Index Option Returns and Generalized Entropy Bounds

Yan Liu*

Texas A&M University, College Station, TX 77843 USA

First draft: July 14, 2012
This revision: December 20, 2014

Abstract

I develop a continuum of new nonparametric bounds. They stem from the solution of an optimization problem that is complementary to the Hansen and Jaganathan (1991) approach and are shown to complete the nonparametric bound universe the literature has so far discovered. Through the lens of these bounds, I estimate rare event distributions using index option returns. Standard disaster models and their perturbations are shown unable to meet the bounds implied by simple static option trading strategies. My results suggest more sophisticated modeling of disaster models in order to reconcile with the index option data.

Keywords: High-order moments, Pricing kernel, Rare disasters, Index options, Nonparametric bounds, Model diagnosis, Pseudo problem

---

* Current Version: December 20, 2014. First posted to SSRN: July 14, 2012. Send correspondence to: Yan Liu, Mays School of Business, Texas A&M University, College Station, TX 77845. Phone: +1 919.428.1118, E-mail: y-liu@mays.tamu.edu. I appreciate the comments of Hengjie Ai, Ravi Bansal, Tim Bollerslev, Yong Chen, Ian Dew-Becker, John Graham, Lars Hansen, Campbell Harvey, Hwagyun Kim, Adam Kolasinski, Ivan Shaliastovich, Lukas Schmid, Michael Stutzer, Caroline Zhu as well as seminar participants at seminars at Duke and Texas A&M University.

1 Introduction

The asset market generates risk and return characteristics that continuously challenge our thinking. To rationalize market abnormalities, economists create models under a few generally accepted economic principles. These models are constantly scrutinized and possibly rejected with the advent of new empirical findings, and new models are again proposed to accommodate these new findings.

In this process, a few important diagnostic tools have been developed by the literature to restrict the behavior of a candidate model. Under the basic no-arbitrage condition, Hansen and Jagannathan (HJ, 1991) construct a bound on the second moment of the stochastic discount factor for a given menu of assets. This nonparametric bound provides a simple way to summarize asset market data, and to help screen candidate models. Snow (1991) extends their work by showing how to bound higher moments of the pricing kernel. Stutzer (1995) proposes an information bound that minimizes the Kullback-Leibler Information Criterion. 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 rely only on the fundamental no-arbitrage condition and provide unique lens through which we can characterize asset market data, diagnose existing asset pricing models, and design new models to explain a larger set of empirical regularities.

I contribute to this literature by first providing a unifying theory on nonparametric bounds. Starting from the no-arbitrage condition alone, I show the existence of a continuum of bounds that restrict the δ-th norm of the pricing kernel, with $\delta \in (-\infty, 0) \cup (0, 1)$. Next, I show that 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 associated with a pricing kernel and show that agents' portfolio choice problems based on this augmented return space impose constraints on moments of the pricing kernel. In a strict duality sense, I show that my approach is complementary to the Hansen and Jagannathan approach. Lastly, I characterize the nonparametric bound universe and discuss its exhaustiveness. In particular, I show that the well-known entropy bound is a special case of the new bound system.

To facilitate the application of my bounds, I propose a new metric termed the *generalized entropy function*. It is a natural generalization of the basic entropy (Stutzer, 1996, Backus, Chernov and Zin, 2011, and Hansen and Sargent, 2008) and encodes all the distributional information of a pricing kernel. I use the system of new nonparametric bounds to make inference on the generalized entropy function of the pricing kernel. Through moment expansions, similar to Martin (2008) and Backus, Chernov and Martin (BCM, 2011), I show how different moments of the pricing kernel contribute to the generalized entropy function and, more importantly, how weighted asset return moments provide information on the entropy function.

For the empirical application, I try to make better inference on disaster models (Rietz, 1988, Barro, 2006, Barro and Ursua, 2008, and Barro et al., 2009). The inherent difficulty for the disaster literature is how to measure events that only happen rarely (i.e., the pseudo problem). Similar to BCM, who evaluate disaster models' performances against index options along several metrics, I also use option data to infer the tail information in the pricing kernel. Unlike BCM, I consider static trading strategies that involve option returns and rely on the newly developed bounds to make inference.

My approach is different from and advantageous over BCM in several aspects. First, no assumption is made on the link between macro fundamentals and the asset market. As a result, my results are robust to model misspecifications. Second, instead of fitting a parametric model and using it to summarize the option cross-section, I take the realized option returns as given and study their implications on the pricing kernel. This again alleviates the goodness-of-fit concern of empirical option pricing models. Finally, a formal hypothesis testing framework is developed to evaluate model performances. This allows one to consider multiple assets simultaneously and generates statistical significance for different model configurations.

Turning to the empirical findings, I first document the unique moment characteristics of option strategy returns from a nonparametric bound perspective. Bounds implied by out-of-the-money (OTM) puts universally dominate bounds implied by either the market index or straddles. This highlights the pricing of jump risks in put options, and is precisely the type of information we need to bear on models with tail risks. Next, I use option implied bounds to confront standard rare disaster models. In accordance with the macro-finance literature, I mark up the permissible parameter region in which all nonparametric bounds are simultaneously satisfied. I show that the new bounds provide the sharpest restrictions on the model compared to existing bounds. Lastly, to take statistical uncertainty into account, I develop a formal testing framework that accommodates multiple assets and different types of bounds. In this framework, I reject the benchmark disaster model and a few alternative specifications.[1] Despite the rejection, the model's ability of approaching asset market bounds does look impressive. Taken as a whole, my results suggest more sophisticated modeling of disaster models — possibly through the time-dependency of disaster occurrences along the lines of Barro and Ursua (2008) and Watcher (2013, 2014) — to reconcile with option returns.

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 the concept of the generalized entropy. 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 Nonparametric Bounds

Hansen and Jaganathan (1991), Snow (1991), Stutzer (1995), Bansal and Lehmann (1997) and Alvarez and Jermann (2005) derive nonparametric bounds under the no-arbitrage condition. Depending on the functional form of the non-linear transformations of the pricing kernel, strong no-arbitrage condition may be needed to guarantee the meaningfulness of the bounds. For instance, the entropy bound (Bansal and Lehmann, 1997, Alvarez and Jermann, 2005) relies on a logarithmic transformation of the pricing kernel. As a result, it only makes sense 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 as the pricing kernel is raised to the second power.[2] To be specific about the asset pricing environment and facilitate discussions, I first introduce some notations that will be used throughout the paper.

[1]More specifically, I reject the benchmark disaster model in Barro (2006) and a few other parameterizations at 5% significance.

[2]The weak no-arbitrage condition requires the pricing kernel to be nonnegative. Hansen and Jaganathan (1991, 1994) include detailed discussions on the corresponding variance bounds when either strict positivity or nonnegativity applies to the hypothesized pricing kernel.

Let $\aleph$ be the collection of gross returns. Conceptually, it includes returns of all tradable 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.,

$$E[MR] = 1, \forall R \in \aleph. \tag{1}$$

Hansen and Jaganathan (1991, 1994) define Q^{++} and Q^{+} to be the set of strictly positive and nonnegative pricing kernels, respectively. Similarly, I define $\aleph^{++} = \{R : R \in \aleph \text{ and } R > 0 \text{ with probability one}\}$ and $\aleph^{+} = \{R : R \in \aleph \text{ and } R \geq 0 \text{ with probability one}\}$. For the same reason as for the entropy bound, I generally require $M \in Q^{++}$ and $R \in \aleph^{++}$ to derive nonparametric bounds. Therefore, I impose these two restrictions for the rest of this section unless otherwise stated. Notice that $M \in Q^{++}$ is an implication of the strong no-arbitrage condition. Also, $R \in \aleph^{++}$ is a weak condition for the gross returns of primitive assets to satisfy due to limited liability. Both assumptions are rather mild and facilitate the applications of nonparametric asset pricing bounds. Nonetheless, it is possible for a portfolio of assets to have negative gross returns in some states if the portfolio has excessive short positions on certain primitive assets. The theories I develop will not apply to these portfolios.

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 $p > 1, q > 1, \frac{1}{p} + \frac{1}{q} = 1$.

Proof. The proof involves simple manipulations of the Euler equation and the application of the 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 Euler equation $E(MR) = 1$. $\square$

The above derivation is different from the proof of the HJ variance bound or, more generally, Snow's high-moment bounds. These bounds place restrictions on the pth moment of the pricing kernel, with p greater than one. As a result, direct applications of the Cauchy-Schwarz inequality or the Hölder's inequality on the no-arbitrage condition suffice.[3] In contrast, for fractional powers of the pricing kernel, the strategy in Proposition 1 is to first create a power transformed gross return (i.e., $R^{1/p}$) to match the pricing kernel. By applying the Hölder's inequality, the $\frac{1}{p}$th power of the product MR is raised to the pth power, which simply becomes the product MR and drops out of the equation via the Euler equation.

Proposition 1 bounds the $\frac{1}{p}$th moment of the pricing kernel by the $-\frac{q}{p}$th moment of a return. As p runs from one to $+\infty$, $1/p$ covers the interval $(0, 1)$. 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 symmetry of M and R in the no-arbitrage condition, we can obtain a continuum of bounds on negative moments of the pricing kernel by switching M with R in Proposition 1.

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

Letting $\delta = \frac{1}{p}$, we rewrite the bounds in Proposition 1 to make them consistent with the notations in Corollary 1:

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

Combining Corollary 1 and equation (2), we find lower bounds on $E(M^{\delta})$ when $\delta < 0$ and upper bounds when $\delta \in (0, 1)$. The change of direction for the inequalities at $\delta = 0$ seems cumbersome, 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.

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 correspond-

[3]For an introduction on Hölder's inequality, see Casella and Berger (2001), Chap.4.

ing to each power δ, similar to Snow (1991) and Bansal and Lehmann (1997). In particular, define $\rho(\delta)$ as

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

then $\rho(\delta)$ gives the sharpest lower (upper) bound on $E(M^{\delta})$ when $\delta \in (-\infty, 0)$ ($\delta \in (0, 1)$). Instead of searching for the optimal return, Bakshi and Chabi-Yo (2014) show how to combine bounds based on multiple assets. Their focus is on combining information from different assets. Mine is on different moment restrictions on the same set of returns.

2.2 Interpreting Bounds

What are the economic stories behind these bounds? What does a negative moment of the return measure? In particular, what do the two sides of these inequalities measure? I provide a utility-based interpretation of these bounds.

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

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

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

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

It is crucial to see the difference between $\aleph^{++}$ and $\aleph^{**}(M)$. The former contains returns of assets that are tradable in the market while the latter contains all positive returns that satisfy the no-arbitrage condition. In other words, $\aleph^{++}$ includes whatever the market has while the potentially much larger $\aleph^{**}$ includes what the market could have. The difference between $\aleph^{++}$ and $\aleph^{**}$ measures the degree of market completeness.

Within this augmented return space $\aleph^{**}(M)$, I seek to solve the portfolio choice problem for an agent with a unit endowment and a risk aversion of $\gamma(\delta)$. This optimization problem can be written as

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

The maximized utility $U_\delta(M)$ depends on the discount factor M, whose information is embedded 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

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

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

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 choice $\tilde{R}_\delta(M)$ is solved as a function of the Lagrange multiplier associated with $E(MR) = 1$, viewed as a budget constraint. Second, the Lagrange multiplier itself is solved using the no-arbitrage condition. □

Now the economic meanings of the new bounds stand out. To ease interpretation, we can rewrite the bounds in Corollary 1 and equation (2) as

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

By Proposition 2, the quantity on the left-hand side is the maximized utility over the augmented return space $\aleph^{**}(M)$ 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 $\aleph^{**}(M) = \aleph^{++}$ or, as a weaker requirement, the optimal choice $\tilde{R}_\delta(M)$ is actually tradable, i.e., $\tilde{R}_\delta(M) \in \aleph^{++}$.

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

From a methodological perspective, it is interesting to compare my approach to that of HJ (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 payoffs. 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. Subsequently, by limiting the asset space to the group of tradable assets, the agent's real-world objective function dictates a lower bar that the starting candidate pricing kernel has to satisfy. In fact, in a strict duality sense, these two approaches are complementary to each other. I rigorously define the duality concept and prove it in the appendix. Intuitively, HJ's approach is more transparent when a certain moment of the pricing kernel (e.g., variance, Sharpe ratio, etc.) is the focus and my approach is more intuitive when a return moment (e.g., a CRRA investor's objective function) is the interest.

Another benefit of developing the utility-based framework in this section is to motivate the development of alternative nonparametric bounds. Replacing the CRRA agent with a CARA agent, I show how to derive the well-known information bound in Stutzer (1995) in the appendix. Stutzer (1995) motivates his bound using the Kullback-Leibler distance in information theory. It is interesting to see how my framework helps rationalize asset pricing restrictions that are developed in a completely different setup.

2.3 Characterizing the Non-parametric Bound Universe

Thus far, I have established bounds for various moments of the pricing kernel, with zero (i.e., $\delta = 0$) being 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 whether 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 utilities converge to the log utility. I only show how bounds in Proposition 1 imply the entropy bound. Essentially the same proof follows for bounds in Corollary 1. I start by rescaling 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,[4] the left-hand side is easily seen to converge to $E(\log(M))$ using the L'Hôspital's rule. A careful application of the rule to the right-hand side will deliver $-E(\log(R))$ as the limit. $\square$

We now have all the pieces to summarize the non-parametric bound universe that the literature has discovered. Figure 1 shows a diagram of this bound universe. When the power s equals one, the expected marginal rate of substitution is bounded within $[\frac{1}{\max R}, \frac{1}{\min R}]$ for a generic $R \in \aleph^{++}$. This seemingly informative bound becomes redundant in the presence of a risk-free rate R_f since $E(M) = 1/R_f$. Starting from $s = 1$ and going right, one encounters the spectrum of Snow's high-moment bounds and the HJ bound is sitting at $s = 2$. Going left, one sees the continuum of bounds I just developed, and the BL/AJ entropy bound fills the hole at $s = 0$. It is intriguing to see the symmetric pattern of these bounds around $s = 1$, particularly in light of the order by which they are discovered by the literature.[5]

[4]We need conditions on moments of M and R to be able to exchange limits and expectations. Dominated convergence will suffice. See Davidson (1994) Part IV for some specific conditions.

[5]Recent papers by Almeida and Garcia (2012, 2013) proposes similar nonparametric bounds. However, there are important differences between their work and mine. First, they follow the Hansen and Jaganathan (1991) approach but use a new objective function to derive bounds. One can insert their optimal solution into the objective function to obtain bounds that are similar to mine but impose a known risk-free rate. I rationalize my bounds using a utility optimization framework, which is a systematic approach to uncover non-parametric bounds (e.g., I show how to derive Stutzer (1995)'s information bound in my framework.) Second, I relate my bounds to the basic entropy

Lastly, we ask if the bound system is complete. This is more than a technical question because we do not want to leave out important information on the pricing kernel that can be learned through asset returns. In particular, given the existence of these one-sided inequalities for essentially all the moments of the pricing kernel, one may wonder whether other bounds, possibly pointing in opposite directions, can further enrich the bound universe. 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 all price a certain asset but has an unbounded limit for a given moment. I leave this proof to the appendix. □

2.4 Discussion

The new continuum of bounds can be extended along several dimensions. First, it can be adapted to study the dynamic behavior of the pricing kernel. Let $M_{t,t+n} = M_{t+1}M_{t+2}\ldots M_{t+n}$ be the time aggregated pricing kernel and $R_{t,t+n} = R_{t+1}R_{t+2}\ldots R_{t+n}$ be a generic multi-period return. Long-horizon asset returns provide bounds on 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 kernel. Different moments shed light on different forms of dynamic dependency. For instance, one natural way to scale an n-period pricing kernel is to take the fractional power $1/m$ on the time aggregated pricing kernel. A bound on the scaled kernel is given by

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

bound, which attracts a lot of attention in the recent asset pricing literature. I generalize the entropy definition and demonstrate how the generalized entropy encodes high-order moment information. These analytical tools allow me to dissect and evaluate state-of-the-art asset pricing models that explore tail information.

When $m \to \infty$, the above bound (properly scaled) converges to the basic multi-period entropy bound. Backus, Chernov and Zin (2011) use this to study time-dependency in discount rates. Setting m at n, the left-hand side becomes

$$E[\exp(\frac{1}{n}\sum_{j=1}^{n}\log M_{t+j})],$$

so information about the average of the log pricing kernel is revealed. Decomposing the pricing kernel into a permanent and a transitory component (Alvarez and Jermann, 2005), the basic entropy leaves the expectation of the permanent component intact while allows the transitory component to decay at a rate of n. My bounds, by allowing one to vary m, shed light on how fast the transitory component decays. Liu (2013) builds on this insight to diagnose dynamic asset pricing models using the generalized entropy bounds.

Second, conditioning information can be incorporated to sharpen bounds on unconditional moments (Bekaert and Liu, 2004, Gallant, Hansen and Tauchen, 1990, and Ferson and Siegel, 2003). Notice that simply adding instruments to the conditional Euler equation (Hansen and Jaganathan, 1990) is different from using returns that are generated by a dynamic trading strategy. The utility-based interpretation of my bounds allow me to take the latter approach. Liu (2013) uses conditioning information to differentiate key predictability assumptions in leading asset pricing models.

3 Bound Informativeness

What do we gain by looking at the bounds for different moments of the pricing kernel? I perform a dissection of the bound system. Much like a doctor performing surgery, I need a "surgical knife" that works on asset pricing 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). I show that it is both economically meaningful and analytically tractable. Equipped with this tool, I apply the cumulant-expansion technique (Backus, Chernov and Martin, 2011 and Martin, 2008) to examine both sides of a bound. Insights are provided on bound informativeness.

3.1 A Useful Quantity

The recent asset pricing literature proposes a convenient measure to study the link between the pricing kernel and asset returns (Bansal and Lehmann, 1997, Alvarez and Jermann, 2005, Backus, Chernov and Martin, 2011, and Martin, 2011).[6] The entropy is used to measure pricing kernel dispersion and is bounded below by the continuously compounded risk premium:

$$L(M) \equiv \log E(M) - E(\log M) \geq E(\log R) - \log(R_f), \tag{8}$$

where $R \in \aleph^{++}$ is an arbitrary return and $R_f = 1/E(M)$ is the gross risk-free rate, assuming one exists. Researchers rely on the entropy bound to gauge the amount of dispersion that an asset pricing model has to generate. However, considering the newly developed continuum of bounds that are relatives of the entropy bound, one would expect to gain additional insights by using alternative 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 study 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:

$$GEF(s; M) \equiv \log E(M) - \frac{1}{s} \log E(M^s) \tag{9}$$

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

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

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

[6] 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).

numeraire invariant. This is an appealing property empirically because we do not need to worry about the adjustment between a nominal and a real pricing kernel.

Globally, as a function defined over the real line, the GEF has several attractive features that will facilitate its applications. First, it can be shown to be an everywhere concave function. Second, GEF is pivotal around $(1, 0)$ in the sense that every GEF has to pass $(1, 0)$ on the two-dimensional plane. This gives a fixation point to anchor the GEF's corresponding to different pricing kernels. Finally, in the familiar lognormal case, $GEF(s; M) = (1 - s)\sigma_M^2/2$ where σ_M^2 is the variance of the log pricing kernel.

The bound system developed in the previous section can be brought in to restrict $GEF(s; M)$. The implied restrictions can be shown as:

$$GEF(s; M) \geq \frac{s-1}{s} \log E(R^{\frac{s}{s-1}}) - \log(R_f), \forall s \in (-\infty, 1). \tag{10}$$

Notice how the two types of bounds in Corollary 1 and equation (2) nicely line up with each other in terms of the directions of inequalities. The undesirable flip in direction at zero disappears once we introduce the generalized entropy function. When $s > 1$, Snow's continuum of high-moment bounds imply

$$GEF(s; M) \leq \frac{s-1}{s} \log E(R^{\frac{s}{s-1}}) - \log(R_f), \forall s \in (1, +\infty). \tag{11}$$

Figure 2 plots a generic GEF with asset market bounds.

Given the convenience offered by the GEF, for the rest of the paper I will focus on bounds given in the form of (10) or (11) unless otherwise specified. I will refer to the system of bounds given in (10) as the *generalized entropy bounds* or, with a slight abuse of terminology, simply entropy bounds.[7] The bounds in (11) are termed the high-moment bounds.

3.2 Expanding the GEF

The Cumulant-Generating Function (CGF) is another recently developed tool to study higher order moments of the pricing kernel (Backus, Chernov and Martin, 2011

[7]Sometimes, to emphasize the difference between the new system of bounds (i.e., the generalized entropy bounds) and the original entropy bound proposed by Bansal and Lehmann (1997) and Alvarez and Jermann (2005), I refer to the original entropy bound as the basic entropy bound.

and Martin, 2008). By Taylor-expanding the log expected pricing kernel into a power series, it shows how higher order moments contribute to the overall entropy. Backus, Chernov and Martin (2011) use the basic entropy as a measure of dispersion and compare a disaster model and an empirical option pricing model through the entropy bound. Assuming a representative agent framework and an iid consumption growth process, Martin (2008) links macro fundamentals (e.g., the wealth-to-consumption ratio) to moments of the consumption growth and performs a calibration exercise.

I contribute to this literature by showing that asset returns provide valuable information about the entire CGF — not just at zero, which corresponds to the basic entropy. This significantly strengthens the link between macro fundamentals and asset market returns, and can potentially help us better distinguish candidate models.

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

$$\begin{aligned} GEF(s;M) &= \sum_{i=1}^{\infty} \frac{\kappa_i(\log M)}{i!} - \frac{1}{s} \sum_{i=1}^{\infty} \frac{\kappa_i(\log M)}{i!} s^i \\ &= \sum_{i=2}^{\infty} \frac{\kappa_i(\log M)}{i!}(1 - s^{i-1}) \\ &= \frac{\kappa_2(\log M)}{2!}(1-s) + \frac{\kappa_3(\log M)}{3!}(1-s^2) + \\ &\quad \frac{\kappa_4(\log M)}{4!}(1-s^3) + \frac{\kappa_5(\log M)}{5!}(1-s^4) + \ldots. \end{aligned} \tag{12}$$

The first two lines Taylor-expand the two parts in the GEF and group similar terms. The last two lines explicitly write out the first few terms in the 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 (ν_1) and excess kurtosis (ν_2) through: $\nu_1 = \kappa_3(\log M)/[\kappa_2(\log M)]^{\frac{3}{2}}$ and $\nu_2 = \kappa_4(\log M)/[\kappa_2(\log M)]^2$ (see Backus, Chernov and Martin, 2011).

The expansion of the GEF reveals that cumulants are weighted by polynomials of s. 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, the GEF puts different weights on different moments. In this way, the GEF conveys information about all the moments of the pricing kernel. In particular, when evaluated at $s = 0$, the GEF equals the basic entropy, which is simply the sum of $\{\kappa_i(\log M)/i!\}_{i=1}^{\infty}$.

I argue that the GEF is especially useful in teasing out 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. At the same time, even moments will mechanically increase as well with the presence of extreme observations 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 basic entropy expansion reflect the inherent asymmetry in jumps. However, given that all moments are equally weighted in the basic entropy expansion, it is difficult to single out the contribution from odd/even moments. I suggest taking large negative s values. A large negative s makes the weights associated with even moments positive and those associated with odd moments negative. Thus, a "net" jump effect is singled out by taking the difference between odd and even moments. In fact, at $s = -1$ odd moments disappear completely from the generalized entropy and we are left with even moments alone. The empirical study of the paper in later sections confirms the usefulness of the generalized entropy bounds at negative powers.

Similar expansions can be applied to returns on the right-hand side of bounds. This is important as it gives us guidance on the selection of the most informative assets. I cumulant-expand the right-hand side of equation (10) as:

$$\begin{aligned} GEF(s;M) &\geq \frac{s-1}{s}\sum_{i=1}^{\infty}\frac{\kappa_i(\log R)}{i!}\left(\frac{s}{s-1}\right)^i - \log R_f \\ &= \underbrace{[E(\log R) - \log R_f]}_{\text{Risk premium}} + \underbrace{\sum_{i=2}^{\infty}\frac{\kappa_i(\log R)}{i!}\left(\frac{s}{s-1}\right)^{i-1}}_{\text{High-order moments}}, \end{aligned} \quad (13)$$

where $\kappa_i(\log R)$ is the i-th return cumulant.

The basic entropy bound, by setting $s = 0$, ignores the term $\sum_{i=2}^{\infty}\frac{\kappa_i(\log R)}{i!}(\frac{s}{s-1})^{i-1}$, which includes high-order moments — including the variance — of returns. My bounds take these moments into account and, by allowing one to vary the power s, help differentiate the contribution from different high-order moments. For instance, 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 provide the tightest bounds, we need returns that possess excess (positive) skewness and kurtosis. Option strategy returns meet these requirements (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 place on the discount factor.

4 Option Market Bounds and Rare Disaster Models

Tail information, long recognized as a potential source to generate economic risk premiums (Rietz, 1988), has recently been elevated to quantitatively explain asset market abnormalities. Barro (2006), Barro and Ursua (2008) and Barro et al. (2009) estimate the tail distribution of the US consumption growth by looking at international macroeconomic data. Based on the exchange economy and representative agent framework, they argue that the calibrated rare event distribution can explain key moments of US asset returns, particularly the equity risk premium. Gabaix (2009), Wachter (2013, 2014) and Gourio (2008) extend the basic disaster model to account for other salient features of asset markets.

At the heart of the rare disaster literature is the so-called *pseudo* problem: given the rare occurrence disasters, one cannot estimate their distributions accurately based on 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, thereby indirectly estimating the disaster distribution. 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 infer the distribution of the consumption growth. I also consider index options but take a different approach. In particular, I use the newly developed nonparametric bounds to study the tail distribution of macro fundamentals. My approach has several advantages. First, it is based on the basic no-arbitrage condition alone and thus free from various sources of model misspecifications. For instance, we do not need to specify how the cash flow varies with the consumption growth, which in itself is empirically challenging. Second, no parametric model is needed to fit the entire 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 features the simultaneous testing of several bounds. I show the discriminatory power of the generalized entropy bounds.

4.1 Data description

For the empirical analysis, I use monthly data on the S&P 500 index, the associated index options and the risk-free rate. The risk-free rate is from Kenneth French's on-line data library. The full sample for the market returns runs from July 1926 to December 2011. The shorter sample, which coincides with the span of the option data from OptionMetrics, is from January 1996 to December 2011, yielding 192 months of data. All nominal returns are converted to ex-post real returns using the Consume Price Index (CPI).

I collect European style S&P 500 index options 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 are used to construct option trading 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 that is 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. Monthly holding period returns are calculated. In this process, option contracts that expire in 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 less affected by liquidity problems (Bondarenko, 2003).

The equally-spaced return series are convenient to handle empirically, especially when we use them to confront asset pricing models. 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. This sensitivity is important as risk-sensitive returns can potentially provide the most informative restrictions on the pricing kernel. Finally, from a more technical point of view, the nonparametric bounds require (gross) returns to have a positive support. However, hold-to-maturity returns are at times extremely high, creating difficulty in constructing a short strategy that always generates positive (gross) returns.

Instead of using raw option returns data, I focus on returns of a few well-known derivative strategies. There are two main reasons for this: 1. Economically meaningful strategies offer clear interpretations of the sources of risks (jump risks and/or volatility risks) that are being traded; 2. The recent literature on option pricing anomalies mainly focus on these trading strategies (Coval and Shuway, 2001, Bondarenko, 2003, Driessen and Maenhout, 2005). To be consistent and comparable with existing studies, I focus on derivative strategies.

In particular, I use the following two strategies as the 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 more sensitive to market jump risks than to volatility risks. On the other hand, an ATM market-neutral straddle generates 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 participants and have been extensively studied by the recent option pricing literature (Coval and Shumway, 2001, Jackwerth, 2000, Bondarenko, 2003). I choose the 96% OTM put as the benchmark since it is less subject to liquidity concerns than deeper OTM puts. The results hold with deeper OTM puts.

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 offsetting short positions in the 92%-OTM put option.[8] For these crash-neutral strategies, 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 of the market returns.[9] It would be interesting to see whether these alternative strategies can provide any additional information beyond what are provided by the benchmark strategies.

Table 1 presents the summary statistics for the returns of various derivative strategies as well as the market index. Consistent with the literature, long positions in these option strategies generate large negative average returns and Sharpe ratios. The flip side would be the potential gains generated by shorting these strategies. Moreover, the returns are highly non-normally distributed, as reflected by the large magnitude of skewness and kurtosis. These moment characteristics of option strategy returns will provide useful information on the pricing kernel.

In spite of the magnitude, the mean returns for my sample are notably smaller and about half the size of those reported by Coval and Shumway (2001), Bondarenko (2003), and Broadie, Chernov and Johannes (2007). This is mainly driven by the instability in estimating the mean returns for these derivative strategies. The aforementioned papers mainly focus on the period before 2005 and many include the 1987 crash episode. In contrast, mine starts in 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, respectively, much larger than their sample averages in early years. A full investigation into the changes in returns is beyond the scope of this paper. To the extent that my sample under-represents the option return population and overestimates mean returns,

[8]For details on the construction of the crash-neutral strategies, see Jackwerth (2000) and Coval and Shumway (2001).

[9]This is only approximately true 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 difficulty in interpreting the crash-neutral strategies. For instance, a short leg in the 92%-OTM put becomes a long leg. To avoid this issue, I also create robust crash-neutral puts and straddles which essentially delete the observations for which the deeper 92%-OTM put is more expensive than the 96%-OTM put. See the descriptions of Table 1 for details.

the bounds constructed below can be regarded as conservative lower bounds on the generalized entropy functions of the pricing kernel.

4.2 Bounds Implied by Option Strategies

With asset market returns and relying on the analytical tools developed in the previous sections, 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 strategy that maximizes 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 (Brandt, 1999, Sahalia and Brandt, 2001). Moreover, since we are considering portfolios that involve highly non-normally distributed option returns, the estimation issue can only get worse. Eventually, estimation uncertainty translates into bound uncertainty and this may significantly affect our inference. To reduce estimation uncertainty, I choose to consider simple static option strategy that has the following form

$$R_P = R_f + \alpha_S(R_S - R_f), \tag{14}$$

where α_S denotes the fraction of wealth allocated to a generic return R_S. Hence, only the tradeoff between a safe asset (R_f) and a return is considered.[10]

Figure 3 plots the bound frontiers (i.e., the right-hand sides) given in inequalities (10) and (11) when the power s equals 2, 0.5, 0, -1, -3 and -8, and Table 2 reports the optimal portfolio weights at a risk-free rate of zero per month.[11] Notice that when $s = 2$, 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 graph to emphasize the flip in the direction of bounds at $s = 1$. Four types of portfolios are shown on this graph: two involve the two

[10] In doing this, I ignore the possible utility gains from combining the market index with derivative strategies. For a CRRA investor with a risk aversion coefficient of more than one, Driessen and Maenhout (2005) show that the allocation to the market index is rarely significant. This indirectly shows the limited utility gains by combining the market index with derivative strategies.

[11] The mean and standard deviation of the risk-free rate for the short sample (1996-2011) is 4bp and 40bp per month, respectively. I therefore center it at zero and extend to ± 3 standard deviations away from the center in Figure 3.

benchmark derivative strategies and, for comparison purposes, the other two that involve 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 the risk-free rate. At $s = 2$, which is the HJ bound, the Sharpe ratio essentially determines the strength of bounds that a given security imposes on the pricing kernel. As a result, in accordance with the rankings of the absolute Sharpe ratios in Table 1, the 96%-OTM put implies the sharpest constraint, and is followed by the market index and lastly the ATM straddle.

At other powers, the bounds belong in the realm of the generalized entropy bounds and hence have utility-based interpretations as discussed previously. In particular, a bound at 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 ($= \frac{1}{1-0.5}$) has a significant short position in the 96%-OTM put: their paper, allowing the market index to be in an investor's choice set, has an estimate of about -10% for α_S while I have an estimate of around -20% by excluding the market index. At $s = 0$, which corresponds to the logrithmic utility case as in the original entropy bound, their estimate is around -15% and mine again roughly doubles their estimate. Putting aside the difference in asset menus and sample periods, both studies show the economic benefits by allowing investors to trade deep OTM put options. The two panels in Figure 3 (i.e., $s = 0.5$ and $s = 0$) illustrate these benefits by highlighting the differences in utility gains among the four candidate strategies.

As s becomes negative, the corresponding risk-aversion coefficient becomes even smaller so the relative gains in expected returns by shorting put options further outweigh the losses in variance and other high-order moments. Consequently, the optimal bounds require even larger positions in the OTM put. Notably, the ATM straddle yields an unimpressive mean relative to the 96%-OTM put and a much higher variance compared to the market index. Accordingly, strategies involving the straddle imply inferior bounds compared to those involving either the put option or the market index.

To gain a deeper insight into the above empirical findings, it is worthwhile to repeat the discussion in Section 2.2. Although the marginal or representative investor determines the market prices of jump and volatility risks,[12], investors with different risk attitudes all reveal valuable information about these prices through their asset allocations. For example, all else being equal, higher prices of jump risks imply more expensive deep OTM puts. With a fixed physical jump distribution, this implies a lower ex-ante and ex-post average return for buying puts for the representative agent. However, for an agent who is less risk averse and hence does not value the put option's hedging ability as much as the representative agent, she treats the increase in put option prices as a lucrative trading opportunity. By shorting more, she increases her expected utility. Following this logic, the above empirical findings (i.e., strategies involving put options imply sharper bounds than alternative strategies) highlight the more important role of priced jump risks than volatility risks in option prices. Notice that this is not saying that volatility risks matter little to investors in all circumstances, as we are only looking at the static asset allocation problems of a power utility agent. To have priced volatility risks appear more important to an investor, we have to study dynamic strategies given the strong predictability in volatilities.[13] This is left to future research.

Besides the two benchmark option strategies, I also consider alternative strategies. They may appear attractive for certain CRRA investors and thereby providing tighter bounds on the pricing kernel. Figure 4 and 5 show the 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 the strategy that shorts the 96%-OTM put option turns out to be the dominating one across all powers that belong to the realm of the generalized entropy bounds. This is somewhat surprising since we would expect its performance to be in between the strategy involving the 92%-OTM put and the one involving the ATM put. A closer look at the admissible portfolio weights reveal that strategies that short the 92%-OTM put have stronger weight restrictions: given that the maximum net return is close to 4 (see Table 1), the maximum proportion one can short in the 92%-OTM put is $1/(1-(4+1)) = -0.25$ at a risk-free rate of zero. This is significantly smaller in magnitude than the allow-

[12]The compensation for jump and volatility risks in the market index is well documented by the option pricing literature. Buraschi and Jackwerth (2001), Coval and Shumway (2001) and Bakshi and Kapadia (2003) show the presence of volatility risk premiums. Bates (2002), Pan (2002) and Ait-Sahalia, Wang and Yared (2001) show the presence of jump risk premiums.

[13]Liu (2013) uses the generalized entropy bounds to study the implications of dynamic strategies on representative agent models.

able 0.45 short position in 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, although irrelevant for the strategy dominance results that our analysis relies on (i.e., strategies involving the 96%-OTM put are more attractive and thus dominating those involving the 92%-OTM put partly because they allow more aggressive short positions), points to the instability of the in-sample asset allocation problem.[14] I will come back to this issue shortly.

Figure 5 shows that bounds based on the benchmark put option strategy dominate those based on the crash-neutral strategies at $s = 2$ and perform well for the other three powers. In particular, the OTM put strategy weakly dominates the crash-neutral put strategy at $s = 0.5$, is on par with the latter at $s = 0$, and is a little less informative at $s = -4$. The key observation is that the differences between bounds based on these two strategies are significantly smaller than the differences between these two strategies and the other two strategies (i.e., ATM straddle and its crash-neural counterpart). Taken as a whole, Figure 4 and 5 suggest the informativeness of deep OTM put strategies (especially strategies with 96%-OTM put) in effectively shaping the admissible region for moments of a candidate pricing kernel. They reveal information about the pricing of jump risks in the economy and empirically provide the sharpest bounds that any valid SDF ought to satisfy.

The optimized bounds found above can be directly used to confront candidate pricing kernels. However, they may appear too stringent for several reasons. First, in-sample asset allocation generates portfolio weight estimates that are too noisy (see Brandt, 2000, Driessen and Maenhout, 2005). What is perhaps more troublesome in our setup is the boundary-dependence of the optimal weights. As seen in 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 traction costs are small for the index option market

[14] Notice that the in-sample asset problems are well-defined both theoretically and numerically in 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 weight to be well within the boundary. However, the boundary-dependence for the optimal weight is exacerbated in our context because the CRRA investor's risk aversion coefficient is sometimes close to zero (e.g., $s = -8$ implies a risk aversion of $1/9 = 1/[1-(-8)]$). As a remedy, I consider more robust weights later on to alleviate this boundary-dependence issue.

(Bakshi, Cao and Chen, 1997) and our long positions in the risk-free asset can serve as margins, microstructure and liquidity issues may become non-negligible if we have excessive short positions in index options. Third, weights for these optimal bounds depend on the prevailing risk-free rate and are thus time-varying. This is cumbersome for our application since a different bound has to calculated for a different risk-free rate.

On further inspection of Figure 4 and 5 we notice that option implied bounds are almost constant across a wide range of bond rates. This motivates us to consider option strategies that are independent of the bond rates. I propose a simple way to create conservative and interest rate independent investment strategies. For the most informative 96%-OTM put, I set a lower threshold of $\alpha_L = -35\%$ for the short position to avoid excessive shorting,[15] and fix the weight on the put option at the optimal zero-interest weight given in Table 2 if the weight does not exceed the threshold, or at α_L otherwise.

In doing this, we end up having the following three OTM put strategies: a long 50% strategy at $s = 2$, a short 20% strategy at $s = 0.5$, and a short 35% strategy at other powers. Figure 6 shows the bounds for these conservative strategies together with those 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% across different levels of the bond rate. This is much smaller than the difference (around 8%) between the bounds implied by the ATM straddle and the OTM put strategy. For the rest of the paper, I use these simple yet efficient strategies to bear on candidate pricing kernels. Had we missed important information by employing these sub-optimal strategies, 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. More specifically, 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 option

[15]This number is chosen to approximately equal the optimal weight at $s = 0$ and a risk-free rate of zero. It is also about 10% away from the boundary value, which helps alleviate the boundary-dependence of the optimal weights.

strategies in the construction of bounds, it is only fair to consider a pricing kernel that features iid shocks. I also adopt a Possion-normal specification for the jump component in consumption growth.[16] This parametric setup has two appealing features: 1. It is flexible enough to match many of the empirical regularities on rare event distributions (Barro, 2006, Barro et al., 2009); 2. It is infinitely divisible, allowing us to "zoom" in on an arbitrarily small frequency. I state the model at the annual frequency but will use its monthly counterpart to match bounds constructed from monthly returns.

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

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

where γ governs investor risk-aversion. The pricing kernel is known to be

$$\log M_{t+1} = \log \beta - \gamma \log G_{t+1}, \tag{15}$$

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

$$\log G_{t+1} = \epsilon_{t+1} + \eta_{t+1}, \tag{16}$$

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

$$\eta_{t+1}|(J = j) \sim \mathcal{N}(j\theta, j\nu^2), J \sim \text{Poisson}(\omega). \tag{17}$$

To derive 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 $E(e^{s\epsilon_{t+1}}) = \exp(\mu s + \sigma^2 s^2/2)$. The MGF for the Poisson-normal part, as shown in Backus, Chernov and Martin (2011), is

$$E(e^{s\eta_{t+1}}) = \exp(\omega[e^{s\theta+(s\nu)^2/2} - 1]). \tag{18}$$

[16] For its applications in the macro-finance literature, see Naik and Lee (1990), Martin (2007) and Backus, Chernov and Zin (2011).

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

$$CGF(s) \equiv \log E(e^{s \log M_{t+1}}) = s \log \beta - \gamma \mu s + \frac{1}{2}\gamma^2 \sigma^2 s^2 + \omega[e^{-\gamma s \theta + (\gamma s \nu)^2/2} - 1]. \quad (19)$$

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

$$r_f = -(\log \beta - \gamma \mu + \frac{1}{2}\gamma^2 \sigma^2 + \omega[e^{-\gamma \theta + (\gamma \nu)^2/2} - 1]). \quad (20)$$

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

$$GEF(s) = \frac{1}{2}\gamma^2 \sigma^2 (1 - s) + \omega[e^{-\gamma \theta + (\gamma \nu)^2/2} - \frac{1}{s} e^{-\gamma s \theta + (\gamma s \nu)^2/2} - \frac{s-1}{s}]. \quad (21)$$

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

$$L(M) = \frac{1}{2}\gamma^2 \sigma^2 + \omega[e^{-\gamma \theta + (\gamma \nu)^2/2} + \gamma \theta - 1]. \quad (22)$$

The 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 economically interesting parameters such as the disaster intensity ω and size θ, I perform a "partial derivative" exercise.

First, as shown in Panel A of Table 4, I fix the two preference parameters (i.e., β and γ), the risk-free rate, and two variance statistics related to the consumption growth. 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). The variance for the normal component of the Poisson-normal shock is fixed at its estimate based on the realized disasters from international macroeconomic data (Barro, 2006). These second moments are estimated with much more precision than first moments (i.e., μ and θ) and my choices agree with the disaster literature.

Next, to illustrate the usefulness of the GEF, I fix the expected loss when a disaster happens at $\omega\theta = -0.006$ and consider three ω and θ combinations that represent *light*, *mild* and *severe* disaster types, respectively, as given in Panel B of Table 4.[17] These three types of disaster distributions imply an increase in disaster size (or,

[17]The mean μ of the normal shock component is used to match the interest rate.

equivalently, a decrease in disaster frequency) 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 that experienced large drops in per capita GDP during World War II, respectively. Note that these three model specifications are difficult to differentiate empirically as they imply the same total volatility in consumption growth, interest rate, and, by a first-order approximation, mean consumption growth rate.[18] It is therefore interesting to see whether GEF can better distinguish these models and, more importantly, whether asset market returns provide support for any of them.

Figure 7 shows the GEF plots for the three disaster models. I focus on the region from -3.5 to 1. As the power s goes above 1 or below -3.5, either the severe disaster case or the light disaster case will go off the charts. I will show later that bounds at these powers are less informative. Starting at $s = 1$ where all three GEF's equal zero and going left, the GEF of the severe disaster type rises more steeply than the other two GEF's and reaches its peak around $s = 0.8$. At its peak, the GEF more than triples that of either the mild or light disaster type. Going further left, it remains the dominating one until s reaches -3, at which the GEF from for the light disaster type catches up. The GEF of the mild disaster type follows a similar pattern, rising faster than the GEF of the light disaster case initially and meeting it at around $s = -2$. Eventually, all three GEF's start rising sharply for large negative powers, with the light disaster case being the dominant one.

To see how different weighting schemes on the cumulants generate the patterns in the GEF, Table 8 presents the contributions from the second to the sixth weighted cumulants (i.e., $\frac{\kappa_j(\log M)}{j!}, j = 2, 3, \ldots, 6$) to the overall entropy.[19] I choose to focus on the first two types of disaster distributions.[20]

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

[18]By equation (20), $r_f = -(\log\beta - \gamma\mu + \frac{1}{2}\gamma^2\sigma^2 + \omega(-\gamma\theta + (\gamma\nu)^2/2))$ through a first-order approximation. Our calibration in Panel A of Table 4 therefore makes sure that the mean consumption growth $\mu + \omega\theta$ is the same across the three model specifications.

[19]Strictly speaking, as in equation (12), 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 the 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) for the derivation of the individual cumulants for shocks following a Poisson-normal distribution.

[20]The severe type displays explosive behavior (compared to the other two types) for the sixth moment and beyond. To offer a better comparison, I focus on the other types.

(i.e., variance) by our calibration. However, the mild disaster case implies higher third to sixth cumulants. As a result, its overall entropy is higher.[21] At $s = 2$, which corresponds to the HJ bound, the signs for the cumulants are reversed and, more importantly, high-order cumulants are magnified compared to the case at $s = 0$. This is because the polynomial coefficients give 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 extensive literature on skewness preferences,[22] it is interesting to see whether even moments alone can rationalize asset market 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 the entropy. Finally, at $s = -3$, odd cumulants show up as negative and cancel out even cumulants. This explains why the light disaster GEF dominates the mild disaster GEF at $s = -3$: the relatively larger disaster size θ for the mild disaster model generates disproportionately high weighted odd cumulants that reduce the overall entropy. In sum, the generalized entropy seems to offer unique insights about the pricing kernel by allowing a flexible weighting of the individual moments of the pricing kernel.

I now confront the rare disaster models with option implied bounds. Analogous to the 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 based on the market Sharpe ratio and by Bansal and Lehmann (1994) to restrict the representative agent's risk aversion based on the equity premium. Note that these papers do not explicitly deal with the uncertainty in the estimation of the bounds/moments. I take this into account in the next section by rigorously testing the bound restrictions.

Again, to sharpen our focus on economically interesting quantities, I choose to consider the triple $(\omega, \theta, \gamma)'$. To avoid a negative volatility, this time I choose to fix the variance σ^2 of the normal shock component.[23] Also, risk aversion is released as a

[21] Of course, higher order cumulants matter, so we need to extrapolate from the patterns seen from the second to the sixth cumulant.

[22] See Kraus and Litzenberger (1976), Rubinstein (1973), and Harvey and Siddique (2000).

[23] 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 (ω, θ, γ) space, a fixed total variance may sometimes imply a negative σ^2. To avoid this, I set σ^2 at 0.035^2 instead. In fact, given the absence of severe disasters

free parameter to match bounds. Other than these changes, our calibrations are the same as those in Panel A of Table 4.

Absent estimation uncertainty, asset pricing bounds essentially delineate a domain in the three-dimensional (ω, θ, γ) space. To ease visualization, I plot the contours on the two-dimensional (ω, γ) plane. Figure 9 shows these contours for four values of disaster size θ. Several interesting patterns emerge from this figure.

First, as expected, a larger magnitude of θ requires a smaller risk aversion. At $\theta = -0.10$, a 4% annual equity risk premium asks for a risk aversion of around 5.5 if disasters occur every 60 years. When θ drops to -0.50, a risk aversion of 2.3 would suffice. Fixing the disaster frequency at 1/60, as θ changes from -0.10 to -0.50, the required risk aversion by the most stringent entropy bound drops from slightly above 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 high level of risk aversion may be regarded as too high to reconcile with many aspects of an individual's risk taking behavior.

Second, returns based on options require much larger risk aversion coefficients than the equity risk premium. In particular, at $1/\omega = 100$ and across various disaster size specifications, the HJ bound implied by longing 50% in the 96%-OTM put typically requires an extra 0.5 units in risk aversion compared to the HJ bound implied by the market. On top of that, the entropy bound at $s = 0$ (by shorting 40% in the 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 horizontal and the vertical axis. For instance, when $\theta = -0.50$ and fixing the risk aversion at 6, the entropy bound at $s = -1$ implies that disasters need to happen on average at least once every 50 years, whereas a duration 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 much worse than the case with only one drop every century. The differences in implications between the generalized entropy bounds and the original entropy bound are hence economically significant. In unreported results, I consider alternative interest rates and entropy bounds at even more

for the US consumption history, it makes more sense to match σ^2 rather than the total variance with the US consumption data.

negative powers. An annual risk-free rate in the range of $(1.00, 1.06)$ results in little change in the contour plots. This is due to the insensitivity of option implied bounds to interest rates, as discussed previously. Entropy bounds at more negative powers do not substantially improve on the parameter frontiers required by the entropy bound at $s = -1$.

To relate my findings to the literature, it is important to emphasize the difference in methodology. In particular, Backus, Chernov and Martin (2011) try to infer rare event information from option data. 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 of option data. Both are based more on convenience than reality. It is difficult to evaluate how small deviations from them could affect the inference on the tail distribution. My approach, on the other hand, is model-free. 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 options. Although it cannot deliver a definitive point estimate, informative bounds provide valuation information about the pricing kernel.

In terms of the empirical findings, Backus, Chernov and Martin (2011) 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 explain the option data.[24] To reconcile my findings with theirs, note that their approach has two sources of error compared to mine. First, a parametric model for the option cross-section may miss important moment characteristics in option returns. Second, they evaluate the pricing kernel against the option pricing model *only* through the lens of the basic entropy. My approach bypasses the first source of error by looking at the raw option data. At the same time, it evaluates a model through the spectrum of the generalized entropies.

4.4 Testing Rare Disaster Models with Option Return Bounds

In this section, I study the statistical significance of the violation of bounds. I start by laying down a testing framework. Suppose the set of parameters governing the model

[24] My results are consistent with the option pricing literature. In particular, Jackwerth (2000) and Bondarenko (2003) find that more frequent crashes are needed to explain put option returns.

is Π. 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

$$\Omega_M(\Pi; S) = \begin{bmatrix} [EM^{s_1}]^{\frac{1}{1-s_1}} \cdot I_{s_1 \in [0,1)} \\ [EM^{s_2}]^{\frac{1}{1-s_2}} \cdot I_{s_2 \in [0,1)} \\ \cdots \\ [EM^{s_N}]^{\frac{1}{1-s_N}} \cdot I_{s_N \in [0,1)} \end{bmatrix}, \tag{23}$$

where $S = (s_1, s_2, \ldots, s_N)'$ 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.e., $I_{s_j \in [0,1)}$'s) adjust the directions of bounds so that the left-hand side (moments of the pricing kernel) always dominate the right-hand side (moments of market returns). Similarly, the transformed return moment vector is given by

$$\Omega_R(R; S) = \begin{bmatrix} E(R_1^{\frac{s_1}{s_1-1}}) \cdot I_{s_1 \in [0,1)} \\ E(R_2^{\frac{s_1}{s_1-1}}) \cdot I_{s_2 \in [0,1)} \\ \cdots \\ E(R_N^{\frac{s_1}{s_1-1}}) \cdot I_{s_N \in [0,1)} \end{bmatrix}, \tag{24}$$

where R_j denotes the return of an asset that is chosen to restrict the s_j-th moment of the pricing kernel. At $s = 0$, the return moment $E(R^{\frac{s}{s-1}})$ should be understood as $-E(\log R)$. For a given parameterization Π of the pricing kernel, the difference between $\Omega_M(\Pi; S)$ and $\Omega_R(R; S)$ should be nonnegative, i.e.,

$$\Theta(S) \equiv \Omega_M(\Pi; S) - \Omega_R(R; S) \geq 0. \tag{25}$$

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

$$\hat{\Theta}(S) = \Omega_M(\Pi; S) - \begin{bmatrix} \frac{1}{T}\sum_{t=1}^{T} R_{t,1}^{\frac{s_1}{s_1-1}} \cdot I_{s_1 \in [0,1)} \\ \frac{1}{T}\sum_{t=}^{T} R_{t,2}^{\frac{s_2}{s_2-1}} \cdot I_{s_2 \in [0,1)} \\ \cdots \\ \frac{1}{T}\sum_{t=1}^{T} R_{t,N}^{\frac{s_N}{s_N-1}} \cdot I_{s_N \in [0,1)} \end{bmatrix}. \tag{26}$$

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. The usual Wald or Likelihood-ratio test does not apply. The difficulty lies in the specification of a null hypothesis that generates easy-to-calculate critical values. Following the multivariate testing literature (Gourieroux, Holly and Monfort 1982, Wolak, 1987, and Patton and Timmermann, 2010), I set the null at $\Theta(S) = 0$, which is least favorable to the alternative $\Theta(S) \geq 0$. To find p-values, I simulate a large number of draws from the estimated empirical limiting distribution of $\hat{\Theta}(S)$ and calculate the fraction of draws that generate an element-wise nonnegative Θ. 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 option return data that are highly skewed and fat-tailed. As an alternative, I bootstrap the historical return data to provide robust p-values. In particular, I 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 $\Omega_M(\Pi; S)$. I calculate the fraction 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 the generalized entropy bounds at different powers with multiple assets. However, a few statistical concerns restrict the way that we can form these tests. First, testing moment restrictions at different powers using the same asset is problematic. This is because the disturbance terms are perfectly correlated in a nonlinear fashion, violating the basic ergodicity assumption that is necessary for most asymptotic theories to work. Second, the estimation of the joint limiting distribution becomes increasingly unstable as we increase the number of the test statistics. Given a few hundred monthly observations, this puts a practical limit on the number of bounds we can simultaneously consider. Facing these issues, I choose to consider two types of bound tests. One is the univariate entropy bound test using the market return only. This also serves as the benchmark test since recent papers studying the equity risk premium impose such a bound (Backus, Chernov and Martin, 2011, Martin, 2008, and Alrevaz and Jermann, 2005). The other type is the joint test of the entropy bound using the market return and a generalized entropy bound using an option trading strategy. The two types of tests are acronymed MKT and MKT+OPT, respectively. As discussed in previous sections, 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 impact of the risk-free rate, I use the following procedure to generate a null hypothesis for the parameter vector Π. First, similar to before, I treat (β, σ^2, ν^2) as known parameters and set them at $(0.99, 0.035^2, 0.2^2)$. Second, I set the real annualized risk-free rate at 0.95, 1.00 and 1.05, which roughly correspond to the lower bound, mean and upper bound of a sensible estimate of the risk-free rate based on the US history. Lastly, given a point of interest for the triple $(\omega, \theta, \gamma)'$, I choose μ to match the required interest rate. In doing this, I give $(\omega, \theta, \gamma)'$ the freedom in meeting bounds from option data and the burden in meeting a target interest rate is transferred onto μ. A large deviation of μ from the mean historical consumption growth rate indicates a failure of the hypothesized parameter vector Π. I choose to report the disaster-adjusted mean consumption growth $\mu + \omega\theta$ instead.[25]

Table 5 reports the testing results for the baseline disaster model with $\omega = 0.02$ and $\theta = -0.35$. This roughly corresponds to the $\omega = 0.017, \theta = -0.38$ estimate based on the empirical distribution of international disasters by Barro (2006), and is close to the baseline parameter choice of the rare disaster literature in Martin (2009) and Backus, Chernov and Zin (2011). At $\gamma = 5$, both the entropy bound (i.e., $s = 0$) for the index and the HJ bound (i.e., $s = 2$) for the optimal option strategy are satisfied, as demonstrated by the positivity of the corresponding M_{diff} statistics. Notably, at $s = 0$ and for the univariate test with the market only (i.e., MKT test), the model implied entropy is in excess of the risk premium by at least 5% per annum. Hence, not surprisingly, the MKT test has p-values well above 0.90 across all interest rate specifications, suggesting no evidence of rejecting the entropy bound based on the market. This is also the case when the HJ bound is added (i.e., MKT + OPT test at $s = 2$), indicating no discriminatory power from the HJ bound. When we include option strategy returns at the basic entropy bound (i.e., $s = 0$), the p-values drop to 10-18%. The reduction in the p-value is impressive. Nonetheless, the model survives at the conventional significance level. When s goes to -1 and -2, the p-values drop to well below 5%, suggesting a strong rejection of the baseline disaster model across all interest rate specifications.

[25] For reasonable ω and θ pairs, the difference between $E_c = \mu + \omega\theta$ and μ is small.

Clearly, the generalized entropy bounds at $s = -1$ and $s = -2$ give us power to reject the baseline model, which cannot be rejected under the basic entropy bound (i.e., $s = 0$). Note that the underlying option strategies are the same under both types of bounds: we short 35% in the 96%-OTM put option. Therefore, compared to the basic entropy bound, the additional discriminatory power provided by the generalized entropy bounds comes from their functional forms or, more specifically, the way they weigh different moments of both the pricing kernel (i.e., left-hand side) and the asset returns (i.e., the right-hand side).

To remedy the rejection, one must increase the amount of dispersion in the generalized entropy function. As we learned from the previous section (see Figure 9), one way to achieve this is to increase the marginal investor's risk aversion. Indeed, when γ is raised to 6, the model is borderline accepted at the 1% significance level, and, at $\gamma = 7$, the p-values show no signs of rejection at all. However, as the 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 $\gamma = 7$, 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 the risk aversion and the risk-free rate.

A look at equation (20) tells us why this is the case. Compared to standard models with normal shocks, disaster models carry an extra term $\omega[e^{-\gamma\theta+(\gamma\nu)^2/2} - 1]$. Since disasters rarely happen (ω is small), this term is small for low γ values. However, as the 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 standard models. 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. Intuitively, in a disaster model framework, the agent's hedging demand for rare event risks is high for even moderately high levels of risk aversion. Given a mean consumption growth rate of 2%, this strong hedging motive requires a low risk aversion to reconcile with the relatively high interest rate. On the other hand, asset market returns (in particular option returns) ask for a high risk aversion to meet the generalized entropy bounds. Hence, there is a tradeoff in determining the representative agent's risk attitude. Taken as a whole, the baseline disaster model is unable to explain the bond market and the option market at the same time.

If the disaster distribution extrapolated from international data cannot explain the US asset market, what can? I next test disaster models with alternative distributional assumptions. Table 6 reports the results at $\omega = 0.02, \theta = -0.10$, which is arguably what the US has experienced.[26] Table 7 and 8 show the results based on further perturbations around the baseline disaster model.

In Table 6, as expected, the rejections from violating the generalized entropy bounds are stronger than those in the baseline case. Indeed, the risk aversion has to go all the way up to 9 to pass the bound test at $s = -1$ at the 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 level of risk aversion, the US type disaster specification can reconcile with the asset market. 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. Finally, 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.

One may wonder whether the demanding entropy bounds at negative powers come from the excessive short positions that we allow investors to hold. Put it differently, given the statistical uncertainties around the optimal allocation rules, 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. I address this concern in two ways.

First, I reduce α_L (i.e., the short positions in the 96%-OTM put) by half and redo all the tests at $\gamma = 5$, the typically assumed risk aversion threshold in disaster models. Table 9 shows 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 the 5% significance level for $s = -1$ and $s = -2$ across all interest rate specifications. 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

[26] 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 has actually experienced. A lower ω value makes the rejections in Table 6 even more stronger.

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.

Second, I consider other prevailing representative agent models. In particular, I test the long-run risk model as parameterized by Bansal, Kiku and Yaron (2012) and the baseline habit model in Campbell and Cochrane (1999). Table 10 shows the results ... (to be added)

To summarize 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 the generalized entropy bounds with negative powers are highlighted. However, I consider these findings as suggestive as opposed to conclusive. First, although I find evidence against static disaster models, their abilities in magnifying pricing kernel dispersion (as measured by the generalized entropy) through tail distortions are impressive. With a risk aversion of ten, even the US type specification can meet all the entropy bounds with a reasonable mean consumption growth rate. This leads one to conjecture that more sophisticated variants of disaster models may survive my tests(see Barro and Ursua, 2008 and Watcher, 2013, 2014). Second, even for the current version of the disaster model, I have not exhausted all possible parameter choices. For instance, the variance ν^2 for the normally distributed individual jump is shown to have a large impact on the risk-free rate when risk aversion is high. Yet I set it at 0.2 for simplicity. A more extensive examination of the seven dimensional parameter vector Π may yield a winner.

Despite these caveats, there are a few important takeaways from the above exercise. First and foremost, confronting a model with the equity risk premium alone is not enough, especially when tail information is the core of the model. In fact, except for $\gamma = 2$ and a risk-free rate below one, 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 alternative tail distributions for the pricing kernel. By subjecting a discount factor to a spectrum of representative option trading strategies, we gain a better sense of its all-around performance. Second, it is crucial to consider the generalized entropy bounds at negative powers, not only because of their informativeness through mean moment restrictions as demonstrated by Figure 9, but also because of the statistical powers they provide to reject candidate models. The unique moment characteristics of option returns, combined with the flexible weighting scheme offered by the generalized entropy bounds, provide a set of exacting moment constraints that help us make better inference on the tail distribution in the fundamentals.

5 Conclusion

Under the fundamental no-arbitrage condition, this paper develops a spectrum of new nonparametric bounds that significantly enrich the existing nonparametric bound universe. These bounds measure the discrepancy between what an optimizing agent could achieve if all admissible assets were tradable and what she can actually achieve in the real-world market, thereby providing an economically meaningful 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 market implied asset pricing bounds. Through moment-expansions on both sides of the new bounds, I show how they provide unique information about the pricing kernel and asset returns.

Equipped with these analytical tools, I study index option returns and their bearings on macro fundamentals through the representative-agent model framework. I find that strategies with short positions in OTM put options dominate both the market index and other representative option strategies in constraining moments of the pricing kernel. This highlights the pricing of jump risks in index options. I then postulate a pricing kernel that follows a standard static disaster model and use option return bounds to differentiate among alternative parameterizations. Both point estimates and formal hypothesis tests suggest the deficiency of standard disaster models in reconciling with option data. Both tail distortion and time-dependency might be needed to meet bounds implied by option returns.

A study of the joint behavior of the time-varying disaster distribution and option returns is an obvious avenue to pursue. This not only helps achieve the unconditional option return bounds but also generates insights into the time-series properties of option returns. The newly developed bound system, in particular an extended version that takes conditioning information into account, is expected to be instrumental.

References

Almeida, C. and R. Garcia, 2012, Assessing misspecified asset pricing models with empirical likelihood estimators, *Journal of Econometrics, 170, 519-537.*

Almeida, C. and R. Garcia, 2013, Robust economic implications of nonlinear pricing kernels, *Working Paper, Getulio Vargas Foundation.*

Alvarez, Fernando and Urban Jermann, 2005, Using asset prices to measure the persistence of the marginal utility of wealth, *Econometrica 73, 1977-2016.*

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.*

Backus, David, Mikhail Chernov and Ian Martin, 2011, Disasters implied by equity index options, *Journal of Finance 66, 1969-2012.*

Backus, David, Mikhail Chernov and Stanley Zin, 2011, Sources of entropy in representative agent models, *Working paper, New York University.*

Bakshi, G. ,C. Cao and Z. Chen, 1997, Empirical performance of alternative option pricing models, *Journal of Finance 52, 2003-2049.*

Bakshi, G. and Kapadia, N., 2003, Delta-hedged gains and the negative market volatility risk premium, *Review of Financial Studies 16, 527-566.*

Bakshi, G. and F. Chabi-Yo, 2014, New entropy restrictions and the quest for better specified asset pricing models, *Working Paper, University of Maryland.*

Bansal, Ravi and Bruce N. Lehmann, 1997, Growth-optimal portfolio restrictions on asset pricing models, *Macroeconomic Dynamics 1, 333-354.*

Barro, Robert J., 2006, Rare disasters and asset markets in the twentieth century, *Quarterly Journal of Economics 121, 823-867.*

Barro, Robert J., Emi Nakamura, Jon Steinsson and Jose F. Ursua, 2009, Crises and recoveries in an empirical model of consumption disasters, *manuscript, June.*

Barro, Robert J. and Jose F. Ursua, 2008, Macroeconomic crises since 1870, *Brookings Papers on Economic Activity, No. 1, 255-350.*

Bates, D., 2002, The market for crash risk, *Unpublished Working Paper, University of Iowa.*

Bekaert, Geert and Jun Liu, 2004, Conditioning information and variance bounds on pricing kernels, *Review of financial studies 17, 339-378.*

Brandt, M., 1999, Estimating portfolio and consumption choice: A conditional euler equations approach, *Journal of Finance 54, 1609-1646.*

Broadie, Mark, Mikhail Chernov and Michael Johannes, 2009, Understanding index option returns, *Review of Financial Studies, 22, 4493-4529.*

Bondarenko, O., 2003, Why are Puts So Expensive?, *Working Paper, University of Illinois, Chicago.*

Buraschi, A. and Jackwerth, J., 2001, The price of a smile: Hedging and spanning in option markets, *Review of Financial Studies 14, 495-527.*

Casella, George and Roger L. Berger, 2001, Statistical Inference, *Duxbury Advanced Series.*

Coval, J. and T. Shumway, 2001, Expected Option returns, *Journal of Finance 56, 983-1009.*

Davidson, James, 1994, Stochastic limit theory, *Oxford University Press.*

Driessen, Joost and Maenhout, Pascal, 2007, An empirical portfolio perspective on option pricing anomalies, *Review of Finance, 11, 561-603.*

Ferson, Wayne and Andrew Siegel, 2003, Stochastic discount factor bounds with conditioning information, *Review of Financial Studies, 16, 567-595.*

Gabaix, X., 2009, Variable raredisasters: an exactly solved framework for ten puzzles in macro-finance. *Unpublished Working Paper, New York University.*

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.*

Gourio F., 2008, Time-series predictability in the disaster model, *Finance Research Letters, 5 (2008), pp. 191203.*

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.*

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.*

Hansen, L.P. and R. Jaganathan, 1991, Implications of security market data for models of dynamic economies, *Journal of Political Economy 99, 225-262.*

Hansen, L.P. and R. Jaganathan, 1994, Assessing specification errors in stochastic discount factor models, *Staff report 167, Federal Reserve Bank of Minneapolis.*

Harvey, Campbell R. and Akhtar Siddque, 2000, Conditional skewness in asset pricing tests, *Journal of Finance 55, 1263-1295.*

Jackwerth, J., 2000, Recovering risk aversion from option prices and realized returns, *Review of Financial Studies 13, 433-451.*

Kraus, Alan and Robert H. Litzenberger, 1976, Skewness preference and the valuation of risk assets, *Journal of Finance 31, 1085-1100.*

Liu, Yan, 2013, Diagnosing Dynamic Asset Pricing Models Using Generalized Entropy Bounds, *Working Paper, Duke University.*

Longstaff, Francis A., and Monika Piazzesi, 2004, Corporate earnings and the equity premium, *Journal of Financial Economics 74, 401-421.*

Martin, Ian, 2007, Disasters and asset pricing: evidence from option markets, *Manuscript, February.*

Martin, Ian, 2008, Disasters and the welfare cost of uncertainty, 2008, *The American Economic Review 98, 74-78.*

Martin, Ian, 2009, Consumption-based asset pricing with higher cumulants, *Manuscript, January.*

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.*

Pan, J., 2002, The jump-risk premia implicit in options: evidence from an integrated time-series study, *Journal of Financial Economics 63, 3-50.*

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.*

Rietz, Thomas A., 1988, The equity premium: a solution, *Journal of Monetary Economics 22, 117-131.*

Rubinstein, Mark E., 1973, The fundamental theorem of parameter-preference security valuation, *Journal of Financial and Quantitative Analysis, Vol. 8, No. 1, 61-69.*

Snow, K.N., 1991, Diagnosing asset pricing models using the distribution of asset returns, *Journal of Finance 46, 955-983.*

Stutzer, Michael, 1995, A Bayesian approach to diagnosis of asset pricing models, *Journal of Econometrics 68, 367-397.*

Wachter, Jessica A., 2013, Can time-varying risk of rare disasters explain aggregate stock market volatility? *Journal of Finance 68, 987-1035.*

Wachter, Jessica A. and Sang Byung Seo, 2014, Option prices in a model with stochastic disaster risk, *Working Paper, University of Pennsylvania.*

Wolak, F.A., 1987, An exact test for multiple inequality and equality constraints in the linear regression model. *Journal of the American Statistical Association 82, 782-793.*

A Tables and Figures

	Mean	Std.	Sharp	Skew.	Kurt.	Max	Min	Corr.
Index (1926-2011)	0.007	0.055	0.113	0.225	10.323	0.384	-0.285	NA
Index (1996-2011)	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 for the monthly returns of the S&P 500 index and eight option strategies. The longer index series is from July 1926 to December 2011 and the shorter index series is from January 1996 to December 2011, which is also the time period for all the option strategies. 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, for which the original 96%-OTM put and ATM straddle are mixed with a short leg on the 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, respectively. They are the original crash neutral series excluding the date on which 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. Under this condition, six observations are deleted from the 192 monthly observations, including two months during which the 92%-OTM put has a higher price than the 96%-OTM put. Skewness and kurtosis are the standardized central third and fourth moments, respectively. The risk-free rate is 60bp annualized for the long sample and 54bp for the short sample.

Power (s)	Market	96%-OTM put	ATM straddle
Panel A: Optimal weights			
2	-1.954	0.489	0.149
0.5	0.941	-0.199	-0.070
0	1.839	-0.341	-0.137
-1	3.450	-0.437	-0.260
-3	5.428	-0.442	-0.461
-8	5.667	-0.442	-0.663
Panel B: Constraints on weights			
$\alpha_S(min)$	-9.589	-0.450	-0.680
$\alpha_S(max)$	5.369	1.175	1.281

Table 2: **Optimal portfolio weights for benchmark option strategies and market index.** Panel A shows the optimal portfolio weights for the optimization problems described in Figure 3 at a fixed interest rate of zero. Panel B shows the range of the portfolio weights that guarantee the positivity of the portfolio returns. For a generic return series $\{R_t\}_{t=1}^T$, the range is given by $[\alpha_S(min), \alpha_S(max)] = [1/(1 - \max_{1 \le t \le T}\{R_t\}), 1/(1 - \min_{1 \le t \le T}\{R_t\})]$.

Power(s)	92%-OTM put	ATM put	R-C-neutral put	R-C-neutral straddle
Panel A: Optimal weights				
2	0.406	0.207	0.253	0.138
0.5	-0.129	-0.088	-0.184	-0.070
0	-0.202	-0.162	-0.331	-0.138
-4	-0.253	-0.306	-0.467	-0.545
Panel B: Constraints on weights				
$\alpha_S(min)$	-0.254	-0.307	-0.468	-0.601
$\alpha_S(max)$	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 problems described in Figure 4 and 5 at a fixed interest rate of zero. Strategies involving the 92%-OTM put, ATM put, crash-neutral put and crash-neutral straddle are described in Table 1. Panel B shows the range of the portfolio weights that guarantee the positivity of the portfolio returns. For a generic return series $\{R_t\}_{t=1}^T$, the range is given by $[\alpha_S(min), \alpha_S(max)] = [1/(1 - \max_{1 \le t \le T}\{R_t\}), 1/(1 - \min_{1 \le t \le T}\{R_t\})]$.

Parameter	Value
Panel A	
β	0.99
γ	5
r_f	0.02
$\sigma^2 + \omega(\theta^2 + \nu^2)$	0.035^2
ν^2	0.2^2
Panel B	
ω_L	0.04
θ_L	-0.15
ω_M	0.02
θ_M	-0.30
ω_S	0.01
θ_S	-0.60

Table 4: **Parameter specifications for disaster models.** This table shows the parameter specifications for the disaster model introduced in Section 4.3. 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 combinations that represent three types of disaster distributions: the light disaster type (ω_L, θ_L), the mild disaster type (ω_M, θ_M), and the severe disaster type (ω_S, θ_S).

			MKT	MKT + OPT				
			$s=0$	$s=2$	$s=0.5$	$s=0$	$s=-1$	$s=-2$
$\gamma=2$	$R_f=0.95$	M_{diff}	-0.081	-1.018	-0.315	-0.500	-0.387	-0.566
	(E_c =-0.025)	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
	$R_f=1.00$	M_{diff}	-0.030	-1.120	-0.263	-0.449	-0.361	-0.532
	($E_c=0.001$)	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
	$R_f=1.05$	M_{diff}	0.019	-1.218	-0.215	-0.400	-0.337	-0.500
	($E_c=0.025$)	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
$\gamma=5$	$R_f=0.95$	M_{diff}	0.057	2.698	-0.213	-0.362	-0.303	-0.444
	($E_c=0.018$)	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
	$R_f=1.00$	M_{diff}	0.108	2.628	-0.162	-0.311	-0.278	-0.409
	($E_c=0.028$)	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
	$R_f=1.05$	M_{diff}	0.157	2.561	-0.114	-0.262	-0.253	-0.376
	($E_c=0.044$)	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
$\gamma=6$	$R_f=0.95$	M_{diff}	0.202	9.103	-0.093	-0.217	-0.223	-0.331
	($E_c=0.039$)	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
	$R_f=1.00$	M_{diff}	0.253	9.088	-0.043	-0.165	-0.197	-0.296
	($E_c=0.048$)	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
	$R_f=1.05$	M_{diff}	0.302	9.074	0.005	-0.116	-0.173	-0.263
	($E_c=0.056$)	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
$\gamma=7$	$R_f=0.95$	M_{diff}	0.485	10.831	0.149	0.067	-0.072	-0.121
	($E_c=0.074$)	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
	$R_f=1.00$	M_{diff}	0.536	10.831	0.199	0.118	-0.046	-0.086
	($E_c=0.081$)	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
	$R_f=1.05$	M_{diff}	0.585	10.831	0.246	0.167	-0.021	-0.052
	($E_c=0.088$)	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 risk-free 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 (generalized) entropy bounds with both the market and the option strategy as the test assets. M_{diff} is the difference between the model implied moment and the sample asset return moment given in equation (26). M_{diff} is calculated for the market return under MKT, and for the option strategy return under MKT + OPT. P^a-value and P^b-value are the p-values generated from the theoretical limiting distribution and a bootstrap procedure, respectively.

			MKT	MKT + OPT				
			$s = 0$	$s = 2$	$s = 0.5$	$s = 0$	$s = -1$	$s = -2$
$\gamma = 2$	$R_f = 0.95$	M_{diff}	-0.089	-1.054	-0.319	-0.507	-0.392	-0.574
	$(E_c = -0.028)$	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
	$R_f = 1.00$	M_{diff}	-0.037	-1.156	-0.268	-0.456	-0.366	-0.540
	$(E_c = -0.003)$	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
	$R_f = 1.05$	M_{diff}	0.012	-1.254	-0.219	-0.407	-0.342	-0.507
	$(E_c = 0.022)$	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
$\gamma = 5$	$R_f = 0.95$	M_{diff}	-0.054	-0.731	-0.298	-0.472	-0.364	-0.518
	$(E_c = -0.004)$	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
	$R_f = 1.00$	M_{diff}	-0.003	-0.831	-0.247	-0.421	-0.338	-0.483
	$(E_c = 0.006)$	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
	$R_f = 1.05$	M_{diff}	0.046	-0.926	-0.198	-0.372	-0.314	-0.451
	$(E_c = 0.016)$	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
$\gamma = 7$	$R_f = 0.95$	M_{diff}	0.010	2.272	-0.252	-0.409	-0.320	-0.387
	$(E_c = 0.006)$	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
	$R_f = 1.00$	M_{diff}	0.061	2.199	-0.201	-0.357	-0.294	-0.352
	$(E_c = 0.013)$	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
	$R_f = 1.05$	M_{diff}	0.110	2.128	-0.152	-0.309	-0.26	-0.319
	$(E_c = 0.020)$	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
$\gamma = 9$	$R_f = 0.95$	M_{diff}	0.167	10.814	-0.123	-0.252	-0.221	0.418
	$(E_c = 0.022)$	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
	$R_f = 1.00$	M_{diff}	0.218	10.813	-0.073	-0.201	-0.195	0.455
	$(E_c = 0.028)$	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
	$R_f = 1.05$	M_{diff}	0.267	10.813	-0.025	-0.152	-0.171	0.490
	$(E_c = 0.033)$	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 risk-free 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 (generalized) entropy bounds with both the market and the option strategy as the test assets. M_{diff} is the difference between the model implied moment and the sample asset return moment given in equation (26). M_{diff} is calculated for the market return under MKT, and for the option strategy return under MKT + OPT. P^a-value and P^b-value are the p-values generated from the theoretical limiting distribution and a bootstrap procedure, respectively.

			MKT	MKT + OPT				
			$s=0$	$s=2$	$s=0.5$	$s=0$	$s=-1$	$s=-2$
$\gamma=2$	$R_f=0.95$	M_{diff}	-0.077	-0.974	-0.312	-0.496	-0.385	-0.563
	$(E_c=-0.023)$	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
	$R_f=1.00$	M_{diff}	-0.026	-1.075	-0.260	-0.444	-0.359	-0.529
	$(E_c=0.003)$	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
	$R_f=1.05$	M_{diff}	0.023	-1.172	-0.212	-0.396	-0.334	-0.497
	$(E_c=0.028)$	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
$\gamma=5$	$R_f=0.95$	M_{diff}	0.213	9.786	-0.076	-0.206	-0.224	-0.339
	$(E_c=0.049)$	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
	$R_f=1.00$	M_{diff}	0.264	9.777	-0.026	-0.154	-0.198	-0.305
	$(E_c=0.059)$	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
	$R_f=1.05$	M_{diff}	0.313	9.768	0.022	-0.106	-0.174	-0.271
	$(E_c=0.069)$	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
$\gamma=5.5$	$R_f=0.95$	M_{diff}	0.379	10.818	0.068	-0.040	-0.137	-0.220
	$(E_c=0.075)$	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
	$R_f=1.00$	M_{diff}	0.430	10.817	0.118	0.011	-0.111	-0.185
	$(E_c=0.084)$	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
	$R_f=1.05$	M_{diff}	0.479	10.817	0.165	0.060	-0.086	-0.151
	$(E_c=0.093)$	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 risk-free 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 (generalized) entropy bounds with both the market and the option strategy as the test assets. M_{diff} is the difference between the model implied moment and the sample asset return moment given in equation (26). M_{diff} is calculated for the market return under MKT, and for the option strategy return under MKT + OPT. P^a-value and P^b-value are the p-values generated from the theoretical limiting distribution and a bootstrap procedure, respectively.

			MKT	MKT + OPT				
			$s=0$	$s=2$	$s=0.5$	$s=0$	$s=-1$	$s=-2$
$\gamma=2$	$R_f=0.95$	M_{diff}	-0.085	-1.039	-0.317	-0.503	-0.389	-0.569
	$(E_c=-0.026)$	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
	$R_f=1.00$	M_{diff}	-0.033	-1.141	-0.266	-0.452	-0.363	-0.535
	$(E_c=-0.001)$	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
	$R_f=1.05$	M_{diff}	0.016	-1.239	-0.217	-0.403	-0.339	-0.502
	$(E_c=0.024)$	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
$\gamma=5$	$R_f=0.95$	M_{diff}	-0.009	-0.011	-0.266	-0.427	-0.336	-0.477
	$(E_c=0.005)$	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
	$R_f=1.00$	M_{diff}	0.043	-0.104	-0.215	-0.376	-0.310	-0.442
	$(E_c=0.015)$	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
	$R_f=1.05$	M_{diff}	0.092	-0.193	-0.167	-0.327	-0.286	-0.409
	$(E_c=0.025)$	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
$\gamma=7$	$R_f=0.95$	M_{diff}	0.159	7.682	-0.136	-0.260	-0.235	-0.274
	$(E_c=0.027)$	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
	$R_f=1.00$	M_{diff}	0.210	7.655	-0.085	-0.208	-0.209	-0.239
	$(E_c=0.035)$	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
	$R_f=1.05$	M_{diff}	0.259	7.629	-0.037	-0.159	-0.184	-0.206
	$(E_c=0.042)$	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
$\gamma=8$	$R_f=0.95$	M_{diff}	0.335	10.803	0.012	-0.083	-0.135	-0.012
	$(E_c=0.046)$	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
	$R_f=1.00$	M_{diff}	0.387	10.803	0.062	-0.032	-0.108	0.024
	$(E_c=0.052)$	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
	$R_f=1.05$	M_{diff}	0.435	10.803	0.109	0.017	-0.084	0.058
	$(E_c=0.058)$	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 risk-free 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 (generalized) entropy bounds with both the market and the option strategy as the test assets. M_{diff} is the difference between the model implied moment and the sample asset return moment given in equation (26). M_{diff} is calculated for the market return under MKT, and for the option strategy return under MKT + OPT. P^a-value and P^b-value are the p-values generated from the theoretical limiting distribution and a bootstrap procedure, respectively.

			MKT	MKT + OPT				
			$s=0$	$s=2$	$s=0.5$	$s=0$	$s=-1$	$s=-2$
$(\omega=0.02,\theta=-0.35)$	$R_f=0.95$	M_{diff}	0.057	2.698	−0.213	−0.286	−0.151	−0.198
	$(E_c=0.018)$	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
	$R_f=1.00$	M_{diff}	0.108	2.628	−0.162	−0.234	−0.125	−0.164
	$(E_c=0.028)$	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
	$R_f=1.05$	M_{diff}	0.157	2.561	−0.114	−0.186	−0.101	−0.131
	$(E_c=0.038)$	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
$\omega=0.02,\theta=-0.10$	$R_f=0.95$	M_{diff}	−0.054	−0.731	−0.298	−0.396	−0.212	−0.272
	$(E_c=-0.004)$	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
	$R_f=1.00$	M_{diff}	−0.003	−0.831	−0.247	−0.345	−0.186	−0.238
	$(E_c=0.006)$	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
	$R_f=1.05$	M_{diff}	0.046	−0.926	−0.198	−0.296	−0.162	−0.205
	$(E_c=0.016)$	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
$\omega=0.04,\theta=-0.15$	$R_f=0.95$	M_{diff}	−0.009	−0.011	−0.266	−0.351	−0.184	−0.232
	$(E_c=0.005)$	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
	$R_f=1.00$	M_{diff}	0.043	−0.104	−0.215	−0.300	−0.158	−0.197
	$(E_c=0.015)$	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
	$R_f=1.05$	M_{diff}	0.092	−0.193	−0.167	−0.251	−0.133	−0.164
	$(E_c=0.025)$	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
$\omega=0.01,\theta=-0.60$	$R_f=0.95$	M_{diff}	0.213	9.786	−0.076	−0.129	−0.072	−0.094
	$(E_c=0.049)$	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
	$R_f=1.00$	M_{diff}	0.264	9.777	−0.026	−0.078	−0.046	−0.059
	$(E_c=0.059)$	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
	$R_f=1.05$	M_{diff}	0.313	9.768	0.022	−0.029	−0.022	−0.026
	$(E_c=0.069)$	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% at $s = 0, -1$ and -2. R_f is the annual risk-free 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 (generalized) entropy bounds with both the market and the option strategy as the test assets. M_{diff} is the difference between the model implied moment and the sample asset return moment given in equation (26). M_{diff} is calculated for the market return under MKT, and for the option strategy return under MKT + OPT. P^a-value and P^b-value are the p-values generated from the theoretical limiting distribution and a bootstrap procedure, respectively.

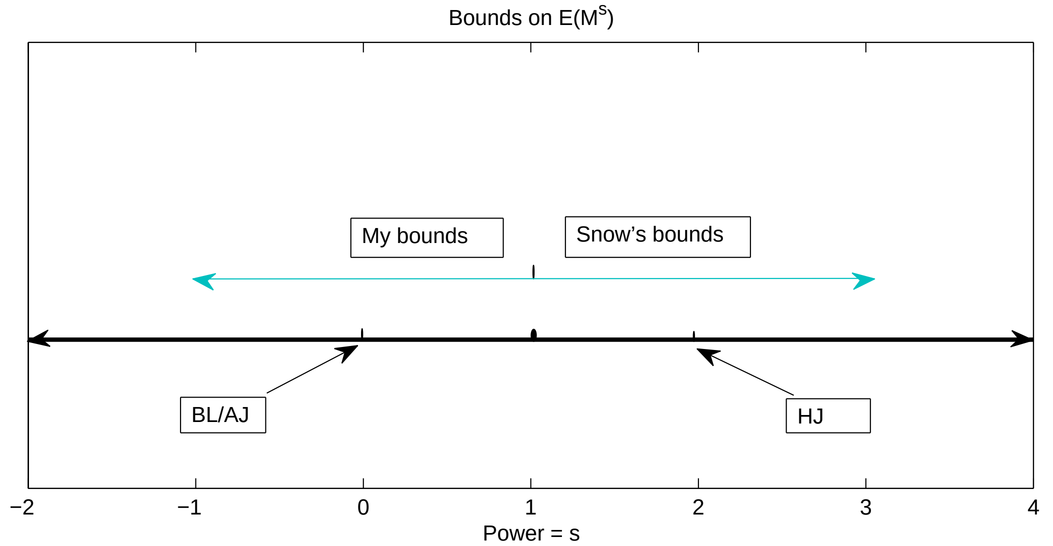

Figure 1: **The non-parametric bound universe**

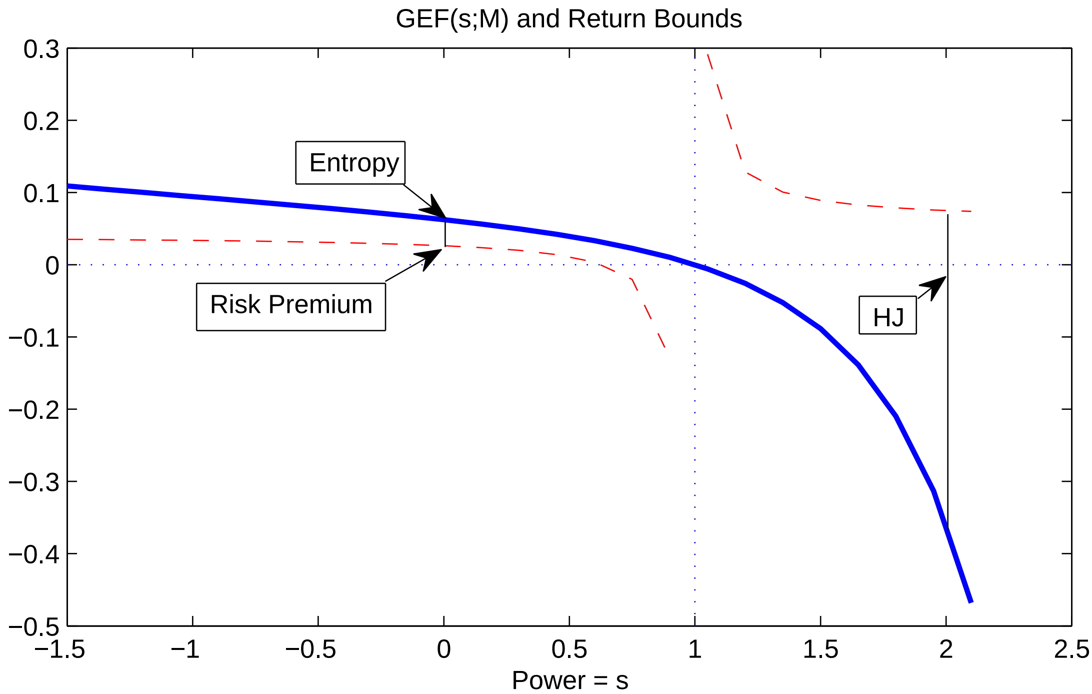

Figure 2: **A typical plot of $GEF(s;M)$ and asset market bounds.** This figure plots a generic GEF and asset market bounds. The thick solid line depicts the GEF and the thin dashed line depicts asset market bounds.

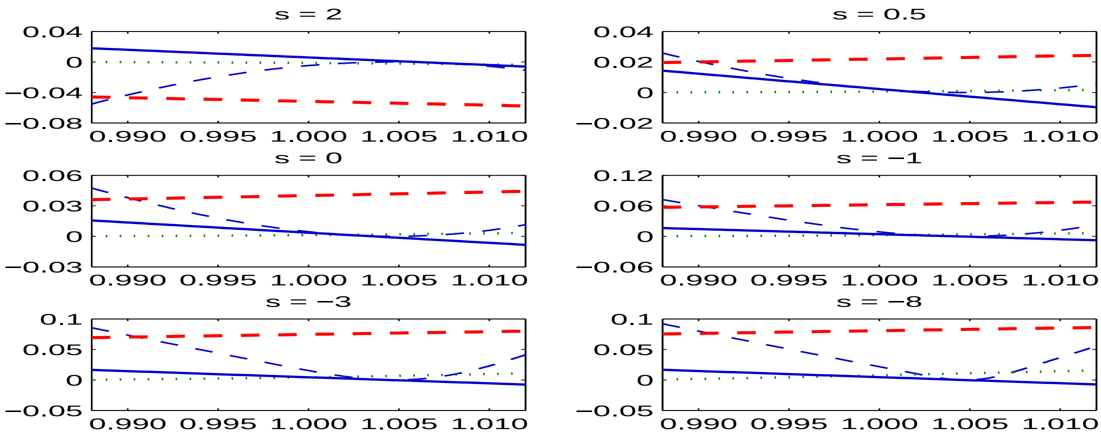

Figure 3: **Bounds implied by benchmark option strategies and market index.** This figure plots the non-parametric bound frontiers (right-hand sides in inequalities (10) and (11)) for the benchmark option trading strategies and the market index across different hypothetical risk-free rates. The estimation is done, at each hypothetical risk-free rate, by conducting a nonlinear search on the optimal portfolio weight α_S to either maximize or minimize the right-hand sides of inequalities (10) and (11). 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 α_S at zero at each level of the hypothetical risk-free rate and the active market strategy involves a search as described above.

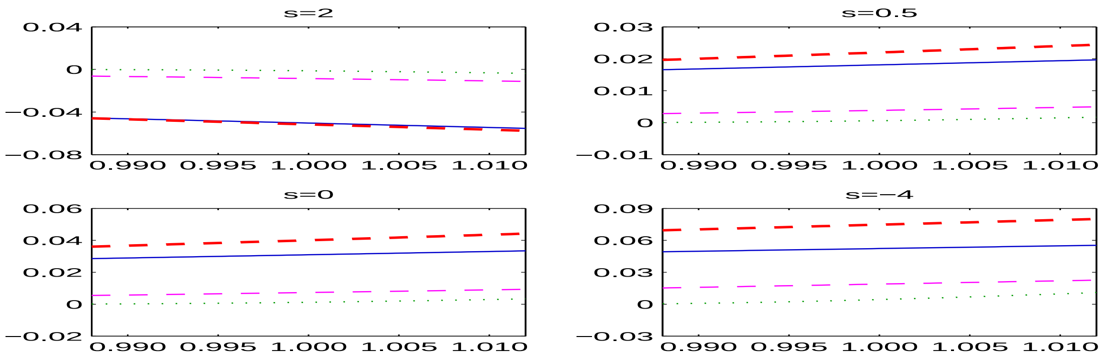

Figure 4: **Bounds implied by benchmark option strategies and two alternative OTM put strategies.** This figure plots the non-parametric bound frontiers (right-hand sides in inequalities (10) and (11)) for the benchmark option trading strategies and the market index across different hypothetical risk-free rates. The estimation is done, at each hypothetical risk-free rate, by conducting a nonlinear search on the optimal portfolio weight α_S to either maximize or minimize the right-hand sides of inequalities (10) and (11). The solid line, thin dashed line, dotted line and thick dashed line depict the frontiers for the 92%-OTM put, ATM put, ATM straddle and 96%-OTM put strategy, respectively.

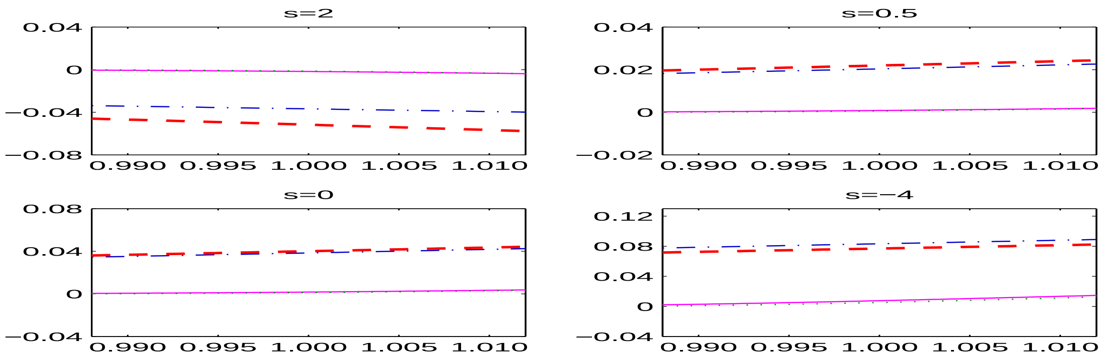

Figure 5: **Bounds implied by benchmark option strategies and two robust crash-neutral strategies.** This figure plots the non-parametric bound frontiers (right-hand sides in inequalities (10) and (11)) for the benchmark option trading strategies and the market index across different hypothetical risk-free rates. The estimation is done, at each hypothetical risk-free rate, by conducting a nonlinear search on the optimal portfolio weight α_S to either maximize or minimize the right-hand sides of inequalities (10) and (11). The two robust crash-neutral strategies are described in Table 1. The solid line, dot-dashed line, dotted line and thick dashed line depict the frontiers for the crash-neutral straddle, crash-neutral put, ATM straddle and 96%-OTM put strategy, respectively.

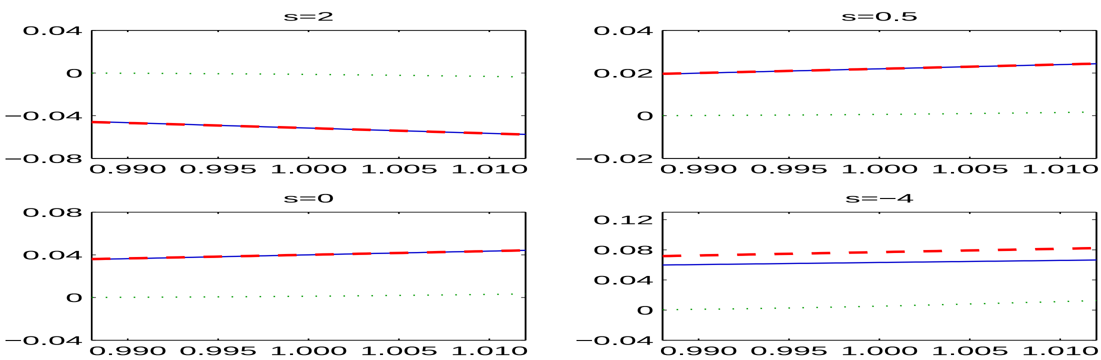

Figure 6: **Bounds implied by optimal and conservative benchmark strategies.** This figure plots the non-parametric bound frontiers (right-hand sides in inequalities (10) and (11)) for the benchmark option trading strategies and the market index across different hypothetical risk-free rates. The estimation is done, at each hypothetical risk-free rate, by conducting a nonlinear search on the optimal portfolio weight α_S to either maximize or minimize the right-hand sides of inequalities (10) and (11). The dotted line and the thick dashed line depict the frontier for the ATM straddle and 96%-OTM put strategy, respectively. The solid lines for $s = 2, 0.5, 0, -4$ depict the frontiers for the risk-free rate independent strategies that long 50%, short 20%, short 35% and short 35% in the 96%-OTM put option strategy, respectively.

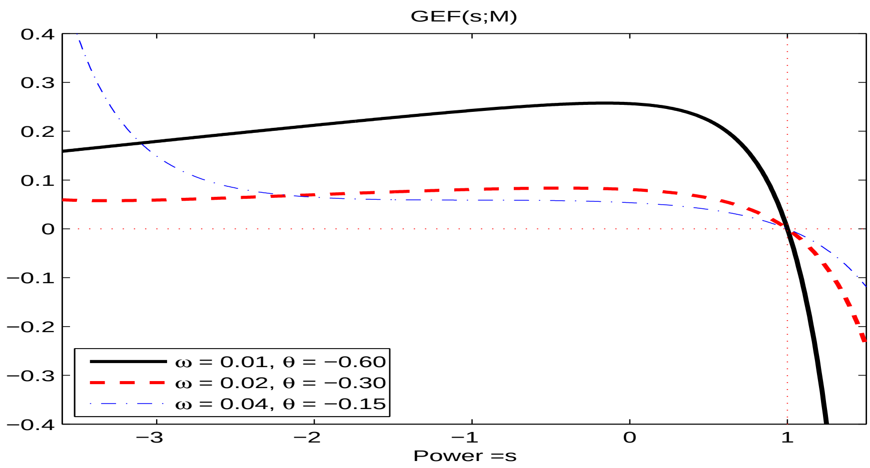

Figure 7: **Generalized entropy function plots for three disaster models.**

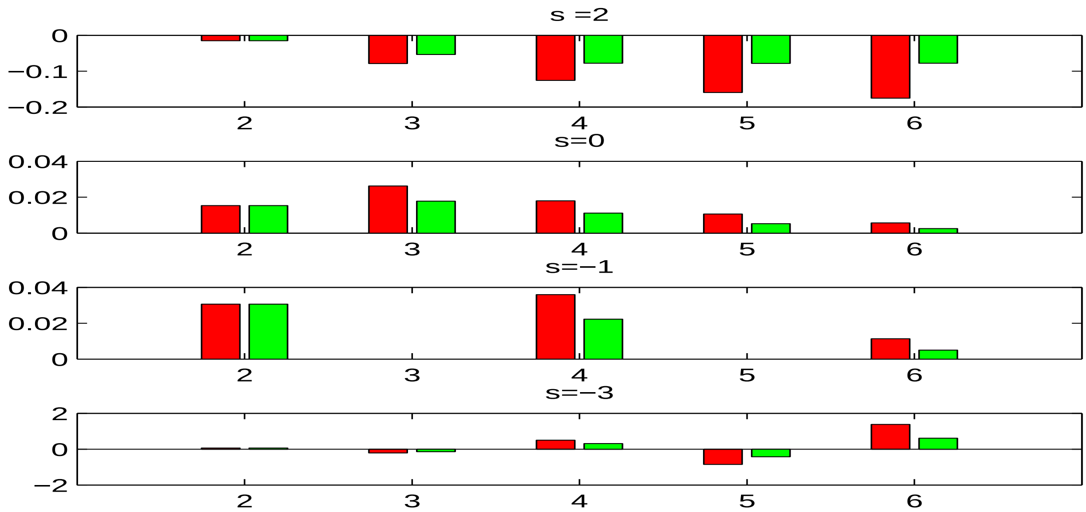

Figure 8: **Weighted cumulants for two disaster models.** This figure displays the second to sixth weighted cumulants for the mild and light disaster model at $s = 2, 0, -1$ and -3. The j-th weighted cumulant is defined as $\frac{\kappa_j(\log M_{t+1})}{j!}(1 - s^{j-1})$ in equation (12). The left (dark) bar and the right (light) bar measure the weighted cumulant for the mild and light disaster model, respectively.

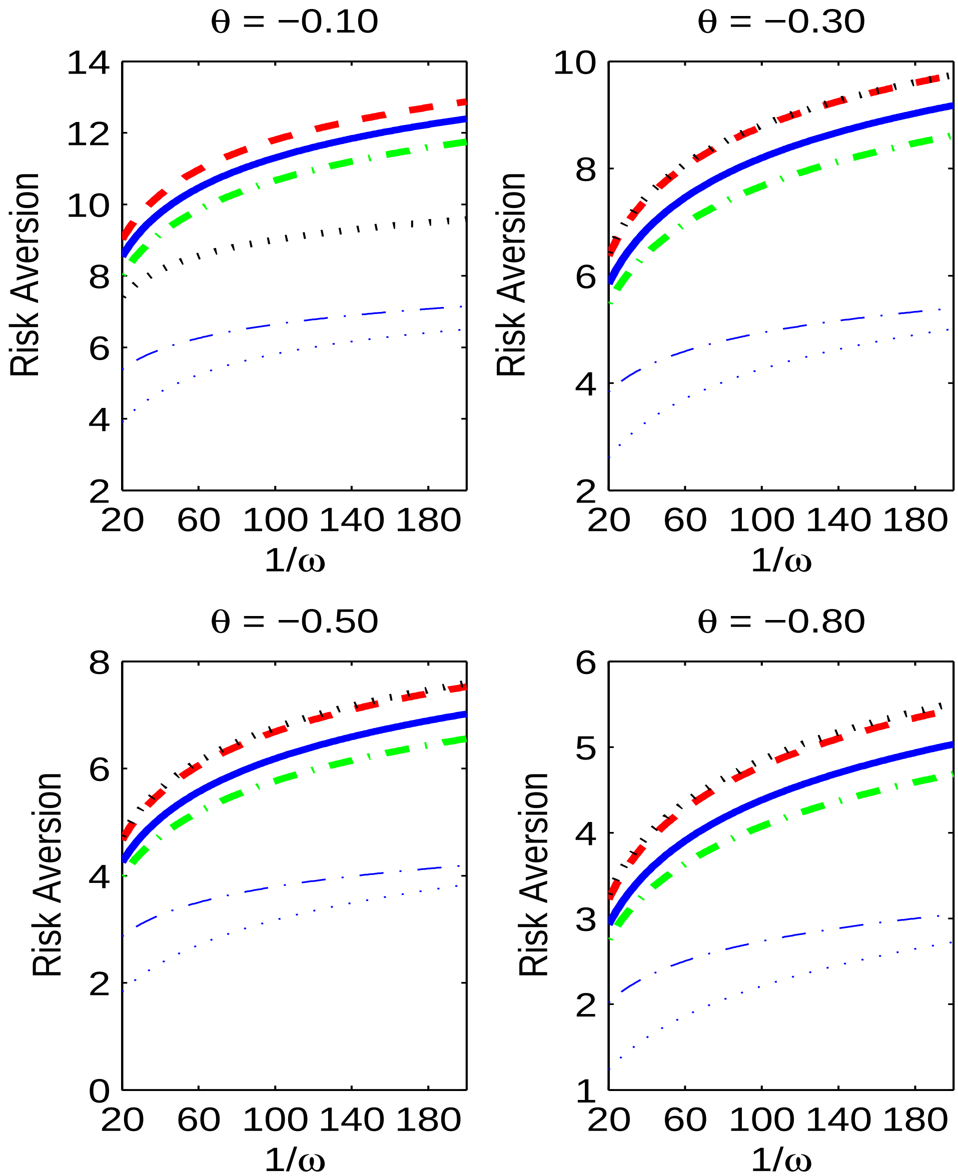

Figure 9: **Risk aversion bounds implied by index option returns.** This figure shows the required risk aversion coefficients corresponding to different disaster frequency ω, disaster size θ and entropy bounds based on option returns. The thin dotted line depicts the required risk aversion in generating 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

$$\begin{array}{ll} \sup\limits_{R} & E[\dfrac{R^{1-\gamma}}{1-\gamma}] \\ s.t. & (1)\ E(MR) = 1, \\ & (2)\ R > 0. \end{array}$$

For simplicity, I succinctly denote $\gamma(\delta)$ by γ. 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 λ and $\mu(s)$, respectively, then the Lagrange function is

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

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 in every state. Assuming an everywhere positive $f(s)$, we arrive at the following solution for $R(s)$

$$R(s) = [\lambda(1-\gamma)]^{-\frac{1}{\gamma}} M(s)^{-\frac{1}{\gamma}}. \tag{27}$$

To express λ as a moment of the pricing kernel, we can multiply both sides of equation (27) by $M(s)f(s)$ and sum across states. This leave us with the following equation for λ

$$[\lambda(1-\gamma)]^{-\frac{1}{\gamma}} E(M^{1-\frac{1}{\gamma}}) = 1. \tag{28}$$

Combining equation (27) and (28), we get the optimal portfolio choice as a function of M only

$$\tilde{R} = M^{-\frac{1}{\gamma}} / E(M^{\frac{\gamma-1}{\gamma}}). \tag{29}$$

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

$$U(M) = \frac{E(\tilde{R}^{1-\gamma})}{1-\gamma} = \frac{[E(M^{\frac{\gamma-1}{\gamma}})]^{\gamma}}{1-\gamma}. \tag{30}$$

Equation (29) and (30) give the optimal solution to this problem.

B.2 Duality Definition and Proof

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

$$\begin{array}{ll} \sup\limits_{R} & E[\frac{R^{1-\gamma}}{1-\gamma}] \\ s.t. & (1)\ E(MR) = 1, \\ & (2)\ R > 0. \end{array}$$

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 δ-th moment of the pricing kernel can be presented by

$$\begin{array}{ll} \inf\limits_{M} & \frac{[E(M^{\delta})]^{\frac{1}{1-\delta}}}{1-\gamma(\delta)} \\ s.t. & (1)\ E(MR) = 1, \\ & (2)\ M > 0. \end{array}$$

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 the following: 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 pricking 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

$$R_{YL}^{-1}(R) \equiv M_{HJ}(R)$$
$$M_{HJ}^{-1}(M) \equiv R_{YL}(M).$$

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

$$M_{HJ}(R) = C(R) \cdot R^{\frac{1}{\delta-1}}, \tag{31}$$

where the normalizing constant $C(R)$ is equal to $1/E(R^{\frac{\delta}{\delta-1}})$. By plugging the formulae in equation (29) and (31) into the duality conditions, it is readily seen that these conditions are satisfied.

B.3 Deriving the Information Bound in Stutzer (1995)

To be added.

B.4 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 δ-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

$$M_n = \begin{cases} n - \alpha_n & \text{if } X \in (0, \dfrac{1}{n}), \\ \\ \dfrac{\alpha_n}{n-1} & \text{if } X \in [\dfrac{1}{n}, 1) \end{cases} \tag{32}$$

where $\{\alpha_n\}_{n=1}^{\infty}$ is a sequence that satisfies $\alpha_n < n$ and $\frac{\alpha_n}{n} \to 0$ (For simplicity, α_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 ∞ since

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

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

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

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.