ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS
ΣΧΟΛΗ ΟΙΚΟΝΟΜΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
SCHOOL OF ECONOMIC SCIENCES
ΤΜΗΜΑ ΟΙΚΟΝΟΜΙΚΗΣ ΕΠΙΣΤΗΜΗΣ
DEPARTMENT OF ECONOMICS

Department of Economics
Athens University of Economics and Business

WORKING PAPER no. 15-2021

Ordering Arbitrage Portfolios and Finding Arbitrage Opportunities

Stelios Arvanitis and Thierry Post

April 2021

Ordering Arbitrage Portfolios and Finding Arbitrage Opportunities

Stelios Arvanitis*and Thierry Post†

April 17, 2021

Concepts are introduced for analyzing arbitrage portfolios in the face of ambiguity about investor risk preferences and initial portfolio holdings. A Stochastic Arbitrage Opportunity is a self-financing overlay portfolio which enhances every feasible host portfolio for all relevant utility functions. The logical alternative to the existence of such opportunities is the solvability of a system of asset pricing restrictions for the base assets. An asymptotic statistical theory is developed to analyze the behavior of the empirical optimal arbitrage portfolio if the latent parameters of the joint payoff distribution are estimated using time-series data. An empirical application to one-month S&P500 index options uncovers significant arbitrage opportunities in the form of short strip strangle spreads, for investors with reasonable levels of relative risk aversion. The optimal option combinations enhance not only passive index portfolios but also portfolios with existing volatility risk exposure.

Key words: Portfolio analysis; Arbitrage portfolios; Higher-order risk; Asset pricing; Incomplete markets; Index options.

*Arvanitis is Associate Professor at the Department of Economics of Athens University of Economics and Business; Athens; Greece; email: stelios@aueb.gr.

†Post is Professor of Finance at the Graduate School of Business of Nazarbayev University; Astana, Kazakhstan; e-mail: thierrypost@hotmail.com

1 Introduction

In portfolio analysis and asset pricing, the Mean-Variance Dominance and Stochastic Dominance rules are used to deal with ambiguity about risk preferences. These partial orders allow for the comparison of investment portfolios without specifying a particular utility function. The focus traditionally is on standard portfolios which are constructed subject to a budget constraint. The present study extends the analysis to self-financing *arbitrage portfolios* which can be added to an existing standard portfolio. A distinguishing feature is that the analysis allows for ambiguity about initial portfolio holdings in addition to ambiguity about risk preferences.

Arbitrage portfolios play a pivotal role in Investments. Active investors can exploit mispricing without making a net investment by buying underpriced securities and short selling overpriced securities. Furthermore, the investment returns of properly constructed arbitrage portfolios are commonly used to proxy for latent risk factors; well-known examples are the long-short equity portfolios by Fama and French.

Korkie and Turtle (2002) develop an analytical framework for arbitrage portfolios based on Mean-Variance Dominance. The present study employs a more general partial order which is based on a Expected Utility, in the spirit of Stochastic Dominance. The analysis is not restricted to specific shapes of the utility function and the return distribution, and can account for tail risk and skewness. This feature is particularly relevant for arbitrage portfolios, which by definition include short positions and may also include open positions in derivative securities.

In the proposed framework, a Stochastic Arbitrage Opportunity (SAO) occurs if a self-financing overlay portfolio can be constructed which enhances all feasible host portfolios for all relevant utility functions. The sets of overlay portfolios, host portfolios and utility functions need to obey only minimal regularity conditions and can be tailored to specific applications.

This concept generalizes the elementary Pure Arbitrage Opportunity (PAO) which is appealing to all active investors regardless of their risk preferences and initial positions. By using side information about the preferences and/or endowment, the concept of SAO enlarges the set of arbitrage opportunities. PAOs can be exploited with arbitrage portfolios which fully hedge exposures to systematic risk factors and fully diversify away non-systematic risk; see, for example, Kim, Korajczyk and Neuhierl (2020). By contrast, SAOs do not require the complete elimination of investment risk, provided that all relevant investors benefit from it.

Several earlier studies have developed generalized arbitrage concepts by imposing restrictions on risk preferences or the market price of risk. Bernardo and Ledoit (2000) and Cochrane and Saa-Requejo (2000) propose pricing bounds for a new base asset which are based on bounds on the gain-loss ratio and bounds on the Sharpe ratio, respectively. Constantinides, Jackwerth and Perrakis (2009) and Post and Rodríguez

Longarela (2021) consider joint restrictions on the prices of a cross-section of base assets which ensure that the market index is not dominated by Second-degree Stochastic Dominance (SSD). The proposed framework facilitates a larger range of restrictions on risk preferences.

The framework also allows for the use of general investment restrictions on the host portfolios and arbitrage portfolios. The restrictions may be externally defined by the demands of regulators and clients of money managers. They may also be used to model proportional transaction costs, as in Jouini and Kallal (2001) and Constantinides, Jackwerth and Perrakis (2009), enhance the robustness for estimation error, as in Jagannathan and Ma (2003) and DeMiguel, Garlappi, Nogales and Uppal (2009), or to mitigate default risk and price impact, as in Post and Rodriguez Longarela (2021).

The SAO concept is designed for the analysis of incomplete capital markets; the analysis therefore eschews assumptions about the existence of a unique pricing kernel and risk-neutral distribution. The existence of SAOs defies equilibrium in a class of asset pricing models. A novel system of asset pricing restrictions is derived which is satisfied if and only SAOs do not exist. The system generalizes the standard no-arbitrage pricing rule and the asset pricing system for incomplete markets by Constantinides, Jackwerth and Perrakis (2009).

The focus is on the cross-sectional choice from a multitude of risky base assets in a single-period model. In theory, the analysis can be extended to multiple periods or continuous time, as in, for example, Bondarenko (2003). However, the large amount of information required for this approach introduces a substantial risk of specification error and a curse of dimensions for empirical estimation. A popular alternative approach represents or approximates multi-period models through the inclusion of conditional portfolios (Hansen and Richard (1987)) and/or timing portfolios (Brandt and Santa-Clara (2006)) as base assets in a single-period model. Furthermore, the existence of arbitrage opportunities in a single-period model generally is a sufficient condition for the existence of arbitrage opportunities in a multi-period model.

A challenge for practical implementation is the empirical estimation of the joint payoff distribution. The optimal estimation method depends on the data dimensions and available side information in specific applications. To estimate investment risk in a statistically consistent way, it is often desirable to consider non-parametric estimation methods, when the data dimensions allow for this approach. It is also desirable to incorporate available side information, for example, in the form of risk factor models and parametric distribution shapes, to mitigate the curse of dimensions for estimating multivariate distributions.

This study includes an asymptotic theory for empirical arbitrage opportunities which are estimated using time-series data. The theory extends and generalizes the statistical analysis of standard portfolios by Post, Karabati and Arvanitis (2018) in important ways. It focuses on arbitrage portfolios and allows for ambiguity about the host portfolio; the assumption framework covers a wide range of parametric and non-parametric estimation methods; the results include the consistent estimation of the entire solution set without the need for non-trivial data-dependent moment selection techniques.

Since the portfolio theory assumes that the conditional payoff distribution is known or estimated with high accuracy, special attention is given to establishing that the adverse effects of estimation error on portfolio choice vanish in large samples. For the analysis of new asset classes and individual securities without a simple factor structure, estimation risk and ambiguity aversion will play an important role. Extensions of the portfolio theory and limit theory to the modeling of ambiguity-averse preferences and the estimation of 'distributions of distributions' are routes for further research.

The proposed framework for arbitrage portfolios and SAOs is applied to the pricing and trading of one-month S&P500 stock index (SPX) options. This application benefits from the known and simple factor structure of option payoffs and the availability of time-series data and side information for estimating the distribution of index returns. The analysis is inspired by several related studies which evaluate SPX options using partial orders.

Constantinides, Jackwerth and Perrakis (2009) develop and test an index option pricing system for all risk-averse index investors in incomplete markets. The pricing system is systematically violated in their empirical study of one-month SPX options, consistent with the empirical option pricing kernel puzzle (Ait-Sahalia and Lo (2000) and Jackwerth (2000)). Follow-up studies by Constantinides, Czerwonko, Jackwerth and Perrakis (2011), Constantinides, Czerwonko and Perrakis (2020) and Post and Rodríguez Longarela (2021) develop trading strategies that seek to exploit these violations, to quantify the economic significance.

The results of these empirical studies are mixed. Constantinides, Czerwonko and Perrakis (2020) find only limited out-of-sample outperformance for active combinations of one-month options compared with weeklies. Post and Rodríguez Longarela (2021) find that significant outperformance can be achieved only for small-scale portfolios which are restricted by the available quote size for the best quoted prices. These empirical results call into question whether the pricing errors are economically significant.

The focus in these studies is on the set of all risk-averse index investors. New light can be shed on the topic by reducing the utility function class and expanding the benchmark set. Czerwonko, Davidson and Perrakis (2021) limit the maximum level of risk aversion to study the overpricing of out-of-the-money (OTM) puts and at-the-money (ATM) straddles, using single-option pricing bounds based on Almost Stochastic Dominance (Leshno and Levy (2002)). Similarly, our empirical study avoids extreme levels of risk aversion to detect mispriced option combinations. In addition, the benchmark set is enriched with standard option combinations, to represent investors who have existing option positions, and to allow for pricing kernels which depend on volatility risk in addition to market risk.

2 Financial Theory

In this section, the concept of SAO is introduced together with a system of dual restrictions on the market prices of the base assets, and discussion of arbitrage portfolio choice and numerical analysis. Section 3 introduces and analyzes empirical counterparts of the theoretical concepts.

2.1 Preliminaries

The focus is on a single-period investment problem with N base assets with payoffs at the investment horizon, $\boldsymbol{x} := (x_1 \cdots x_N)^{\mathrm{T}} \in \mathbb{R}^N$, $N < \infty$. To represent or approximate certain multi-period problems, the base assets may include dynamic portfolios which are periodically re-balanced based on conditioning information. Riskfree base assets may be included.

The joint probability distribution function of the payoffs is given by $\mathcal{F} : \mathcal{X} \to [0, 1]$, where $\mathcal{X} := [a, b]$, $-\infty < a < b < \infty$, includes the supports of the base assets. The distribution function is generally conditional on the prevailing stage of business cycle and/or financial market conditions. A general factor model is assumed: $\boldsymbol{x} = g(\boldsymbol{f}) + \boldsymbol{\varepsilon}$, where $\boldsymbol{f}$ is a vector of M common risk factors, $g : \mathbb{R}^M \to \mathcal{X}^N$ is a potentially non-linear mapping to the asset payoffs, and $\boldsymbol{\varepsilon}$ is a vector of mutually independent residuals. Without limiting the number and nature of the risk factors, this structure is not restrictive, and allows for the non-informative specification $\boldsymbol{f} = \boldsymbol{x}$.

The prices of the base assets are taken as given, and are collected in the price vector $\boldsymbol{p} \in \mathbb{R}^N$. The analysis can account for proportional transaction costs in the form of a bid-ask spread by introducing long positions and short positions as separate base assets with different prices, as in Jouini and Kallal (2001) and Constantinides, Jackwerth and Perrakis (2009).

Risk preferences are represented by utility functions $u : \mathbb{R} \to \mathbb{R}$. Instead of specifying a particular functional form, a set of relevant utility functions, $\mathcal{U}$, is defined using general functional properties. The utility functions are assumed to be continuously differentiable, strictly increasing and strictly concave. Additional assumptions can be employed to improve the discriminatory power, for example, higher-order risk aversion properties or a range for the Relative Risk Aversion (RRA) coefficient.

The set $\mathcal{U}$ is assumed to be uniformly bounded and convex. This assumption is not very restrictive, given that the support of the payoff distribution is bounded and the proposed stochastic order is invariant to linear transformation and mixing of utility functions. For example, the set of all aforementioned utility functions can be represented using $\mathcal{U}_2$, the bounded set of strictly positive mixtures of elementary Russell and Seo (1989) utility functions $v_{2;\phi}(x) = -(\phi - x)_+$, $\phi \in \mathcal{X}$.

The portfolio possibilities are described by two distinct portfolio sets: a benchmark set of host portfolios, $\mathrm{K} \subseteq \mathbb{R}^N$, and an overlay set of arbitrage portfolios, $\Delta \subseteq \mathbb{R}^N$. Korkie and Turtle (2002) refer to these portfolio sets as the Investment Opportunity Set and the Self-Financing Investment Opportunity Set, respectively. The asset positions $\boldsymbol{\kappa} \in \mathrm{K}$ and $\boldsymbol{\delta} \in \Delta$ are treated as decision variables. The asset positions translate into portfolio payoffs $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\kappa}$ and $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\delta}$, respectively. The payoffs are not treated as the decision variables, because the set of available payoff vectors may be latent in empirical applications, the payoff vectors are of infinite dimensions for continuous distributions, and the payoff vectors may not have unique prices in disequilibrium.

The benchmark set K is assumed to be closed and bounded but it need not be convex. In some cases, K consists of a single benchmark index, for example, a general market index or a tailor-made style index. In other cases, the host portfolio is ambiguous, because the portfolio composition of a given investor is not fully known, or, alternatively, multiple, heterogeneous investors are analyzed simultaneously. In such cases, the benchmark set may include a multitude of host portfolios, for example, the universe of all passive mutual funds for a specific asset class and geographical region. The benchmark set could also consist of a continuum of convex mixtures of the base assets. Since the host portfolios are standard portfolios, the normalized asset positions $\boldsymbol{\kappa} \in \mathrm{K}$ must obey the normalized budget restriction: $\boldsymbol{p}^{\mathrm{T}}\boldsymbol{\kappa} = 1$.

The overlay portfolios $\boldsymbol{\delta} \in \Delta$ are combinations of short positions and long positions with a net investment of zero, or $\boldsymbol{p}^{\mathrm{T}}\boldsymbol{\delta} = 0$. The overlays can be added to a host portfolio $\boldsymbol{\kappa} \in \mathrm{K}$ to form a combined standard portfolio $\boldsymbol{\lambda} = (\boldsymbol{\kappa} + \boldsymbol{\delta})$. An overlay could be constructed using active investing by the investor who chooses also the host portfolio or, alternatively, by an external specialized intermediary such as a hedge fund or private equity fund. In some cases, the feasible overlays may depend on the choice of the host portfolio, or $\Delta(\boldsymbol{\kappa})$, $\boldsymbol{\kappa} \in \mathrm{K}$. In these cases, the analysis may focus on the intersection $\Delta = \cap_{\boldsymbol{\kappa} \in \mathrm{K}} \Delta(\boldsymbol{\kappa})$, to detect universal improvement possibilities which occur for every host.

For analytical convenience, it is assumed that the overlay set is a bounded and convex polytope which is characterized by R linear restrictions: $\Delta := \left\{\boldsymbol{\delta} \in \mathbb{R}^N : \mathbf{A}\boldsymbol{\delta} \leq \boldsymbol{a}\right\}$. Here, $\mathbf{A}$ is a $(R \times N)$ matrix of left-hand-side coefficients for R linear restrictions, and $\boldsymbol{a}$ is the corresponding $(R \times 1)$ vector of right-hand-side coefficients. The restrictions may be externally defined by the demands of regulators and clients of money managers or self-imposed to reduce default risk, price impact and sensitivity to estimation error. The restrictions include the aforementioned zero-investment constraint, $\boldsymbol{p}^{\mathrm{T}}\boldsymbol{\delta} = 0$. Moreover, the system of inequalities is assumed to be consistent ($\Delta \neq \emptyset$), and to include at least the 'passive' solution, or $\mathbf{0}_N \in \Delta$. The passive solution is assumed to be the default choice if qualified arbitrage opportunities do not exist.

The set of all feasible combined portfolios is given by the vector subspace sum $\Lambda_0 := \mathrm{K} + \Delta$. The analysis does not attempt to identify individual standard portfolios $\boldsymbol{\lambda} \in \Lambda_0$ which stand out as being particularly appealing, for two reasons. First, it is generally not possible to identify a combined portfolio which is superior to all host portfolios for all relevant utility functions. Second, the benchmark set is a superset of the relevant portfolios for multiple investors; some elements of this superset may be infeasible for some of the investors. Instead, the analysis identifies arbitrage portfolios $\boldsymbol{\delta} \in \Delta$ which are appealing overlays for all feasible host portfolios and all relevant utility functions.

2.2 Stochastic Arbitrage Opportunities

A partial order for arbitrage portfolios is introduced and used to generalize the notion of PAO to SAO. To allow for compact notation, the expected utility increment for $(u, \boldsymbol{\kappa}, \boldsymbol{\delta}) \in \mathcal{U} \times \mathrm{K} \times \Delta$ is denoted by $D(u, \boldsymbol{\kappa}, \boldsymbol{\delta}, \mathcal{F}) := \mathbb{E}_{\mathcal{F}}\left[u(\boldsymbol{x}^{\mathrm{T}}(\boldsymbol{\kappa}+\boldsymbol{\delta}))\right] - \mathbb{E}_{\mathcal{F}}\left[u(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\kappa})\right]$. Furthermore, the pair $(\mathcal{U}, \mathrm{K})$ and the triplet $(\mathcal{U}, \mathrm{K}, \mathcal{F})$ are sometimes replaced by $(\cdot)$ when the meaning is clear from the context.

Definition 2.2.1 (Strict Arbitrage Opportunities). *An overlay portfolio $\boldsymbol{\delta} \in \Delta$ (strictly) stochastically enhances a given host portfolio $\boldsymbol{\kappa} \in \mathrm{K}$, or $(\boldsymbol{\kappa}+\boldsymbol{\delta}) \succ_{(\mathcal{U},\mathcal{F})} \boldsymbol{\kappa}$, if $D(u, \boldsymbol{\kappa}, \boldsymbol{\delta}, \mathcal{F}) > 0$ for all $u \in \mathcal{U}$. It is a (strict) stochastic arbitrage opportunity if such enhancement is achieved for all $\boldsymbol{\kappa} \in \mathrm{K}$. The set of all feasible (strict) SAOs is given by*

$$\Delta^{\succ}_{(\mathcal{U},\mathrm{K},\mathcal{F})} := \{\boldsymbol{\delta} \in \Delta : D(u, \boldsymbol{\kappa}, \boldsymbol{\delta}, \mathcal{F}) > 0 \; \forall \, (u, \boldsymbol{\kappa}) \in \mathcal{U} \times \mathrm{K}\}. \tag{1}$$

If a SAO $\boldsymbol{\delta} \in \Delta^{\succ}_{(\cdot)}$ exists, then every host portfolio $\boldsymbol{\kappa} \in \mathrm{K}$, regardless of its efficiency inside the benchmark set K, is inefficient in the combined portfolio set Λ_0, in the sense that it is not optimal for any utility function, and inferior to the combined portfolio $\boldsymbol{\kappa}+\boldsymbol{\delta}$.

The SAO set has a number of useful properties. Since the utility functions $u \in \mathcal{U}$ are concave, and the overlay set Δ is convex, $\Delta^{\succ}_{(\cdot)}$ is convex, so that $(\boldsymbol{\delta}_1, \boldsymbol{\delta}_2) \in \Delta^{\succ}_{(\cdot)} \Rightarrow c\boldsymbol{\delta}_1 + (1-c)\boldsymbol{\delta}_2 \in \Delta^{\succ}_{(\cdot)}$ for all $c \in [0,1]$. In addition, since $\{\mathbf{0}_N\} \in \Delta$, it follows that $\boldsymbol{\delta} \in \Delta^{\succ}_{(\cdot)} \Rightarrow c\boldsymbol{\delta} \in \Delta^{\succ}_{(\cdot)}$ for all $c \in \langle 0, 1]$. SAOs therefore are not disconnected portfolios but part of a continuous neighborhood of SAOs.

Since enhancement is defined using strict inequalities and strict risk aversion, the SAO set may be open. By definition, the SAO set does not include the passive solution $\boldsymbol{\delta} = \{\mathbf{0}_N\}$, as $D(u, \boldsymbol{\kappa}, \mathbf{0}_N, \mathcal{F}) = 0$ for all $(u, \boldsymbol{\kappa}) \in \mathcal{U} \times \mathrm{K}$. To facilitate numerical analysis and statistical analysis, Section 2.4 introduces an alternative definition based on weak inequalities and weak risk aversion.

The set of SAOs naturally increases as the restrictions on the overlays (Δ) are loosened and/or the restrictions on the benchmarks (K) and/or risk preferences ($\mathcal{U}$) are tightened. The sharpest results are obtained when redundant host portfolios and utility functions are excluded.

Perhaps surprisingly, the analysis is not invariant to the inclusion of non-optimal elements in K. The transitivity of the dominance relation entails $\left((\boldsymbol{\kappa}_1+\boldsymbol{\delta}) \succ_{(\mathcal{U},\mathcal{F})} \boldsymbol{\kappa}_1\right) \wedge \left(\boldsymbol{\kappa}_1 \succ_{(\mathcal{U},\mathcal{F})} \boldsymbol{\kappa}_2\right) \Rightarrow \left((\boldsymbol{\kappa}_1+\boldsymbol{\delta}) \succ_{(\mathcal{U},\mathcal{F})} \boldsymbol{\kappa}_2\right)$, but not $\left((\boldsymbol{\kappa}_1+\boldsymbol{\delta}) \succ_{(\mathcal{U},\mathcal{F})} \boldsymbol{\kappa}_1\right) \wedge \left(\boldsymbol{\kappa}_1 \succ_{(\mathcal{U},\mathcal{F})} \boldsymbol{\kappa}_2\right) \Rightarrow \left((\boldsymbol{\kappa}_2+\boldsymbol{\delta}) \succ_{(\mathcal{U},\mathcal{F})} \boldsymbol{\kappa}_2\right)$, as the enhancement relation relies on the dependence structure between the overlay and the host in addition to the two marginal distributions. For this reason, the benchmark set generally cannot be replaced without consequences by the subset of its optimal elements.

Nevertheless, non-optimal elements may be excluded to sharpen the analysis, if the host portfolios are believed to be constructed using optimization. Riskless alternatives are a case in point. Although some base assets and arbitrage portfolios may be riskless, it is generally desirable to exclude riskless host portfolios. If one of the host portfolios $\boldsymbol{\kappa} \in \mathrm{K}$ is riskfree ($\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\kappa} = z$), and $\mathcal{U} = \mathcal{U}_2$, then the concept of SAO reduces to the basic concept of a PAO, or an arbitrage portfolio $\boldsymbol{\delta} \in \Delta$ which generates only non-negative payoffs $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\delta} \geq 0$, and at least one strictly positive payoff $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\delta} > 0$.

This set reduction is generally undesirable and avoidable. A riskfree alternative is generally not optimal in K for any utility function, if some of the risky host portfolios feature risk premiums and/or pricing errors. In addition, the existence of enhancement possibilities is trivial for riskless host portfolios, if at least one arbitrage portfolio has a positive expected payoff.

In case of the singleton specification $\mathrm{K} = \{\boldsymbol{\kappa}\}$ and $\mathcal{U} = \mathcal{U}_2$, the enhanced portfolios $\boldsymbol{\lambda} \in \left(\{\boldsymbol{\kappa}\} + \Delta^{\succ}_{(\cdot)}\right)$ dominate the benchmark by SSD, or $\boldsymbol{\lambda} \succ_{(\mathcal{U}_2,\mathcal{F})} \boldsymbol{\kappa}$, and the risk arbitrage concept of Post and Rodriguez Longarela (2021) is obtained. The definition of SAO is much more general and can expand the set of arbitrage opportunities by imposing additional risk preference assumptions, and avoid specification error for the initial portfolio holdings by using multiple host portfolios.

A SAO should not be confused with a stochastic bound (Arvanitis, Post and Topaloglou (2020)), or a combined portfolio $\boldsymbol{\lambda} \in \Lambda_0$ which dominates every host, or $\boldsymbol{\lambda} \succ_{(\mathcal{U}_2,\mathcal{F})} \boldsymbol{\kappa}$ for all $\boldsymbol{\kappa} \in \mathrm{K}$. A SAO is an arbitrage portfolio $\boldsymbol{\delta} \in \Delta$, while a bound is a standard portfolio $\boldsymbol{\lambda} \in \Lambda_0$. In addition, exact bounds generally do not exist if K includes multiple portfolios with diverse risk profiles, which is the reason for the use of approximate bounds in Arvanitis, Post and Topaloglou (2020). The definition of SAO is much less demanding than the definition of a stochastic bound, because the host portfolio enters both on the left-hand side and the right-hand-side of the pairwise order $(\boldsymbol{\kappa}+\boldsymbol{\delta}) \succ_{(\mathcal{U},\mathcal{F})} \boldsymbol{\kappa}$. Instead of looking for a portfolio which is superior to every benchmark, the search is for a portfolio adjustment which is an improvement for every host portfolio. This subtle adjustment yields a large improvement in discriminatory power in relevant applications.

2.3 Asset Pricing Restrictions

One of the first principles of asset pricing states that a PAO does not exist if and only if a non-negative pricing kernel $m : \mathbb{R}^N \to \mathbb{R}$ exists; see, for example, Dybvig and Ross (2008). This section generalizes this principle from PAOs to SAOs.

Some additional concepts are introduced to facilitate the derivation and interpretation. A common specification of the pricing kernel is $m(\boldsymbol{x}) = d(u, \boldsymbol{\kappa})u'(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\kappa})$, $(u, \boldsymbol{\kappa}) \in \mathcal{U} \times \mathrm{K}$, where the scalar $d(u, \boldsymbol{\kappa}) > 0$ captures time preferences and the marginal utility function $u'(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\kappa})$ captures risk preferences. The set $\mathcal{M}_{(\mathcal{U},\mathrm{K})} := \left\{m(\boldsymbol{x}) = d(u, \boldsymbol{\kappa})u'(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\kappa}); (u, \boldsymbol{\kappa}) \in \mathcal{U} \times \mathrm{K}\right\}$ includes all candidate pricing kernels which take this shape for some permissible utility function and feasible host portfolio. To invoke standard results from game theory and duality theory, the analysis uses the convex hull conv $\left(\mathcal{M}_{(\mathcal{U},\mathrm{K})}\right)$ and the dual polyhedral cone $\Delta^* := \left\{\boldsymbol{q} = \mathbf{A}^{\mathrm{T}}\boldsymbol{\sigma}; \boldsymbol{a}^{\mathrm{T}}\boldsymbol{\sigma} \leq 0; \boldsymbol{\sigma} \geq \mathbf{0}_R\right\}$.

Theorem 2.3.1 (Existential condition). *(Strict) SAOs do not exist for a given utility function class* $\mathcal{U}$*, benchmark set* K *and overlay set* Δ*, or* $\Delta^{\succ}_{(\mathcal{U},\mathrm{K},\mathcal{F})} = \emptyset$*, if and only if*

$$\left(\mathbb{E}_{\mathcal{F}}\left[m(\boldsymbol{x})\boldsymbol{x}\right] \in \Delta^*\right) \text{ for some } m \in \text{conv}\left(\mathcal{M}_{(\mathcal{U},\mathrm{K})}\right). \tag{2}$$

In other words, the Present Value of the base asset payoffs, or $\mathbb{E}_{\mathcal{F}}\left[m(\boldsymbol{x})\boldsymbol{x}\right]$, must lie inside the dual polyhedral cone Δ^*, for some permissible kernel $m \in$conv $\left(\mathcal{M}_{(\cdot)}\right)$, to exclude the existence of SAOs. The convex hull conv $\left(\mathcal{M}_{(\cdot)}\right)$ is used here instead of $\mathcal{M}_{(\cdot)}$, to establish a general necessary and sufficient condition. Without the convexification, a duality gap may arise between the arbitrage conditions and pricing conditions.

For a riskfree host portfolio ($\boldsymbol{\kappa} \in \mathrm{K} : \boldsymbol{x}^{\mathrm{T}}\boldsymbol{\kappa} = x$), marginal utility is constant, or $u'(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\kappa}) = c$, and the existential condition restricts only the expected value of the base assets, consistent with marginal risk-neutrality. Since some of the arbitrage portfolios generally have a positive expected payoff, the inclusion of a riskless host portfolio generally does not affect the existential condition, despite the potentially large reduction of the SAO set $\Delta^{\succ}_{(\mathcal{U},\mathrm{K},\mathcal{F})}$.

The asset pricing system for incomplete markets with proportional transactions costs by Constantinides, Jackwerth and Perrakis (2009) is obtained as a special case of (2) for a particular specification of $\mathcal{U}$, K and Δ. Specifically, their system is obtained by setting $\mathcal{U} = \mathcal{U}_2$, $\mathrm{K} = \{\boldsymbol{\kappa}\}$, a general market index, and furthermore, treating long positions (entered at the ask price) and short position (entered at the bid price) as separate base assets, and including corresponding sign restriction for the portfolio positions in Δ: $\delta_i \geq 0$

($\delta_i \leq 0$) for the base assets which are traded at the ask (bid) price. This pricing kernel system, in turn, is the logical complement to the condition that the benchmark portfolio is undominated by any alternative portfolio $\boldsymbol{\lambda} \in \Lambda_0$ by SSD, as shown in Post and Rodríguez Longarela (2021).

The proposed framework allows for generalizing the analysis by alternative specifications of $\mathcal{U}$, K and Δ. Naturally, the asset pricing restrictions become tighter (looser) as the overlay set is enlarged (reduced) and/or the benchmark set and utility function class are reduced (enlarged).

The empirical application in Section 4 enlarges the benchmark set by introducing standard option combinations as additional benchmarks, and reduces the utility function class by restricting the variability of the pricing kernel. Since the value of options depends on market volatility, the inclusion of host portfolios with written option positions is similar to the use of pricing kernels which are increasing in volatility as in Christoffersen, Heston and Jacobs (2013).

In the empirical application, the asset pricing restrictions are systematically violated, due to the existence of empirical strict SAOs in every monthly sample, for all general specifications of $\mathcal{U}$ and K. Instead of testing for the existence of qualified pricing kernels, the focus is on quantifying the economic significance of pricing errors in terms of the out-of-sample investment performance of selected empirical SAOs.

2.4 Arbitrage Portfolio Choice

The attention is now turned to the selection of a specific SAO which optimizes a given objective function. Using the strict SAO set $\Delta^{\succ}_{(\cdot)}$ as the feasible set can be analytically inconvenient, as it may be an empty set or an open set. To facilitate numerical analysis, the feasible set is closed by using weak inequalities ($D(u, \boldsymbol{\kappa}, \boldsymbol{\delta}, \mathcal{F}) \geq 0$) and closure of the utility function class ($\mathrm{cl}\,(\mathcal{U})$).

Definition 2.4.1 (Weak Arbitrage Opportunities). *An overlay portfolio $\boldsymbol{\delta} \in \Delta$ (weakly) stochastically enhances a given host portfolio $\boldsymbol{\kappa} \in$ K, or $(\boldsymbol{\kappa} + \boldsymbol{\delta}) \gtrsim_{(\mathcal{U},\mathcal{F})} \boldsymbol{\kappa}$, if $D(u, \boldsymbol{\kappa}, \boldsymbol{\delta}, \mathcal{F}) \geq 0$ for all $u \in \mathrm{cl}(\mathcal{U})$. It is a (weak) stochastic arbitrage opportunity if such enhancement is achieved for all $\boldsymbol{\kappa} \in$ K. The set of all feasible (weak) SAOs is given by*

$$\Delta^{\gtrsim}_{(\mathcal{U},\mathrm{K},\mathcal{F})} := \left\{ \boldsymbol{\delta} \in \Delta : D(u, \boldsymbol{\kappa}, \boldsymbol{\delta}, \mathcal{F}) \geq 0 \;\forall\, (u, \boldsymbol{\kappa}) \in \mathrm{cl}\,(\mathcal{U}) \times \mathrm{K} \right\}. \tag{3}$$

By construction, the weak set is a superset of the strict set: $\Delta^{\gtrsim}_{(\cdot)} \supseteq \Delta^{\succ}_{(\cdot)}$. Since the passive solution $\{\mathbf{0}_N\}$ is included, the weak SAO set is non-empty ($\Delta^{\gtrsim}_{(\cdot)} \neq \emptyset$). It is also convex and closed.

The enlargement of the feasible set $(\Delta^{\succsim}_{(\cdot)} - \Delta^{\succ}_{(\cdot)})$ is inconsequential for the analysis. One specific subset of this enlargement consists of trivial solutions: $\Delta^{\equiv}_{\mathcal{F}} := \left\{\boldsymbol{\delta} \in \Delta : \boldsymbol{x}^{\mathrm{T}}\boldsymbol{\delta} = 0 \,\forall \boldsymbol{x} : \mathcal{F}(\boldsymbol{x}) > 0\right\}$. It includes only the passive solution $\{\mathbf{0}_N\}$ if the payoffs to the base assets are linearly independent, but that is not the case in the index options application, as two distinct option combinations may have identical cash flows. Including these trivial solutions is inconsequential, as they cannot be optimal for any utility function if strict SAOs exist; moreover, if strict SAOs do not exist, then the default choice is the passive solution $\{\mathbf{0}_N\}$ which is equivalent to any trivial solution $\boldsymbol{\delta} \in \Delta^{\equiv}_{\mathcal{F}}$.

Including the non-trivial weak SAOs $\left(\Delta^{\succsim}_{(\cdot)} - \Delta^{\succ}_{(\cdot)} - \Delta^{\equiv}_{\mathcal{F}}\right)$ is also inconsequential. Any of these portfolios is accompanied by strict SAOs its close proximity which are indistinguishable from it in terms of the value for continuous objective functions. For example, mixing $\boldsymbol{\delta} \in \left(\Delta^{\succsim}_{(\cdot)} - \Delta^{\succ}_{(\cdot)} - \Delta^{\equiv}_{\mathcal{F}}\right)$ with $\mathbf{0}_N$ always yields a feasible strict SAO: for any $(u, c) \in \mathcal{U} \times (0, 1)$, we have that $D\left(u, \boldsymbol{\kappa}, \left(\left(c\boldsymbol{\delta} + (1-c)\mathbf{0}_N\right), \mathcal{F}\right) > cD\left(u, \boldsymbol{\kappa}, \boldsymbol{\delta}, \mathcal{F}\right) \geq 0\right.$, and thus, $(c\boldsymbol{\delta} + (1-c)\mathbf{0}_N) \in \Delta^{\succ}_{(\cdot)}$. As a result, $\left(\Delta^{\succsim}_{(\cdot)} - \Delta^{\equiv}_{\mathcal{F}}\right) \neq \emptyset \Leftrightarrow \Delta^{\succ}_{(\cdot)} \neq \emptyset$. In addition, as $c \uparrow 1$, the mixture becomes indistinguishable from the real thing.

An appealing specification for the objective function is the expected utility for some given utility function, or $G(v, \boldsymbol{\kappa}, \boldsymbol{\delta}, \mathcal{F}) := \mathbb{E}_{\mathcal{F}}\left[v(\boldsymbol{x}^{\mathrm{T}}(\boldsymbol{\kappa} + \boldsymbol{\delta}))\right]$, $v \in \mathrm{cl}(\mathcal{U})$. The choice of the host portfolio from the benchmark set K or a closed subset $\underline{\mathrm{K}} \subseteq \mathrm{K}$ may also be part of the optimization problem, to achieve additional improvements for investors who hold non-optimal elements of K. Approximate solutions are defined by a pre-defined tolerance parameter $\epsilon \geq 0$.

The following optimization problem, optimal solutions and approximate solutions are considered:

$$\max_{\Delta^{\succsim}_{(\cdot)} \times \underline{\mathrm{K}}} G(v, \boldsymbol{\kappa}, \boldsymbol{\delta}, \mathcal{F}); \tag{4}$$

$$\boldsymbol{\delta}^{*}_{(\mathcal{U},\mathrm{K})}(v, \mathcal{F}) := \arg\max_{\Delta^{\succsim}_{(\cdot)}} \left(\max_{\underline{\mathrm{K}}} G(v, \boldsymbol{\kappa}, \boldsymbol{\delta}, \mathcal{F})\right). \tag{5}$$

$$\Delta^{(\epsilon)}_{(\mathcal{U},\mathrm{K})}(v, \mathcal{F}) := \left\{\boldsymbol{\gamma} \in \Delta^{\succsim}_{(\cdot)} : \max_{\Delta^{\succsim}_{(\cdot)} \times \underline{\mathrm{K}}} G(v, \boldsymbol{\kappa}, \boldsymbol{\delta}, \mathcal{F}) - \max_{\underline{\mathrm{K}}} G(v, \boldsymbol{\kappa}, \boldsymbol{\gamma}, \mathcal{F}) \leq \epsilon\right\}. \tag{6}$$

The solution may not be unique, because v may be globally or locally linear. In this case, it will be clear from the context whether $\boldsymbol{\delta}^{*}_{(\cdot)}(v, \mathcal{F})$ represents an arbitrary selection from the solution set or a continuous representation of the entire set. The approximate solution set $\Delta^{(\epsilon)}_{(\cdot)}(v, \mathcal{F})$ contains a continuum of solutions for every $\epsilon > 0$ if an optimal solution exists.

The optimal solution generally features binding inequalities (equivalences), or $D(u, \boldsymbol{\kappa}, \boldsymbol{\delta}^{*}_{(\cdot)}(v, \mathcal{F}), \mathcal{F}) = 0$, for some host portfolios $\boldsymbol{\kappa} \in \mathrm{K}$ and utility functions $u \in \mathrm{cl}(\mathcal{U})$. Such equivalences occur for some $u \in \mathcal{U}$

if the optimal solution is not a strict SAO. In addition, equivalences for strict SAOs may occur for any $u \in (\mathrm{cl}(\mathcal{U}) - \mathcal{U})$, including possible trivial utility functions $u \in \mathcal{U}^{=} := \{u \in \mathrm{cl}\,(\mathcal{U}) : u(a) = u(b)\}$.

The existence of these binding inequalities raises the question whether the solution is robust to small perturbations of the probability distribution $\mathcal{F}$. The effect of the binding inequalities is however mitigated by the feature that approximate solutions $\boldsymbol{\delta} \in \Delta^{(\epsilon)}_{(\cdot)}(v, \mathcal{F})$ generally exist without non-trivial equivalence relations, for arbitrary $\epsilon > 0$, as is shown below.

If strict SAOs exist, then the optimum for a given utility function $v \in \mathcal{U}$ is superior to all host portfolios for that specific utility function, or $D(v, \boldsymbol{\kappa}, \boldsymbol{\delta}^*_{(\cdot)}(v, \mathcal{F}), \mathcal{F}) > 0$ for all $(v, \boldsymbol{\kappa}) \in \mathcal{U} \times \mathrm{K}$. It is generally harmless to assume that the same condition also applies for optima of any non-trivial utility function $v \in (\mathrm{cl}\,(\mathcal{U}) - \mathcal{U}^{=})$:

Definition 2.4.2 (Joint Enhancement). *The set of optimal SAOs $\boldsymbol{\delta}^*_{(\cdot)}(v, \mathcal{F})$, $v \in \mathrm{cl}\,(\mathcal{U})$, jointly enhance the benchmark set* K *if:*

$$\left(D(v, \boldsymbol{\kappa}, \boldsymbol{\delta}^*_{(\cdot)}(v, \mathcal{F}), \mathcal{F}) > 0\right) \ \forall\, (v, \boldsymbol{\kappa}) \in (\mathrm{cl}\,(\mathcal{U}) - \mathcal{U}^{=}) \times \mathrm{K}. \tag{7}$$

If $\mathcal{U}$ is closed, then $\mathrm{cl}\,(\mathcal{U}) - \mathcal{U}^{=} = \mathcal{U}$, and this condition is naturally satisfied if strict SAOs exist. If $\mathcal{U}$ is open, existence does not guarantee Joint Enhancement, in some rare cases. For example, if all arbitrage portfolios $\boldsymbol{\delta} \in \Delta$ have non-positive expected payoffs, then equivalence is inevitable for $v(x) = x$. Joint Enhancement should therefore be seen as a tight sufficient condition but not a necessary condition for existence (unless $\mathcal{U}$ is closed).

Joint Enhancement implies the existence of robust approximate solutions. To show this, mixtures of multiple optimal SAOs are considered, using mixing weights $w \in \mathcal{W}$, where $\mathcal{W}$ is the set of strictly monotone Borel measures defined on $\mathrm{cl}\,(\mathcal{U})$. A mixed optimal SAO is defined using the following Lebesgue integral: $\boldsymbol{\gamma}_{(\mathcal{U},\mathrm{K})}(w, \mathcal{F}) := \int_{(\mathrm{cl}(\mathcal{U}) - \mathcal{U}^{=})} \boldsymbol{\delta}^*_{(\cdot)}(u, \mathcal{F}) dw\,(u)\,,\ w \in \mathcal{W}$. Let w_v denote the degenerate measure at v.

Lemma 2.4.3 (Robust Neighbors). *If the base assets jointly enhance the benchmark set* K*, then, for any $\boldsymbol{\delta}^*_{(\cdot)}(v)$, $v \in (\mathrm{cl}\,(\mathcal{U}) - \mathcal{U}^{=})$, and tolerance $\epsilon > 0$, there exists some arbitrage portfolio $\boldsymbol{\gamma} = \boldsymbol{\gamma}_{(\mathcal{U},\mathrm{K})}\left((1 - c_\epsilon)\, w_v + c_\epsilon w, \mathcal{F}\right)$, $c_\epsilon \in (0, 1)$, $w \in \mathcal{W}$, which has the following properties: (i)* $D(u, \boldsymbol{\kappa}, \boldsymbol{\gamma}, \mathcal{F}) > 0$ for all $(u, \boldsymbol{\kappa}) \in (\mathrm{cl}\,(\mathcal{U}) - \mathcal{U}^{=}) \times \mathrm{K}$; *(ii)* $\boldsymbol{\gamma} \in \Delta^{(\epsilon)}_{(\cdot)}(v, \mathcal{F})$.

2.5 Numerical analysis

By Definition 2.4.1, (weak) SAOs are solutions to the following inequality system:

$$D(u, \boldsymbol{\kappa}, \boldsymbol{\delta}, \mathcal{F}) \geq 0 \; \forall \, (u, \boldsymbol{\kappa}) \in \mathrm{cl}(\mathcal{U}) \times \mathrm{K}; \qquad (8)$$

$$\boldsymbol{\delta} \in \Delta.$$

This system is analytically challenging for two reasons. First, closed-form solutions generally do not exists for the portfolio payoff distribution $\mathcal{F}_{\boldsymbol{\lambda}}(x) := \int_{\{\boldsymbol{x}:\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\lambda}\leq x\}} d\mathcal{F}(\boldsymbol{x})$ and the portfolio-level expected utility $\mathbb{E}_{\mathcal{F}}\left[u(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\lambda})\right] = \int u(x) d\mathcal{F}_{\boldsymbol{\lambda}}(x)$, even for relatively simple distributions such as the multivariate log-normal distribution. Second, the system has a semi-infinite structure due to the need to evaluate the first set of model restrictions for every $(u, \boldsymbol{\kappa}) \in \mathrm{cl}(\mathcal{U}) \times \mathrm{K}$, although the number of decision variables (M) is finite.

These complications generally call for some sort of discretization of $\mathcal{F}$, $\mathcal{U}$ and K to allow for numerical analysis using finite mathematical programming problems. The optimal computational strategy generally depends on the specific shape of $\mathcal{F}$, $\mathcal{U}$ and K in the application in question. A number of general approaches are discussed below.

Continuous distributions $\mathcal{F}$ can be discretized using, for example, Monte-Carlo simulation methods or lattice models. The trade-off between the achieved accuracy and the required computer burden naturally depends on the dimensions of the distribution. The use of factor models can be effective to reduce these dimensions in many applications, including the analysis in Section 4 below.

Since expected utility is a linear function of $u \in \mathrm{cl}(\mathcal{U})$, and $\mathrm{cl}(\mathcal{U})$ is a convex set, $\mathrm{cl}(\mathcal{U})$ may be replaced with the set of its extreme elements, or $\mathcal{U}^{(0)} : \mathrm{cl}(\mathcal{U}) = \mathrm{conv}\left(\mathcal{U}^{(0)}\right)$, without loss of generality. For many relevant specifications of $\mathcal{U}$, the extreme elements $u \in \mathcal{U}^{(0)}$ are low-dimensional functions, which reduces the numerical complexity of searching over $\mathrm{cl}(\mathcal{U})$.

For example, for SSD, the elementary function set $\mathcal{U}_2^{(0)}$ consists of the elementary Russell and Seo (1989) utility functions $v_{2;\phi}(x) = -(\phi - x)_+$, $\phi \in [a, b]$. The one-dimensional parameter space $[a, b]$ can easily be discretized with an arbitrary level of precision. The enhancement constraints $D(v_{2;\phi}, \boldsymbol{\kappa}, \boldsymbol{\delta}, \mathcal{F}) \geq 0$ are not smooth. However, they can be relaxed to an equivalent system of linear inequalities, for discrete distributions, as in Rockafellar and Uryasev (2000) and Dentcheva and Ruszczynski (2003, 2006), or, evaluated using iterative Mixed-Integer Linear Programming algorithms as in Klein Haneveld and van der Vlerk (2006), Kunzi-Bay and Mayer (2006) and Fábián, Mitra and Roman (2011).

Similarly, for general n-th degree Stochastic Dominance, the elementary utility functions are $v_{n;\phi}(x) = -(\phi - x)_{+}^{n-1}$, and $D(v_{n;\phi}, \boldsymbol{\kappa}, \boldsymbol{\delta}, \mathcal{F}) \geq 0$ is a convex non-smooth polynomial constraint which can be handled using Convex Polynomial Programming; see, for example, Post and Kopa (2017) for $n = 3$. For Infinite-degree Stochastic Dominance (Thistle (1993)), the elementary function set $\mathcal{U}_{\infty;0}$ consist of negative exponential functions $v_{\infty;\phi}(x) := \left(1 - e^{-\phi x}\right)$, for $\phi > 0$, and $v_{\infty;\phi}(x) := \frac{1}{b}x$, for $\phi = 0$, and the enhancement constraints are smooth and convex.

If the benchmark set K is continuous, then it may be approximated using a discrete subset of host portfolios $\boldsymbol{\kappa}_i \in \mathrm{K}$, $i = 1, \cdots, K$, for example, using a grid over the portfolio weight space. This approach can benefit from assuming that the host portfolios are efficient, so that K is the efficient set of an underlying portfolio set and the discretization uses a discrete subset of the efficient set, which is generally much smaller than the benchmark set.

One way to select efficient portfolios is to identify the optimizers for a number of representative utility functions, such as the elementary functions $\mathcal{U}^{(0)}$. Alternatively, efficient host portfolios can be selected from a set of candidates by using portfolio optimality tests and portfolio efficiency tests along the lines of Bawa, Bodurtha, Rao and Suri (1985) and Post (2003). A spanning test along the lines of Arvanitis, Hallam, Post and Topaloglou (2019) can be used to verify that the optimizers exactly or approximately span the entire benchmark set K. If spanning is rejected, the solutions for additional utility functions can be added as additional host portfolios.

2.6 Numerical example

A numerical example is developed to illustrate the key concepts and results from this section.

The investment universe consists of $N = 3$ base assets with prices $\boldsymbol{p} = \mathbf{1}_3$ and payoffs which follow a simple factor structure based on $M = 2$ mutually independent standard uniform factors, $f_1, f_2 \sim U(0, 1)$: $x_1 = 1; x_2 = \frac{5}{8} + f_1$; $x_3 = \frac{5}{8} + f_2$.

The market is incomplete due to the existence of investment constraints. The benchmark set K consists of a subset of convex mixtures of these three assets. Two distinct specifications are considered: $\mathrm{K}_1 := \{\boldsymbol{\kappa} \in \mathbb{R}_+^3 : \mathbf{1}_3^{\mathrm{T}}\boldsymbol{\kappa} = 1\}$; $\mathrm{K}_2 := \{\boldsymbol{\kappa} \in \mathrm{K}_1 : \boldsymbol{\kappa}_2 \geq 0.5; \boldsymbol{\kappa}_3 = 0\}$.

The overlay set Δ consists of all convex mixtures of three extreme arbitrage portfolios: $\boldsymbol{\delta}_1 = \mathbf{0}_3$; $\boldsymbol{\delta}_2 = \left(-\frac{1}{4}\ \frac{1}{4}\ 0\right)^{\mathrm{T}}$; $\boldsymbol{\delta}_3 = \left(0\ -\frac{1}{4}\ \frac{1}{4}\right)^{\mathrm{T}}$. The payoffs to these extreme portfolios amount to $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\delta}_1 = 0$; $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\delta}_2 = \frac{1}{4}f_1 - \frac{3}{32}$; $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\delta}_3 = \frac{1}{4}\left(f_2 - f_1\right)$. It is straightforward to see that none of the arbitrage portfolios is a pure arbitrage possibility, and to find positive pricing kernels which correctly price the base assets. The question is whether generalized arbitrage opportunities $\boldsymbol{\delta} \in \Delta^{\succ}_{(\mathcal{U},\mathrm{K},\mathcal{F})}$ or restricted pricing kernels $m \in$conv $\left(\mathcal{M}_{(\mathcal{U},\mathrm{K})}\right)$ exist and can be identified. All risk averse investors are considered: $\mathcal{U} = \mathcal{U}_2$.

Due to the independence of the factors, diversification is an important source of portfolio enhancement. For the specification $\mathrm{K} = \mathrm{K}_1$, the arbitrage portfolios add no diversification benefits beyond what can be achieved using the benchmark set, and SAOs do not exist. For example, the host portfolio $\boldsymbol{\kappa}_1 = \left(\frac{1}{4}\ \frac{3}{8}\ \frac{3}{8}\right)^{\mathrm{T}}$ is a maximizer of $\mathbb{E}_{\mathcal{F}}\left[u(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\kappa})\right]$, and cannot be enhanced, for $u(x) = -\frac{1}{3}x^{-3}$; the arbitrage portfolios are redundant, as they relax only non-binding investment constraints, for this utility function.[1]

It follows that $m = \mathbb{E}_{\mathcal{F}}\left[\left(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\kappa}_1\right)^{-4}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\kappa}_1\right]^{-1}(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\kappa}_1)^{-4}$ is a qualified pricing kernel. The solution to arbitrage portfolio optimization problem (4) is simply $\mathbf{0}_N$, for every host portfolio set $\underline{\mathrm{K}} \subseteq \mathrm{K}$ and utility function $v \in \mathrm{cl}\,(\mathcal{U}_2)$.

The conclusions change when the benchmark set is reduced to $\mathrm{K} = \mathrm{K}_2$. The additional investment restrictions have the effect of relaxing the conditions for existence of SAOs, and tightening the conditions for existence of qualified pricing kernels. Some arbitrage portfolios now offer unique diversification benefits which lead to enhancements for all risk averters. It follows that no qualified pricing kernel exists. For example, $\boldsymbol{\delta}_3$ is a strict SAO and the optimal SAO $\boldsymbol{\delta}^{*}_{(\mathcal{U},\mathrm{K})}(v,\mathcal{F})$ for $v(x) = -\frac{1}{3}x^{-3}$ and $\underline{\mathrm{K}} = \mathrm{K}$; this arbitrage portfolio relaxes the portfolio constraint $\boldsymbol{\kappa}_3 = 0$ which is binding for all host portfolio and utility functions.

This particular SAO has a non-trivial equivalence relation with every feasible host portfolio, for $u(x) = x$, due to $\mathbb{E}_{\mathcal{F}}\left[\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\delta}_3\right] = 0$, and therefore it is not a robust weak SAO, and it may not be detected using numerical optimization due to rounding errors. However, due to the availability of $\boldsymbol{\delta}_2$ in the overlay set and $\mathbb{E}_{\mathcal{F}}\left[\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\delta}_2\right] = \frac{1}{32} > 0$, Joint Enhancement (7) occurs. Hence, the neighborhood of the optimal SAO $\boldsymbol{\delta}_3$ includes robust SAOs which are qualified substitutes for $\boldsymbol{\delta}_3$. For example, the Robust Neighbor $\boldsymbol{\delta}_4 = 0.01\boldsymbol{\delta}_2 + 0.99\boldsymbol{\delta}_3$ is a strict SAO without binding equivalence relations for any $u \in \mathrm{cl}\,(\mathcal{U}_2)$.

3 Statistical Theory

This section introduces empirical counterparts for the theoretical concepts and develops an asymptotic theory for these constructs. The focus is on the consistent estimation of the optimal solution set, as non-optimal arbitrage portfolios are irrelevant for asset pricing and portfolio choice. Interestingly, the Joint Enhancement property (see Definition 2.2.1) facilitates approximation of the optimal solution set without the use of non-

[1] The optimum can be found in this case in an analytical way. The risky assets x_2 and x_3 are mutually i.i.d. and subject to identical weight constraints. Thus, they can be optimally held in equal proportion, and we are left with the simpler problem of mixing riskfree asset x_1 and risky alternative $x^* := \frac{1}{2}(x_2 + x_3) = \frac{5}{8} + \frac{1}{2}(f_1 + f_2)$. The combined risky alternative x^* obeys a triangular distribution with mean $\frac{9}{8}$ and variance $\frac{1}{24}$. Different mixtures of x_1 and x^* yield a different distributional location and scale but the same triangular shape. Hence, the optimal mixing weight of x^* is found using the Merton fraction $w^* := (\frac{9}{8}-1)/(4\cdot\frac{1}{24}) = \frac{3}{4}$ based on the means and variances and a level of Relative Risk Aversion of 4.

trivial data-dependent moment selection techniques.

To facilitate compact presentation, the analysis first establishes the main results based on a general and sufficient high-level assumption for the empirical process of the SD conditions. This assumption is subsequently validated using detailed and testable lower–level assumptions. Separate sets of lower-level assumptions are presented for non-parametric and parametric estimation. Special attention is given to lower-level assumptions which are relevant for the analysis of stock index options in Section 4. A concluding section discusses various extensions which seem of interest for other applications.

3.1 High-level statistical assumption framework

To simplify notation, $\mathcal{F}$ now denotes the latent joint distribution of $\boldsymbol{f}, \boldsymbol{\varepsilon}$, and/or $\boldsymbol{x} = g(\boldsymbol{f}) + \boldsymbol{\varepsilon}$. An estimator $\mathcal{F}_T$ is constructed based on a time series sample $(\boldsymbol{x}_t, \boldsymbol{f}_t)_{t=1,\ldots,T}$ and possible conditioning observables. The estimation may be based on a well-specified (semi-) parametric model or a fully non-parametric method, depending on the data properties and available side information.

The limit theory evolves as $T \to \infty$. The squiggly arrow $\rightsquigarrow$ denotes convergence in distribution. The analysis also uses Painleve-Kuratowski (PK) convergence $\left(\overset{\text{PK}}{\rightsquigarrow}\right)$ for sequences of compact subsets of the overlay set $\triangle$. With high probability (w.h.p.) refers to probability converging to one.

A high-level assumption is used for the asymptotic behavior of the empirical process $D(u, \boldsymbol{\kappa}, \boldsymbol{\delta}, \mathcal{F}_T - \mathcal{F})$, after proper scaling. Let $\ell^\infty(A)$ denote the set of bounded real functions defined on set $A \neq \emptyset$, equipped with the sup norm.

Assumption 3.1.1 (Empirical Process Convergence). *For any $(u, \boldsymbol{\kappa}, \boldsymbol{\delta}) \in cl(\mathcal{U}) \times K \times \triangle$, there exists some positive real sequence $m_T \to \infty$, for which*

$$m_T D(u, \boldsymbol{\kappa}, \boldsymbol{\delta}, \mathcal{F}_T - \mathcal{F}) \rightsquigarrow \mathcal{G}(u, \boldsymbol{\kappa}, \boldsymbol{\delta}),$$

where $\mathcal{G}(\cdot,\cdot,\cdot)$ denotes a tight stochastic process with sample paths in $\ell^\infty(cl(\mathcal{U}) \times K \times \triangle)$.

The empirical process $D(u, \boldsymbol{\kappa}, \boldsymbol{\delta}, \mathcal{F}_T - \mathcal{F})$ depends on the utility function in question; sufficient primitive conditions for the time-series process and the structure of the CDF estimator are discussed in Section 3.3. Specific expressions for the scaling factor m_T and limiting process $\mathcal{G}$ are then obtained.

The scaling factor m_T may be the usual rate $\sqrt{T}$ or a slower rate due to the non-existence of sufficiently high moments for the asymptotic representations of (semi-) parametric estimators or, alternatively, the dependence of the procedure on tuning parameters like the bandwidth in non-parametric settings.

The limiting process $\mathcal{G}$ is zero-mean Gaussian, under the lower-level assumptions in Section 3.3, but, more generally, it may also be a zero-mean stable process, or a projection of such processes on cones associated with the parameters $u, \boldsymbol{\kappa}, \boldsymbol{\delta}$; see Section 3.4.

3.2 Asymptotic Results

Using the estimator $\mathcal{F}_T$, the following empirical counterpart of the weak SAO set is constructed:

$$\Delta^{\gtrsim}_{(\mathcal{U},\mathrm{K},\mathcal{F}_T)} := \left\{ \boldsymbol{\delta} \in \triangle : (\boldsymbol{\kappa} + \boldsymbol{\delta}) \gtrsim_{(\mathcal{U},\mathcal{F}_T)} \boldsymbol{\kappa}, \ \forall \boldsymbol{\kappa} \in \mathrm{K} \right\}. \tag{9}$$

The following result is obtained for this empirical set:

Theorem 3.2.1 (PK Consistency). *If Empirical Process Convergence (Assumption 3.1.1) holds, then the following results are obtained: (i) If* $\left(\Delta^{\gtrsim}_{(\cdot)} - \Delta^{\equiv}_{\mathcal{F}}\right) \neq \emptyset$, *and Joint Enhancement (7) occurs, then for any* $v \in (\mathrm{cl}\,(\mathcal{U}) - \mathcal{U}^{=})$*, and for any* $c_T \to 0$ *such that* $m_T c_T \to \infty$*, we have that* $\Delta^{\gtrsim}_{(\mathcal{U},\mathrm{K},\mathcal{F}_T)} \cap \Delta^{(c_T)}_{(\mathcal{U},\mathrm{K})}(v, \mathcal{F}) \overset{PK}{\rightsquigarrow} \boldsymbol{\delta}^*_{(\mathcal{U},\mathrm{K})}(v, \mathcal{F})$*. (ii) If* $\left(\Delta^{\gtrsim}_{(\cdot)} - \Delta^{\equiv}_{\mathcal{F}}\right) = \emptyset$*, then* $\Delta^{\gtrsim}_{(\mathcal{U},\mathrm{K},\mathcal{F}_T)} \overset{PK}{\rightsquigarrow} \Delta^{\equiv}_{\mathcal{F}}$*.*

The theorem distinguishes between the case in which non-trivial weak SAOs jointly enhance all host portfolios and the case in which non-trivial weak SAOs do not exist. In the former case, the empirical weak SAO set contains w.h.p. elements of the set of approximate optimizers $\Delta^{(c_T)}_{(\mathcal{U},\mathrm{K})}(v, \mathcal{F})$ which converge to the optimal set $\boldsymbol{\delta}^*_{(\mathcal{U},\mathrm{K})}(v, \mathcal{F})$, for any nontrivial $v \in (\mathrm{cl}\,(\mathcal{U}) - \mathcal{U}^{=})$, and any approximation error c_T that converges to zero at a sufficiently slow rate. In the latter case, the limit contains only the trivial weak SAOs.

If the utility class is closed, or $\mathcal{U} = \mathrm{cl}(\mathcal{U})$, then these two cases are jointly exhaustive, because the existence of non-trivial weak SAOs then implies Joint Enhancement. However, if the utility class is open, then the conditions do not cover the case where non-trivial weak SAOs exist but Joint Enhancement does not occur. In this case, it is possible that optimal SAOs do not appear in the empirical set with asymptotically positive probability. This exceptional case is further discussed in Section 3.4.

The empirical counterpart of optimization problem (4) maximizes the goal function $\max_{\underline{\mathrm{K}}} G(v, \boldsymbol{\kappa}, \boldsymbol{\delta}, \mathcal{F}_T) = \max_{\boldsymbol{\kappa} \in \underline{K}} \mathbb{E}_{\mathcal{F}_T}\left[v(\boldsymbol{x}^{\mathrm{T}}(\boldsymbol{\kappa} + \boldsymbol{\delta}))\right],\ v \in cl(\mathcal{U})$, subject to the empirical enhancement condition $\boldsymbol{\delta} \in \Delta^{\gtrsim}_{(\mathcal{U},\mathrm{K},\mathcal{F}_T)}$. The following result shows that the empirical solutions approximate the original as $T \to \infty$:

Theorem 3.2.2 (Empirical Solution Properties). *If Empirical Process Convergence (Assumption 3.1.1) holds, then the following results are obtained, for any* $v \in cl(\mathcal{U}) - \mathcal{U}^{=}$*: i)*

$$\max_{\boldsymbol{\delta} \in \Delta^{\gtrsim}_{(\mathcal{U},\mathrm{K},\mathcal{F}_T)}} \left(\max_{\underline{\mathrm{K}}} G(v, \boldsymbol{\kappa}, \boldsymbol{\delta}, \mathcal{F}_T) \right) \rightsquigarrow \sup_{\boldsymbol{\delta} \in \Delta^{0}_{(\mathcal{U},\mathrm{K},\mathcal{F})}} \left(\max_{\underline{\mathrm{K}}} G(v, \boldsymbol{\kappa}, \boldsymbol{\delta}, \mathcal{F}) \right), \tag{10}$$

where, $\Delta^0_{(\mathcal{U},\mathrm{K},\mathcal{F})} := \boldsymbol{\delta}^*_{(\mathcal{U},\mathrm{K})}(v,\mathcal{F})$ *if* $\left(\Delta^{\gtrsim}_{(\cdot)} - \Delta^{\equiv}_{\mathcal{F}}\right) \neq \emptyset$ *and Joint Enhancement (7) occurs, and* $\Delta^0_{(\mathcal{U},\mathrm{K},\mathcal{F})} := \Delta^{\equiv}_{\mathcal{F}}$ *if* $\left(\Delta^{\gtrsim}_{(\cdot)} - \Delta^{\equiv}_{\mathcal{F}}\right) = \emptyset$. *ii) If* $\left(\Delta^{\gtrsim}_{(\cdot)} - \Delta^{\equiv}_{\mathcal{F}}\right) \neq \emptyset$ *and Joint Enhancement (7) occurs, then every limit of any subsequence of elements of* $\boldsymbol{\delta}^*_{(\mathcal{U},\mathrm{K})}(v,\mathcal{F}_T)$ *lies in* $\boldsymbol{\delta}^*_{(\mathcal{U},\mathrm{K})}(v,\mathcal{F})$. *iii) If* $\left(\Delta^{\gtrsim}_{(\cdot)} - \Delta^{\equiv}_{\mathcal{F}}\right) = \emptyset$*, then* $\boldsymbol{\delta}^*_{(\mathcal{U},\mathrm{K})}(v,\mathcal{F}_T) \overset{PK}{\rightsquigarrow} \Delta^{\equiv}_{\mathcal{F}}$.

According to Part (i), the empirical optimal value of the goal function approximates its latent population value based on (a) the set of optimal weak SAOs when non-trivial weak SAOs exist and jointly enhance the host portfolios, and (b) the set of trivial weak SAOs when non-trivial weak SAOs do not exist. Part (ii) means that any portfolio at which the empirical solutions accumulate will be an optimal SAO when non-trivial weak SAOs exist. Part (iii) says that the empirical solutions approximate the set of trivial weak SAOs w.h.p. when non-trivial weak SAOs do not exist.

The theorem is reassuring about the possibilities of two types of decision errors which can occur in empirical portfolio optimization. The probability of suboptimal choice, or a material opportunity loss for the objective function, vanishes in large samples, as the empirical weak SAO set includes w.h.p. the set of optimal SAOs. In addition, the probability of selecting an arbitrage portfolio which is not a SAO, and thus leads to deterioration $D(u,\boldsymbol{\kappa},\boldsymbol{\delta},\mathcal{F}) < 0$ for some host portfolio and utility function, also vanishes asymptotically, because the empirical enhancement $D(u,\boldsymbol{\kappa},\boldsymbol{\delta},\mathcal{F}_T)$ diverges from zero in large samples if $D(u,\boldsymbol{\kappa},\boldsymbol{\delta},\mathcal{F}) < 0$.

3.3 Lower-Level Examples

Empirical Process Convergence (Assumption 3.1.1) is now motivated using more specific assumptions, notably ones which are relevant for the analysis of stock index options in Section 4. Since the cash flows of the options are fully determined by the value of the underlying index, a perfect factor model applies: $\boldsymbol{x} = g(\boldsymbol{f})$ and $\boldsymbol{\varepsilon} := \mathbf{0}_{N\times 1}$. Without loss of generality, the lower-level examples are based on the class of mixed utility functions $\mathcal{U} = \mathcal{U}_2$ (see Section 2.1); the results carry over directly to convex subclasses $\mathcal{U} \subset \mathcal{U}_2$.

Conditioning variables are used to capture the time-variation of the payoff distribution. Given an L-vector of conditioning variables $\boldsymbol{Z}$, $\mathcal{F}$ is the joint CDF of $\boldsymbol{f}$ and $\boldsymbol{x}$ conditional on the σ-algebra generated by $\boldsymbol{Z}$. $\mathbb{E}_{\mathcal{F}}[\cdot]$ denotes integration w.r.t. the conditional distribution; $\mathbb{E}[\cdot]$ denotes unconditional expectation. $\mathcal{F}$ has a density function f. $\|\cdot\|$ denotes the Euclidean norm on $\mathbb{R}^n$, and $\|\cdot\|_F$ the Frobenius norm on $\mathbb{R}^{n\times n}$ with n depending on the context.

The analyst has at her disposal an observable time series $(\boldsymbol{x}_t, \boldsymbol{f}_t, \boldsymbol{Z}_t)_{t=1,\ldots,T}$, to estimate $\mathcal{F}$. The conditional distribution is evaluated at specific values of the conditioners $\boldsymbol{z}$ which characterize the prevailing market conditions at the portfolio formation date. These values lie in the support of $\boldsymbol{Z}$ and are known to the analyst.

Parametric estimation

A well-specified d-dimensional parametric model takes the form of a family of densities $\{f_\theta, \theta \in \Theta\}$ with parameter space $\Theta \subseteq \mathbb{R}^d$. The model is well specified in the sense that the latent population density arises as a special case $f = f_{\theta_0}$ for $\theta_0 \in \Theta$. Using the sample and the model, a parameter estimator θ_T is constructed which yields the density estimatorf_{θ_T} and the CDF estimator $\mathcal{F}_T(\boldsymbol{x}) := \int_{\prod_{i=1}^{M}(-\infty, \boldsymbol{x}_i]} f_{\theta_T}(\boldsymbol{y})\, d\boldsymbol{y}$.

The following composite lower-level assumption suffices to validate Empirical Process Convergence (Assumption 3.1.1) and to invoke PK Convergence (Theorem 3.2.1) and Empirical Solution Properties (Theorem 3.2.2) in this parametric setting:

Assumption 3.3.1 (Parametric Primitive Assumptions). *i) The stochastic process* $(\boldsymbol{f}_t, \boldsymbol{Z}_t)_{t \in \mathbb{Z}}$ *is stationary and strongly mixing, with mixing coefficients* $(a_j)_{j \in \mathbb{N}}$ *that asymptotically satisfy* $\sum_{j=0}^{\infty} j^{1+\frac{2}{2+\delta}} a_j^{\frac{\delta}{2+\delta}} < +\infty$ *for some* $\delta > 0$. *ii)* θ_0 *lies in the interior of a closed and convex set* $\Theta^\star \subseteq \Theta$ *which also contains* θ_T *w.h.p. Whenever* θ_T *lies in* $\Theta^\star$, *it admits the asymptotic representation* $\theta_T = \theta_0 + H_T(\theta_T^\star)\, s_T(\theta_0) + o_p(1)$. *Here,* $\theta_T^\star \in \Theta^\star$ *lies on the line that connects* θ_T *and* θ_0, $H_T(\theta) := \frac{1}{T}\sum_{t=1}^{T} H(\theta, \boldsymbol{f}_t, \boldsymbol{Z}_t)$ *for some symmetric matrix-valued* $H : \Theta^\star \times \mathbb{R}^{M+L} \to \mathbb{R}^{d \times d}$ *which is continuous w.r.t.* θ, *and* $s_T(\theta) := \frac{1}{T}\sum_{t=1}^{T} s(\theta, \boldsymbol{f}_t, \boldsymbol{Z}_t)$ *for the score function* $s : \Theta^\star \times \mathbb{R}^{M+L} \to \mathbb{R}^d$. *Furthermore,* $\mathbb{E}\left[\sup_{\Theta^\star} \|H(\theta, \boldsymbol{f}_0, \boldsymbol{Z}_0)\|_F\right] < \infty$, $\mathbb{E}\left[s(\theta_0, \boldsymbol{f}_0, \boldsymbol{Z}_0)\right] = 0_{d \times 1}$, $\mathbb{E}\left[\|s(\theta_0, \boldsymbol{f}_0, \boldsymbol{Z}_0)\|^{2+\delta}\right] < +\infty$. *iii) For all* $\theta \in \Theta^\star$, *the support of* f_θ *is independent of* θ, $f_\theta(\boldsymbol{x}) > 0$ *and* $f_\theta(\boldsymbol{x})$ *is continuously differentiable w.r.t.* θ, *for all* $\boldsymbol{x}$ *inside the common support, and* $\sup_{\Theta^\star} \mathbb{E}\left[\|l_\theta\|^{2+\delta}\right] < +\infty$, *where* $l_\theta := \frac{1}{f_\theta}\frac{\partial f_\theta}{\partial \theta}$.

Part (i) holds for many dynamic models that are commonly used in Finance such as ARMA or GARCH-type models under standard assumptions (see, for example, Mikosch and Straumann (2006), Section 4). Part (ii) is standard for a variety of M-estimators derived from optimization of smooth enough objective functions. For example, if θ_T is the Maximum Likelihood Estimator, it would follow under (i), if the average log-likelihood function has an integrable supremum over Θ, its expectation is uniquely maximized at θ_0, and if (iii) were extended to two-times continuous differentiability for all $\boldsymbol{x}$. Part (iii) holds for many parametric statistical models with the notable exception of families of uniform distributions.

Proposition 3.3.2 (Parametric Validation). *If the Parametric Primitive Assumptions (Assumption 3.3.1) hold, then Empirical Process Convergence (Assumption 3.1.1) holds with* $m_T = \sqrt{T}$ *and* $\mathcal{G}(u, \boldsymbol{\kappa}, \boldsymbol{\delta})$

equal to a zero-mean Gaussian process with the following covariance kernel:

$$
\begin{gathered}
V_{\mathcal{G}}\left(u, \boldsymbol{\kappa}, \boldsymbol{\delta}, u^{*}, \boldsymbol{\kappa}^{*}, \boldsymbol{\delta}^{*}\right):= \\
\mathbb{E}_{\mathcal{F}}\left[\left(u\left(g\left(\boldsymbol{f}_{0}\right)^{T}(\boldsymbol{\kappa}+\boldsymbol{\delta})\right)-u\left(g\left(\boldsymbol{f}_{0}\right)^{T} \boldsymbol{\kappa}\right)\right) l_{\theta_{0}}\left(\boldsymbol{f}_{0}\right)\right] \times \\
\mathbb{E}\left[H\left(\theta_{0}, \boldsymbol{f}_{0}\right)\right] \sum_{t \in \mathbb{Z}} Cov\left(s\left(\theta_{0}, \boldsymbol{f}_{0}\right), s\left(\theta_{0}, \boldsymbol{f}_{t}\right)\right) \mathbb{E}\left[H\left(\theta_{0}, \boldsymbol{f}_{0}\right)\right] \times \\
\mathbb{E}_{\mathcal{F}}\left[\left(u^{*}\left(g\left(\boldsymbol{f}_{0}\right)^{T}\left(\boldsymbol{\kappa}^{*}+\boldsymbol{\delta}^{*}\right)\right)-u^{*}\left(g\left(\boldsymbol{f}_{0}\right)^{T} \boldsymbol{\kappa}^{*}\right)\right) l_{\theta_{0}}^{T}\left(\boldsymbol{f}_{0}\right)\right].
\end{gathered}
$$

In this case, m_T has the standard parametric rate $\sqrt{T}$. The covariance kernel of the limiting process depends on the conditional covariance between the empirical enhancement conditions and the score of the log-density at the true parameters, as well as on the limiting variance of the parameter estimator. The first part represents the effect of the local uncertainty about $\mathcal{F}$ due to the unknown θ_0, and the second part captures the effect of the estimation error.

Non-parametric estimation

The Nadaraya-Watson (NW) estimator of $\mathcal{F}$ is used as a representative non-parametric estimator. It is widely used in the theory and empirical applications of non-parametric regression. The results can easily be extended to the more general class of local polynomial estimators, at the cost of heavier notation, definitions and derivations. Using the NW estimator $\mathcal{F}_T$, the empirical enhancement $D(x, \boldsymbol{\kappa}, \boldsymbol{\delta}, \mathcal{F}_T)$ amounts to

$$
\begin{gathered}
\mathbb{E}_{\mathcal{F}_{T}}\left[u\left(g(\boldsymbol{f})^{\mathrm{T}}(\boldsymbol{\kappa}+\boldsymbol{\delta})\right)-u\left(g(\boldsymbol{f})^{\mathrm{T}} \boldsymbol{\kappa}\right)\right]= \\
\sum_{t=1}^{T} w_{\boldsymbol{z}, t}\left[u\left(g\left(\boldsymbol{f}_{t}\right)^{\mathrm{T}}(\boldsymbol{\kappa}+\boldsymbol{\delta})\right)-u\left(g\left(\boldsymbol{f}_{t}\right)^{\mathrm{T}} \boldsymbol{\kappa}\right)\right]
\end{gathered},
$$

where $w_{\boldsymbol{z},t} := \mathcal{K}\left(\frac{\boldsymbol{z}_t-\boldsymbol{z}}{b_T}\right) / \sum_{t=1}^{T} \mathcal{K}\left(\frac{\boldsymbol{z}_t-\boldsymbol{z}}{b_T}\right)$, $\mathcal{K}$ is a convolution kernel function on $\mathbb{R}^n$ (see Giné and Nickl (2016)), and b_T is a bandwidth parameter.

If the smooth kernel function $\mathcal{K}$ is replaced by the discrete uniform kernel $\frac{1}{T}$, due to the absence of conditioning information, then the simple Empirical CDF is obtained. This estimator appears in several applications of portfolio analysis and portfolio optimization with SD constraints in the OR/MS literature. It seems inappropriate here, given the need for conditional CDF estimates in the options application.

The following lower-level assumption validates Empirical Process Convergence (Assumption 3.1.1) and allows for invoking PK Convergence (Theorem 3.2.1) and Empirical Solution Properties (Theorem 3.2.2).

Assumption 3.3.3 (Non-parametric Primitive Assumptions). *i) Assumption A.3.3.1.i) holds. ii)* $\mathcal{K}$ *is a positive symmetric bounded and Lipschitz continuous convolution kernel on* $\mathbb{R}^L$ *such that* $\int_{\mathbb{R}^L} \mathcal{K}(\boldsymbol{y})\, d\boldsymbol{y} = 1$,

$\int_{\mathbb{R}^L} \|\boldsymbol{y}\|^2 \mathcal{K}(\boldsymbol{y})\, d\boldsymbol{y} < +\infty$. *ii) For all $t \in \mathbb{Z}^\star$, the distribution of $(\boldsymbol{Z}_0, \boldsymbol{Z}_t)$ has a density $f_{0,t}$, for which there exists $C > 0$ independent of t, and for all $(\boldsymbol{u}, \boldsymbol{y}) \in \mathbb{R}^{2L}$, $|f_{0,t}(\boldsymbol{u}, \boldsymbol{y}) - f_{\boldsymbol{Z}}(\boldsymbol{u}) f_{\boldsymbol{Z}}(\boldsymbol{y})| \leq C$, where $f_{\boldsymbol{Z}}$ is the density of $\boldsymbol{Z}_0$. iii) As $T \to \infty$, $b_T \downarrow 0$, $Tb_T^L \to \infty$, $Tb_T^{L+4} \to 0$. iv) The joint distribution of $(\boldsymbol{f}_0, \boldsymbol{Z}_0)$ has a continuous and bounded density $f_{\boldsymbol{f},\boldsymbol{Z}}(\boldsymbol{f}, \boldsymbol{y})$ that is twice differentiable w.r.t. $\boldsymbol{y}$ with uniformly bounded and Lipschitz continuous second derivatives that are integrable w.r.t. $\boldsymbol{f}$. v) For some $c > 0$, the function $(0, c] \ni b \to \frac{\mathcal{K}\left(\frac{\boldsymbol{y}}{b}\right)}{b^L}$ is Lipschitz continuous, with Lipschitz coefficient $\kappa(\boldsymbol{y})$ such that $\int_{\mathbb{R}^{M+L}} \kappa^2(\boldsymbol{y}) f_{\boldsymbol{f},\boldsymbol{Z}}(\boldsymbol{f}, \boldsymbol{y})\, d\boldsymbol{f} d\boldsymbol{y} < +\infty$.*

Assumption 3.3.3.ii)-v) imposes usual conditions in non-parametric statistics (see El Machkouri, Fan, and Reding, 2020, and references therein). Assumption 3.3.3.v) holds whenever the kernel is differentiable and $\sup_{b \in (0,c]} \frac{\left\|\mathcal{K}'\left(\frac{\boldsymbol{y}}{b}\right)\right\|_F}{b^{L+1}} < +\infty$, where $\mathcal{K}'$ is the Jacobian. This condition holds, for example, for the Gaussian kernel.

Proposition 3.3.4 (Parametric Validation). *If the Non-parametric Primitive Assumptions (Assumption 3.3.3) hold, then Empirical Process Convergence (Assumption 3.1.1) holds with $m_T = \sqrt{Tb_T^L}$ and $\mathcal{G}(u, \boldsymbol{\kappa}, \boldsymbol{\delta})$ equal to a zero-mean Gaussian process with the following covariance kernel:*

$$V_{\mathcal{G}}(u, \boldsymbol{\kappa}, \boldsymbol{\delta}, u^* \boldsymbol{\kappa}^*, \boldsymbol{\delta}^*) := M_{\boldsymbol{z}}\, Cov_{\mathcal{F}}\left(\mathcal{D}(u, \boldsymbol{\kappa}, \boldsymbol{\delta}), \mathcal{D}(u^*, \boldsymbol{\kappa}^*, \boldsymbol{\delta}^*)\right) ,$$

where $M_{\boldsymbol{z}} := f_{\boldsymbol{Z}}(\boldsymbol{z}) \int_{\mathbb{R}^L} \mathcal{K}^2(z)\, dz$, $\mathcal{D}(u, \boldsymbol{\kappa}, \boldsymbol{\delta}) := \ u\left(g(\boldsymbol{f}_0)^T (\boldsymbol{\kappa} + \boldsymbol{\delta})\right) - u\left(g(\boldsymbol{f}_0)^T \boldsymbol{\kappa}\right)$.

Here, m_T has a slower than $\sqrt{T}$, non-parametric rate due to the presence of the bandwidth parameter. The covariance kernel of the limiting process depends on the conditional covariance between the empirical enhancement conditions, as well as on the density of the conditioning variables and the L_2 norm of the kernel.

3.4 Extensions

The limit theory depends critically on Joint Enhancement (7), if $\mathcal{U}$ is open. When this assumption is violated, Robust Neighbors may not exist, and convergence of the empirical solutions to the optimal SAOs may fail. In such cases, the results of Theorems 3.2.1-3.2.2 may be rectified by using slack variables. Specifically, if the empirical enhancement conditions include (potentially) data-dependent slacks which converge to zero at a strictly slower rate than the rate at which m_T diverges to infinity, it is possible to show that, when strict SAOs exist, a) $\Delta^{\gtrsim}_{(\mathcal{U},\mathrm{K},\mathcal{F}_T)} \overset{\mathrm{PK}}{\rightsquigarrow} \Delta^{\gtrsim}_{(\mathcal{U},\mathrm{K},\mathcal{F})}$, a result which is result stronger than Theorem 3.2.1.(i), and b) Theorem 3.2.2 holds.

The previous paragraphs did not explicitly consider the discretization schemes for $\mathrm{cl}\,(\mathcal{U})$ and K which were proposed in Section 2.5. However, it is straightforward to show that the current limit theory is also valid for any scheme that allows for the PK convergence of the solution set of the discretized SAO system to $\Delta^{\gtrsim}_{(\mathcal{U},\mathrm{K},\mathcal{F}_T)}$, either as $T \to \infty$, or even for fixed T.

The factors-to-payoffs mapping g could be allowed to depend on unknown low-dimensional parameters. The effect of these additional parameters can be studied using conditions of smoothness on g as a function of those parameters, and asymptotic representation conditions of the parameter estimators, similar to Assumption 3.3.1.(ii) in the parametric example. The parameter estimation error would not affect our consistency results if, for example, the rates of convergence of the parameters asymptotically dominate or are equivalent to m_T.

The lower-level analysis assumes a perfect factor model, or $\boldsymbol{\varepsilon} = \mathbf{0}_{N\times 1}$. This assumption can be relaxed as long as a filter $(\tilde{\boldsymbol{\varepsilon}}_t)_{t=1,\dots,T}$ is available that approximates $(\boldsymbol{\varepsilon}_t)_{t=1,\dots,T}$. The joint time series $(\tilde{\boldsymbol{\varepsilon}}_t, \boldsymbol{f}_t, \boldsymbol{Z}_t)_{t=1,\dots,T}$, can then be used to estimate $\mathcal{F}$. To mitigate the curse of dimension for the derivation of the filter and the estimation of $\mathcal{F}$, side information for the joint distribution of $(\boldsymbol{\varepsilon}_t, \boldsymbol{f}_t, \boldsymbol{Z}_t)_{t=1,\dots,T}$ is generally required. One possible assumption is that the processes $(\boldsymbol{f}_t)$ and (ε_t) are independent conditionally on $(\boldsymbol{Z}_t)$, $(\boldsymbol{f}_t)$ adheres to a low dimensional parametric model, and (ε_t) follows some elliptical joint distribution with zero mean and cross sectional and temporal covariance matrices that, as functions of $\boldsymbol{Z}$, also depend on low-dimensional parameters.

Assumption 3.3.1.(b) can be extended to cases were the parametric rates are slower than $\sqrt{T}$ and the limiting processes are not Gaussian, without affecting the results. A typical example would be the case of GARCH-type martingale differences factors with innovations that lie in the domain of attractions of stable laws (see, for example, Mikosch and Straumann, 2006).

The NW estimator in the non-parametric example can be replaced by the implied probability distribution of the Smoothed Empirical Likelihood method (see Kitamura, Tripathi, and Ahn, 2004) whenever side information about $\mathcal{F}$ is available in the form of conditional moment inequalities. In such cases, the limiting processes would be projections of Gaussian processes similar to Proposition 3.3.4 on parameter cones.

Beyond consistent estimation, our framework can be used for further steps of statistical inference about SAOs. For example, let $\Delta^\star \subset \Delta$ be closed and disjoint to $\Delta^{\equiv}_{\mathcal{F}}$. Then a consistent and asymptotically conservative resampling test about the null hypothesis of $\Delta^\star \cap \Delta^{\gtrsim}_{(\mathcal{U},\mathrm{K},\mathcal{F})} \neq \emptyset$ can be based on the statistic $\sup_{\Delta^\star} \inf_{\mathcal{U}\times\mathrm{K}} D(u, \boldsymbol{\kappa}, \boldsymbol{\delta}, \mathcal{F}_T)$.

4 Equity Index Options Combinations

The proposed framework for analyzing arbitrage portfolios is applied to the pricing and trading of one-month SPX options. The analysis is inspired by Constantinides, Jackwerth and Perrakis (2009) and related studies. The focus traditionally is on all risk-averse index investors, which amounts to using the utility function class $\mathcal{U} = \mathcal{U}_2$, and a singleton benchmark set $\mathrm{K} = \boldsymbol{\kappa}$ which consist only of an index portfolio. New light can be shed on this topic by reducing $\mathcal{U}$ and enlarging K.

4.1 Model specification

4.1.1 Investment possibilities

The N base assets are based on $N' = \frac{1}{2}N$ distinct securities: a one-month T-bill, a stock portfolio of index constituents, and $(N' - 2)$ one-month index options of different types (calls and puts) and with different strikes. All positions are scaled by the number of index units ($100 times the index) held by the investor. Long positions and short positions in the securities are treated as separate base assets, for analytical convenience. Long positions are entered at the best available ask prices $p_i = p_i^{ask}$, $i = 1, \cdots, N'$; short positions at best bid prices $p_i = p_{i-N'}^{bid}$, $i = N' + 1, \cdots, N$. The corresponding sign restrictions are: $\lambda_i \geq 0$, $i = 1, \cdots, N'$, for long positions, and $\lambda_i \leq 0$, $i = N' + 1, \cdots, N$, for short positions. The current value of the index is denoted by y_0, and the value on the expiration date by $y_1 = (1 + r_{SPX})y_0$, where r_{SPX} is the net index return in decimal form.

Constantinides, Czerwonko and Perrakis (2020) introduce the restriction that the total open position in options is smaller than or equal to the number of index units: $\sum_{i=3}^{N'} \lambda_i - \sum_{i=N'+3}^{N} \lambda_i \leq 1$. This joint position limit is expected to enhance the robustness for estimation error of the probability distribution, as in Jagannathan and Ma (2003) and DeMiguel, Garlappi, Nogales and Uppal (2009).

Three alternative specifications of the benchmark set K are used. The singleton specification K_1 includes the index portfolio only, as in the earlier studies. Two additional benchmark portfolios are introduced to represent investors with initial written option positions. K_2 includes the index portfolio and a second benchmark which combines the index portfolio with writing one at-the-money call option, similar to the Chicago Board Options Exchange (CBOE) S&P500 BuyWrite Index (BXM). K_3 includes the index portfolio, BXM and a third benchmark which combines the bill with writing one at-the-money put, similar to the CBOE S&P500 PutWrite Index (PUT). Introducing host portfolios with written option positions amounts to taking into consideration pricing kernels which are increasing functions of market volatility as in Christoffersen, Heston and Jacobs (2013).

BXM and PUT are closely related but differ in a subtle way. The difference between their gross payoffs is limited to the difference between the dividend payment to BXM and the interest payment to PUT, by definition. The difference in the net payoffs can be larger than that, because the bid prices of the call and put may diverge due to bid-ask spreads and/or pricing errors. Furthermore, the call premium of BXM is reinvested in the index, while the put premium of PUT is reinvested in the bill.

An important role in the empirical results is played by 'call front spreads'. These option combinations are constructed by buying calls with a lower strike price (k_1) and writing calls with a higher strike price ($k_2 > k_1$); the long position consists of fewer options than the short position ($\lambda_1 < |\lambda_2|$), but it has the same value, so that the combined position is self-financing. The combined payoff is positive if $k_1 < y_1 < y^* := \left(\frac{\lambda_1 k_1 + \lambda_2 k_2}{\lambda_1 + \lambda_2}\right)$ and becomes negative if $y_1 > y^*$. Since this option combination has unlimited risk if the index moves significantly higher, its effect critically depends on the host portfolio to which it is added.

For illustration, assume that the value of the index is $y_0 = 3000$, the ask price of a call with strike $k_1 = 2970$ is $p_1 = 80$, and the bid price of a call with strike $k_2 = 3030$ is $p_2 = 40$. The largest feasible front spread of these two option series consists of buying $\lambda_1 = 0.33$ of the 2970 series and writing $|\lambda_2| = 0.67$ of the 3030 series. Figure 1 includes the payoff diagrams which are obtained if this call front spread ('CFS') is combined with a host portfolio which is equal to either SPX or BXM. In the latter case, the combined payoff decreases for $y_1 > 3030$.

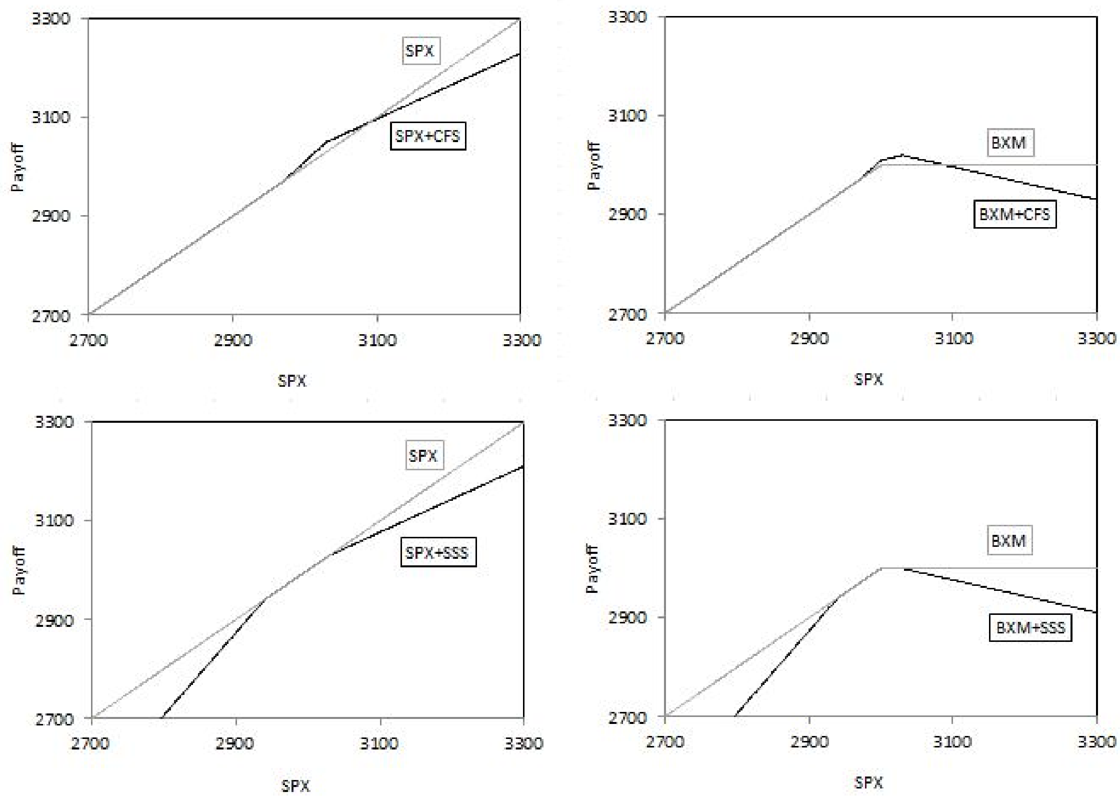
[Insert Figure 1 about here.]

Another relevant option combination is the 'short strip strangle', which is constructed by writing λ_1 OTM puts and writing $\lambda_2 < \lambda_1$ out-of-the money calls. The payoff is negative if one of the two options expires in the money ($y_1 < k_1$ or $y_1 > k_2$). The two option premiums can be used to buy bills, stocks or other options. Figure 1 includes an example of a short strip strangle ('SSS') based on $|\lambda_1| = 0.67$ written puts with strike $k_1 = 2940$ and $|\lambda_2| = 0.33$ written calls with strike $k_2 = 3030$. Again, the effect of the option combination depends on the relevant host portfolio.

4.1.2 Conditional distribution estimator

A unique feature of this application is the known and perfect single-factor structure ($M = 1$): the payoffs to all options and options combinations at the expiration date are completely determined by the future index

Figure 1: Example payoff diagrams. The top diagrams show the payoff diagram for two combinations of a hypothetical call front spread (CFS) and a host portfolio. The CFS consists of a short leg of 0.33 calls with a strike of 2970 and a short leg of 0.67 calls with a strike of 3030. The left-hand-side panel shows the diagram for the combination of CFS and SPX; the right-hand-side panel shows a payoff diagram which is obtained if BXM is the host portfolio. The bottom diagrams similarly show payoff diagrams for combinations of host portfolios and a hypothetical short strip strangle (SSS) based on 0.67 written puts with strike 2940 and 0.33 written calls with strike 3030. The option premiums are not included in these diagrams.



value $f_1 = y_1$. Using k for the strike price, the options payoff is $x = \max(y_1 - k, 0)$ for call options, and $x = \max(k - y_1, 0)$ for put options. This simple factor structure avoids the curse of dimensions for the estimation of the joint payoff distribution, despite the large number of base assets (options).

The univariate probability distribution of index returns r_{SPX} is estimated using a time series of monthly observations. The distribution is estimated every month using a moving window of 240 non-overlapping monthly observations prior to the portfolio formation date. To account for the effect of the prevailing market conditions, a conditional estimator is used.

Following Constantinides, Jackwerth and Perrakis (2009), $L = 3$ conditioning variables are used: the annualized T-bill yield (Z_1), dividend yield (Z_2) and at-the-money implied volatility (IV; Z_3). The IV is used here instead of the CBOE S&P500 Volatility Index, to synchronize the market volatility measure with the option prices. All conditioning information is publicly available on the portfolio formation date, to avoid possible forward-looking bias.

Full non-parametric estimation of the conditional distribution seems unrealistic given the limited number of monthly observations. Instead, a semi-parametric approach is used based on a Cornish-Fisher density expansion of a conditional normal distribution. A parsimonious specification is used with $d = 4$ latent parameters $\boldsymbol{\theta} \in \mathbb{R}^4$ which are estimated every month.

The conditional mean is given by $\mu(\boldsymbol{Z}_t; \boldsymbol{\theta}_t) = \frac{28}{365}\left(Z_{1,t} - Z_{2,t} + \theta_{1,t}\right)$, where $\theta_{1,t} \in [0, 1]$ is the annualized (cum-dividend) market risk premium in decimal form; the conditional standard deviation is $\sigma(\boldsymbol{Z}_t; \boldsymbol{\theta}_t) = \sqrt{\frac{28}{365}}\left(1 - \theta_{2,t}\right) Z_{3,t}$,where $\theta_{2,t} \in [0, 1]$ is the volatility risk premium expressed as a fraction of IV. The Cornish-Fisher shape parameters determine the estimated skewness level, $\theta_{3,t} \in [-3, 3]$, and excess kurtosis level, $\theta_{4,t} \in [-3, 6]$.

The parameters $\boldsymbol{\theta}_t$ are estimated every month using Maximum Likelihood in the estimation window. For every formation date $t = 1, ..., 354$, the estimation problem is $\hat{\boldsymbol{\theta}}_t := \arg\max_{\boldsymbol{\theta} \in \Theta} \sum_{s=t-240}^{t-1} \ln\left(f\left(r_{SPX;s} | \boldsymbol{Z}_{s-1}; \boldsymbol{\theta}\right)\right)$, where $\Theta = [0, 1] \times [0, 1] \times [-3, 3] \times [-3, 6]$. The estimates are interior points of the parameter set Θ in every month. The average value (across all 354 months) for $\hat{\theta}_{1,t}$ is around 0.08; the average of $\hat{\theta}_{2,t}$ hovers around 0.2; skewness and excess kurtosis estimates $\widehat{\theta}_{3,t}$ and $\widehat{\theta}_{4,t}$ around -0.5 and 1, respectively. These numbers are consistent with the stylized facts about monthly index returns and index option prices.

For numerical reasons, the conditional normal distribution is discretized using 100 grid points with equal spacing which range from $\mu(\boldsymbol{Z}_t; \widehat{\boldsymbol{\theta}}_t) - 5\sigma(\boldsymbol{Z}_t; \widehat{\boldsymbol{\theta}}_t)$ to $\mu(\boldsymbol{Z}_t; \widehat{\boldsymbol{\theta}}_t) + 5\sigma(\boldsymbol{Z}_t; \widehat{\boldsymbol{\theta}}_t)$. The Cornish-Fisher expansion is applied to the discretized distribution and results in 100 grid points with unequal spacing.

4.1.3 Risk Preferences

The default specification for the risk preferences is the set of all risk-averse utility functions ($\mathcal{U} = \mathcal{U}_2$), as in the earlier studies. This specification is quite general, as it restricts neither the higher-order risk preferences nor the degree of risk aversion. The flip size of this general specification is that it may obscure arbitrage portfolios which are appealing for all standard utility functions but not for some pathological risk preferences such as those discussed in Leshno and Levy (2002). To enlarge the set of SAOs, utility subsets $\mathcal{U} \subset \mathcal{U}_2$ are also considered.

The higher-order risk aversion properties of prudence ($u'''(x) \geq 0$) and Decreasing Absolute Risk Aversion ($(\ln(u'))'' \geq 0$) are generally accepted and known to improve the discriminatory power in various portfolio choice and asset pricing applications. These properties are however relatively ineffective for the analysis of index options. Ritchken and Kuo (1989) demonstrate that higher-order risk aversion does do not affect the theoretical upper bound for the price of index call options. The theoretical *lower* bound *is* affected by higher-order risk aversion but it is almost never violated, because index options tend to be relatively expensive.

A more promising approach is to restrict the degree of risk aversion, similar to the restrictions on the pricing kernel in Bernardo and Ledoit (2000) and Cochrane and Saa-Requejo (2000), and the Almost Stochastic Dominance restrictions on the slope or curvature of the utility functions in Leshno and Levy (2002) and Czerwonko, Davidson and Perrakis (2021).

The chosen specification is based on Infinite-degree Stochastic Dominance; see Section 2.5. Its extreme utility set $\mathcal{U}_\infty^{(0)}$ is approximated using the discrete set $\widehat{\mathcal{U}}_\infty^{(0)} := \{v_{\infty;\phi_1}(x), \cdots, v_{\infty;\phi_S}(x)\}$. The selected levels of Absolute Risk Aversion (ARA), ϕ_s, $s = 1, \cdots, S$, are based on the relation between ARA and Relative Risk Aversion (RRA), or $ARA(x) = RRA(x)/x$, $x \in [a, b]$, and a given lower bound $\underline{RRA}$ and upper bound $\overline{RRA}$. Specifically, the ARA is bounded by $\phi_1 = \underline{RRA}/b$ and $\phi_S = \overline{RRA}/a$, and the remaining $(S-2)$ values are obtained using a logarithmic spacing on the interval $[\phi_1, \phi_S]$. The full utility set is $\widehat{\mathcal{U}}_\infty := \mathrm{conv}\left(\widehat{\mathcal{U}}_\infty^{(0)}\right)$.

The chosen parameter values are $\underline{RRA} = 0.1$, $\overline{RRA} = 10$, and $S = 10$. The upper bound $\overline{RRA} = 10$ is chosen to exceed risk aversion levels which seem justified based on theoretical and empirical arguments; see, for example, Meyer and Meyer (2005).

This formulation is analytically convenient, because $\widehat{\mathcal{U}}_\infty$ is closed, which ensures the Joint Enhancement property (Definition 2.4.2) if strict SAOs exist, and, moreover, the enhancement condition $D(u, \boldsymbol{\kappa}, \boldsymbol{\delta}, \mathcal{F}) \geq 0$ needs to be verified only at the discrete set of elementary functions $\widehat{\mathcal{U}}_\infty^{(0)}$, and the resulting restrictions $D(v_{\infty;\phi_s}(x), \boldsymbol{\kappa}, \boldsymbol{\delta}, \mathcal{F}) \geq 0$, $s = 1, \cdots, S$, are smooth and convex.

4.2 Empirical analysis

4.2.1 Data

A master data set of option prices was constructed using intraday quotes from the CBOE tape. The data set consists of bid and ask prices of one-month SPX options which expire in the months of February 1990 to July 2019, a total of 354 months. The master file was kindly provided by Michal Czerwonko, to allow for comparison with the aforementioned earlier studies.

In the present study, options are selected from the master file based on several data filters: (i) the strike price is less than 10% out of the money and less than 10% in the money; (ii) the bid price is at least 15 cents in 1990 prices; (iii) the ask price exceeds the bid price; (iv) the bid and ask prices are consistent with the Put Call Parity, to exclude pure arbitrage possibilities. The number of qualified option series ($\frac{1}{2}N-2$) ranges from about 20 at the start of the sample in 1990 to about 200 towards the end of the sample in 2019.

To avoid problems with data synchronization, the concurrent value of the SPX (y_0) is estimated every month from the option prices by inverting the Put Call Parity. Similarly, a concurrent value for the at-the-money IV is computed every month by inverting the Black and Scholes valuation formula for the qualified option series which is closest to being at the money. The BMX and PUT benchmarks are constructed using the bid prices of the qualified option series which are closest to the money.

4.2.2 Trading strategy

Active portfolios are formed every month, from January 1990 to June 2019, by adding arbitrage portfolios to host portfolios, based on the conditional CDF estimator of index returns. The formation date is 28 days prior to the options expiration date; the options are held until expiration.

On every formation date, the empirical counterpart of optimization problem (4) is solved, for both $\mathcal{U}=\mathcal{U}_2$ and $\mathcal{U}=\widehat{\mathcal{U}}_\infty$. The objective function is the expected portfolio payoff, which amount to using $v(x)=x$.

Out-of-sample investment performance is evaluated using the annualized realized total holding period return (HPR) during the 28-day holding period, and is evaluated from February 1990 through July 2019.

4.2.3 Results for all levels of risk aversion

Table 1 summarizes the results which are obtained for $\mathcal{U} = \mathcal{U}_2$. The first three columns summarize the investment performance of the three benchmark portfolios (SPX, BXM and PUT). The remaining nine columns summarize the results which are obtained using arbitrage portfolios. For each of the three benchmark sets (K_1, K_2 and K_3), three combined portfolios are constructed by adding the optimal arbitrage portfolio to one of the three host portfolios.

Shown are the central moments of HPR together with the Certainty Equivalent Return (CER) for a power utility functions for various values of the RRA coefficient. Also shown are the median values (across all 354 months) of the optimal option positions $|\lambda_i|$ and the moneyness measure k_i/y_0.

The analysis for index investors ($K = K_1$) shows moderate out-of-sample outperformance, consistent with the earlier studies. Average return is enhanced by some 72 basis points per annum, and investment risk is reduced. The improvement of average return is statistically only marginally significant, but the combined effect of return enhancement and risk reduction is an improvement of the CER with high levels of statistical confidence, for medium to high levels of risk aversion.

The outperformance is achieved predominantly using call front spreads. The typical spread is constructed by buying near-the-money calls and writing ITM calls. Open put positions are rare, as buying puts is expensive and writing puts increases downside risk for index investors.

For a BXM investor with an initial covered call position, call front spreads introduce additional downside risk, as shown in the bottom-left panel of Figure 1 above. As a result, the optimal arbitrage portfolios for $K = K_2$ include fewer open call positions than for $K = K_1$. In addition, the typical option portfolio also includes a put bear spread position to reduce downside risk. The remaining outperformance is not statistically significant (except for the CER for RRA=8).

The results for $K = K_3$ are very similar to those for $K = K_2$, both in terms of the composition of the the optimal option combination and the investment performance of the combined portfolio. Apparently, the difference between the net payoff to BXM and PUT is too small to have a material effect here.

[Insert Table 1 about here.]

Table 1: Optimal combined portfolios for all levels of risk aversion. Out-of-sample investment performance is summarized for combined portfolios constructed from various host portfolios (SPX, BXM, PUT) and arbitrage portfolios. The arbitrage portfolios are formed every month, from January 1990 to June 2019, based on a conditional CDF estimator of index returns and a given benchmark set $\mathrm{K} = \mathrm{K}_1, \mathrm{K}_2, \mathrm{K}_3$. The formation date is 28 days prior to the options expiration date; the options are held until expiration. On every formation date, the empirical counterpart of optimization problem (4) is solved for utility function class $\mathcal{U} = \mathcal{U}_2$ and objective function equal to expected portfolio payoff. The realized HPR is computed for every 28-day holding period from February 1990 through July 2019. Average returns and CERs which are significantly higher than those of the host portfolio using a paired t-test with significance level of 10%, 5% and 1% are indicated using one, two and three asterisks, respectively. The median of the optimal option positions $|\lambda_i|$ are computed across all 354 holding periods. The median of the moneyness measure k_i/y_0 is computed if option positions are opened in at least half of the months.

Benchmark set					K_1			K_2			K_3		
Host portfolio		SPX	BXM	PUT	SPX	BXM	PUT	SPX	BXM	PUT	SPX	BXM	PUT
Moments	Mean	10.80	8.85	8.73	11.51*	9.56*	9.43*	11.20	9.26	9.13	11.22	9.28	9.15
(ppa)	StDev	15.01	9.96	9.64	14.30	9.90	9.62	14.17	9.68	9.41	14.13	9.63	9.35
	Skew	-1.24	-3.47	-3.51	-1.36	-3.16	-3.15	-1.49	-3.56	-3.55	-1.57	-3.77	-3.77
	Kurt	5.41	19.24	19.30	5.81	17.48	17.25	6.19	20.18	19.94	6.59	22.10	21.87
CER	RRA=1	9.58	8.30	8.21	10.40*	9.02*	8.92*	10.11	8.73	8.64	10.14	8.76	8.66
(ppa)	RRA=2	8.42	7.76	7.70	9.35**	8.49*	8.43*	9.08*	8.22	8.16	9.10*	8.25	8.18
	RRA=4	5.80	6.47	6.51	6.96**	7.26*	7.27*	6.70*	7.02	7.03	6.72*	7.04	7.05
	RRA=8	-0.61	3.09	3.41	1.17**	4.09*	4.34*	0.88**	3.84*	4.10*	0.80**	3.80*	4.07*
Median	Buy Call	0	0	0	0.22	0.22	0.22	0.01	0.01	0.01	0.00	0.00	0.00
option	Buy Put	0	0	0	0	0	0	0.03	0.03	0.03	0.04	0.04	0.04
positions	Write Call	0	1	0	0.67	1.67	0.67	0.52	1.52	0.52	0.34	1.34	0.34
(#)	Write Put	0	0	1	0	0	1	0.03	0.03	1.03	0.06	0.06	1.06
Median	Buy Call				1.01	1.01	1.01	0.95	0.95	0.95	1.01	1.01	1.01
option	Buy Put							1.04	1.04	1.04	1.01	1.01	1.01
moneyness	Write Call		0		1.05	1.02	1.05	1.03	1.01	1.01	1.03	1.03	1.03
	Write Put			0				0.98	0.98	0.99	0.96	0.96	0.98

4.2.4 Results for restricted levels of risk aversion

Table 2 summarizes the results for $\mathcal{U} = \widehat{\mathcal{U}}_{\infty}$. Restricting the level of risk aversion leads to major change of the optimal option combinations compared with the results in Table 1. For $\mathrm{K} = \mathrm{K}_1$, the bulk of the option combinations are short strip strangles which are constructed by writing OTM puts and ITM calls.

The chosen option combinations differ in important ways from those in Czerwonko, Davidson and Perrakis (2021): their put front spreads combine written OTM puts with long ATM puts (instead of written ITM calls); their at-the-money (ATM) straddles and strangles do not use OTM puts and ITM calls. These differences arise because they fix the moneyness levels in advance to study the systematic overpricing of given combinations, while the strikes are an endogenous part of our optimization problem, as in Constantinides, Czerwonko, and Perrakis (2020).

The option combinations increase average return to 148 basis point per annum in excess of the stock index, roughly doubling the outperformance seen in Table 1. The improvement is statistically significant at every conventional level of statistical significance. The writing of puts naturally increases downside risk (witness the negative skewness and excess kurtosis), but the additional risk does not worsen the CER for the relevant levels of RRA and hence the strangles appear to be genuine SAOs.

The effect of the specification of the benchmark set K largely disappears. In contrast to the writing of calls, the effect of the writing of puts on the combined portfolio is similar for all three host portfolios. Investors with plausible levels of risk aversion agree about the appeal of writing puts, even if they are already endowed with an open option position (BXM and PUT).

In conclusion, the economic significance of index option pricing errors seems materially higher than in the earlier studies. A simple buy-and-hold option strategy based on past public information suffices to stochastically enhance the stock index. If extreme levels of risk aversion are excluded, the outperformance becomes economically and statistically significant and robust to the inclusion of pre-existing option positions in the benchmark.

[Insert Table 2 about here.]

5 Conclusions

This study generalizes the concept of pure arbitrage using partial information about utility functions and initial portfolio positions; derives asset pricing restrictions for the base assets which are necessary and

Table 2: Optimal combined portfolios for restricted levels of risk aversion. Out-of-sample investment performance is summarized for combined portfolios constructed from various host portfolios (SPX, BXM, PUT) and arbitrage portfolios. The arbitrage portfolios are formed every month, from January 1990 to June 2019, based on a conditional CDF estimator of index returns and a given benchmark set $\mathrm{K} = \mathrm{K}_1, \mathrm{K}_2, \mathrm{K}_3$. The formation date is 28 days prior to the options expiration date; the options are held until expiration. On every formation date, the empirical counterpart of optimization problem (4) is solved for utility function class $\mathcal{U} = \widehat{\mathcal{U}}_\infty$ and objective function equal to expected portfolio payoff. The realized HPR is computed for every 28-day holding period from February 1990 through July 2019. Average returns and CERs which are significantly higher than those of the host portfolio using a paired t-test with significance level of 10%, 5% and 1% are indicated using one, two and three asterisks, respectively. The median of the optimal option positions $|\lambda_i|$ are computed across all 354 holding periods. The median of the moneyness measure k_i/y_0 is computed if option positions are opened in at least half of the months.

Benchmark set					K_1			K_2			K_3		
Host portfolio		SPX	BXM	PUT	SPX	BXM	PUT	SPX	BXM	PUT	SPX	BXM	PUT
Moments	Mean	10.80	8.85	8.73	12.27***	10.33***	10.20***	12.25***	10.31***	10.18***	12.13***	10.19***	10.06***
(ppa)	StDev	15.01	9.96	9.64	14.97	10.72	10.44	15.00	11.06	10.80	15.08	11.00	10.73
	Skew	-1.24	-3.47	-3.51	-1.93	-4.13	-4.17	-2.32	-4.63	-4.65	-2.30	-4.71	-4.74
	Kurt	5.41	19.24	19.30	8.88	25.52	25.82	11.97	30.71	30.81	11.58	31.31	31.48
CER	RRA=1	9.58	8.30	8.21	11.04***	9.67***	9.58***	10.99***	9.60***	9.50***	10.86***	9.49***	9.39***
(ppa)	RRA=2	8.42	7.76	7.70	9.84***	9.02***	8.96***	9.75***	8.88***	8.82***	9.61***	8.77***	8.72***
	RRA=4	5.80	6.47	6.51	7.01***	7.42**	7.45**	6.73**	7.06*	7.11*	6.57**	6.97*	7.02*
	RRA=8	-0.61	3.09	3.41	-0.50	2.84	3.18	-1.90	1.52	1.96	-2.11	1.45	1.90
Median	Buy Call	0	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
option	Buy Put	0	0	0	0.04	0.04	0.04	0.05	0.05	0.05	0.06	0.06	0.06
positions	Write Call	0	1	0	0.07	1.07	0.07	0.02	1.02	0.02	0.01	1.01	0.01
(#)	Write Put	0	0	1	0.81	0.81	1.81	0.90	0.90	1.90	0.87	0.87	1.87
Median	Buy Call				1.02	1.02	1.02	1.01	1.01	1.01	1.02	1.02	1.02
option	Buy Put				1.06	1.06	1.06	1.07	1.07	1.07	1.07	1.07	1.07
moneyness	Write Call		0		0.98	0.99	0.98	0.96	0.98	0.96	0.97	0.98	0.97
	Write Put			0	0.93	0.93	0.97	0.94	0.94	0.97	0.93	0.93	0.97

sufficient conditions for excluding generalized arbitrage opportunities; develops a statistical theory for the empirical solution set for arbitrage portfolio choice which applies for a range of parametric and non-parametric estimators for the joint payoff distribution.

The study also aims to contribute to research on the valuation and trading of index options by exploring the economic significance of index option pricing errors in terms of the out-of-sample profitability of optimized option combinations. The empirical results show that the arbitrage opportunities for one-month index options are materially larger than documented in previous studies and robust to the choice of the host portfolio when plausible constraints are imposed on the level of risk aversion. These findings are consistent with, and deepen, the option pricing kernel puzzle (Ait-Sahalia and Lo (2000) and Jackwerth (2000)).

The use of a multi-period model is unlikely to change the conclusions because the existence of arbitrage opportunities in a single-period model generally is a sufficient condition for the existence of arbitrage opportunities in a multi-period model. Exposure to volatility risk also seems an unlikely explanation, given that the results are robust to the use of host portfolios (BXM, PUT) which are short volatility. Since the analysis does not assume the existence of a unique pricing kernel or risk-neutral distribution, the results are furthermore immune for the critique by Linn, Shive and Shumway (2018) on the traditional approach of estimating the pricing kernel using the ratio of an estimated physical density to a separately estimated risk-neutral density.

A maintained assumption is that the investor knows the conditional payoff distribution or can estimate it with high accuracy using time-series data and side information. This assumption seems plausible for the univariate distribution of SPX returns which drives the payoffs to SPX options, as it arguably is the most studied distribution in Finance. For the analysis of new asset classes or individual securities without a simple factor structure, estimation risk and ambiguity aversion will naturally play a bigger role. A challenge for further research is to extend the present framework to ambiguity-averse preferences and the estimation of 'distributions of distributions' to better deal with such investment problems.

References

[1] Ait-Sahalia, Y. and A. Lo, 2000, Nonparametric Risk Management and Implied Risk Aversion, Journal of Econometrics 94, 9-51.

[2] Aliprantis C.D., and K.C. Border, 2006, Infinite dimensional analysis: a hitchhiker's guide, Springer.

[3] Arvanitis, S., M. Hallam, Th. Post and N. Topaloglou, 2019, Stochastic Spanning, Journal of Business & Economic Statistics, 37:4, 573-585.

[4] Arvanitis, S., Th. Post, and N. Topaloglou, 2021, Stochastic Bounds for Reference Sets in Portfolio Analysis, forthcoming in Management Science.

[5] Bawa, V.S., J.N. Bodurtha Jr., M.R. Rao and H.L. Suri, 1985, On Determination of Stochastic Dominance Optimal Sets, Journal of Finance 40, 417–431.

[6] Bernardo, A.E., and O. Ledoit, 2000, Gain, Loss, and Asset Pricing, Journal of Political Economy 108, 144–172.

[7] Brandt, M.W. and P. Santa-Clara, 2006, Dynamic Portfolio Selection by Augmenting the Asset Space, Journal of Finance 61, 2187-221.

[8] Bondarenko, O., 2003, Statistical arbitrage and securities prices, Review of Financial Studies 16(3), 875–919.

[9] Christoffersen, P., S. Heston, and K. Jacobs, 2013, Capturing Option Anomalies with a Variance-Dependent Pricing Kernel. Review of Financial Studies 26(8), 1963-2006.

[10] Cochrane, J.H. and J. Saa-Requejo, 2000, Beyond Arbitrage: Good-Deal Asset Price Bounds in Incomplete Markets, Journal of Political Economy 108, 79-119.

[11] Constantinides, G.M., M. Czerwonko, J.C. Jackwerth, and S. Perrakis, 2011, "Are Options on Index Futures Profitable for Risk Averse Investors? Empirical Evidence," Journal of Finance 66, 1407–1437.

[12] Constantinides, G.M., M. Czerwonko and S. Perrakis, 2020, Mispriced Index Option Portfolios, forthcoming in Financial Management.

[13] Constantinides, G.M., J.C. Jackwerth, and S. Perrakis, 2009, Mispricing of S&P 500 Index Options, Review of Financial Studies 22, 1247–1277.

[14] Czerwonko, M., R. Davidson, and S. Perrakis, 2021, Tail Risk, Almost Stochastic Dominance and Index Option Anomalies, working paper.

[15] DeMiguel, V., L. Garlappi, F.J. Nogales and R. Uppal, 2009, A Generalized Approach to Portfolio Optimization: Improving Performance by Constraining Portfolio Norms, Management Science 55, 798-812.

[16] Dentcheva, D., 2001, Continuity of multifunctions characterized by Steiner selections, Nonlinear Analysis: Theory, Methods & Applications, 47(3), 1985-1996.

[17] Dentcheva, D. and A. Ruszczynski, 2003, Optimization with Stochastic Dominance Constraints, SIAM Journal on Optimization 14, 548–566.

[18] Dentcheva, D. and A. Ruszczynski, 2006, Portfolio optimization with stochastic dominance constraints, Journal of Banking and Finance 30, 433-451.

[19] Dybvig, P.A. and S.A. Ross, 2008, Arbitrage, in Steven F. Durlauf and Lawrence E. Blume, eds., The New Palgrave: A Dictionary of Economics, Volume 1, Second Edition, Palgrave MacMillan, New York.

[20] El Machkouri, M., X. Fan, and L. Reding, 2020, On the Nadaraya–Watson kernel regression estimator for irregularly spaced spatial data, Journal of Statistical Planning and Inference, 205, 92-114.

[21] Fábián, C.I., G. Mitra, and D. Roman, 2011, Processing second-order stochastic dominance models using cutting-plane representations, Mathematical Programming 130, 33-57.

[22] Giné, E. and R. Nickl, 2016, Mathematical foundations of infinite-dimensional statistical models, Vol. 40, Cambridge University Press.

[23] Jackwerth, J.C., 2000, Recovering risk aversion from option prices and realized returns, Review of Financial Studies 13, 433–451.

[24] Jagannathan, R., T. Ma, 2003, Risk reduction in large portfolios: Why imposing the wrong constraints helps, Journal of Finance 58, 1651-1684.

[25] Jouini, E. and H. Kallal, 2001, Efficient trading strategies in the presence of market frictions. Review of Financial Studies 14, 343-369.

[26] Kim, S., R.A. Korajczyk, and A. Neuhierl, 2020, Arbitrage Portfolios, forthcoming in Review of Financial Studies, hhaa102, https://doi.org/10.1093/rfs/hhaa102.

[27] Kitamura, Y., G. Tripathi, and H. Ahn, 2004, Empirical likelihood-based inference in conditional moment restriction models, Econometrica, 72(6), 1667-1714.

[28] Knight, K., 1999, Epi-convergence in distribution and stochastic equi-semicontinuity, Unpublished manuscript, 37, 28-72.

[29] Korkie, B. and H.J. Turtle, 2002, A Mean-Variance Analysis of Self-Financing Portfolios, Management Science 48 (3), pp. 427-443.

[30] Linn, M, S. Shive, and T. Shumway, 2018, Pricing Kernel Monotonicity and Conditional Information, Review of Financial Studies, 31, 493–531.

[31] Leshno, M. and H. Levy, 2002, Preferred by "All" and Preferred by "Most" Decision Makers: Almost Stochastic Dominance, Management Science 48(8): 1074-1085.

[32] Klein Haneveld, W.K. and M.H. van der Vlerk, 2006, Integrated chance constraints: reduced forms and an algorithm, Computational Management Science 3, 245-269.

[33] Kunzi-Bay, A. and J. Mayer, 2006, Computational aspects of minimizing conditional value-at-risk, Computational Management Science 3, 3-27.

[34] Meyer, D.J., and J. Meyer, 2005, Relative Risk Aversion: What Do We Know?, Journal of Risk and Uncertainty 31, 243–262.

[35] Mikosch, T., and D. Straumann, 2006, Stable limits of martingale transforms with application to the estimation of GARCH parameters, Annals of Statistics, 34(1), 493-522.

[36] Molchanov, I., 2006, Theory of random sets, Springer Science & Business Media.

[37] Post, Th., 2003, Empirical Tests for Stochastic Dominance Efficiency, Journal of Finance 58, 1905-1932.

[38] Post, Th., S. Karabatı, and S. Arvanitis, 2018, Portfolio optimization based on stochastic dominance and empirical likelihood, Journal of Econometrics 206 (1), 167-186.

[39] Post, Th. and M. Kopa, 2017, Portfolio Choice Based on Third-Degree Stochastic Dominance, Management Science 63 (10), 3381–3392.

[40] Post, Th. and I. Rodriguez Longarela, 2021, Risk Arbitrage Opportunities for Stock Index Options, Operations Research 69 (1), 100-113.

[41] Rio, E., 2017, Asymptotic theory of weakly dependent random processes (Vol. 80), Berlin: Springer.

[42] Ritchken, P. and S. Kuo, 1989, On Stochastic Dominance and Decreasing Absolute Risk Averse Option Pricing Bounds, Management Science 35, 51-59.

[43] Rockafellar R.T. and S. Uryasev, 2000, Optimization of Conditional Value-at-Risk, Journal of Risk 2, 21-41.

[44] 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.

[45] Salinetti, G., and R.J. Wets, 1977, On the relations between two types of convergence for convex functions, Journal of Mathematical Analysis and Applications, 60(1), 211-226.

[46] Thistle, P., 1993, Negative Moments, Risk Aversion, and Stochastic Dominance, Journal of Financial and Quantitative Analysis 28(2), 301-311.

[47] Van Der Vaart, A. W., and J.A. Wellner, 1996, Weak convergence and empirical processes, Springer, New York, NY.

6 Appendices

6.1 Appendix A: Proofs for Financial Theory

Proof of Theorem 2.3.1. By Definition 2.2.1, $\boldsymbol{\delta} \notin \Delta_{(\cdot)}^{\succ}$ iff $\left(\mathbb{E}_{\mathcal{F}}\left[u(\boldsymbol{x}^{\mathrm{T}}(\boldsymbol{\kappa}+\boldsymbol{\delta}))-u(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\kappa})\right] \leq 0\right)$, for some $(u, \boldsymbol{\kappa}) \in \mathcal{U} \times \mathrm{K}$. Due to concavity of the utility function, this condition is equivalent to $\boldsymbol{\lambda} = \boldsymbol{\kappa}$ being the optimum for $\sup_{\boldsymbol{\lambda} \in \Theta} \mathbb{E}_{\mathcal{F}}\left[u(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\lambda})\right]$, for feasible set $\Theta := \{\boldsymbol{\lambda} \in \Lambda_0 : \boldsymbol{\lambda} = \boldsymbol{\kappa} + \theta\boldsymbol{\delta}; \theta \in [0,1]\}$. The necessary and sufficient Karush-Kuhn-Tucker condition therefore implies $\boldsymbol{\delta} \notin \Delta^{\succ}_{(\cdot)}$ iff $\mathbb{E}_{\mathcal{F}}\left[u'(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\kappa})\boldsymbol{x}\right]^{\mathrm{T}}\boldsymbol{\delta} \leq 0$ for some $(u,\boldsymbol{\kappa}) \in \mathcal{U} \times \mathrm{K}$. Since the scaling of marginal utility by the scalar $d(u,\boldsymbol{\kappa})$ does not affect the optimality condition, this condition is equivalent to $\mathbb{E}_{\mathcal{F}}\left[m(\boldsymbol{x})\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\delta}\right] \leq 0$ for some $m \in \mathcal{M}_{(\cdot)}$. Hence, $\Delta^{\succ}_{(\cdot)} = \emptyset$ if and only if

$$\sup_{\Delta} \inf_{\mathcal{M}_{(\cdot)}} \left(\mathbb{E}_{\mathcal{F}}\left[m(\boldsymbol{x})\boldsymbol{x}^{\mathrm{T}}\right]\boldsymbol{\delta}\right) \leq 0 \Leftrightarrow \tag{11}$$

$$\sup_{\Delta} \inf_{\mathrm{conv}\left(\mathcal{M}_{(\cdot)}\right)} \left(\mathbb{E}_{\mathcal{F}}\left[m(\boldsymbol{x})\boldsymbol{x}^{\mathrm{T}}\right]\boldsymbol{\delta}\right) \leq 0 \Leftrightarrow \tag{12}$$

$$\inf_{\mathrm{conv}\left(\mathcal{M}_{(\cdot)}\right)} \sup_{\Delta} \left(\mathbb{E}_{\mathcal{F}}\left[m(\boldsymbol{x})\boldsymbol{x}^{\mathrm{T}}\right]\boldsymbol{\delta}\right) \leq 0 \Leftrightarrow \tag{13}$$

$$\inf_{\mathrm{conv}\left(\mathcal{M}_{(\cdot)}\right)} \left\{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{\sigma} : \mathbb{E}_{\mathcal{F}}\left[m(\boldsymbol{x})\boldsymbol{x}\right] = \mathbf{A}^{\mathrm{T}}\boldsymbol{\sigma}; \boldsymbol{\sigma} \geq \mathbf{0}_R\right\} \leq 0 \Leftrightarrow \tag{14}$$

$$\left(\mathbb{E}_{\mathcal{F}}\left[m(\boldsymbol{x})\boldsymbol{x}\right] \in \Delta^*\right) \text{ for some } m \in \mathrm{conv}\left(\mathcal{M}_{(\cdot)}\right). \tag{15}$$

The convexification of the feasible set of inner minimization problem, $\mathcal{M}_{(\cdot)}$, in (12) is allowed because the objective function is a bilinear map $\Delta \times \mathcal{M}_{(\cdot)} \to \mathbb{R}$. The reversal of the order of the optimization operators in (13) is allowed by the Kneser-Fan Minimax Theorem, because the feasible sets for both optimization operators are convex. The equivalent formulation (14) is based on the dual formulation of the embedded, linear maximization problem over Δ.□

Proof of Lemma 2.4.3. Part (i). For every $(u,\boldsymbol{\kappa},w) \in (\mathrm{cl}\,(\mathcal{U}) - \mathcal{U}^{=}) \times \mathrm{K} \times W$, we have that

$$\begin{aligned} D(u,\boldsymbol{\kappa},\boldsymbol{\gamma},\mathcal{F}) &\geq (1-c_\epsilon)\, D(u,\boldsymbol{\kappa},\boldsymbol{\delta}^*_{(\cdot)}(v,\mathcal{F}),\mathcal{F}) + c_\epsilon D(u,\boldsymbol{\kappa},\boldsymbol{\gamma}_{(\cdot)}(w,\mathcal{F}),\mathcal{F}) \\ &\geq c_\epsilon D(u,\boldsymbol{\kappa},\boldsymbol{\gamma}_{(\cdot)}(w,\mathcal{F}),\mathcal{F}) \\ &\geq c_\epsilon \int_{\mathrm{cl}(\mathcal{U})} D(u,\boldsymbol{\kappa},\boldsymbol{\delta}^*_{(\cdot)}(v,\mathcal{F}),\mathcal{F})dw\,(v) > 0. \end{aligned}$$

The first and third weak inequality follow from the concavity property of expected utility for $u \in \mathrm{cl}\,(\mathcal{U})$. The final strict inequality follows from the Joint Enhancement property (7). Part (ii). The compactness of $\mathrm{cl}\,(\mathcal{U})$ and K imply that $\max_{\overline{\mathrm{K}}} \mathbb{E}_{\mathcal{F}}\left[u(\boldsymbol{x}^{\mathrm{T}}(\boldsymbol{\kappa}+\boldsymbol{\delta}))\right]$ is continuous in $\boldsymbol{\delta}$ for all $u \in \mathrm{cl}\,(\mathcal{U})$. For any $(c,w) \in (0,1) \times \mathcal{W}$, we have that $(1-c)\,w_v + cw \in \mathcal{W}$. Since Δ is compact, and $\boldsymbol{\delta}^*_{(\cdot)}(u,\mathcal{F})$ is continuous in u due to Corollary 2.E of Salinetti and Wets (1977) and Theorem 3.8 of Dentcheva (2001), as $c \to 1$, $\boldsymbol{\gamma}_{(\cdot)}((1-c)\,w_v + cw,\mathcal{F}) \to$

$\boldsymbol{\delta}^*_{(\cdot)}(v,\mathcal{F})$, using Corollary 15.7 of Aliprantis and Border (2006). Thereby, for any $\epsilon > 0$, there exists some $c_\epsilon \in (0,1)$ for which $\max_{\underline{\mathrm{K}}} \mathbb{E}_{\mathcal{F}}\left[v\left(\boldsymbol{x}^{\mathrm{T}}\left(\boldsymbol{\kappa}+\boldsymbol{\delta}^*_{(\cdot)}(v,\mathcal{F})\right)\right)\right] - \max_{\underline{\mathrm{K}}} \mathbb{E}_{\mathcal{F}}\left[v\left(\boldsymbol{x}^{\mathrm{T}}\left(\boldsymbol{\kappa}+\boldsymbol{\gamma}\right)\right)\right] < \epsilon$, and thus (ii) applies.□

6.2 Appendix B: Proofs for Statistical Theory

Proof of Theorem 3.2.1. First notice that any element of $\Delta^{=}_{\mathcal{F}}$ belongs to $\Delta^{\gtrsim}_{(\mathcal{U},\mathrm{K},\mathcal{F}_T)}$. Section 2.4 shows that when non-trivial weak SAOs do not exist, then $\Delta^{\gtrsim}_{(\mathcal{U},\mathrm{K},\mathcal{F})} = \Delta^{=}_{\mathcal{F}}$. Hence, if $\boldsymbol{\delta} \notin \triangle^{=}_{\mathcal{F}}$, then $\exists u \in \mathcal{U}, \boldsymbol{\kappa} \in K$ for which $D(u,\boldsymbol{\kappa},\boldsymbol{\delta},\mathcal{F}) < 0$. Assumption 3.1.1 then implies that $\limsup_{T\to\infty} \mathbb{P}\left[D(u,\boldsymbol{\kappa},\boldsymbol{\delta},\mathcal{F}_T) \geq 0\right] = 0$, which establishes Theorem 3.2.1.(ii). Suppose next that non-trivial weak SAOs do not exist. Let $c_T = o(1)$ and such that $m_T c_T \to \infty$. Assumption 3.1.1, and Corollary 2.E of Salinetti and Wets (1977) imply that we need to consider only fixed $u \in \mathrm{cl}(\mathcal{U})$ in our derivations. For any $\boldsymbol{\delta} \in \boldsymbol{\delta}^*_{(\mathcal{U},\mathrm{K})}(v,\mathcal{F}) \subseteq \Delta^{(c_T)}_{(\mathcal{U},\mathrm{K})}(v,\mathcal{F})$ that does not have any trivial equivalencies, we have that $\min_K D(u,\boldsymbol{\kappa},\boldsymbol{\delta},\mathcal{F}) > 0$ for all $u \in (\mathrm{cl}(\mathcal{U}) - \mathcal{U}^{=})$, thus $\liminf_{T\to\infty} \mathbb{P}\left[\min_K D(u,\boldsymbol{\kappa},\boldsymbol{\delta},\mathcal{F}_T) \geq 0\right] = 1$ for all $u \in \mathrm{cl}(\mathcal{U})$ due to Assumption 3.1.1. Suppose then that $\boldsymbol{\delta}$ has non-trivial equivalences. Then for any $u \in (\mathrm{cl}(\mathcal{U}) - \mathcal{U}^{=})$ that does not correspond to some equivalence or to some $u \in \mathcal{U}^{=}$, the previous analysis holds. If $u \in (\mathrm{cl}(\mathcal{U}) - \mathcal{U}^{=})$ and u corresponds to some non-trivial equivalence, then consider the strong SAO $\boldsymbol{\gamma}_T := \boldsymbol{\gamma}_{(\mathcal{U},\mathrm{K})}\left((1-c_T)\,w_u + c_T w, \mathcal{F}\right)$, as in Lemma 2.4.3. Notice that

$$\begin{aligned}
&\mathbb{P}\left[m_T \min_K D(u,\boldsymbol{\kappa},\boldsymbol{\gamma}_T,\mathcal{F}_T-\mathcal{F}) \geq -m_T \min_K D(u,\boldsymbol{\kappa},\boldsymbol{\gamma}_T,\mathcal{F})\right] \\
&= \mathbb{P}\left[m_T \min_K D(u,\boldsymbol{\kappa},\boldsymbol{\gamma}_T,\mathcal{F}_T-\mathcal{F}) \geq -m_T \min_K \mathbb{E}_{\mathcal{F}}\left[u'\left(\boldsymbol{x}^{\mathrm{T}}(\boldsymbol{\kappa}+\boldsymbol{\gamma}_T)^\star\right)X^{\mathrm{T}}\right]\boldsymbol{\gamma}_T\right],
\end{aligned}$$

where the second equality follows from the MVT and dominated convergence, and $(\boldsymbol{\kappa}+\boldsymbol{\gamma}_T)^\star$ lies between $\boldsymbol{\kappa}+\boldsymbol{\gamma}_T$ and $\boldsymbol{\kappa}$. Since u does not correspond to a trivial equivalence, we have that there exists some $c > 0$ independent of T such that $\min_{i=1,\ldots,N}\left[\min_K \mathbb{E}_{\mathcal{F}}\left[u'\left(\boldsymbol{x}^{\mathrm{T}}(\boldsymbol{\kappa}+\boldsymbol{\gamma}_T)^\star\right)X^{\mathrm{T}}\right]\right] > c$ and such that $m_T \min_K \mathbb{E}_{\mathcal{F}}\left[u'\left(\boldsymbol{x}^{\mathrm{T}}(\boldsymbol{\kappa}+\boldsymbol{\gamma}_T)^\star\right)X^{\mathrm{T}}\right]\boldsymbol{\gamma}_T \geq c\sum_{i=1}^{N}\left|\boldsymbol{\gamma}_{i,T}\right| > 0$. Hence the last probability in the previous display is greater than or equal to

$$\mathbb{P}\left[m_T \min_K D(u,\boldsymbol{\kappa},\boldsymbol{\gamma}_T,\mathcal{F}_T-\mathcal{F}) \geq -m_T c\sum_{i=1}^{N}\left|\int_{\mathrm{cl}(\mathcal{U})} \boldsymbol{\delta}^*_{i,(\cdot)}(v,\mathcal{F})\,d\left[(1-c_T)\,w_u + c_T w\right]\right|\right].$$

Notice that $\left|\int_{\mathrm{cl}(\mathcal{U})} \boldsymbol{\delta}^*_{i,(\cdot)}(v,\mathcal{F})\,d\left[(1-c_T)\,w_u + c_T w\right]\right| = c_T\left|\int_{\mathrm{cl}(\mathcal{U})} \boldsymbol{\delta}^*_{i,(\cdot)}(v,\mathcal{F})\,dw\right|$. Thereby, the probability of the last display is greater than or equal to

$$\mathbb{P}\left[m_T \min_K D(u,\boldsymbol{\kappa},\boldsymbol{\gamma}_T,\mathcal{F}_T-\mathcal{F}) \geq -m_T c_T c\sum_{i=1}^{N}\left|\int_{\mathrm{cl}(\mathcal{U})} \boldsymbol{\delta}^*_{i,(\cdot)}(v,\mathcal{F})(v)\,dw\right|\right],$$

and due to Assumption 3.1.1 the CMT and the Portmanteau Theorem, the $\liminf$ of the latter is greater than or equal to $\mathbb{P}\left[\min_K \mathcal{G}\left(u, \boldsymbol{\kappa}, \boldsymbol{\delta}\right) \geq -\infty\right] = 1$. Hence, $\liminf_{T\to\infty} \mathbb{P}\left[\boldsymbol{\gamma}_T \in \Delta^{\gtrsim}_{(\mathcal{U},\mathrm{K},\mathcal{F}_T)}\right] = 1$ and $\boldsymbol{\gamma}_T \to \boldsymbol{\delta}$. The previous then imply that in all cases $\boldsymbol{\delta}$ lies in the $\liminf$ of $\Delta^{\gtrsim}_{(\mathcal{U},\mathrm{K},\mathcal{F}_T)}$ w.h.p. Any other element of $\Delta^{(c_T)}_{(\mathcal{U},\mathrm{K})}(v, \mathcal{F})$ that lies in the empirical SAO set with asymptotically positive probability will necessarily converge to some element of $\boldsymbol{\delta}^*_{(\mathcal{U},\mathrm{K})}(v, \mathcal{F})$. The previous establish Theorem 3.2.1.(i), since there cannot exist accumulation points of sequences of elements of $\Delta^{(c_T)}_{(\mathcal{U},\mathrm{K})}(v, \mathcal{F})$ that lie outside $\boldsymbol{\delta}^*_{(\mathcal{U},\mathrm{K})}(v, \mathcal{F})$.□

Proof of Theorem 3.2.2. By Assumption 3.1.1 we obtain that $\mathbb{E}_{\mathcal{F}_T}\left[v(\boldsymbol{x}^{\mathrm{T}}\left(\boldsymbol{\kappa} + \boldsymbol{\delta}\right))\right]$ converges in probability to $\mathbb{E}_{\mathcal{F}}\left[v(\boldsymbol{x}^{\mathrm{T}}\left(\boldsymbol{\kappa} + \boldsymbol{\delta}\right))\right]$ uniformly in $\boldsymbol{\kappa}, \boldsymbol{\delta}$. Then the results in i), ii), iii) follow via Theorem 3.2.1, employing Skorokhod representations applicable due to Theorem 3.7.25 of Giné and Nickl (2016), using Proposition 3.2, Ch. 5 of Molchanov (2006) and then reverting to the original probability space. Specifically for iii), due to the aforementioned uniform convergence, $\mathbb{E}_{\mathcal{F}_T}\left[v(\boldsymbol{x}^{\mathrm{T}}\left(\boldsymbol{\kappa} + \boldsymbol{\delta}\right))\right] - \mathbb{E}_{\mathcal{F}_T}\left[v(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\kappa})\right] \leq 0$ w.h.p., for any $\boldsymbol{\delta} \in \Delta$. Hence any $\boldsymbol{\delta} \in \Delta^{\equiv}_{\mathcal{F}}$ lies in $\boldsymbol{\delta}^*_{(\mathcal{U},\mathrm{K})}(v, \mathcal{F}_T)$ w.h.p.□

Proof of Proposition 3.3.2. Notice first that due to Assumption 3.3.1.i)-ii), the ergodic theorem and Theorem 4.2 of Rio (2017),

$$\sqrt{T}\left(\theta_T - \theta_0\right) \rightsquigarrow N\left(0_d, \mathbb{E}\left[H\left(\theta, \boldsymbol{f}_0\right)\right] \sum_{t\in\mathbb{Z}} \mathrm{Cov}\left(s\left(\theta_0, \boldsymbol{f}_0\right), s\left(\theta_0, \boldsymbol{f}_t\right)\right) \mathbb{E}\left[H\left(\theta, \boldsymbol{f}_0\right)\right]\right).$$

Furthermore, and due to the previous, Assumption 3.3.1.iii), the MVT and dominated convergence, for any $x, \boldsymbol{\kappa}, \boldsymbol{\delta}$, w.h.p. and for some $\theta^\star_T \in \Theta^\star$ on the line that connects θ_T and θ_0,

$$\begin{gathered}\sqrt{T} D\left(u, \boldsymbol{\kappa}, \boldsymbol{\delta}, \mathcal{F}_{\theta_T} - \mathcal{F}\right) \\ = \mathbb{E}_{\mathcal{F}_{\theta^\star_T}}\left[\left(u\left(\boldsymbol{x}^{\mathrm{T}}\left(\boldsymbol{\kappa} + \boldsymbol{\delta}\right)\right) - u\left(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\kappa}\right)\right) l^{\mathrm{T}}_{\theta^\star_T}\left(\boldsymbol{f}\right)\right] \sqrt{T}\left(\theta_T - \theta_0\right)\end{gathered}.$$

The uniform boundedness of $\mathrm{cl}\left(\mathcal{U}\right)$ implies that the difference $u\left(\boldsymbol{x}^{\mathrm{T}}\left(\boldsymbol{\kappa} + \boldsymbol{\delta}\right)\right) - u\left(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\kappa}\right)$ is almost everywhere uniformly bounded. Due to this, Assumption 3.3.1.iii), and the Cauchy-Schwarz inequality, there exists a universal constant $k > 0$, such that, for any $y \in \mathbb{R}^d$,

$$\begin{gathered}\left|\sup_{\Theta^\star} \mathbb{E}\left[\left(u\left(\boldsymbol{x}^{\mathrm{T}}\left(\boldsymbol{\kappa} + \boldsymbol{\delta}\right)\right) - u\left(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\kappa}\right)\right) l^{\mathrm{T}}_{\theta^\star_T}\left(\boldsymbol{f}_0\right)\right] y\right| \leq \\ \sup_{\Theta^\star} \mathbb{E}\left[\left|u\left(\boldsymbol{x}^{\mathrm{T}}\left(\boldsymbol{\kappa} + \boldsymbol{\delta}\right)\right) - u\left(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\kappa}\right)\right| \left\|l_\theta\left(\boldsymbol{f}_0\right)\right\|\right] \left\|y\right\| \leq \\ \leq k \sup_{\Theta^\star} \mathbb{E}\left[\left\|l_\theta\right\|^{2+\delta}\right] \left\|y\right\|.\end{gathered}, \tag{16}$$

Due to the compactness of Θ, the weak convergence above, uniform integrability, and (16) we obtain the fidi part of convergence in Assumption 3.1.1. The compactness of $\triangle \times K$, Theorem 2.7.5 of van der Vaart and

Wellner (1996), and (16), imply asymptotic tightness in probability. Hence, Theorem 3.3.2 follows.□

Proof of Proposition 3.3.4. The uniform boundedness of $\mathrm{cl}\,(\mathcal{U})$ implies that the difference $u\left(\boldsymbol{x}^{\mathrm{T}}\left(\boldsymbol{\kappa}+\boldsymbol{\delta}\right)\right)-u\left(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\kappa}\right)$ is almost everywhere uniformly bounded. Assumption 3.3.3.iv) implies, via dominated convergence and the smoothness of the utilities, that $\mathbb{E}_{\mathcal{F}}\left[\left[u(\boldsymbol{x}^{\mathrm{T}}\left(\boldsymbol{\kappa}+\boldsymbol{\delta}\right))\right]^{p}\right]$ is continuous in $\boldsymbol{z}$ for $p=1,2,2+\delta$, where δ is as in Assumption A.3.3.1.i, and that $f_{\boldsymbol{Z}}$ and $\mathbb{E}_{\mathcal{F}}\left[u(\boldsymbol{x}^{\mathrm{T}}\left(\boldsymbol{\kappa}+\boldsymbol{\delta}\right))\right]$ are twice differentiable in $\boldsymbol{z}$ with bounded Jacobian matrices. Thus Assumption 3.3.3.i)-iv) implies the validity of Corollary 1 of El Machkouri, Fan, and Reding, (2020), for every finite linear combination of $\left(\frac{1}{b^{L}}\mathcal{K}\left(\frac{\boldsymbol{Z}-\boldsymbol{z}}{b}\right)\ \left(u\left(\boldsymbol{x}^{\mathrm{T}}\left(\boldsymbol{\kappa}+\boldsymbol{\delta}\right)\right)-u\left(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\kappa}\right)\right)\ \right)_{u,\boldsymbol{\kappa},\boldsymbol{\delta}}$ thereby establishing the required FiDi convergence for the weak convergence needed via the Cramer-Wold device. In order to establish tightness, notice that the compactness of $\triangle\times K$, Theorem 2.7.5 of van der Vaart and Wellner (1996) and Assumption 3.3.3.v) along with the fact that $b_{T}\downarrow 0$ and the Cauchy-Schwarz inequality imply that the bracketing numbers of the class$\left\{\frac{1}{b^{L}}\mathcal{K}\left(\frac{\boldsymbol{Z}-\boldsymbol{z}}{b}\right)\left[u\left(\boldsymbol{x}^{\mathrm{T}}\left(\boldsymbol{\kappa}+\boldsymbol{\delta}\right)\right)-u\left(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\kappa}\right)\right],u,b,\boldsymbol{\kappa},\boldsymbol{\delta}\right\}$ w.r.t. the $L_{2}\left(\mathcal{F}\right)$ metric are asymptotically equivalent to $\frac{c}{\varepsilon^{D}}$ as $\varepsilon\downarrow 0$ for some $D>0$. The class also has an $L_{2}\left(\mathcal{F}\right)$ integrable envelope due to uniform boundedness and the existence of $\int_{\mathbb{R}^{L+M}}\mathcal{K}^{2}\left(\boldsymbol{y}\right)f_{\boldsymbol{f},\boldsymbol{Z}}\left(\boldsymbol{f},\boldsymbol{y}\right)d\boldsymbol{f}d\boldsymbol{y}$. Fidi convergence and the previous imply then the weak convergence needed, via Theorem 8.1 of Rio (2017).□

ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS
ΣΧΟΛΗ ΟΙΚΟΝΟΜΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
SCHOOL OF ECONOMIC SCIENCES
ΤΜΗΜΑ ΟΙΚΟΝΟΜΙΚΗΣ ΕΠΙΣΤΗΜΗΣ
DEPARTMENT OF ECONOMICS

Department of Economics
Athens University of Economics and Business

List of Recent Working Papers

2020

- **01-20** Incubated Early Stage Startuppers in Athens and their Initiatives during the Crisis (2010-2016), Ioannis Besis and Ioanna Sapfo Pepelasis
- **02-20** Alternative Monetary Policy Rules in an Imperfectly Competitive Dynamic Stochastic General Equilibrium Model, George Alogoskoufis and Stelios Giannoulakis
- **03-20** Quantifying Qualitative Survey Data: New Insights on the (Ir)Rationality of Firms' Forecasts, Alexandros Botsis, Christoph Görtz and Plutarchos Sakellaris
- **04-20** Can Country-Specific Interest Rate Factors Explain the Forward Premium Anomaly?, Efthymios Argyropoulos, Nikolaos Elias, Dimitris Smyrnakis and Elias Tzavalis
- **05-20** Signaling product quality through prices in an oligopoly model with costly consumer search, Minghua Cheny, Konstantinos Serfes and Eleftherios Zacharias
- **06-20** Thinking ahead of the next big crash: Clues from Athens in classical times, George C. Bitros
- **07-20** Financial crises, firm-level shocks, and large downturns: Evidence from Greece, Stelios Giannoulakis and Plutarchos Sakellaris
- **08-20** Notes on the Demand Side Secular Stagnation, George D. Demopoulos and Nicholas A. Yannacopoulos
- **09-20** Institutions and Macroeconomic Performance: Core vs Periphery Countries in the Eurozone, Tryfonas Christou, Apostolis Philippopoulos and Vanghelis Vassilatos
- **10-20** Macroeconomic Policy Lessons for Greece, George Economides, Dimitris Papageorgiou and Apostolis Philippopoulos
- **11-20** Energy Transitions and the role of the EU ETS: The case of Greece, Andriana Vlachou and Georgios Pantelias
- **12-20** Measuring the Default Risk of Small Business Loans: Improved Credit Risk Prediction using Deep Learning, Yiannis Dendramis, Elias Tzavalis and Aikaterini Cheimarioti
- **14-20** Saving Democracy by Switching to Corporate-like Governance, George C. Bitros
- **15-20** The impact of the lockdown on the Greek economy and the role of the Recovery Fund, George Economides, Apostolis Philippopoulos and Vanghelis Vassilatos

2021

- **01-21** Historical Cycles of the Economy of Modern Greece From 1821 to the Present, George Alogoskoufis
- **02-21** Greece Before and After the Euro: Macroeconomics, Politics and the Quest for Reforms, George Alogoskoufis
- **03-21** Commodity money and the price level, George C. Bitros

- **04-21** **Destabilizing asymmetries in central banking: With some enlightenment from money in classical Athens, George C. Bitros**
- **05-21** **Exploring the Long-Term Impact of Maximum Markup Deregulation, Athanasios Dimas and Christos Genakos**
- **06-21** **A regularization approach for estimation and variable selection in high dimensional regression models, Y. Dendramis, L. Giraitis, G. Kapetanios**
- **07-21** **Tax Competition in the Presence of Environmental Spillovers, Fabio Antoniou, Panos Hatzipanayotou, Michael S. Michael, Nikos Tsakiris**
- **08-21** **Firm Dynamics by Age and Size Classes and the Choice of Size Measure, Stelios Giannoulakis and Plutarchos Sakellaris**
- **09-21** **Measuring the Systemic Importance of Banks, Georgios Moratis, Plutarchos Sakellaris**
- **10-21** **Firms' Financing Dynamics Around Lumpy Capacity Adjustments, Christoph Görtz, Plutarchos Sakellaris, John D. Tsoukalas**
- **11-21** **On the provision of excludable public goods General taxes or user prices? George Economides and Apostolis Philippopoulos**
- **12-21** **Asymmetries of Financial Openness in an Optimal Growth Model, George Alogoskoufis**
- **13-21** **Evaluating the impact of labour market reforms in Greece during 2010-2018 , Georgios Gatopoulos, Alexandros Louka, Ioannis Polycarpou, Nikolaos Vettas**
- **14-21** **From the Athenian silver to the bitcoin standard: Private money in a state-enforced free banking model, George C. Bitros**

Πατησίων 76, 104 34 Αθήνα. Τηλ.: 210 8203303-5. E-mail: econ@aueb.gr / www.aueb.gr
76, Patission Street, Athens 104 34 Greece. Tel.: (+30) 210 8203303-5

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/

Πατησίων 76, 104 34 Αθήνα. Τηλ.: 210 8203303-5. E-mail: econ@aueb.gr / www.aueb.gr
76, Patission Street, Athens 104 34 Greece. Tel.: (+30) 210 8203303-5