Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Portfolio Analysis Using Stochastic Dominance, Relative Entropy, and Empirical Likelihood

Thierry Post, Valerio Poti

informs
http://pubsonline.informs.org/journal/mnsc/

MANAGEMENT SCIENCE
Vol. 63, No. 1, January 2017, pp. 153–165
ISSN 0025-1909 (print), ISSN 1526-5501 (online)

Portfolio Analysis Using Stochastic Dominance, Relative Entropy, and Empirical Likelihood

Thierry Post,[a] **Valerio Potì**[b]

[a] Graduate School of Business, Koç University, 34450 Sarıyer/Istanbul, Turkey; [b] UCD Michael Smurfit Graduate Business School, University College Dublin, Belfield, Dublin 4, Ireland

Contact: thierrypost@hotmail.com (TP); valerio.poti@ucd.ie (VP)

Received: October 17, 2014
Accepted: July 25, 2015
Published Online in Articles in Advance: March 2, 2016

https://doi.org/10.1287/mnsc.2015.2325

Abstract. This study formulates portfolio analysis in terms of stochastic dominance, relative entropy, and empirical likelihood. We define a portfolio inefficiency measure based on the divergence between given probabilities and the nearest probabilities that rationalize a given portfolio for some admissible utility function. When applied to a sample of time-series observations in a blockwise fashion, the inefficiency measure becomes a likelihood ratio statistic for testing inequality moment conditions. The limiting distribution of the test statistic is bounded by a chi-squared distribution under general sampling schemes, allowing for conservative large-sample testing. We develop a tight numerical approximation for the test statistic based on a two-stage optimization procedure and piecewise linearization techniques. A Monte Carlo simulation study of the empirical likelihood ratio test shows superior small-sample properties compared with various generalized method of moments tests. An application analyzes the efficiency of a passive stock market index in data sets from the empirical asset pricing literature.

History: Accepted by Manel Baucells, decision analysis.

Keywords: stochastic dominance • relative entropy • empirical likelihood • convex programming • utility theory • portfolio theory • asset pricing

1. Introduction

A growing body of literature presents mathematical programming problems and statistical inference methods for portfolio analysis based on criteria of stochastic dominance (SD), including Bawa et al. (1985), Shalit and Yitzhaki (1994), Post (2003), Kuosmanen (2004), Levy (2006), Roman et al. (2006), Post and Versijp (2007), Scaillet and Topaloglou (2010), and Linton et al. (2014). These methods have important applications in active portfolio management and empirical asset pricing. For example, Roman et al. (2013) and Hodder et al. (2015) present practical applications to active security selection and asset allocation.

The existing methods generally take the probability distribution as given and search over the admissible utility functions or, equivalently, the lower partial moments of the distribution. However, in many applications, the relevant distribution is not completely known, due to heterogeneous beliefs, subjective distortion, and/or estimation error. To account for incomplete information about the probability distribution, we develop an approach to stochastic efficiency analysis based on relative entropy and empirical likelihood.

Our approach is reminiscent of the early work by Pearman and Kmietowicz (1986) and Keppe and Weber (1989) on algorithms for pairwise first-order SD (FSD) and second-order SD (SSD) based on linear partial information about probabilities. A distinguishing feature of our analysis is the use of information-theoretical divergence measures and statistical hypothesis tests. In addition, we account for mixtures of prospects and higher-order risk aversion in order to achieve more discriminating power than is possible based on pairwise lower-order SD rules.

We first consider the case where the analyst knows the population distribution but not the decision-maker's subjective beliefs or probability distortion, as in a controlled choice experiment. We define a theoretical inefficiency measure based on the divergence between given probabilities and the nearest probabilities that can rationalize a given portfolio for some admissible utility function. The measure minimizes relative entropy (RE) over preferences and probabilities subject to inequality moment restrictions.

Next, we consider the case where the population distribution is latent and the analyst has access to a random time series of observations. When applied to the empirical distribution, our inefficiency measure becomes a likelihood ratio (LR) statistic for moment restrictions in the spirit of the nonparametric empirical likelihood (EL) methodology (see the review by Owen 2001). To account for serial dependence and stochastic volatility, we employ the blockwise EL (BEL; Kitamura 1997). The LR statistic allows for large-sample testing based on critical values from a chi-squared distribution.

Attractively, the EL approach avoids the statistical estimation and numerical inversion of the error covariance matrix, which can be problematic for a broad cross section of mutually dependent alternatives. Still, computing our LR statistic is generally a challenging high-dimensional nonconvex optimization problem. Our problem is more complex than standard EL problems, because the system of equations is underdetermined and involves moment inequality conditions and many parameter restrictions. To reduce the computational burden, we develop a tight numerical approximation based on a two-step optimization procedure and piecewise linearization techniques.

We focus on decreasing absolute risk aversion (DARA) SD (Vickson 1975 and Bawa 1975) arguably the most appealing of all SD criteria. DARA SD (DSD) imposes the standard regularity conditions of nonsatiation, risk aversion, and DARA. In many applications, relaxing these regularity conditions leads to a substantial loss of discriminating power and economic meaning. For example, Basso and Pianca (1997) show that general Nth order SD (NSD) allows for financial option prices that are inconsistent with DARA and Post et al. (2015) show that NSD underestimates the pricing errors of small-cap stocks for DARA investors.

Our empirical application analyzes the efficiency of a passive stock market index in data sets from the empirical asset pricing literature. One way to view this application is as a test for capital market equilibrium in a representative-investor model. Almeida and Garcia (2012) measure misspecification of an asset pricing model using relative entropy. A distinguishing feature of our SD-based approach is the use of the general regularity conditions for a DARA representative investor without the specification of a specific functional form. The application can alternatively be interpreted as a revealed preference analysis of those individual investors who adopt a passive strategy of broad diversification.

In portfolio management, our method can be used at several stages of the investment process. The LR test can establish whether a passive benchmark index is significantly inefficient in a given investment universe. In addition, the test can be used for in-sample back testing and out-of-sample performance evaluation of active investment strategies.

2. Stochastic Efficiency

We consider N distinct base assets with gross returns $X_1, \ldots, X_N \in \mathbb{R}_{++} := (0, +\infty)$ that are treated as random variables with a discrete, state-dependent joint distribution characterized by S mutually exclusive and exhaustive states of the world (or "scenarios") with probabilities p_s and realizations $x_{1,s}, \ldots, x_{N,s}$, $s = 1, \ldots, S$.

The probabilities are characterized by a probability mass function (PMF) with the set of scenarios as support; $p_s = p(s)$, $p \in \mathscr{P} := \{p\colon \{1, \ldots, S\} \to [0,1]; \sum_{s=1}^{S} p(s) = 1\}$. Depending on the purpose of the analysis, the state probability p_s may be thought of as the objective probability, the subjective probability, or the estimated probability of scenario $s = 1, \ldots, S$.

The portfolio possibilities are represented by the unit simplex $\mathscr{X} := \{\sum_{i=1}^{N} \lambda_i X_i\colon \sum_{i=1}^{N} \lambda_i = 1; \lambda_i \geq 0\ i = 1, \ldots, N\}$. To allow for general linear restrictions, the base assets $X_1, \ldots, X_N$ can be thought of as the vertices of a general polyhedral choice set; see Post and Versijp (2007, §V) for a general treatment. To allow for dynamic intertemporal choice problems, the base assets could be periodically rebalanced portfolios of individual securities. We evaluate a given and feasible portfolio $Y \in \mathscr{X}$ with realizations $y_1, \ldots, y_S$.

Investor preferences are described by decreasing absolute risk aversion (DARA) utility functions $u \in \mathscr{U} := \{u \in \mathscr{C}^3\colon u'(x) > 0; u''(x) \leq 0; \ln(u'(x))'' \geq 0\ \forall x \in \mathbb{R}_{++}\}$. Our analysis is invariant to positive linear transformations of utility. Motivated by this result, the mathematical optimization strategy in §5 employs an empirical normalization of marginal utility.

Definition 1 (Stochastic Efficiency). A given portfolio $Y \in \mathscr{X}$ is stochastically efficient for a given PMF $p \in \mathscr{P}$ if it is the optimum for some admissible utility function $u \in \mathscr{U}$, or $Y = \arg\max_{X \in \mathscr{X}} \sum_{s=1}^{S} p_s u(x_s)$.

This definition does not require that Y is a unique optimum, because we do not require that utility is strictly concave ($u''(x) < 0$). In theory, it is therefore possible that Y is classified as efficient despite being dominated by an alternative portfolio $X \in \mathscr{X}$, for example, a mean-preserving antispread of Y. Specifically, it is possible that $\sum_{s=1}^{S} p_s u(x_s) = \sum_{s=1}^{S} p_s u(y_s)$ for some $u \in \mathscr{U}$ and $\sum_{s=1}^{S} p_s u(x_s) > \sum_{s=1}^{S} p_s u(y_s)$ for other $u \in \mathscr{U}$. However, it follows from Kopa and Post (2015, Lemma 1) that dominance is equivalent to $\sum_{s=1}^{S} p_s u(x_s) > \sum_{s=1}^{S} p_s u(y_s)$ for all strictly concave $u \in \mathscr{U}$. Since every concave function can be approximated with arbitrary high accuracy by some strictly concave function, Definition 1 excludes dominance relations that are robust to minimal changes to the outcomes or probabilities.

Consider the following example with two assets ($N = 2$) and two scenarios ($S = 2$) with equal probability: $X_1 = [2\ 2]^{\mathrm{T}}$; $X_2 = [1\ 3]^{\mathrm{T}}$; $p = [0.5\ 0.5]^{\mathrm{T}}$. In this case, $Y = X_2$ is optimal for a risk-neutral investor and hence efficient by Definition 1. At the same time, Y is dominated by X_1 because it is more risky. The divergence of the two criteria occurs because Y is the (nonunique) optimum for some concave utility functions but nonoptimal for all strictly concave utility functions. If we change the probabilities to $p' = [0.5 + \epsilon\ 0.5 - \epsilon]^{\mathrm{T}}$ with $\epsilon > 0$ for an infinitesimal constant, then Y is both inefficient and dominated; if we use $p'' = [0.5 - \epsilon\ 0.5 + \epsilon]^{\mathrm{T}}$, then Y is both efficient and not dominated.

Following Post (2003), our analysis formulates portfolio efficiency in terms of a finite number of optimality conditions:

Lemma 1 (Optimality Conditions). *A given portfolio $Y \in \mathscr{X}$ is optimal for given utility function $u \in \mathscr{U}$ and PMF $p \in \mathscr{P}$ if and only if it obeys the following first-order optimality conditions*:

$$\sum_{s=1}^{S} p_s u'(y_s)(x_{i,s} - y_s) \leq 0, \quad i = 1, \ldots, N. \tag{1}$$

This formulation is analytically convenient because it avoids an explicit search over all feasible portfolios $X \in \mathscr{X}$ for a dominating alternative. Such a search would introduce the need to account for the ranking of the returns to the optimal portfolio, a task that generally requires integer programming. For the SSD criterion, Kuosmanen (2004), Roman et al. (2006), and Kopa and Post (2015) deal with this task using large linear programming (LP) relaxations with $\mathscr{O}(S^2)$ model variables and constraints. Programs for the stronger higher-order SD and DSD criteria however are not available to the best of our knowledge.

Our framework could be used to elicit investor preferences in the spirit of revealed preference analysis. Similar to Afriat's theorem (Afriat 1967, 1972), our optimality conditions (1) define a supporting hyperplane; the marginal utility levels $u'(y_s), s = 1, \ldots, S$, can be seen as Afriat numbers. Varian (1983) develops revealed preference analysis for investor behavior; Varian (1985) and Kuosmanen et al. (2007) develop statistical tests. A distinguishing feature of our study is the use of the concepts of SD, RE, and EL. In addition, we focus on decision support and asset pricing rather than revealed preference analysis.

We will refer to the product term $d_s := p_s u'(y_s)$ as the "decision weight" for state $s = 1, \ldots, S$. The computational strategy in §5 relies on the log decomposition $\ln(d_s) = \ln(p_s) + \ln(u'(y_s)), s = 1, \ldots, S$. The log state probability $\ln(p_s)$ will be used for a linear formulation of RE; log marginal utility $\ln(u'(y_s))$ will be used for a linear formulation of DARA.

3. Relative Entropy

Our analysis starts with a given PMF $q \in \mathscr{P}$ and measures the divergence to the nearest PMF $p \in \mathscr{P}$ that rationalizes the evaluated portfolio $Y \in \mathscr{X}$. One way to interpret the problem is as one of searching for subjectively distorted probabilities $p_s = g(q_s)$ in the neighborhood of the objective probabilities $q_s = \mathbb{P}(Y = y_s)$, where $g(\cdot)$ is an unknown distortion function. Section 4 explores an alternative interpretation, one of searching for the objective probabilities $p_s = \mathbb{P}(Y = y_s)$ in the neighborhood of estimated probabilities $q_s = \hat{p}_s$.

Definition 2 (Kullback–Leibler Information Criterion). The divergence of PMF $q \in \mathscr{P}$ with respect to PMF $p \in \mathscr{P}$ amounts to

$$KLIC(p \mid q) := \sum_{s=1}^{S} q_s \ln\left(\frac{q_s}{p_s}\right). \tag{2}$$

Relative entropy is a well-known information-theoretic measure of the dissimilarity between two discrete distributions. It has a number of known desirable properties, for example, it is a convex function of q, is always nonnegative, and equals zero if and only if $p = q$. Note that $KLIC(p \mid q)$ is not a metric because it is not symmetric and does not satisfy the triangle inequality.

Definition 3 (Stochastic Inefficiency Measure). For given PMF $q \in \mathscr{P}$, the deviation of portfolio $Y \in \mathscr{X}$ from stochastic efficiency is measured by the smallest divergence with respect to a PMF $p \in \mathscr{P}$ that obeys the optimality conditions for some admissible utility function $u \in \mathscr{U}$:

$$SIM(q) := \min_{(p,u)\in\mathscr{A}} KLIC(p \mid q), \tag{3}$$

$$\mathscr{A} := \left\{(p, u) \in \mathscr{P} \times \mathscr{U}: \sum_{s=1}^{S} p_s u'(y_s)(x_{i,s} - y_s) \leq 0,\ i = 1, \ldots, N\right\}. \tag{4}$$

Proposition 1 (Efficiency Condition). *A given portfolio $Y \in \mathscr{X}$ is stochastically efficient relative to given PMF $q \in \mathscr{P}$ if and only if $SIM(q) = 0$; the portfolio is stochastically inefficient if and only if $SIM(q) > 0$.*

The inefficiency measure $SIM(q)$ amounts to the minimum relative entropy (MRE) relative to a PMF that rationalizes the evaluated portfolio for some admissible utility function. For an empirical distribution with $\hat{p}_s = S^{-1}, s = 1, \ldots, S$, the MRE problem becomes a maximum entropy problem and $SIM(\hat{p})$ resembles an LR test statistic (see §4).

The optimal solution for $p \in \mathscr{P}$ is a unique property of the RE approach to SD analysis. The solution can be seen as "implied probabilities," or probabilities that are implied by the composition of the evaluated portfolio and the maintained assumptions. The implied probabilities have various possible applications, including robustness analysis, recentering the simulation process of bootstrap procedures, and detecting outliers or Peso problems.

Theoretically, it is possible that the MRE problem has no feasible solution, because the probabilities are required to be nonnegative. In this case, it is not possible to reweight the scenarios in a way that rationalizes the evaluated portfolio for an admissible utility function, an extremely robust form of stochastic inefficiency. This situation however seems pathological when the number of scenarios far exceeds the number of assets ($S \gg N$) and we have never encountered it in our simulations and applications.

4. Empirical Likelihood

We will now consider the case where the return distribution is a latent stochastic process with continuous CDF $\mathcal{F}\colon \mathbb{R}^N_{++} \to [0,1]$ with a finite covariance matrix of full rank (N). Let $g(\mathbf{x} \mid u) := u'(y)(\mathbf{x} - \mathbf{1}_N y)$ denote a vector-valued moment function on $\mathbb{R}^N_{++} \times \mathcal{U}$ with expectation $\mathcal{G}(u) := \mathbb{E}[g(\mathbf{x} \mid u)] = \int g(\mathbf{x} \mid u)\, d\mathcal{F}(\mathbf{x})$. Our null hypothesis is portfolio efficiency and the alternative is inefficiency:

$$\mathcal{H}_0\colon\ \mathcal{G}(u) \le \mathbf{0}_N \quad u \in \mathcal{U}, \tag{5}$$

$$\mathcal{H}_1\colon\ \mathcal{G}(u) \nleq \mathbf{0}_N \quad \forall\, u \in \mathcal{U}. \tag{6}$$

The analyst observes a time series $(\mathbf{x}_t)_{t=1}^{T}$, $\mathbf{x}_t := (x_{1,t}, \ldots, x_{N,t})^{\mathrm{T}}, t = 1, \ldots, T$. The empirical realizations represent historical scenarios and we will continue to use the notation from the previous sections with $S = T$. The return sequence $(\mathbf{x}_t)_{t\in\mathbb{N}}$ is a stationary, weakly dependent dynamic process. Various stationary autoregressive integrated moving average, generalized autoregressive conditional heteroskedasticity, and stochastic volatility models meet this requirement.

Following the BEL procedure (Kitamura 1997), we subdivide the original time series into $M := (T - B + 1)$ maximally overlapping blocks of B consecutive observations. The block length B grows with T but at a lower rate, for example, $(B^{-1} + B^2/T) \to 0$ as $T \to \infty$. For serially IID observations, $B = 1$ yields the standard EL method.

The BEL procedure assigns probabilities $\rho \in \mathcal{P}_M := \{\rho\colon \{1, \ldots, M\} \to [0,1]; \sum_{b=1}^{M} \rho(b) = 1\}$ to the data blocks rather than the individual observations. Observation $t = 1, \ldots, T$ is included in all blocks with indices from $\underline{b}_t := \max(1, t - B + 1)$ to $\bar{b}_t := \min(t, M)$. The PMF space is therefore given by

$$\mathcal{P}_T := \left\{ p \in \mathcal{P}\colon p_t = \sum_{b=\underline{b}_t}^{\bar{b}_t} B^{-1} \rho_b,\ t = 1, \ldots, T;\ \rho \in \mathcal{P}_M \right\}. \tag{7}$$

For a given PMF $p \in \mathcal{P}_T$, we can construct the following estimator for $\mathcal{G}(u)$:

$$\hat{\mathcal{G}}_T(p, u) := \sum_{t=1}^{T} p_t g(\mathbf{x}_t \mid u) = \sum_{t=1}^{T} \sum_{b=\underline{b}_t}^{\bar{b}_t} B^{-1} \rho_b g(\mathbf{x}_t \mid u). \tag{8}$$

The unconstrained maximum likelihood (ML) estimate of $\mathcal{F}(\mathbf{x})$ equals the empirical distribution function $\hat{\mathcal{F}}_T(\mathbf{x}) := T^{-1} \sum_{t=1}^{T} \mathbb{I}(\mathbf{x}_t \le \mathbf{x})$, with associated mass function $\hat{p}_t := T^{-1}, t = 1, \ldots, T$. The ML estimate of $\mathcal{G}(u)$ amounts to

$$\hat{\mathcal{G}}_T(\hat{p}, u) = T^{-1} \sum_{t=1}^{T} g(\mathbf{x}_t \mid u). \tag{9}$$

Applying the RE measure (2) and inefficiency measure (3) to the empirical distribution gives the following empirical measures:

$$KLIC_T(p) := \sum_{t=1}^{T} T^{-1} \ln\left(\frac{T^{-1}}{p_t}\right)$$

$$= -\ln(M) - \sum_{b=1}^{M} M^{-1} \ln(\rho_b), \tag{10}$$

$$SIM_T := \min_{(p,u)\in\mathcal{A}_T} KLIC_T(p), \tag{11}$$

$$\mathcal{A}_T := \{(p, u) \in \mathcal{P}_T \times \mathcal{U}\colon \hat{\mathcal{G}}_T(p, u) \le \mathbf{0}_N\}. \tag{12}$$

The empirical RE measure (10) is a monotone decreasing transformation of the LR:

$$KLIC_T(p) = -M^{-1} \ln(LR_T(p)), \tag{13}$$

$$LR_T(p) := \frac{\prod_{t=1}^{T} p_t}{\prod_{t=1}^{T} \hat{p}_t} = \frac{\prod_{b=1}^{M} \rho_b}{M^{-M}}. \tag{14}$$

Hence, minimizing $KLIC_T(p)$ amounts to maximizing $LR_T(p)$ and the empirical inefficiency measure SIM_T amounts to an LR test statistic for moment restrictions with a known asymptotic sampling distribution.

Proposition 2 (Asymptotic Distribution). *Let $z \sim \chi^2_{N-1}$, a central chi-squared distribution with $N - 1$ degrees of freedom, and $c(\alpha) := \inf_c\{c\colon \mathbb{P}[z \ge c] \le \alpha\}$, the associated critical value for a given significance level $\alpha < 0.5$. The following asymptotic size and power properties apply*:

$$\lim_{T\to\infty} \mathbb{P}((2M/B) \cdot SIM_T \ge c(\alpha) \mid H_0) \le \alpha, \tag{15}$$

$$\lim_{T\to\infty} \mathbb{P}((2M/B) \cdot SIM_T > c(\alpha) \mid H_1) = 1. \tag{16}$$

It is difficult to establish the uniform validity of asymptotic approximations for our general system of moment inequalities and parameter inequalities. However, we know that the asymptotic null distribution is bounded from above by a chi-squared distribution. We can use this distribution for conservative statistical inference. In large samples, we can reject the null for a given significance level $\alpha < 0.5$ if $SIM_T > c(\alpha) \cdot (B/(2M))$.

The bounding distribution describes the case where the inequality moment conditions are binding ($\mathcal{G}(u) = \mathbf{0}_N$) and thus ignores the search over Lagrange multipliers for nonactive assets. In addition, the bounding distribution assumes a unique optimizing utility function and thus ignores the minimization over $\mathcal{U}_0 := \{u \in \mathcal{U}\colon \mathcal{G}(u) \le \mathbf{0}_N\}$. Ignoring nonactive assets and multiple optimizing utility functions increases the p-value and the chi-squared distribution represents a least-favorable distribution that maximizes the p-value under the null.

Future research could focus on improving the approximation of the null distribution of the LR statistic by using the bootstrap or subsampling. The implied probabilities could be used to recenter the simulation process. An attractive feature of this approach is that it preserves the known (historical) scenarios by calibrating the unknown probabilities, hence avoiding unrealistic scenarios (such as negative gross returns).

Another possible extension uses Bayesian statistics. The LR test uses a frequentist inference framework. It is tempting to see the implied probabilities as posterior probabilities. However, a proper Bayesian definition of posterior probabilities would require us to use a likelihood function based on $KLIC(q \mid p)$ rather than $KLIC(p \mid q)$. This orientation will require another computational strategy, because $KLIC(q \mid p) = \sum_{s=1}^{S} p_s \ln(p_s/q_s)$ involves a multiplicative structure for the probabilities and their logs. Still, we may expect a limited effect in large samples under the null, because $p = q \Longrightarrow KLIC(p \mid q) = KLIC(q \mid p) = 0$.

5. Computational Strategy

Computing the test statistic SIM_T generally is a high-dimensional nonconvex optimization problem. The standard computational strategy for EL problems does not apply here, because our system of equations generally is not identified and uses moment inequality conditions and parameter restrictions. We propose a tractable two-step optimization procedure to approximate SIM_T by splitting the original problem into two subproblems with lower dimensions. In large samples, the subproblems are near independent and the approximation achieves machine precision levels in our experience.

Step 1. Let $\mathbf{p} := (p_1 \cdots p_T)^{\mathrm{T}}$ and $\nabla\mathbf{u} := (u'(y_1) \cdots u'(y_T))^{\mathrm{T}}$. Minimize the divergence with respect to (normalized) decision weights $\mathbf{d} := \mathbf{p} \otimes \nabla\mathbf{u}$ rather than the state probabilities ($\mathbf{p}$):

$$SIM_{1;T} := \min_{\mathbf{d}\in\mathcal{D}_T}\left\{T^{-1}\sum_{t=1}^{T}\ln\left(\frac{T^{-1}}{d_t}\right)\middle| T^{-1}\sum_{t=1}^{T} d_t y_t = 1\right\}, \tag{17}$$

$$\mathcal{D}_T := \{\mathbf{p}\otimes\nabla\mathbf{u}\colon (p,u)\in\mathcal{A}_T\} \tag{18}$$

$$= \left\{\mathbf{d}\in\mathbb{R}_+^T\colon \sum_{t=1}^{T} d_t(x_{i,t}-y_t)\le 0,\quad i=1,\ldots,N\right\}. \tag{19}$$

Step 2. Minimize the relative entropy subject to the optimal first-step solution for the decision weights:

$$SIM_{2;T} := \min_{\substack{(p,u)\in\mathcal{A}_T:\\ \mathbf{p}\otimes\nabla\mathbf{u}=\mathbf{d}_T^*}}\left(T^{-1}\sum_{t=1}^{T}\ln\left(\frac{T^{-1}}{p_t}\right)\right), \tag{20}$$

$$\mathbf{d}_T^* := \arg\min_{\mathbf{d}\in\mathcal{D}_T}\left\{T^{-1}\sum_{t=1}^{T}\ln\left(\frac{T^{-1}}{d_t}\right)\middle| T^{-1}\sum_{t=1}^{T} d_t y_t = 1\right\}. \tag{21}$$

We propose to use $SIM_{2;T}$ as an approximation for SIM_T in large samples. If $M \gg N$, then the system of N joint inequalities is highly underdetermined and the solution space for the decision weights $\mathcal{D}_T$ is approximately orthogonal to the solution space for the state probabilities $\mathcal{A}_{\mathcal{P};T} := \{p \in \mathcal{P}_T\colon (p,u) \in \mathcal{A}_T, u \in \mathcal{U}\}$ and the solution space for the utility function $\mathcal{A}_{\mathcal{U};T} := \{u \in \mathcal{U}\colon (p,u) \in \mathcal{A}_T, p \in \mathcal{P}_T\}$, or, equivalently, $\mathcal{D}_T \approx \{\mathbf{p} \otimes \nabla\mathbf{u}\colon p \in \mathcal{A}_{\mathcal{P};T}, u \in \mathcal{A}_{\mathcal{U};T}\} =: \mathcal{D}_T^{\perp}$.

Proposition 3 (Orthogonal Solution Spaces). *If the solution space $\mathcal{D}_T$ is orthogonal to the solution spaces $\mathcal{A}_{\mathcal{P};T}$ and $\mathcal{A}_{\mathcal{U};T}$, then the approximation is perfect; $\mathcal{D}_T = \mathcal{D}_T^{\perp} \Longrightarrow SIM_{2;T} = SIM_T$.*

The first optimization problem (17) is relatively small and convex; it resembles a standard EL problem. The optimization problem in Step 2 is however nonconvex because of the multiplicative equality constraints $\mathbf{p} \otimes \nabla\mathbf{u} = \mathbf{d}_T^*$. We propose a tight linear approximation based on (i) the log decomposition $\ln(d_t) = \ln(p_t) + \ln(u'(y_t)), t = 1, \ldots, T$; (ii) local piecewise-linear (PWL) approximations to the logarithmic function $\ln(p_t), t = 1, \ldots, T$; (iii) the PWL representation of log marginal utility $\ln(u'(y_t)), t = 1, \ldots, T$ of Post et al. (2015).

To model the state probabilities, we introduce three sets of model variables: π_t and $\tilde{\pi}_t$ capture p_t and $\ln(p_t)$, $t = 1, \ldots, T$, respectively; ϖ_b captures ρ_b, $b = 1, \ldots, M$. To connect the level variables π_t, $t = 1, \ldots, T$, with the associated log variables $\tilde{\pi}_t$, $t = 1, \ldots, T$, we use a local PWL approximation to the logarithmic function.

Lemma 2 (Log State Probability). *Let $\hat{p}_{t,k} \in (0,1]$, $k = 1, \ldots, K$, denote a set of data-dependent sampling points for p_t, $t = 1, \ldots, T$. We can build the following PWL upper envelope function*:

$$h(p_t) = \min_{k=1,\ldots,K}(\ln(\hat{p}_{t,k}) + (\hat{p}_{t,k})^{-1}(p_t - \hat{p}_{t,k})). \tag{22}$$

By construction, $\ln(p_t) \le h(p_t)$ and $\ln(p_t) \approx h(p_t)$ if $p_t \approx \hat{p}_{t,k}$ for some $k = 1, \ldots, K$. (Without proof.)

To implement this approach, our application uses $K = 9$: $\hat{p}_{t,k} = T^{-1}2^{(k-5)}$, $k = 1, \ldots, 9$. This specification is based on our experience that $2^{-3} \le (T^{-1}p_t) \le 2^3$ for the large majority of scenarios and $2^{-4} \le (T^{-1}p_t) \le 2^4$ in almost all cases.

For modeling marginal utility, we use the rank order and the ranked values of the evaluated returns y_t, which are known in any given sample. We use $R(t)$ for the rank order of y_t and $y^{[r]}$, $r = 1, \ldots, T$ for the rth ranked value, so that $y^{[1]} < \cdots < y^{[T]}$. We use ordered data only for computational purposes and neither the concepts in §3 nor the statistical theory in §4 requires ordered data.

The model variables $\tilde{\beta}_r$, $r = 1, \ldots, T$, capture $\ln(u'(y^{[r]}))$, $r = 1, \ldots, T$. In addition, for the PWL representation of $\ln(u'(y))$, we introduce α_r, $r = 1, \ldots, T$, to capture the behavior of the absolute risk aversion (ARA) coefficient $a(y) := -\ln(u'(y))'$.

Lemma 3 (Log Marginal Utility). *For any utility function $u \in \mathcal{U}$, we can represent the levels of log marginal utility $\ln(u'(y^{[r]}))$, $r = 1, \ldots, T$, using a decreasing and convex PWL function that is linear in a finite number of parameters*:

$$\ln(u'(y^{[r]})) = \sum_{s=r}^{T-1} \alpha_s (y^{[s+1]} - y^{[r]}) + \alpha_T, \quad r = 1, \ldots, T; \tag{23a}$$

$$\alpha_s \geq 0, \quad s = 1, \ldots, T-1. \tag{23b}$$

(The lemma is based on Post et al. (2015, Proposition 2).)

The following LP problem in canonical form uses the above lemmas and model variables to approximate $SIM_{2;T}$:

$$\widehat{SIM}_{2;\,T} := \min\left(-\ln(T) - T^{-1} \sum_{t=1}^{T} \tilde{\pi}_t\right) \tag{24a}$$

$$\text{s.t.} \quad \tilde{\pi}_t + \sum_{s=R(t)}^{T-1} \alpha_s (y^{[s+1]} - y^{[R(t)]}) + \alpha_T = \ln(d_t^*), \quad t = 1, \ldots, T; \tag{24b}$$

$$-\tilde{\pi}_t + (\hat{p}_{t,\,k} B)^{-1} \sum_{b=\underline{b}_t}^{\bar{b}_t} \varpi_b \geq -\ln(\hat{p}_{t,\,k}) + (\hat{p}_{t,\,k})^{-1} \hat{p}_{t,\,k}, \quad t = 1, \ldots, T;\ k = 1, \ldots, K; \tag{24c}$$

$$B^{-1} \sum_{t=1}^{T} \sum_{b=\underline{b}_t}^{\bar{b}_t} \varpi_b = 1; \tag{24d}$$

$$\alpha_r \geq 0, \quad r = 1, \ldots, T-1; \tag{24e}$$

$$\varpi_b \geq 0, \quad b = 1, \ldots, M. \tag{24f}$$

Proposition 4 (Second-Stage Approximation). *As K increases, $\widehat{SIM}_{2;T}$ approaches $SIM_{2;T}$ from below.*

We may use the error terms $(\pi_t^* - \exp(\tilde{\pi}_t^*))$ or $(d_t^* - \exp(\tilde{\pi}_t^* + \tilde{\beta}_t^*))$, $t = 1, \ldots, T$, to diagnose the goodness of the linear approximation and possibly refine the specification of the sampling points. In our experience with empirical applications, the above-mentioned specification with $K = 9$ yields only miniscule errors and works satisfactory.

The two-step optimization procedure is perfectly manageable with standard computer hardware and solver software for the typical problem dimensions of empirical asset pricing data sets ($N \leq 100$ and $T \leq 1{,}000$). The total run time of all computations for our simulations and applications spanned several working days on a standard desktop PC with a 2.93 GHz quad-core Intel i7 processor, 16 GB of RAM, and using MATLAB with the external MOSEK solver.

6. Simulation Experiment

We analyze the small-sample properties of our LR test in a Monte Carlo simulation experiment from Post and Versijp (2007, §VI-C). The experiment is based on an investment problem with one riskless asset and 10 risky assets with investment returns drawn from the empirical distribution of monthly excess returns to 10 U.S. beta-decile stock portfolios. The empirical distribution is highly asymmetric (witness, for example, a highly negative skewness for the low-beta portfolios and a highly positive skewness for the high-beta portfolios) and hence the mean-variance (M-V) criterion does not apply.

The statistical size (relative frequency of false rejection) is measured by the rejection rate for a stochastically efficient mean-lower partial moment portfolio (MLPMP). This portfolio is found by minimizing the second-order lower partial moment for a return threshold of −/−10% subject to the constraint that the expected return is greater than or equal to 0.83% per month. The MLPMP can be shown to be DSD efficient but M-V inefficient. The statistical power (relative frequency of false acceptance) is measured by the rejection rate for the inefficient equal-weighted portfolio (EWP).

Since the simulated observations are serially independent and identically distributed (IID), we use a standard EL LR test ($B = 1$, $M = T$). We compare the LR test with a standard generalized method of moments (GMM) J-test for M-V efficiency. We also run three J-tests for stochastic efficiency: (i) the two-stage GMM J-test for third-order SD (TSD) efficiency of Post and Versijp (2007) with the inverse covariance matrix of returns as the initial weighting matrix; (ii) an iterated GMM (IGMM) J-test for TSD efficiency (with 10 iterations) as a numerical approximation to the continuously updating GMM (CUGMM); (iii) a two-stage GMM J-test for DSD efficiency that imposes the additional restriction of DARA.

Table 1 summarizes our results. For the M-V J-test, the rejection rate under the null exceeds the nominal level and increases with the sample size, reflecting the M-V inefficiency of the MLPMP. By contrast, the size functions of the SD tests show a decreasing pattern. Despite the use of asymptotically conservative critical values, the SD tests overreject in small samples, presumably because the left tail of the distribution (which determines the efficiency classification of the MLPMP) converges at a relatively low rate.

The performance of the two-step GMM is indistinguishable from that of the IGMM. Our choice of the initial weighting matrix removes most of the differences between the two methods. Unreported results show that the performance of the two-step method deteriorates when the procedure starts from the identity matrix rather than the inverse covariance matrix of returns. In addition, the IGMM often encounters oscillation rather than convergence of the optimal utility gradient. It is particularly challenging that the error

Table 1. Simulation Experiment

		Size (rejection rate MLPMP)			Power (rejection rate EWP)		
Test	T	$\alpha = 0.025$	$\alpha = 0.050$	$\alpha = 0.100$	$\alpha = 0.025$	$\alpha = 0.050$	$\alpha = 0.100$
GMM M-V J-test	120	0.067	0.121	0.200	0.125	0.192	0.295
	240	0.069	0.108	0.189	0.140	0.222	0.322
	480	0.065	0.118	0.204	0.251	0.352	0.472
	960	0.150	0.180	0.260	0.610	0.740	0.860
Two-stage GMM TSD J-test	120	0.049	0.096	0.162	0.090	0.132	0.222
	240	0.063	0.086	0.143	0.119	0.183	0.261
	480	0.035	0.061	0.104	0.182	0.266	0.378
	960	0.031	0.070	0.110	0.465	0.545	0.657
IGMM TSD J-test	120	0.049	0.097	0.161	0.090	0.132	0.222
	240	0.062	0.085	0.144	0.119	0.183	0.262
	480	0.035	0.060	0.102	0.181	0.266	0.378
	960	0.030	0.060	0.110	0.469	0.561	0.663
Two-stage GMM DSD J-test	120	0.059	0.106	0.176	0.103	0.154	0.238
	240	0.068	0.094	0.153	0.140	0.212	0.285
	480	0.041	0.063	0.110	0.220	0.306	0.405
	960	0.033	0.067	0.101	0.595	0.658	0.797
EL DSD LR test	120	0.073	0.135	0.214	0.134	0.206	0.298
	240	0.071	0.116	0.175	0.149	0.241	0.316
	480	0.049	0.084	0.149	0.240	0.342	0.465
	960	0.037	0.073	0.119	0.608	0.706	0.831

Notes. We generate random data sets of returns to 10 risky assets by drawing from the empirical distribution of monthly excess returns to beta-decile stock portfolios. We analyze the statistical properties of five alternative tests for portfolio efficiency: (i) a GMM J-test for M-V efficiency; (ii) a two-stage GMM J-test for TSD efficiency; (iii) an IGMM J-test for TSD efficiency; (iv) a two-stage GMM J-test for DSD efficiency; (v) our LR test for DSD efficiency. In all five tests, the model variables are restricted by requiring that the portfolio optimality condition (1) holds with equality for the T-bill. Statistical size is measured by the rejection rate for the efficient MLPMP; statistical power by the rejection rates for the inefficient EWP. We consider sample sizes (T) of 120, 240, 480, and 960 observations. For every sample size, we generate 1,000 random samples from the relevant distribution. We apply all five tests to the same data sets in order to reduce the effect of simulation error. We use nominal significance levels (α) of 2.5%, 5%, and 10% and reject the null if the test statistic exceeds the relevant percentile of the chi-squared distribution with nine degrees of freedom.

covariance matrix explodes for utility functions that place a large weight on tail risk, reflecting the high cross-sectional tail correlation.

Comparing the rejection rates of the J-test for TSD efficiency and the J-test for DSD efficiency, we find that using the DSD criterion increases power compared with using the TSD criterion. This pattern reflects the important difference between DARA ($\ln(u'(x))'' \geq 0$) and prudence ($u'''(x) \geq 0$). In a material number of simulation runs, the TSD criterion fails to detect inefficiency of the EWP by allowing for a utility gradient with increasing absolute risk aversion ($\ln(u'(x))'' < 0$). The additional power of the DSD criterion seems to justify the additional computational burden of the DARA restrictions.

Using the LR test instead of the J-test leads to further improvements, consistent with the theoretical arguments of Newey and Smith (2004). The power converges quickly to levels close to that of the M-V efficiency test. Rejection rates above 50% arise at conventional significance levels for sample sizes of roughly 480 or more monthly observations, which is well in the range of standard data sets (our application below uses $T > 1{,}000$).

Our conclusions contrast with those of Davidson and Duclos (2013), who find no improvements for an LR test for pairwise FSD between two independent prospects relative to the paired t-test of Kaur et al. (1994). This disparity seems related to the use of different SD criteria and choice sets. Pairwise lower-order SD criteria are very sensitive to sampling variation compared with criteria that account for mixtures of prospects and higher-order risk aversion, making it difficult to distinguish between alternative statistical methods. In addition, the EL method has a comparative advantage when the number of moment conditions (N) is large. Whereas the paired t-test used by Davidson and Duclos does not involve an error covariance matrix, the GMM J-tests in our simulation experiment and empirical application need to estimate a high-dimensional error covariance matrix ($N = 10$).

7. Market Portfolio Efficiency

We examine whether the Center for Research in Security Prices (CRSP) all-share index is stochastically efficient relative to five different sets of 10 benchmark portfolios from the Kenneth French data library (http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html). The portfolios are formed by

Table 2. Tests of Market Efficiency

Panel A: LR test of market portfolio efficiency

Test portfolios	M-V GMM	TSD Two-stage GMM	TSD IGMM	DSD Two-stage GMM	DSD EL	DSD BEL $B=3$	DSD BEL $B=12$
10 ME	6.33	6.43	6.30	6.98	7.43	7.61	7.95
	(0.71)	(0.70)	(0.71)	(0.64)	(0.59)	(0.57)	(0.54)
10 BtM	11.76	10.31	10.26	10.86	11.18	11.33	11.59
	(0.23)	(0.33)	(0.33)	(0.29)	(0.26)	(0.25)	(0.24)
10 R1-1	34.66	30.55	30.52	33.54	28.55	28.59	33.67
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
10 R12-2	62.26	47.33	47.31	50.55	41.98	42.24	46.33
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
10 R60-13	11.84	11.69	11.84	14.05	13.32	13.83	14.13
	(0.22)	(0.23)	(0.22)	(0.12)	(0.15)	(0.13)	(0.12)

Panel B: Regression analysis of implied probabilities

Test portfolios	*C*	*DIF*	MKT	SMB	HML	UMD	*LAG*	R^2 (%)
10 ME	0.00	−0.01	−0.07	0.11	−0.02	−0.03	−0.01	31.6
	(−1.48)	(−21.9)	(−1.58)	(1.61)	(−0.26)	(−0.63)	(−0.29)	
10 BtM	0.00	−0.08	0.00	0.07	−0.06	−0.05	0.01	41.1
	(−1.90)	(−27.0)	(0.07)	(0.85)	(−0.83)	(−0.93)	(0.33)	
10 R1-1	−0.01	−0.59	0.12	−0.02	0.04	−0.07	−0.02	66.5
	(−5.19)	(−45.6)	(2.31)	(−0.23)	(0.50)	(−1.16)	(−0.97)	
10 R12-2	−0.02	−0.94	0.14	−0.13	−0.08	−0.08	0.01	66.7
	(−5.22)	(−40.1)	(1.95)	(−1.09)	(−0.69)	(−0.99)	(0.62)	
10 R60-13	−0.01	−0.11	0.03	−0.03	−0.06	−0.06	−0.05	46.6
	(−2.35)	(−29.3)	(0.46)	(−0.37)	(−0.72)	(−1.03)	(−2.13)	

Notes. Panel A shows results for seven tests for market portfolio efficiency relative to five different sets of 10 portfolios that are formed based on a given stock characteristic: ME, BtM, R1-1, R12-2, or R60-13. The CRSP all-share index is our market portfolio and the T-bill is our risk-free asset. We use monthly gross returns from the first available month to December 2014 ($1{,}008 \le T \le 1{,}062$). The analysis includes a GMM J-test for M-V efficiency, two-stage GMM and IGMM J-tests for TSD efficiency, a two-stage GMM J-test for DSD efficiency, our EL LR test and BEL LR tests with block lengths of $B = 3, 12$. The model variables of all tests are restricted to be consistent with the equity premium by requiring a zero pricing error for the T-bill. The table reports the statistic $T \cdot J_T$ for the J-tests and $(2M/B) \cdot SIM_T$ for the LR tests together with asymptotic p-values from the chi-squared distribution (in parentheses). Panel B uses OLS regression analysis in each of the five data sets to explain the variation in $\ln(p_t^*/q_t)$ using six regressors: our difficulty measure (*DIF*), the four factors of the Carhart (1997) factor model (MKT, SMB, HML, and UMD), and the one-month lagged value of the dependent variable (*LAG*). T-statistics are shown in parentheses below the coefficient estimates.

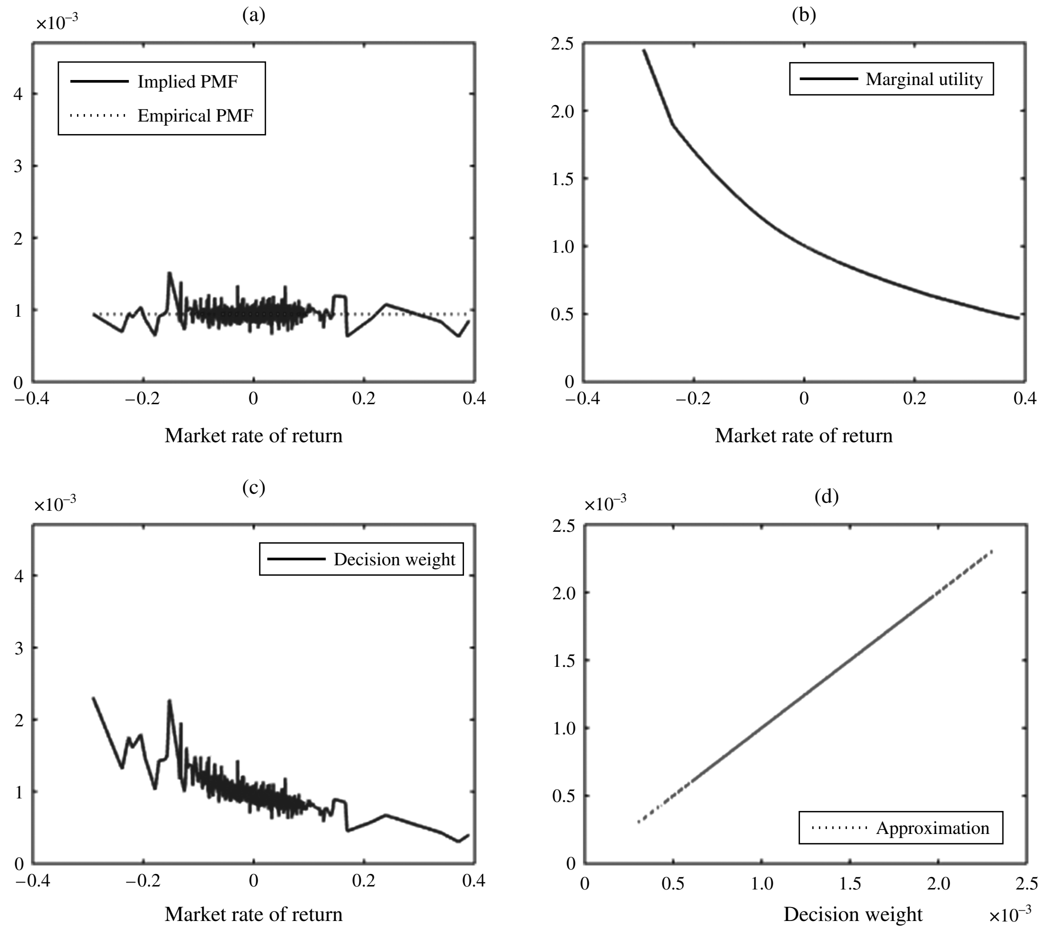
sorting stocks on a given characteristic: market capitalization of equity (ME), book-to-market ratio (BtM), short-term price reversal (R1-1), price momentum (R12-2), or long-term price reversal (R60-13). We use the Treasury bill as the risk-free asset. The sample period starts in July 1926 for the ME, BtM, and R1-1 data sets; January 1927 for the R12-2 data set; and January 1931 for the R60-13 data set. We use monthly gross returns from the starting month to December 2014, leading to five large data sets ($1{,}008 \le T \le 1{,}062$).

As in the simulation experiment, we use a GMM J-test for M-V efficiency, two-stage GMM and IGMM J-tests for TSD efficiency, a two-stage GMM J-test for DSD efficiency, and our EL LR test. We now also include BEL LR tests with block lengths of $B = 3$ and $B = 12$ to account for possible dynamic patterns. Using large data sets naturally reduces the statistical differences between the GMM and EL methods. Nevertheless, the EL method still has a comparative advantage in large samples due to the additional information in the form of implied probabilities and relatively affordable computations.

Panel A of Table 2 summarizes the results of seven different efficiency tests. We cannot reject market portfolio efficiency at conventional significance levels in the ME, BtM, and R60-13 data sets. In the R1-1 and R12-2 data sets, we can convincingly reject the null with all tests. The evidence against the null based on the DSD tests is always stronger than based on the TSD tests and sometimes even stronger than based on the M-V test (in the ME and R60-13 data sets). The BEL method yields only slightly lower p-values than the EL method. The robustness to the block length presumably reflects the use of low-frequency returns to diversified and periodically rebalanced portfolios, which are relatively close to being serially IID.

Panel B of Table 2 investigates the EL implied probabilities in the five data sets using ordinary least squares

Figure 1. LR Test Results for 10 ME Portfolios



Notes. Shown are (a) the implied probability π_s^* and empirical probability $q_t = 1/T$ vs. the market rate of return y_t, $t = 1, \ldots, T$; (b) the marginal utility levels $\exp(\tilde{\beta}_t^*)$ vs. the market rate of return y_t, $t = 1, \ldots, T$; (c) the decision weights δ_t^* vs. the market rate of return y_t, $t = 1, \ldots, T$; and (d) the decision weights δ_t^*, $t = 1, \ldots, T$ vs. the approximation $\exp(\tilde{\pi}_t^* + \tilde{\beta}_t^*)$, $t = 1, \ldots, T$. The data set comprises monthly rates of return from July 1926 to December 2014 ($T = 1{,}062$) on the 10 portfolios based on individual stocks' market capitalization of equity, as well as on the risk-free rate and market value-weighted market portfolio proxy, downloaded from the data library of Kenneth French.

(OLS) regression analysis. The regressand is the log normalized value $\ln(p_t^*/q_t)$, $t = 1, \ldots, T$. The key regressor is a measure of a given scenario's "difficulty level" that is constructed using an ancillary regression analysis. The "market model" links the portfolio returns to the T-bill rate and the market return: $x_{i,t} = x_{F,t} + \beta_i(y_t - x_{F,t}) + \varepsilon_i$, $i = 1, \ldots, 10$, where β_i is the "market beta." Under the null of M-V efficiency, the residuals must have zero means, that is, $\alpha_i := \mathbb{E}[\varepsilon_i] = 0$, $i = 1, \ldots, 10$. Motivated by this requirement, we define the following difficulty measure for a given scenario $t = 1, \ldots, T$:

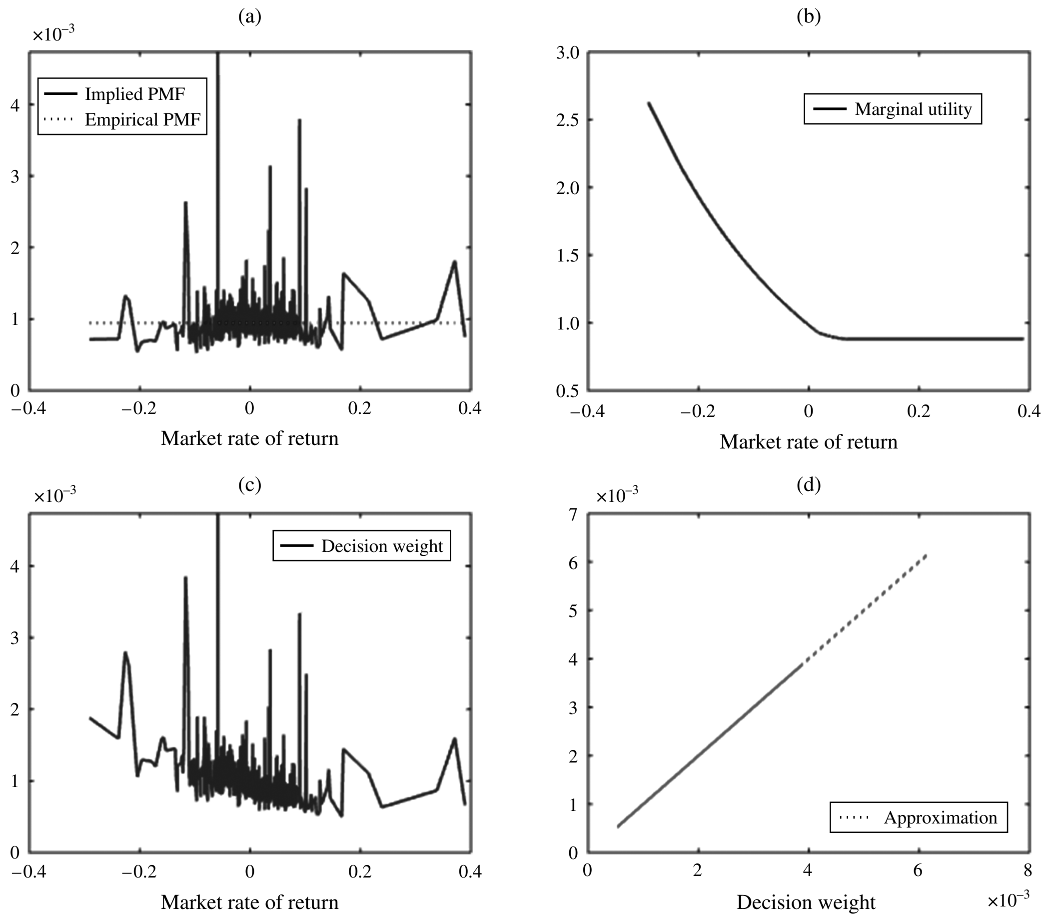
$$DIF_t := \sqrt{T}\frac{\sum_{i=1}^{10} \hat{\varepsilon}_{i,t}\hat{\alpha}_i}{\sum_{i=1}^{10}(\hat{\alpha}_i)^2}. \tag{25}$$

In this expression, $\hat{\varepsilon}_{i,t} := (x_{i,t} - x_{F,t}) - \hat{\beta}_i(y_t - x_{F,t})$ is the OLS regression residual, $\hat{\alpha}_i := T^{-1}\sum_{t=1}^{T}\hat{\varepsilon}_{i,t}$ is "Jensen's alpha," and $\hat{\beta}_i := \sum_{t=1}^{T}(x_{i,t} - x_{F,t})(y_t - x_{F,t})/\sum_{t=1}^{T}(y_t - x_{F,t})^2$ is the estimated beta. If DIF_t takes a high value for a given scenario, then that scenario "confirms" the overall evidence against the null, and we may expect a low implied probability.

To examine the role of standard stock return factors, the regressors also include the four factors of Carhart (1997): excess market return (MKT); "small-minus-big" (SMB) return; "high-minus-low" (HML) return; "up-minus-down" (UMD) return. A final regressor is the one-month lagged value of the regressand (*LAG*).

DIF explains about one-third to two-thirds of the variation of the implied probabilities. Not surprisingly, the explained percentage is higher in the data sets with a higher LR statistic. The three nonmarket factors (SMB, HML, and UMD) have minimal explanatory power (also if *DIF* is removed from the analysis). Apparently, our efficiency test does not amount to an indirect application of known factor models. The lagged value of implied probability also has little explanatory power, consistent with the notion that our

Figure 2. LR Test Results for 10 R12-2 Portfolios



Notes. Shown are (a) the implied probability π_s^* and empirical probability $q_t = 1/T$ vs. the market rate of return y_t, $t = 1, \ldots, T$; (b) the marginal utility levels $\exp(\tilde{\beta}_t^*)$ vs. the market rate of return y_t, $t = 1, \ldots, T$; (c) the decision weights δ_t^* vs. the market rate of return y_t, $t = 1, \ldots, T$; and (d) the decision weights δ_t^*, $t = 1, \ldots, T$ vs. the approximation $\exp(\tilde{\pi}_t^* + \tilde{\beta}_t^*)$, $t = 1, \ldots, T$. The data set comprises monthly rates of return from January 1927 to December 2014 ($T = 1{,}056$) on the 10 portfolios based on individual stocks' one-year price momentum (R12-2), as well as on the risk-free rate and market value-weighted market portfolio proxy, downloaded from the data library of Kenneth French.

data are close to serial IIDness and the limited effect of using BEL instead of EL.

Figures 1 and 2 show more detailed results for the EL LR test in the ME and R12-2 data sets, respectively. We select these two particular data sets because they represent the best empirical fit (ME data set) and worst empirical fit (R12-2 data set) among the five data sets.

Consistent with the low value of the LR statistic for the ME data set, Figure 1(a) shows that the implied probabilities (π_t^*) are close to the empirical probabilities ($q_t = T^{-1}$). The highest weight is assigned to the scenario of November 2003 (*DIF* < 0) and amounts to roughly 2.5 times the lowest weight, which is assigned to April 1981 (*DIF* > 0). By contrast, the implied probabilities diverge strongly from the empirical probabilities in Figure 2(b), reflecting a poor overall fit in the R12-2 data set. The weight for July 1944 (*DIF* < 0) is about 10 times the weight for September 2011 (*DIF* > 0) in this data set.

Figure 1(b) shows the optimal marginal utility levels ($\exp(\tilde{\beta}_t^*)$) for the ME data set. Marginal utility takes a two-piece exponential shape with a small drop in the ARA coefficient at a return level of about −/−20%. This shape resembles the gradient of a standard exponential function. Similar shapes are found for the BtM and R60-13 data sets. Apparently, explaining these data sets does not require nonstandard utility functions. By contrast, in Figure 2(b) for the R12-2 data set, the gradient displays a sharp drop in the ARA coefficient at a return level of about 0%, suggesting an important role for downside risk. A similar pattern is found for the R1-1 data set.

Figure 1(c) shows the optimal decision weights (δ_t^*) for the ME data set. Because of the limited differences between the implied and empirical probabilities, the decision weights show limited volatility and preserve the general decreasing and convex pattern of the utility gradient. By contrast, in data sets with a poor overall fit, the decision weights are very volatile and do not resemble a well-behaved gradient; see, for example, Figure 2(c).

Our PWL approximation to the logarithmic function (see §5) is very precise in this application. The model variables for the levels of implied probability show a near-perfect alignment with the associated model variables for the logs. Figures 1(d) and 2(d) illustrate the goodness of the approximation by plotting δ_t^* against $\exp(\tilde{\pi}_t^* + \tilde{\beta}_t^*)$ for the ME data set and R12-2 data set, respectively. Similar high precision is found for the other three data sets.

8. Concluding Remarks

The EL LR test seems a useful complement to existing tests for stochastic efficiency. The GMM J-test by Post and Versijp (2007) requires the statistical estimation and numerical inversion of the error covariance matrix, which adversely affects its small-sample properties. Our simulation study indeed shows superior properties in small samples of the LR test relative to the J-tests. In addition, our test generates additional information in the form of implied probabilities, which have several application possibilities. Furthermore, using our two-step procedure, computing the LR statistic is more affordable than solving the large QP problems needed for the J-test for DSD efficiency.

Scaillet and Topaloglou (2010) and Linton et al. (2014) rely on LP problems that search for a portfolio that is not SD dominated by the evaluated portfolio (Scaillet and Topaloglou) or that dominates the evaluated portfolio (Linton et al. 2014). Compared with our portfolio optimality conditions (1), these programs require a very large number of model variables and constraints in order to capture the return ranking of the solution portfolio. Furthermore, programs are only available for the FSD and SSD criteria and not for the stronger higher-order SD and DSD criteria.

We focus on the criterion of DSD efficiency in this study, because DARA and portfolio diversification are essential in our application area. Nevertheless, the analysis can be extended to the SSD and TSD efficiency criteria by relaxing the model constraints that impose (log-)convexity of marginal utility. An analysis of FSD efficiency however requires another approach, because the portfolio optimality conditions (1) are not sufficient when risk seeking is allowed. An analysis without mixtures can be based on the global optimality conditions in Theorem 1 of Post et al. (2015) rather than our local optimality conditions.

In particular, pairwise lower-order SD criteria seem analytically convenient because these criteria require a simple search over a finite number of known two-piece linear utility functions (see Russell and Seo 1989). Unfortunately, widening the admissible utility set and narrowing the admissible choice set decreases the number of cases that are outside the null and the ability to detect those cases in finite samples. Notably, pairwise lower-order SD relations are rare and easy to distort with minor perturbations in the left tail of the probability distribution (see Kroll and Levy 1980, Nelson and Pope 1991).

Acknowledgments

T. Post acknowledges generous research support from the College of Administrative Sciences and Economics and the Graduate School of Business of Koç University. V. Potì acknowledges generous research support from UCD Michael Smurfit Graduate Business School and University College Dublin.

Appendix

Proof of Lemma 1. The utility maximization problem $\max_{X\in\mathscr{X}} \sum_{s=1}^{S} p_s u(x_s)$ is a standard optimization problem of maximizing a quasiconcave objective function over a convex polytope. The usual Karush–Kuhn–Tucker first-order optimality conditions are given by

$$\sum_{s=1}^{S} p_s u'(y_s) x_{i,s} \le c, \quad i = 1, \ldots, N. \tag{26}$$

The usual complementary slackness conditions apply: the inequalities (1) must be always binding for assets that are included in the evaluated portfolio ($\lambda_i > 0$) but they may be nonbinding for "inactive assets" ($\lambda_i = 0$). Aggregating the N inequalities using the portfolio weights, we find that

$$\sum_{i=1}^{N} \lambda_i \sum_{s=1}^{S} p_s u'(y_s) x_{i,s} = \sum_{i=1}^{N} \lambda_i c \tag{27}$$

$$\Leftrightarrow \sum_{s=1}^{S} p_s u'(y_s) y_s = c. \tag{28}$$

Equality (28) uses the complementary slackness conditions, $\sum_{i=1}^{N} \lambda_i x_{i,s} = y_s$ and $\sum_{i=1}^{N} \lambda_i = 1$. Subtracting the left-hand side (LHS) and right-hand side (RHS) of (28) from the LHS and RHS of (26), respectively, gives (1). □

Proof of Proposition 1. The proof follows from $[KLIC(p \mid q) > 0 \Leftrightarrow p \neq q]\ \forall\, p, q \in \mathscr{P}$, and the necessary and sufficient optimality conditions (Lemma 1). If $SIM(q) = 0$, then $p^* = q$ and the optimality conditions are satisfied under the reference distribution q for some admissible $u \in \mathscr{U}$ (efficiency). If $SIM(q) > 0$, then $p^* \neq q$ and the optimality conditions are violated under the reference distribution q for all admissible $u \in \mathscr{U}$ (inefficiency). □

Proof of Proposition 2. Let

$$\mathscr{H}_0^{=}(u)\colon \mathscr{G}(u) = \mathbf{0}_N,$$
$$\mathscr{H}_0^{=}\colon \mathscr{G}(u) = \mathbf{0}_N \quad u \in \mathscr{U},$$
$$SIM_T^{=}(u) := \min_{p\in\mathscr{P}_T}\{KLIC_T(p)\colon \hat{\mathscr{G}}_T(p,u) = \mathbf{0}_N\},$$
$$SIM_T^{=} := \min_{u\in\mathscr{U}} SIM_T^{=}(u).$$

The following inequalities prove (15):

$$\lim_{T\to\infty} \mathbb{P}(SIM_T \ge c(\alpha) \mid \mathscr{H}_0)$$
$$\le \lim_{T\to\infty} \mathbb{P}(SIM_T \ge c(\alpha) \mid \mathscr{H}_0^{=}) \tag{29}$$
$$\le \lim_{T\to\infty} \mathbb{P}(SIM_T^{=} \ge c(\alpha) \mid \mathscr{H}_0^{=}) \tag{30}$$
$$\le \lim_{T\to\infty} \mathbb{P}(SIM_T^{=}(u) \ge c(\alpha) \mid \mathscr{H}_0^{=}(u)) = \alpha \quad \forall\, u \in \mathscr{U}. \tag{31}$$

Inequality (29) uses that $\mathscr{H}_0^{=}$ is the least-favorable case that maximizes the p-value under $\mathscr{H}_0$; inequality (30) uses $SIM_T \leq SIM_T^{=}$; inequality (31) uses $SIM_T^{=} \leq SIM_T^{=}(u)$, $u \in \mathscr{U}$. Finally, if $u \in \mathscr{U}$ rationalizes the evaluated portfolio ($\mathscr{H}_0^{=}(u)$), then $(2M/B) \cdot SIM_T^{=}(u) = -2\ln(LR_T(p^*))$ is a BEL LR statistic, where $p^* \in \mathscr{P}_T$ is the implied PMF. Using Wilks' theorem, the LR statistic obeys an asymptotic central chi-squared distribution, that is,

$$(2M/B) \cdot SIM_T^{=}(u) \xrightarrow{d} \chi^2_{N-1}.$$

B^{-1} represents an adjustment factor to account for the use of overlapping blocks. The relevant number of degrees of freedom is $(N-1)$ because one of the binding moment equalities conditions (1) is always redundant.

We now turn to (16). For any given $u \in \mathscr{U}$, let

$$SIM_T^{\leq}(u) := \min_{p \in \mathscr{P}_T}\{KLIC_T(p)\colon \hat{\mathscr{G}}_T(u,p) \leq \mathbf{0}_N\}.$$

A standard BEL LR test for given $u \in \mathscr{U}$ yields, for sufficiently small $c > 0$,

$$\lim_{T\to\infty} \mathbb{P}(SIM_T^{\leq}(u) \geq c \mid \mathscr{H}_1) = 1$$
$$\Leftrightarrow \quad \lim_{T\to\infty} \mathbb{P}(SIM_T^{\leq}(u) < c \mid \mathscr{H}_1) = 0. \tag{32}$$

Post et al. (2015) show that $\mathscr{U}$ can be represented by piecewise-exponential functions with a different exponent for every subinterval. We may approximate $\mathscr{U}$ with a countable set $\hat{\mathscr{U}} \subset \mathscr{U}$ by partitioning the return interval and parameter space for the exponent. We can then approximate SIM_T with

$$\widehat{SIM}_T := \min_{u \in \hat{\mathscr{U}}}\{SIM_T^{\leq}(u)\}.$$

We can obtain an arbitrary good approximation, that is, $|\widehat{SIM}_T - SIM_T| < \varepsilon$, with $\varepsilon \in (0,c)$ for an arbitrary small number, with a sufficiently fine partition. It follows that

$$\lim_{T\to\infty} \mathbb{P}(SIM_T < y \mid \mathscr{H}_1)$$
$$\leq \lim_{T\to\infty} \mathbb{P}(\widehat{SIM}_T < y + \varepsilon \mid \mathscr{H}_1) \tag{33}$$
$$= \lim_{T\to\infty} \mathbb{P}\left(\bigcup_{u\in\hat{\mathscr{U}}}\{SIM_T^{\leq}(u) < y+\varepsilon\} \,\middle|\, \mathscr{H}_1\right) \tag{34}$$
$$\leq \lim_{T\to\infty} \sum_{u\in\hat{\mathscr{U}}} \mathbb{P}(SIM_T^{\leq}(u) < y + \varepsilon \mid \mathscr{H}_1) = 0 \tag{35}$$
$$\Leftrightarrow \quad \lim_{T\to\infty} \mathbb{P}(SIM_T \geq y \mid \mathscr{H}_1) = 1, \tag{36}$$

for every $y \in (0, c-\varepsilon)$. Inequality (33) follows from $\widehat{SIM}_T \geq SIM_T$; equality (34) and inequality (35) use the inclusion-exclusion principle for the union of sets and equality (32). Since $\lim_{T\to\infty}(2M/B)^{-1}c(\alpha) = 0$ for every $\alpha < 0.5$, we find

$$\lim_{T\to\infty} \mathbb{P}((2M/B) \cdot SIM_T > c(\alpha) \mid \mathscr{H}_1) = 1. \quad \square$$

Proof of Proposition 3. Our proof is based on the following arguments:

$$SIM_T = \min_{p\in\mathscr{P}_T}\left(-\ln(T) - T^{-1}\sum_{t=1}^{T}\ln(p_t)\right) \tag{37}$$
$$= -\ln(T) - T^{-1}\max_{(p,u)\in\mathscr{A}_T}\left(\sum_{t=1}^{T}\ln(p_t u'(y_t)) - \sum_{t=1}^{T}\ln(u'(y_t))\right) \tag{38}$$
$$= -\ln(T) - T^{-1}\left[\max_{(p,u)\in\mathscr{A}_T}\left(\sum_{t=1}^{T}\ln(p_t u'(y_t))\right) - \max_{u\in\mathscr{U}_T}\left(\sum_{t=1}^{T}\ln(u'(y_t))\right)\right] \tag{39}$$
$$= -\ln(T) - T^{-1}\left[\max_{\mathbf{d}\in\mathscr{D}_T}\left(\sum_{t=1}^{T}\ln(d_t)\right) - \max_{u\in\mathscr{U}_T}\left(\sum_{t=1}^{T}\ln(u'(y_t))\right)\right] \tag{40}$$
$$= -\ln(T) - \max_{\substack{(p,u)\in\mathscr{A}_T:\\ \mathbf{p}\otimes\nabla\mathbf{u}=\mathbf{d}_T^*}}\left(T^{-1}\sum_{t=1}^{T}\ln(p_t)\right) = SIM_{2;T}. \tag{41}$$

Equality (39) uses that $\max_{(p,u)\in\mathscr{A}_T}(\sum_{t=1}^{T}\ln(p_t u'(y_t))) = \max_{p\in\mathscr{P}_T}(\sum_{t=1}^{T}\ln(p_t)) + \max_{u\in\mathscr{U}_T}(\sum_{t=1}^{T}\ln(u'(y_t)))$ if $\mathscr{D}_T = D_T^{\perp}$. Equality (40) uses (18)–(19). In Equality (41), $\mathbf{d}_T^*$ uses the normalization $T^{-1}\sum_{t=1}^{T} d_t y_t = 1$; see (21). This normalization is harmless because the utility gradient $\nabla\mathbf{u}$ can be multiplied by an arbitrary positive scalar and, in addition, $u'(y_t) > 0$ and $y_t > 0$, $t = 1,\ldots,T$, so that $T^{-1}\sum_{t=1}^{T} d_t y_t = T^{-1}\sum_{t=1}^{T} p_t u'(y_t) y_t > 0$. $\square$

Proof of Proposition 4. LP problem (24) is a reduced and rewritten version of

$$\widehat{SIM}_{2;T} = \min\left(-\ln(T) - T^{-1}\sum_{t=1}^{T}\tilde{\pi}_t\right) \tag{42}$$
$$\text{s.t.}\quad \tilde{\pi}_t + \tilde{\beta}_{R(t)} = \ln(d_t^*), \quad t = 1,\ldots,T; \tag{43}$$
$$\tilde{\pi}_t \leq \ln(\hat{p}_{t,k}) + (\hat{p}_{t,k})^{-1}(\pi_t - \hat{p}_{t,k}), \quad t = 1,\ldots,T;\ k = 1,\ldots,K; \tag{44}$$
$$\tilde{\beta}_r = \sum_{s=r}^{T-1}\alpha_s(y^{[s+1]} - y^{[r]}) + \alpha_T, \quad r = 1,\ldots,T; \tag{45}$$
$$\pi_t = \sum_{b=\underline{b}_t}^{\bar{b}_t} B^{-1}\varpi_b, \quad t = 1,\ldots,T; \tag{46}$$
$$\sum_{t=1}^{T}\pi_t = 1; \tag{47}$$
$$\alpha_r \geq 0, \quad r = 1,\ldots,T-1; \tag{48}$$
$$\varpi_b \geq 0, \quad b = 1,\ldots,M. \tag{49}$$

We can derive (24) by substituting the RHS of (45) for $\tilde{\beta}_{R(t)}$, $t = 1,\ldots,T$ in (43) and the RHS of (46) for π_t, $t = 1,\ldots,T$ in (44), eliminate redundant variables and constraints, and rearrange the remaining constraints in canonical form for LP.

Assume that $p_t u'(y_t) = d_t^*$, $t = 1,\ldots,T$ for some $(p,u) \in \mathscr{A}_T$. Using Lemmas 2 and 3, the following specification for the model variables gives a feasible solution to (43)–(49):

$$\pi_t = p_t, \quad t = 1,\ldots,T;$$
$$\tilde{\pi}_t = \ln(p_t), \quad t = 1,\ldots,T;$$
$$\varpi_b = \rho_b, \quad b = 1,\ldots,B;$$
$$\tilde{\beta}_r = \ln(u'(y^{[r]})), \quad r = 1,\ldots,T;$$
$$\alpha_r = \frac{\ln(u'(y^{[r]})/u'(y^{[r+1]}))}{y^{[r+1]} - y^{[r]}} - \frac{\ln(u'(y^{[r+1]})/u'(y^{[r+2]}))}{y^{[r+2]} - y^{[r+1]}}, \quad r = 1,\ldots,T-2;$$
$$\alpha_{T-1} = \frac{\ln(u'(y^{[T-1]})/(u'(y^{[T]})))}{y^{[T]} - y^{[T-1]}};$$
$$\alpha_T = \ln(u'(y^{[T]})).$$

It follows that $\widehat{SIM}_{2;T} \leq SIM_{2;T}$. As we increase K, the PWL approximation to the logarithmic function improves and $\widehat{SIM}_{2;T}$ approaches $SIM_{2;T}$ from below. $\square$

References

Afriat SN (1967) The construction of a utility function from expenditure data. *Internat. Econom. Rev.* 8:67–77.

Afriat SN (1972) Efficiency estimation of production functions. *Internat. Econom. Rev.* 13:568–598.

Almeida C, Garcia R (2012) Assessing misspecified asset pricing models with empirical likelihood estimators. *J. Econometrics* 170:519–537.

Basso A, Pianca P (1997) Decreasing absolute risk aversion and option pricing. *Management Sci.* 43:206–216.

Bawa VS (1975) Optimal rules for ordering uncertain prospects. *J. Financial Econom.* 2:95–121.

Bawa VS, Bodurtha JN Jr, Rao MR, Suri HL (1985) On determination of stochastic dominance optimal sets. *J. Finance* 40: 417–431.

Carhart M (1997) On persistence in mutual fund performance. *J. Finance* 52:57–82.

Davidson R, Duclos J-Y (2013) Testing for restricted stochastic dominance. *Econometric Rev.* 32:84–125.

Hodder JE, Jackwerth JC, Kolokolova O (2015) Improved portfolio choice using second-order stochastic dominance. *Rev. Finance* 19:1623–1647.

Kaur A, Prakasa Rao, BL, Singh H (1994) Testing for second-order stochastic dominance of two distributions. *Econometric Theory* 10:849–866.

Keppe HJ, Weber M (1989) Stochastic dominance with incomplete information on probabilities. *Eur. J. Oper. Res.* 43:350–355.

Kitamura Y (1997) Empirical likelihood methods with weakly dependent processes. *Ann. Statist.* 25:2084–2102.

Kopa M, Post T (2015) A general test for SSD portfolio efficiency. *Oper. Res. Spectrum* 37:703–734.

Kroll Y, Levy H (1980) Sampling errors and portfolio efficient analysis. *J. Financial Quant. Anal.* 15:655–688.

Kuosmanen T (2004) Efficient diversification according to stochastic dominance criteria. *Management Sci.* 50:1390–1406.

Kuosmanen T, Post T, Scholtes S (2007) Testing for productive efficiency in case of errors-in-variables. *J. Econometrics* 136: 131–162.

Levy H (2006) *Stochastic Dominance: Investment Decision Making Under Uncertainty*, 2nd ed. (Springer Science, New York).

Linton OB, Post T, Whang Y-J (2014) Testing for the stochastic dominance efficiency of a given portfolio. *Econometrics J.* 17: 59–74.

Nelson RD, Pope R (1991) Bootstrapped insights into empirical applications of stochastic dominance. *Management Sci.* 37: 1182–1194.

Newey WK, Smith RJ (2004) Higher order properties of GMM and generalized empirical likelihood estimators. *Econometrica* 72:219–255.

Owen AB (2001) *Empirical Likelihood* (CRC Press, New York).

Pearman AD, Kmietowicz ZW (1986) Stochastic dominance with linear partial information. *Eur. J. Oper. Res.* 23:57–63.

Post T (2003) Empirical tests for stochastic dominance efficiency. *J. Finance* 58:1905–1932.

Post T, Versijp P (2007) Multivariate tests for stochastic dominance efficiency of a given portfolio. *J. Financial Quant. Anal.* 42:489–515.

Post T, Fang Y, Kopa M (2015) Linear tests for decreasing absolute risk aversion stochastic dominance. *Management Sci.* 61: 1615–1629.

Roman D, Darby-Dowman K, Mitra G (2006) Portfolio construction based on stochastic dominance and target return distributions. *Math. Programming* 108:541–569.

Roman D, Mitra G, Zverovich V (2013) Enhanced indexation based on second-order stochastic dominance. *Eur. J. Operational Res.* 228:273–281.

Russell WR, Seo TK (1989) Representative sets for stochastic dominance rules. Fomby TB, Seo TK eds. *Studies in the Economics of Uncertainty: In Honor of Josef Hadar* (Springer Verlag, New York), 59–76.

Scaillet O, Topaloglou N (2010) Testing for stochastic dominance efficiency. *J. Bus. Econom. Statist.* 28:169–180.

Shalit H, Yitzhaki S (1994) Marginal conditional stochastic dominance. *Management Sci.* 40:670–684.

Varian HR (1983) Nonparametric tests of models of investor behavior. *J. Financial Quant. Anal.* 18:269–278.

Varian HR (1985) Nonparametric analysis of optimizing behavior with measurement error. *J. Econometrics* 30:445–458.

Vickson RG (1975) Stochastic dominance tests for decreasing absolute risk aversion. I. discrete random variables. *Management Sci.* 21:1438–1446.