

# Empirical Likelihood Estimators for Stochastic Discount Factors

March 17, 2008

## Abstract

Hansen and Jagannathan (HJ, 1991) provided bounds on the volatility of Stochastic Discount Factors (SDF) that proved extremely useful to diagnose and test asset pricing models. This nonparametric bound reflects a duality between the mean-standard deviation frontier for SDFs and the mean-variance frontier for portfolios of asset returns. We extend this fundamental contribution by proposing information bounds that minimize general convex functions of SDFs directly taking into account higher moments of returns. These Minimum Discrepancy bounds reflect a duality with finding the optimal portfolio of asset returns with a general HARA utility function. The maximum utility portfolio implies SDF estimators that are based on implied probabilities associated with the class of Generalized Empirical Likelihood estimators. We analyze the implications of these information bounds for the pricing of size portfolios and the performance evaluation of hedge funds.

**Keywords:** Stochastic Discount Factor, Information-Theoretic Bounds, Generalized Minimum Contrast Estimators, Implicit Utility Maximizing Weights.

**JEL Classification:** C1,C5,G1

## 1 Introduction

Hansen and Jagannathan (HJ, 1991) provided bounds on the volatility of Intertemporal Marginal Rates of Substitution (IMRS) or Stochastic Discount Factors (SDF) that proved extremely useful as diagnostics of asset pricing models, as well as for other applications such as predictability issues, variance spanning tests, and performance measurement [see Bekaert and Liu (2004) and references therein]. Their methodology consisted in unconditionally and linearly projecting the SDF into the payoff space of a number of securities to characterize a region for its volatility as a function of its mean. This nonparametric bound reflects a duality between the mean-standard deviation frontier for SDFs and the mean-variance frontier for portfolios of asset returns.

A more general objective is to build frontiers that do not rely on linear projections and go beyond the variance. Snow (1991) extends the HJ analysis by obtaining bounds on higher moments of the SDFs. Stutzer (1995) proposes an information bound based on the minimization of the Kullback-Leibler information criterion and shows that it is equivalent to finding the optimal portfolio of asset returns with a negative exponential CARA utility function.<sup>1</sup> Bansal and Lehmann (1997) show a similar duality with logarithmic utility.

The main contribution of this paper is to derive information bounds that result from the portfolio optimization of a HARA utility investor and show that they correspond to minimum contrast measures of the Cressie Read (1984) family that have been recently used in the econometric literature as one step alternatives to Generalized Method of Moments (GMM) estimators [see Smith (1997), Newey and Smith (2004), Kitamura (2006), and Schennach (2007)].<sup>2</sup> Therefore, exploiting this duality, we obtain more general information bounds and corresponding theoretically-sound estimators of *nonparametric* stochastic discount factors implicit in the portfolio problems.<sup>3</sup>

Some particular cases of portfolio implied SDF bounds have been previously studied in the finance literature. However, they do not in general stress the link with corresponding econometric estimators (an exception being Stutzer (1995)), nor do they use empirically

---

<sup>1</sup>Several studies used volatility-minimizing metrics. Jensen (1968) derived a risk-adjusted measure of portfolio performance based on the CAPM (see also Roll (1978), Mayers and Rice (1979), and Dybvig and Ross (1985)). Admati and Ross (1985), and Admati et al. (1986) assume maximization of a CARA utility function to measure investment performance. More recently, McCulloch and Rossi (1990) test the APT theory with a negative exponential CARA utility by *explicitly solving* an optimal portfolio problem and comparing certainty equivalents.

<sup>2</sup>Recently, Newey and Smith (2004) showed that any member in the Cressie Read discrepancy family, a subset of the Minimum Discrepancy (MD) estimators [Corcoran (1998)], has a dual member in the class of Generalized Empirical Likelihood (GEL) estimators [Smith (1997)]. Our HARA portfolio problems are formulated via the GEL estimation procedure, dual to the Cressie Read discrepancy family.

<sup>3</sup>In fact, Newey and Smith (2004), and Kitamura (2006) provide evidence on the superior higher-order properties of GEL/MD estimators when compared to GMM estimators.

the implied SDFs themselves. The log utility maximizing portfolio, called growth portfolio by Bansal and Lehmann (1997), corresponds to the empirical likelihood (EL) criterion.<sup>4</sup> Another prominent particular case is the exponential tilting (ET) criterion of Stutzer (1995) and its corresponding optimum portfolio of a CARA investor.<sup>5</sup> Log utility and CARA correspond to specific values of the parameters of a HARA utility function, especially the risk aversion parameters. We explore new utility functions, in particular with higher risk aversion, and derive the corresponding information bounds called Cressie-Read (CR) bounds through a unified methodology.

The general goal of our approach is to obtain nonparametric<sup>6</sup> estimates of bounds on the stochastic discount factors, as in Hansen and Jagannathan (1991), but that account for higher moments in the distribution of returns. This is important for pricing assets with non-normal returns or evaluating strategies that create non-normal returns. All ET, EL and CR measures put implicit weights on a potentially infinite number of higher order moments. This is in contrast with Snow (1991) who uses Holder's inequality to obtain mean-  $n^{th}$ -moment SDF frontiers, for arbitrary values of  $n$ . These frontiers taken together are more informative than standard HJ bounds, but fail to provide criteria that would help in the construction of SDF bounds that integrate more than a pair of moments.

Another way of introducing higher moments in the SDF bound is to find the SDFs that are linear functions of the payoffs and say the payoffs squared and solve the asset pricing equation. This is the approach followed by Chabi-Yo (2008). He proposes both unconditional and conditional bounds. The projection on squared payoffs links the bounds to skewness and kurtosis of returns, therefore allowing for a characterization of the bound in the joint space of mean, variance, skewness and kurtosis.<sup>7</sup> He extends studies of Gallant, Hansen, and Tauchen (1990, hereafter GHT), Ferson and Siegel (2001, 2003), and Bekaert and Liu (2004) who incorporate conditioning information in deriving the mean-variance frontier.<sup>8</sup>

---

<sup>4</sup>The term Empirical Likelihood has been coined by Owen (1988) to describe a procedure that uses likelihood as a measure of distance to develop an estimation procedure that is robust against distributional assumptions as GMM, but with the good properties of a parametric likelihood procedure. Kitamura (2006) provides an excellent survey of these methods.

<sup>5</sup>Bernardo and Ledoit (2000), and Cochrane and Sa-Requejo (2000) propose bounds under the assumption of incomplete markets. Cerny (2004) derives the discount factor restrictions implied by standard utility functions.

<sup>6</sup>We use the word nonparametric as opposed to parametric models such as CAPM, CCAPM, APT, among others. Our bounds are solutions to utility maximization problems, and once we choose a particular utility function (corresponding to a certain discrepancy measure), the only parameters appearing are Lagrange Multiplier representing portfolio weights.

<sup>7</sup>It is worth noting that the price of the volatility contract is essential in constructing volatility bounds that differ from the HJ bound in an unconditional setting.

<sup>8</sup>Balduzzi and Kallal (1997) add risk variables in addition to asset returns and obtain more stringent bounds than the HJ bounds. Kan and Zhou (2006) tighten the HJ bound by assuming that the pricing

To assess the empirical implications of these generalized discount factor bounds, we begin by examining size portfolios as in Snow(1991) and Stutzer (1995). We find that the bounds based on the HARA utility with a medium risk aversion coefficient differ significantly from the bounds produced by HJ, and EL or ET divergence measures. To measure the relative information gain from including small firms in the investor’s portfolio we use the ratio of the bound for the large and small firms over the bound for the large firms only. The EL criterion, corresponding to the log utility function, produces the highest gain. This can be rationalized by the fact that EL puts more weight on higher moments than ET. Otherwise we confirm and extend the results in Snow (1991).

The EL, ET or other information-theoretic bounds provide a set of so-called implied probabilities.<sup>9</sup> These probabilities weight optimally the time series of asset payoffs, indicating the importance of each return observation in building the SDF bounds compared with the uniform empirical probabilities of  $(1/T)$ , where  $T$  is the number of available observations in the sample. These implied probabilities are used to compute the time series of the implicit nonparametric SDFs for all discrepancy measures and for each SDF mean. An important consequence of our methodology is to guarantee that the implied SDFs remain always positive, contrary to the linear SDF in Hansen and Jagannathan (1991) without positivity constraints. In addition, the implied SDFs are instructive for the discounting weights they put on certain extreme events in the sample.

Several authors have proposed nonparametric models of the SDFs to address the violations associated with simpler linear models such as the CAPM. Bansal et al. (1993), Bansal and Viswanathan (1993), and Chapman (1997) propose SDFs where the pricing kernel is not linear in the market return. However, Dittmar (2002) remarks that the models in these papers are ad-hoc specifications of either the set of priced factors or the form of nonlinearity. The pricing kernels do not follow directly from assumptions on investors’ preferences or return distributions, as it does in the CAPM reference model. Instead, Dittmar (2002) starts with an approximation of an unknown marginal utility function by a Taylor series expansion but restricts the polynomial terms in the expansion by imposing decreasing absolute prudence. Therefore, the risk factor obtains endogenously from preference assumptions and is a sole function of aggregate wealth.

Our approach is to use the most common utility functions found in the portfolio liter-

---

kernel is a reduced-form function of a finite set of state-variables.

<sup>9</sup>See Back and Brown (1993) for defining implied probabilities in GMM estimators, Owen (1988) for EL, Kitamura and Stutzer (1997) for ET, and Imbens et al. (1998) for Cressie Read discrepancy estimators. Brown and Newey (2002) provide implied probabilities for Generalized Empirical Likelihood (GEL) estimators. Smith (2004) shows how these probabilities can be used to obtain efficient moment estimation as in Brown and Newey (1998).

ature and to derive the SDFs that are consistent with these utility functions.<sup>10</sup> It happens that these implied SDFs correspond to one step estimators that have been recently proposed in the econometric literature to generalize GMM [see Kitamura (2006) and references therein].

The stochastic discount factor approach has also been widely used to measure the performance of managed portfolios. With this approach, abnormal performance is measured by the expected product of a portfolio's returns and a stochastic discount factor. The evaluation can proceed unconditionally or conditionally to a set of lagged instruments. Ferson and Siegel (2003) evaluate the conditional performance of a set of U.S. equity mutual funds using several SDF models. Other papers have analyzed the performance of hedge funds using the stochastic discount factor and GMM (Generalized Method of Moments) approach of Hansen and Jagannathan (1991, 1997).<sup>11</sup> An important common approach to all these studies is the use of GMM to estimate the parametric models of the SDF, either linear or nonlinear.<sup>12</sup> In other words, some stochastic discount factors include higher moments of returns but their estimators are based on a quadratic objective function where only the first two moments are taken into account.

In contrast with previous studies, we exploit the generality of the Minimum Discrepancy functions defining our SDF bounds and their corresponding implied SDFs to evaluate the performance of hedge funds. These funds use dynamic trading strategies whose returns exhibit nonlinear patterns. To capture these nonlinearities and measure an alpha performance of hedge funds, Agarwal and Naik (2004) introduce linear portfolios of options along with other risk factors.<sup>13</sup> Diez de los Rios and Garcia (2006) estimate and test the presence of nonlinearities in hedge fund returns and determine whether hedge funds provide value to investors.

Nonparametric empirical likelihood methods and other information-theoretic estimators have been proposed since they provide the same first-order asymptotic properties as

---

<sup>10</sup>Brandt (1999) proposes a conditional Euler equation approach to obtain optimal consumption and portfolio weights consistent with a representative agent maximizing a CRRA utility function [see extensions in Ait Sahalia and Brandt (2002), Brandt and Santa Clara (2006) and Brandt et al. (2007)]. Despite some intersection between our unconditional analysis of HARA utility maximization problems and Brandt's work, we introduce fairly general Minimum Discrepancy information bounds previously unexplored in asset pricing problems.

<sup>11</sup>See in particular, Ferson and Siegel (2001), Farnsworth et al. (2002), Bailey, Li and Zhang (2004), Kazemi and Schneeweis (2003) and Wang and Zhang (2005).

<sup>12</sup>Exceptions include Chen and Knez (1996), Dahlquist and Soderlind (1999), and Farnsworth et al. (2003) who, similarly to Hansen and Jagannathan (1991), estimate a "primitive-efficient" SDF, which is essentially a non-parametric linear SDF implied by a linear projection on the payoff space of benchmark instruments.

<sup>13</sup>See also Mitchell and Pulvino (2001) and Fung and Hsieh (2001) for characterizing hedge fund returns as returns from option-based trading strategies.

GMM but with several higher-order asymptotic advantages and also good finite-sample properties.<sup>14</sup> In this sense, our analysis complements the above mentioned studies by using estimation criteria that account explicitly for higher moments of returns and not simply the first two moments. These higher moments are induced by the option-like strategies and are a defining characteristic of hedge fund returns. We find that the various criteria almost always agree on the best and worst fund categories but vary in the ranking of the intermediate ones. The alpha valuations obtained with the implied SDFs are reasonable and in line with the results produced by other methods that have been used to account for the nonlinearities in hedge fund returns. In addition, the strict positivity of our nonparametric implied SDFs inherit all the good properties listed by Chen and Knez (1996) for strictly positive performance measures.

The rest of the paper is organized as follows. In section 2, we describe how the Minimum Discrepancy bounds are derived starting from a portfolio optimization problem with a HARA utility function. We show how the EL and ET measures can be obtained as special cases of this general problem. We also derive the respective SDFs from the measure-implied probabilities. In Section 3 we explore the differences between these various bounds for a set of size portfolios. We also compare the relative information gains brought about by the various measures. Section 4 contains a performance evaluation of hedge funds. We describe our semi-nonparametric approach to assessing performance and contrast it with the parametric approaches used in the literature. Section 5 concludes.

## 2 Minimum Discrepancy Estimators and SDF Bounds

### 2.1 MD Estimators

Consider a model expressed through a set of moment conditions  $E[g(z_i, \beta)] = 0, i = 1, \dots, T$ , where  $\{z_i\}_{i=1, \dots, T}$  represents a time series of random vectors,  $g(\cdot, \beta)$  is a vector in  $R^m$ , and  $\beta$  is an unknown vector of parameters within a set  $B \in R^K$ . Minimum Discrepancy (MD) Estimators (Corcoran (1998)) seek to estimate  $\beta$  by providing a set of probabilities  $\pi$  that will reweight the sample empirical probabilities  $\frac{1}{T}$  to satisfy the moment conditions exactly. The discrepancy between the empirical probabilities  $\frac{1}{T}$  and  $\pi$  is measured through a convex divergence function  $\phi$  in the spirit of Csiszàr (1967).<sup>15</sup>

The parameter vector  $\beta$  and the probability measure  $\pi$  are estimated by:

$$\{\beta_{MD}, \pi_{MD}\} = \underset{\beta \in B, \{\pi_1, \dots, \pi_T\}}{\text{arg min}} \frac{1}{T} \sum_{i=1}^T \phi(\pi_i), \text{ subj. to } \sum_{i=1}^T \pi_i g(z_i, \beta) = 0_K, \sum_{i=1}^n \pi_i = 1 \quad (1)$$

<sup>14</sup>See Imbens (1997), Imbens et al. (1998), Kitamura (2001), Newey and Smith (2004), and Kitamura and Otsu (2005).

<sup>15</sup>See Kitamura (1996, 2006) for a continuous time formulation of this problem.

where  $0_K$  represents an  $m$ -dimensional vector of zeros.

By solving this problem one essentially obtains a point estimate  $\hat{\beta}_{MD}$  for the vector of parameters that satisfies the moment conditions and minimizes the discrepancy  $\phi$  between the implied probabilities  $\hat{\pi}_{MD}$  and the uniform weights  $\frac{1}{T}$ . Moreover, these implied probabilities (Back and Brown (1993)) can be used to obtain an efficient estimation of moment conditions (Brown and Newey (1998)) and probability distribution functions via bootstrapping schemes (Brown and Newey (2002)).

Newey and Smith (2004) showed that under the Cressie Read (1984) family of discrepancies, the solution for the optimization problem (1) can be obtained by looking at its dual version. In fact, the dual optimization problem for MD estimators defines the important class of Generalized Empirical Likelihood (GEL) estimators proposed by Smith (1997). This class encompasses a large set of one-step alternative estimators to the two-step GMM estimator, which we analyze in our work, including EL [Owen (1988), Imbens (1997)], ET [Kitamura and Stutzer (1997), Imbens et al. (1998)], and CUE [Hansen et al. (1996)], among others.<sup>16</sup>

Changing from the original problem (1) to the dual problem represents a huge advantage from an operational point of view, since while the inner optimization on the original MD problem is based on the space of discrete probability measures with dimension  $T$  (sample dimension), the inner optimization on the dual GEL problem is based on a space of dimension  $m$  (image of function  $g$ ):

$$\left\{ \hat{\beta}, \hat{\lambda} \right\} = \underset{\beta \in B}{arg \min} \sup_{\alpha \in \mathfrak{R}, \lambda \in \Lambda(\beta)} \alpha - \sum_{i=1}^T \frac{1}{T} \phi^*(\alpha + \lambda' g(z_i, \beta)) \quad (2)$$

where  $\phi^*(y) = \sup_{x \in \mathfrak{R}} (xy - \phi(x))$  denotes the convex conjugate of  $\phi(\cdot)$ , and  $\lambda$  is an  $m$ -dimensional vector of auxiliary parameters representing Lagrange Multipliers coming from the restrictions present in problem (1) (see Newey and Smith (2004); Kitamura (2006)). In addition, for each value  $\beta$  of the original parameter vector,  $\Lambda(\beta) \in \mathfrak{R}^m$  represents the domain where function  $\phi^*$  is well defined and concave.

Under the Cressie Read family, the discrepancy function is given by  $\phi(\pi) = \frac{(\pi)^{\gamma+1} - 1}{\gamma(\gamma+1)}$ ,

---

<sup>16</sup>Most of these one-step alternative estimators have been proposed under the assumption of i.i.d random variables  $\{z_i\}_{i=1, \dots, T}$ , with the exception of Kitamura and Stutzer (1997) who proposed ET under serially dependent data, and Imbens (1997) which suggests alternative procedures to deal with dependence. However, generalizations of MD/GEL estimators to explicitly consider serially dependent data are available and were shown to work well [see for instance, Smith (2004) and Kitamura (2006)]. In this case, estimation is preceded by a smoothing procedure of the moment conditions using kernel function weighting methods.

and the dual problem becomes<sup>17</sup>:

$$\left\{ \hat{\beta}_{GEL}, \hat{\lambda}_{GEL} \right\} = \arg \min_{\beta \in B} \sup_{\lambda \in \Lambda(\beta)} \frac{1}{T} \sum_{i=1}^T -\frac{1}{\gamma + 1} (1 + \gamma \lambda' g(z_i, \beta))^{\frac{\gamma+1}{\gamma}} = \arg \min_{\beta \in B} \sup_{\lambda \in \Lambda(\beta)} \sum_{i=1}^T M(\lambda' g(z_i, \beta)) \quad (3)$$

Theorem 2.2 in Newey and Smith (2004) states that the solutions to problems (1) and (3) are coincident under the Cressie Read family, meaning that  $\hat{\beta}_{MD} = \hat{\beta}_{GEL}$ . Moreover, the theorem shows how to recover the implied probabilities  $\hat{\pi}_{MD}$  via the first derivatives of the function  $M$ , evaluated at the optimal set of parameters  $\hat{\beta}_{GEL}$  and Lagrange Multipliers  $\hat{\lambda}_{GEL}$  that solve (2):

$$\hat{\pi}_{MD}^i = \frac{M_1(\hat{\lambda}'_{GEL} g(z_i, \hat{\beta}_{GEL}))}{\sum_{j=1}^T M_1(\hat{\lambda}'_{GEL} g(z_j, \hat{\beta}_{GEL}))} \quad (4)$$

with  $M_1(v) = \frac{dM(v)}{dv}$ .

Equation (4) will be particularly important since we later use a simplified version to construct the nonparametric SDFs that will generalize Hansen and Jagannathan's (1991) linear SDF. Note that the optimal Lagrange Multipliers  $\hat{\lambda}_{GEL}$ 's play a fundamental role to determine the probabilities  $\pi$  that will optimally reweight the sample empirical probabilities, and consequently define our generalized implied SDFs.

## 2.2 Minimum Discrepancy SDF Bounds

Here, we will be interested in providing SDF bounds that account for higher moments of returns generalizing the linear projection approach of Hansen and Jagannathan (1991).

Given a set of asset returns, the HJ minimum variance SDF bound imposes a restriction on the space of SDFs that are able to price those returns *by directly constraining their variances*. A natural generalization of this bound consists in proposing restrictions on the space of SDFs based on *discrepancy functions* more general than a quadratic one, taking into account higher moments of the family of SDFs pricing the original returns. This is precisely the approach we follow here. Interestingly, the duality existing between the mean-variance SDF sub-space and the mean-variance portfolio frontier will be extended, linking restrictions on higher moments of SDFs to information on higher moments of returns.

In order to define our bounds, let  $R$  denote the vector of returns whose realizations are given by a time series of asset returns  $\{R_i\}_{i=1, \dots, T}$  in a  $K$ -dimensional space. We propose MD bounds by making use of the definition of an SDF as a set of moment conditions, the Euler equations:

$$E(mR) = 1_K \quad (5)$$

---

<sup>17</sup>Here,  $\gamma = 0$  and  $\gamma = -1$  should be understood as limiting cases for function  $\phi(\pi)$ .

where  $1_K$  represents a  $K$ -dimensional vector of ones. On its sample form, for an SDF with mean  $a$ , the Euler equation becomes:

$$\frac{1}{T} \sum_{i=1}^T m_i \left( R_i - \frac{1}{a} \right) = 0_K \quad (6)$$

Now, given a convex and homogeneous discrepancy function  $\phi$ , our Minimum Discrepancy SDF bound will be given by:

$$\bar{m} = \arg \min_{\{m_1, \dots, m_T\}} \frac{1}{T} \sum_{i=1}^T \phi(m_i), \text{ subject to } \frac{1}{T} \sum_{i=1}^T m_i \left( R_i - \frac{1}{a} \right) = 0, \frac{1}{T} \sum_{i=1}^T m_i = a \quad (7)$$

where condition  $\sum_{i=1}^T m_i \left( R_i - \frac{1}{a} \right) = 0$  essentially imposes the Euler equation to any SDF  $m$  (with mean  $a$ ) pricing  $R$ . This optimization problem restricts the space of possible SDFs pricing  $R$ , with restrictions coming directly from this general discrepancy function  $\phi$ <sup>18</sup>.

We are interested in transforming problem (7) to an authentic Minimum Discrepancy problem in the space of sample probability measures rather than sample SDFs. By the homogeneity of function  $\phi$ , each SDF  $m$  with mean  $a$  can be associated with a corresponding probability measure  $\pi(m)$ , via a simple linear transformation  $\pi(m) = \frac{m}{T \cdot a}$ . Once this transformation is performed, the corresponding equivalent MD bound problem is given by:

$$\bar{\pi} = \arg \min_{\{\pi_1, \dots, \pi_T\}} \frac{(aT)^\alpha}{T} \sum_{i=1}^T \phi(\pi_i), \text{ subject to } \sum_{i=1}^T \pi_i \left( R_i - \frac{1}{a} \right) = 0, \sum_{i=1}^T \pi_i = 1 \quad (8)$$

where  $\alpha$  denotes the degree of homogeneity of function  $\phi$ .

In particular, when  $\phi$  belongs to the Cressie Read discrepancy family, the duality results provided in Section 2.1 by Equations (1)-(3) simplify the problem of obtaining MD SDF bounds to the following GEL problem:

$$\hat{\lambda}_{GEL} = \sup_{\lambda \in \Lambda_{CR}} \frac{(aT)^{(\gamma+1)}}{T} \sum_{i=1}^T -\frac{1}{\gamma+1} \left( 1 + \gamma \lambda' \left( R_i - \frac{1}{a} \right) \right)^{\frac{(\gamma+1)}{\gamma}} \quad (9)$$

where  $\gamma$  is a fixed hyperparameter appearing in the Cressie Read discrepancy function  $\phi(\pi) = \frac{(\pi)^{\gamma+1}-1}{\gamma(\gamma+1)}$ , and  $\Lambda_{CR} = \{\lambda \in R^K, \text{ s.t. for } i = 1, \dots, T : (1 + \gamma \lambda' (R_i - \frac{1}{a}))^{\frac{1-\gamma}{\gamma}} > 0\}$ .

---

<sup>18</sup>Restrictions on the space of SDFs more general than those imposed by variance have been analyzed by Snow (1991), Stutzer (1995), Bansal and Lehmann (1997), Bernardo and Ledoit (2000), Cochrane and Saa-Requejo (2000), and Cerny (2004). However, connections with the econometric literature of one-step estimators alternative to GMM have not been proposed, an exception being Stutzer (1995). While he analyzes the specific case of the ET estimator, we generalize his work by considering the whole class of Cressie Read discrepancy functions.

From an econometric point of view,  $\hat{\lambda}_{GEL}$  represents the solution of a specific GEL estimator. Indeed, observe that different choices of  $\gamma$  lead to distinct well known one-step alternative estimators to GMM. For instance, when  $\gamma = -1$ ,  $\phi(\pi) = -\ln \pi$  and the EL estimator is obtained; when  $\gamma = 0$ ,  $\phi(\pi) = \pi \ln \pi$ , and ET is obtained; and when  $\gamma = 1$ , CUE is obtained.

Most importantly, note that via a simplified version of Equation (4), these Lagrange Multipliers allow us to directly recover the MD implied probabilities that optimally reweight the sample:

$$\hat{\pi}_{MD}^i = \frac{M_1(\hat{\lambda}'_{GEL} (R_i - \frac{1}{a}))}{\sum_{j=1}^T M_1(\hat{\lambda}'_{GEL} (R_j - \frac{1}{a}))} \quad (10)$$

with  $M_1(v) = \frac{dM(v)}{dv}$ , and  $M(v) = -\frac{1}{\gamma+1}(1 + \gamma v)^{\frac{\gamma+1}{\gamma}}$ . From these probabilities it is straightforward to obtain the corresponding SDF  $m_{MD}$  that minimizes the CR discrepancy:

$$m_{MD}^i = T * a * \frac{(1 + \gamma \hat{\lambda}'_{GEL} (R_i - \frac{1}{a}))^{\frac{1}{\gamma}}}{\sum_{j=1}^T (1 + \gamma \hat{\lambda}'_{GEL} (R_j - \frac{1}{a}))^{\frac{1}{\gamma}}} \quad (11)$$

Note that the general form for the MD SDF solving the bound problem will be a hyperbolic function of the original returns  $R$ :

$$m_{MD}(R) = \beta * \left(1 + \gamma \hat{\lambda}'_{GEL} \left(R - \frac{1}{a}\right)\right)^{\frac{1}{\gamma}} \quad (12)$$

Particular yet important cases will appear for EL, ET and CUE. For EL,  $\gamma = -1$  and the optimal SDF becomes:

$$m_{EL}(R) = \beta * \frac{1}{(1 - \hat{\lambda}'_{EL} (R - \frac{1}{a}))} \quad (13)$$

For ET,  $\gamma = 0$  and the optimal SDF will be an exponential function of excess returns:

$$m_{ET}(R) = \beta * e^{\hat{\lambda}'_{ET} (R - \frac{1}{a})} \quad (14)$$

Finally, for CUE,  $\gamma = 1$  and similarly to HJ, the optimal SDF will be a linear function of excess returns:

$$m_{CUE}(R) = \beta * \left(1 + \hat{\lambda}'_{CUE} \left(R - \frac{1}{a}\right)\right) \quad (15)$$

On the other hand, note that there is also room for an interesting economic interpretation for problem (9) as an optimal portfolio problem: The solution of our MD bound for each Cressie Read estimator will correspond to an optimal portfolio problem based on a specific HARA-type utility function

$$u(W) = -\frac{1}{\gamma+1} (1 + \gamma W)^{\frac{\gamma+1}{\gamma}}, \quad (16)$$

with  $W > -\frac{1}{\gamma}$  sufficient to guarantee the concavity and strict monotonicity of function  $u$ .

For the econometric estimators EL, ET, and CUE discussed above, we immediately identify corresponding optimal portfolio problems based on respectively logarithmic, exponential, and quadratic utility functions. In particular, Stutzer (1995) proposed a portfolio interpretation for the ET estimator based on a standard two-period model of optimal portfolio choices (see Huang and Litzenberger (1988)). Based on the same two-period model, we extend this interpretation to the whole Cressie Read family. Suppose an investor distributes his/her initial wealth  $W_0$  putting  $\lambda_j$  units of wealth on the risky asset  $R_j$  and the remaining  $W_0 - \sum_{j=1}^K \lambda_j$  in a risk-free asset paying  $r_f = \frac{1}{a}$ . Terminal wealth is then  $W = W_0 * r_f + \sum_{j=1}^K \lambda_j * (R_j - r_f)$ . Assume in addition that this investor maximizes the HARA utility function provided above in equation (16), solving the following optimal portfolio problem:

$$\Omega = \sup_{\lambda \in \Lambda} E(u(W)) \quad (17)$$

where  $\Lambda = \{\lambda : u(W(\lambda)) \text{ is strictly increasing and concave}\}$ . Note that by scaling the original vector  $\lambda$  to be  $\tilde{\lambda} = \frac{\lambda}{(1 + \frac{\gamma W_0}{a})}$ , we can decompose the utility function in  $u(W) = u(W_0 * r_f) * \left(1 + \gamma \tilde{\lambda} (R - \frac{1}{a})\right)^{\frac{\gamma+1}{\gamma}}$ . This decomposition essentially shows that solving the GEL optimal problem in (9) will measure the gain when switching from a total allocation of wealth at the risk-free asset paying  $r_f$  to an optimal (in the utility  $u$  sense) diversified allocation that includes both risky assets and the risk-free asset.

Now, in order to complete our characterization of Minimum Discrepancy SDF bounds, we provide an operational algorithm to obtain such bounds when there is no risk-free asset in the space of returns. Similarly to HJ, the idea is to propose a grid of possible meaningful values for the SDF mean, say fixing a set  $A = \{a_1, a_2, \dots, a_J\}$ , and to solve the optimization problem in (9) for each  $a_l \in A$ , obtaining a corresponding optimal weights vector  $\hat{\lambda}_{GEL}(a_l)$  for each SDF mean. The MD SDF bound is given by the following expression:

$$I_{CR}(a_l, \gamma) = \frac{(a_l T)^{(\gamma+1)}}{T} \sum_{i=1}^T -\frac{1}{\gamma+1} \left(1 + \gamma \hat{\lambda}_{GEL}(a_l)' \left(R_i - \frac{1}{a_l}\right)\right)^{\frac{\gamma+1}{\gamma}}, l = 1, 2, \dots, J \quad (18)$$

Alternatively, we can go back to the basic definition of the bound as a minimum discrepancy/distance problem, and write the solution by first obtaining the optimal implied probabilities appearing in Equation (10)  $\hat{\pi}_{MD}^i(a_l)$ , and substitute then in the sampled divergence function  $\phi$ , obtaining:

$$I_{CR}(a_l, \gamma) = \frac{(a_l T)^{(\gamma+1)}}{T} \sum_{i=1}^T \frac{\hat{\pi}_{MD}^i(a_l)^{\gamma+1} - 1}{\gamma(\gamma+1)}, l = 1, 2, \dots, J \quad (19)$$

As a matter of fact, expressions in (18) and (19) should be equivalent according to Theorem 2.2 by Newey and Smith (2004).

### 3 An Empirical Illustration with Size Portfolios

Snow (1991) explored the implications of looking at higher-moment bounds for size portfolios. The advantage was to analyze the small firm effect in a nonparametric way and to test the methodology with asset returns that exhibit skewness and kurtosis. He concludes that, for certain time periods, higher moments of small firms place more severe restrictions on stochastic discount factors than do lower moments and that small firms contain information about the distribution of the SDF that is not found in large firms.

We perform a similar analysis with the various bounds we described in the previous section. We analyze for each criterion - EL, ET and CR - the bound that is produced by including only large firms, only small firms and both large and small firms. We identify small firms as the three lowest size portfolios in the ten decile size portfolios constructed by Fama and French. The large firms are similarly defined as the three largest size portfolios. The returns are monthly over the period 1926:7-2006:11.<sup>19</sup>

Let us start with the ET information bound in panel a of Figure 1. We find that small firms impose stricter restrictions than large firms in a region for the SDF mean between 0.992 and 0.998. This region corresponds to annual values of the interest rate between 2.4% and 9.7%. This result is similar to what Snow (1991) reports for the three norms he selected - 3, 2 and 3/2. Adding small firms to large firms produces a frontier that is always above the frontiers of large or small firms taken alone. The EL bounds on panel b of Figure 1 behave similarly to the ET bounds, but the interval of SDF means for which the small firms are more restrictive is slightly larger.

For the Cressie-Read criterion, which corresponds to the maximization of a HARA utility function, we need to choose the parameter that sets the risk aversion of the agent. We should remember that EL corresponds to  $\gamma = -1$ , while ET is obtained when  $\gamma = 0$ . In Figure 2, we produce the bounds for values of 1 and 3 for  $\gamma$ . The first value corresponds to a quadratic utility function, and to the CUE estimator by Hansen et al. (1996). So the bounds should be similar to the Hansen-Jagannathan variance bounds. Indeed, the plots in Panels a and b of Figure 2 confirm this. The shapes of the CR information bound with  $\gamma = 1$  and of the HJ bound are almost the same for the small, large and small and large firms except that the small firms add less information for higher mean values of the SDF. When agents display more risk aversion,  $\gamma = 3$ , the bound curves for all firms and large

---

<sup>19</sup>The returns were obtained from the data library on the web site of Kenneth French [http : //mba.tuck.dartmouth.edu/pages/faculty/ken.french/data\\_library.html](http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html).

firms only cross. There is now a loss of information associated with the small firms.

A good way to measure the relative information gain from including small firms is to plot the ratio of the bound for the large and small firms over the bound for the large firms only. In Figure 3, we report this relative information gain measure for all five criteria. The EL criterion, corresponding to the log utility function, produces the highest gain. This can be rationalized by the fact that EL puts more weight on higher moments than ET. A fourth-order expansion in Appendix A shows that the weights on these higher moments are decreasing functions of  $\gamma$  (see Figure 4). Therefore, the less risk averse investor will tend to include relatively more small firms in his portfolio. It is interesting to note that the largest differences appear to occur around a SDF mean of 0.992 (an interest rate close to 10%). This corresponds to states of the world where small firms will tend to experience more difficulties because of high interest rates.

Another way of comparing the various criteria is to analyze the SDFs produced by each of them. In Figure 5, we compare the SDF graphs of ET and EL for all firms and for a specific value of the SDF mean of 0.996, the middle point of the interval where the small and large firm bound curves intersected. This corresponds to an interest rate of 4.8%, a value a bit higher than the average over the 1926-2006 sample. The EL SDF appears more volatile than the ET SDF, with very high values attributed to crashes. Recall that the SDF is constructed by multiplying the implied probabilities by a constant. The measure tilts the probabilities to satisfy the pricing moment condition and therefore accommodates large negative values with high probabilities. ET produces a similar SDF but the implied probabilities are lower for these events.

In Figure 6, we include the SDF graphs for HJ, CR ( $\gamma = 1$ ) and CR ( $\gamma = 3$ ). The first important remark is that for one observation the HJ stochastic discount factor is negative. It is known that the HJ criterion can produce negative values for the SDF without a positivity constraint. This cannot happen when we find a SDF that corresponds to the solution of an optimal portfolio problem. Indeed, the CR measure with  $\gamma = 1$ , which is the closest to HJ since it results from finding the optimal portfolio with a quadratic utility function, attributes a value of zero to the same observation.

#### 4 Performance Evaluation of Hedge Funds

Performance measurement and portfolio choice have evolved hand in hand for a long time. The Treynor (1965), Sharpe (1966) and Jensen (1968) early measures have followed in the footsteps of modern portfolio theory. More recently, researchers have used a stochastic discount factor approach to evaluate the performance of portfolio managers [see in particular Glosten and Jagannathan (1994), Chen and Knez (1996), Farnworth et al. (2002), and

Dahlquist and Soderlind (1999)]. Although most of the stochastic discount factor models are parametric (linear CAPM or APT-based models, consumption-based asset pricing models, numeraire portfolios), some are nonparametric such as the so-called primitive efficient portfolio in Farnworth et al. (2002). In this approach, a GMM estimation procedure is used to estimate the alpha of the managed portfolios. Most often a conditional evaluation is performed to verify that a positive performance is not simply the reflection of some publicly available information and is really attributable to the ability or superior information of the manager.

We maintain and extend the SDF approach with several nonparametric SDFs that are consistent with utility functions most often used in portfolio optimization. In other words, we align the performance measure and the portfolio decision criterion. However, we assess performance unconditionally for several reasons. First, the conditional approach refers to managed portfolios but the conditioning information is always public since the true private information, which is in fact used to actually manage the portfolios, is unobservable. Therefore, with respect to this private information, we observe the portfolio returns always unconditionally. It means in particular that this absence of conditioning will generate fatter tails. Indeed, through the consideration of higher order moments, Garcia, Renault and Tsafack (2007) extend the argument of Clark (1973) to note that, in general, the more information we condition the returns of a portfolio upon, the thinner are the tails of the resulting distribution. Since our SDFs incorporate higher moments, we can better capture the effect of this unobservable private information. A second argument has to do with the difficulties associated with the conditional tests. Dahlquist and Soderlind (1999) report that the power of the tests decreases with the conditioning information and that numerical problems are encountered when one wants to maintain the positivity of the SDF in a conditional setting. Our unconditional SDFs remain positive. To verify whether a positive performance is not simply the reflection of some publicly available information, we can always regress the resulting alphas on some observable variables.

#### **4.1 Stochastic Discount Factors Extracted from Risk Factors**

In order to obtain the implied SDF for each IT estimator to be used in the performance evaluation experiment, we must solve an optimal portfolio problem under the particular utility function defined by that estimator. To derive the implied SDFs, we solve the optimal portfolio problems with monthly time-series (January 1996 to March 2004) of five relevant market risk factors, following Hasanhodzic and Lo (2007). These five factors provide a reasonable set of risk exposures for a typical hedge fund (see Diez and Garcia, 2007 and references therein). We first include the CRSP value-weighted NYSE, AMEX, and NAS-

DAQ combined index as a stock market measure; we use the return on an equally weighted portfolio of British, German and Japanese one-month eurocurrency deposits to capture any exposure to an exchange rate (FX) factor; we use the return on the Lehman U.S. Corporate AA Intermediate Bond Index to capture bond market risk; to capture a credit risk factor we use the return on the Lehman U.S. Corporate BAA Intermediate Bond Index in excess of the return on the Lehman U.S. Treasury index; and finally we include the return on the Goldman Sachs Commodity index as the fifth factor.

Table 1 presents the optimal weights given by each IT estimator to the risk factors. Numbers between parenthesis represent the percentage of the optimal portfolio given to each factor. Interestingly, as can be observed, the optimal weights for all IT estimators given to the Credit Risk factor are the highest among all risk factors, followed by the Bond factor. Those two factors cover around 85% of the optimal portfolio allocations of all estimators. Note that, according to the table, all the IT estimators are actually heavily buying U.S. Treasury and Corporate bonds. In addition, commodities represent only tiny portions of their optimal portfolios, while the S&P index is sold (around 5% of the portfolios), and the weights are not too different on the currency factor. While ET and EL sell currency, the Cressie Read ones buy currency.

An important issue regarding the characterization of hedge fund strategies is the presence of nonlinearities in the form of option-like features. Several papers have addressed this issue. Fung and Hsieh (2001) analyze trend-following strategies and show that their payoffs are related to those of an investment in a lookback straddle. Mitchell and Pulvino (2001) show that returns to risk arbitrage are similar to those obtained from selling uncovered index put options. Agarwal and Naik (2004) extend these results and show that, in fact, a wide range of equity-oriented hedge fund strategies exhibit this non-linear payoff structure. In particular, they use a stepwise regression procedure to identify the significant risk factors. To account for non-linearities, they include option-based risk factors that consist of returns obtained by buying, and selling one month later, liquid put and call options on the Standard & Poor's (S&P) 500 index. More recently, Hasanhodzic and Lo (2007) introduce as a risk factor the first difference in the end-of-month value of the CBOE volatility index (VIX), which can be interpreted as a proxy for the return on the portfolio of options used to compute the VIX, while Diez and Garcia (2007) propose a procedure to find the optimal nonlinear features in a set of hedge fund returns. Given this evidence it is important to include in the set of risk factors returns on option portfolios along with the returns on the risk factors. We include four option factors used in Agarwal and Naik (2004), one out-of-the-money put factor, one out-of-the-money call factor, one at-the-money put factor and one at-the-money call factor. Table 2 presents the optimal weights given by each IT

estimator once the option factors are added to the originally adopted risk factors. Overall, the weights attributed to the original risk factors are not very distinct from the no-options case. However, put options appear to be especially important in the optimal portfolios, being responsible for around 4% of the total allocation of all IT estimators, except for CR ( $\gamma = 1$ ). For this estimator, which is closer to the HJ mean-variance estimator, options do not appear to be important having a total participation of only 1% of the portfolio. This is consistent with the idea that once kurtosis and skewness are significantly weighted in the utility function, then options become more important in the optimal portfolio (see Appendix A).

Figures 7 and 8 exhibit the time series of SDFs implied by the optimal portfolio weights, with and without the option factors. The EL and ET estimators in figure 7 provide very similar patterns for the dynamics of the SDFs with coinciding peaks and troughs, but the variance of the EL SDF is higher than the variance of the ET SDF. The option factors change the SDFs significantly only at specific times associated with particular market conditions. These differences are more subdued with the HJ and CR ( $\gamma = 1$ ) estimators as shown in figure 8. We summarize the relative information gain for the various estimators when options are included in figure 9. It is observed that the relative gain is a declining function of the risk aversion coefficient in the HARA utility function<sup>20</sup>. Also the relative gains increase with the mean of the SDFs.

We will now use these SDFs to assess the performance of several indices of hedge funds built from the TASS database. It provides monthly returns and net asset value data on 4,606 funds beginning in February 1977. For building these hedge fund indexes, our sample starts in January 1996 and ends in March 2004.<sup>21</sup> The individual funds are classified into eleven categories: 1) convertible arbitrage; 2) fixed-income arbitrage; 3) event driven; 4) equity market neutral; 5) long-short equity; 6) global macro; 7) emerging markets; 8) dedicated short bias; 9) managed futures. Appendix B gives a brief description of the typical strategies followed in each category.

Being able to access the optimal portfolio weights is very useful to analyze results of the performance that each estimator will attribute to the different hedge fund categories. Indeed, based on the time series of risk factors' returns, we can directly link the behavior of the IT implied SDFs to market conditions and macroeconomic factors. Figure 10 shows an example on how EL and CR ( $\gamma = 3$ ), two of our extreme IT estimators, price the equity market neutral hedge fund category. Note that the behavior of the implied SDFs is very

---

<sup>20</sup>By also observing the implied time series of each SDF, it is observed that the SDFs variances also decline with risk aversion.

<sup>21</sup>We choose to start in 1996 to have a reasonable representation for all categories of funds. Also, the TASS database does not give any information on exited funds prior to 1994.

different in moments of market crises and booms. In the graph we point to two significant events, as the October 1998 crisis and the credit market improvement in the second half of 2001. This graph is illustrative of the in-depth chronological analysis that can be conducted for the implied SDFs.

These differences in the dynamics of the SDFs translate directly in the performance evaluation of various categories of hedge funds. In Table 3, we report the alphas associated with the various estimators, with and without the option factors. For comparison we also report the performance evaluation corresponding to a linear model of the risk factors, a model with the Agarwal and Naik (2004) put and call option factors, and the three Diez and Garcia (2007) estimators with different volatilities.

In general, the estimators agree on the two extreme categories, the best (convertible arbitrage, C1) and the worse (managed futures, C9), but vary a bit more across the other categories. However the results are relatively robust regarding the positive or negative assessment of performance. The presence of options can change this sign from positive to negative as in the long-short equity hedge (C5). It means that if nonlinearities are present they may just reflect the use of derivatives that hedge funds pay for on the market and not the timing ability of the hedge fund managers.

For some categories such as equity market neutral (C4), we notice an important difference between the assessment of the IT estimators and all the previous estimators. It is consistently negative for the IT estimators, while it is slightly positive for the other estimators except for the high volatility one. However, overall, the evaluation is surprisingly consistent across estimators obtained with our new methodology and the previous ones.

## 5 Conclusion

We extend results on stochastic discount factor variance bounds of Hansen and Jagannathan (1991) by proposing more general Minimum Discrepancy (MD) bounds based on the minimization of discrepancy convex functions. Solutions to these MD problems naturally imply nonparametric nonlinear SDFs that take into account higher moments of the distributions of assets returns. Based on a duality argument of Newey and Smith (2004), we relate the problem of finding general MD bounds to that of solving an optimal portfolio problem, therefore giving economic content to the non-linear implied SDFs. When specializing to the Cressie Read family of discrepancies, our bounds are obtained as solutions to optimal portfolio problems based on HARA utility functions. Moreover, each portfolio problem corresponds to a Generalized Empirical Likelihood estimator as proposed by Smith (1997), with special cases corresponding to important one-step econometric estimators alternative to GMM, namely, Empirical Likelihood, Exponential Tilting, and CUE estimators. We

analyze the use and performance of our implied SDFs in two empirical problems, first revisiting the size portfolio problem of Snow (1991), and then studying hedge fund performance evaluation with different metrics based on our implied SDFs. Results indicate that this new class of higher order SDF bounds has a strong potential to be used in a large number of financial problems, specially those involving assets with nonlinear payoffs.

## References

- [1] Admati A., S. Bhattacharya, P. Pfleiderer, and S.A. Ross (1986). On Timing and Selectivity. *Journal of Finance*, **41**, 3, 715-730.
- [2] Admati A., and S.A. Ross (1985). Measuring Investment Performance in a Rational Expectations Equilibrium Model. *Journal of Business*, **58**, 1, 1-26.
- [3] Agarwal V., and N.Y. Naik (2004). Risk and Portfolio Decisions Involving hedge Funds. *Review of Financial Studies*, **17**, 1, 63-98.
- [4] Ait Sahalia Y., and M. Brandt (2002). Variable Selection for Portfolio Choice. *Journal of Finance*, **56**, 4, 1297-1351.
- [5] Back K., and D.P. Brown (1993). Implied Probabilities in GMM Estimators. *Econometrica*, **61**, 971-975.
- [6] Bailey W., H. Li, and X. Zhang (2004). Hedge Fund Performance Evaluation: A Stochastic Discount Factor Approach. *Working Paper*, Cornell University.
- [7] Balduzzi P. and H. Kallal (1997). Risk Premia and Variance Bounds. *Journal of Finance*, **52**, 1913-1949.
- [8] Bansal R. and B. Lehman (1997). Growth-Optimal Portfolio Restrictions on Asset Pricing Models. *Macroeconomic Dynamics*, **1**, 333-354.
- [9] Bansal R., D. Hsieh, and S. Viswanathan (1993). A New Approach to International Arbitrage Pricing. *Journal of Finance*, **48**, 1719-1747.
- [10] Bansal R. and S. Viswanathan (1993). No Arbitrage and Arbitrage Pricing: A New Approach. *Journal of Finance*, **48**, 1231-1262.
- [11] Bekaert G. and J. Liu (2004). Conditioning Information and Variance Bounds on Pricing Kernels. *Review of Financial Studies*, **17**, 2, 339-378.

- [12] Bernardo A. and O. Ledoit (2000). Gain, Loss, and Asset Pricing. *Journal of Political Economy*, **108**, 1, 144-172.
- [13] Brandt M. (1999). Estimating Portfolio and consumption Choice: A Conditional Euler Equations Approach. *Journal of Finance*, **54**, 1609-1646.
- [14] Brandt M., and P. Santa Clara (2006). Dynamic Portfolio Selection by Augmenting the Asset Space. *Journal of Finance*, **61**, 2187-2217.
- [15] Brandt M., P. Santa Clara, and R. Valkanov (2007). Parametric Portfolio Policies: Exploiting Characteristics in the Cross Section of Equity Returns. *Working Paper*, Duke University.
- [16] Brown B.W., and W.K. Newey (1998). Efficient Semiparametric Estimation of Expectations. *Econometrica*, **66**, 453-464.
- [17] Brown B.W., and W.K. Newey (2002). Generalized Method of Moments, Efficient Bootstrapping, and Improved Inference. *Journal of Business and Economic Statistics*, **20**, 507-517.
- [18] Cerny A. (2003). Generalized Sharpe Ratios and Asset Pricing in Incomplete Markets. *European Finance Review*, **7** , 191-233.
- [19] Clark P.K. (1973). A Subordinated Stochastic Process Model with finite Variance for Speculative Prices. *Econometrica*, **41**, 135-156.
- [20] Chabi-Yo F. (2008). Conditioning Information and Variance Bounds on Pricing Kernels with Higher-Order Moments: Theory and Evidence. *Review of Financial Studies*, **21**, 181-231.
- [21] Chapman D. (1997). Approximating the Asset Pricing Kernel. *Journal of Finance*, **52**, 1382-1410.
- [22] Chen Z. and P. Knez (1996). Portfolio Performance Measurement: Theory and Applications. *Review of Financial Studies*, **9**, 511-555.
- [23] Cochrane J. and J. Sa-Requejo (2000). Beyond Arbitrage: Good-deal Asset Price Bounds in Incomplete Markets. *Journal of Political Economy*, **108**, 1, 79-119.
- [24] Corcoran S.A. (1998). Bartlett Adjustment of Empirical Discrepancy Statistics. *Biometrika*, **85**, 967-972.

- [25] Cressie N., and T.R.C. Read (1984). Multinomial Goodness-of-fit Tests. *Journal of the Royal Statistics Society Series B*, **46**, 440-464.
- [26] Csiszàr I. (1967). Information-type Measures of Difference of Probability Distributions and Indirect Observations. *Studia Scientiarum Mathematicarum Hungarica*, **2**, 299-318.
- [27] Dahlquist M. and P. Soderlind (1999). Evaluating Portfolio Performance with Stochastic Discount Factors. *Journal of Business*, **72**, 3, 347-383.
- [28] Diez de los Rios A. and R. Garcia (2006). Assessing and valuing the Nonlinear Structure of Hedge Fund Returns. *Working Paper*, University of Montreal.
- [29] Dittmar R. (2002). Nonlinear Pricing Kernels, Kurtosis Preference, and the Cross-Section of Assets Returns. *Journal of Finance*, **57**, 1, 369-403.
- [30] Dybvig P.H., and C. Rogers (1997). Recovery of Preferences from Observed Wealth in a Single Realization. *Review of Financial Studies*, **10**, 151-174.
- [31] Dybvig P.H., and S. Ross (1985). Differential Information and Performance Measurement Using a Security Market Line. *Journal of Finance*, **11**, 2, 383-399.
- [32] Farnsworth H., W. Ferson, D. Jackson, and S. Todd (2002). Performance Evaluation with Stochastic Discount Factors. *Journal of Business*, **75**, 473-504.
- [33] Ferson W. and A. Siegel (2001). The Efficient Use of Conditioning Information in Portfolios. *Journal of Finance*, Vol. **56**, 3, 967-982.
- [34] Ferson W. and A. Siegel (2003). Stochastic Discount Factor Bounds with Conditioning Information. *Review of Financial Studies*, **16**, 567-595.
- [35] Fung W., and D.A. Hsieh (2001). The Risk in Hedge Fund Strategies: Theory and Evidence from Trend Followers. *Review of Financial Studies*, **14**, 2, 313-341.
- [36] Gallant R., L.P. Hansen, and G. Tauchen (1990). Using Conditional Moments of Assets Payoffs to Infer the Volatility of Intertemporal Marginal Rates of Substitution. *Journal of Econometrics*, **45**, 141-179.
- [37] Garcia R., E. Renault, and G. Tsafack (2007). Proper Conditioning for Coherent VaR in Portfolio Management. *Management Science*, forthcoming.
- [38] Glosten L.R. and R. Jagannathan (1994). A Contingent Claim Approach to Performance Evaluation. *Journal of Empirical Finance*, **1**, 133-160.

- [39] Hansen L., J. Heaton, and A. Yaron (1996). Finite-sample Properties of Some Alternative GMM Estimators. *Journal of Business and Economic Statistics*, **14**, 262-280.
- [40] Hansen L.P., and R. Jagannathan (1991). Implications of Security Market Data for Models of Dynamic Economies. *Journal of Political Economy*, **99**, 2, 225-262.
- [41] Hansen L.P., and R. Jagannathan (1997). Assessing Specification Errors in Stochastic Discount Factor Models. *Journal of Finance*, **52**, 2, 557-589.
- [42] Hasanhodzic J. and A.W. Lo (2007) Can Hedge-Fund Returns be Replicated?: The Linear Case. *Journal of Investment Management*, **5**, 2, 5-45.
- [43] Imbens G. (1997). One-Step Estimators for Over-Identified Generalized Method of Moments Models. *Review of Economic Studies*, **64**, 359-383.
- [44] Imbens G., R.H. Spady, and P. Johnson (1998). Information Theoretic Approaches to Inference in Moment Condition Models. *Econometrica*, **66**, 2, 333-357.
- [45] Jensen M. (1968). The Performance of Mutual Funds in the Period 1945-1964. *Journal of Finance*, **23**, 2, 389-416.
- [46] Kan R., and G. Zhou (2006). A New Variance Bound on the Stochastic Discount Factor. *Journal of Business*, **79**, 2, 941-961.
- [47] Kazemi H., and T. Scheneeweis (2003). Conditional Approach to Performance Evaluation of Hedge Funds. *Working Paper*, University of Massachusetts.
- [48] Kitamura Y. (2001). Asymptotic Optimality of Empirical Likelihood for Testing Moment Restrictions. *Econometrica*, **69**, 1661-1672.
- [49] Kitamura Y. (2006). Empirical Likelihood Methods in Econometrics: Theory and Practice. Working Paper, Yale University.
- [50] Kitamura and Otsu (2005). Minimax Estimation and Testing for Moment Condition Models Via Large Deviations. *Preliminary Draft*, Yale University.
- [51] Kitamura Y. and M. Stutzer (1997). An Information-theoretic Alternative to Generalized Method of Moment Estimation. *Econometrica*, **65**, 861-874.
- [52] Kitamura Y. and M. Stutzer (2002). Connections Between Entropic and Linear Projections in Asset Pricing Estimation. *Journal of Econometrics*, **107**, 159-174.
- [53] Lhabitant F-S. (2004). *Hedge Funds: Quantitative Insights.*, John Wiley & Sons.

- [54] Mayers D., and E. Rice (1979). Measuring Portfolio Performance and the Empirical Content of Asset Pricing Models. *Journal of Financial Economics*, **7**, 3-28.
- [55] McCulloch R., and P.E. Rossi (1990). Posterior, Predictive, and Utility-Based Approaches to Testing the Arbitrage Pricing Theory. *Journal of Financial Economics*, **28**, 7-38.
- [56] Mitchell M. and T. Pulvino (2001). Characteristics of Risk and Return in Risk Arbitrage. *Journal of Finance*, **56**, 6, 2135-2175.
- [57] Newey W. and R. Smith (2004). Higher Order Properties of GMM and Generalized Empirical Likelihood Estimators. *Econometrica*, **72**, 219-255.
- [58] Nevo A. (2002). Sample Selection and Information-theoretic Alternatives to GMM. *Journal Econometrics*, **107**, 149-157.
- [59] Owen A. B. (1988). Empirical Likelihood Ratio Confidence Intervals for A Single Functional. *Biometrika*, **75**, 2, 237-249.
- [60] Roll R. (1978). Ambiguity When Performance is Measured by the Securities Market Line. *Journal of Finance*, **33**, 4, 1051-1069.
- [61] Schennach S. (2007). Point Estimation with Exponentially Tilted Empirical likelihood. *Annals of Statistics*, **35**, 2, 634-672.
- [62] Sharpe W.F. (1966). Mutual Fund Performance. *Journal of Business*, **39**, 1, 119-138.
- [63] Singleton K. (2006). *Empirical Dynamic Asset Pricing.*, Princeton University Press.
- [64] Smith R.J. (1997). Alternative Semi-parametric Likelihood Approaches to Generalized Method of Moments Estimation. *Economic Journal*, **107**, 503-510.
- [65] Smith R.J. (2004). GEL Criteria for Moment Condition Models. *Working Paper*, University of Warwick.
- [66] Snow K. (1991). Diagnosing Asset Pricing Models Using the Distribution of Asset Returns. *Journal of Finance*, **46**, 955-983.
- [67] Stutzer M. (1995). A Bayesian Approach to Diagnosis of Asset Pricing Models. *Journal of Econometrics*, **68**, 367-397.
- [68] Treynor J.L. (1965). How to Rate Management of Investment Funds. *Harvard Business Review*, **43**, 1, 63-75.

- [69] Wang Z. and X. Zhang (2005). Empirical Evaluation of Asset Pricing Models: Arbitrage and Pricing Errors over Contingent Claims. Working paper, Cornell University.

## 6 Appendix A - Taylor Expansion of the HARA Utility Function Implied by the Cressie Read Estimators

According to Eq. (9) in Section 2, the utility function that is maximized to obtain the solution of the Cressie Read Bounds and their implied SDFs is given by:

$$u(v) = -\frac{1}{\gamma + 1}(1 + \gamma v)^{\frac{\gamma+1}{\gamma}} \quad (20)$$

where  $v = \lambda * (R - \frac{1}{a})$ , and  $a$  represents the SDF mean.

Now we are interested in performing a Taylor expansion around  $E[(R - \frac{1}{c})]$  to analyze the effects of the coefficient of risk aversion  $\gamma$  on the importance of skewness and kurtosis. To that end, we use the corresponding third and fourth derivatives of  $u$  that are respectively given by:

$$u_3(v) = -(1 - \gamma)(1 + \gamma v)^{-2 + \frac{1}{\gamma}} \quad (21)$$

$$u_4(v) = -(1 - \gamma)(1 - 2\gamma)(1 + \gamma v)^{-3 + \frac{1}{\gamma}} \quad (22)$$

Figure 4 plots these two derivatives as functions of  $\gamma$  to give information on how each CR estimator weights higher moments of returns. Functions are depicted for both positive and negative values of  $v$ , since in practice,  $v$  it is a random variable which directly depends on risk factors' excess returns with respect to the SDF mean. First, note that skewness weights are an increasing function with  $\gamma$ . This indicates, for instance, that while EL ( $\gamma = -1$ ) and ET ( $\gamma = 0$ ) give negative weights to skewness, CUE (CR( $\gamma = 1$ )) gives zero weight, and CR( $\gamma = 3$ ) gives positive weights. Regarding kurtosis, it is an increasing function of  $\gamma$  up to  $\gamma = 1$  and then it becomes decreasing with  $\gamma$ . That means that all CR estimators give negative weights to kurtosis, except for CUE ( $\gamma = 1$ ), which is the one that solves a quadratic utility optimization problem, thus giving zero weight to kurtosis.

## 7 Appendix B - Definitions of TASS Hedge Fund Categories

These definitions are based on the book by Lhabitant (2004).

### **C1 Convertible arbitrage**

A typical strategy in this category is to be long in the convertible bond and short in the common stock of the same company. Profits are generated from both positions. The principal is usually protected from market fluctuations.

### **C2 Fixed-income arbitrage**

The goal is to exploit price anomalies between related interest rate securities, such as interest rate swaps, U.S. and non-U.S. government bonds, and mortgage-backed securities.

### **C3 Event driven**

This strategy aims at making profits by using price movements related to special pending events such as mergers, liquidations, bankruptcies, or reorganizations. In risk arbitrage, the hedge fund manager usually invests long in the stock of the company being acquired and short in the stock of the acquiring company.

### **C4 Equity market neutral**

This investment strategy aims at balancing long and short positions to ensure a negligible market exposure in a broad sense. A fund may be neutral to a specific exchange rate, a stock index, a series of interest rates, or other factors.

### **C5 Long-short equity**

Long/short strategies involve the combined purchase and sale of two securities. The main source of return comes from the spread in performance between the stocks on the long side (which should appreciate in value) and the shorted stocks (which should decrease in value). The strategies can be based on value, growth, or size.

### **C6 Global macro**

Global macro funds do not hedge anything. They make directional bets based on their forecasts of market directions according to economic trends or particular events. They are not specialized, and carry long and short positions in any of the major world capital or derivative markets. The portfolios include stocks, bonds, currencies, and commodities. Most funds invest globally in both developed and emerging markets.

### **C7 Emerging markets**

These funds take positions in all types of securities in emerging markets around the world. Investments in emerging market equities are primarily long, since many emerging markets do not allow short selling and no viable futures markets exist to hedge market risk.

### **C8 Dedicated short bias**

Dedicated short hedge funds seek to profit from a decline in the value of stocks by taking short positions. These funds are rare nowadays, since they migrated to the long/short category, where they still have a systematic short bias.

### **C9 Managed futures**

These funds, often referred to as commodity trading advisers (CTAs), invest in financial and commodity futures markets and currency markets around the world. A large proportion are trend followers (buy in an up market and sell in a down market). Others use discretionary (judgmental) or systematic (based on technical information) strategies.

<b>IT Estimators</b>	<b>Bond</b>	<b>Credit</b>	<b>Commodity</b>	<b>S&amp;P</b>	<b>Currency</b>
Exponential Tilting	39.8 (37.8%)	56.2 (53.5%)	-0.1 (0.1%)	-5.7 (5.4%)	-3.4(3.2%)
Empirical Likelihood	35.8(43.7%)	34.6 (42.3%)	0.7(0.9%)	-3.7(4.5%)	-7.1(8.6%)
Cressie Read ( $\gamma = 1$ )	46.7(29.9%)	88.8(56.8%)	-2.5(1.6%)	-7.5(4.8%)	10.8 (6.9%)
Cressie Read ( $\gamma = 3$ )	15.8(28.0%)	30.3(53.9%)	-1.7(3.0%)	-3.7(6.6%)	4.8(8.5%)

Table 1: Optimal Portfolio Weights for IT Estimators (SDF Mean = 0.9962).

<b>Estim.</b>	<b>Bond</b>	<b>Credit</b>	<b>Commod.</b>	<b>S&amp;P</b>	<b>Curr.</b>	<b>ATM Call</b>	<b>OTM Call</b>	<b>ATM Put</b>	<b>OTM Put</b>
ET	36.7 <sub>(29.4%)</sub>	65.6 <sub>(52.6%)</sub>	-0.2 <sub>(0.1%)</sub>	-12.8 <sub>(10.2%)</sub>	-3.2 <sub>(2.5%)</sub>	-0.4 <sub>(0.3%)</sub>	0.9 <sub>(0.7%)</sub>	-2.6 <sub>(2.1%)</sub>	2.6 <sub>(2.1%)</sub>
EL	32.7 <sub>(33.4%)</sub>	45.7 <sub>(46.7%)</sub>	1.0 <sub>(1.0%)</sub>	-8.9 <sub>(9.1%)</sub>	-3.6 <sub>(3.7%)</sub>	-0.9 <sub>(0.9%)</sub>	1.2 <sub>(1.3%)</sub>	-1.9 <sub>(2.0%)</sub>	1.8 <sub>(1.9%)</sub>
CR (1)	37.6 <sub>(27.2%)</sub>	82.6 <sub>(59.8%)</sub>	-1.2 <sub>(0.9%)</sub>	-12.3 <sub>(8.9%)</sub>	3.1 <sub>(2.2%)</sub>	-0.1 <sub>(0.1%)</sub>	0.9 <sub>(0.6%)</sub>	0.3 <sub>(0.2%)</sub>	0.1 <sub>(0.1%)</sub>
CR (3)	13.4 <sub>(27.1%)</sub>	31.4 <sub>(63.3%)</sub>	-0.5 <sub>(1.0%)</sub>	-2.0 <sub>(4.0%)</sub>	-0.2 <sub>(0.3%)</sub>	-0.1 <sub>(0.1%)</sub>	0.4 <sub>(0.8%)</sub>	1.0 <sub>(2.0%)</sub>	-0.7 <sub>(1.4%)</sub>

Table 2: Optimal Portfolio Weights for IT Estimators when Option Factors are Included (SDF Mean = 0.9962).

<b>IT Estimators</b>	<b>C1</b>	<b>C2</b>	<b>C3</b>	<b>C4</b>	<b>C5</b>	<b>C6</b>	<b>C7</b>	<b>C8</b>	<b>C9</b>
Exponential Tilting	4.29	1.00	2.02	-2.11	0.14	-0.19	2.29	-0.48	-4.26
ET with Options	4.06	1.71	1.78	-2.15	-1.05	-0.28	0.38	0.63	-4.30
Empirical Likelihood	3.77	0.84	1.75	-2.19	0.66	0.23	2.22	-0.48	-2.71
EL with Options	3.93	2.00	2.46	-2.22	-0.28	-0.45	2.05	-1.09	-4.28
Hansen and Jagannathan	4.55	2.04	2.43	-1.52	-0.10	0.03	2.46	0.08	-5.39
HJ with Options	3.92	2.66	1.19	-1.58	-1.62	0.06	-0.71	2.58	-4.36
Cressie Read ( $\gamma = 1$ )	4.39	0.92	2.39	-1.09	1.54	-1.09	3.31	-4.16	-5.79
CR ( $\gamma = 1$ ) with Options	3.35	1.10	1.06	-1.44	-0.80	-1.15	-0.41	-1.01	-4.71
Cressie Read ( $\gamma = 3$ )	3.60	0.27	1.62	-0.54	0.86	-3.10	0.45	-2.70	-6.71
CR ( $\gamma = 3$ ) with Options	3.58	0.10	1.90	-0.74	2.15	-2.66	2.55	-5.82	-6.70
<b>Previous Estimators</b>									
Linear Factor Model	3.72	0.35	1.66	0.02	0.49	-1.51	-0.50	1.10	-4.15
Diez and Garcia $\sigma = 5\%$	5.72	2.09	4.71	2.25	2.84	1.29	8.18	-0.82	-8.10
Diez and Garcia $\sigma = 15\%$	4.78	0.57	2.24	0.51	0.93	-1.32	1.90	0.79	-6.04
Diez and Garcia $\sigma = 25\%$	2.44	-2.67	-1.45	-1.40	-1.91	-7.63	-4.78	3.71	-1.73
Agarwal and Naik	3.58	0.40	1.59	0.22	0.98	-1.58	-1.03	1.30	-3.96

Table 3: Hedge Funds Performance Evaluation under Different Estimators.

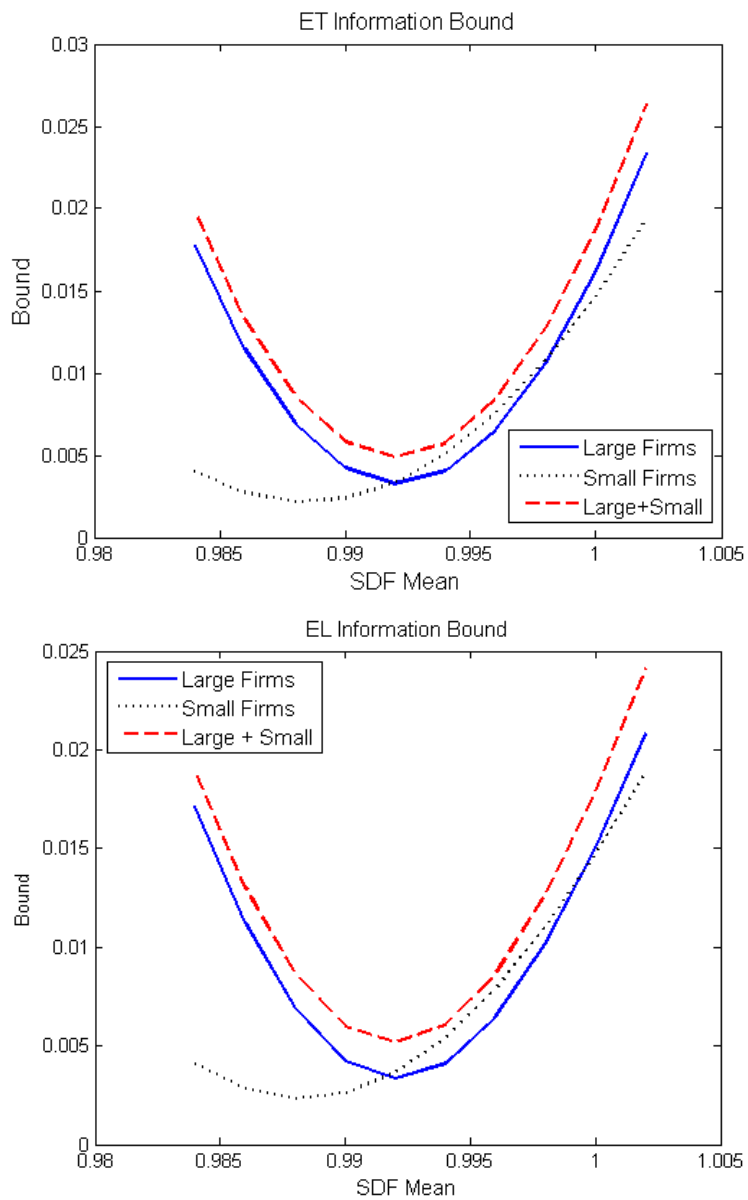


Figure 1: ET and EL Information Bounds for Size Portfolios

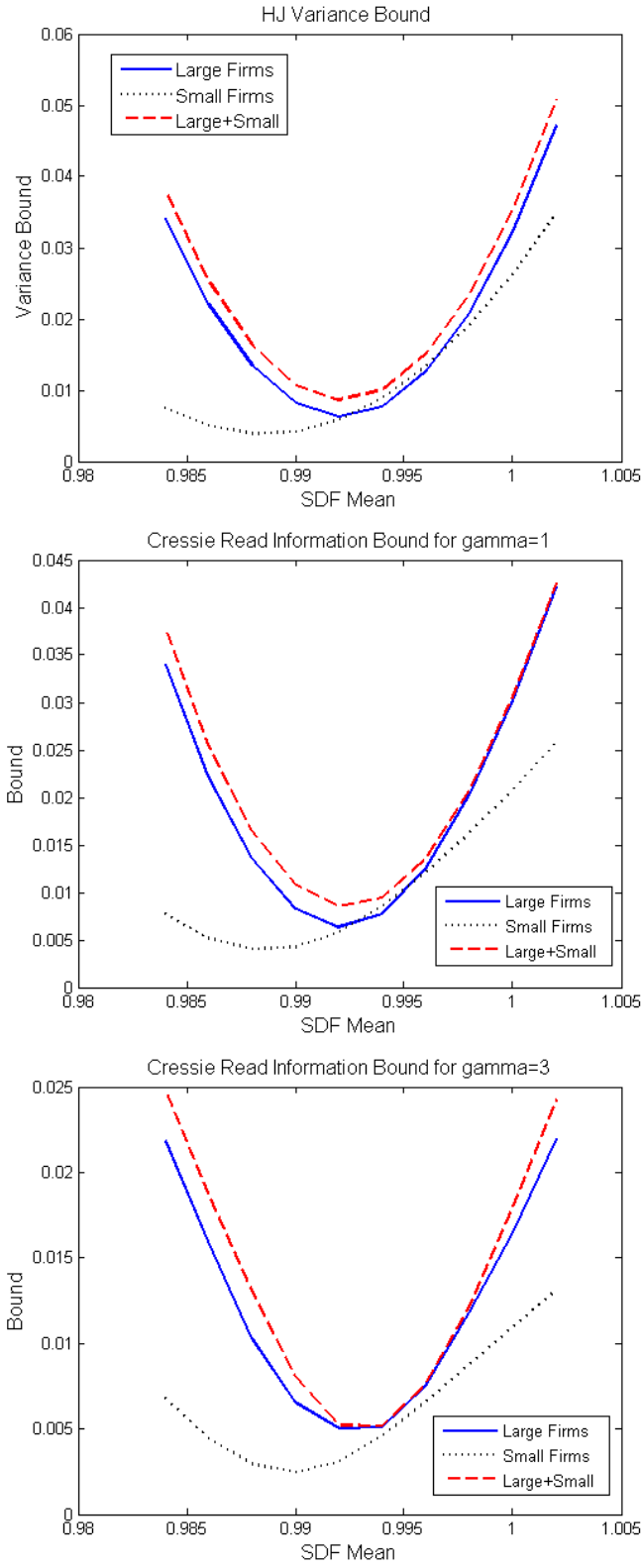


Figure 2: HJ and CR Information Bounds for Size Portfolios

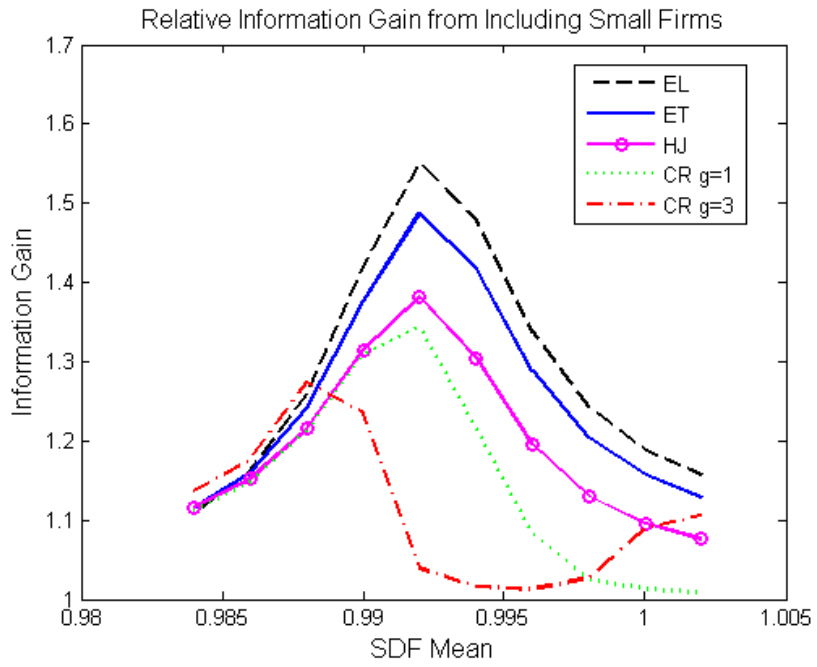


Figure 3: Information Gain from Including Small Firms Portfolios

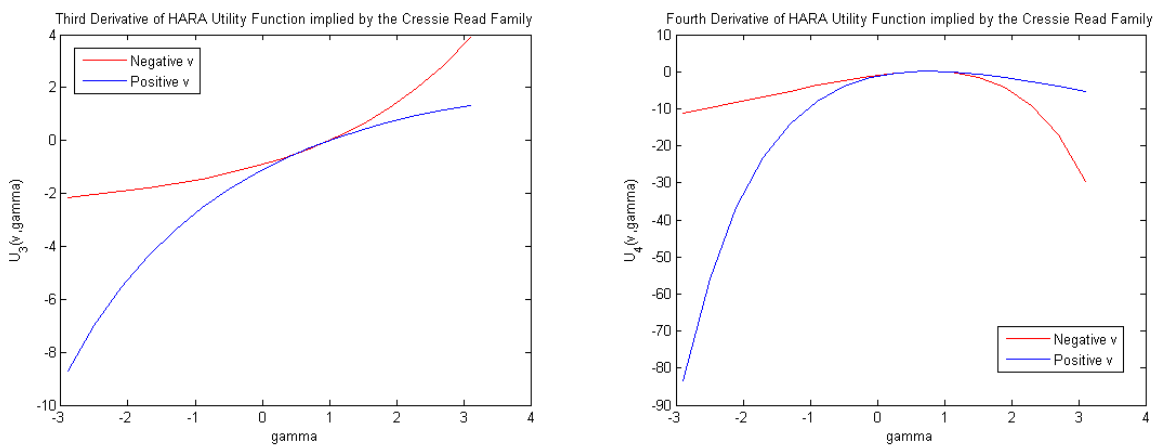


Figure 4: Skewness and Kurtosis Weights on Cressie Read Estimators

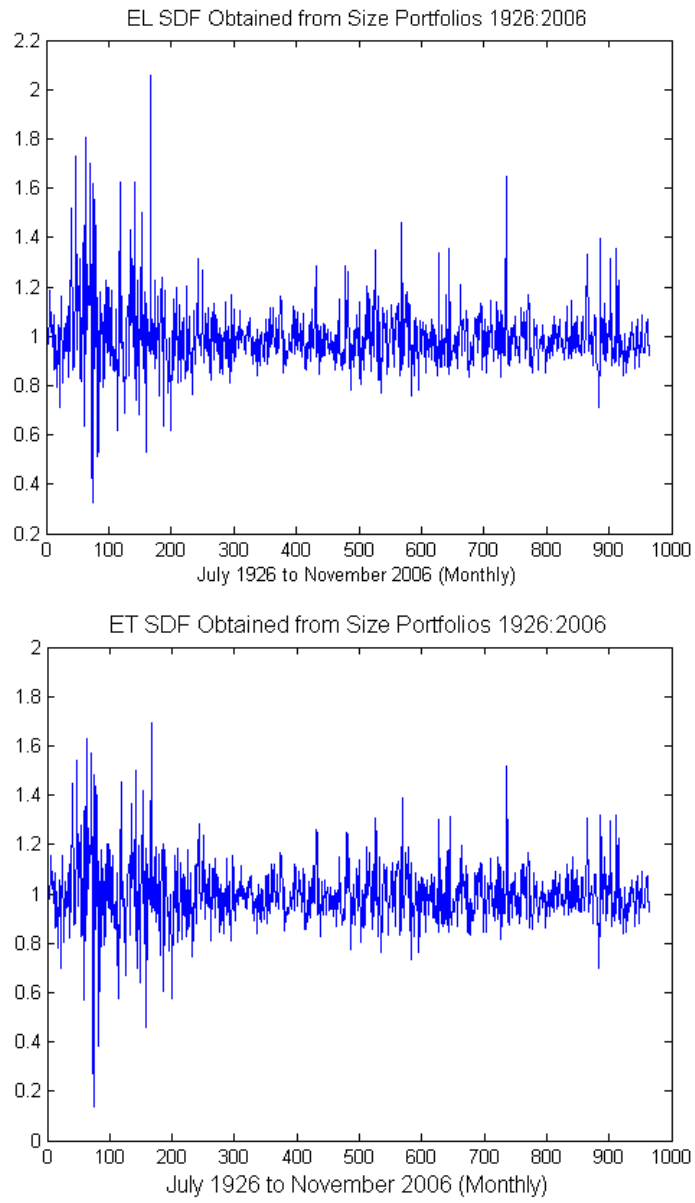


Figure 5: EL and ET Stochastic Discount Factors for Size Portfolios

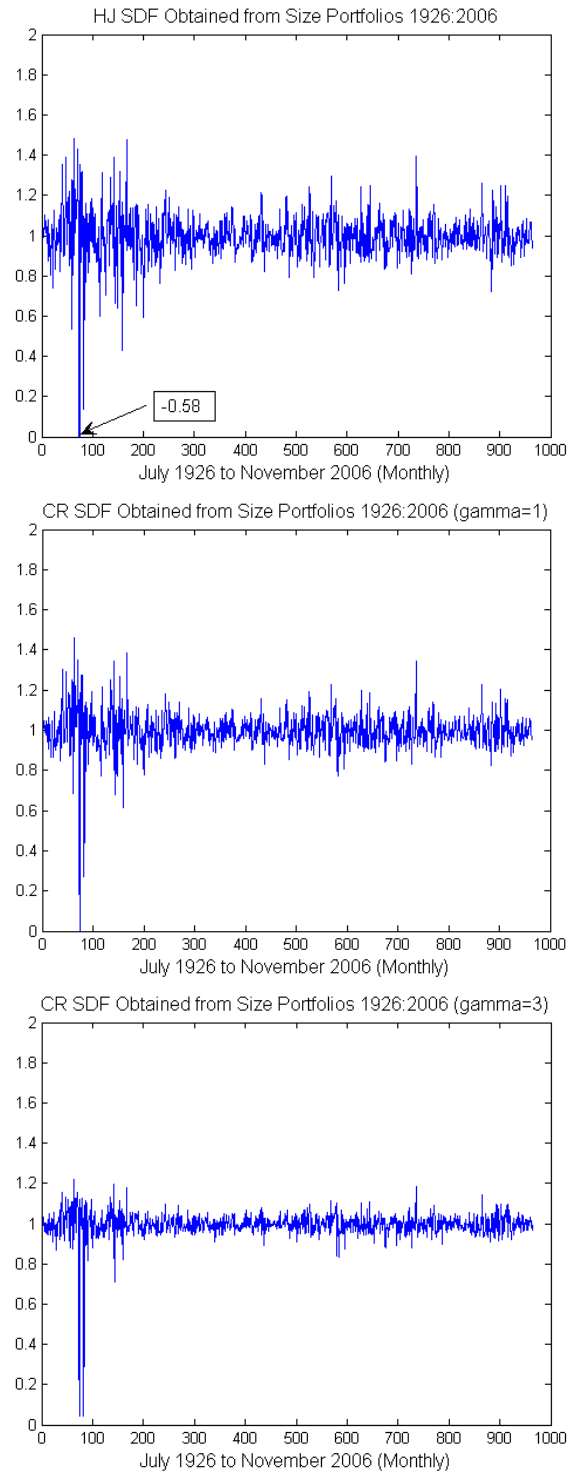


Figure 6: HJ and CR Stochastic Discount Factors for Size Portfolios

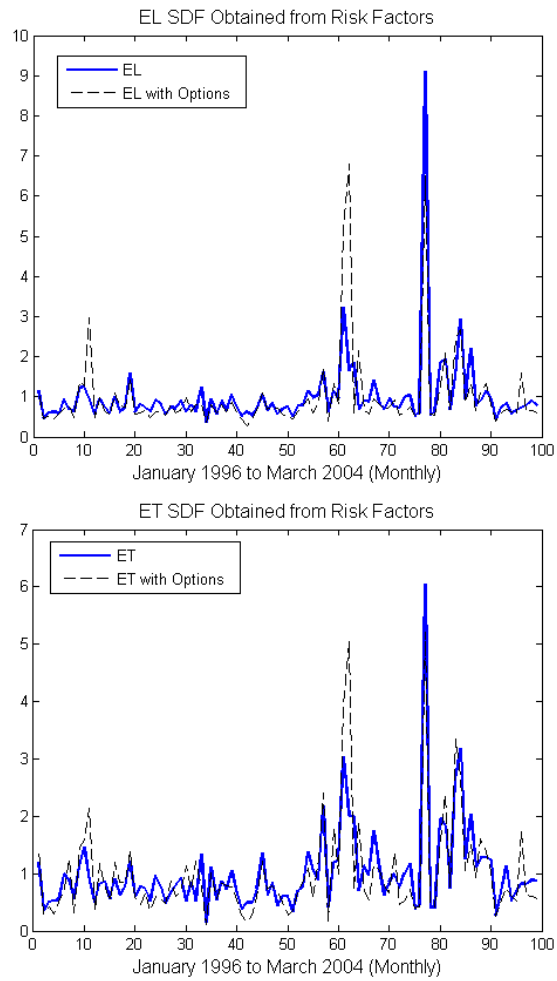


Figure 7: EL and ET Stochastic Discount Factors Extracted from Market Risk Factors

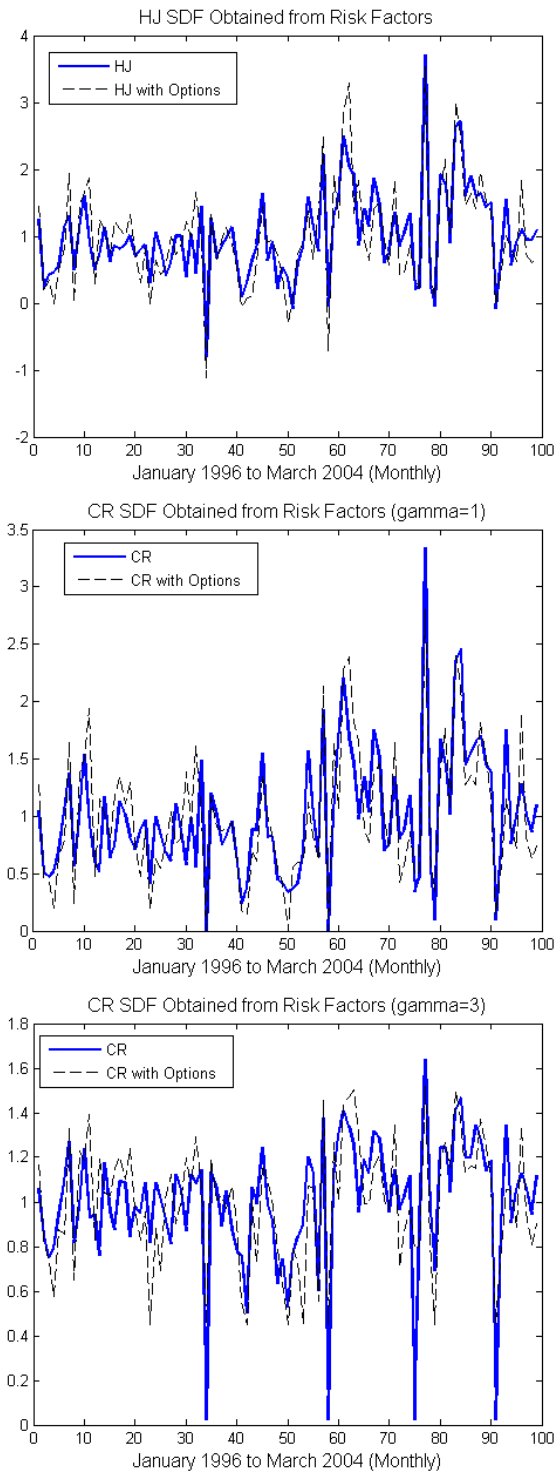


Figure 8: HJ and CR Stochastic Discount Factors Extracted from Market Risk Factors

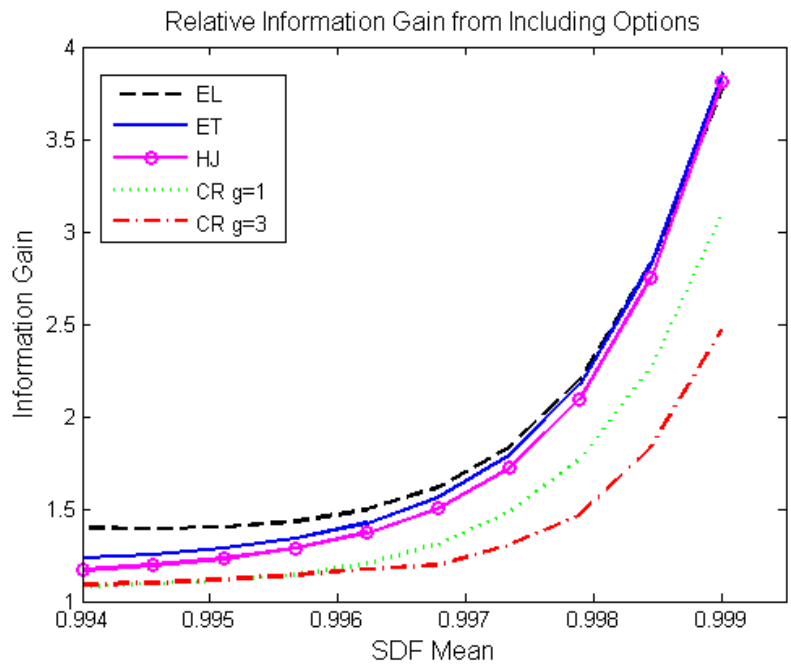


Figure 9: Information Gain from Including Options

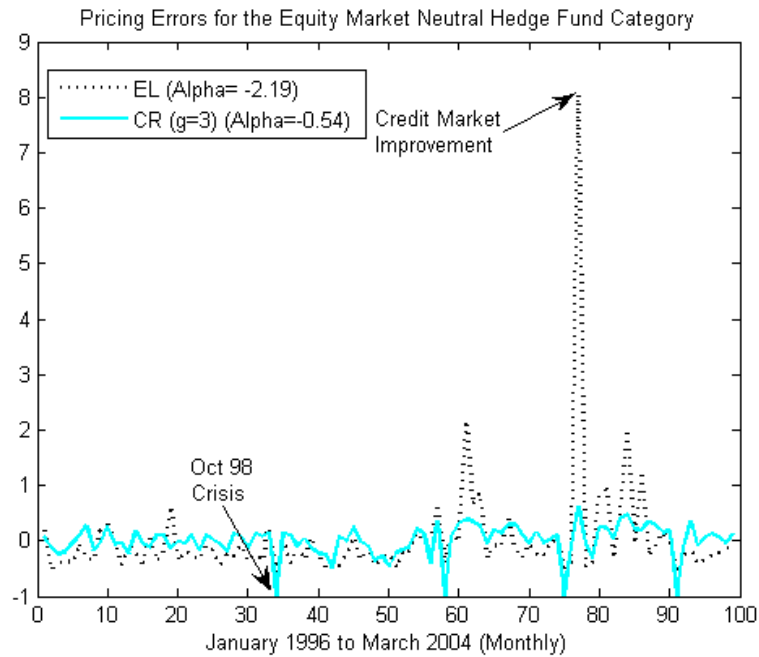


Figure 10: Comparisons Across Estimators: Time Series of Equity Market Neutral Pricing Errors