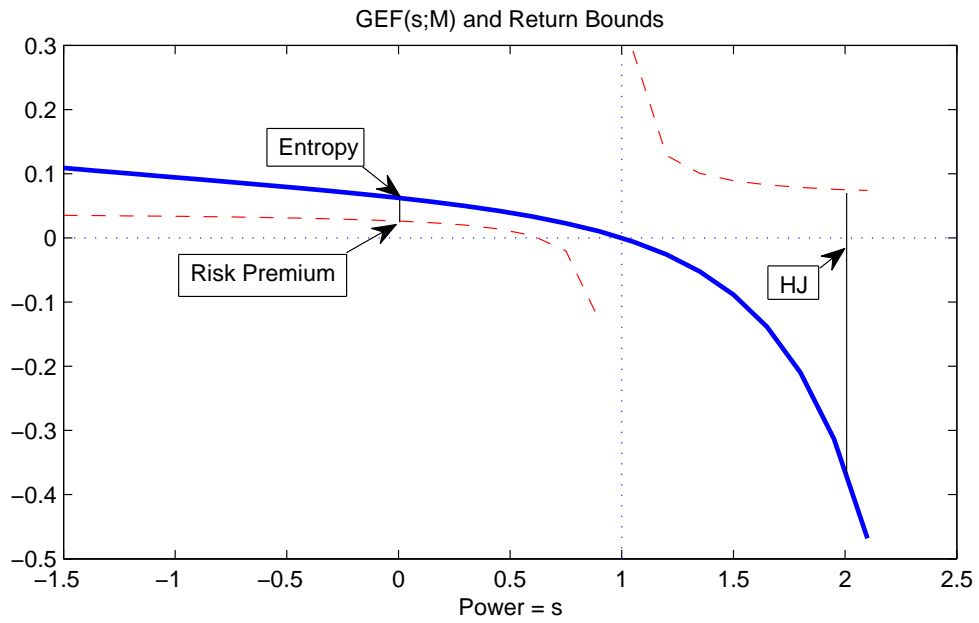


Figure 2: **A typical plot of  $GEF(s; M)$  and asset market bounds.** This figure shows the plot of a generic GEF and asset market bounds. The thick solid line depicts the GEF and the thin dashed line depicts asset market bounds.



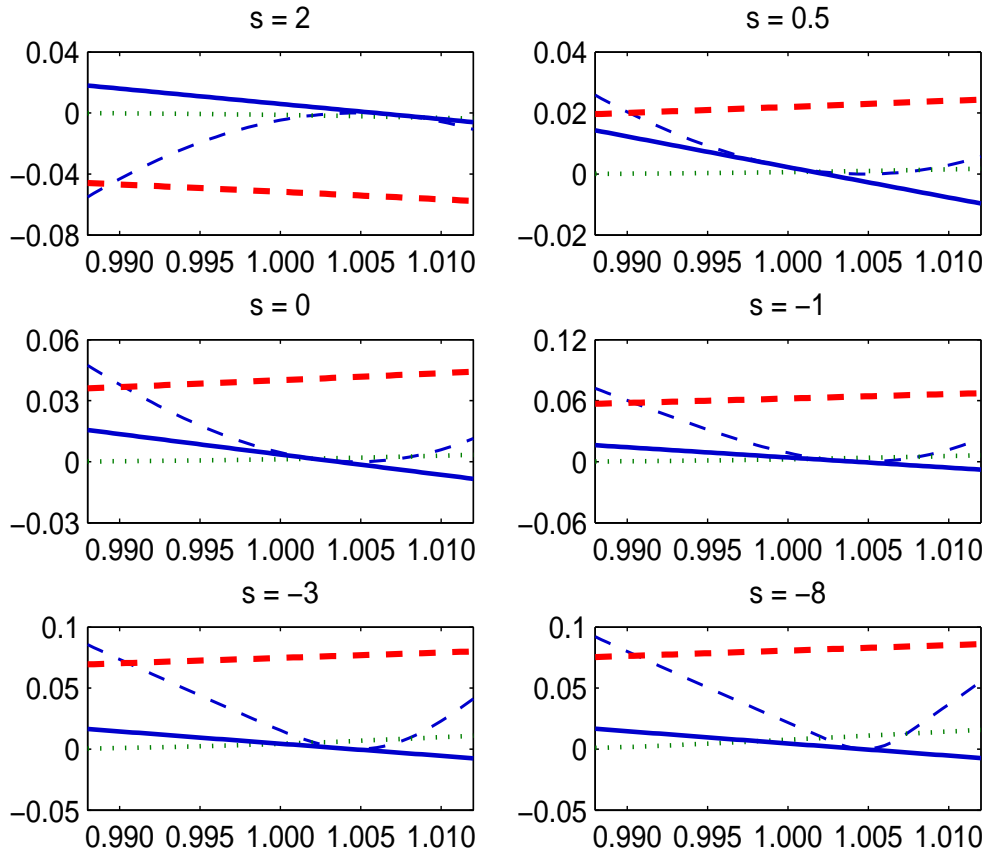


Figure 3: **Bounds implied by benchmark strategies and the index.** This figure plots the non-parametric bound frontiers (right hand side in inequality (11)) for the benchmark trading strategies and the market index across different hypothetical riskfree rates. The estimation is generally done, at each hypothetical riskfree rate, by conducting a nonlinear search on the optimal portfolio weight  $\alpha_S$  to either maximize or minimize the right hand side in inequalities (11) and (12). The solid line, thin dashed line, dotted line and thick dashed line depict the frontiers for the passive market strategy, active market strategy, ATM straddle strategy and 96%-OTM put option strategy, respectively. The passive market strategy simple sets  $\alpha_S$  at zero at every interest rate level and the active market strategy involves a search as described above.

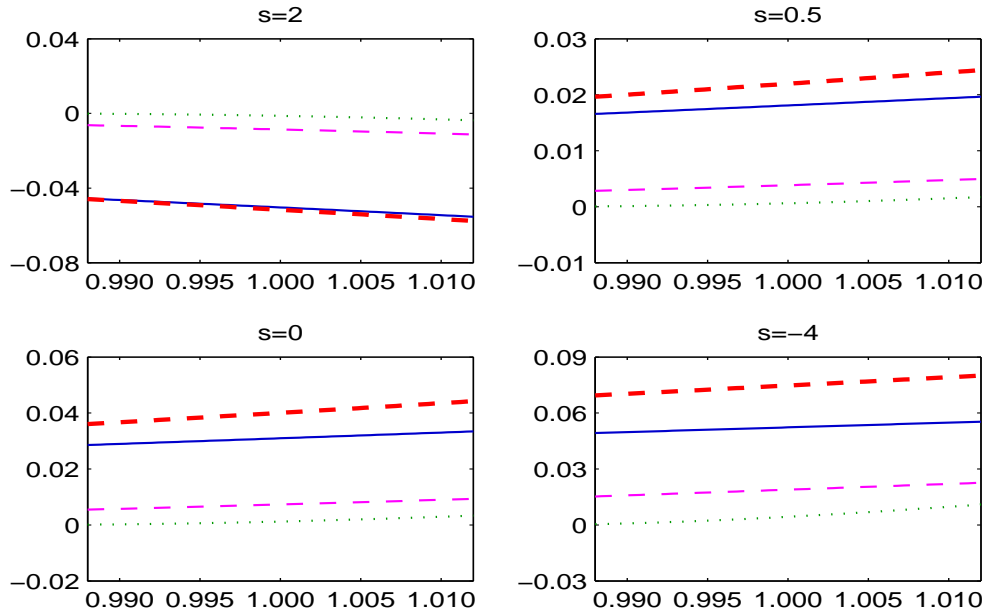


Figure 4: **Bounds implied by benchmark strategies and two alternative OTM put strategies.** This figure plots the non-parametric bound frontiers (right hand side in inequality (11)) for the benchmark trading strategies and two alternative OTM put option strategies across different hypothetical riskfree rates. The estimation is generally done, at each hypothetical riskfree rate, by conducting a non-linear search on the optimal portfolio weight  $\alpha_S$  to either maximize or minimize the right hand side in inequalities (11) and (12). The solid line, thin dashed line, dotted line and thick dashed line depict the frontiers for the 92%-OTM put option strategy, ATM put option strategy, ATM straddle strategy and 96%-OTM put option strategy, respectively.

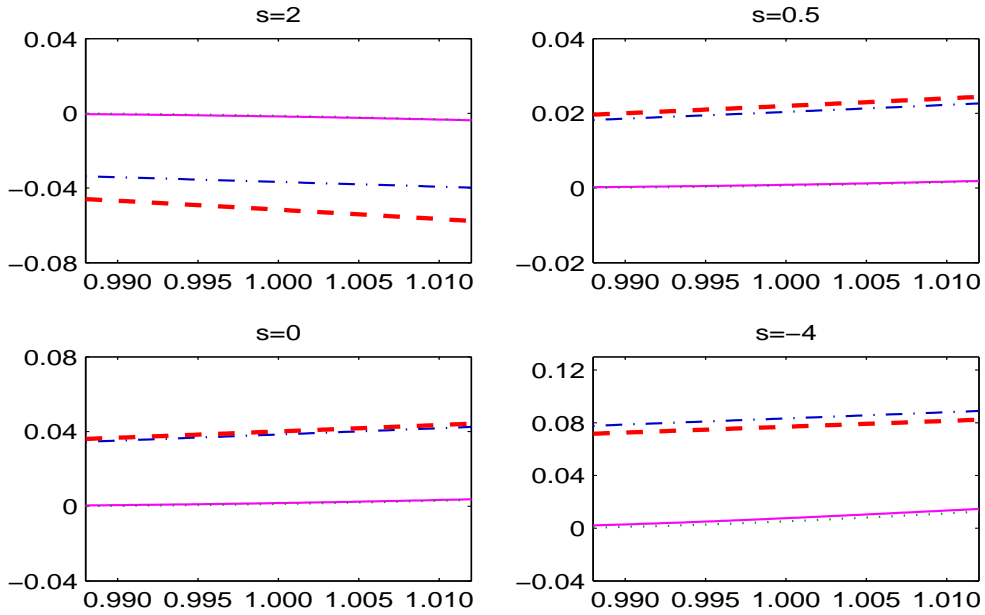


Figure 5: **Bounds implied by benchmark strategies and two crash-neutral strategies.** This figure plots the non-parametric bound frontiers (right hand side in inequality (11)) for the benchmark trading strategies and two crash-neutral strategies across different hypothetical riskfree rates. The estimation is generally done, at each hypothetical riskfree rate, by conducting a nonlinear search on the optimal portfolio weight  $\alpha_S$  to either maximize or minimize the right hand side in inequalities (11) and (12). Note that the two robust crash-neutral return series are used instead of the original full-sample series. The solid line, dot-dashed line, dotted line and thick dashed line depict the frontiers for the crash-neutral straddle strategy, crash-neutral put option strategy, ATM straddle strategy and 96%-OTM put option strategy, respectively.

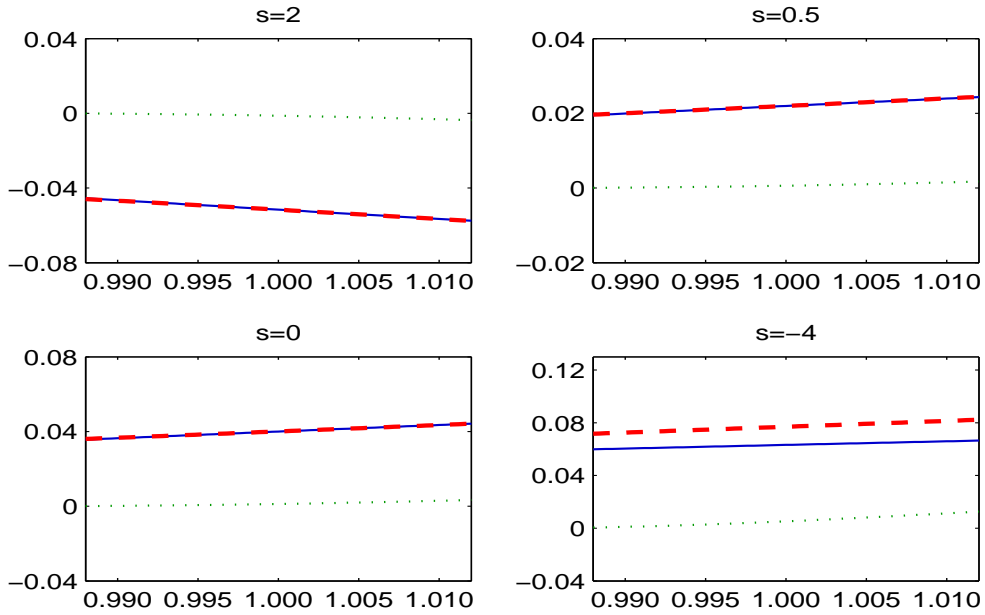


Figure 6: **Bounds implied by optimal and conservative benchmark strategies.** This figure plots the non-parametric bound frontiers (right hand side in inequality (11)) for the benchmark trading strategies and conservative benchmark strategies across different hypothetical riskfree rates. For the optimal benchmark strategies, the estimation is generally done, at each hypothetical riskfree rate, by conducting a non-linear search on the optimal portfolio weight  $\alpha_S$  to either maximize or minimize the right hand side in inequalities (11) and (12). The dotted line and the thick dashed line depict the frontier for the ATM straddle strategy and 96%-OTM put option strategy, respectively. The solid lines at  $s = 2, 0.5, 0, -4$  depict the interest-rate independent longing 50%, shorting -20%, shorting -35% and shorting -35% on the 96%-OTM put option strategies, respectively.

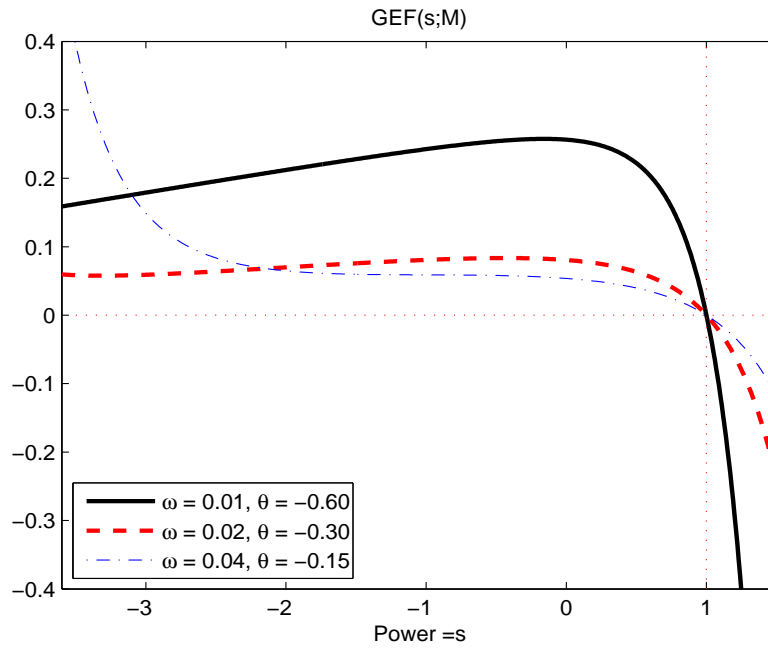


Figure 7: Generalized entropy function plots for three disaster models.

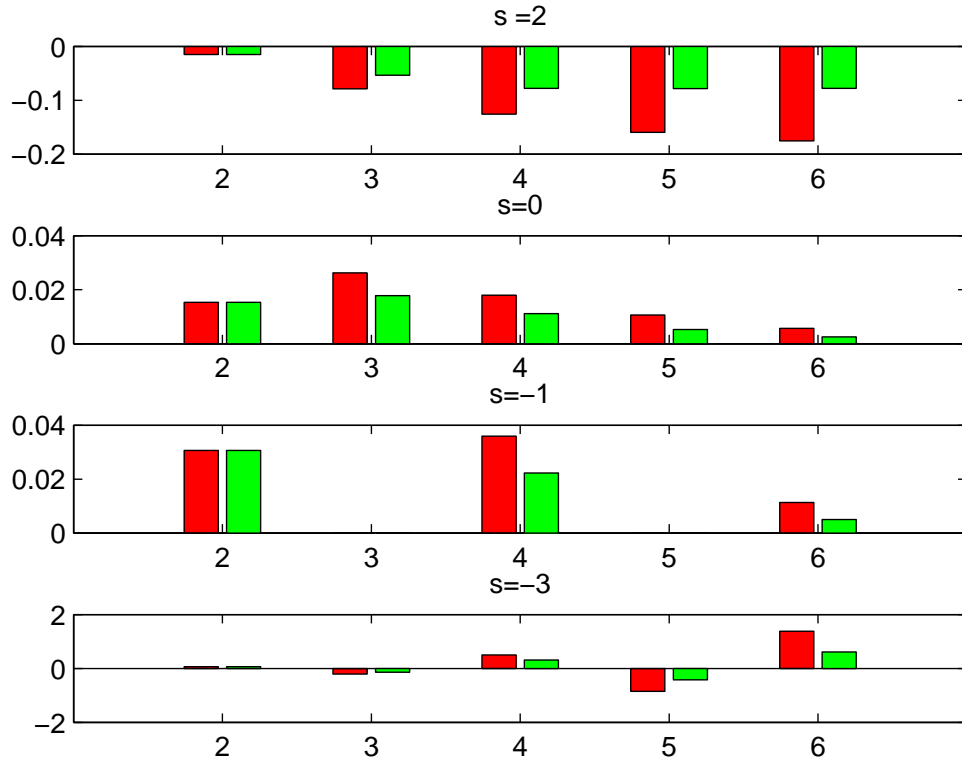


Figure 8: **Weighted cumulants for two disaster models** This figure displays the second to sixth weighted cumulants for the mild and light disaster model at  $s = 2, 0, -1$  and  $-3$ . The  $j$ -th weighted cumulant is defined as  $\frac{\kappa_j(\log M_{t+1})}{j!}(1 - s^{j-1})$  as in equation 13. The left (dark) bar and the right (light) bar measure the weighted cumulant for the mild and light disaster model, respectively.

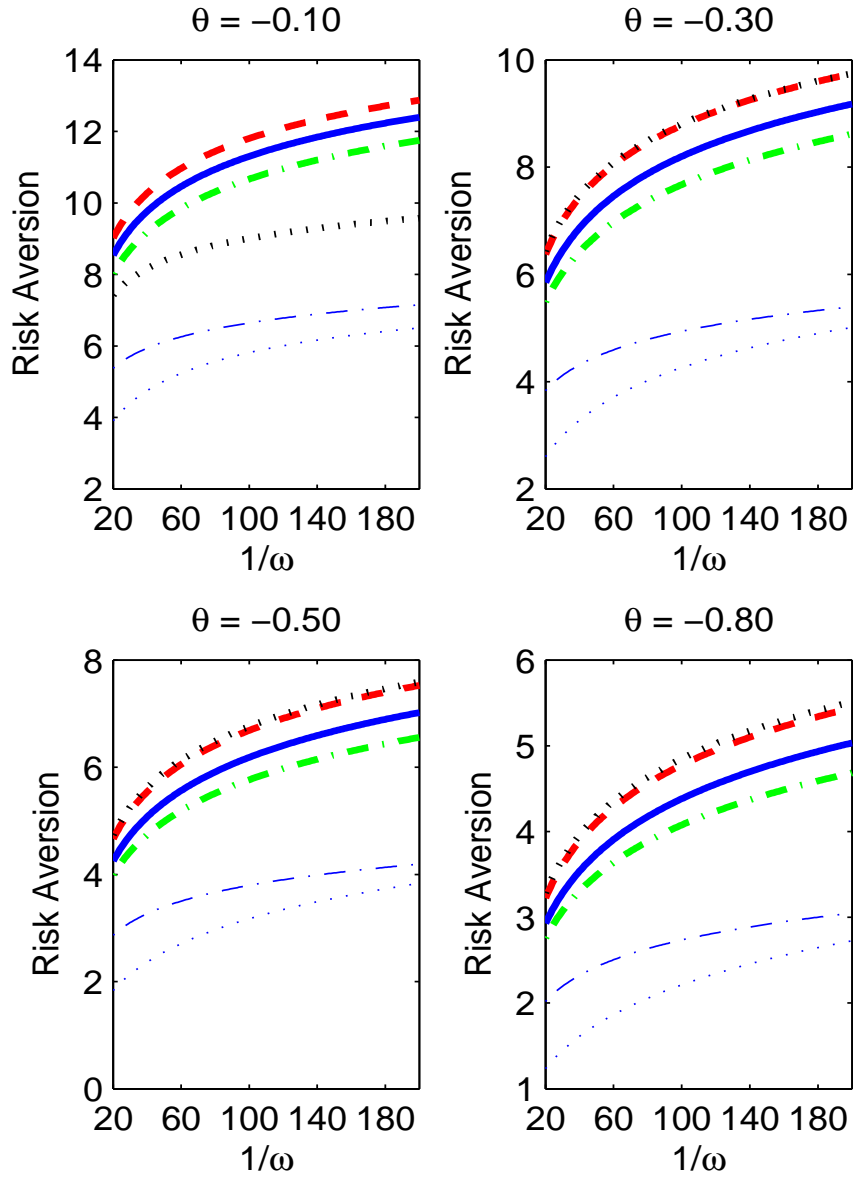


Figure 9: **Risk aversion bounds implied by index option return.** This figure shows the required risk aversion coefficients corresponding to different disaster frequency  $\omega$ , disaster size  $\theta$  and option return implied entropy bounds. The thin dotted line depicts the required risk aversion in generating a 4% annual equity risk premium. The thin dash-dotted line, thick dash-dotted line, thick solid line, thick dashed line and thick dotted line depict the required risk aversion coefficients in satisfying the entropy bounds at power  $s = 2, 0.5, 0, -1$  and  $-2$ , respectively.

## B Proofs

### B.1 Proof of Proposition 2.

The optimization problem we want to solve is

$$\begin{aligned} \sup_R \quad & E\left[\frac{R^{1-\gamma}}{1-\gamma}\right] \\ \text{s.t.} \quad & (1) E(MR) = 1, \\ & (2) R > 0. \end{aligned}$$

For simplicity, I succinctly denote  $\gamma(\delta)$  by  $\gamma$ . Moreover, to save space, I only solve the case when  $\gamma \in (0, 1)$ . For  $\gamma \in (1, \infty)$  the maximization problem is essentially a minimization problem and a similar proof follows. For  $\gamma \in (0, 1)$ , the maximization problem will be well-defined if all moments of  $M$  are assumed to exist. This is because

$$\begin{aligned} E(R^{1-\gamma}) &= E(R^{1-\gamma} M^{1-\gamma} M^{\gamma-1}) \\ &\leq [E([(MR)^{1-\gamma}]^{\frac{1}{1-\gamma}})]^{1-\gamma} \cdot [E(M^{\gamma-1})^{\frac{1}{\gamma}}]^\gamma \\ &= E(M^{\frac{\gamma-1}{\gamma}})^\gamma \end{aligned}$$

Note that I am using the same trick as in the proof of the new bounds. Also, for  $\gamma \in (1, \infty)$  a lower bound for  $E(R^{1-\gamma})$  exists so the corresponding minimization problem is also well-defined.

Let the state density function be  $f(s)$  and let the Lagrange multipliers associated with  $E(MR) = 1$  and  $R(s) > 0$  be  $\lambda$  and  $\mu(s)$ , respectively, then the Lagrange function is

$$\mathcal{L}(R(s), \lambda, \mu(s)) = \frac{1}{1-\gamma} \int R(s)^{1-\gamma} f(s) ds - \lambda \left( \int M(s)R(s)f(s) ds - 1 \right) - \mu(s)R(s)$$

It is easy to see that the objective function  $\frac{1}{1-\gamma} \int R(s)^{1-\gamma} f(s) ds$  is concave in  $R(s)$ . Additionally, the constraint  $\int M(s)R(s)f(s) ds = 1$  is linear in  $R(s)$ . Under these two conditions, the Kuhn-Tucker first-order conditions are both necessary and sufficient for a maximum of this problem. The first-order condition for the argument  $R(s)$  is

$$R(s)^{-\gamma} f(s) - \lambda M(s) f(s) - \mu(s) = 0$$

Since returns need to have a positive support, the Lagrange multiplier associated with the positivity constraint  $\mu(s)$  will be zero for every state. Assuming an everywhere positive  $f(s)$ , we arrive at the following solution for  $R(s)$

$$R(s) = [\lambda(1-\gamma)]^{-\frac{1}{\gamma}} M(s)^{-\frac{1}{\gamma}}. \quad (29)$$



To express  $\lambda$  as a moment of the pricing kernel, we can multiply both sides of equation (29) by  $M(s)f(s)$  and sum across states. This leave us with the following equation for  $\lambda$

$$[\lambda(1 - \gamma)]^{-\frac{1}{\gamma}} E(M^{1-\frac{1}{\gamma}}) = 1 \quad (30)$$

Combining equation (29) and (30), we get the optimal portfolio choice as a function of  $M$  only

$$\tilde{R} = M^{-\frac{1}{\gamma}} / E(M^{\frac{\gamma-1}{\gamma}}). \quad (31)$$

Note that by assumption  $M \in Q^{++}$ , so  $\tilde{R} \in \aleph^{++}$ . This validates the earlier step in setting  $\mu(s)$  to zero. Finally, by plugging the optimal choice  $\tilde{R}$  into the objective function, we have

$$U(M) = \frac{E(\tilde{R}^{1-\gamma})}{1 - \gamma} = \frac{[E(M^{\frac{\gamma-1}{\gamma}})]^\gamma}{1 - \gamma}. \quad (32)$$

Equation (31) and (32) give the optimal solution of this problem.

## B.2 Duality definition and proof

For an optimizing investor with a risk-aversion coefficient of  $\gamma$ , her optimization problem is

$$\begin{aligned} \sup_R \quad & E\left[\frac{R^{1-\gamma}}{1-\gamma}\right] \\ \text{s.t.} \quad & (1) E(MR) = 1, \\ & (2) R > 0. \end{aligned}$$

Denote the maximized objective function by  $U_{YL}(M)$  and the optimal choice variable by  $R_{YL}(M)$ , respectively. Note that they are both functionals on  $M$  and can be thought of as operators: they operate on any pricing kernel defined on  $\aleph^{++}$  and yield a value function and a choice return variable. Symmetrically, a Hansen-Jaganathan type of optimization on the  $\delta$ -th moment of the pricing kernel can be presented by

$$\begin{aligned} \inf_M \quad & \frac{[E(M^\delta)]^{\frac{1}{1-\delta}}}{1 - \gamma(\delta)} \\ \text{s.t.} \quad & (1) E(MR) = 1, \\ & (2) M > 0. \end{aligned}$$

where  $\gamma(\delta) = \frac{1}{1-\delta}$  is what I will term the dual parameter transformation. Similarly, let  $U_{HJ}(R)$  and  $M_{HJ}(R)$  be the associated functionals (operators). Then a duality between these two optimization problems is satisfied iff the following conditions hold:

$$\begin{aligned} R_{YL}(M_{HJ}(R)) &\equiv R \\ M_{HJ}(R_{YL}(M)) &\equiv M \end{aligned}$$

In words, these relationships say: 1. The pricing kernel that satisfies the HJ problem with a given return  $R$  is the only kernel that can yield an optimal choice of  $R$  in my optimization problem; 2. The return that is the optimal choice under my optimization scheme for a given pricing kernel  $M$  is the only return that can yield  $M$  as the optimal choice in the HJ problem. If two operators satisfy these above duality conditions, then inverse operators can be defined straightforwardly as

$$\begin{aligned} R_{YL}^{-1}(R) &\equiv M_{HJ}(R) \\ M_{HJ}^{-1}(M) &\equiv R_{YL}(M) \end{aligned}$$

Given the duality definition, it is easy to see that HJ and my optimization are indeed dual problems. To see this, we only need to work out  $M_{HJ}(R)$ . Similar to Proposition 1, it can be shown that

$$M_{HJ}(R) = C(R) \cdot R^{\frac{1}{\delta-1}} \quad (33)$$

where the normalizing constant  $C(R)$  is equal to  $1/E(R^{\frac{\delta}{\delta-1}})$ . By plugging the formulae in equation (31) and (33) into the duality conditions, it is readily seen that they are satisfied.

### B.3 Proof of Proposition 3.

I prove by giving an example. I construct a sequence of pricing kernels that can all price a riskless bond but have either explosive or degenerate  $\delta$ -th moment in the limit.

Let the state space be  $(0, 1)$  and let  $X$  be a random variable that is uniformly distributed on  $(0, 1)$ :  $X \sim U(0, 1)$ . Let  $\{M_n\}_{n=1}^{\infty}$  be a sequence of pricing kernels that are defined by

$$M_n = \begin{cases} n - \alpha_n & \text{if } X \in (0, \frac{1}{n}), \\ \frac{\alpha_n}{n-1} & \text{if } X \in [\frac{1}{n}, 1) \end{cases} \quad (34)$$

where  $\{\alpha_n\}_{n=1}^{\infty}$  is a sequence that satisfies  $\alpha_n < n$  and  $\frac{\alpha_n}{n} \rightarrow 0$  (For simplicity,  $\alpha_n$  can be set at the constant one). Pricing kernels defined in such a way can be understood as describing economies with rare disasters. Rare events happen with a probability  $\frac{1}{n}$  and the state price is high in disaster states. Note that a one-period riskless bond has a gross return of one, since  $E(M_n) = 1$  for any  $n$ . Notice that  $E(M_n^\delta)$  goes to  $\infty$  since

$$E(M_n^\delta) \geq (n - \alpha_n)^\delta \frac{1}{n} \rightarrow \infty$$

for any  $\delta > 1$ . However, if a riskless bond is the only security, then return moments are all equal to one. Therefore, no upper bound can be imposed on  $E(M^\delta)$ . Similarly, for  $\delta \in (0, 1)$ , we have

$$E(M_n^\delta) = (n - \alpha_n)^\delta \frac{1}{n} + \left(\frac{\alpha_n}{n-1}\right)^\delta \left(1 - \frac{1}{n}\right) \rightarrow 0,$$

so no lower bound (except the trivial zero bound) exists for  $\delta \in (0, 1)$ . Lastly, if  $\delta \in (-\infty, 0)$  then no upper bound exists.