



## **Department of Economics**

**Athens University of Economics and Business**

**WORKING PAPER no. 07-2024**

### **Norm Constrained Empirical Portfolio Optimization with Stochastic Dominance: Robust Optimization Non-Asymptotics**

**Stelios Arvanitis**

**July 2024**

The Working Papers in this series circulate mainly for early presentation and discussion, as well as for the information of the Academic Community and all interested in our current research activity.

The authors assume full responsibility for the accuracy of their paper as well as for the opinions expressed therein.

# Norm Constrained Empirical Portfolio Optimization with Stochastic Dominance: Robust Optimization Non-Asymptotics

*Stelios Arvanitis\**

## Abstract

The present note provides an initial theoretical explanation of the way norm regularizations may provide a means of controlling the non-asymptotic probability of False Dominance classification for empirically optimal portfolios satisfying empirical Stochastic Dominance restrictions in an iid setting. It does so via a dual characterization of the norm-constrained problem, as a problem of Distributional Robust Optimization. This enables the use of concentration inequalities involving the Wasserstein distance from the empirical distribution, to obtain an upper bound for the non-asymptotic probability of False Dominance classification. This leads to information about the minimal sample size required for this probability to be dominated by a predetermined significance level.

*JEL Codes:* C44, C58, G11. *MSC2020:* 62C, 62P20, 62E17.

*Keywords:* Portfolio optimization, Stochastic dominance,  $\ell_p$  norm regularization, Convex Duality, Wasserstein Distance, Distributionally robust optimization, Concentration inequality, Non-asymptotics, False dominance classification.

## 1 Introduction

Portfolio optimization based on stochastic dominance (SD) restrictions is a non-parametric generalization of the standard mean-variance approach, for optimal portfolio weights' selection regarding investment strategies outside the realm of satiation and/or elliptical returns' distributions.

---

\*Department of Economics, Athens University of Economics and Business, Greece. Email: stelios@aueb.gr

A recent literature in operations research and econometrics have addressed the analytical challenges for numerical optimization and statistical inference, largely making SD portfolio optimization empirically applicable. Indicative such applications include Hodder et al. (2015) (10); Constantinides et al. (2020) (8); Post and Rodriguez-Longarela (2018) (16), Arvanitis and Post (2024) (1). There portfolio selection is usually performed via the optimization of an empirical criterion under the constraint that the choice set is comprised by portfolios that empirically dominate a benchmark portfolio. The empirically optimal portfolio by construction dominates the benchmark in the sample, it is however susceptible to the decision error of False Dominance (FD) classification in the population.

Under general sampling schemes this decision error becomes asymptotically negligible. Controlling the probability of this error for fixed sample size is however important in applications, especially when the sample size is not particularly large compared to the dimensionality of the portfolios considered. In some of the applications mentioned above, a heuristic used in the underlying empirical optimization seems to improve the out-of-sample properties of the optimal portfolio. The optimization problem is augmented by a restriction on the distance of the portfolio sought compared to the benchmark.

The present note provides an initial theoretical explanation of the way suchlike restrictions may provide a means of controlling the fixed sample size probability of FD classification in an iid setting. It does so first, via the conditional on particular events characterization of the norm-constrained problem, as a problem of Distributional Robust Optimization (DRO). There strong convex duality results (see Gao, Chen, and Kleywegt (2017) (9)) enable the representation of the original problem as a problem of conservative optimization over a Wasserstein ball centered at the empirical distribution, as long as the criterion has a Lipschitz continuity property. Second, using this characterization along with concentration inequalities involving the Wasserstein distance from the empirical distribution, an upper bound for the non-asymptotic probability of FD classification is obtained, leading among others to considerations about the minimal sample size required for this probability to be dominated by a predetermined significance level.

The structure of the note is the following: the second section analyzes the SD framework for portfolio optimization and its' regularized formulation. The third section derives the DRO characterization, the non-asymptotic bounds for the probability of FD, and briefly discusses some paths for future research.

## 2 Stochastic dominance and portfolio optimization framework

$(\mathbf{X}_t)_{t \in \mathbb{Z}}$  is a stationary process with values in some subset of  $\mathbb{R}^d$ . The random vector  $\mathbf{X}_t$  represents the one period stationary returns of  $d$  financial assets, and  $\mathcal{X} \subset \mathbb{R}^d$  is the pointwise bounded from below support of its' latent joint distribution. Boundedness from above is considered plausible for moderate observation frequencies. The researcher has at her disposal an observable sample from the process,  $(\mathbf{X}_t)_{t=1, \dots, T}$ ;  $\mathbb{P}_T$  denotes the empirical distribution of the sample.

A portfolio on  $\mathbf{X}_0$  is any real linear function on  $\mathbb{R}^d$ ; the elements of its representing vector are the portfolio weights. Alternative portfolios are evaluated inside the expected utility paradigm, using utility functions  $u : \mathcal{X} \rightarrow \mathbb{R}$  that are increasing, continuous, and concave. These populate the closed (in the topology of uniform convergence on compacta), convex set  $\mathcal{U}_2$  while  $\mathcal{U}_2^*$  denotes the set of Russell-Seo utilities-see Russell and Seo (1989) (18), i.e. the utilities of the form  $u(x) = -(z-x)_+$ ,  $z \in \mathcal{X}$ , those constitute the extreme points of  $\mathcal{U}_2$ ;  $\mathcal{U}_2^* - \{0\}$  does not contain the constant utility at zero-corresponding to the threshold that equals the pointwise infimum of  $\mathcal{X}$ .  $\mathcal{U}_2 - \mathcal{U}_2^-$  analogously denotes  $\mathcal{U}_2$  without constant utilities.

The analysis involves a set of portfolios  $\Lambda \subseteq \mathbb{R}^d$ . It is considered convex and compact. In what follows  $\lambda$ ,  $\tau$  denote respectively a typical element of  $\Lambda$  and a distinguished benchmark portfolio inside  $\Lambda$ .

The above enable the definition of a stochastic dominance relation on the sets of prospects, via  $\mathcal{U}_2$ : in the stationary framework considered,  $\lambda$  is said to dominate  $\tau$  w.r.t. the utility class  $\mathcal{U}_2$  iff  $D(u, \tau, \lambda, \mathbb{P}) := \mathbb{E}_{\mathbb{P}}(u(\lambda' \mathbf{X}_0)) - \mathbb{E}_{\mathbb{P}}(u(\tau' \mathbf{X}_0)) \geq, \forall u \in \mathcal{U}_2$ -here  $\mathbb{E}_{\mathbb{P}}$  denotes integration w.r.t.  $\mathbb{P}$ . Thus,  $\lambda$  is preferred over  $\tau$  by every utility in the considered class, this is what is known in the literature as the second order stochastic dominance of  $\lambda$  over the benchmark  $\tau$ ;  $\lambda \underset{\mathbb{P}, 2}{\succeq} \tau$ . The definition remains invariant if  $\mathcal{U}_2$  is replaced by  $\mathcal{U}_2^*$ -see Russell and Seo (1989) (18). Furthermore,  $\Lambda_{\mathbb{P}}^{\succeq} := \left\{ \lambda \in \Lambda; \lambda \underset{\mathbb{P}, 2}{\succeq} \tau \right\}$  is the non-empty convex set of portfolios that dominate the benchmark in the population. Non-emptiness holds due to reflexivity of the dominance relation, and convexity follows from the concavity of the utility functions at hand, the linearity of the portfolio formation and the monotonicity of the integral. Substituting the latent  $\mathbb{P}$  with  $\mathbb{P}_T$  in  $\underset{\mathbb{P}, 2}{\succeq}$ , the empirical analogue  $\Lambda_{\mathbb{P}_T}^{\succeq}$  is obtained.

Consider a choice  $\lambda_{\mathbb{P}_T} \in \Lambda_{\mathbb{P}_T}^{\succeq}$ . Controlling the probability of FD for  $\lambda_{\mathbb{P}_T}$ , i.e.  $\mathbb{P}(\lambda \underset{\mathbb{P}_T, 2}{\succeq} \tau) / \lambda \not\underset{\mathbb{P}, 2}{\succeq} \tau$ , can be of particular empirical interest, as FD can lead to suboptimal portfolio choices. This can asymptotically-as  $T \rightarrow \infty$ -vanish as long as the probabilistic properties of the sampling scheme ensure that  $\mathbb{P}_T \rightsquigarrow \mathbb{P}$  where

$\rightsquigarrow$  denotes convergence in distribution and the existence of some  $\epsilon > 0$  such that  $\sup_T \mathbb{E}_{\mathbb{P}_T}(\|\mathbf{X}_0\|_2^{1+\epsilon}) < +\infty$ , where  $\|\cdot\|_2$  denotes the Euclidean norm. This is due to that  $\mathcal{U}_2^*$  is uniformly Lipschitz, and due to uniform integrability. The weak convergence is easily establishable in contexts of stationarity and ergodicity for  $(\mathbf{X}_t)_{t \in \mathbb{Z}}$ ; boundedness for the  $1 + \epsilon$  moments of the normed process would follow in this stationary and ergodic context along which a financially plausible assumption of  $\mathbb{E}_{\mathbb{P}}(\|\mathbf{X}_0\|_2^{1+\epsilon}) < +\infty$ . The question of controlling this probability is also of interest for fixed-and potentially realistically large enough  $T$ . This is what is investigated in the subsequent analysis.

Every choice  $\boldsymbol{\lambda}_{\mathbb{P}} \in \boldsymbol{\Lambda}_{\mathbb{P}}^{\succ}$  can be represented as a solution-albeit in some cases trivial-of the optimization problem  $\max_{\boldsymbol{\lambda} \in \boldsymbol{\Lambda}_{\mathbb{P}}^{\succ}} \mathbb{E}_{\mathbb{P}}(u(\boldsymbol{\lambda}'\mathbf{X}_0))$  for some  $u \in \mathcal{U}_2$ . More importantly for a given non-constant  $u \in \mathcal{U}_2$  any solution, say  $\boldsymbol{\lambda}(u, \mathbb{P})$ , to the optimization problem  $\max_{\boldsymbol{\lambda} \in \boldsymbol{\Lambda}_{\mathbb{P}}^{\succ}} \mathbb{E}_{\mathbb{P}}(u(\boldsymbol{\lambda}'\mathbf{X}_0))$  can be of economic interest; any such latent portfolio is perceivable as the best a risk averter investor with preferences represented by the particular  $u$  can achieve in terms of expected utility, if she insists on working with portfolios that would be weakly preferred by every risk averter to the benchmark. This is a problem of portfolio optimization augmented with stochastic dominance (second order) SD conditions.

Latency of  $\mathbb{P}$  implies generally latency of  $\boldsymbol{\lambda}(u, \mathbb{P})$ . The latter can be statistically approximated by its empirical analogue;  $\boldsymbol{\lambda}(u, \mathbb{P}_T)$ , i.e. the solution to the empirical portfolio optimization augmented with empirical stochastic dominance conditions  $\max_{\boldsymbol{\lambda} \in \boldsymbol{\Lambda}_{\mathbb{P}_T}^{\succ}} \mathbb{E}_{\mathbb{P}_T}(u(\boldsymbol{\lambda}'\mathbf{x}))$ . Hence the analysis that follows considers an arbitrary yet fixed  $u$  and asks whether there is a modification of the optimization problem that enables the non-asymptotic investigation of the probability of FD for its solutions.

## 2.1 Regularized formulation of portfolio optimization

A modification used in practice augments the expected utility criterion with an additive regularization term that depends on the  $\ell^p$  distance between the portfolio sought and the benchmark. The intuition is that when the (Lagrange) multiplier of the aforementioned distance is chosen optimally, then in order for a portfolio that lies "away" from the benchmark to solve the optimization problem, it would have to "strongly" satisfy the empirical dominance conditions at least in some neighborhood of  $u$ .

As mentioned above, the  $\ell_p$ -distance from the benchmark portfolio weights is considered here,  $\|\boldsymbol{\lambda} - \boldsymbol{\tau}\|_p := (\sum_{i=1}^d |\lambda_i - \tau_i|^p)^{1/p}$ , for the case where  $p \geq 1$ , and  $\max_{i=1, \dots, d} |\lambda_i - \tau_i|$  for  $p = +\infty$ . The regularized optimization portfolio is then de-

fined by:

$$\boldsymbol{\lambda}(u, \mathbb{P}_T, p, \xi_T) \in \arg \max_{\boldsymbol{\Lambda}_{\mathbb{P}_T^{\xi_T}}} (\mathbb{E}_{\mathbb{P}_T}(u(\boldsymbol{\lambda}'\mathbf{x})) - \xi_T \|\boldsymbol{\lambda} - \boldsymbol{\tau}\|_p), \quad (1)$$

where the random variable  $\xi_T \geq 0$  assumes the role of the regularization multiplier. Its optimal selection is expected to influence the non-asymptotic properties of the probability of FD. The modified problem additionally thus depends on both the choice of the multiplier  $\xi_T$  and the norm order  $p$ .

It is possible that  $\boldsymbol{\Lambda}$  already includes (explicit or implicit) pre-existing norm constraints; then the analysis studies the effect on the probability of FD of tightening of the existing norm constraints.

### 3 Results

The issue of the derivation of non-asymptotic properties for the portfolio solutions of the empirical regularized problem is considered here, with a view towards the fixed  $T$  properties of the probability of FD.

In the first part of this section the regularized problem is translated to a problem of distributionally robust optimization (DRO) using a dual formulation of regularized problems involving Lipschitz criteria to problems of conservative optimization in Wasserstein neighborhoods of  $\mathbb{P}_T$ . There, the Lagrange multiplier  $\xi_T$  plays an important role in the DRO formulation below, as it determines the radius of pessimism and at least locally the slacks for the dominance constraints.

In the second part this conservative representation is exploited in order to non-asymptotically bound the probability of FD characterization via concentration inequalities.

Some further notation will be also useful: for  $\mathbb{Q}$  an arbitrary distribution on  $\mathbb{R}^d$ , and  $q(p)$  such that  $\frac{1}{p} + \frac{1}{q} = 1$ , the first Wasserstein distance between  $\mathbb{Q}$  and the empirical distribution  $\mathbb{P}_T$  is defined by  $\mathcal{W}(\mathbb{P}_T, \mathbb{Q}; p) := \min_{\gamma \in \Gamma(\mathbb{P}_T, \mathbb{Q})} \int_{\mathbb{R}^d \times \mathbb{R}^d} \|z - z^*\|_q d\gamma(z, z^*)$ , where  $\Gamma(\mathbb{P}_T, \mathbb{Q})$  denotes the set of Borel probability distributions on  $\mathbb{R}^d \times \mathbb{R}^d$  that have respective "marginals"  $\mathbb{P}_T, \mathbb{P}$ , and also have finite first moment (see Gao, Chen, and Kleywegt (2017) (9)).  $\mathcal{W}$  metrizes weak convergence (see Rahimian and Mehrota (2019) (17)). For  $\epsilon > 0$ ,  $\mathcal{M}_p(\mathbb{P}_T, \epsilon) := \{\mathbb{P} : \mathcal{W}(\mathbb{P}_T, \mathbb{Q}; p) \leq \epsilon\}$  is the Wasserstein closed ball centered at  $\mathbb{P}_T$  with radius  $\epsilon$ . Also,  $d(p) := d^l$ ,  $l := \max(\frac{1}{2} - \frac{1}{p}, \frac{1}{p} - \frac{1}{2}) =$

$$\begin{cases} \frac{1}{p} - \frac{1}{2}, & p \leq 2 \\ \frac{1}{2} - \frac{1}{p}, & p > 2. \end{cases}$$

### 3.1 DRO formulation

A characterization of the regularized problem in (1) as a distributionally robust optimization (DRO) problem is obtained here. It is based on strong convex duality results for robust optimization (see for example Lemma 1 of Gao, Chen, and Kleywegt (2017) (9)).

The event  $\mathcal{E}_1 := \{\exists \boldsymbol{\lambda}(u, \mathbb{P}_T, p, \xi_T) : \inf_{\mathbb{Q} \in \mathcal{M}_p(\mathbb{P}_T, \xi_T)} D(u, \boldsymbol{\tau}, \boldsymbol{\lambda}(u, \mathbb{P}_T, \xi_T), \mathbb{Q}) > 0\}$  characterizes the samples for which the empirical regularized program has non-trivial solutions. Furthermore, for every sample realization inside the event defined as  $\mathcal{E}_2 := \{\forall v \in \mathcal{U}_2^* - \{0\}, \forall \boldsymbol{\lambda}(v, \mathbb{P}_T, p, \xi_T) : \inf_{\mathbb{Q} \in \mathcal{M}_p(\mathbb{P}_T, \xi_T)} D(v, \boldsymbol{\tau}, \boldsymbol{\lambda}(v, \mathbb{P}_T, \xi_T), \mathbb{Q}) > 0\}$ , every choice of the objective utility results to non-trivial empirical solutions. It is noted that  $\mathcal{E}_2$  is a subset of  $\mathcal{E}_1$  due to the Russell-Seo utility representation-see Russell and Seo (1989) (18). Then:

**Proposition 1.** (*Distributional Robustness*). *Suppose that  $u$  has a unital Lipschitz coefficient. Then:*

1. *Conditional on  $\mathcal{E}_1$ , there exists some non-empty open subset  $\mathcal{U}_2(u)$  of  $\mathcal{U}_2$ , for which Problem (1) is equivalent to:*

$$\max_{\boldsymbol{\Lambda}_{\mathbb{P}_T}^{\succeq} \cap \boldsymbol{\Lambda}_u^{\succeq}} \inf_{\mathbb{Q} \in \mathcal{M}_p(\mathbb{P}_T, \xi_T)} \mathbb{E}_{\mathbb{Q}}(u(\boldsymbol{\lambda}' \boldsymbol{x})); \quad (2)$$

$$\boldsymbol{\Lambda}_u^{\succeq} := \left\{ \boldsymbol{\lambda} \in \Lambda : \inf_{\mathbb{Q} \in \mathcal{M}_p(\mathbb{P}_T, \xi_T)} D(v, \boldsymbol{\tau}, \boldsymbol{\lambda}, \mathbb{Q}) > 0, \forall v \in \mathcal{U}_2(v) \right\}. \quad (3)$$

2. *Conditional on  $\mathcal{E}_2$ , Problem (1) is equivalent to:*

$$\sup_{\boldsymbol{\Lambda}_{\xi_T}^{\succeq}} \inf_{\mathbb{Q} \in \mathcal{M}_p(\mathbb{P}_T, \xi_T)} \mathbb{E}_{\mathbb{Q}}(u(\boldsymbol{\lambda}' \boldsymbol{x})); \quad (4)$$

$$\text{where } \boldsymbol{\Lambda}_{\xi_T}^{\succeq} := \left\{ \boldsymbol{\lambda} \in \Lambda : \inf_{\mathbb{Q} \in \mathcal{M}_p(\mathbb{P}_T, \xi_T)} D(v, \boldsymbol{\tau}, \boldsymbol{\lambda}, \mathbb{Q}) > 0, \forall v \in \mathcal{U}_2^* - \{0\} \right\}. \quad (5)$$

*Proof.* Theorem 1 of Gao, Chen, and Kleywegt (2017) (9) implies that

$$\begin{aligned} G_{\mathbb{P}_T}(\boldsymbol{\tau} + (\boldsymbol{\lambda} - \boldsymbol{\tau})) - \xi_T(c) \|\boldsymbol{\lambda} - \boldsymbol{\tau}\|_p &= \inf_{\mathbb{Q} \in \mathcal{M}_p(\mathbb{P}_T, \xi_T)} G_{\mathbb{Q}}(\boldsymbol{\tau} + (\boldsymbol{\lambda} - \boldsymbol{\tau})) \\ D(u, \boldsymbol{\tau}, \boldsymbol{\tau} + (\boldsymbol{\lambda} - \boldsymbol{\tau}), \mathbb{P}_T) &= \inf_{\mathbb{Q} \in \mathcal{M}_p(\mathbb{P}_T, \xi_T)} D(u, \boldsymbol{\tau}, \boldsymbol{\tau} + (\boldsymbol{\lambda} - \boldsymbol{\tau}), \mathbb{Q}) + \xi_T \|\boldsymbol{\lambda} - \boldsymbol{\tau}\|_p, \end{aligned}$$

for  $G_{\mathbb{P}}(\boldsymbol{\lambda}) := \mathbb{E}_{\mathbb{P}}(u(\boldsymbol{\lambda}' \boldsymbol{X}))$ . This and translation in  $\mathbb{R}^d$  by  $\boldsymbol{\tau}$ , directly imply that Problem (1) is equivalent to  $\max_{\boldsymbol{\Lambda}} \inf_{\mathbb{Q} \in \mathcal{M}_p(\mathbb{P}_T, \xi_T)} G_{\mathbb{Q}}(\boldsymbol{\lambda})$ . Using a scaling and translation



argument that is allowed for von Neyman-Morgenstern preferences,  $\mathcal{U}_2$  can be chosen uniformly bounded due to the uniform Lipschitz property of  $\mathcal{U}_2^*$  and the Russell-Seo representation-see Russell and Seo (1989) (18). This and Lemma 2.7.5 of van der Vaart and Wellner (1996) (21) along with the compactness of the support, imply that  $\mathcal{U}_2$  is totally bounded, and thereby equicontinuous. This and uniform integrability then imply that  $D(u, \boldsymbol{\tau}, \boldsymbol{\lambda}, \mathbb{P}) - \xi_T \|\boldsymbol{\lambda} - \boldsymbol{\tau}\|_p$  is jointly continuous in  $(u, \boldsymbol{\lambda})$ . Since there exists a non trivial solution to (1), it must satisfy  $D(u, \boldsymbol{\tau}, \boldsymbol{\lambda}, \mathbb{P}_T) \geq \xi_T \|\boldsymbol{\lambda} - \boldsymbol{\tau}\|_p$ . This then implies the existence of  $\mathcal{U}_2(v)$ .  $\mathcal{U}_2(v)$  is open as a union of open sets; it is obtained as the inverse image of  $(0, +\infty)$  w.r.t. the regularized criterion evaluated at the solution, on  $\mathcal{U}_2$ . Finally, similarly to Theorem 1 of Arvanitis and Post (2024) (2), for the portfolio defined as  $\boldsymbol{\gamma} = \boldsymbol{\gamma}_{(u, \boldsymbol{\tau})}((1 - \delta^*)w_u + \delta^*w)$ , with  $\boldsymbol{\gamma}_{(u, \boldsymbol{\tau})}(w) := \int_{\mathcal{U}_2^* - \{0\}} \boldsymbol{\lambda}(u, \mathbb{P}_T, p, \xi_T) dw(v)$ , where  $w$  lies in the set of non-degenerate Borel measures on  $\mathcal{U}_2$ , and  $w_u$  the degenerate measure at  $u$ ,  $\delta^* \in (0, 1)$ , we have that  $\boldsymbol{\gamma} \in \boldsymbol{\Lambda}_{\xi_T}^{\succ}$  due to the concavity of the utilities involved; then by the definition of the portfolios involved, and due the Lipschitz continuity of  $D(u, \boldsymbol{\tau}, \cdot, \mathbb{P}_T)$  and of the  $\ell_p$  norm, setting  $\boldsymbol{\lambda}_T := \boldsymbol{\lambda}(u, \mathbb{P}_T, p, \xi_T)$  it is obtained that

$$\begin{aligned} 0 &\leq \mathbb{E}_{\mathbb{P}_T}(u(\boldsymbol{\lambda}'_T - \xi_T \|\boldsymbol{\lambda}_T - \boldsymbol{\tau}\|_p - \mathbb{E}_{\mathbb{P}_T}(u(\boldsymbol{\gamma}' \mathbf{X}))) + \xi_T \|\boldsymbol{\gamma} - \boldsymbol{\tau}\|_p \\ &\leq \delta^* (\sup_{\mathbb{Q} \in \mathcal{M}_p(\mathbb{P}_T, \xi_T)} \mathbb{E}_{\mathbb{Q}}(\|\boldsymbol{x}\|_2) + \xi_T d^{\frac{1}{p} - \frac{1}{2}}) \text{diam}(\boldsymbol{\Lambda}), \end{aligned}$$

and (4) holds since  $\delta^*$  can be chosen arbitrarily small.  $\square$

Lipschitz coefficient unitarity holds for example in the case of portfolio choice via maximization of expected return; then  $u$  is the identity. More generally, given that utility rescaling does not affect preferences and optimal choice, if  $u$  is non-trivial and has a bounded derivative, then the Lipschitz coefficient can be always be equal to one. The regularized version of the objective function is equal, due to duality, to a robust expected value of  $u(\cdot' \boldsymbol{x})$ ; actually this is the most conservative expectation over the Wasserstein ball centered at the ecdf, with radius formed by the Lagrange multiplier. The dual representation of  $\mathcal{W}$  (see Kantorovich (1958) (11)) implies that  $\mathcal{M}_p(\mathbb{P}_T, \xi_T)$  is convex; then Sion's Minimax Theorem-see Sion (1958) (19)-implies that  $\max_{\boldsymbol{\Lambda}_{\mathbb{P}_T}^{\succ}} \inf_{\mathbb{Q} \in \mathcal{M}_p} \mathbb{E}_{\mathbb{Q}}(u(\boldsymbol{\lambda}' \mathbf{X})) = \inf_{\mathbb{Q} \in \mathcal{M}_p} \max_{\boldsymbol{\Lambda}_{\mathbb{P}_T}^{\succ} } \mathbb{E}_{\mathbb{Q}}(u(\boldsymbol{\lambda}' \mathbf{X}))$ . This conforms to the solution characterization as the most conservative (over the aforementioned ball) maximizer of expected utility that stochastically dominates the benchmark.

Due to the equi-continuity properties of  $\mathcal{U}_2$ , and if there exists an empirically optimal portfolio that strictly dominates the benchmark, i.e.  $\mathcal{E}_1$  holds, the empirical problem has a representation where the regularization term, permeates as a positive slack to a neighborhood of SD conditions around  $D(u, \boldsymbol{\tau}, \boldsymbol{\lambda}, \mathbb{P}_T)$ ; each of these conditions then, by the same duality property, has a conservative characterization given in

(3). Hence, regularizing the objective implies a local regularization for the dominance conditions. Under the stronger  $\mathcal{E}_2$ , the representation uses the regularization term as a positive slack on the totality of non-trivial SD conditions. This implies that there exist elements of  $\boldsymbol{\lambda}(u, \mathbb{P}_T, p, \xi_T)$  that satisfy enhanced versions of the SD inequalities, resulting into strong properties regarding the probability of FD classifications, as the following result reveals. It is noted that due to the Russell-Seo representation-see Russell and Seo (1989) (18), the linearity of  $D$  w.r.t.  $u$  and the properties of the infimum:

$$\boldsymbol{\Lambda}_{\xi_T}^{\succeq} := \left\{ \boldsymbol{\lambda} \in \Lambda : \inf_{\mathbb{Q} \in \mathcal{M}_p(\mathbb{P}_T, \xi_T)} D(v, \boldsymbol{\tau}, \boldsymbol{\lambda}, \mathbb{Q}) > 0, \forall v \in \mathcal{U}_2 - \mathcal{U}_2^{\neq} \right\}.$$

Under  $\mathcal{E}_2$  the regularization permeates every SD condition involving a non-constant utility.

### 3.2 Non-asymptotic bounds for the False Dominance probability

In order to derive non asymptotic bounds for the probability of selecting a portfolio that does not dominate the benchmark in the population (FD), the event  $\mathcal{E}_3 := \{\exists \boldsymbol{\lambda} \in \boldsymbol{\Lambda}_{\mathbb{P}_T}^{\succeq} : D(u, \boldsymbol{\tau}, \boldsymbol{\lambda}, \mathbb{P}) < 0\}$  is also utilized: it corresponds to the samples for which FD classifications actually occurs, when regularization does not permeate SD inequalities. Furthermore, for  $\tau > 0$  let  $h(\tau) := \frac{1 + \ln \mathbb{E}_{\mathbb{P}}[\exp(\tau \|\mathbf{x}\|_2^2)]}{\tau}$ , and for  $m > 2$  let  $C(d) := 2 \times 3^{d - \log_3(d)} \mathbb{I}(d - \log_3(d) < \frac{\log_3(T)}{2}) + 4 \mathbb{E}_{\mathbb{P}}[\|\mathbf{x}\|_2^m] \mathbb{I}(d - \log_3(d) \geq \frac{\log_3(T)}{2})$ , where  $\mathbb{I}$  denotes the indicator function. Then the following result is obtained via the use of concentration inequalities involving the Wasserstein distance from the empirical distribution:

**Proposition 2** (False Dominance Classification and Opportunity Loss). *Suppose that  $(\mathbf{X}_t)_{t \in \mathbb{Z}}$  is iid, that  $d > 2$ , that for some  $\tau > 0$ ,  $\mathbb{E}_{\mathbb{P}}(\exp(\tau \|\mathbf{x}\|_2^2)) < +\infty$ , and that  $u$  has unital Lipschitz coefficient. Then,*

1. for any  $T \geq 1$ , and if  $\xi_T > 2C(d)d(p) \sup_{\Lambda} \|\boldsymbol{\lambda}\|_2 T^{-\frac{1}{d}}$

$$\mathbb{P}(\boldsymbol{\Lambda}_{\xi_T}^{\succeq} - \boldsymbol{\Lambda}_{\mathbb{P}}^{\succeq} \neq \emptyset \mid \mathcal{E}_2) \leq \exp \left( - \frac{\left(1 - C(d)T^{-\frac{d+1}{d}}\right)^2 T \xi_T^2}{8d^2(p) \inf_{\tau > 0} h^2(\tau) \sup_{\Lambda} \|\boldsymbol{\lambda}\|_2^2} \right). \quad (6)$$

Consequently, if the event  $\mathcal{E}_2$  holds w.h.p. and  $T \xi_T^2 \rightarrow \infty$ , then the probability of FD classification for the elements of  $\boldsymbol{\Lambda}_{\xi_T}^{\succeq}$  converges to zero.

2. There exists some  $T_0 \geq 1$  such that for any  $T \geq T_0$ , if  $\xi_T > 2d(p) \sup_{\Lambda} \|\boldsymbol{\lambda}\|_2 \max \left\{ C(d)T^{-\frac{1}{d}}, \sqrt{2} \inf_{\tau>0} h(\tau) \sqrt{\frac{\ln T}{T}} \right\}$ ,

$$\mathbb{P}(\mathcal{E}_3 \setminus \mathcal{E}_2) > \mathbb{P}(\Lambda_{\xi_T}^{\succ} - \Lambda_{\mathbb{P}}^{\succ} \neq \emptyset \setminus \mathcal{E}_2). \quad (7)$$

3. Finally, if  $\boldsymbol{\lambda}^*$  denotes the optimal solution to the population problem and  $\boldsymbol{\lambda}_T := \boldsymbol{\lambda}(u, \mathbb{P}_T, p, \xi_T)$ , then for any for any  $T \geq 1$ , and if  $\xi_T > C(d)T^{-\frac{1}{d}}$ ,

$$\begin{aligned} & \mathbb{P}(|\mathbb{E}_{\mathbb{P}_T}(u(\boldsymbol{\lambda}'_T \mathbf{x})) - \mathbb{E}_{\mathbb{P}}(u(\boldsymbol{\lambda}'^* \mathbf{x}))| > d(p) \mathbb{E}_{\mathbb{P}_T} \|\mathbf{x}\|_2 \|\boldsymbol{\lambda}_T - \boldsymbol{\lambda}^*\|_p + \|\boldsymbol{\lambda}^*\|_2 \xi_T) \\ & \leq \exp \left( - \frac{\left(1 - C(d)T^{-\frac{d+1}{d}}\right)^2 T \xi_T^2}{2 \inf_{\tau>0} h^2(\tau)} \right). \end{aligned} \quad (8)$$

*Proof.* Conditional on  $\mathcal{E}_2$ , the event  $\Lambda_{\xi_T}^{\succ} - \Lambda_{\mathbb{P}}^{\succ} \neq \emptyset$  is equivalent to the inequalities' system  $\inf_{\Lambda_{\xi_T}^{\succ}, \mathcal{U}_2^*} \inf_{\mathbb{P} \in \mathcal{M}_p(\mathbb{P}_T, \xi_T)} D(u, \boldsymbol{\tau}, \boldsymbol{\lambda}, \mathbb{P}) \geq 0$  and  $\inf_{\Lambda_{\xi_T}^{\succ}, \mathcal{U}_2^*} D(u, \boldsymbol{\tau}, \boldsymbol{\lambda}, \mathbb{P}) < 0$ . Due to the dual representation of  $\mathcal{W}$  (see Kantorovich (1958)), the Hilbert projection Theorem, and the equivalence between the  $\ell_p$  norms in  $\mathbb{R}^d$ ,

$$\begin{aligned} & \mathbb{P}(\Lambda_{\xi_T}^{\succ} - \Lambda_{\mathbb{P}}^{\succ} \neq \emptyset) \\ & \leq \mathbb{P} \left( \inf_{\Lambda_{\xi_T}^{\succ}, \mathcal{U}_2^*} \inf_{\mathbb{P} \in \mathcal{M}_p(\mathbb{P}_T, \xi_T)} D(v, \boldsymbol{\tau}, \boldsymbol{\lambda}, \mathbb{P}) - \inf_{\Lambda_{\xi_T}^{\succ}, \mathcal{U}_2^*} D(v, \boldsymbol{\tau}, \boldsymbol{\lambda}, \mathbb{P}) \geq - \inf_{\Lambda_{\xi_T}^{\succ}, \mathcal{U}_2^*} D(v, \boldsymbol{\tau}, \boldsymbol{\lambda}, \mathbb{P}) \right) \\ & \leq \mathbb{P} \left( \inf_{\mathbb{P} \in \mathcal{M}_p(\mathbb{P}_T, \xi_T)} \sup_{\Lambda_{\xi_T}^{\succ}, \mathcal{U}_2^*} |D(v, \boldsymbol{\tau}, \boldsymbol{\lambda}, \mathbb{P}) - D(v, \boldsymbol{\tau}, \boldsymbol{\lambda}, \mathbb{P})| \geq - \inf_{\Lambda_{\xi_T}^{\succ}, \mathcal{U}_2^*} D(v, \boldsymbol{\tau}, \boldsymbol{\lambda}, \mathbb{P}) \right) \\ & \leq \mathbb{P} \left( 2 \sup_{\Lambda} \|\boldsymbol{\lambda}\|_2 d^{\max(0, \frac{1}{p} - \frac{1}{2})} \inf_{\mathbb{P} \in \mathcal{M}_p(\mathbb{P}_T, \xi_T)} \mathcal{W}(\mathbb{P}, \mathbb{P}; p) \geq - \inf_{\Lambda_{\xi_T}^{\succ}, \mathcal{U}_2^*} D(v, \boldsymbol{\tau}, \boldsymbol{\lambda}, \mathbb{P}) \right) \\ & \leq \mathbb{P} \left( \mathcal{W}(\mathbb{P}, \mathbb{P}_T; p) \geq \left( 2d^{\max(0, \frac{1}{p} - \frac{1}{2})} \sup_{\Lambda} \|\boldsymbol{\lambda}\|_2 \right)^{-1} \left( \xi_T - \inf_{\Lambda_{\xi_T}^{\succ}, \mathcal{U}_2^*} D(v, \boldsymbol{\tau}, \boldsymbol{\lambda}, \mathbb{P}) \right) \right) \\ & \leq \mathbb{P} \left( \mathcal{W}(\mathbb{P}_T, \mathbb{P}; \frac{1}{2}) > \frac{1}{2d(p) \sup_{\Lambda} \|\boldsymbol{\lambda}\|_2} \left( \xi_T - \inf_{\Lambda_{\xi_T}^{\succ}, \mathcal{U}_2^*} D(v, \boldsymbol{\tau}, \boldsymbol{\lambda}, \mathbb{P}) \right) \right) \\ & \leq \mathbb{P} \left( \mathcal{W}(\mathbb{P}_T, \mathbb{P}; \frac{1}{2}) > \frac{\xi_T}{2d(p) \sup_{\Lambda} \|\boldsymbol{\lambda}\|_2} \right). \end{aligned}$$

Due to the Bobkov-Godse equivalence (see Theorem 1.3 of Bobkov and Gotze, (1999) (3)-see also Corollary 2.4 and Particular Case 2.5 of Bolley and Villani, (2005) (5)), and Theorem 1.1 of Bolley et al. (2007) (6), the existence of the exponential squared moment for  $\|\mathbf{X}_0\|_2$ , and relation (14) of Boissard (2011) (4), for any  $t^* > 0$ , we have that

$$\mathbb{P} \left( \mathcal{W} \left( \mathbb{P}_T, \mathbb{P}; \frac{1}{2} \right) > t^* + \mathbb{E} \left( \mathcal{W} \left( \mathbb{P}_T, \mathbb{P}; \frac{1}{2} \right) \right) \right) \leq \exp \left( \frac{-Tt^{*2}}{2 \inf_{\tau>0} h^2(\tau)} \right).$$

Setting in the above  $t^* := t - \mathbb{E} \left( \mathcal{W} \left( \mathbb{P}_T, \mathbb{P}; \frac{1}{2} \right) \right)$  for  $t > \mathbb{E} \left( \mathcal{W} \left( \mathbb{P}_T, \mathbb{P}; \frac{1}{2} \right) \right)$ , and observing that  $(t - \mathbb{E} \left( \mathcal{W} \left( \mathbb{P}_T, \mathbb{P}; \frac{1}{2} \right) \right)) \geq \left(1 - C(d)T^{-\frac{d+1}{d}}\right)^2 t^2$ , whenever  $t > \mathbb{E} \left( \mathcal{W} \left( \mathbb{P}_T, \mathbb{P}; \frac{1}{2} \right) \right)$ , we obtain from the proof of Theorem 2.3 of Lei (2020) (12) (where we identify 4 as an upper bound for the constant  $c_p$ ), as well as from the proof of Theorem 3.1 of Lei (2020) (12), where first, we identify  $c \leq 1 - \frac{\log_3(d)}{d}$  due to the equivalence between the Euclidean and the max norm, and the form of the packing number for closed max-balls, and, second, let  $m \rightarrow \infty$  in the case where  $d - \log_3(d) < \log_3(T)$  due to the existence of the exponential squared moment, that if  $t > C(d)T^{-\frac{1}{d}}$ , then

$$\mathbb{P} \left( \mathcal{W} \left( \mathbb{P}_T, \mathbb{P}; \frac{1}{2} \right) > t \right) \leq \exp \left( - \frac{\left(1 - C(d)T^{-\frac{d+1}{d}}\right)^2 Tt^2}{2 \inf_{\tau>0} h^2(\tau)} \right).$$

Setting in the previous  $t := \frac{\xi_T}{2d(p) \sup_{\Lambda} \|\boldsymbol{\lambda}\|_2}$  establishes (6).

Furthermore, due to Corollary 2 of Maesono (1989) (15), and for any  $v$ , for which  $D(v, \boldsymbol{\tau}, \boldsymbol{\lambda}, \mathbb{P}) < 0$ , we obtain that there exist positive constants  $C_1, C_2 > 0$ , independent of  $T$ , and a positive asymptotically negligible sequence  $\gamma(T)$  that satisfies  $\sqrt{T}\gamma(T) \rightarrow \infty$ , such that eventually,

$$\begin{aligned} & \frac{C_1}{\sqrt{T(1+T^{-1})}} - \frac{C_2\gamma(T)}{T(1+T^{-1})} - o\left(T^{-\frac{1}{2}}\right) + \exp(-cTD^2(v, \boldsymbol{\tau}, \boldsymbol{\lambda}, \mathbb{P})) \\ & \geq \mathbb{P}(D(v, \boldsymbol{\tau}, \boldsymbol{\lambda}, \mathbb{P}_T) - D(v, \boldsymbol{\tau}, \boldsymbol{\lambda}, \mathbb{P}) \geq -D(v, \boldsymbol{\tau}, \boldsymbol{\lambda}, \mathbb{P})) \geq \mathbb{P}(\boldsymbol{\Lambda}_{\mathbb{P}_T}^{\succeq} - \boldsymbol{\Lambda}_{\mathbb{P}}^{\succeq} \neq \emptyset), \end{aligned}$$

as long as  $\xi_T > (2\sqrt{2}d(p) \inf_{\tau>0} h(\tau) \sup_{\Lambda} \|\boldsymbol{\lambda}\|_2) \sqrt{\frac{\frac{1}{2} \ln T + \ln(1+T^{-1}) - \ln C_1}{T}}$ , and thus the second result in (7) follows. For the final result notice that due to the triangle inequality, the Lipschitz continuity property of  $v$ , the boundedness of the support, the Cauchy-Schwarz inequality, and the equivalence between the  $\ell_p$  norms in  $\mathbb{R}^d$ ,

$$\begin{aligned} & |\mathbb{E}_{\mathbb{P}_T}(u(\boldsymbol{\lambda}'_T \mathbf{X}_0)) - \mathbb{E}_{\mathbb{P}}(u(\boldsymbol{\lambda}'^* \mathbf{X}_0))| \\ & \leq |\mathbb{E}_{\mathbb{P}_T}(u(\boldsymbol{\lambda}'_T \mathbf{X}_0)) - \mathbb{E}_{\mathbb{P}_T}(u(\boldsymbol{\lambda}'^* \mathbf{X}_0))| + |\mathbb{E}_{\mathbb{P}_T}(u(\boldsymbol{\lambda}'^* \mathbf{X}_0)) - \mathbb{E}_{\mathbb{P}}(u(\boldsymbol{\lambda}'^* \mathbf{X}_0))|, \end{aligned}$$

with

$$|\mathbb{E}_{\mathbb{P}_T}(v(\boldsymbol{\lambda}'_T \mathbf{X}_0)) - \mathbb{E}_{\mathbb{P}_T}(v(\boldsymbol{\lambda}'^* \mathbf{X}_0))| \leq d(p) \mathbb{E}_{\mathbb{P}_T}(\|\mathbf{X}_0\|_2) \|\boldsymbol{\lambda}_T - \boldsymbol{\lambda}^*\|_p,$$

and

$$|\mathbb{E}_{\mathbb{P}_T}(u(\boldsymbol{\lambda}^* \mathbf{X}_0)) - \mathbb{E}_{\mathbb{P}}(u(\boldsymbol{\lambda}^* \mathbf{X}_0))| \leq \|\boldsymbol{\lambda}^*\|_2 \mathcal{W}(\mathbb{P}_T, \mathbb{P}; \frac{1}{2}),$$

and thereby

$$\begin{aligned} \mathbb{P}\left(|\mathbb{E}_{\mathbb{P}_T}(u(\boldsymbol{\lambda}'_T \mathbf{X}_0)) - \mathbb{E}_{\mathbb{P}}(u(\boldsymbol{\lambda}' \mathbf{X}_0))| > d(p) \mathbb{E}_{\mathbb{P}_T}(\|X\|_2) \|\boldsymbol{\lambda}_T - \boldsymbol{\lambda}^*\|_p + \|\boldsymbol{\lambda}^*\|_2 \xi_T\right) \\ \leq \mathbb{P}(\mathcal{W}(\mathbb{P}_T, \mathbb{P}; \frac{1}{2}) > \xi_T), \end{aligned}$$

and the result in (8) follows by setting  $t := \xi_T$  in (3.2).  $\square$

The results rely first on the iidness of the sample, an assumption that is compatible with our data frequency; it can be extended to m-dependent processes as well as to a class of Markov processes that possess contractive transition kernels (see Boissard (2011) (4)).

Second, they rely on the existence of some square-exponential moment for  $\|\mathbf{X}_0\|_2^2$ . This is equivalent to the existence of the moment generating function of  $\|\mathbf{X}_0\|_2^2$  in a neighborhood of zero, a condition that fails whenever  $\|\mathbf{X}_0\|$  follows a distribution with the right-tail behavior of the log-normal distribution. The exponential moment existence holds whenever  $X$  is bounded, or more generally whenever its squared elements follow sub-Gaussian distributions (see indicatively Chapter 2 of Vershynin (2018) (22)). The maximal moment parameter  $\tau$  can be estimated via the ratio  $\frac{(\kappa+1)\mathbb{E}_{\mathbb{P}_T}(\|\mathbf{X}_0\|_2^{2\kappa})}{\mathbb{E}_{\mathbb{P}_T}(\|\mathbf{X}_0\|_2^{2\kappa+2})}$ , due to the power series representation of the exponential moment and the properties of the ratio test for real series. Given this the optimization resulting to  $\inf_{\tau>0} h^2(\tau)$  can be empirically approximated. The choice of some non-optimal  $\tau$  can also be considered at the cost of a potentially less efficient probability bound, and a larger regularization parameter.

The probability bounds in the first and third cases of the theorem decline exponentially fast in  $T\xi_T^2$ , and hold for all  $T$  as long as the regularization parameter dominates a sequence of order  $\exp(-\frac{\ln T}{d})$ ; this declines slowly when the base asset dimensionality is large. This low rate of asymptotic negligibility for the multiplier can be circumvented at either the cost of some positive large multiplicative constant in front of the probability bound, or at the cost that the results hold eventually for large enough  $T$  that also depends on the regularizer (see for example Bolley et al. (2007) (6)).

The requirement of existence of squared exponential moments, can also be circumvented at the cost of loss of exponentially decreasing probability bounds; the existence of regular (polynomial) moments of some order would imply eventually polynomially decreasing probability bounds in  $T\xi_T^2$  (see also Bolley et al. (2007) (6)).

The first result says that under  $\mathcal{E}_2$ , the probability that there exist empirically enhanced portfolios that are non dominant in the population, is bounded above by the exponential

$$\exp\left(-\frac{\left(1-C(d)T^{-\frac{d+1}{d}}\right)^2 T\xi_T^2}{8d^2(p)\inf_{\tau>0}h^2(\tau)\sup_{\Lambda}\|\boldsymbol{\lambda}\|_2^2}\right).$$

The bound depends on the regularization coefficient, the base assets dimensionality, the size of the portfolio space, through  $d(p)$  on the choice of the  $\ell_p$  norm, and the squared exponential moment parameter; e.g, if  $d - \log_3(d) \geq \frac{\log_3(T)}{2}$  and  $\xi_T = 2cC(d)d(p)\sup_{\Lambda}\|\boldsymbol{\lambda}\|_2 T^{-\frac{1}{d}}$ , for  $c > \max(1, \frac{1}{8d(p)\inf_{\tau>0}h(\tau)})$  and  $p$  slightly less than 2, then the result implies that the probability of FD error is eventually bounded above by  $\exp(-c^*T^{1-\frac{2}{d}})$  for some (estimable) positive constant  $c^*$ .

Thus, for a given significance level  $\alpha \in (0, 1)$ , and if  $T \geq \left(-\frac{\ln(\alpha)}{c^*}\right)^{\frac{d}{d-2}}$ , the probability of FD is thus bounded above by  $\alpha$ . The same upper bound on the probability of FD holds whenever the regularization parameter is greater than the maximum between  $\frac{\sqrt{-8\ln(\alpha)d(p)\inf_{\tau>0}h(\tau)}}{|1-C(d)T^{-\frac{d+1}{d}}|}\sup_{\Lambda}\|\boldsymbol{\lambda}\|_2 T^{-\frac{1}{2}}$  and  $2C(d)d(p)\sup_{\Lambda}\|\boldsymbol{\lambda}\|_2 T^{-\frac{1}{d}}$ . Tracing

the proof shows that analogous bounds would hold for all  $\Lambda_{\mathbb{P}_T}^{\succeq}$ , as long as the weaker  $\mathcal{E}_1$  holds, yet failure of population dominance only happens inside  $\mathcal{U}_2^*(u)$ .

The second result also conditions on  $\mathcal{E}_2$ . It says that for the appropriate choice of the regularization multiplier the probability of FD for the members of the enhanced set, will eventually become smaller than the probability of false dominance for the whole of  $\Lambda_{\mathbb{P}_T}^{\succeq}$ , whenever failure of population dominance happens due to the complement of  $\mathcal{U}_2^*(u)$ . It shows the significance of the positive slacks implied by the regularization parameters on the SD conditions, for the small sample mitigation of the particular decision error.

The third result corresponds to an oracle inequality that relates the opportunity loss entailed by the empirical regularized problem, with the  $\ell_p$ -deviation between the empirical and the population solution and the regularization factor. If the empirical solution is consistent, it implies asymptotic negligibility for the opportunity loss w.h.p.

If instead of the regularized problem, the norm-constrained Lagrangean formulation is used,

$$\max_{\Lambda_p(c) \cap \Lambda_{\mathbb{P}_T}^{\succeq}} \mathbb{E}_{\mathbb{P}_T}(u(\boldsymbol{\lambda}'\boldsymbol{x})); \quad (9)$$

$$\Lambda_p(c) := \left\{ \boldsymbol{\lambda} \in \Lambda : \|\boldsymbol{\lambda} - \boldsymbol{\tau}\|_p \leq c \right\}, \quad (10)$$

for  $c > 0$ , then by the Lagrangian Duality Theorem, see Ch. 8, Theorem 1 of Luenberger (1997) (14), and Corollary 2.3.1 of Burke (1991) (7), analogous results would hold as long as  $\max_{\lambda \in \Lambda_{\frac{\gamma}{p}, T}} \|\lambda - \tau\|_p - c = \Omega\left(T^{-\frac{1}{d}}\right)$ .

### 3.3 Discussion

Analogous results would be analogously derivable to the above for other stochastic dominance relations, if the relevant set of utility functions that define the relation is equi-continuous w.r.t. the topology of locally uniform convergence. This could accommodate cases of preferences with non-global disposition towards risk like appropriate sets of S-shaped utilities associated with the Prospect Stochastic Dominance relation-see Levy and Levy (2013) (13).

The iid framework employed in Proposition 2 can be restrictive for economic data of high to moderate observation frequencies. The extension of such-like results to the more econometrically plausible framework of ergodicity and strong mixing could be benefited by analogous extensions of the Wasserstein distance empirical concentration inequalities to such settings. This is an interesting path for further research. Optimality of the probability bounds is of simultaneous interest for the non-asymptotic control of the probability of FD in realistically large samples.

### References

- [1] Arvanitis, S. and Post, T., 2024. Generalized stochastic arbitrage opportunities. *Management Science*, 70(7), pp.4629-4648.
- [2] Arvanitis, S. and Post, T., 2024. Stochastic Arbitrage Opportunities: Set Estimation and Statistical Testing. *Mathematics*, 12(4), p.608.
- [3] Bobkov, S.G. and Götze, F., 1999. Exponential integrability and transportation cost related to logarithmic Sobolev inequalities. *Journal of Functional Analysis*, 163(1), pp.1-28.
- [4] Boissard, E., 2011. Simple bounds for the convergence of empirical and occupation measures in 1-Wasserstein distance. *Electronic Journal of Probability*, 16, 2296-2333.
- [5] Bolley, F. and Villani, C., 2005. Weighted Csiszár-Kullback-Pinsker inequalities and applications to transportation inequalities. In *Annales de la Faculté des sciences de Toulouse: Mathématiques* (Vol. 14, No. 3, pp. 331-352).

- 
- [6] Bolley, F., Guillin, A. and Villani, C., 2007. Quantitative concentration inequalities for empirical measures on non-compact spaces. *Probability Theory and Related Fields*, 137(3), pp.541-593.
  - [7] Burke, J.V., 1991. An exact penalization viewpoint of constrained optimization. *SIAM Journal on control and optimization*, 29(4), pp.968-998.
  - [8] Constantinides, G.M., Czerwonko, M. and Perrakis, S., 2020. Mispriced index option portfolios. *Financial Management*, 49(2), pp.297-330.
  - [9] Gao, R., Chen, X. and Kleywegt, A.J., 2017. Distributional robustness and regularization in statistical learning. arXiv preprint arXiv:1712.06050.
  - [10] Hodder E., Jackwerth J. C., Kolokolova O., 2015. Improved portfolio choice using second-order stochastic dominance. *Rev. Finance* 19(4):1623–1647.
  - [11] Kantorovich, L.V., 1958. On methods of analysis of some extremal problems in planning production. *RAND*.
  - [12] Lei, J., 2020. Convergence and concentration of empirical measures under Wasserstein distance in unbounded functional spaces. *Bernoulli*, 26(1), 767-798.
  - [13] Levy, M. and Levy, H., 2013. Prospect theory: Much ado about nothing?. In *Handbook of the Fundamentals of Financial Decision Making: Part I* (pp. 129-144).
  - [14] Luenberger, D.G., 1997. *Optimization by vector space methods*. John Wiley & Sons.
  - [15] Maesono, Y., 1989. Lower bounds for the normal approximation of a sum of independent random variables. *Australian Journal of Statistics*, 31(3), pp.475-485.
  - [16] Post, T. and Longarela, I.R., 2021. Risk arbitrage opportunities for stock index options. *Operations Research*, 69(1), pp.100-113.
  - [17] Rahimian, H. and Mehrotra, S., 2019. Distributionally robust optimization: A review. arXiv preprint arXiv:1908.05659.
  - [18] Russell, W.R. and T.K. Seo, 1989. Representative sets for stochastic dominance rules, in: *Studies in the economics of uncertainty: In honor of Josef Hadar*, Springer-Verlag.



- 
- [19] Sion, M., 1958, On general minimax theorems, *Pacific Journal of Mathematics*, 8(1), pp.171-176.
  - [20] Sunklodas, J., 1999. On a lower bound of the uniform distance in the central limit theorem for  $\phi$ -mixing random variables. *Acta Applicandae Mathematica*, 58(1), pp.327-341.
  - [21] Vaart, A.W. and Wellner, J.A., 1996. Weak convergence. In *Weak convergence and empirical processes* (pp. 16-28). Springer, New York, NY.
  - [22] Vershynin, R, 2018, *High-dimensional probability: An introduction with applications in data science* (Vol. 47). Cambridge University Press.



**Department of Economics  
Athens University of Economics and Business**

**List of Recent Working Papers**

**2022**

- 01-22 Is Ireland the most intangible intensive economy in Europe? A growth accounting perspective, Ilias Kostarakos, KieranMcQuinn and Petros Varthalitis
- 02-22 Common bank supervision and profitability convergence in the EU, Ioanna Avgeri, Yiannis Dendramis and Helen Louri
- 03-22 Missing Values in Panel Data Unit Root Tests, Yiannis Karavias, Elias Tzavalis and Haotian Zhang
- 04-22 Ordering Arbitrage Portfolios and Finding Arbitrage Opportunities, Stelios Arvanitis and Thierry Post
- 05-22 Concentration Inequalities for Kernel Density Estimators under Uniform Mixing, Stelios Arvanitis
- 06-22 Public Sector Corruption and the Valuation of Systemically Important Banks, Georgios Bertsatos, Spyros Pagratis, Plutarchos Sakellaris
- 07-22 Finance or Demand: What drives the Responses of Young and Small Firms to Financial Crises? Stelios Giannoulakis and Plutarchos Sakellaris
- 08-22 Production function estimation controlling for endogenous productivity disruptions, Plutarchos Sakellaris and Dimitris Zaverdas
- 09-22 A panel bounds testing procedure, Georgios Bertsatos, Plutarchos Sakellaris, Mike G. Tsionas
- 10-22 Social policy gone bad educationally: Unintended peer effects from transferred students, Christos Genakos and Eleni Kyrkopoulou
- 11-22 Inconsistency for the Gaussian QMLE in GARCH-type models with infinite variance, Stelios Arvanitis and Alexandros Louka
- 12-22 Time to question the wisdom of active monetary policies, George C. Bitros
- 13-22 Investors' Behavior in Cryptocurrency Market, Stelios Arvanitis, Nikolas Topaloglou and Georgios Tsomidis
- 14-22 On the asking price for selling Chelsea FC, Georgios Bertsatos and Gerassimos Sapountzoglou
- 15-22 Hysteresis, Financial Frictions and Monetary Policy, Konstantinos Giakas
- 16-22 Delay in Childbearing and the Evolution of Fertility Rates, Evangelos Dioikitopoulos and Dimitrios Varvarigos
- 17-22 Human capital threshold effects in economic development: A panel data approach with endogenous threshold, Dimitris Christopoulos, Dimitris Smyrnakis and Elias Tzavalis
- 18-22 Distributional aspects of rent seeking activities in a Real Business Cycle model, Tryfonas Christou, Apostolis Philippopoulos and Vangelis Vassilatos

## 2023

- 01-23 Real interest rate and monetary policy in the post Bretton Woods United States, George C. Bitros and Mara Vidali
- 02-23 Debt targets and fiscal consolidation in a two-country HANK model: the case of Euro Area, Xiaoshan Chen, Spyridon Lazarakis and Petros Varthalitis
- 03-23 Central bank digital currencies: Foundational issues and prospects looking forward, George C. Bitros and Anastasios G. Malliaris
- 04-23 The State and the Economy of Modern Greece. Key Drivers from 1821 to the Present, George Alogoskoufis
- 05-23 Sparse spanning portfolios and under-diversification with second-order stochastic dominance, Stelios Arvanitis, Olivier Scaillet, Nikolas Topaloglou
- 06-23 What makes for survival? Key characteristics of Greek incubated early-stage startup(per)s during the Crisis: a multivariate and machine learning approach, Ioannis Besis, Ioanna Sapfo Pepelasis and Spiros Paraskevas
- 07-23 The Twin Deficits, Monetary Instability and Debt Crises in the History of Modern Greece, George Alogoskoufis
- 08-23 Dealing with endogenous regressors using copulas; on the problem of near multicollinearity, Dimitris Christopoulos, Dimitris Smyrnakis and Elias Tzavalis
- 09-23 A machine learning approach to construct quarterly data on intangible investment for Eurozone, Angelos Alexopoulos and Petros Varthalitis
- 10-23 Asymmetries in Post-War Monetary Arrangements in Europe: From Bretton Woods to the Euro Area, George Alogoskoufis, Konstantinos Gravas and Laurent Jacque
- 11-23 Unanticipated Inflation, Unemployment Persistence and the New Keynesian Phillips Curve, George Alogoskoufis and Stelios Giannoulakis
- 12-23 Threshold Endogeneity in Threshold VARs: An Application to Monetary State Dependence, Dimitris Christopoulos, Peter McAdam and Elias Tzavalis
- 13-23 A DSGE Model for the European Unemployment Persistence, Konstantinos Giakas
- 14-23 Binary public decisions with a status quo: undominated mechanisms without coercion, Efthymios Athanasiou and Giacomo Valletta
- 15-23 Does Agents' learning explain deviations in the Euro Area between the Core and the Periphery? George Economides, Konstantinos Mavrigiannakis and Vangelis Vassilatos
- 16-23 Mild Explocivity, Persistent Homology and Cryptocurrencies' Bubbles: An Empirical Exercise, Stelios Arvanitis and Michalis Detsis
- 17-23 A network and machine learning approach to detect Value Added Tax fraud, Angelos Alexopoulos, Petros Dellaportas, Stanley Gyoshev, Christos Kotsogiannis, Sofia C. Olhede, Trifon Pavkov
- 18-23 Time Varying Three Pass Regression Filter, Yiannis Dendramis, George Kapetanios, Massimiliano Marcellino
- 19-23 From debt arithmetic to fiscal sustainability and fiscal rules: Taking stock, George Economides, Natasha Miouli and Apostolis Philippopoulos
- 20-23 Stochastic Arbitrage Opportunities: Set Estimation and Statistical Testing, Stelios Arvanitis and Thierry Post
- 21-23 Behavioral Personae, Stochastic Dominance, and the Cryptocurrency Market, Stelios Arvanitis, Nikolas Topaloglou, and Georgios Tsoomidis
- 22-23 Block Empirical Likelihood Inference for Stochastic Bounding: Large Deviations Asymptotics Under  $m$ -Dependence, Stelios Arvanitis and Nikolas Topaloglou
- 23-23 A Consolidation of the Neoclassical Macroeconomic Competitive General Equilibrium Theory via Keynesianism (Part 1 and Part 2), Angelos Angelopoulos

**24-23 Limit Theory for Martingale Transforms with Heavy-Tailed Noise, Stelios Arvanitis and Alexandros Louka**

**2024**

- 01-24 Market Timing & Predictive Complexity, Stelios Arvanitis, Foteini Kyriazi, Dimitrios Thomakos**
- 02-24 Multi-Objective Frequentistic Model Averaging with an Application to Economic Growth, Stelios Arvanitis, Mehmet Pinar, Thanasis Stengos, Nikolas Topaloglou**
- 03-24 State dependent fiscal multipliers in a Small Open Economy, Xiaoshan Chen, Jilei Huang, Petros Varthalitis**
- 04-24 Public debt consolidation: Aggregate and distributional implications in a small open economy of the Euro Area, Eleftherios-Theodoros Roumpanis**
- 05-24 Intangible investment during the Global Financial Crisis in the EU, Vassiliki Dimakopoulou, Stelios Sakkas and Petros Varthalitis**
- 06-24 Time will tell! Towards the construction of instantaneous indicators of different agent-types, Iordanis Kalaitzoglou, Stelios Arvanitis**



## **Department of Economics Athens University of Economics and Business**

The Department is the oldest Department of Economics in Greece with a pioneering role in organising postgraduate studies in Economics since 1978. Its priority has always been to bring together highly qualified academics and top quality students. Faculty members specialize in a wide range of topics in economics, with teaching and research experience in world-class universities and publications in top academic journals.

The Department constantly strives to maintain its high level of research and teaching standards. It covers a wide range of economic studies in micro-and macroeconomic analysis, banking and finance, public and monetary economics, international and rural economics, labour economics, industrial organization and strategy, economics of the environment and natural resources, economic history and relevant quantitative tools of mathematics, statistics and econometrics.

Its undergraduate program attracts high quality students who, after successful completion of their studies, have excellent prospects for employment in the private and public sector, including areas such as business, banking, finance and advisory services. Also, graduates of the program have solid foundations in economics and related tools and are regularly admitted to top graduate programs internationally. Three specializations are offered: 1. Economic Theory and Policy, 2. Business Economics and Finance and 3. International and European Economics. The postgraduate programs of the Department (M.Sc and Ph.D) are highly regarded and attract a large number of quality candidates every year.

For more information:

<https://www.dept.aueb.gr/en/econ/>