Assessing Misspecified Asset Pricing Models with Empirical Likelihood Estimators

Caio Almeida *
Graduate School of Economics
Getulio Vargas Foundation

René Garcia †
Edhec Business School

February 13, 2009

Abstract

Hansen and Jagannathan (1997) compare misspecified asset pricing models based on least-square projections on a family of admissible stochastic discount factors. We extend their fundamental contribution by considering Minimum Discrepancy (MD) projections where misspecification is measured by convex functions that can explicitly take into account higher moments of asset returns. The MD problems are solved on dual spaces with the interpretation of optimal portfolio problems based on HARA utility functions, producing a family of estimators that captures the least-square problem as a particular case. We use our proposed methodology to test the Consumption Asset Pricing Model and illustrate, under several different discrepancy functions and regions of the parametric space, the relation between the parametric proxy model, and the closest admissible SDF. On the estimation problem, not surprisingly, all MD estimators clearly reject the CCAPM model. However, some of these estimators lead to admissible SDFs that are very distinct from the one implied by the least-square solution. By their pricing implications, this rich set of optimal MD SDFs represent useful tools to diagnose missing factors in asset pricing models.

Keywords: Stochastic Discount Factor, Euler Equations, Generalized Minimum Contrast Estimators, Model Selection, Cressie Read Discrepancies.

JEL Classification: C1,C5,G1

1 Introduction

Asset pricing models provide approximations of reality that are useful to explain economic stylized facts. Hansen and Jagannathan (1997), hereafter HJ, suggested that an appropriate way to compare their performance was to evaluate functions of their implied pricing errors on corresponding Euler equations. They proposed a useful test for comparisons of possibly misspecified asset pricing models, based on a least-square projection of a proxy on a family of admissible stochastic discount factors. This test has been used in a number of empirical papers as a tool for model diagnostics as well as model selection (see for instance, Jagannathan and Wang (1996), Hodrick and Zhang (2001), Wang and Zhang (2005), Chen and Ludvigson (2008), Chen et al. (2008), and Kan and Robotti (2008) among many others).

The idea of adopting least-square theory to determine admissible discount factors that are close to asset pricing proxies is intuitive. First, it provides an easy interpretation of the degree of misspecification of a model as a maximum pricing error measure in the space of payoffs. It is also easy to implement by using duality theory in convex optimization problems (see Luemberger (1969)). Nonetheless, the quadratic metric has some important drawbacks too. It implies that misspecification is provided by a quadratic form on the pricing errors of the primitive securities that fails to take into account moments of the payoffs (returns) distributions other than mean and variance.

There is a large body of research indicating the importance of considering skewness and kurtosis when pricing assets[1]. There are also several studies showing that the second moment of certain types of asset returns might not be finite, and advocating the use of stable Paretian distributions to model returns[2]. By either suggesting the inclusion of more moments or by suggesting the adoption of metrics that take into account the inexistence of some moments, these studies point to the use of alternative metrics that go beyond the first two moments of distributions. Interestingly, in econometrics, there is also a considerable literature proposing increasingly more sophisticated Empirical Likelihood-type estimators that are robust against distributional assumptions and that possess good properties analogous to those of parametric likelihood procedures (see Kitamura (2001, 2006))[3].

---

[1]See for instance Kraus and Litzenberger (1976), Rubinstein (1973), Baroni-Adesi (1985), Harvey and Siddique (2000), Dittmar (2002), and Vanden (2006), among many others.

[2]Mandelbrot (1963), and Fama (1965) suggested the use of stable Paretian distributions to capture leptokurtic returns of certain assets. Madan et al. (1998) and Carr and Wu (2003) adopt stable processes for underlying assets in option pricing models. Tokat et al. (2003) analyze the effect of using stable processes in asset allocation problems. Beaulieu and Dufour (2005) propose linear asset pricing tests when disturbances follow stable asymmetric distributions. Ibragimov (2004) verifies the validity of several economic models under heavy-tailed distributions.

[3]For instance, Owen (1988) proposed the Empirical Likelihood estimator, and Kitamura and Stutzer

In particular, Stutzer (1995) and Kitamura and Stutzer (2002) have suggested the use of relative entropy to develop a research program that parallels that of HJ (1991, 1997).

Based on the motivation provided above, we propose alternative methods to measure the degree of misspecification of asset pricing models that make use of the theory of Minimum discrepancy (MD) estimators (Corcoran (1998)). The idea is to consider general convex functions ϕ to calculate the distance between a certain asset pricing proxy y and the family M of admissible stochastic discount factors (SDFs) that prices a set of underlying primitive securities x. We formulate this problem within a Minimum discrepancy framework where the goal is to obtain a SDF m^*, or a probability measure p^*, that is admissible (i.e., satisfies the moment conditions by pricing primitive securities) and that is the closest possible[4] to the asset pricing proxy, by minimizing either $\phi(1 + m - y)$ (Additive case) or $\phi(|\frac{m}{y}|)$ (Multiplicative case). We make use of duality theory (see Kitamura (2006), and Borwein and Lewis (1991)) to estimate this Minimum discrepancy probability measure and its distance to the proxy y, by solving finite-dimensional problems with interpretations of optimal portfolio problems.

When the Minimum discrepancy problems are specialized to the class of Cressie Read (1984) discrepancies, we show that, under our formulation, the dual optimization problems reduce to a class of Generalized Empirical Likelihood (GEL) estimators (Smith (1997)) where the proxy model y appears only in the discrepancy function ϕ and not in the moment conditions as usual. This formulation makes clear that we are interested, like HJ (1997), in measuring the degree of misspecification of a model y with respect to a family of admissible SDFs M that is invariant to changes in the model vector of parameters θ. Moreover, this family will be also invariant with respect to changes of models and should depend only on the primitive payoffs x and their prices q.

By looking at the first-order conditions of the dual quadratic problem in HJ (1997), a nice interpretation for their least-square solution is obtained. They showed that the admissible SDF that is closest to the proxy y is given by y subtracted by a linear combination of primitive asset payoffs $(\lambda'_{HJ} \cdot x)$ that is the smallest linear correction (in the least squares sense) for y to become admissible. It happens that in our MD problems we have similar interpretations of solutions as proxy corrections to become admissible SDFs: The solutions to our Additive MD problems give additive correction terms to the proxy y that are nonlinear

---

(1997) the Exponential Tilting Estimator (see also Imbens et al. (1998)). Smith (1997) proposed a large class of Generalized Empirical Likelihood estimators, later shown to be equivalent to the subset of Minimum Discrepancy estimators with Cressie Read discrepancies (Newey and Smith (2004)). Recently, Smith (2007) developed tests for conditional moment restrictions models based on a kernel-weighted version of the CressieRead power divergence family of discrepancies.

[4]in the spirit of Csiszar's I-divergencies (1975, 1991).

2

functions of the optimal linear combination of primitive asset's payoffs $(\lambda'_{MD,ad} \cdot x)$, which are the smallest correction (in the ϕ divergence additive sense) for y to become an admissible SDF. On the other hand, the solutions to our Multiplicative MD problems provide multiplicative correction terms that are nonlinear on both the proxy y and primitive assets' payoffs $(\lambda'_{MD,mu} \cdot x)$, which are the smallest correction (in the ϕ divergence multiplicative sense) for y to become an admissible SDF.

Another important issue relates to model choice and parameter estimation. HJ (1997) suggest, as a first step, the estimation of model parameters by minimizing the HJ distance between the model and the family of admissible SDFs. Then, assuming the existence of a set of model candidates whose parameters have been previously estimated using the HJ distance, HJ suggest selecting the model with smallest distance[5]. Similarly in our MD problems, for any fixed discrepancy function ϕ, our theory, supported by the theory of GEL estimators (see Newey and Smith (2004); Kitamura (2006)), suggests the estimation of model parameters by minimizing the discrepancy chosen, and the subsequent use of minimum discrepancy distances to rank candidate asset pricing models[6].

Our empirical application consists in carefully analyzing the Consumption CAPM (Breeden (1979)), by testing its ability to price a set of primitive securities (bond and S&P 500) at different regions of the parametric space. The discrepancy between asset pricing proxies and admissible SDFs is measured by different functions belonging to the Cressie Read (1984) family, namely: The Pearson's Chi-Square, EL (Owen (1984)), Hellinger's distance, ET (Kitamura and Stutzer (1997)), Euclidean Likelihood or CUE (Hansen et al. (1996)) and two other discrepancies with high positive values of the Cressie Read parameter γ ($CR(\gamma = 2)$, and $CR(\gamma = 5)$). Based on a grid for the risk aversion parameter of the CCAPM model, we show that most of the discrepancies agree on the choice of θ but once we increase too much the Cressie Read parameter γ, the MD problems stop satisfying the moment conditions (introducing pricing errors), and end up optimally choosing different θ's. We discuss the empirical findings across Cressie Read discrepancies relating them to

---

[5]Hodrick and Zhang (2001) use the HJ (1997) distance to compare ten different asset pricing models based on the 25 Fama and French (1997) test assets. Recently, Kan and Robotti (2008) suggest a more formal selection by developing a test to compare the HJ distance between models, and obtain the asymptotic distribution of this test, for any combination of correclty specified, misspecified, and /or nested /non-nested model candidates.

[6]Formal statistical comparison tests are not developed in this paper. In that matter, we refer the reader to a recent literature developing formal tests for model choice with empirical likelihood estimators (see Kitamura (2000), Ramalho and Smith (2002), Hong et al. (2003), Kitamura (2006a), and Chen et al. (2007), among others). In particular, Kitamura (2000) proposes nonparametric likelihood ratio tests based on the Exponential Tilting estimator to compare (possibly misspecified) moment-based econometric models. Hong et al. (2003) extend Kitamura's model selection tests to the whole family of GEL estimators. The framework in these two papers can be adapted to our above-mentioned GEL problems to derive formal statistical tests.

the size of the pricing errors, and to the magnitude of the admissible SDFs.

After analyzing the CCAPM model based on calibrations of the risk aversion parameter, we perform estimation of the model using both our additive and multiplicative formulations of the MD problems. For all adopted discrepancy functions and under the two formulations the estimated risk aversion parameters are high. Additive estimators obtain values around 37, while multiplicative estimators obtain values around 64. For each CR discrepancy and risk-aversion coefficient estimated, we also obtain the closest admissible SDFs to the CCAPM model. We show that these SDFs vary across Cressie read discrepancies, and specially across the type of estimation problem (additive and multiplicative). Multiplicative estimators present a better goodness of fit with admissible SDFs very close to their CCAPM counterparts. In addition, multiplicative estimators with high gamma values in general achieve pricing errors smaller than their additive counterparts.

The rest of the paper is organized as follows. Section 2 introduces the market structure, defines admissible SDFs, and presents the original Hansen and Jagannathan framework for model estimation and selection. Section 3 formulates our generalization to the HJ methodology that considers minimum discrepancy optimization problems. It presents the main theorem that provides a family of metrics that contains HJ (1997) as a particular case. It also defines implied probabilities, presents their relation to admissible stochastic discount factors, and provides some particular model selection procedures based on known discrepancy functions belonging to the Cressie Read family. Section 4 is empirical. It explains the adopted asset pricing model, the consumption CAPM, introduces the dataset adopted and provides estimation results. Section 5 presents a discussion on the results, analyzing the relation between implied admissible SDFs, pricing errors, and discrepancies adopted. Section 6 concludes.

2 Stochastic Discount Factors and Asset Pricing Proxies

Following the lead of Harrison and Kreps (1979), Chamberlein and Rothschild (1983), and Hansen and Jagannathan (1997) we model portfolio payoffs as elements of a Hilbert space, and use a continuous linear functional on that space to represent prices assigned to those payoffs.

We assume that assets are purchased at a certain time t and that the payoffs are received at a time $T > t$. Let Γ_T represent the sigma-algebra that represents the conditioning information at date T, and L^2 denote the space of all square integrable (i.e., finite second moments) random variables that are measurable with respect to Γ_T. Assume there exists a set of n primitive securities whose payoffs are represented by a vector $x \in \Re^n$, with $x \in L^2$

and in addition having a nonsingular second moment matrix Exx'[7]. A payoff p will be any square integrable variable which is obtained as a linear combination of the payoffs of the n primitive securities:

$$P \equiv \{x \cdot c : c \in \Re^n\} \tag{1}$$

We further assume that the payoffs in P satisfy the Law of One Price and that the pricing functional π is continuous and linear on P.

An admissible SDF will be any square-integrable random variable m that correctly prices all asset payoffs $p \in P$

$$\pi(p) = E(mp) \tag{2}$$

An asset pricing model y will be an approximation for an admissible SDF, and will possibly price some payoffs in P with error:

$$\pi_y(p) = E(yp) \tag{3}$$

where the error is measured by the difference $\pi(p) - \pi_y(p)$.

2.1 Hansen and Jagannathan's (1997) Least-Squares Approximation of Proxies

Given a a proxy asset pricing model $y(\theta)$, parameterized by a vector of parameters $\theta \in \Re^k$, HJ (1997) suggest to measure its degree of misspecification by obtaining the least-squares projection of this proxy into the space of admissible SDFs M:

$$\delta_{HJ}(\theta)^2 = \min_{m \in M} \|m - y(\theta)\|^2 = \min_{m \in M} E\{(m - y(\theta))^2\} \tag{4}$$

This problem can be rewritten by noticing that $m \in M$ can be reexpressed as $m \in L^2$ satisfying the moment condition (2) for the particular set of primitive securities:

$$\delta_{HJ}(\theta)^2 = \min_{m \in L^2} E\{(m - y(\theta))^2\} \text{ subject to } E(mx) = \pi(x) = q \tag{5}$$

Making use of Lagrange multipliers the problem becomes:

$$\delta_{HJ}(\theta)^2 = \min_{m \in L^2} \sup_{\lambda \in \Re^n} E\{(m - y(\theta))^2 - 2\lambda'(mx - q)\} \tag{6}$$

---

[7]We assume the existence of the second moments to be able to work in a Hilbert space. For a treatment of the case with inexistent moments (not performed here), the payoffs should be in a Banach (L^j) space (see Royden (1988)), and, in principle, the existence of a linear pricing functional could be questioned. However, we refer to Araujo and Monteiro (1989) who provide a proof of existence of equilibrium in L^j spaces, therefore guaranteeing the existence of a linear pricing functional in such spaces. In this case, our methodology would also work and we conjecture that the interpretations provided in Section 3 would still be valid.

By fixing the Lagrange multipliers and solving the minimization on the variable m Hansen and Jagannathan (1997) obtained the following dual optimization problem:

$$\delta_{HJ}(\theta)^2 = \max_{\lambda \in \Re^n} E\{y^2 - (y - \lambda'x)^2 - 2\lambda'q\} \tag{7}$$

whose solution is given by the (square of) Hansen and Jagannathan's distance:

$$\delta_{HJ}(\theta)^2 = E\{(Exy - q)'Exx'^{-1}(Exy - q)\}^{\frac{1}{2}} \tag{8}$$

2.1.1 Interpreting the Primal and Dual Problems

The dual optimization problem (7) is nicely interpreted by HJ (1997) as an optimal portfolio problem with a quadratic utility function. The Lagrange multipliers represent the portfolio weights on the different primitive securities payoffs. Stutzer (1995) explores this portfolio interpretation in a nonparametric setting based on ET obtaining a CARA (exponential) utility function, and Almeida and Garcia (2008) generalize Stutzer's interpretation in a nonparametric setting with general Cressie Read discrepancy functions providing a portfolio interpretation with HARA (Hyperbolic Absolute Risk Aversion) utility functions. In the present paper, we will also obtain portfolio interpretations to our dual MD problems as we shall see in Section 3.4.

The first-order conditions from problem (7) also give an interesting interpretation, this time for the solution of the primal problem:

$$q = E\{(y - \lambda'_{HJ} \cdot x)x\} \tag{9}$$

Equation (9) shows that the optimal Lagrange multipliers λ_{HJ} that solve this problem find the smallest correction in the mean square sense to the proxy y such that it becomes an admissible SDF. Under what we call our additive Minimum Discrepancy problems, we will observe results that will paralell this one but where the correction to the proxy will be non-linear in the primitive payoffs x.

HJ (1997) also interpret the primal problem (4) as a maximum pricing error problem per unit norm. A linear functional $\pi_a = \pi - \pi_y$ representing the approximate pricing errors is defined. They show that δ_{HJ} is the norm of this functional, and moreover that this norm is achieved by a special payoff $\tilde{p}$ obtained with the application of the Riezs representation theorem to the functional π_a.

2.1.2 Model Estimation Based on the HJ Distance

HJ (1997) suggest estimating the parameter vector θ by minimizing the HJ distance:

$$\underset{\theta \in \Re^k}{argmin} \; \delta_{HJ}(\theta) \qquad (10)$$

as an estimator alternative to the GMM (Hansen (1982)):

$$\underset{\theta \in \Re^k}{argmin} \; g(\theta)' \cdot W \cdot g(\theta) \qquad (11)$$

where $g(\theta) = E(y(\theta).x)$ represents the moment conditions, and W is an $n \times n$ symmetric positive definite matrix that might depend on the sample observations. Note that the HJ estimator in equation (10) as well as the GMM estimator in (11) are special cases of the minimum distance estimators with a quadratic norm. In an asset pricing context, HJ (1997) showed that the main difference between these two estimators is that in general the optimal matrix W in Equation (11) (see Hansen and Singleton (1982)) will depend on the particular model proxy y adopted, while this normalizing matrix is fixed at $(Exx')^{-1}$ in the case of the HJ estimator. This apparently small distinction is in fact very important. For instance, suppose we decided to adopt the GMM criterion to select among possibly misspecified models. The metric adopted to measure misspecification in this case would be given by Equation (11) with $g(\theta)$ being the pricing errors. As the weighting matrix changes with each asset pricing model y, this GMM metric will weight pricing errors differently across models. In this case, the HJ distance should be preferable since it gives weights to the pricing errors that are invariant to the asset pricing proxy y.

3 Minimum Discrepancy Approximation of Proxies

3.1 The Additive Minimum Discrepancy Problem

Given a proxy asset pricing model y, and a convex discrepancy function ϕ, similarly to HJ (1997), the idea posed by the Minimum Discrepancy problem is to find an admissible SDF which is as close as possible to y in the ϕ discrepancy sense:

$$\delta_{MD}^{ad}(\theta) = \underset{m \in L^2}{min} \; E\{\phi(1 + m - y(\theta))\} \text{ subject to } E(mx) = q \qquad (12)$$

This problem should be of interest when the underlying primitive securities include assets with non-Gaussian returns. In such cases it is not clear that the penalty for a proxy asset pricing model y should only depend on the second moments of the pricing errors. Looking at more general discrepancies will probably be more appropriate when looking at assets

with nonlinear or asymmetric payoffs such as options, mortgages, credit derivatives, other exotic but liquid instruments, and also equities with skewed and fat tailed returns.

In what follows, we make use of arguments found in Borwein and Lewis (1991) to solve our discrepancy problem based on a simpler unconstrained optimization problem on its dual space. The corresponding dual optimization problem is given by:

$$v_{MD}^{ad}(\theta) = \max_{\lambda \in \Re^n} \lambda'q - E\{\phi^*(\lambda'x)\} \tag{13}$$

where ϕ^* denotes the convex conjugate of ϕ, calculated by the following expression:

$$\phi^*(z) = \sup_w zw - \phi(w) \tag{14}$$

Newey and Smith (2004) show that when the discrepancy function is chosen within the Cressie-Read (1984) family the dual problem belongs to the class of GEL estimators. In a recent paper, Almeida and Garcia (2008) generalize Hansen and Jagannathan (HJ, 1991) nonparametric bounds to consider higher moments of returns[8] by specializing the Cressie Read discrepancy problem to a nonparametric setting. In that context, implied probabilities[9] work as the optimal admissible SDF (optimal in the CR divergence sense). In their work, implied probabilities are hyperbolic functions of linear combinations of the primitive assets payoffs. The linear combination weights come from an optimal HARA portfolio problem that corresponds to the GEL dual problem, and involves only the primitive securities (with no parametric model).

Our results parallel those in Almeida and Garcia (2008), with the important difference that we consider the explicit existence of a parametric model y in order to generalize HJ (1997). While their optimization problems obtain nonparametric admissible SDFs, as the theorem below will show, our optimization problems will obtain admissible SDFs that combine parametric aspects (coming from y) with nonparametric aspects coming from an optimal linear combination of primitive assets' payoffs (optimal in the divergence sense). However, to be able to incorporate the parametric model and at the same time keep the moment conditions compatible with those appearing in Hansen and Jagannathan (1997), we formulate our MD problem in a slightly different form from those appearing in Newey and Smith (2004) and Almeida and Garcia (2008). First, we formulate the moment conditions by pricing the payoffs instead of excess returns. This has a direct impact on the dual

---

[8]See also Snow (1991) for a generalization of Hansen and Jagannathan (1991) bounds that similarly to Almeida and Garcia (2008), takes into account higher moments of returns.

[9]Back and Brown (1993) provide GMM implied probabilities, 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.

optimization problem. Second, to measure misspecification of the asset pricing proxy y with respect to an additive distance to an admissible SDF m (like HJ (1997) do), we need to introduce a translation of the function ϕ by $1 - y$. This translation guarantees that function ϕ achieves its minimum when the asset pricing proxy y is an admissible SDF, a very desirable property when the goal is to measure the degree of misspecification. In addition, the inclusion of y in the divergence function (and not in the moment condition explicitly) allows us to interpret this MD problem as genuinely a generalization of HJ (1997), that is a way to measure the "'distance"' (discrepancy) of the proxy y to a fixed family M of admissible SDFs.

The next theorem provides the type of optimization problems that will have to be solved to find the discrepancy of y with respect to the family M, and the corresponding admissible SDF closest to y, when the discrepancy belongs to the Cressie Read (1984) family of discrepancies, most adopted in the current econometric literature.

Theorem 1. *Let $y(\theta)$ represent the asset pricing proxy, parameterized by a vector of parameters $\theta \in \Theta$. Let the discrepancy function belong to the class of Cressie Read functions: $\phi(\pi) = \frac{\pi^{\gamma+1}-1}{\gamma(\gamma+1)}$ with $\gamma \in \Re$. In this case for a fixed vector of parameters θ, the optimization problem (12) specializes to:*

$$\delta_{CR}^{ad}(\theta) = \min_{m \in L^2} E\left\{ \frac{(1 + m - y(\theta))^{\gamma+1} - 1}{\gamma(\gamma+1)} \right\} \text{ subject to } E(mx) = q \tag{15}$$

Then the GEL problem dual to the MD problem is given by:

$$v_{CR}^{ad}(\theta) = \max_{\lambda \in \Re^n} \lambda'q - E\left\{ \frac{(\gamma\lambda' \cdot x)^{\frac{\gamma+1}{\gamma}}}{\gamma+1} + (y(\theta) - 1)\lambda' \cdot x + \frac{1}{\gamma(\gamma+1)} \right\} \tag{16}$$

and the admissible SDF which is closest to the asset pricing proxy y is given by:

$$m_{CR}^{ad}(\theta) = y(\theta) - 1 + (\gamma\lambda_{ad}' \cdot x)^{\frac{1}{\gamma}} \tag{17}$$

where λ_{ad} is the solution of the optimization problem (16).

Proof:

Let $\tilde{\phi}(m) = \phi(1 + m - y) = \frac{(1+m-y)^{\gamma+1}-1}{\gamma(\gamma+1)}$, and note that it is a convex function.

According to Borwein and Lewis (1991), if we find the convex conjugate of $\tilde{\phi}$, $\tilde{\phi}^*$, we can use it in Equation (13) to write the dual optimization problem that has the same solution as the primal MD problem in Equation (12) with a Cressie and Read divergence. To obtain the convex conjugate we apply Equation (14) to $\tilde{\phi}$. Letting $H(x) = zx - \tilde{\phi}(x)$, and differentiating to obtain its supremum in x we obtain $x^{sup} = y - 1 + (\gamma z)^{\frac{1}{\gamma}}$ and consequently

9

$\tilde{\phi}^*(z) = (y-1)z + \frac{(\gamma z)^{\frac{\gamma+1}{\gamma}}}{\gamma+1} + \frac{1}{\gamma(\gamma+1)}$. Applying $\tilde{\phi}^*$ in Equation (13) gives the optimization problem (16). The first-order conditions of this optimization problem with respect to λ are:

$$\frac{\partial v_{CR}^{ad}}{\partial \lambda} = q - E\{(y + (\gamma \lambda_{ad}' \cdot x)^{\frac{1}{\gamma}}).x\} = 0 \tag{18}$$

showing that m_{CR}^{ad} in Equation (17) is an admissible SDF that minimizes the MD problem (12) when the divergence is a member of the Cressie-Read family. ⋄

We assume that there exists a risk-free asset on the set of primitive securities paying interest rate equal to r_f. The existence of such an asset is also assumed by HJ (1997, assumption 1.2) and it is important to guarantee that our discrepancy problems are well posed in the sense that the mean of any admissible SDF will be equal to $\frac{1}{r_f}$. Of course, if in practice such an asset does not exist, we can augment the primitive securities payoff space by a synthetic risk-free asset. We provide a corollary to Theorem 1 that simplifies the dual optimization problem by taking into account the existence of this risk-free asset.

Corollary 2. *Assuming that there is a risk-free asset among the primitive securities then the dual optimization problem in Equation (16) can be simpliefied to (by also eliminating the constant term):*

$$v_{CR}^{ad,co}(\theta) = \max_{\tilde{\lambda} \in \Re^{n-1}} \tilde{\lambda}' q^{co} - E\left\{ \frac{(1 + \gamma \tilde{\lambda}' \cdot x)^{\frac{\gamma+1}{\gamma}}}{\gamma+1} + (y(\theta)-1)\tilde{\lambda}' \cdot x \right\} \tag{19}$$

where q^{co} is the vector of prices of the $n-1$ remaining primitive securities other than the risk-free asset. The corresponding admissible SDF that solves this problem is given by:

$$m_{CR}^{ad,co}(\theta) = y(\theta) - 1 + (1 + \gamma \tilde{\lambda}_{ad}' \cdot x)^{\frac{1}{\gamma}} \tag{20}$$

Proof:

To prove this corollary just observe that the risk-free asset has a constant payoff equal to 1, which allows the separation of the maximization in two parts:

$$v_{CR}^{ad,co}(\theta) = \max_{\tilde{\lambda} \in \Re^{n-1}, \alpha \in \Re} \frac{1}{r_f}.\alpha + \tilde{\lambda}' q^{co} - E\left\{ \frac{(\gamma\alpha + \gamma\tilde{\lambda}' \cdot x)^{\frac{\gamma+1}{\gamma}}}{\gamma+1} + (y(\theta)-1)(\tilde{\lambda}' \cdot x + \alpha) \right\} \tag{21}$$

By taking the derivative of equation (21) with respect to α, eliminating $\tilde{\lambda}$ and equating to zero, we obtain the concentrated value $\alpha^* = \frac{(\frac{1}{r_f} - E_\mu[y-1])^\gamma}{\gamma}$. For the particular case where the proxy model y prices the risk-free asset, α^* becomes $\frac{1}{\gamma}$ and by substituting α^* in (21) and by eliminating constant terms (not depending on $\tilde{\lambda}$) the result follows. ⋄

3.2 The Multiplicative Minimum Discrepancy Problem

As mentioned earlier in the introduction, other variations of the asset pricing model's evaluation problem could be proposed based on minimum discrepancy estimators. In particular, we can cast the problem in the traditional structure of the GMC or MD estimators:

$$\tilde{\delta}_{MD}^{mu}(\theta) = \min_{m \in L^2} E\{\phi(m)\} \text{ subject to } E(my(\theta)x) = q \tag{22}$$

where ϕ is a convex function.

As in some problems, the admissible SDFs achieve negative values in some states of the nature (see HJ (1991, 1997)), to guarantee that we will still have a MD problem based on a convex function when this happens, we could also propose a variation of the problem that composes function $\phi(.)$ with the absolute value function $|.|$:

$$\tilde{\delta}_{MD}^{mu}(\theta) = \min_{m \in L^2} E\{\phi(|m|)\} \text{ subject to } E(my(\theta)x) = q \tag{23}$$

Note that the absolute value function $|.|$ composed with ϕ maintains $\phi(|.|)$ as a convex function even if the SDF m achieves negative values. This way with Equation (22) we will be able to solve problems whose discrepancies are calculated with respect to a family of nonnegative SDFs, while with Equation (23) we will solve the corresponding problems with families of SDFs that admit negative values. In what follows, we will derive the results for the more general problem given by Equation (23) and then we will specialize to the case given by Equation (22).

The interpretation of the estimation problem posed by either (22) or (23)is in principle different from that given to the additive MD problem described in section 3.1. Here model y is represented by the moment condition $E(m.y.x) = q$ and our goal is to find among all the probability measures (or SDFs, see Section 3.3) m such that $m*y$ is an admissible SDF pricing the primitive securities with payoffs x, the one that is closest to the true unknown measure μ in the ϕ-discrepancy sense.

Apparently the main difference (in structure) between problems (12) and (23) lies in the family of probability measures that satisfy the moment conditions. While in the problem described by Eq. (12) there is a fixed family of admissible SDFs (the ones that price x), in the multiplicative problem described here the family of m's satisfying the moment condition clearly depends on y since the moment condition asks for $m * y$ to be an admissible SDF. However, this problem can be easily transformed to an equivalent problem that will present a fixed family of admissible SDFs and will be more comparable to the additive MD problem. Assuming that the proxy y is different of zero almost surely, by renaming the product $m*y$ to a new random variable $\tilde{m}$, we obtain

$$\tilde{\delta}_{MD}^{mu}(\theta) = \min_{\tilde{m} \in L^2} E\left\{\phi\left(\left|\frac{\tilde{m}}{y(\theta)}\right|\right)\right\} \text{ subject to } E(\tilde{m}x) = q \qquad (24)$$

The transformed problem has a structure similar to the additive problem in (12) and consequently similar to the original Hansen and Jagannathan (1997) analysis. It proposes estimating distances of a proxy y to a fixed family of admissible SDFs (the ones that satisfy the Euler equations).

This problem should be of interest since it also proposes a way of selecting models based on metrics that take into account higher moments of admissible SDFs other than the mean and the variance. Moreover, in contrast to problem (12), its solution involves a nonlinear term in the proxy y that might become an important element responsible for differences in the discrepancy measured in different points of the parametric space Θ, as we shall see later. Similarly to problem (12) and using the same techniques, we now prove a theorem that will give the corresponding dual optimization problem that should be equivalent to the discrepancy problem (24) as well as the admissible SDF that will solve the optimization problem, when the discrepancy function is within the Cressie Read family.

Theorem 3. *Let $y(\theta)$ represent the asset pricing proxy, parameterized by a vector of parameters $\theta \in \Theta$. Let the discrepancy function belong to the class of Cressie Read functions: $\phi(\pi) = \frac{\pi^{\gamma+1}-1}{\gamma(\gamma+1)}$ with $\gamma \in \Re$. In this case for a fixed vector of parameters θ, the optimization problem (24) specializes to:*

$$\tilde{\delta}_{CR}^{mu}(\theta) = \min_{m \in L^2} E\left\{\frac{|\frac{\tilde{m}}{y(\theta)}|^{\gamma+1}-1}{\gamma(\gamma+1)}\right\} \text{ subject to } E(\tilde{m}x) = q \qquad (25)$$

Then the dual GEL problem to the MD problem is given by:

$$\tilde{v}_{CR}^{mu}(\theta) = \max_{\lambda \in \Re^n} \lambda'q - E\left\{\frac{|\gamma y \lambda' \cdot x|^{\frac{\gamma+1}{\gamma}}}{\gamma+1} + \frac{1}{\gamma(\gamma+1)}\right\} \qquad (26)$$

and the admissible SDF which is closest to the asset pricing proxy y is given by:

$$\tilde{m}_{CR}^{mu}(\theta) = (-1)^{I_{\{\gamma y \lambda'_{mu} \cdot x < 0\}}}.y|\gamma y \lambda'_{mu} \cdot x|^{\frac{1}{\gamma}} \qquad (27)$$

where λ_{mu} is the solution of the optimization problem (26), and $I\{.\}$ represents a set indicator function.

Proof:

Let $\tilde{\phi}(m) = \phi(|\frac{\tilde{m}}{y}|) = \frac{|\frac{\tilde{m}}{y}|^{\gamma+1}-1}{\gamma(\gamma+1)}$. Then $\tilde{\phi}$ is a convex function defined in the interval $[0, \infty)$. Now, we mimick the proof of Theorem 1. First, we find the convex conjugate of $\tilde{\phi}$, $\tilde{\phi}^*$, and then use it in Equation (13) to write the dual optimization problem that

has the same solution as the primal MD problem in Equation (25). To obtain the convex conjugate we apply Equation (14) to $\tilde{\phi}$. Letting $H(x) = zx - \phi(\tilde{x})$, and differentiating to obtain its supremum in x we obtain $x^{sup} = (-1)^{I_{\{\gamma \cdot y \cdot z < 0\}}} |\gamma \cdot y \cdot z|^{\frac{1}{\gamma}}$ and consequently $\tilde{\phi}^*(z) = \frac{|\gamma y z|^{\frac{\gamma+1}{\gamma}}}{\gamma+1} + \frac{1}{\gamma(\gamma+1)}$. Applying $\tilde{\phi}^*$ in Equation (13) gives optimization problem (26). The first-order conditions of this optimization problem with respect to λ are:

$$\frac{\partial \tilde{v}_{CR,mu}}{\partial \lambda} = q - E\{((-1)^{I_{\{\gamma y \lambda'_{mu} \cdot x < 0\}}} \cdot (|\gamma y \lambda'_{mu} \cdot x|)^{\frac{1}{\gamma}} y).x\} = 0 \qquad (28)$$

showing that $m_{CR,mu}$ in Equation (27) is an admissible SDF that minimizes the MD problem (25). $\diamond$

Similarly to the MD additive problem if we consider the existence of a risk-free asset r_f we can concentrate one of the Lagrange Multipliers and obtain a simplified result. The next corollary treats this case:

Corollary 4. *Assuming that there is a risk-free asset among the primitive securities then the dual optimization problem in Equation (26) can be simpliefied to (by also eliminating the constant term):*

$$v_{CR}^{mu,co}(\theta) = \max_{\tilde{\lambda} \in \Re^{n-1}} \tilde{\lambda}' q^{co} - E\left\{\frac{|1 + \gamma y \tilde{\lambda}' \cdot x|^{\frac{\gamma+1}{\gamma}}}{\gamma + 1}\right\} \qquad (29)$$

where q^{co} is the vector of prices of the $n - 1$ remaining primitive securities other than the risk-free asset. The corresponding admissible SDF that solves this problem is given by:

$$m_{CR}^{mu,co}(\theta) = (-1)^{I_{\{1 + \gamma y \lambda'_{mu} \cdot x < 0\}}} y |1 + \gamma y \tilde{\lambda}'_{mu} \cdot x|^{\frac{1}{\gamma}} \qquad (30)$$

Proof:
To prove this corollary, we follow essentially the same steps appearing in the proof of Corollary 2. $\diamond$

As we suggested in the beginning of this section, an important alternative formulation of the MD multiplicative problem concerns the case where we deal with only nonnegative SDFs, which will be another way of generalizing the least-square problem with positivity constraint solved by HJ (1997)[10]. This corresponds to the problem in Equation (22) where we eliminate the absolute value function that is composing with the convex ϕ function. In this case, the problem becomes completely compatible with the usual GMC literature. We derive the corresponding dual problem and MD admissible SDF in the next corollary to Theorem 3 and Corollary 4.

_____

[10]HJ (1997) actually solve the problem with nonnagativity constraint.

Corollary 5. *If in the Theorem 3 we solve the MD problem (22) instead of problem (23), and also assume the existence of a risk-free asset, the corresponding dual optimization problem and the MD admissible SDF are given by:*

$$\tilde{v}_{CR}^{mu,co}(\theta) = \max_{\tilde{\lambda} \in \Re^{n-1}} \tilde{\lambda}' q^{co} - E\left\{ \frac{(1 + \gamma y \tilde{\lambda}' \cdot x)^{\frac{\gamma+1}{\gamma}}}{\gamma + 1} \right\} \tag{31}$$

$$\tilde{m}_{CR}^{mu,co}(\theta) = y(1 + \gamma y \tilde{\lambda}_{mu}' \cdot x)^{\frac{1}{\gamma}} \tag{32}$$

Proof:

By eliminating the absolute value function in Theorem 3, calculating the convex conjugate of the Cressie Read function $\phi(\frac{\tilde{m}}{y})$, and eliminating the constant term, we obtain $\tilde{\phi}^*(z) = \frac{(\gamma y z)^{\frac{\gamma+1}{\gamma}}}{\gamma+1}$. By concentrating the LM of the risk-free asset out we obtain $\tilde{\phi}^{*,co}(z) = \frac{(1+\gamma y z)^{\frac{\gamma+1}{\gamma}}}{\gamma+1}$, which according to Borwein and Lewis (1991) leads to the dual maximization problem appearing in Equation (31). The first-order conditions of this problem with respect to $\tilde{\lambda}$ give the admissible SDF in Equation (32). ◇

Note that the correction to the proxy y in this last case comes as a hyperbolic function that depends on the optimal linear combination of primitive securities payoffs x and on the proxy itself. In particular, by making the proxy y equal to the constant 1, we obtain the nonparametric admissible SDFs from Almeida and Garcia (2008).

3.3 GMC Estimators and MD SDF Problems

We first formalize the concept of a Generalized Minimum Contrast estimator by following the notation of Section 3 in Kitamura (2006)[11].

Suppose that the econometrician observes IID realizations of a $\Re^n$ random variable z with probability law μ, and that he is interested in the model defined by the following set of moment conditions:

$$E[g(z, \theta)] = \int g(z, \theta) d\mu = 0, \ \theta \in \Re^k \tag{33}$$

Define a function D that will make use of function ϕ to measure the discrepancy between two probability measures P and Q:

$$D(P, Q) = \int \phi\left(\frac{dP}{dQ}\right) dQ \tag{34}$$

---

[11]where he relates GEL estimators to Generalized Minimum Contrast estimators. See also Corcoran (1998) for a sample version of minimum discrepancy estimators.

Given a general convex function ϕ, a Generalized Minimum Contrast estimator will seek to estimate the parameter vector θ by finding a probability measure π that satisfies the moment conditions in (33) and that minimizes the discrepancy with respect to the true unknown probability measure μ. This sentence is formalized by the following optimization problem:

$$v(\theta) = \inf_{\theta \in \Theta} \inf_{\pi \in P(\theta)} D(\pi, \mu) \tag{35}$$

where $P(\theta) = \left\{ \pi \text{ is a probability measure in } \Re^n \text{ satisfying: } \int g(z, \theta) d\pi = 0 \right\}$.

3.3.1 GMC and the MD Additive SDF Problem

Let us now give the direct interpretation of our MD additive problem from the point of view of GMC estimators. First note that in our problem, μ is the unknown probability law generating the observable payoffs x of the primitive securities. Given the family M of admissible SDFs m, we want to find a probability measure $\pi = \frac{m}{E_\mu(m)}$ that satisfies the nonparametric moment conditions imposed by the Euler equations for the primitive securities, and that in addition, is as close as possible to the asset pricing proxy y in the following sense: By constructing new probability measures $\tilde{\pi} = \frac{1+m-y}{E_\mu(1+m-y)}$ we want the π that will generate the $\tilde{\pi}$ closest (in the ϕ divergence sense) to the unknown probability μ. Of course, by construction, any of those $\tilde{\pi}$ will satisfy a transformed moment condition $E_\mu(\tilde{\pi}x) = r_f.q - E_\mu((y-1)x)$, where r_f represents the risk-free rate[12]. Define $\tilde{P}(\theta) = \{$probabilities $\tilde{\pi}$ such that: $\pi = r_f.(\tilde{\pi} + y(\theta) - 1)$ is a probability, and $\int (x - r_f.q) d\pi = 0\}$. Then our MD additive problem is equivalent to the following GMC problem:

$$v(\theta) = \inf_{\theta \in \Theta} \inf_{\tilde{\pi} \in \tilde{P}(\theta)} D(\tilde{\pi}, \mu) \tag{36}$$

The solution to this problem, according to Corollary 2 will be given by

$$\tilde{\pi}_{CR}^{ad} = \frac{1 + m_{CR}^{ad} - y(\theta)}{E_\mu(1 + m_{CR}^{ad} - y(\theta))} = \frac{(1 + \gamma \lambda_{ad}' \cdot x)^{\frac{1}{\gamma}}}{E_\mu(1 + \gamma \lambda_{ad}' \cdot x)^{\frac{1}{\gamma}}} \tag{37}$$

And in the particular case where the proxy correctly prices the risk-free asset, equation (37) reduces to:

$$\tilde{\pi}_{CR}^{ad} = (1 + \gamma \lambda_{ad}' \cdot x)^{\frac{1}{\gamma}} \tag{38}$$

The probability measure $\tilde{\pi}_{CR}^{ad}$ in its corresponding sample version, generates $\pi_{CR,i}^{ad}$'s that

---

[12]Here we assume that there is a risk-free asset r_f and also that the SDF proxy y correctly prices this risk-free asset.

are denominated implied probabilities.

Those implied probabilities are useful in a variety of applications. Brown and Newey (1998) obtain an efficient estimation of moment conditions based on implied probabilities. Brown and Newey (2002) suggest their use on the estimation of probability distribution functions via bootstrapping schemes. Smith (2004) shows how these probabilities can be used to obtain efficient moment estimation for GEL estimators. Antoine et al. (2007) show, in a context of Euclidean Likelihood, how implied probabilities contain important information coming from overidentifying restrictions that can be used to decrease the variance of the estimator.

In an asset pricing context, Almeida and Garcia (2008) show that implied probabilities can be used to derive nonparametric bounds for stochastic discount factors that by construction take into account higher moments of returns from primitive securities. The main idea relating implied probabilities and admissible SDFs is that any positive SDF is composed by a probability measure discounted by the risk-free rate. In particular, when the risk-free rate is assumed to be constant, and admitting that implied probabilities might achieve negative values for certain members of the Cressie Read family, we can extend the previous statement to the following one: Under Cressie Read discrepancy problems, any admissible SDF will be a linear transformation of a probability measure. As any Cressie Read discrepancy becomes a homogeneous function with an appropriate affine transformation, the MD problem for SDFs and for probability measures are equivalent what makes the implied probabilities exactly a normalized version of the admissible SDF m_{CR} in Equation (17), translated by $1 - y$ as shown in Equation (37).

On the corresponding sample version (of size T) of the Additive MD problem, we obtain:

$$\pi_{CR,i}^{ad} = T \frac{(1 + \gamma \tilde{\lambda}_{ad}' \cdot x_i)^{\frac{1}{\gamma}}}{\sum_{j=1}^{T}(1 + \gamma \tilde{\lambda}_{ad}' \cdot x_j)^{\frac{1}{\gamma}}}, \ i = 1, ..., T \tag{39}$$

3.3.2   GMC and the MD Multiplicative SDF Problem

Our MD multiplicative problem follows a standard GMC problem. In fact, given the family M of admissible SDFs m, we want to find a probability measure $\pi = \frac{m}{E_\mu(m)}$ that satisfies the nonparametric moment conditions imposed by the Euler equations for the primitive securities, and that in addition, is as close as possible to the asset pricing proxy y in the following sense: By constructing new probability measures $\tilde{\pi} = \frac{\frac{m}{y}}{E_\mu(\frac{m}{y})}$ we want the π that will generate the $\tilde{\pi}$ closest (in the ϕ divergence sense) to the unknown probability μ. Those $\tilde{\pi}$ will of course depend on θ and will satisfy the original moment condition $E_\mu(\tilde{\pi}.y.x) = r_f.q$, where r_f represents the risk-free rate. Defining $\tilde{P}(\theta) = \{$probabilities $\tilde{\pi}$ such that: $\int (y.x -$

$r_f.q)d\tilde{\pi} = 0\}$, we obtain the following GMC problem:

$$\tilde{v}(\theta) = \inf_{\theta \in \Theta} \inf_{\tilde{\pi} \in \tilde{P}(\theta)} D(\tilde{\pi}, \mu) \tag{40}$$

The solution to this problem, according to Corollary 5 will be given by

$$\tilde{\pi}_{CR}^{mu} = \frac{\frac{m_{CR}^{mu}}{y(\theta)}}{E_\mu(\frac{m_{CR}^{mu}}{y(\theta)})} = \frac{(1 + \gamma y(\theta)\lambda'_{mu} \cdot x)^{\frac{1}{\gamma}}}{E_\mu(1 + \gamma(\theta)\lambda'_{mu} \cdot x)^{\frac{1}{\gamma}}} \tag{41}$$

And in the particular case where the proxy correctly prices the risk-free asset, equation (41) reduces to:

$$\tilde{\pi}_{CR}^{mu} = (1 + \gamma y(\theta)\lambda'_{mu} \cdot x)^{\frac{1}{\gamma}} \tag{42}$$

On the sample version of the Multiplicative MD problem, we obtain:

$$\tilde{\pi}_{CR,i}^{mu} = T\frac{(1 + \gamma y_i(\theta)\lambda'_{mu} x_i)^{\frac{1}{\gamma}}}{\sum_{j=1}^{T}(1 + \gamma y_j(\theta)\lambda'_{mu} x_j)^{\frac{1}{\gamma}}}, \quad i = 1, ..., T \tag{43}$$

Note that in both the additive and multiplicative MD problems, the probabilities may achieve negative values. It is documented in the literature that in general for members of the Cressie Read family with negative parameter γ the probabilities are naturally positive. However, for positive γ's these probabilities might achieve negative values. As in our particular financial environment negative probabilities will imply negative states of the nature for the stochastic discount factor (what in a complete market setting would imply the existence of arbitrage), it may be interesting to restrict the original discrepancy problem to strictly positive admissible SDFs[13]. On the other hand, in an incomplete market setting, having an admissible SDF as solution to the MD problem with negative values in some states does not rule out absence of arbitrage, as the only implication of no-arbitrage is that it should exist at least one positive admissible SDF, but it does not have to be exactly the one that solves our optimization problem (see Cochrane (2000)).

3.4 Interpreting the Dual Optimization Problems

In a seminal work, 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)). He showed that the ET entropy minimization problem corresponds to an optimal portfolio problem with a CARA utility function. Based on the same two-period model, Almeida and Garcia (2008) extended his interpretation to the whole Cressie Read

---

[13]This can be obtained by imposing that $1 + \gamma\lambda'_{mu}x > 0$ in the additive case, and that $1 + \gamma y(\theta)\lambda'_{mu}x > 0$ in the multiplicative case (here also assuming that the proxy $y > 0$).

family in a nonparametric setting. Here we extend their interpretation to the additive and multiplicative dual optimization problems in a semi-parametric setting.

Both dual MD problems will admit interesting economic interpretations as optimal portfolio problems. Let us first analyze the solution of our MD multiplicative problem for Cressie Read estimators since it will present a very similar interpretation to that appearing in Almeida and Garcia (2008): According to Corollary 5 it corresponds to an optimal portfolio problem based on a specific HARA-type utility function:

$$u(W, y(\theta)) = -\frac{1}{\gamma+1} \left(1 + \gamma y(\theta)W\right)^{\left(\frac{\gamma+1}{\gamma}\right)}, \tag{44}$$

with W and y such that $1 + \gamma y(\theta)W > 0$ a sufficient condition to guarantee the concavity and strict monotonicity of function u.

Suppose an investor distributes his/her initial wealth W_0 putting λ_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 (44), solving the following optimal portfolio problem:

$$\Omega = \sup_{\lambda \in \Lambda} E\left(u(W)\right) \tag{45}$$

where $\Lambda = \{\lambda : u(W(\lambda))$ is strictly increasing and concave$\}$, and expectation is taken with respect to W and y. By conditioning on y and by scaling the original vector λ to be $\tilde{\lambda} = \frac{\lambda}{(1+\gamma y(\theta)W_0.rf)}$, we can decompose the utility function in $u(W) = u(W_0 * r_f) * \left(1 + \gamma y(\theta)\tilde{\lambda}\left(R - \frac{1}{a}\right)\right)^{\left(\frac{\gamma+1}{\gamma}\right)}$. This decomposition essentially shows that solving a version of the GEL optimal problem appearing in Corollary 5 for excess returns (instead of payoffs) 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, conditional on the stochastic process describing y. Of course by integrating on the distribution of y we would obtain an average of the gains from diversifying the allocation. In Almeida and Garcia (2008) the HARA function does not depend on y what simplifies the problem to that equivalent to a degenerate distribution of y. Here $y(\theta)$ works as a normalizing factor for the returns.

When considering the additive dual MD optimization problems appearing in Corollary 2, we can interpret them as allocation problems among the n primitive securities in a way

to maximize the utility function U defined by:

$$U(W) = -\frac{1}{\gamma + 1} (1 + \gamma W)^{(\frac{\gamma+1}{\gamma})} - (y(\theta) - 1).W, \qquad (46)$$

Note that the utility is composed by two terms, one linear in wealth and the other, in principle, given by a HARA utility exactly as in the multiplicative MD problem, except that it does not depend on the proxy model y. If we again take expectations conditioning on the proxy y, the linear term would correspond to a risk-neutral economy with stochastic discount factor $1 - y$. In fact, here $1 - y$ works as a penalty factor to avoid the resulting marginal utility from the HARA part to be too far from y.

3.5 Some Special Cressie Read Discrepancies

In this section we specialize the results in Theorems 1 and 3 to provide the dual (portfolio-type) optimization problems and corresponding admissible SDFs solutions for some special discrepancies in the Cressie Read family frequently adopted in the econometric literature. We begin by investigating the relation between Euclidean Likelihood and the HJ (1997) distance under both the additive and multiplicative forms of our discrepancy problems. In a sequence, we provide, again for both forms, the optimization problems and solutions under EL (CR with $\gamma = -1$) and ET (CR with $\gamma = 0$) discrepancies. The other three additional discrepancies adopted in the empirical section, namely Pearson's Chi-Square (Cressie Read with $\gamma = -2$), Hellinger's distance (CR with $\gamma = -\frac{1}{2}$), and CR with $\gamma = 2$ can be obtained directly by application of Theorem 1 with their specific gammas.

3.5.1 Hansen and Jagannathan Distance Derived from Euclidean Likelihood

Euclidean likelihood or CUE is obtained by fixing $\gamma = 1$ on the Cressie Read discrepancy. By using this value of gamma in the Corollary 2 above and dropping the constant terms, we obtain the following optimization problem:

$$v_{CUE}^{ad,co}(\theta) = \max_{\lambda \in \Re^{n-1}} \lambda'q - E\{\frac{1}{2}(1+\lambda'\cdot x)^2 + (y-1)\lambda'x\} = \max_{\lambda \in \Re^{n-1}} \lambda'q - E\{\frac{1}{2}(\lambda'\cdot x)^2 + y\lambda'x\} \ (47)$$

whose first order conditions obtains $m_{CUE} = y + \lambda'x$, precisely the linear correction term obtained by HJ (1997). By comparing Equation (47) to Equation (7) we note that the two problems are equivalent. Thus, under our proposed MD problem of additive type the HJ distance becomes one element within the particular Cressie Read family.

Note that under the multiplicative MD formulation proposed in Section 3.2 this equivalence to the Hansen and Jagannathan (1997) distance is not maintained. In fact, by taking $\gamma = 1$ on Corollary 5, we obtain the following dual optimization problem:

$$\tilde{v}^{mu,co}_{CUE}(\theta) = \max_{\lambda \in \Re^{n-1}} \lambda'q - E\{\frac{(1 + y\lambda' \cdot x)^2}{2}\} \tag{48}$$

Note that the asset pricing model y appears as a quadratic term in (48) that is not eliminated like in the additive case, and could not reproduce HJ (1997) since it would have to appear linearly in the dual optimization problem. The admissible SDF that solves this optimization problem is, according to Corollary 5, given by $\tilde{m}_{CR} = y(1 + y\lambda'_{mu} \cdot x)$. Observe that while in HJ (1997) the solution to the problem that asks what is the least squares distance of a proxy y to the family M of admissible SDFs pricing primitive securities x is given by the distance between y and the admissible SDF $y + \lambda'_{ad} \cdot x$, here the admissible SDF which is closest to y in the multiplicative CR divergence sense is quadratic in y: In fact it can be seen as y added to a correcting term that is quadratic on y and that also depends on the primitive securities payoffs $(y^2\lambda'_{mu} \cdot x)$ that guarantees it to be admissible. By looking at these two results obtained in this quadratic case, it appears to be the case that the CR $(\gamma = 1)$ multiplicative distance will be more sensitive to changes in the proxy y than the corresponding HJ distance, or CR $(\gamma = 1)$ additive distance.

3.5.2 Empirical Likelihood $(\gamma = -1)$

In this limiting interesting case, the Cressie Read discrepacy converges to $\phi(\pi) = -ln(\pi)$. Our MD problem under the additive form becomes:

$$\delta^{ad}_{EL}(\theta) = \min_{m \in L^2} E\{-ln(1 + m - y(\theta))\} \text{ subject to } E(mx) = q \tag{49}$$

Noting that the expression $\frac{(1+\gamma x)^{\frac{\gamma+1}{\gamma}}}{\gamma+1} + \frac{1}{\gamma(\gamma+1)}$ converges to $-1 - ln(1-x)$ when $\gamma \to -1$, and applying the results in Corollary 2, the dual optimization problem becomes:

$$v^{ad,co}_{EL}(\theta) = \max_{\tilde{\lambda} \in \Re^{n-1}} \tilde{\lambda}'q^{co} - E\left\{-ln(1 - \tilde{\lambda}' \cdot x) + (y(\theta) - 1)\tilde{\lambda}' \cdot x\right\} \tag{50}$$

The corresponding admissible SDF that solves this problem is given by:

$$m^{ad,co}_{EL}(\theta) = y(\theta) - 1 + \frac{1}{(1 - \tilde{\lambda}'_{ad} \cdot x)} \tag{51}$$

Similarly, the EL in the multiplicative case of Corollary 5 becomes:

$$v^{mu,co}_{EL}(\theta) = \max_{\tilde{\lambda} \in \Re^{n-1}} \tilde{\lambda}'q^{co} - E\left\{-ln(1 - y\tilde{\lambda}' \cdot x)\right\} \tag{52}$$

The corresponding admissible SDF that solves this problem is given by:

$$m_{EL}^{mu,co}(\theta) = \frac{y}{(1 - y\tilde{\lambda}_{mu}' \cdot x)} \tag{53}$$

3.5.3 Exponential Tilting ($\gamma = 0$)

The ET discrepancy is also a limiting case on the Cressie Read family studied by Kitamura and Stutzer (1997). The Cressie Read discrepacy converges in this case to $\phi(\pi) = \pi ln\pi$, whose convex conjugate is e^{z-1}.

Then our MD problem under the additive form becomes:

$$\delta_{ET}^{ad}(\theta) = \min_{m \in L^2} E\{(1 + m - y(\theta))ln(1 + m - y(\theta))\} \text{ subject to } E(mx) = q \tag{54}$$

Noting that the expression $\frac{(1+\gamma x)^{\frac{\gamma+1}{\gamma}}}{\gamma+1} + \frac{1}{\gamma(\gamma+1)}$ converges to e^x when $\gamma \to 0$ and applying the results in Corollary 2, the dual optimization problem becomes:

$$v_{ET}^{ad,co}(\theta) = \max_{\tilde{\lambda} \in \Re^{n-1}} \tilde{\lambda}' q^{co} - E\left\{e^{\tilde{\lambda}' \cdot x} + (y(\theta) - 1)\tilde{\lambda}' \cdot x\right\} \tag{55}$$

The corresponding admissible SDF that solves this problem is given by:

$$m_{ET}^{ad,co}(\theta) = y(\theta) - 1 + e^{\tilde{\lambda}_{ad}' \cdot x} \tag{56}$$

The multiplicative problem based on Corollary 5 is given by:

$$v_{ET}^{mu,co}(\theta) = \max_{\tilde{\lambda} \in \Re^{n-1}} \tilde{\lambda}' q^{co} - E\left\{e^{y\tilde{\lambda}' \cdot x}\right\} \tag{57}$$

The corresponding admissible SDF that solves this problem is given by:

$$m_{ET}^{mu,co}(\theta) = ye^{y\tilde{\lambda}_{mu}' \cdot x} \tag{58}$$

3.6 Model Estimation Based on Minimum Discrepancy Bounds

Researchers have been using the HJ (1997) distance to estimate asset pricing models by finding the parameter vector θ^* that minimizes this distance. Similarly to their approach and following Kitamura (2006) and the whole literature in Minimum Discrepancy estimators, we propose estimating the above asset pricing models by finding the parameter vector θ_{MD} that minimizes any specific discrepancy function either described by Equation (12) or (22)[14]:

---

[14]or also by the corresponding versions with the absolute value function that would consider negative admissible SDFs. See Equation (23) for the MD multiplicative case with negative SDFs.

$$\theta_{MD}^{ad} = \underset{\theta \in \Re^k}{argmin} \; \delta_{MD}^{ad}(\theta) \tag{59}$$

or

$$\theta_{MD}^{mu} = \underset{\theta \in \Re^k}{argmin} \; \delta_{MD}^{mu}(\theta) \tag{60}$$

Note that these problems are composed optimization problems and for any fixed θ_0 in the parameter space, the inner problem will deliver a minimum discrepancy coming from the admissible SDF closest (in ϕ sense) to $y(\theta)$, and a set of Lagrange Multipliers representing portfolio weights from a HARA utility function (see Section 3.4). In the empirical section instead of estimating models based in (59) or (60), in order to understand the mechanics of the MD problems, we begin by taking a step behind and concentrate on the analysis of the inner optimization problems of the additive MD problem. By calibrating values of θ in the parametric space we analyze the sensitivity of the different divergencies to changes in parameters values of a specific asset pricing model. In a sequence, we also analyze the calibration process of the multiplicative problem, and then estimate both problems and compare their results.

4   Empirical Application

HJ (1997) illustrated the usefulness of their least-square projection by analyzing the degree of misspecification of the canonical consumption-based asset pricing model of Breeden (1979) and Lucas (1978) for various values of the preference parameters. We will perform a similar analysis here but considering several Cressie Read discrepancy functions: Pearson's, EL, Hellinger's, ET, CUE, and two other discrepancies with high positive values of the Cressie Read parameter γ ($CR(\gamma = 2)$, and $CR(\gamma = 5)$).

The CCAPM SDF is given by:

$$m_t^{ccapm} = \beta E \left(\frac{C_{t+1}}{C_t} \right)^{-\theta}. \tag{61}$$

where C_t denotes the time t aggregate consumption in the economy considered.

We use the annual (1890- 1985) time-series data on stocks and bonds of Campbell and Shiller (1989) updated to 2004 and the corresponding aggregate consumption annual series[15]. Similarly to HJ (1997), we propose a small grid for values of the risk aversion coefficient θ to analyze the sensitivity of different discrepancy functions to changes in the parametric space. We concentrate on the risk-aversion parameter as it is the most important parameter in the CCAPM and since it is the one generating nonlinearities in the model.

---

[15]The dataset is available on Shiller's website, http://www.econ.yale.edu/ shiller/data.htm.

For this reason, for each value of θ, β is fixed to a value that guarantees that the mean of the CCAPM SDF proxy is always equal to 0.98, the averaged value of the historical 1-month Treasury Bill.

For fixed values of the parameters β and θ, and given a time series of consumption growth rates we can compute the SDF m^{ccapm}. Once we know the SDF proxy, it is possible to compute pricing errors, to estimate the discrepancy distance δ_{CR} and Lagrange multipliers for any fixed discrepancy in the Cressie Read family (fixed γ).

We chose values for the risk aversion parameter θ from two very distinct regions of the parametric space. Small values of θ ($\theta = 1, 5$) will correspond to small volatility CCAPM SDFs that will have more difficulty in pricing the stock returns (S&P 500). On the other hand, high values of θ ($\theta = 20, 50$) will generate more volatile CCAPM SDFs that will have variation compatible with the extreme variation of equity returns. This behavior description for the CCAPM SDF is compatible with the equity premium puzzle first document by Mehra and Prescott (1985) and reexpressed in terms of SDF's variance bounds by HJ (1991).

4.1 Lagrange Multipliers (portfolio weights) and Implied SDFs

We start by analyzing results obtained with the Pearson, EL, Hellinger, ET, CUE, and $CR(\gamma = 2)$ discrepancies. For each fixed discrepancy and parameter θ value, we solve the MD additive optimization problem proposed in Equation (16) to find the Lagrange Multipliers (LM), the corresponding implied admissible SDF, and the value that minimizes the discrepancy.

Table 1 presents the LM estimated with each CR discrepancy. As noticed in section 3.4, those LM estimates correspond to optimal portfolio weights from the maximization of a HARA utility function (plus a linear term) when the agent can invest in a short-term bond and/or the S&P 500. The dual HARA functions obtained from the Cressie Read discrepancies need an inversion of sign in LM weights to be strictly increasing. In this sense, negative weights in the table indicate that the agent is buying an asset while positive ones indicate selling it. We observe that for all values of the parameter θ within the grid, all discrepancies agree on the signs of the weights attributed to the bond and the S&P: they all sell the bond and buy the S&P. According to the results appearing in Corollary 2, the admissible SDFs that solve the concentrated additive MD problem should be negatively correlated to the S&P returns whenever the nonparametric term given by $(1 + \gamma\tilde{\lambda}'_{ad}x)^{\frac{1}{\gamma}}$ dominates the parametric term coming from the CCAPM. This is exactly what can be observed in Table 2 and figures 1, 2, 3, and 4.

Table 2 presents the correlation of admissible SDFs with the S&P 500 returns. In the

last column it presents the correlation of the parametric CCAPM model with the S&P 500 returns. For all discrepancies, the correlation is decreasing in absolute value with the value of θ. This is well illustrated in Figures 1 to 4. They present for each value of θ and each discrepancy function, the CCAPM SDF (dashed line) and the corresponding admissible SDF (solid line) that is closest to the CCAPM SDF in that region of the parametric space. Note that for small values of θ (see Figures 1 and 2), the nonparametric part of the admissible SDF generates more variability than the CCAPM (compare the solid and dashed lines). For larger values of θ (see Figures 3 and 4), the CCAPM term shows high variability and the correlation between admissible SDf and S&P returns goes from high negative values to small negative values, which is precisely the correlation of the CCAPM SDF with the S&P returns (see last column of Table 2).

Still observing the portfolio weights (or LMs) in Table 1, we can see that the weights for the S&P are not very sensitive to changes in the parameter value θ while the weights in the short-term bond clearly decrease with θ. This has an intuitive interpretation: since when increasing the value of θ we increase the variability of the CCAPM SDF, any admissible SDF that will solve the MD problem should present the nonparametric term $(1 + \gamma \tilde{\lambda}'_{ad} x)^{\frac{1}{\gamma}}$ with volatility of the magnitude of the parametric term (the CCAPM SDF). The way to achieve this high volatility is to keep higher weights on the S&P and lower weights in the bond.

4.2   Discrepancy Measures and Implied Probabilities

We next move to the analysis of the minimum discrepancy values obtained by solving the dual maximization problems that will capture the degree of misspecifation of the CCAPM model in each region of the parametric space. Table 3 presents the minimizing values for the discrepancy functions adopted. Note that all Cressie Read discrepancies achieve their smallest value (considering the parameter grid) when $\theta = 50$. In principle, if we had to choose a parameter value based on any of these discrepancy problems we would choose the same as HJ (1997), which corresponds to our CUE quadratic problem. However, the behavior of the implied admissible SDFs for each discrepancy function varies a lot, specially for smaller values of the parameter θ (see again Figures 1 to 4). For instance, while Cressie Read estimators with non-positive γ (Pearson, EL, Hellinger, and ET) produce SDFs that are positively skewed with respect to the constant 1 (have more extreme positive values) the corresponding estimators with positive γ (CUE and CR($\gamma = 2$)) produce SDFs that are negatively skewed with respect to the constant 1. Also, SDFs implied by CR estimators with increasing γ become less and less volatile in general. Of course, these differences become more subtle once we increase the parameter value to a high risk-aversion coefficient

on the CCAPM model, but still, a careful observation of the implied admissible SDFs reveals that estimators with higher γ get closer to the CCAPM SDF with smaller values of the parameter θ (see for instance, the picture in the right bottom of Figure 2).

In order to further analyze this point, based on equation 37 we computed the implied probabilities corresponding to those admissible SDFs. Here we have a total of 115 annual observations what generates constant empirical probabilities equal to $\pi^{emp} = \frac{1}{115} = 0.0087$. Figures 5 to 8 show these implied probabilities. Note that for any fixed value of the CCAPM parameter θ, the variability of those probabilities around the empirical probability (dashed line) is a decreasing function of the Cressie Read parameter γ. In particular, for small values of θ the implied probabilities of the CR ($\gamma = 2$) estimator are already very close to the empirical probabilities. This suggests that perhaps for higher values of the Cressie Read parameter γ it might happen that the estimator will minimize the discrepancy on smaller values of the CCAPM risk aversion parameter θ. This would be interesting since apparently the SDFs (both the CCAPM and the corresponding admissible ones) appear to be much better behaved from an economic viewpoint for the value of $\theta = 5$ than for $\theta = 50$. In fact, Almeida and Garcia (2008) showed that while the CCAPM model is only accepted in the admissible region of the HJ (1991) variance bounds for very high values of the risk aversion coefficient (θ), when analyzing Minimum Discrepancy bounds with Cressie Read estimators with high positive values of the Cressie Read parameter γ, the CCAPM becomes admissible with much smaller values of the risk aversion coefficient (θ).

To further investigate this point we implemented an estimator with very high γ: The Cressie Read with $\gamma = 5$. Figures 9 and 10 show respectively the implied SDFs and corresponding implied probabilities for all values of the parameter θ. Note in figure 9 how in this case, even for $\theta = 5$, the implied admissible SDF is already very close to the corresponding CCAPM SDF. Now compare the implied probabilities of this estimator with the implied probabilities of the previously analyzed estimators. We can safely conclude that the implied probabilities here are very close to the empirical probabilities even for $\theta = 1$. In fact, by looking at Table 4 we observe that the CR discrepancy for $\gamma = 5$ is minimized when $\theta = 20$, confirming our intuition that estimators with high values of CR γ will pick up smaller values of the parameter θ. However, a question remains to be answered: How come the estimator is able to pick up admissible SDFs so close to the CCAPM SDF when the parameter θ is small, like 1 or 5?

4.3 Pricing Errors

To try to answer this question we calculate the pricing errors obtained with each implied SDF and also with the CCAPM SDF. These errors are presented in Table 5. First note

that for the Cressie Read estimators EL, Hellinger, ET, and CUE the pricing errors are not presented since they are practically zero (all smaller than 10^{-7}). Also note that we should expect that the CCAPM SDF will present in general the highest pricing errors since all the other (implied) SDFs are trying to correct the CCAPM SDF to become an admissible one. For all the CR estimators shown we can easily conclude that the moment condition of the S&P returns becomes binding exactly when the γ parameter in absolute value is high. For Pearson's ($\gamma = -2$) the pricing errors are still acceptable (between $\frac{1}{10}$ and $\frac{1}{6}$ of the CCAPM errors). For CR ($\gamma = 2$) they are already high (around 25% of the original CCAPM errors) but it is precisely for the CR estimator with highest γ that the errors appear to be very high around 60% of the original CCAPM errors). This happens because to keep the HARA function $(1 + \gamma \lambda' x)^{\frac{\gamma+1}{\gamma}}$ concave for high values of γ imposes a very tight restriction since $(1 + \gamma \lambda' x)$ must be positive. This constraint forces the optimization problem to choose a truncated version of an admissible SDF as solution.

A way to avoid high pricing errors for high values of the Cressie Read parameter γ is to redefine the original minimum discrepancy function with a composition of the convex function ϕ with the absolute value function $\phi(|1 + m - y|)$. This function is still convex (just check by calculating the second derivative) and its solution will admit negative values of $(1 + \gamma \lambda' x)$ while still keeping concave the HARA function from the dual optimization problem. Of course there is a cost to all that: For the region where $(1 + \gamma \lambda' x)$ is negative both the admissible SDF as well as the corresponding implied probabilities will achieve negative values. In order to illustrate this point we present the solution of the Cressie Read ($\gamma = 5$) estimator allowing the SDF (and the implied probabilities) to become negative. Figures 11 and 12 show respectively the implied SDFs and corresponding implied probabilities for all values of the parameter θ. In this case, we produce zero pricing errors exactly as the estimators with low values of gamma, but we see that a few points (5 in a total of 115 observations) become negative. Interestingly the solution to this MD problem also chooses $\theta = 20$, exactly as the restricted CR ($\gamma = 5$) problem described above. In fact, by looking at the implied admissible SDFs we observe that apart from these 5 negative values the CCAPM and the admissible SDF are practically the same for $\theta = 20$.

4.4 Comparison between the Additive and the Multiplicative CR Cases

4.4.1 Implied Admissible SDFs

Using the same grid for the CCAPM parameter theta, we obtain admissible SDFs and corresponding implied probabilities under the multiplicative case and compare to results shown in the additive case. Figures 13-16 present additive (dotted line) and multiplicative (solid line) admissible SDFs for all analyzed Cressie Read discrepancies and for the four

values of the risk aversion coefficient theta (1, 5, 20, and 50). By observing the pictures we note that the two types of admissible SDFs (additive and multiplicative) have similar shapes, specially for small values of the parameter theta like 1 or 5 (see Figures 13 and 14). However, a clear difference appears in extreme values of the SDFs: in general, the multiplicative SDFs are more positively skewed than the additive SDFs. The differences in extremes are more exacerbated for high values of theta like 20 or 50 (see Figures 15 and 16). A heuristic explanation for these differences can be constructed by looking at the analytical SDF formulas in Equations (20, additive) and (32, multiplicative). First, note that the asset pricing model proxy y (the CCAPM) only enters as an additive term on the additive SDF and it enters in two ways on the multiplicative SDF, as a term multiplying the primitive securities within a hyperbolic function, and as an extra linear factor multiplying this hyperbolic function. Intuitevely we could expect that whenever the proxy y is above one, the extra linear factor will make the multiplicative SDF to achieve higher values than the additive one, exactly as observed in the figures. On the other hand, when the proxy y is below one, apparently the hyperbolic term becomes dominant in the multiplicative SDF.

To better analyze the differences in goodness of fit of the two types of estimators, we present additional graphs in Figure 17 for a fixed risk aversion coefficient of 20, including the additive, the multiplicative and the CCAPM SDFs[16]. The graphs indicate that the positive skewness of the multiplicative SDF, which is specially pronounced on extreme values (higher than 2), makes the multiplicative estimator worse than the additive in capturing peaks of the proxy model (see the left half of the SDF graphs including observations 1 to 60). On the other hand, the multiplicative estimator appears to be better in capturing the values of CCAPM that are not too far from the mean of 0.98 (see the right half of the SDF graphs including observations 60 to 115).

4.4.2 Implied Probabilities

From section 3.3 we see that the implied probabilities under the additive estimator are obtained by a hyperbolic function of the underlying assets returns, while under the multiplicative estimator they are hyperbolic functions of returns weighted by the proxy model. For small values of the risk aversion parameter (theta equal to 1, or 5) the implied probabilities under these two types of estimators are very similar. Therefore we do not report them here. These probabilities begin to achieve different shapes across estimators (add and mul) only in the region of the parametric space that is closer to the optimal parameter values,

---

[16]We only present pictures for a high value of theta because in this region of the parametric space, the admissible SDFs begin to approximate the corresponding CCAPM proxy providing a clearer perspective of the differences in goodness of fit between the two types of estimators. Note also that we only include the Pearson, CUE and CR (g=2) cases because the three other provide very similar pictures.

meaning the region for high values of theta (see section 3.6 for details on the estimation). Figure 18 shows, for a fixed risk aversion coefficient of 20, the implied probabilities for the additive and the multiplicative estimators, for some representative CR discrepancies (Pearson, CUE, and CR (g=2)). Note that the multiplicative implied probabilities are clearly less volatile than the additive ones. However, the multiplicative still have some extreme values that are higher than the additive ones (see especially observations 7 and 56 under CUE, 7 under Pearson, and 57 under CR (g=2)). Recall that the objective of implied probabilities is to satisfy the moment conditions and simultaneously be as close as possible to the homogeneous empirical probabilities. In this sense, the results presented in Figure 18 confirm the higher stability of the multiplicative estimator when compared to the additive one.

4.4.3 Pricing Errors

As we have shown before another important aspect to be analyzed is the existence or not of pricing errors for certain elements of the Cressie Read family. We observed that for a fixed value of the parameter theta, high values of gamma restrict the dual estimator to search for optimal weights in smaller regions of the portfolio space since the HARA function $(1+\gamma\lambda'x)^{\frac{\gamma+1}{\gamma}}$ should continue to be concave. It is interesting to note that under the multiplicative estimator however there is a change in the HARA function to $(1+\gamma y\lambda'x)^{\frac{\gamma+1}{\gamma}}$, including an extra term on the proxy model y that may change the way gamma imposes a restriction. In fact, Table 6 presents pricing errors under the multiplicative estimator for respectively the Pearson, the CR (g=2), and the CR (g=5) estimators (for all other CR estimators the pricing errors are, similarly to the additive case, practically zero). Note that in general, pricing errors are smaller than in the additive case. For instance, for the Pearson estimator they are zero in the whole grid of the parameter theta (except for theta=5, but the errors are small there). For the other two CR estimators, we observe that the pricing errors of the short-term bond increase under the multiplicative estimator but are still small in absolute values, achieving at most 46 basis points except for one observation under CR (g=5, theta=50) where the error is high (128 basis points). On the other hand the pricing errors of the equity returns that are the highest in absolute value under both additive and multiplicative estimators, decrease significantly by more than 20 percent, with a few exceptions. Results here indicate that under the CCAPM model the existence of an extra proxy term within the HARA function improves the estimator ability to price the underlying primitive assets.

4.5 Model Estimation

We follow the estimation proposal from Section 3.6, and solve the optimization problems under the additive and multiplicative problems for all the CR discrepancies previously analyzed. Figures 19 and 20 present the results of the estimation process, that is, the estimated risk aversion coefficient theta and the corresponding admissible SDF that is closest to the estimated parametric model (the CCAPM). Observing the pictures we see that all the CR discrepancies under both additive and multiplicative estimators, obtain very high values for the risk aversion coefficient confirming the failure of the CCAPM model to explain the equity premium with an acceptable risk aversion coefficient. The additive estimators achieve parameters of the order of 37, ranging from a minimum of 37.2 (under the CR (g=2) estimator) to a maximum of 43.0 (under the Pearson estimator). The Pearson estimator produces too high a value when compared to the other CR additive estimators (all were between 37.2 and 37.8). The reason for such difference in values might rely on the fact that only for the additive Pearson estimator we had to make use of the composed absolute function on the optimization problem since under the original problem the HARA function was not well defined under certain portfolio weights and certain returns of primitive assets. The multiplicative estimators produce parameter values around 64, ranging from a minimum of 62.9 (under the CR (g=2)) to a maximum of 65.9 under the Pearson estimator. The multiplicative estimators produce more homogeneous values in accordance to the stability analysis performed before under the fixed grid of theta parameters. Note that the multiplicative estimators also are more emphatic in rejecting the CCAPM model producing risk aversion coefficients that are more than fifty percent higher than the ones produced under the additive estimators. Another interesting aspect of the estimation process is that, in accordance to results obtained by Almeida and Garcia (2008) when analyzing the CCAPM model based on nonparametric MD bounds, the authors found that high values of gamma tend to attenuate the rejection of the CCAPM model. Similarly here under a more formal estimation procedure we observe that higher values of gamma like 2 produce smaller values for the estimated risk aversion parameter theta, under both additive and multiplicative estimators.

Also in the pictures we can confirm that the multiplicative admissible SDFs have a better fit to the corresponding estimated CCAPM SDF than the additive SDFs, except that they miss some extreme points, exactly as analyzed when we calibrated the risk aversion parameter theta before. In addition, the additive admissible SDFs clearly achieve some negative values while the multiplicative ones do not. Additional controls to guarantee that the additive SDFs become always positive might be implemented but with the cost of possibly introducing/increasing pricing errors on the primitive securities returns.

5 Discussion

A few lessons can be learned from the empirical application we just described. While applying our new discrepancy measures may appear like beating a dead horse, the exercise is in fact very instructive in terms of misspecification. Indeed, the Cressie-Read discrepancy measures all concur in showing that the CCAPM model becomes compatible with return data only for high values of the risk-aversion. However, our approach allows us to recover the implied nonparametric SDF or equivalently the pricing errors and this has information about the potential source of misspecification. Indeed, as usual in econometrics, a good misspecification test should provide some insight about the source of the specification problem and suggest some direction for improving the model. For a regression model, analyzing and testing some characteristics of the residuals will reveal some missing variables. Similarly, pricing errors can be suggestive of directions to improve the asset pricing model.

We discussed in section 4.3 that forcing the model to achieve zero pricing errors may not be the best way to obtain economically meaningful results. For EL, Hellinger, ET, and CUE criteria, pricing errors are zero but implied SDFs are not satisfying from en economic point of view with huge discounting factors as we increase θ. For a high γ, reducing pricing errors to zero means obtaining a number of negative values for the implied probabilities. However, we have seen that allowing a higher gamma will produce pricing errors but will choose a lower value of theta and generate more reasonable implied SDFs for values of θ such as 5 (see Figure 9). One can extract the pricing errors and relate them to variables entering in more elaborate asset pricing models such as recursive utility or habit formation models for consumption-based asset pricing models or in factor models when extending a CAPM model. One can also analyze the distribution of the pricing errors to determine the potential presence of skewness and kurtosis and identify the periods where mispricing is more prevalent.

The richness of our framework in terms of discrepancy measures may be construed as a hurdle since a criterion has to be chosen to pick a gamma among possible values or even to choose between many families of discrepancies. In Almeida and Garcia (2008), we discuss robustness issues related to diagnosing asset pricing models and performance evaluation. For diagnosing models, we showed that increasing γ allows to lower the risk aversion parameter that will make the CCAPM model admissible. Intuitively, the increased admissibility of the CCAPM with higher values of γ comes from the fact that a lower risk aversion parameter θ in the CCAPM model will be needed to accommodate the weighted returns, since relatively lower probabilities will be assigned to extreme states, especially to good states, exactly as the theory of marginal utility suggests. In the case of our

misspecification measure and the corresponding statistics that come with it, varying the γ will tell us to what extent the model assessment is dependent upon the discrepancy measure chosen. Allowing for this robustness analysis is in our view a good feature of our approach since it can tell us in which direction to improve the asset pricing models at hand.

6 Conclusion

We extend the least-square projection proposed by Hansen and Jagannathan (1997) to measured the degree of misspecification of asset pricing models by suggesting more general projections based on the minimization of discrepancy convex functions. Solutions to these Minimum Discrepancy (MD) problems naturally imply semiparametric and nonlinear SDFs that take into account higher moments of the distributions of assets returns. We relate the problem of finding general MD projections of asset pricing models onto the family of admissible SDFs to that of solving an optimal portfolio problem. When specializing to the Cressie Read family of discrepancies, our projections are obtained as solutions to optimal portfolio problems based on HARA utility functions added to a linear term on the asset pricing proxy that imposes the proxy as an imperfect SDF benchmark. We also relate the MD admissible SDFs to the implied probabilities from the econometric literature (see Newey and Smith (2004)), showing that in our context those probabilities are a normalized version of the admissible SDFs translated by an affine function of the asset pricing proxy model.

We apply our methodology to empirically analyze the CCAPM model, making use of a number of well-known Cressie Read discrepancies, namely Pearson's, EL, Hellinger's, ET, CUE, and two other discrepancies with high positive values of the Cressie Read parameter γ ($CR(\gamma = 2)$, and $CR(\gamma = 5)$). Based on a grid for the risk aversion parameter of the CCAPM model, we show that most of the discrepancies agree on the choice of θ but once we increase too much the Cressie Read parameter γ, the MD problems stop satisfying the moment conditions (introducing pricing errors), and end up choosing θ's in regions of the parametric space other than the one chosen by the previous estimators. We also perform estimation of the CCAPM model based on all the mentioned discrepancies and also in two different formulations of the MD problems (additive and multiplicative). We discuss the empirical findings across Cressie Read discrepancies relating them to the size of the pricing errors, and to the magnitude of the admissible SDFs. Our results indicate that this new class of higher-order SDF projections has a strong potential to be used as a tool to estimate and rank asset pricing models, specially when estimation is based on assets with nonlinear payoffs.

References

[1] Antoine B., H. Bonnal, and E. Renault (2007). On the Efficient Use of the Informational Content of Estimating Equations: Implied Probabilities and Euclidean Empirical Likelihood. *Journal of Econometrics*, **138**, 461-487.

[2] Almeida C. and R. Garcia (2008). Empirical Likelihood Estimators for Stochastic Discount Factors. Working Paper, Getulio Vargas Foundation and EDHEC Business School.

[3] Araujo A. and P.K. Monteiro (1989). Equilibrium without Uniform Conditions. *Journal of Economic Theory*, **48**, 2, 416-427.

[4] Back K., and D.P. Brown (1993). Implied Probabilities in GMM Estimators. *Econometrica*, **61**, 971-975.

[5] Baroni-Adesi G. (1985). Arbitrage Equilibrium with Skewed Asset Returns. *Journal of Financial and Quatitative Analysis*, **20**, 3, 299-313.

[6] Beaulieu M.-C. and J.-M. Dufour (2005). Exact Multivariate Tests of Asset Pricing Models with Stable Asymmetric Distributions. Working Paper at CIRANO.

[7] Borwein J.M. and A.S. Lewis (1991). Duality Relationships for Entropy-Like Minimization Problems. *SIAM Journal of Control and Optimization*, **29**, 2, 325-338.

[8] Breeden D. (1979). An Intertemporal Asset Pricing Model with Stochastic Consumption and Investment Opportunities. *Journal of Financial Economics*, **7**, 3, 265-296.

[9] Brown B.W., and W.K. Newey (1998). Efficient Semiparametric Estimation of Expectations. *Econometrica*, **66**, 453-464.

[10] 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.

[11] Campbell J. and R. Shiller (1989). The Dividend-Price Ratio and Expectations of Future Dividends and Discount factors. *Review of Financial Studies*, **1**, 3, 195-228.

[12] Carr P. and L. Wu (2003). The Finite Moment Log Stable Process and Option Pricing. *Journal of Finance*, **58**, 12, 753-778.

[13] Chamberlein G. and M. Rothschild (1983). Arbitrage, Factor Structure, and Mean-Variance Analysis on Large Asset Markets. *Econometrica*, **51**, 5, 1281-1304.

[14] Chen X., J. Favilukis, and S. Ludvigson (2008). An Estimation of Economic Models with Recursive Preferences. Working Paper, at Yale, LSE, and NYU.

[15] Chen X., H. Hong, and M. Shum (2007). Nonparametric Likelihood Ratio Model Selection Tests Between Parametric Likelihood and Moment Condition Models. *Journal of Econometrics*, **141**, 109-140.

[16] Chen X., and S. Ludvigson (2008). Land of Addicts? An Empirical Investigation of Habit-Based Asset Pricing Models. forthcoming at *Journal of Applied Econometrics*.

[17] Cochrane J.H. (2000). *Asset Pricing*, Princeton University Press.

[18] Corcoran S.A. (1998). Bartlett Adjustment of Empirical Discrepancy Statistics. *Biometrika*, **85**, 967-972.

[19] Cressie N., and T.R.C. Read (1984). Multinomial Goodness-of-fit Tests. *Journal of the Royal Statistics Society Series B*, **46**, 440-464.

[20] Csiszàr I. (1975). I-Divergence Geometry of Probability Distributions and Minimization Problems. *The Annals of Probability*, **3**, 1, 146-158.

[21] Csiszàr I. (1991). Why Least Squares and Maximum Entropy? An Axiomatic Approach to Inference for Linear Inverse Problems. *The Annals of Statistics*, **19**, 4, 2032-2066.

[22] Dittmar R. (2002). Nonlinear Pricing Kernels, Kurtosis Preference, and the Cross-Section of Assets Returns. *Journal of Finance*, **57**, 1, 369-403.

[23] 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.

[24] Hansen L.P. (1982). Large Sample Properties of Generalized Method of Moments Estimators. *Econometrica*, **50**, 4, 1029-1054.

[25] 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.

[26] Hansen L.P., and R. Jagannathan (1997). Assessing Specification Errors in Stochastic Discount Factor Models. *Journal of Finance*, **52**, 2, 557-589.

[27] Hansen L.P., and K. Singleton (1982). Generalized Instrumental Variables Estimation of Nonlinear rational Expectations Models. *Econometrica*, **50**, 5, 1269-1286.

[28] Harrison M. and D. Kreps (1979). Martingales and Arbitrage in Multiperiod Security Markets. *Journal of Economic Theory*, **20**, 381-408.

[29] Harvey C. and A. Siddique (2000). Conditional Skewness in Asset Pricing Tests. *Journal of Finance*, **55**, 3, 1263-1295.

[30] Huang C.F. and R. Litzemberger (1988). Foundations for Financial Economics. *North-Holland*.

[31] Hodrick R. and X. Zhang (2001). Evaluationg the Specification Errors of Asset Pricing Models. *Journal of Financial Economics*, **62**, 327-376.

[32] Hong H., B. Preston, and M. Shum (2003). Generalized Empirical Likelihood-Based Model Selection Criteria for Moment Condition Models. *Econometric Theory*, **19**, 923-943.

[33] Ibragimov R. (2004). On the Robustness of Economic Models to Heavy-tailedness Assumptions. Working Paper, Yale University.

[34] Imbens G. (1997). One-Step Estimators for Over-Identified Generalized Method of Moments Models. *Review of Economic Studies*, **64**, 359-383.

[35] Imbens G., R.H. Spady, and P. Johnson (1998). Information Theoretic Approaches to Inference in Moment Condition Models. *Econometrica*, **66**, 2, 333-357.

[36] Jagannathan R. and Z. Wang (1996). The Conditional CAPM and the Cross-Section of Expected Returns. *Journal of Finance*, **51**, 1, 3-53.

[37] Kan R. and C. Robotti (2008). Model Comparison Using the Hansen-Jagannathan Distance. forthcoming at *Review of Financial Studies*.

[38] Kitamura Y. (2001). Asymptotic Optimality of Empirical Likelihood for Testing Moment Restrictions. *Econometrica*, **69**, 1661-1672.

[39] Kitamura Y. (2000). Comparing Misspecified Dynamic Econometric Models Using Nonparametric Likelihood. Working Paper, University of Pennsylvania.

[40] Kitamura Y. (2006a). A Likelihood-Based Approach to the Analysis of a Class of Nested and Non-Nested Models. Working Paper, Yale University.

[41] Kitamura Y. (2006). Empirical Likelihood Methods in Econometrics: Theory and Practice. Working Paper, Yale University.

[42] Kitamura Y. and M. Stutzer (1997). An Information-theoretic Alternative to Generalized Method of Moment Estimation. *Econometrica*, **65**, 861-874.

[43] Kitamura Y. and M. Stutzer (2002). Connections Between Entropic and Linear Porjections in Asset Pricing Estimation. *Journal of Econometrics*, **107**, 159-174.

[44] Kraus A. and R.H. Litzenberger (1976). Skewness Preference and teh Valuation of Risk Assets. *Journal of Finance*, **31**, 4, 1085-1100.

[45] Lucas R. (1978). Asset Prices in an Exchange Economy. *Econometrica*, **46**, 6, 1429-45.

[46] Luemberger D. (1969). *Optimization by Vector Space Methods*. John Wiley & Sons.

[47] Madan D., P. Carr, and E.C. Chang (1998). The Variance Gamma Process and Option Pricing. *European Finance Review*, **2**, 79-105.

[48] Mandelbrot B. (1963). The Variation of Certain Speculative Prices. *Journal of Business*, **26**, 394-419.

[49] Mehra R., and E. Prescott (1985). The Equity Premium: A Puzzle. *Journal of Monetary Economics*, **15**, 145-161.

[50] Newey W. and R. Smith (2004). Higher Order Properties of GMM and Generalized Empirical Likelihood Estimators. *Econometrica*, **72**, 219-255.

[51] Owen A. B. (1988). Empirical Likelihood Ratio Confidence Intervals for A Single Functional. *Biometrika*, **75**, 2, 237-249.

[52] Ramalho J.J.S., and R. J. Smith (2002). Generalized Empirical Likelihood Non-nested Tests. *Journal of Econometrics*, **107**, 99-125.

[53] Royden H.L. (1988). *Real Analysis*, Prentice-Hall.

[54] Rubinstein M. (1973). The Fundamental Theorem of Parameter-Preference Security Valuation. *Journal of Financial and Quantitative Analysis*, **8**, 1, 61-9.

[55] Smith R.J. (1997). Alternative Semi-parametric Likelihood Approaches to Generalized Method of Moments Estimation. *Economic Journal*, **107**, 503-510.

[56] Smith R.J. (2004). GEL Criteria for Moment Condition Models. *Working Paper*, University of Warwick.

[57] Smith R.J. (2007). Efficient Information Theoretic Inference for Conditional Moment Restrictions. *Journal of Econometrics*, **138**, 2, 430-460.

[58] Snow K. (1991). Diagnosing Asset Pricing Models Using the Distribution of Asset Returns. *Journal of Finance*, **46**, 955-983.

[59] Stutzer M. (1995). A Bayesian Approach to Diagnosis of Asset Pricing Models. *Journal of Econometrics*, **68**, 367-397.

[60] Tokat Y., S.T. Rachev, and E. Schwartz (2003). The Stable Non-Gaussian Asset Allocation: A Comparison with the Classical Gaussian Approach. *Journal of Economic Dynamics & Control*, **27**, 937-969.

[61] Vanden J.M. (2006). Option Coskewness and Capital Asset Pricing. *Review of Financial Studies*, **19**, 4, 1279-1320.

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

Table 1: Lagrange Multipliers for the CCAPM Under Different CR Discrepancies.
Risk factors are composed by annual returns over the period 1891 to 2004. The Bond and Stock risk factors are represented respectively by a short-term bond and S&P 500 returns as in Campbell and Shiller (1989). Cressie Read estimators solve HARA utility maximization problems whose portfolios are linear combinations of the listed risk factors with an extra linear term including the asset pricing proxy model y (the CCAPM), and where the Lagrange Multipliers are the portfolio weights. A fixed SDF mean equal to 0.98 is adopted.

CCAPM Parameter (θ)	CR Discrepancies					
	Pearson's	**EL**	**Hellinger's**	**ET**	**CUE/HJ**	**CR ($\gamma = 2$)**
$\theta = 1$						
Bond	0.5050	1.0665	1.1658	1.1968	1.0917	0.7201
S&P	-1.1105	-1.4252	-1.4991	-1.477	-1.3828	-0.8419
$\theta = 5$						
Bond	0.8363	1.2279	1.3527	1.4005	1.3070	0.8625
S&P	-1.0330	-1.4109	-1.4846	-1.468	-1.3703	-0.8419
$\theta = 20$						
Bond	0.8363	1.4414	1.5735	1.6290	1.5544	0.8625
S&P	-1.0330	-1.3165	-1.3741	-1.365	-1.2690	-0.8419
$\theta = 50$						
Bond	0.8027	0.1691	0.1105	0.0373	-0.1245	0.5986
S&P	-1.0295	-1.2847	-1.3274	-1.3312	-1.2265	-0.8482

Table 2: Correlation Between Implied SDFs and the S&P 500 Returns Under Different CR Discrepancies.
Risk factors are composed by annual returns over the period 1891 to 2004. The Bond and Stock risk factors are represented respectively by a short-term bond and S&P 500 returns as in Campbell and Shiller (1989). Cressie Read SDFs are obtained from the first-order condition of HARA utility maximization problems whose portfolios are linear combinations of the listed risk factors with an extra linear term including the asset pricing proxy model y (the CCAPM). A fixed SDF mean is set equal to 0.98.

CCAPM Parameter (θ)	CR Discrepancies						
	Pearson's	**EL**	**Hellinger's**	**ET**	**CUE/HJ**	**CR ($\gamma = 2$)**	**CCAPM**
1	-0.8045	-0.8804	-0.9197	-0.9443	-0.9656	-0.9413	-0.0318
5	-0.6617	-0.7593	-0.7858	-0.8013	-0.8133	-0.7065	-0.0385
20	-0.3104	-0.3335	-0.3312	-0.3280	-0.3213	-0.2597	-0.0413
50	-0.0978	-0.1111	-0.1095	-0.1079	-0.1050	-0.0793	-0.0071

Table 3: Measuring Misspecification of the CCAPM via Different CR Discrepancies.
Risk factors are composed by annual returns over the period 1891 to 2004. The Bond and Stock risk factors are represented respectively by a short-term bond and S&P 500 returns as in Campbell and Shiller (1989). Cressie Read estimators solve HARA utility maximization problems whose portfolios are linear combinations of the listed risk factors with an extra linear term including the asset pricing proxy model y (the CCAPM). A fixed SDF mean is set equal to 0.98.

CCAPM Parameter (θ)	CR Discrepancies					
	Pearson's	**EL**	**Hellinger's**	**ET**	**CUE/HJ**	**CR ($\gamma = 2$)**
1	0.0343	0.0377	0.0382	0.0380	0.0356	0.0291
5	0.0333	0.0371	0.0376	0.0374	0.0351	0.0288
20	0.0295	0.0323	0.0327	0.0325	0.0306	0.0259
50	**0.0277**	**0.0313**	**0.0316**	**0.0315**	**0.0299**	**0.0244**

Table 4: Specification Errors and Lagrange Multipliers (Weights) for the CCAPM Under a CR Discrepancy with High γ.

Risk factors are composed by annual returns over the period 1891 to 2004. The Bond and Stock risk factors are represented respectively by a short-term bond and S&P 500 returns as in Campbell and Shiller (1989). The Cressie Read estimator solves a HARA utility maximization problems whose portfolios are linear combinations of the listed risk factors with an extra linear term including the asset pricing proxy model y (the CCAPM). A fixed SDF mean is set equal to 0.98.

CCAPM Parameter (θ)	Discrepancy —— Lagrange Multipliers (weights)		
	CR ($\gamma = 5$)	**LM Bond**	**LM S&P**
1	0.0343	0.2765	-0.3207
5	0.0350	0.3515	-0.3409
20	**0.0335**	0.6471	-0.3207
50	0.0341	-0.4647	-0.3207

Table 5: Pricing Errors for Different CR Discrepancies under the Additive Estimator.

The pricing error for the implied SDFs is defined by $E_\mu((1+R)m_{CR} - 1)$, where R represents a gross return (either short-term bond or S&P 500) and m_{CR} represents the admissible SDF implied by a certain Cressie Read discrepancy. The pricing error for the CCAPM is defined by $E_\mu((1+R)(\frac{C_{t+1}}{C_t})^{-\theta} - 1)$, where C_t represents aggregate consumption at time t. Risk factors are composed by annual returns over the period 1891 to 2004. The Bond and Stock risk factors are represented respectively by a short-term bond and S&P 500 returns as in Campbell and Shiller (1989). Cressie Read estimators solve HARA utility maximization problems whose portfolios are linear combinations of the listed risk factors with an extra linear term including the asset pricing proxy model y (the CCAPM), and where the Lagrange Multipliers are the portfolio weights. A fixed SDF mean is set equal to 0.98.

CCAPM Parameter	CCAPM	Pearson's	CR ($\gamma = 2$)	CR ($\gamma = 5$)
$\theta = 1$				
Bond	-0.0001	-0.0013	0.0002	-0.0001
S&P	0.0514	0.0036	0.0167	0.0376
$\theta = 5$				
Bond	-0.0009	0.0002	-0.0001	-0.0008
S&P	0.0503	0.0087	0.0159	0.0366
$\theta = 20$				
Bond	-0.0020	-0.0009	-0.0012	-0.0004
S&P	0.0457	0.0042	0.0110	0.0325
$\theta = 50$				
Bond	0.0035	0.0043	0.0037	0.0005
S&P	0.0483	0.0071	0.0128	0.0314

Table 6: Pricing Errors for Different CR Discrepancies under the Multiplicative Estimator.

The pricing error for the implied SDFs is defined by $E_\mu((1+R)m_{CR}-1)$, where R represents a gross return (either short-term bond or S&P 500) and m_{CR} represents the admissible SDF implied by a certain Cressie Read discrepancy. Risk factors are composed by annual returns over the period 1891 to 2004. The Bond and Stock risk factors are represented respectively by a short-term bond and S&P 500 returns as in Campbell and Shiller (1989). Cressie Read estimators solve HARA utility maximization problems whose portfolios are linear combinations of the listed risk factors returns weighted by the asset pricing proxy model y (the CCAPM), and where the Lagrange Multipliers are the portfolio weights. A fixed SDF mean is set equal to 0.98.

CCAPM Parameter	Pearson's	CR ($\gamma = 2$)	CR ($\gamma = 5$)
$\theta = 1$			
Bond	0.0000	-0.0005	0.0010
S&P	0.0000	0.0091	0.0305
$\theta = 5$			
Bond	0.0015	-0.0019	0.0009
S&P	0.0045	0.0017	0.0278
$\theta = 20$			
Bond	0.0000	-0.0046	0.0030
S&P	0.0000	0.0157	0.0340
$\theta = 50$			
Bond	0.0000	-0.0037	0.0128
S&P	0.0000	-0.0025	0.0262
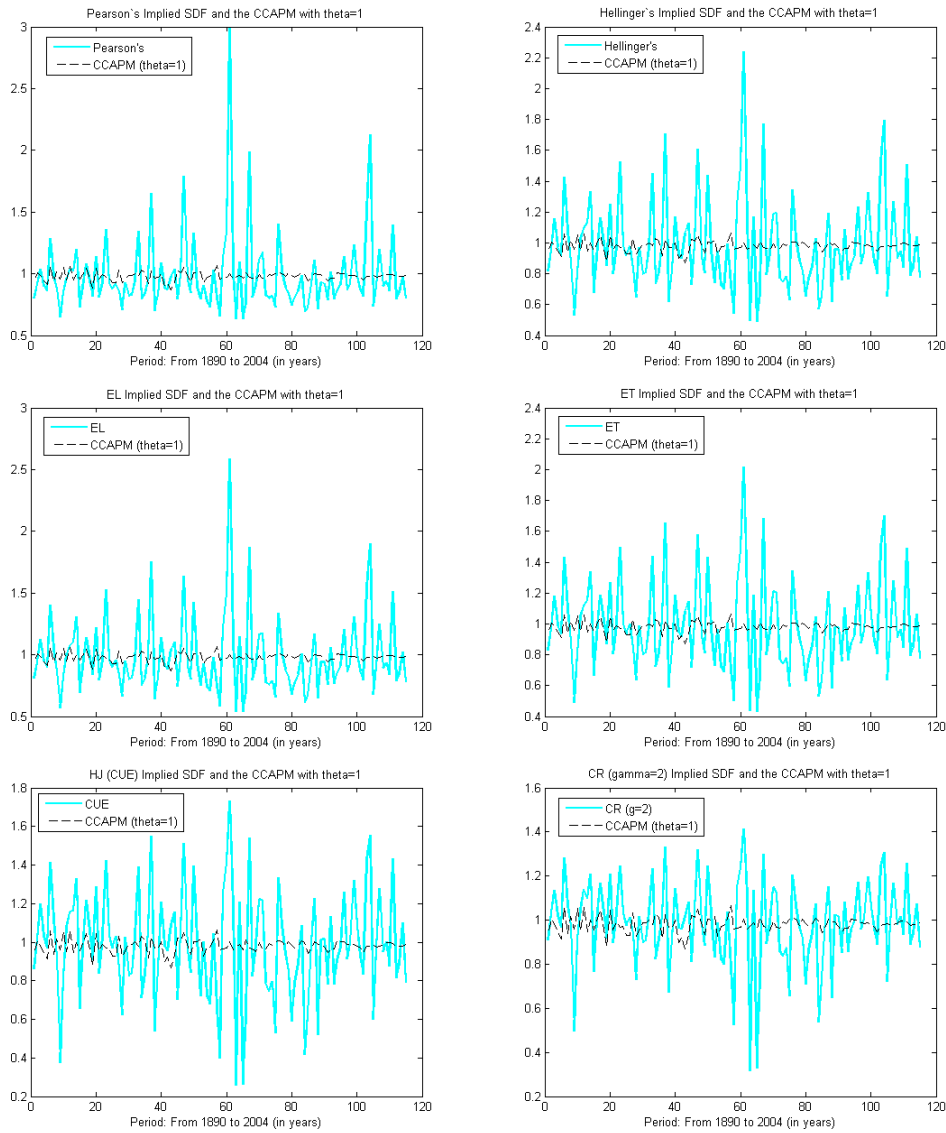
Figure 1: Admissible SDFs and the CCAPM with $\theta = 1$

This picture presents Admissible SDFs under different CR discrepancies and the CCAPM SDF with a risk aversion coefficient (θ) of equal to 1. The MD problems are solved based on annual returns on S&P 500 and a short-term bond over the period 1890 to 2004. The CCAPM SDF is based on consumption groth data over the period 1890 to 2004. All SDF means are fixed at 0.98.
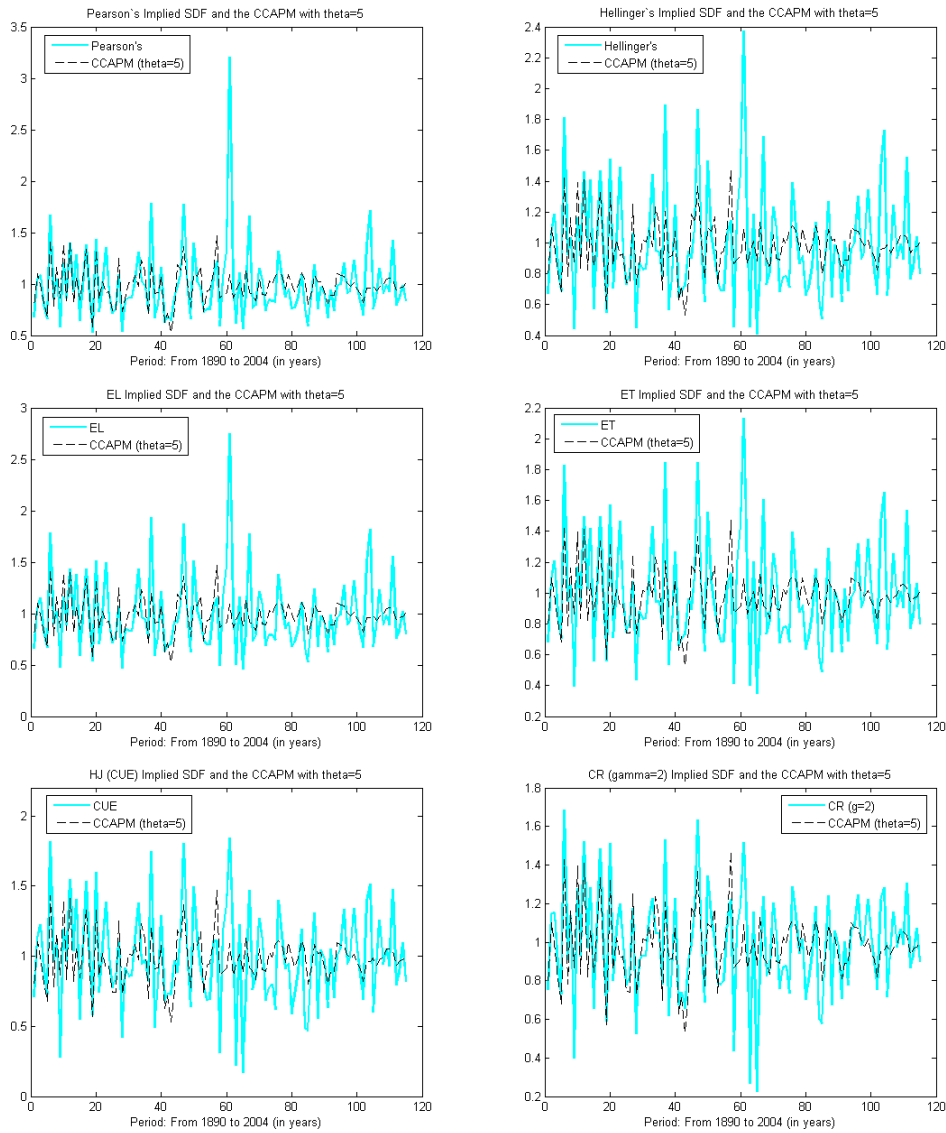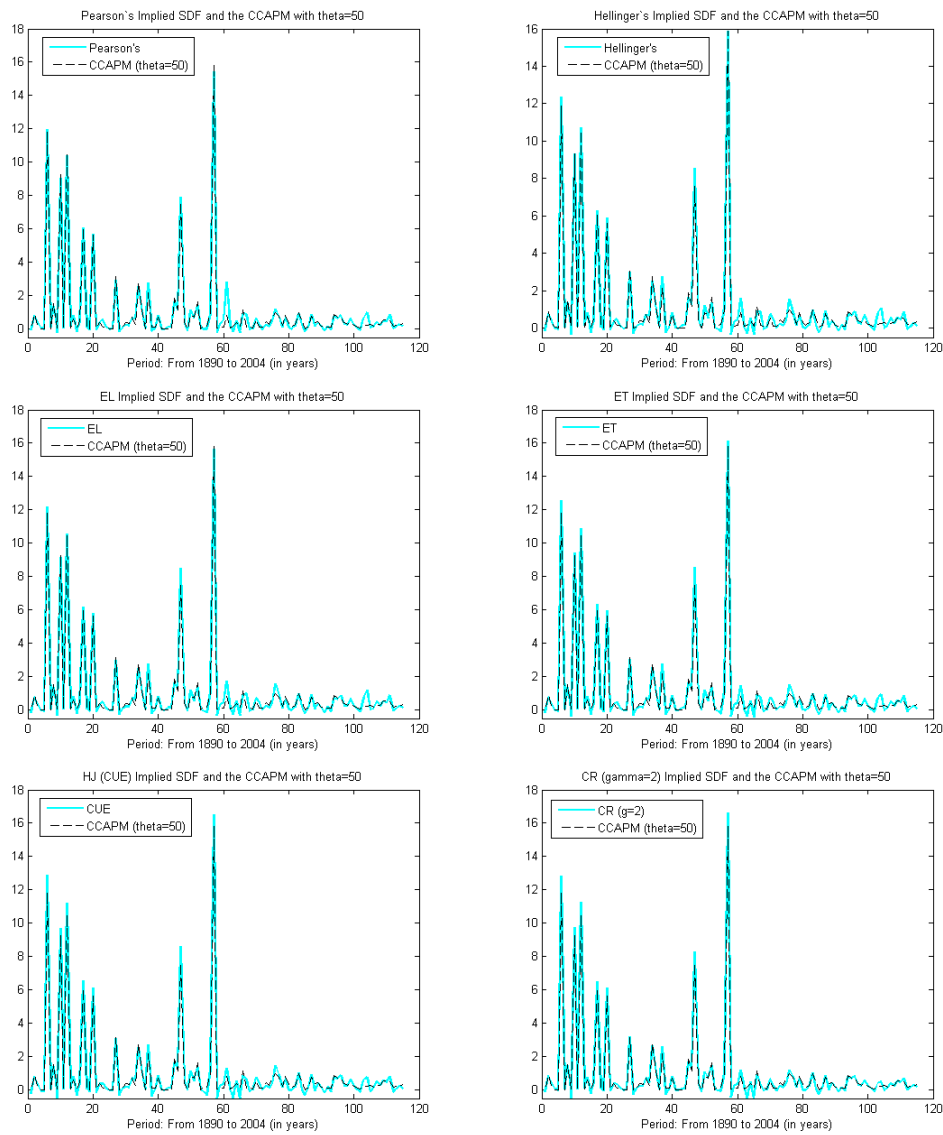
Figure 2: Admissible SDFs and the CCAPM with $\theta = 5$

This picture presents Admissible SDFs under different CR discrepancies and the CCAPM SDF with a risk aversion coefficient (θ) of equal to 5. The MD problems are solved based on annual returns on S&P 500 and a short-term bond over the period 1890 to 2004. The CCAPM SDF is based on consumption groth data over the period 1890 to 2004. All SDF means are fixed at 0.98.
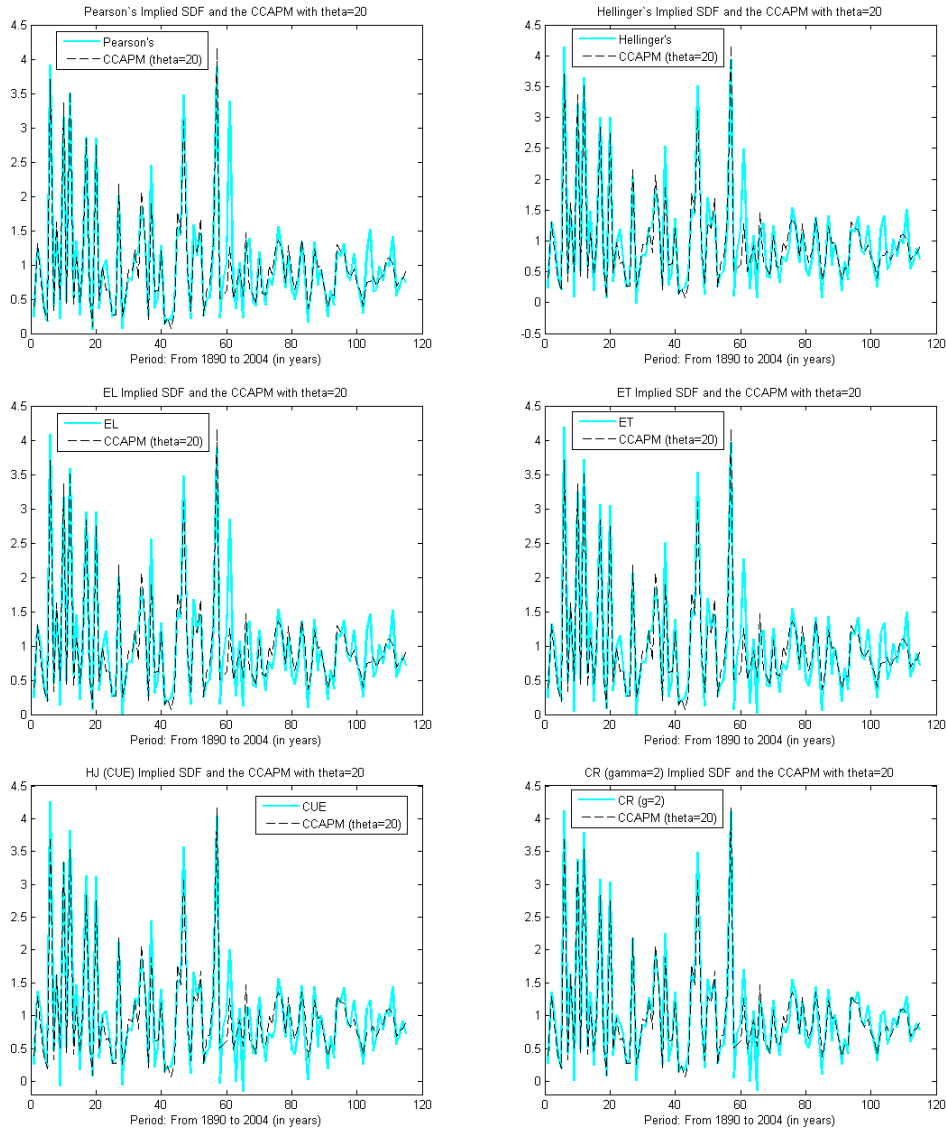
Figure 3: Admissible SDFs and the CCAPM with $\theta = 20$
This picture presents Admissible SDFs under different CR discrepancies and the CCAPM SDF with a risk aversion coefficient (θ) of equal to 20. The MD problems are solved based on annual returns on S&P 500 and a short-term bond over the period 1890 to 2004. The CCAPM SDF is based on consumption groth data over the period 1890 to 2004. All SDF means are fixed at 0.98.

Figure 4: Admissible SDFs and the CCAPM with $\theta = 50$

This picture presents Admissible SDFs under different CR discrepancies and the CCAPM SDF with a risk aversion coefficient (θ) of equal to 50. The MD problems are solved based on annual returns on S&P 500 and a short-term bond over the period 1890 to 2004. The CCAPM SDF is based on consumption groth data over the period 1890 to 2004. All SDF means are fixed at 0.98.
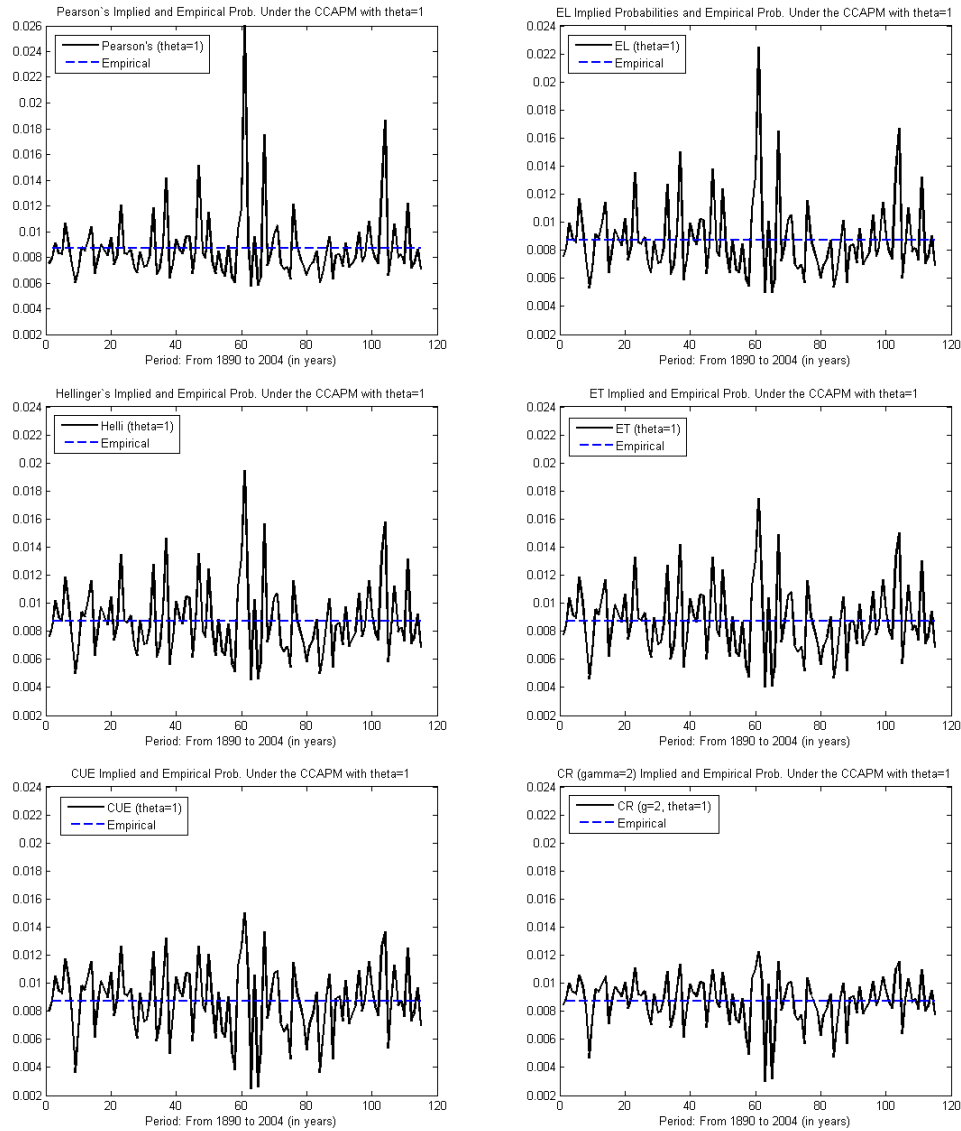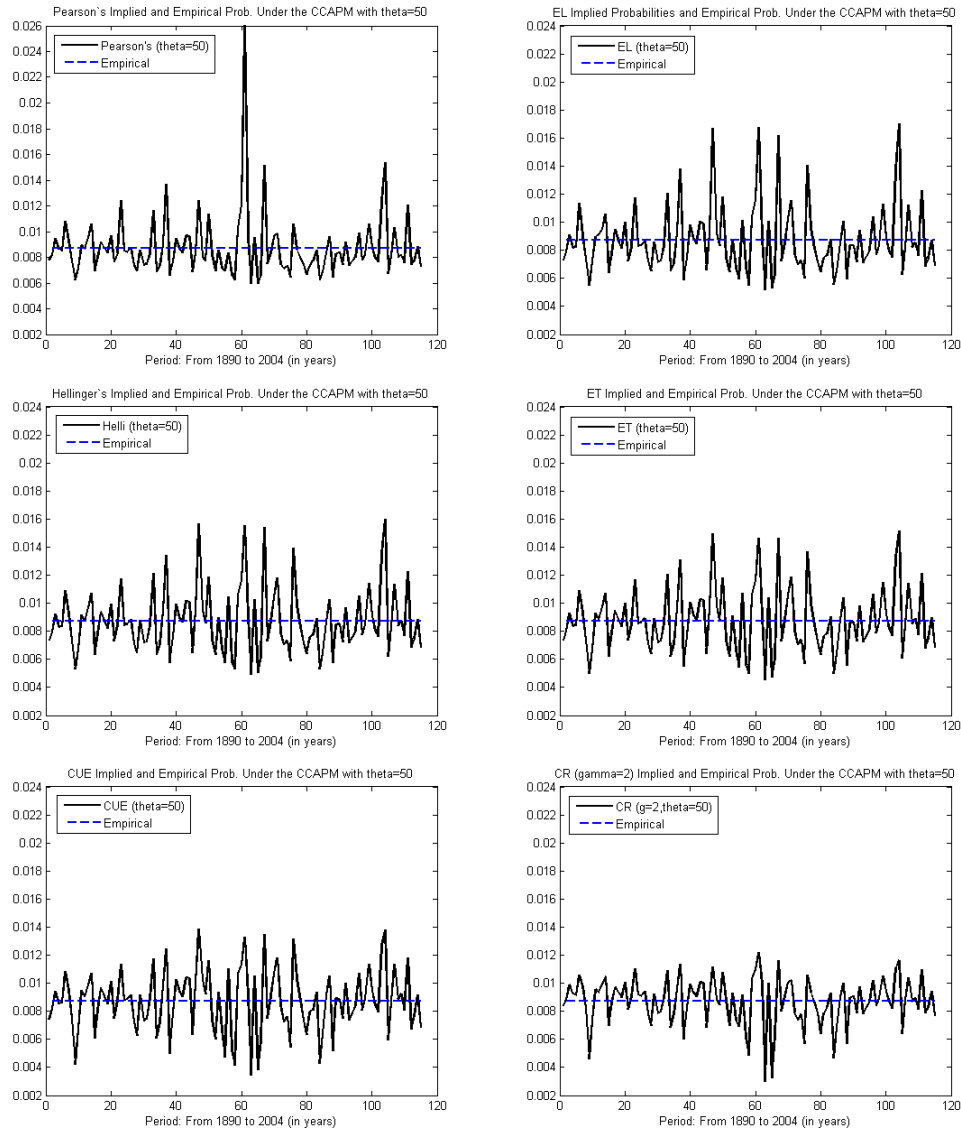
Figure 5: Implied Probabilities under the CCAPM with $\theta = 1$
This picture presents implied probabilities for different discrepancy measures from the Cressie Read family, when measuring the degree of misspecification of the CCAPM asset pricing model for a fixed riskl aversion coefficient $\theta = 1$. Results are based on annual returns from S&P 500 and a short-term bond over the period 1890 to 2004. The implied probabilities are obtained by solving the dual optimization problems that have a portfolio interpretation in terms of maximization of a HARA utility function. Implied probabilities are hyperbolic functions of the returns.
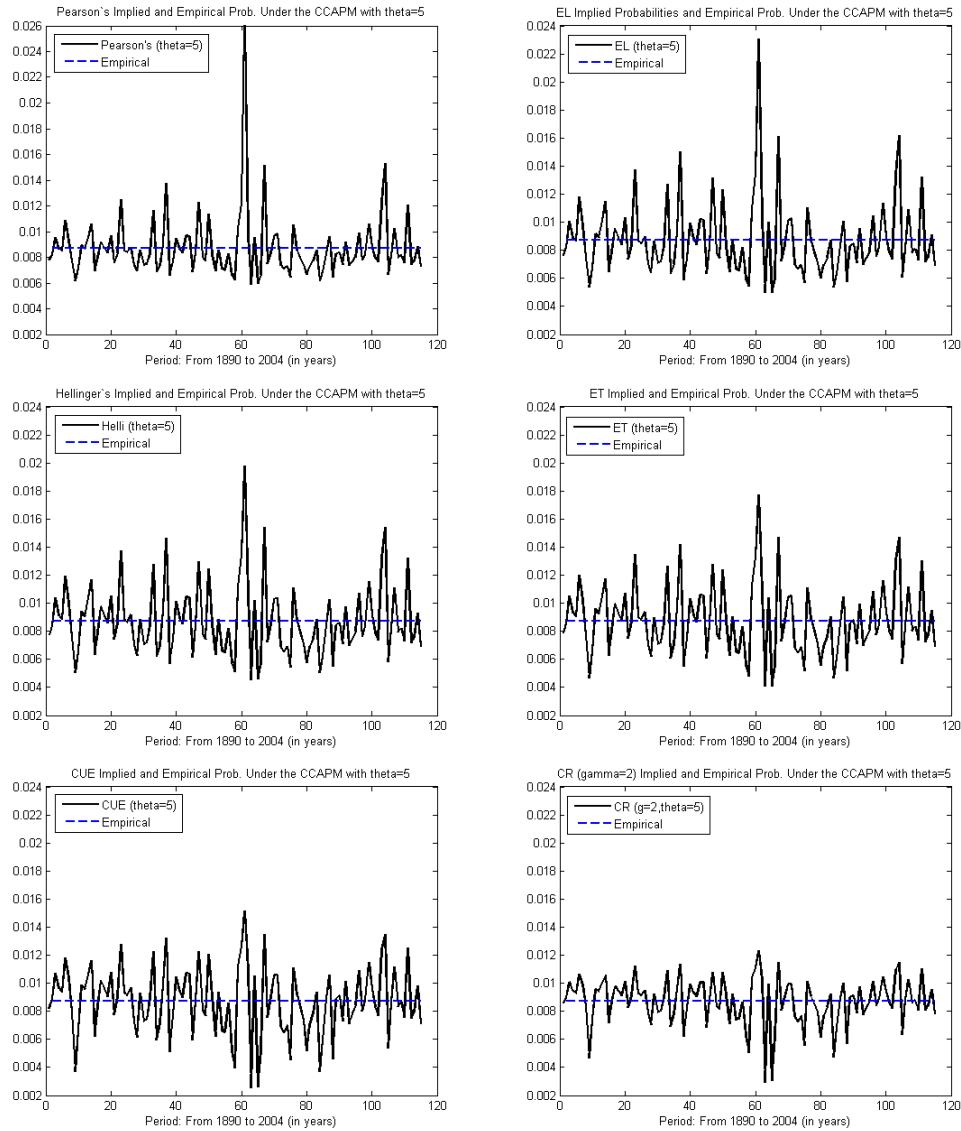
Figure 6: Implied Probabilities under the CCAPM with $\theta = 5$

This picture presents implied probabilities for different discrepancy measures from the Cressie Read family, when measuring the degree of misspecification of the CCAPM asset pricing model for a fixed riskl aversion coefficient $\theta = 5$. Results are based on annual returns from S&P 500 and a short-term bond over the period 1890 to 2004. The implied probabilities are obtained by solving the dual optimization problems that have a portfolio interpretation in terms of maximization of a HARA utility function. Implied probabilities are hyperbolic functions of the returns.
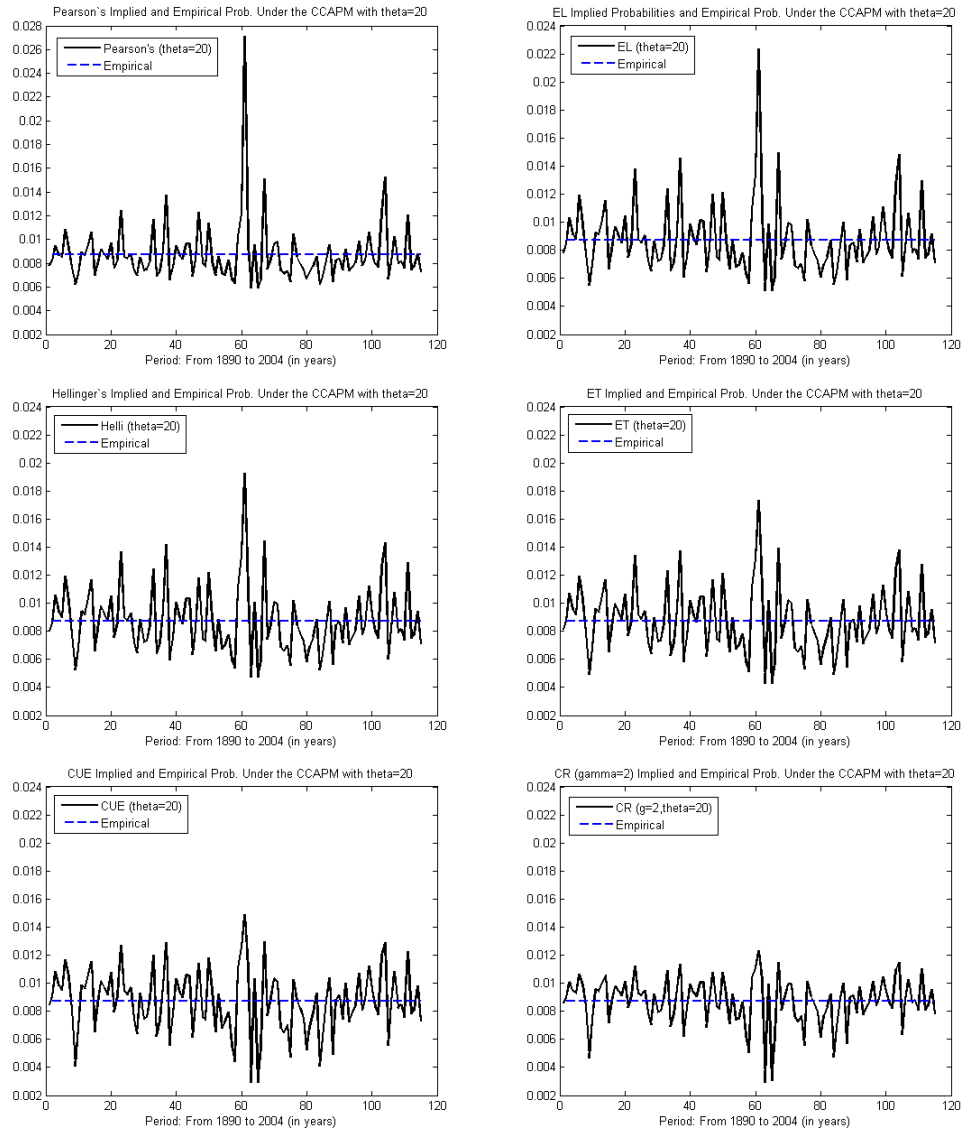
Figure 7: Implied Probabilities under the CCAPM with $\theta = 20$
This picture presents implied probabilities for different discrepancy measures from the Cressie Read family, when measuring the degree of misspecification of the CCAPM asset pricing model for a fixed riskl aversion coefficient $\theta = 20$. Results are based on annual returns from S&P 500 and a short-term bond over the period 1890 to 2004. The implied probabilities are obtained by solving the dual optimization problems that have a portfolio interpretation in terms of maximization of a HARA utility function. Implied probabilities are hyperbolic functions of the returns.

Figure 8: Implied Probabilities under the CCAPM with $\theta = 50$
This picture presents implied probabilities for different discrepancy measures from the Cressie Read family, when measuring the degree of misspecification of the CCAPM asset pricing model for a fixed riskl aversion coefficient $\theta = 50$. Results are based on annual returns from S&P 500 and a short-term bond over the period 1890 to 2004. The implied probabilities are obtained by solving the dual optimization problems that have a portfolio interpretation in terms of maximization of a HARA utility function. Implied probabilities are hyperbolic functions of the returns.
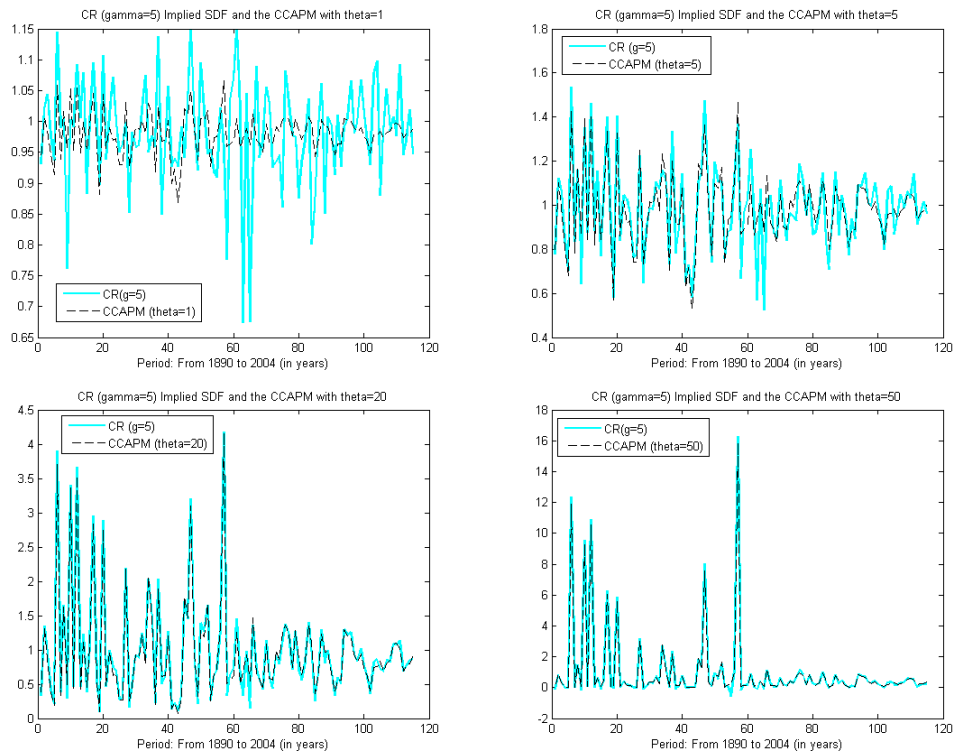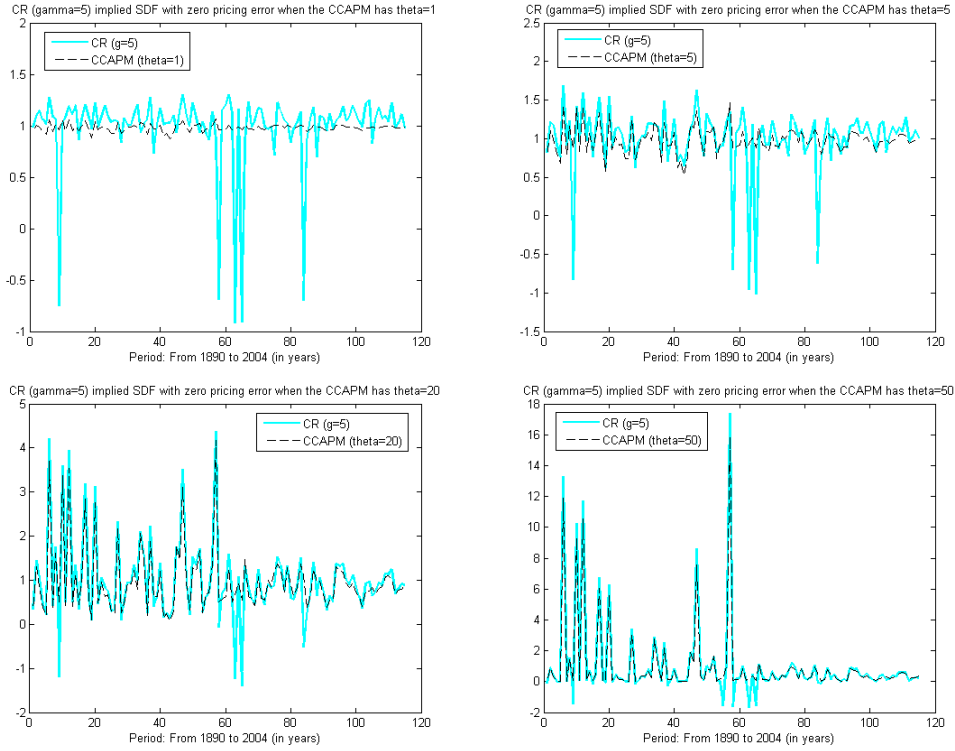
Figure 9: CR ($\gamma = 5$) SDFs and the CCAPM

This picture presents, for different regions of the parametric space of the CCAPM model, SDFs that are as close as possible to admissible ones under the CR ($\gamma = 5$) discrepancy criterion. The CR problem is solved based on annual returns on S&P 500 and a short-term bond over the period 1890 to 2004. The CCAPM SDF is based on consumption groth data over the period 1890 to 2004. All SDF means are equal to 0.98.
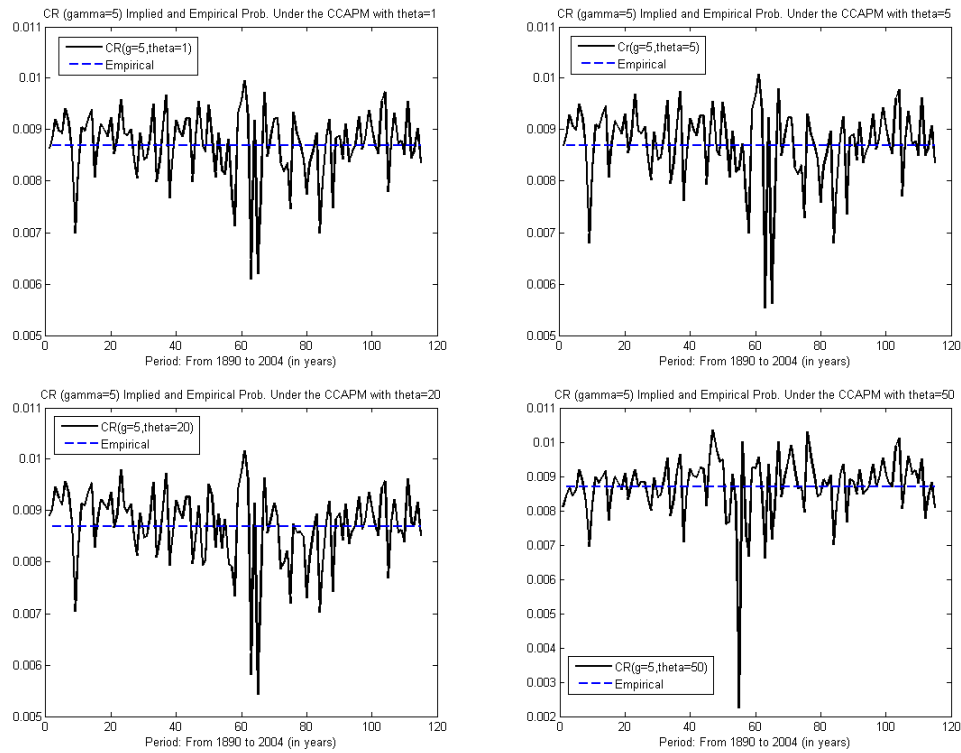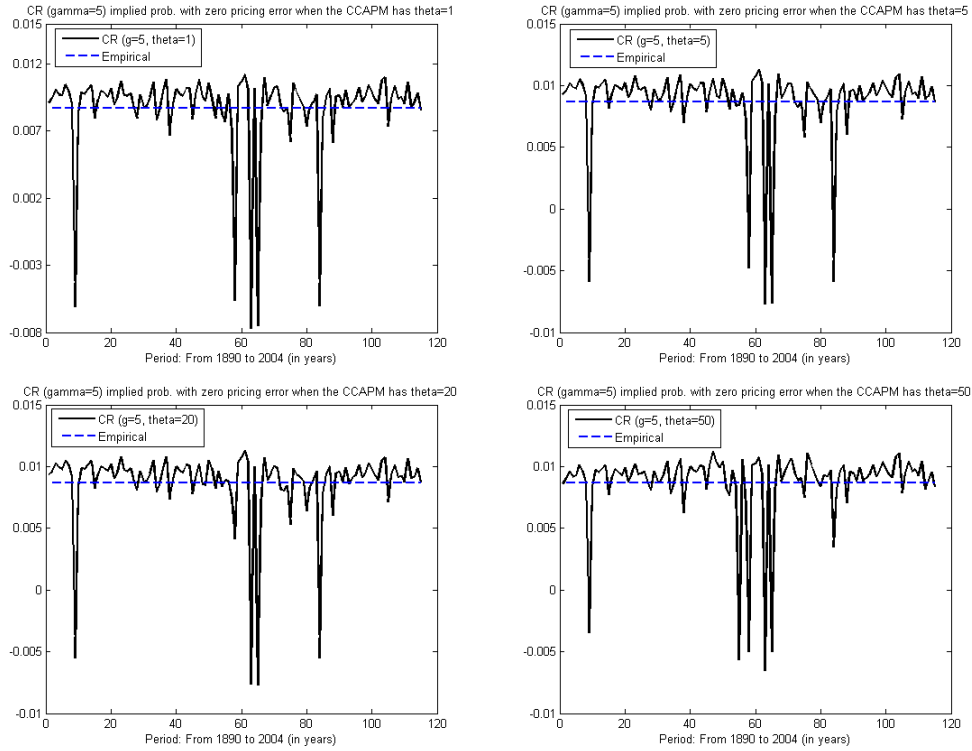
Figure 10: CR ($\gamma = 5$) Implied Probabilities under the CCAPM
This picture presents, for different regions of the parametric space of the CCAPM model, implied probabilities for the Cressie Read discrepancy measure with a coefficient $\gamma = 5$. Results are based on annual returns from S&P 500 and a short-term bond over the period 1890 to 2004. The implied probabilities are obtained by solving the dual optimization problems that have a portfolio interpretation in terms of maximization of a HARA utility function. Implied probabilities are a hyperbolic function of the returns.

Figure 11: CR ($\gamma = 5$) SDFs with Zero Pricing Errors and the CCAPM
This picture presents, for different regions of the parametric space of the CCAPM model, admissible SDFs under the CR ($\gamma = 5$) discrepancy criterion composed with the absolute value fuction that achieve negative values in some states. The CR problem is solved based on annual returns on S&P and a short-term bond over the period 1890 to 2004. The CCAPM SDF is based on consumption groth data over the period 1890 to 2004. All SDF means are equal to 0.98.

Figure 12: CR ($\gamma = 5$) Implied Probabilities with Zero Pricing Errors under the CCAPM

This picture presents, for different regions of the parametric space of the CCAPM model, implied probabilities for the Cressie Read discrepancy measure with a coefficient $\gamma = 5$ composed with the absolute value fuction. Results are based on annual returns from S&P 500 and a short-term bond over the period 1890 to 2004. The implied probabilities are obtained by solving the dual optimization problems that have a portfolio interpretation in terms of maximization of a HARA utility function composed with the absolute value fuction. Implied probabilities are a hyperbolic function of the returns.
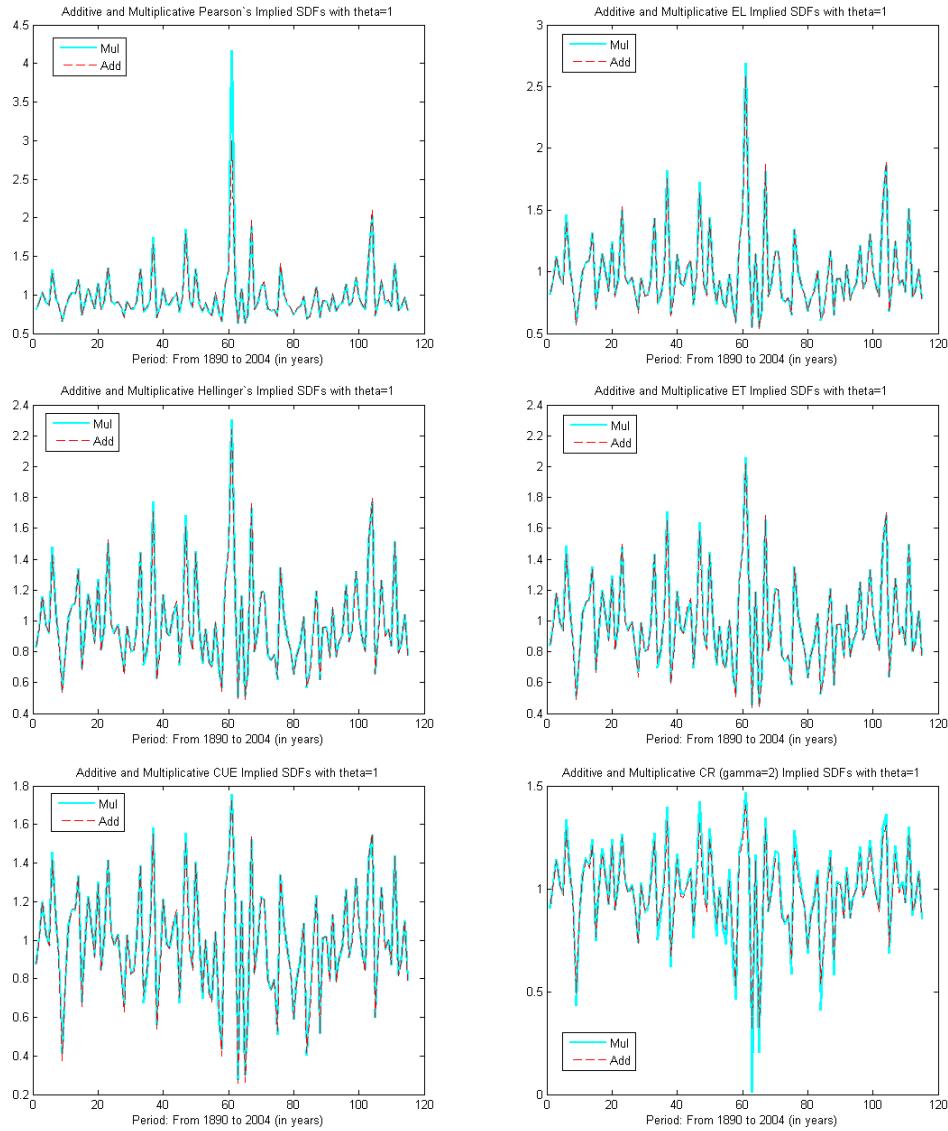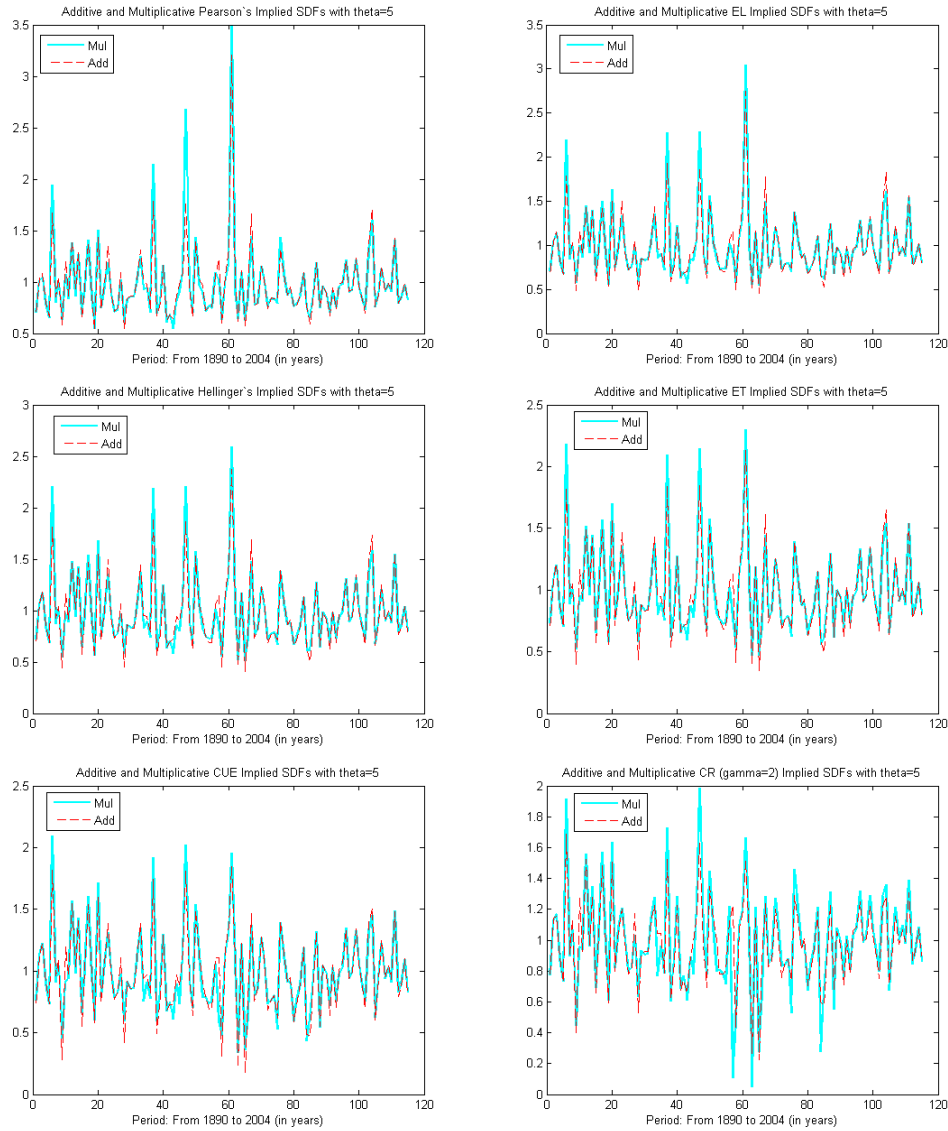
Figure 13: Comparison of Admissible SDFs under the Additive and Multiplicative Cases with $\theta = 1$

This picture presents admissible SDFs for different discrepancy measures from the Cressie Read family, under two formulations (additive and multiplicative) of the misspecification problem of the CCAPM model with a fixed risk aversion coefficient $\theta = 1$. Results are based on annual returns from S&P 500 and a short-term bond over the period 1890 to 2004. The SDFs are obtained by solving the dual optimization problems that have portfolio interpretations in terms of maximization of HARA utility functions.
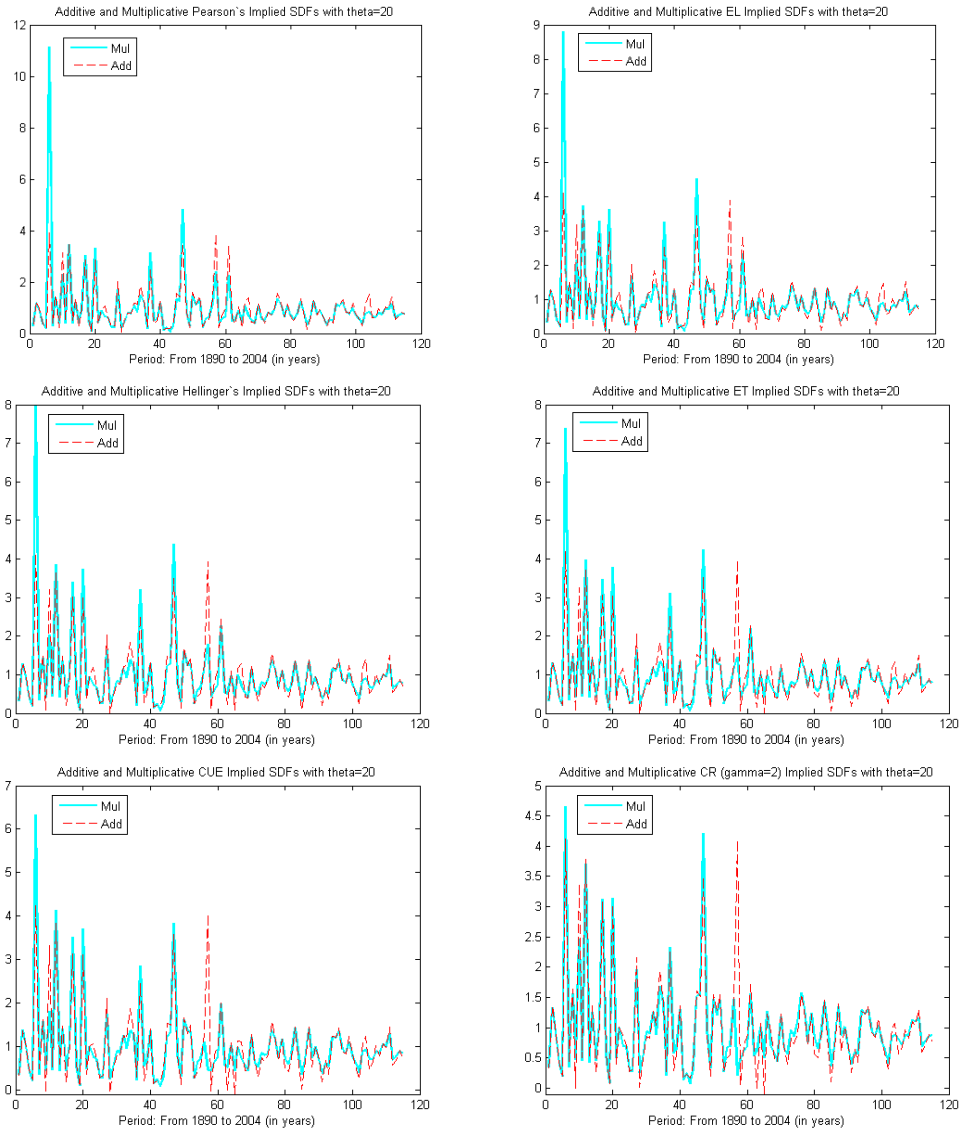
Figure 14: Comparison of Admissible SDFs under the Additive and Multiplicative Cases with $\theta = 5$
This picture presents admissible SDFs for different discrepancy measures from the Cressie Read family, under two formulations (additive and multiplicative) of the misspecification problem of the CCAPM model with a fixed risk aversion coefficient $\theta = 5$. Results are based on annual returns from S&P 500 and a short-term bond over the period 1890 to 2004. The SDFs are obtained by solving the dual optimization problems that have portfolio interpretations in terms of maximization of HARA utility functions.
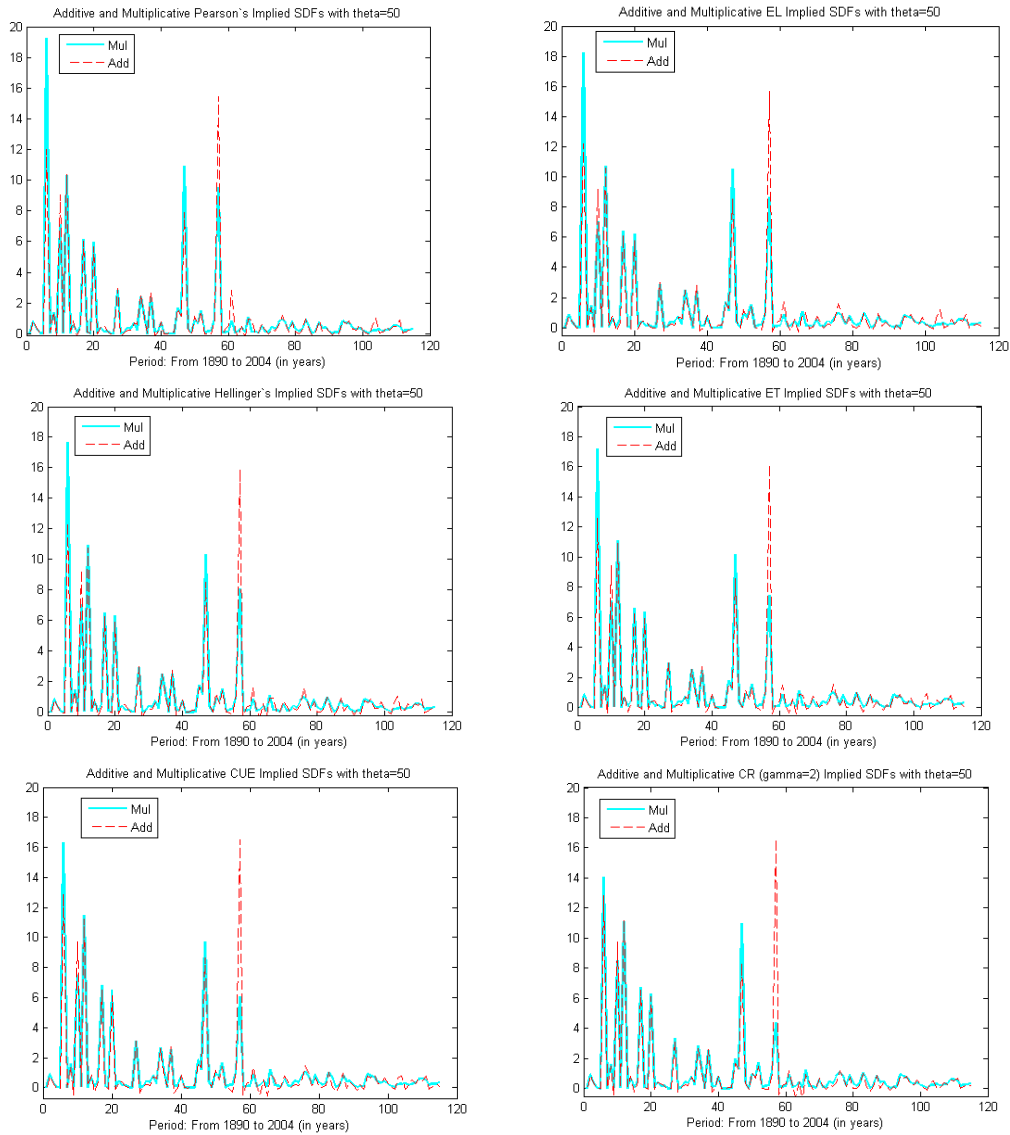
Figure 15: Comparison of Admissible SDFs under the Additive and Multiplicative Cases with $\theta = 20$

This picture presents admissible SDFs for different discrepancy measures from the Cressie Read family, under two formulations (additive and multiplicative) of the misspecification problem of the CCAPM model with a fixed risk aversion coefficient $\theta = 20$. Results are based on annual returns from S&P 500 and a short-term bond over the period 1890 to 2004. The SDFs are obtained by solving the dual optimization problems that have portfolio interpretations in terms of maximization of HARA utility functions.
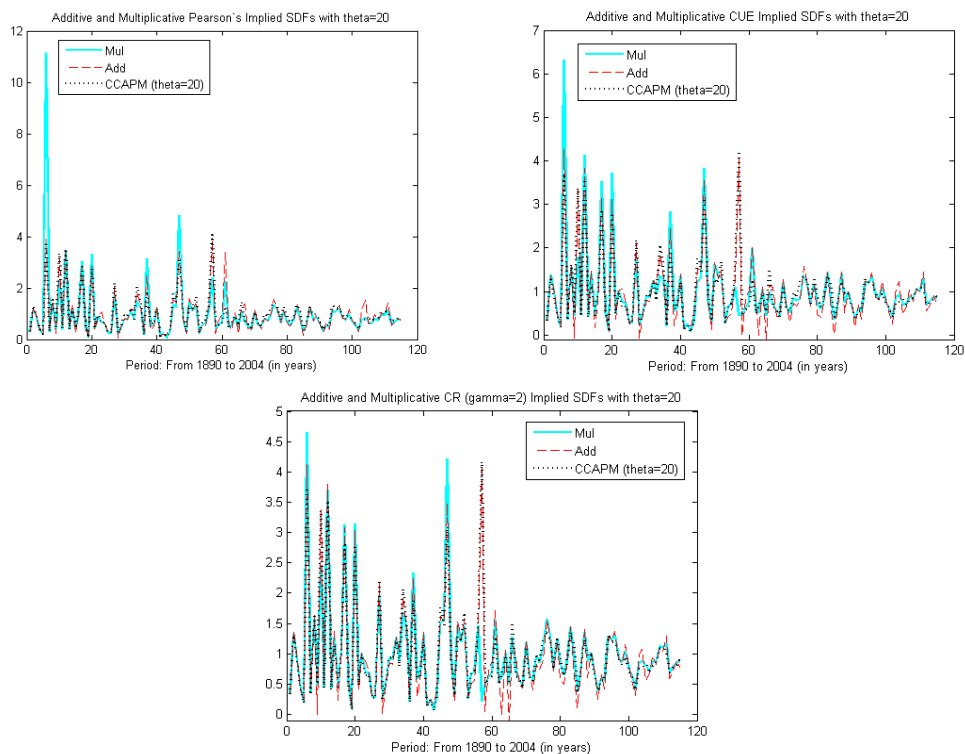
Figure 16: Comparison of Admissible SDFs under the Additive and Multiplicative Cases with $\theta = 50$

This picture presents admissible SDFs for different discrepancy measures from the Cressie Read family, under two formulations (additive and multiplicative) of the misspecification problem of the CCAPM model with a fixed risk aversion coefficient $\theta = 50$. Results are based on annual returns from S&P 500 and a short-term bond over the period 1890 to 2004. The SDFs are obtained by solving the dual optimization problems that have portfolio interpretations in terms of maximization of HARA utility functions.

Figure 17: Comparison of Admissible SDFs under the Additive and Multiplicative Cases, including the CCAPM Proxy

This picture presents the CCAPM with a fixed risk aversion coefficient $\theta = 20$ and the corresponding admissible SDFs for different discrepancy measures from the Cressie Read family, under two formulations (additive and multiplicative). Results are based on annual returns from S&P 500 and a short-term bond over the period 1890 to 2004. The CR SDFs are obtained by solving the dual optimization problems that have portfolio interpretations in terms of maximization of HARA utility functions.
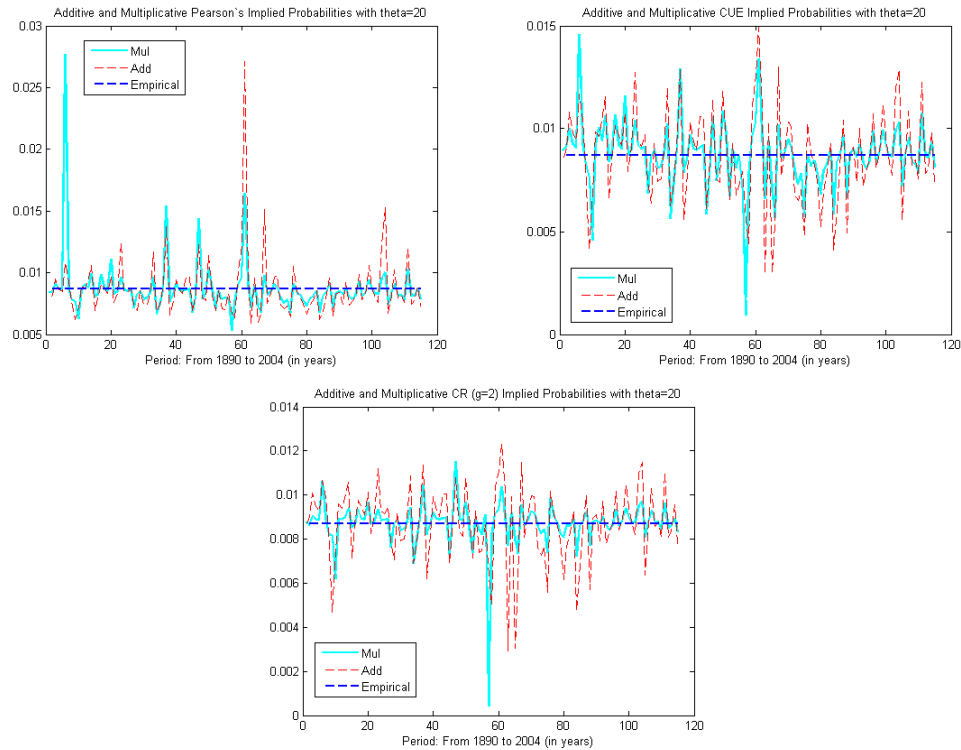
Figure 18: Comparison of Implied Probabilities under the Additive and Multiplicative Cases, including the Empirical Probabilities
This picture presents the implied probabilities obtained when evaluating the CCAPM with a fixed risk aversion coefficient $\theta = 20$ with different discrepancy measures from the Cressie Read family, under two formulations of the estimators (additive and multiplicative). Results are based on annual returns from S&P 500 and a short-term bond over the period 1890 to 2004. The CR implied probabilities are obtained by solving the dual optimization problems that have portfolio interpretations in terms of maximization of HARA utility functions.
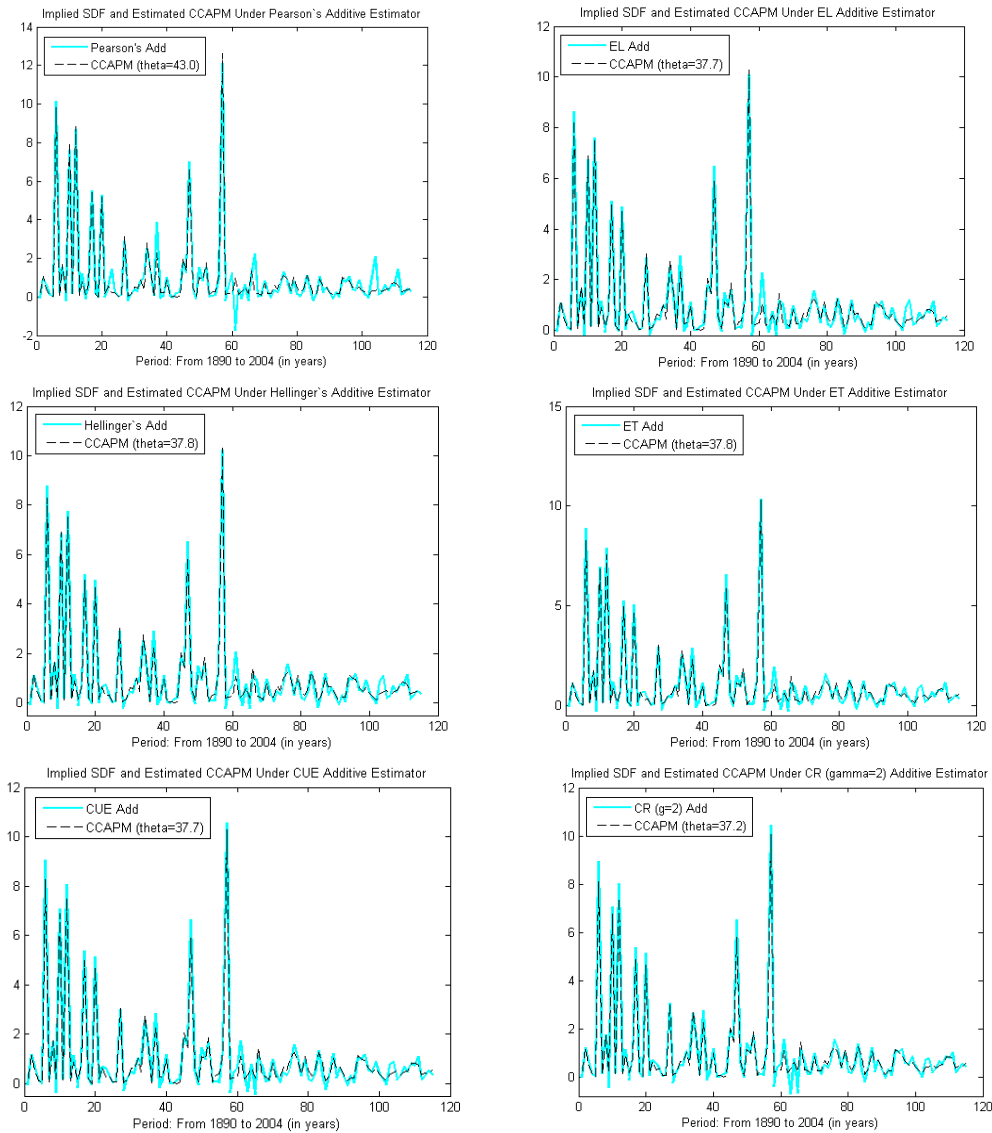
Figure 19: Estimating the CCAPM under Additive Minimum Discrepancy Problems
This picture presents admissible SDFs and corresponding estimated CCAPM SDFs for different discrepancy measures from the Cressie Read family, under additive discrepancy problems. Results are based on annual returns from S&P 500 and a short-term bond over the period 1890 to 2004. The SDFs are obtained by solving a double optimization problem where we search for the parameter theta that minimizes the Cressie Read discrepancy which contains a linear term in the CCAPM asset pricing proxy.
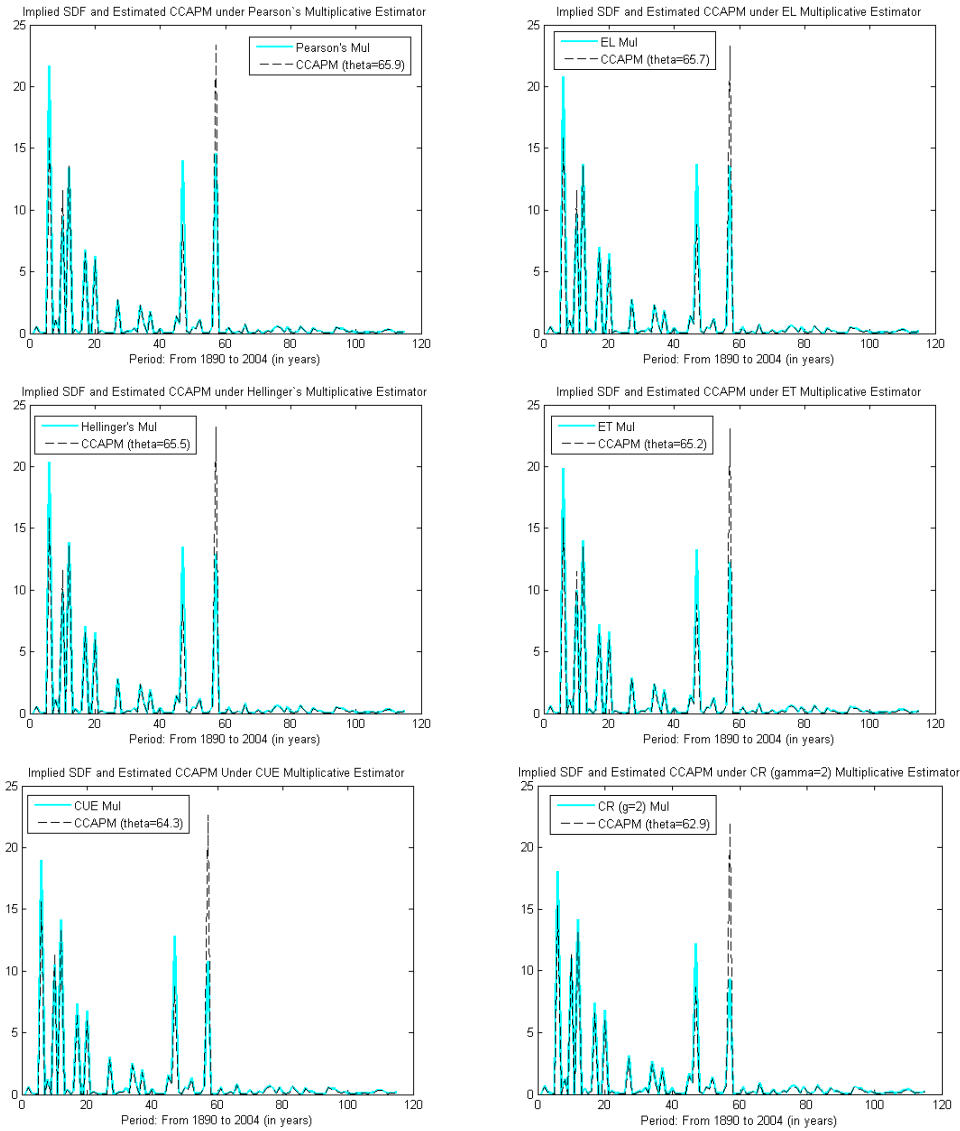
Figure 20: Estimating the CCAPM under Multiplicative Minimum Discrepancy Problems
This picture presents admissible SDFs and corresponding estimated CCAPM SDFs for different discrepancy measures from the Cressie Read family, under multiplicative discrepancy problems. Results are based on annual returns from S&P 500 and a short-term bond over the period 1890 to 2004. The SDFs are obtained by solving a double optimization problem where we search for the parameter theta that minimizes the Cressie Read discrepancy which contains a non-linear term in the CCAPM asset pricing proxy.