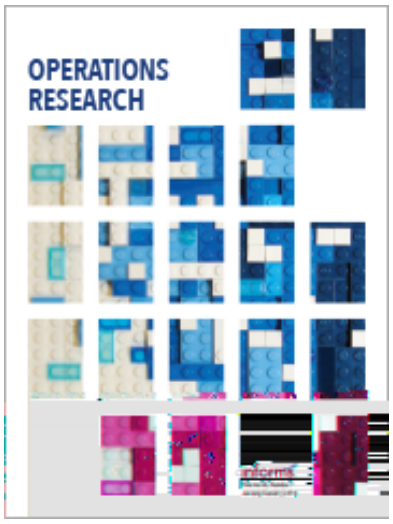
This article was downloaded by: [5.78.95.66] On: 20 March 2024, At: 07:31
Publisher: Institute for Operations Research and the Management Sciences (INFORMS)
INFORMS is located in Maryland, USA



Operations Research

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

Optimal Impact Portfolios with General Dependence and Marginals

Andrew W. Lo, Lan Wu, Ruixun Zhang, Chaoyi Zhao

To cite this article:
Andrew W. Lo, Lan Wu, Ruixun Zhang, Chaoyi Zhao (2024) Optimal Impact Portfolios with General Dependence and Marginals. Operations Research

Published online in Articles in Advance 20 Mar 2024

. https://doi.org/10.1287/opre.2023.0400

Full terms and conditions of use: https://pubsonline.informs.org/Publications/Librarians-Portal/PubsOnLine-Terms-and-Conditions

This article may be used only for the purposes of research, teaching, and/or private study. Commercial use or systematic downloading (by robots or other automatic processes) is prohibited without explicit Publisher approval, unless otherwise noted. For more information, contact permissions@informs.org.

The Publisher does not warrant or guarantee the article's accuracy, completeness, merchantability, fitness for a particular purpose, or non-infringement. Descriptions of, or references to, products or publications, or inclusion of an advertisement in this article, neither constitutes nor implies a guarantee, endorsement, or support of claims made of that product, publication, or service.

Copyright © 2024, INFORMS

Please scroll down for article—it is on subsequent pages

With 12,500 members from nearly 90 countries, INFORMS is the largest international association of operations research (O.R.) and analytics professionals and students. INFORMS provides unique networking and learning opportunities for individual professionals, and organizations of all types and sizes, to better understand and use O.R. and analytics tools and methods to transform strategic visions and achieve better outcomes.
For more information on INFORMS, its publications, membership, or meetings visit http://www.informs.org

https://pubsonline.informs.org/journal/opre

OPERATIONS RESEARCH
Articles in Advance, pp. 1–15
ISSN 0030-364X (print), ISSN 1526-5463 (online)

Downloaded from informs.org by [5.78.95.66] on 20 March 2024, at 07:31 . For personal use only, all rights reserved.

Contextual Areas

Optimal Impact Portfolios with General Dependence and Marginals

Andrew W. Lo,[a,b,c,d] Lan Wu,[e,f] Ruixun Zhang,[e,f,g,h,*] Chaoyi Zhao[e,*]

[a] MIT Sloan School of Management, Cambridge, Massachusetts 02142; [b] Laboratory for Financial Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02142; [c] Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts 02142; [d] Santa Fe Institute, Santa Fe, New Mexico 87501; [e] School of Mathematical Sciences, Peking University, Beijing 100871, China; [f] Laboratory for Mathematical Economics and Quantitative Finance, Peking University, Beijing 100871, China; [g] Center for Statistical Science, Peking University, Beijing 100871, China; [h] National Engineering Laboratory for Big Data Analysis and Applications, Peking University, Beijing 100871, China

*Corresponding author

Contact: alo@mit.edu (AWL); lwu@pku.edu.cn (LW); zhangruixun@pku.edu.cn, https://orcid.org/0000-0002-7670-8393 (RZ); zhaochaoyi@pku.edu.cn, https://orcid.org/0000-0002-5921-3612 (CZ)

Received: July 23, 2023
Revised: November 2, 2023; February 5, 2024
Accepted: February 12, 2024
Published Online in Articles in Advance: March 2, 2024

Area of Review: Financial Engineering

https://doi.org/10.1287/opre.2023.0400

Copyright: © 2024 INFORMS

Abstract. We develop a mathematical framework for constructing optimal impact portfolios and quantifying their financial performance by characterizing the returns of impact-ranked assets using induced order statistics and copulas. The distribution of induced order statistics can be represented by a mixture of order statistics and uniformly distributed random variables, where the mixture function is determined by the dependence structure between residual returns and impact factors—characterized by copulas—and the marginal distribution of residual returns. This representation theorem allows us to explicitly and efficiently compute optimal portfolio weights under any copula. This framework provides a systematic approach for constructing and quantifying the performance of optimal impact portfolios with arbitrary dependence structures and return distributions.

Funding: Research funding from the China National Key R&D Program [Grant 2022YFA1007900], the China National Natural Science Foundation [Grants 12271013, 72342004], the Fundamental Research Funds for the Central Universities (Peking University), and the MIT Laboratory for Financial Engineering is gratefully acknowledged.

Supplemental Material: The online appendix is available at https://doi.org/10.1287/opre.2023.0400.

Keywords: impact investing • environmental, social, and governance (ESG) • induced order statistics • copula • representation theorem • portfolio theory

1. Introduction

Impact investing, broadly defined as investments that consider not only financial objectives but also other goals that support certain social priorities and agendas, has drawn an increasing amount of attention in recent years. This form of investing has been given other names, such as "sustainable" and "green" investing, and examples include divestment from "sin stocks," "socially responsible" investing (SRI), the use of environmental, social, and governance (ESG) investment criteria, and private equity funds seeking social impact alongside financial returns.[1]

The construction of impact portfolios typically involves ranking assets based on a measure of impact, such as a company's ESG score, its amount of carbon emissions, or the prospect of developing a disease-modifying drug. Popular methods include negative screening, imposing filters so that certain companies are excluded from the investable universe; positive screening, selecting companies for high values of certain attributes; relative weighting, over- or under-weighting companies within an industry based on a certain measure; and full ESG factor integration in which ESG information is combined with other fundamental and technical factors of a company to improve the investment process.[2]

Despite the rapid growth in popularity and assets under the management of impact investing, it is unclear whether they are adding or removing value from an investor's point of view.[3] In addition, there are ongoing controversies regarding whether ESG portfolios and other impact investing products may be violations of the fiduciary duties of the managers of those products.[4] These challenges call for a general framework for constructing optimal impact portfolios and for measuring and disclosing the financial impact of impact investing.

1.1. Main Results

In this article, we develop a framework to quantify the distribution of asset returns after ranking, screening,

Downloaded from informs.org by [5.78.95.66] on 20 March 2024, at 07:31 . For personal use only, all rights reserved.

and weighting based on a measure of impact. We demonstrate how to construct optimal impact portfolios, and explicitly quantify their performance in terms of the well-known information ratio (see, for example, Grinold and Kahn 2019). Our framework allows for arbitrary marginal distributions of the asset returns and the impact factor, as well as arbitrary dependence structures between the two, characterized by copulas.

We formalize impact investing as the sorting and selection within an investment universe of N assets based on an impact factor, X_i, for asset i, so that higher values of X_i correspond to greater impact, for example, lower carbon emissions, higher ESG score, etc. We consider a linear multifactor model for asset returns and focus on the residual returns in excess of all known factors, denoted by $\boldsymbol{\theta} \equiv (\theta_1, \theta_2, ;;;, \theta_N)^\top$. The impact on investment performance is therefore determined by the joint distribution of the vector $\boldsymbol{X} \equiv (X_1, X_2, ;;;, X_N)^\top$ of impact factors with the residual returns, $\boldsymbol{\theta}$.

Our fundamental tool to quantify the distribution of $\boldsymbol{\theta}$ after ranking based on $\boldsymbol{X}$ is the *induced order statistic*, which uses random variables that are ranked not by their own values ($\boldsymbol{\theta}$ in our case) but by the values of other random variables ($\boldsymbol{X}$ in our case). However, the existing literature on induced order statistics often relies on strong distributional assumptions, and their properties under general distributions are rarely discussed.[5]

Our main result is a *representation theorem*, which shows that, for a *general* bivariate distribution of $(X_i, \theta_i)^\top$, the distribution of induced order statistics of θ_i ranked by X_i can be represented by a mixture of order statistics and uniformly distributed random noise, and the mixture function is determined by the copula of X_i and θ_i and the marginal distribution of θ_i. These results completely characterize the distribution—both in nite sample and asymptotically—of induced order statistics and allow for the ef cient computation of optimal portfolio weights. To the best of our knowledge, we are the rst to provide results for induced order statistics under general copulas and marginal distributions.

When the distribution of $(X_i, \theta_i)^\top$ is discontinuous, our construction proof shows that the representation holds if and only if the copula is linearly interpolating in the unde ned regions given by Sklar's (1959) theorem.[6] This generality allows us to derive weights for optimal impact portfolios when the impact factor, $\boldsymbol{X}$, represents discrete rating values, which is often the case in practice, and when residual returns, $\boldsymbol{\theta}$, include jumps or take discrete values in high-frequency data.

The representation theorem has several important applications. To understand the impact of general dependence structures between the residual return, $\boldsymbol{\theta}$, and the impact factor, $\boldsymbol{X}$, we apply the representation theorem under different copulas. In particular, we can quantify the impact of asymmetric tail dependence between asset returns and the impact factor—modeled by the Archimedean copula—which is commonly observed in many impact investing contexts.[7] We nd that stronger tail dependence implies larger weights for assets on that tail in optimal impact portfolios.

In addition, we use the representation theorem to study the impact of the marginal distribution of residual returns on impact portfolios. It is well documented that stock returns are skewed and heavy-tailed. We show that a larger positive (negative) skewness leads to smaller weights for higher (lower) impact-ranked assets. Heavy-tailed returns may lead to nonmonotonic weights for the optimal impact portfolio, because of the extreme levels of risk from assets with the most extreme impact values.

1.2. Related Literature

Our work contributes to a rapidly growing literature on quantitative models for green nance. For example, Blasberg et al. (2021) propose the carbon default swap to manage the exposure to transition risk. De Angelis et al. (2023) model the effect of impact investors on mitigating climate risks by raising the cost of capital of the most carbon-intensive companies. Aïd and Biagini (2023) discuss the impact of regulations on the carbon emission market. Gobet and Lage (2023) model the optimal transition path of credit portfolios. Our work focuses on impact investing from the perspective of portfolio construction. The literature on green portfolio construction and sustainable asset pricing includes, for example, Pástor et al. (2021), Pedersen et al. (2021), Sorensen et al. (2021), and Flora and Tankov (2023). Although this literature incorporates green constraints into the traditional asset-pricing framework, the interaction between these constraints and asset returns is not explicitly modeled.

Our results extend the recent work of Lo and Zhang (2023), which models the dependence between impact factor and asset returns when they are jointly normally distributed with a constant correlation. However, as demonstrated empirically in Online Appendix EC.2, impact factors and returns do not follow normal distributions in practice. Although the joint normality assumption yields simple analytical results, it is unclear how impact portfolios and their performances are affected by deviations from normality (e.g., skewed and heavy-tailed distributions) and asymmetric dependencies between asset returns and impact factors. In this article, we provide a complete characterization of these interactions and a framework to construct impact portfolios that account for them. Empirical results in Online Appendix EC.2 show that allowing for general

Downloaded from informs.org by [5.78.95.66] on 20 March 2024, at 07:31 . For personal use only, all rights reserved.

dependence and marginal distributions can significantly improve the financial performance of impact portfolios.

As mentioned above, we contribute to the theory of induced order statistics by characterizing their distributions under general dependence structures. These statistics, whose term was first coined by Bhattacharya (1974), are also referred to as *concomitants* of the order statistics (David 1973).[8] Lo and MacKinlay (1990) apply these same statistical tools to quantify data-snooping biases in testing financial asset-pricing models. To our knowledge, we are the first to combine induced order statistics with copulas, and the latter are widely used in credit risk models (Giesecke 2003, 2004; Cont and Kan 2011; Filiz et al. 2012; Cont and Minca 2013; Brigo et al. 2014). In this article, we derive the distribution of induced order statistics under arbitrary copulas, and provide a novel application in the context of impact investing.

We present the main results under the assumption that $(X_i, \theta_i)^\top$ is independent and identically distributed (IID) across assets (Assumption 1) for three reasons. First, the analytical results under the IID assumption yield considerable intuition that is mathematically clear yet fundamental to describing the distribution of induced order statistics and the performance of impact portfolios. Second, the main results can be generalized to allow for cross-sectional heterogeneity with more mathematical complexity (see Online Appendix EC.1.3.2), but both the representation theorem and its associated intuition carry over. Finally, our characterization of induced order statistics under general marginal distributions and dependence is new even under this IID assumption, which may yield applications in other domains and contexts beyond impact investing.

In fact, the theory of induced order statistics for bivariate random variables with arbitrary copulas can potentially be applied much more broadly to contexts that involve ranking bivariate pairs by one covariate. This includes portfolio selection based on hundreds of new factors and anomalies in the "Factor Zoo" discussed in the recent literature (Harvey et al. 2016, Feng et al. 2020, Hou et al. 2020), as well as the pricing and portfolio construction of credit bonds, credit default swaps (CDS), CDS indices, and so on, because investors may form portfolios by ranking these assets based on their credit ratings.

1.3. Outline

The remainder of this article is organized as follows. We describe our framework in Section 2. Section 3 shows the representation theorem. We then discuss the impact of general dependence structures in Section 4 and marginal distributions in Section 5. We conclude in Section 6, and provide additional technical details, empirical studies, and proofs in the online appendix.

2. The Framework

We consider a world of N assets whose returns in excess of the risk-free rate satisfy the following linear multifactor model:

$$r_{it} = \alpha_i + \beta_{i1}\Lambda_{1t} + \beta_{i2}\Lambda_{2t} + \cdots + \beta_{iK}\Lambda_{Kt} + \varepsilon_{it}, \quad i = 1, 2, \ldots, N; \quad t = 1, 2, \ldots, T; \tag{1}$$

where Λ_{kt} is the k-th factor return (in excess of the risk-free rate), $k = 1, 2, \ldots, K$, α_i and β_{ik} are the excess return and factor betas, respectively, and ε_{it} is the idiosyncratic return component independent of all factors. The ε_{it}'s are zero-mean random variables that are IID both across assets i and over time t. To allow for potential mispricings, we assume that α_i are random variables that are IID across assets i, and they follow a nondegenerate distribution.[9]

We define the following quantity to be the *residual return* of asset i at time t, for $i = 1, 2, \ldots, N$ and $t = 1, 2, \ldots, T$:

$$\theta_{it} \equiv \alpha_i + \varepsilon_{it}; \tag{2}$$

Therefore, given a specific time t, the unconditional distribution of θ_{it} is IID across asset i. Given a specific asset i, the unconditional distribution of θ_{it} is identically distributed but dependent over time t because of the common term, α_i.

Under suitable restrictions on $\{\alpha_i, \beta_{ik}\}$ and the definitions of the factor returns $\{\Lambda_k\}$, the linear multifactor model in (1) is consistent with a number of asset-pricing models, such as the Capital Asset Pricing Model (CAPM) (Sharpe 1964, Lintner 1965), the Intertemporal CAPM (Merton 1973), the Arbitrage Pricing Theory (APT) (Ross 1976), and the Fama-French multifactor models (Fama and French 1993). In particular, all of these asset-pricing models imply that $\alpha_i = 0$, and that the expected returns of the assets are simply the sum of the risk-free rate plus all the risk premia multiplied by the asset's corresponding risk exposures or β_{ik}'s.

However, to measure the impact of impact investing, we take no position as to whether any particular asset-pricing model holds. We allow for the possibility of superior investment performance of individual securities, but also include the conventional case of equilibrium or no-arbitrage pricing if we set the mean and variance of the alphas to zero. The implications of our model are broadly applicable to an equilibrium asset-pricing setup where "alpha" may be reinterpreted as omitted factors which investors are either unaware of or unable to access as easily as portfolio managers (see Lo and Zhang 2023, section 4.4).

2.1. Impact Factors and Induced Order Statistics

Impact investors rank assets according to a specific impact variable,[10] and form portfolios based on the impact variable. We assume that there exists a specific

Downloaded from informs.org by [5.78.95.66] on 20 March 2024, at 07:31 . For personal use only, all rights reserved.

impact factor for security i at time t (e.g., the ESG score), X_{it}, which may be correlated with the residual returns of the i-th asset, θ_{it}. Although θ_{it} is not observable at time t, its corresponding impact factor, X_{it}, is observable. Let $\boldsymbol{X}_t = (X_{1t}, ;;;, X_{Nt})^\top$ and $\boldsymbol{\theta}_t = (\theta_{1t}, ;;;, \theta_{Nt})^\top$. In particular, we assume the following:

Assumption 1. *For a specific time t,*

$$\begin{matrix} X_{1t} \\ \theta_{1t} \end{matrix}, \begin{matrix} X_{2t} \\ \theta_{2t} \end{matrix}, ;;;, \begin{matrix} X_{Nt} \\ \theta_{Nt} \end{matrix} \overset{\text{IID}}{\sim} F(\cdot, \cdot); \tag{3}$$

Assumption 1 states that the impact factors and residual returns are IID across different assets. Independence across assets is justi able because we are modeling residual returns. The assumption of identical distributions allows for mathematical tractability and implies that the relationship between X_{it} and θ_{it} is homogeneous. In our main article, we employ this IID assumption to derive concise theoretical results and useful mathematical intuition. However, the IID assumption can be relaxed to model cross-sectional heterogeneity (such as industry effects) by mixtures of homogeneous subgroups. We relax this assumption in Online Appendix EC.1.3.2. Moreover, because the majority of this article discusses the static case, we omit the subscript t for notational simplicity except when considering a dynamic model in Online Appendix EC.1.2.3.

The bivariate vectors, $(X_i, \theta_i)^\top, i = 1, 2, ;;;, N$, are ranked according to their rst components (e.g., ESG scores), X_i:

$$\begin{matrix} X_{1:N} \\ \theta_{[1:N]} \end{matrix}, \begin{matrix} X_{2:N} \\ \theta_{[2:N]} \end{matrix}, ;;;, \begin{matrix} X_{N:N} \\ \theta_{[N:N]} \end{matrix}, \tag{4}$$

where $X_{1:N} \le X_{2:N} \le \cdots \le X_{N:N}$ are the order statistics of $\boldsymbol{X}$. The notation $\theta_{[i:N]}$ represents the i-th *induced order statistic* of $\boldsymbol{\theta}$ induced by another variable, $\boldsymbol{X}$. Hereinafter, we use the notation

$$\boldsymbol{\theta}_{[X]} = (\theta_{[1:N]}, \theta_{[2:N]}, ;;;, \theta_{[N:N]})^\top \tag{5}$$

to denote the vector of induced order statistics of $\boldsymbol{\theta}$ ranked by values of $\boldsymbol{X}$.

The framework in (1)–(4) is different from the impact investing model proposed in Lo and Zhang (2023) in several important ways. First, we generalize Lo and Zhang (2023) by considering the interaction between the impact factor and residual returns, $\boldsymbol{\theta}$, instead of the alpha as in Lo and Zhang (2023). This is also consistent with the framework of the fundamental law of active management in Grinold (1989). Second, Lo and Zhang (2023) consider the case where $(X_i, \theta_i)^\top$ follows a bivariate normal distribution, which implies that the dependence between the impact factor and asset return is described by a single correlation coef cient. Here, we consider a much richer set of dependence structures modeled by copulas (see Section 4). Third, the general bivariate distribution F in (3) allows for the analysis of different marginal distributions of X_i and θ_i. In particular, asset returns in practice include important deviations from normality such as nonzero skewness and heavy-tailedness which have important implications for impact portfolio construction, as shown in Section 5. Our empirical results in Online Appendix EC.2 demonstrate that allowing for general marginal distributions and dependence structures between X_i and θ_i can greatly enhance the performance of impact portfolios.

2.2. Impact Portfolio Construction

Impact investing essentially involves selecting securities based on the impact factor, $\boldsymbol{X}$. We consider a portfolio based on the N assets, and $= (w_1, w_2, ;;;, w_N)^\top$ represents the vector of asset weights for the N assets ranked by $\boldsymbol{X}$. In other words, we hold a fraction of our portfolio, w_i, in the asset with the i-th smallest impact factor, $X_{i:N}$. The residual return of the portfolio is therefore

$$\theta_p \equiv w_1\theta_{[1:N]} + w_2\theta_{[2:N]} + \cdots + w_N\theta_{[N:N]} = {}^\top\boldsymbol{\theta}_{[X]}; \tag{6}$$

Equation (6) shows that the residual return of the portfolio, θ_p, is determined by the distribution of $\boldsymbol{\theta}_{[X]}$ and the choice of . In addition, impact investors use the distribution of $\boldsymbol{\theta}_{[X]}$ to determine the portfolio weights, . Therefore, our goal is to characterize the distribution of $\boldsymbol{\theta}_{[X]}$, given the joint distribution F in Assumption 1.

To quantify the performance of the impact portfolio, let $\boldsymbol{\mu}$ and Σ be the expectation and covariance matrix of the residual returns of the N impact-ranked assets:

$$\boldsymbol{\mu} \equiv \mathbb{E}(\boldsymbol{\theta}_{[X]}), \quad \Sigma \equiv \mathrm{Cov}(\boldsymbol{\theta}_{[X]}); \tag{7}$$

Then, the expectation and variance of the portfolio s residual return are

$$\mathbb{E}(\theta_p) = {}^\top\boldsymbol{\mu}, \quad \mathrm{Var}(\theta_p) = {}^\top\Sigma \ ; \tag{8}$$

The optimal portfolio weights and their corresponding performance metrics can be derived from standard portfolio theory. For completeness, we explicitly provide these constructions based on the moments of residual returns given by (7), which serves as the foundation for our results in subsequent sections. We consider impact investors who use either the information ratio or the mean-variance utility as the objective function to form portfolios. The following proposition from standard portfolio theory characterizes the optimal portfolios with these objectives.

Proposition 1 (Optimal Portfolio). *Under the multifactor model of* (1), *if investors construct portfolios based on N assets with frictionless borrowing and lending at the risk-free rate, and if they maximize the information ratio,* $\mathrm{IR} = \mathbb{E}(\theta_p) > \sqrt{\mathrm{Var}(\theta_p)}$, *or the mean-variance utility,* $\mathbb{E}(\theta_p) \quad 0;5\lambda\mathrm{Var}(\theta_p)$, *with a constant risk-aversion*

Downloaded from informs.org by [5.78.95.66] on 20 March 2024, at 07:31 . For personal use only, all rights reserved.

parameter $\lambda > 0$, *the optimal portfolio weights and the optimal information ratio will be given by*

$$^{*} \propto \Sigma^{-1}\boldsymbol{\mu}, \quad \textit{and} \quad \mathrm{IR}^{*} = \sqrt{\boldsymbol{\mu}^{\top}\Sigma^{-1}\boldsymbol{\mu}}, \tag{9}$$

where $\boldsymbol{\mu}$ *and* Σ *are given by* (7).

Proposition 1 characterizes the optimal portfolio weights. This is useful for both relative weighting strategies, when companies are reweighted based on the rank of a certain measure,[11] and exclusionary investing, when filters are imposed to exclude certain companies from the investable universe. We discuss the trade-off between performance and impact for exclusionary investing in more detail in Online Appendix EC.1.2.6, and provide an extension to accommodate investors who care about both active and passive returns in Online Appendix EC.1.1. In this article, without loss of generality, we set the proportional constant to be one, that is, $^{*} = \Sigma^{-1}\boldsymbol{\mu}$.

One advantage of our framework is that one needs to estimate a much smaller number of parameters in portfolio construction. In Markowitz portfolio theory, one needs to estimate $N(N+1)/2$ parameters for the covariance matrix (Sharpe 1963). In comparison, our framework circumvents this issue because both $\boldsymbol{\mu}$ and Σ depend solely on the bivariate distribution F defined in (3). With certain distributional assumptions for F, one only needs to estimate a small number of parameters of F. This leads to more robust impact portfolios in practice, as demonstrated empirically by Lo et al. (2022).

In this article, we focus on the residual return of the portfolio rather than its raw return for the following reasons. First, institutional investors usually assess a portfolio's *relative* performance in comparison with benchmark portfolios (such as an index fund) or baseline asset-pricing models (such as the Fama-French factor models), rather than a portfolio's raw returns. This is the typical target of interest for ESG and sustainable investment products.[12] Second, more broadly, the literature on active investing also focuses on the residual returns to derive active investment performance metrics such as the information ratio.[13] Our choice follows both strands of the literature. Third, although residual returns of a portfolio are not directly realizable by holding individual assets, they are realizable when using index funds and factor exchange-traded funds to neutralize a portfolio's exposure to existing asset-pricing factors. Investors can realize the excess returns in this case.

3. The Distribution of Induced Order Statistics

The key to constructing impact portfolios using Proposition 1 lies in the characterization of the distribution of induced order statistics, $\boldsymbol{\theta}_{[X]}$, and, in particular, their first and second moments. In this section, we employ copulas to derive a representation theorem for the distribution of induced order statistics as a mixture of order statistics of uniform random variables and uniform random noise. The mixture is determined by the dependence structure between residual returns and impact factors, which is characterized by copulas, as well as the marginal distribution of residual returns. We present an exact finite-sample representation theorem in Section 3.1, and provide an asymptotic version in Section 3.2.

3.1. The Representation Theorem

Sklar's (1959) well-known theorem decomposes the joint distribution of an arbitrary pair of random variables—$(X_i, \theta_i)^{\top}$ in our case—into a *copula* function that characterizes their dependence structure, and two marginal distributions. To be specific, we denote the marginal distribution functions of X_i and θ_i by F_X and F_θ, respectively. Sklar's theorem states that there exists a copula $C(\cdot,\cdot)$ such that the bivariate distribution function of $(X_i, \theta_i)^{\top}$, F, can be represented as

$$F(x,y) = C(F_X(x), F_\theta(y)), \quad \forall x, y \in \mathbb{R}; \tag{10}$$

If F_X and F_θ are continuous, C is unique; otherwise, C is uniquely determined on $\overline{\mathrm{R}}_X \times \overline{\mathrm{R}}_\theta$, where $\overline{\mathrm{R}}_X$ and $\overline{\mathrm{R}}_\theta$ are the closures of the ranges of F_X and F_θ, respectively. In this article, we refer to C as a copula of F when it satisfies (10).

Next, we present our main representation theorem that explicitly shows the relationship between the distribution of $\boldsymbol{\theta}_{[X]}$ and the three components of F—F_X, F_θ, and C—under a general F.

3.1.1. Representation Theorem Under Smoothness Condition.

We first present the key result under a smoothness condition of F to build intuition, and then provide the general representation theorem without any smoothness assumptions.

Theorem 1 (The Representation Theorem). *Under* Assumption 1, *if both* F_X *and* F_θ *are continuous functions, and* C *is a copula of* F *with a density, we have*

$$\begin{aligned}&(\theta_{[1:N]}, \theta_{[2:N]}, \ldots, \theta_{[N:N]})\\ &\overset{d}{=} (g(U_{1:N}, V_1), g(U_{2:N}, V_2), \ldots, g(U_{N:N}, V_N)),\end{aligned} \tag{11}$$

where "$\overset{d}{=}$" *denotes equality in distribution. Function* g *is defined as*

$$g(u,w) \equiv F_\theta^{-1} \circ g_u(w), \tag{12}$$

where $g_u(w)$ *is the inverse function of* $v \mapsto \frac{\partial C}{\partial u}(u,v)$,[14] *and* $\circ$ *represents function composition. In addition,* $U_1, U_2, \ldots, U_N$ *and* $V_1, V_2, \ldots, V_N$ *are mutually independent and IID random variables, each following* Uniform$(0,1)$, *and* $U_{1:N} \le U_{2:N} \le \cdots \le U_{N:N}$ *are the order statistics of* $U_1, U_2, \ldots, U_N$.

Theorem 1 provides an explicit representation for the distribution of induced order statistics, $\boldsymbol{\theta}_{[X]}$, by a

Downloaded from informs.org by [5.78.95.66] on 20 March 2024, at 07:31 . For personal use only, all rights reserved.

mixture of order statistics, $U_{i:N}$, and IID uniform noise, V_i, for any smooth bivariate distribution, F. The mixture function given by (12), $g(\cdot,\cdot)$, is constructed using F_θ and C but does not depend on F_X. This theorem explicitly establishes a relationship between the distribution of $\boldsymbol{\theta}_{[X]}$ and the joint distribution of $\boldsymbol{X}$ and $\boldsymbol{\theta}$.

The representation given by Theorem 1 requires the smoothness of F, speci cally the continuity of F_X and F_θ, which are assumptions also adopted by prior studies (Lo and Zhang 2023), and the differentiability of C, which are satis ed by many commonly used copulas. However, from both the mathematical and practical perspectives, we are still interested in whether this representation holds for a more general F.

3.1.2. Representation Theorem for General Distributions. To extend the representation theorem to general distributions F, we require a few technical results. First, when F_θ is discontinuous, we need to specify the de nition of the inverse function, $F_\theta{}^{1}(\cdot)$, in the mixture function (12). Second, when C is nondifferentiable, the function $g_u(\cdot)$ in (12), which is the inverse of $v \longmapsto \frac{@C}{@u}(u,v)$, is not well de ned. We adopt the concept of modi ed partial Dini derivative rst proposed by Fang et al. (2020), instead of the traditional de nition of derivatives, to address this issue (see Online Appendix EC.1.3.1 for further details).

Finally, when F_X is discontinuous, we demonstrate that the representation holds only when the copula of F is linearly interpolating in an undetermined region of the copula. To be precise, according to Sklar's theorem, the copula of F satisfying (10) is uniquely determined on $\overline{\mathsf{R}}_X \times \overline{\mathsf{R}}_\theta$. For the undetermined regions, we de ne a notion of linear interpolation as follows.

Definition 1 (Linearly Interpolating Copula). Let $\overline{\mathsf{R}}_X^c = [0,1]\backslash\overline{\mathsf{R}}_X$. A copula $C(u,v)$ is linearly interpolating on $\overline{\mathsf{R}}_X^c \times \overline{\mathsf{R}}_\theta$ with respect to u if

$$C(u,v) = \frac{F_X(d) \quad u}{F_X(d) \quad F_X(d\ )} C(F_X(d\ ),v) + \frac{u \quad F_X(d\ )}{F_X(d) \quad F_X(d\ )} C(F_X(d),v) \tag{13}$$

holds for any $(u,v) \in (F_X(d\ ),F_X(d)) \times \overline{\mathsf{R}}_\theta$ and $d \in \Delta_X$, where Δ_X is the set of all discontinuity points of F_X, and $F_X(d\ ) = \lim_{x\to d} F_X(x)$.[15]

In De nition 1, we say a copula is linearly interpolating on $\overline{\mathsf{R}}_X^c \times \overline{\mathsf{R}}_\theta$ with respect to u when it is a linear function of $u \in \overline{\mathsf{R}}_X^c$ given any value of $v \in \overline{\mathsf{R}}_\theta$. Online Appendix EC.1.3.1 provides the intuition behind this concept.

With these technical results in hand, we can proceed to the representation theorem of the distribution of $\boldsymbol{\theta}_{[X]}$ for a very general joint distribution function, F.

Theorem 2 (Representation Theorem for General F). *Under* Assumption 1, *we have*

$$\begin{aligned}&(\theta_{[1:N]},\theta_{[2:N]},;;;,\theta_{[N:N]})\\ &\overset{d}{=}(g(U_{1:N},V_1),g(U_{2:N},V_2),;;;,g(U_{N:N},V_N))\end{aligned} \tag{14}$$

for any $N \geq 1$ ***if a d*** ***if*** *the copula* $C(u,v)$ *is linearly interpolating on* $\overline{\mathsf{R}}_X^c \times \overline{\mathsf{R}}_\theta$ *with respect to* u. *Here, the function* g *is defined as*

$$g(u,w) \equiv F_\theta{}^{1} \circ g_u(w), \tag{15}$$

where $g_u(w)$ *is the inverse function of* $v \longmapsto \mathcal{D}_1 C(u,v)$, *and* $\mathcal{D}_1 C(u,v)$ *is defined by* (EC.30) *in Online Appendix EC.1.3.1. Other notation is defined as in Theorem* 1.

Theorem 2 gives a necessary and suf cient condition for the representation—the copula, $C(u,v)$, must be linearly interpolating on $\overline{\mathsf{R}}_X^c \times \overline{\mathsf{R}}_\theta$ with respect to u. Proposition EC.9 in Online Appendix EC.1.3.1 provides details of the existence and uniqueness of this linearly interpolating copula, as well as how to construct it. Note that, when F_X is continuous, we have $\overline{\mathsf{R}}_X = [0,1]$ and $\overline{\mathsf{R}}_X^c = \emptyset$. Therefore, Theorem 1 can be viewed as a special case of Theorem 2 where no interpolation is needed for the copula on $\overline{\mathsf{R}}_X^c \times \overline{\mathsf{R}}_\theta$.

3.1.3. Implications. Representation Theorems 1 and 2 have several important implications.

First, they demonstrate that the distribution of induced order statistics, $\boldsymbol{\theta}_{[X]}$, can always be represented as a mixture of order statistics, $U_{i:N}$, and IID uniformly distributed random variables, V_i. In addition, the mixture function, g, is explicitly given by (15). This not only provides insight into the source of the randomness of $\boldsymbol{\theta}_{[X]}$ but also allows us to analyze the distribution of induced order statistics using the many established tools for order statistics.

Second, the representation theorems allow us to calculate the moments of $\boldsymbol{\theta}_{[X]}$, which is required for constructing optimal impact portfolios. By specifying the dependence structure between $\boldsymbol{X}$ and $\boldsymbol{\theta}$, along with their marginal distributions, we can construct a linearly interpolating copula and then use Theorem 2 to ef ciently compute the moments of $\boldsymbol{\theta}_{[X]}$ through numerical integration. In particular, the rst two moments of $\boldsymbol{\theta}_{[X]}$ are given by the following proposition, which shows that the computational complexity of computing the rst two moments of one asset (pair) does not scale with the number of assets N:

Downloaded from informs.org by [5.78.95.66] on 20 March 2024, at 07:31 . For personal use only, all rights reserved.

Proposition 2. *Under the assumptions of Theorem 2, if the copula $C(u, v)$ is linearly interpolating on $\overline{\mathrm{R}}_X^c \times \overline{\mathrm{R}}_\theta$ with respect to u, we have*

$$\mathbb{E}(\theta_{[i:N]}) = \int_0^1 \int_0^1 g(u,w) \cdot \frac{N! u^{i-1}(1-u)^{N-i}}{(i-1)!(N-i)!} \mathrm{d}u\mathrm{d}w,$$

$$\mathbb{E}(\theta_{[i:N]}^2) = \int_0^1 \int_0^1 [g(u,w)]^2 \cdot \frac{N! u^{i-1}(1-u)^{N-i}}{(i-1)!(N-i)!} \mathrm{d}u\mathrm{d}w,$$

$$\mathbb{E}(\theta_{[i:N]}\theta_{[j:N]}) = \int_0^u \int_0^1 \int_0^u \int_0^1 g(u,w) \cdot g(z,s) \cdot \frac{N! u^{i-1}(z-u)^{j-i-1}(1-z)^{N-j}}{(i-1)!(j-i-1)!(N-j)!} \mathrm{d}z\mathrm{d}u\mathrm{d}w\mathrm{d}s,$$

for $i, j = 1, 2, \ldots, N$ and $i < j$, where function g is defined by (15).

Third, the representation theorems have a broad range of potential practical applications because of their ability to characterize the distribution of $\boldsymbol{\theta}_{[X]}$ under any bivariate distribution between the impact factors and the residual returns. For example, we allow the impact factors, $\boldsymbol{X}$, to take discrete rating values.[16] The residual returns, $\boldsymbol{\theta}$, can also be discontinuous, which can describe prices with jumps, daily price limits, and high-frequency data with discrete price levels.[17] We also allow nondifferentiable dependence structures, C, between $\boldsymbol{X}$ and $\boldsymbol{\theta}$.[18]

Fourth, the representation theorems assert that the distribution of $\boldsymbol{\theta}_{[X]}$ is typically determined by the dependence structure, that is, the copula C, and the marginal distribution of residual returns, $\boldsymbol{\theta}$, but *not* the marginal distribution of the impact factors, $\boldsymbol{X}$. This is because the mixture function, g, given by (15), depends solely on C and F_θ. The specific form of F_X does not affect the representation, and only the discontinuity points of F_X are necessary to determine the interpolation of C (see Definition 1 in Online Appendix EC.1.3). This implies that impact investors who follow Proposition 1 to construct optimal impact portfolios should only focus on the *rank* of impact factors, rather than their specific distribution. It is well known that considerable noise exists in impact measurements such as the ESG score (Berg et al. 2022). The fact that portfolio weights depend only on the *rank* suggests that our methodology is more robust against noise and outliers in impact measurements.

We generalize the representation theorems to allow for cross-sectional heterogeneity in Online Appendix EC.1.3.2 and highlight the similarities and differences in terms of the implications above. We also use the generalized representation theorem to model markets with a mixture of heterogeneous subgroups and show how that affects the optimal impact portfolio.

3.2. Asymptotic Version

Theorems 1 and 2 characterize the exact finite-sample distribution of $\boldsymbol{\theta}_{[X]}$ for N assets. In practice, the universe of investable assets can be very large, in which case we can approximate the distribution of $\boldsymbol{\theta}_{[X]}$ as N increases without bound. As in Section 3.1, we first present the results for a smooth F and then the general version.

Theorem 3. *Under Assumption 1, assume that both F_X and F_θ are continuous functions, C is a copula of F with a density, and $\frac{\partial C}{\partial u}(u, v)$ is a continuous function of u. For any fixed $m \geq 1$, any constants $\xi_1 < \xi_2 < \cdots < \xi_m$, and any sequence $1 < i_1(N) < i_2(N) < \cdots < i_m(N) < N$ such that $i_k(N)/N \to \xi_k \in (0,1)$ for $k = 1, 2, \ldots, m$ as $N \to +\infty$, we have*

$$(\theta_{[i_1(N):N]}, \theta_{[i_2(N):N]}, \ldots, \theta_{[i_m(N):N]}) \xrightarrow{d} (g(\xi_1, V_1), g(\xi_2, V_2), \ldots, g(\xi_m, V_m)); \tag{16}$$

Here, $\xrightarrow{d}$ denotes convergence in distribution, and random variables V_i and function g are defined as in Theorem 1.

Theorem 3 is the asymptotic counterpart of Theorem 1. We follow both the theory of order statistics (David and Nagaraja 2003) and that of impact investing (Lo and Zhang 2023) to consider asymptotic convergence where the number of investable assets, N, grows, and assets $i_1(N), i_2(N), \ldots, i_m(N)$ converge to assets ranking at $100\xi_1, 100\xi_2, \ldots, 100\xi_m$ percentiles among all assets in the universe. Practitioners who invest in a large number of assets or target assets ranking at specific percentiles can use this theorem to guide their impact investing strategies.

The following result provides the asymptotic version of the representation theorem for general F.

Theorem 4. *Under Assumption 1, let C be a copula of F and $\mathcal{D}_1 C(u, v)$ be a continuous function of u. We have*

$$(\theta_{[i_1(N):N]}, \theta_{[i_2(N):N]}, \ldots, \theta_{[i_m(N):N]}) \xrightarrow{d} (g(\xi_1, V_1), g(\xi_2, V_2), \ldots, g(\xi_m, V_m)) \tag{17}$$

*for any fixed $m \geq 1$, any constants $\xi_1 < \xi_2 < \cdots < \xi_m$, and any sequence $1 < i_1(N) < i_2(N) < \cdots < i_m(N) < N$ such that $i_k(N)/N \to \xi_k \in (0,1)$ for $k = 1, 2, \ldots, m$ as $N \to +\infty$, **if and only if** the copula $C(u, v)$ is linearly interpolating on $\overline{\mathrm{R}}_X^c \times \overline{\mathrm{R}}_\theta$ with respect to u. Here, "$\xrightarrow{d}$" denotes convergence in distribution, and random variables V_i and function g are defined as in Theorem 2.*

Compared with Theorems 1 and 2, the asymptotic version has the following implications.

First, similar to Theorem 2, the linearly interpolating copula is necessary and sufficient for the asymptotic representation. One can easily construct the linearly interpolating copula and calculate the moments of the asymptotic distribution numerically using Theorem 4.

Downloaded from informs.org by [5.78.95.66] on 20 March 2024, at 07:31 . For personal use only, all rights reserved.

Second, unlike the exact finite-sample version, the asymptotic version reveals that the asymptotic distribution of induced order statistics is no longer a mixture of order statistics and random noise. Instead, it becomes a mixture of asset quantiles, ξ_i, and uniform noise, V_i. The mixture function, g, remains the same as in the finite-sample version and is determined by the marginal distribution of $\boldsymbol{\theta}$ and the dependence structure, C.

Third, because ξ_i in (17) are constants, the randomness of the asymptotic distribution of $(\theta_{[i_1(N):N]}, \theta_{[i_2(N):N]}, \ldots, \theta_{[i_m(N):N]})$ arises solely from V_i. Therefore, the induced order statistics $\theta_{[i_1(N):N]}, \theta_{[i_2(N):N]}, \ldots, \theta_{[i_m(N):N]}$ are asymptotically mutually independent. This highlights an advantage of our framework compared with the standard mean-variance portfolio construction—the estimation of the inverse of Σ is more stable when the number of assets N is large. Proposition 1 shows that the optimal weight depends on Σ^{-1}, which can be estimated stably in our framework because the asymptotic independence of the induced order statistics implies that Σ is approximately diagonal.

Finally, as with the finite-sample version, the asymptotic version also shows that the asymptotic distribution of $\boldsymbol{\theta}_{[X]}$ is typically determined by C and F_θ, but not F_X. This provides a road map for investigating the distribution of induced order statistics, and therefore the optimal impact portfolios. In Sections 4 and 5, we study how different copulas and marginal distributions of $\boldsymbol{\theta}$ affect the optimal impact portfolio weights and their corresponding returns, respectively. In Online Appendix EC.1.2, we discuss the special case where $(X_i, \theta_i)^\top$ is jointly normally distributed. Most of the results in this article are derived based on the powerful representation theorems, Theorems 1 and 4.

4. General Dependence via Copulas

In this section, we consider general dependence structures between $\boldsymbol{X}$ and $\boldsymbol{\theta}$. In particular, based on our representation theorems, we demonstrate how the moments of $\boldsymbol{\theta}_{[X]}$ depend on the characteristics of the copula. We focus on specific copulas, including the Gaussian and Archimedean copula (which encompasses the Clayton copula and Gumbel copula), to model tail dependence. The fundamental copulas studied by Yang et al. (2006) and the elliptical copula are discussed in Online Appendices EC.1.4.3 and EC.1.4.4, respectively.

4.1. Gaussian Copula

The Gaussian copula, which is constructed from multivariate normal distributions, is one of the most widely used copulas in the literature (Cont and Kan 2011).

Definition 2 (Gaussian Copula). The bivariate Gaussian copula with parameter $\rho \in (-1, 1)$ is defined as

$$C_\rho^{\mathrm{Ga}}(u, v) \equiv \boldsymbol{\Phi}_\rho(\Phi^{-1}(u), \Phi^{-1}(v)), \quad u, v \in [0, 1],$$

where Φ is the distribution function of $\mathsf{N}(0, 1)$, and $\boldsymbol{\Phi}_\rho$ is the bivariate distribution function of a bivariate normal distribution with mean 0, variance 1, and correlation ρ.

The parameter ρ captures the degree of dependence between the impact factor and the residual return.

The following result characterizes the distribution of $\boldsymbol{\theta}_{[X]}$ under the Gaussian copula using Theorems 1 and 2.

Proposition 3. *Under Assumption 1, if F_X is continuous and the copula of F is a Gaussian copula with parameter $\rho \in (-1, 1)$, we have*

$$\begin{aligned}&(\theta_{[1:N]}, \ldots, \theta_{[N:N]})\\ &\overset{d}{=} F_\theta^{-1} \circ \Phi\left(\rho Y_{1:N} + \sqrt{1-\rho^2}Z_1\right), \ldots,\\ &F_\theta^{-1} \circ \Phi\left(\rho Y_{N:N} + \sqrt{1-\rho^2}Z_N\right), \end{aligned} \tag{18}$$

where $Z_1, Z_2, \ldots, Z_N \overset{\mathrm{IID}}{\sim} \mathsf{N}(0, 1)$, $Y_{1:N} \le Y_{2:N} \le \cdots \le Y_{N:N}$ are the order statistics of $Y_1, Y_2, \ldots, Y_N \overset{\mathrm{IID}}{\sim} \mathsf{N}(0, 1)$, and random variables $\{Y_i\}_{i=1}^N$ and $\{Z_i\}_{i=1}^N$ are mutually independent.

Proposition 3 shows that, under the Gaussian copula, the distribution of $\theta_{[i:N]}$ is determined by $\rho Y_{i:N} + \sqrt{1-\rho^2}Z_i$, which is a *linear* mixture of order statistics of standard normal random variables and IID standard normal noise. Proposition 3 also implies that the distribution of $\theta_{[i:N]}$ is linked to $\rho Y_{i:N} + \sqrt{1-\rho^2}Z_i$ via a special function:

$$Q(x) \equiv F_\theta^{-1} \circ \Phi(x), \quad x \in \mathbb{R}; \tag{19}$$

In fact, $Q(x)$ is precisely the quantile-quantile plot (Q-Q plot) of F_θ versus the standard normal distribution, a widely used statistical tool. Financial practitioners often use the Q-Q plot to visualize the heaviness of the tails of random variables. In particular, if $\boldsymbol{\theta}$ is normally distributed with mean μ_θ and variance σ_θ^2, we have $F_\theta(y) = \Phi((y - \mu_\theta)/\sigma_\theta)$, and $Q(x) = F_\theta^{-1} \circ \Phi(x) = \mu_\theta + \sigma_\theta x$, which is a linear Q-Q plot. In this case, it is straightforward to verify that Proposition 3 reduces to the case of a bivariate normal distribution (see Proposition EC.4 in Online Appendix EC.1.2).

The following result gives the asymptotic distribution of $\boldsymbol{\theta}_{[X]}$ under the Gaussian copula, which is a corollary of the asymptotic version of Theorems 3 and 4.

Proposition 4. *Under the assumptions of Proposition 3, for any fixed $m \ge 1$, any constants $\xi_1 < \xi_2 < \cdots < \xi_m$, and any sequence $1 < i_1(N) < i_2(N) < \cdots < i_m(N) < N$ such that $i_k(N)/N \to \xi_k \in (0, 1)$ for $k = 1, 2, \ldots, m$ as $N \to +\infty$, we have*

$$\begin{aligned}&(\theta_{[i_1(N):N]}, \ldots, \theta_{[i_m(N):N]})\\ &\xrightarrow{d} F_\theta^{-1} \circ \Phi\left(\rho\Phi^{-1}(\xi_1) + \sqrt{1-\rho^2}Z_1\right), \ldots,\\ &F_\theta^{-1} \circ \Phi\left(\rho\Phi^{-1}(\xi_m) + \sqrt{1-\rho^2}Z_m\right), \end{aligned}$$

where the notation follows that in Proposition 3.

Downloaded from informs.org by [5.78.95.66] on 20 March 2024, at 07:31 . For personal use only, all rights reserved.

By combining Proposition 4 and the optimal weights given by (9), the optimal weight for assets ranked at quantile ξ under the Gaussian copula can be approximated by

$$w^{\mathrm{Ga}}(\xi) \equiv \frac{\mathbb{E}\left[F_\theta^{-1} \circ \Phi\left(\rho\Phi^{-1}(\xi) + \sqrt{1-\rho^2}Z\right)\right]}{\mathrm{Var}\left[F_\theta^{-1} \circ \Phi\left(\rho\Phi^{-1}(\xi) + \sqrt{1-\rho^2}Z\right)\right]}, \quad \xi \in (0,1), \tag{20}$$

where Z is a standard normal random variable. There is no need to calculate the covariances among induced order statistics because $\theta_{[i_1(N):N]}, \theta_{[i_2(N):N]}, \ldots, \theta_{[i_m(N):N]}$ are asymptotically mutually independent. Investors can use (20), which can be easily computed numerically, to approximate the optimal weights when the total number of assets, N, is large.

4.2. Archimedean Copula

Although the Gaussian copula is widely used, it has been documented that the tail dependencies between asset returns and impact factors for both tails are asymmetric in many contexts applicable to impact investing (Bax et al. 2023), a feature that cannot be captured by the Gaussian copula. Therefore, in this section, we construct optimal portfolios under another important family of copulas—the Archimedean copula family—which has been proven to be useful for modeling tail dependence in financial data (McNeil et al. 2015).

Definition 3 (Archimedean Copula). The Archimedean copula with generator function ψ is defined as

$$C_\psi^{\mathrm{Ar}}(u,v) \equiv \phi(\psi(u) + \psi(v)), \quad u,v \in [0,1],$$

where the generator function $\psi : [0,1] \to [0,+\infty]$ is a continuous, strictly decreasing, and strictly convex function such that $\psi(1) = 0$, and its inverse $\phi \equiv \psi^{-1}$ is defined on $[0,+\infty]$.[19]

The dependence structure of an Archimedean copula is determined by its generator function, ψ, instead of the single parameter ρ of the Gaussian copula. This allows the Archimedean copula to capture a rich set of dependence structures.[20] The following proposition characterizes the distribution of $\boldsymbol{\theta}_{[X]}$ under the Archimedean copula using Theorems 1 and 2. The corresponding asymptotic result can also be derived using Theorems 3 and 4, which we omit because of space constraints.

Proposition 5. *Under Assumption* 1, *if F_X is continuous and the copula of F is an Archimedean copula with generator function ψ (let $\phi = \psi^{-1}$), and if the first- and second-order derivatives of ϕ, ϕ' and ϕ'', both exist on $[0,+\infty)$, we have*

$$\begin{pmatrix} \theta_{[1:N]} \\ \theta_{[2:N]} \\ \vdots \\ \theta_{[N:N]} \end{pmatrix} \overset{d}{=} \begin{pmatrix} F_\theta^{-1} \circ \phi\left[\phi'^{-1}\left[\phi' \circ \phi^{-1}(U_{1:N}) \cdot V_1\right]\right] \phi^{-1}(U_{1:N}) \\ F_\theta^{-1} \circ \phi\left[\phi'^{-1}\left[\phi' \circ \phi^{-1}(U_{2:N}) \cdot V_2\right]\right] \phi^{-1}(U_{2:N}) \\ \vdots \\ F_\theta^{-1} \circ \phi\left[\phi'^{-1}\left[\phi' \circ \phi^{-1}(U_{N:N}) \cdot V_N\right]\right] \phi^{-1}(U_{N:N}) \end{pmatrix},$$

where the notation is the same as in Theorem 1.

Given an Archimedean copula with any generator function ψ, we can use Proposition 5 to derive the distribution of $\boldsymbol{\theta}_{[X]}$ explicitly. Like the result for the Gaussian copula, Proposition 5 demonstrates that the distribution of $\boldsymbol{\theta}_{[X]}$ under the Archimedean copula is also a mixture of order statistics, $U_{i:N}$, and IID uniform noise, V_i, although the mixture is more complicated than the linear mixture for the Gaussian copula given by Proposition 3.

Archimedean copulas are widely used for characterizing various tail dependencies. To derive optimal impact portfolios when the dependence between the impact factors and residual returns is asymmetric in its tails, we consider two special cases of the Archimedean copula: the Clayton and Gumbel copulas. The Clayton copula characterizes lower tail dependence (i.e., small values of the residual returns, $\boldsymbol{\theta}$, generally appear together with small values of the impact factors, $\boldsymbol{X}$), and the Gumbel copula characterizes upper tail dependence (i.e., large values of $\boldsymbol{\theta}$ generally appear together with large values of $\boldsymbol{X}$).[21] One can expect that the two copulas yield opposite results. For brevity, we discuss the Clayton copula below and relegate the discussion of the Gumbel copula to Online Appendix EC.1.4.2.

4.3. Clayton Copula

The Clayton copula, which models lower tail dependence, is defined as follows:

Definition 4 (Clayton Copula). The Clayton copula with parameter $\gamma \in (0,+\infty)$ is defined as an Archimedean copula with generator function

$$\psi_\gamma^{\mathrm{Cl}}(u) \equiv \frac{u^{-\gamma} - 1}{\gamma}, \quad u \in [0,1]; \tag{21}$$

The Clayton copula parameterizes the Archimedean copula generated by (21), and the parameter γ measures the strength of dependence between $\boldsymbol{X}$ and $\boldsymbol{\theta}$.

By replacing the generator function of the Archimedean copula in Proposition 5 by $\psi_\gamma^{\mathrm{Cl}}$, we can derive the distribution of its induced order statistics, $\boldsymbol{\theta}_{[X]}$ (for brevity, we write the i-th entry only):

$$\theta_{[i:N]} \overset{d}{=} F_\theta^{-1}\left[\left(1 + U_{i:N}^{-\gamma} V_i^{-\gamma/(\gamma+1)} - U_{i:N}^{-\gamma}\right)^{-1/\gamma}\right]; \tag{22}$$

This result allows us to calculate the moments of $\boldsymbol{\theta}_{[X]}$ numerically. In addition, by taking the limits of (22) with respect to γ, we can show that the right-hand side of (22) converges to $F_\theta^{-1}(U_{i:N})$ as γ increases without bound, and converges to $F_\theta^{-1}(V_i)$ as γ approaches zero. Therefore, γ determines the relative importance of the order statistics, $U_{i:N}$, and the uniform noise, V_i, in the representation.

We use numerical examples to investigate the optimal impact portfolios under the Clayton copula using (22) and Proposition 2. Figure 1 displays the