Naive Diversification Preferences and their Representation

Enrico G. De Giorgi* **Ola Mahmoud†**

October 11, 2018‡

Abstract

A widely applied diversification paradigm is the *naive diversification* choice heuristic. It stipulates that an economic agent allocates equal decision weights to given choice alternatives independent of their individual characteristics. This article provides mathematically and economically sound choice theoretic foundations for the naive approach to diversification. We axiomatize naive diversification by defining it as a preference for equality over inequality, derive its relationship to the classical diversification paradigm, and provide a utility representation. In particular, we (i) prove that the notion of permutation invariance lies at the core of naive diversification and that an economic agent is a naive diversifier if and only if his preferences are convex and permutation invariant; (ii) derive necessary and sufficient conditions on the utility functions that give rise to preferences for naive diversification; (iii) show that naive diversification preferences arise when decision makers only consider beliefs that imply some weak form of independence, which is closely related to correlation neglect.

Keywords: naive diversification, convex preferences, permutation invariant preferences, exchangeability, inequality aversion, majorization, Dalton transfer, Lorenz order.

JEL Classification: C02, D81, G11.

---

*Department of Economics, School of Economics and Political Science, University of St. Gallen, Bodanstrasse 6, 9000 St. Gallen, Switzerland, Tel. +41 71 224 24 30, Fax. +41 71 224 28 94, email: enrico.degiorgi@unisg.ch.

†Faculty of Mathematics and Statistics, School of Economics and Political Science, University of St. Gallen, Bodanstrasse 6, 9000 St. Gallen, Switzerland and Center for Risk Management Research, University of California, Berkeley, Evans Hall, CA 94720-3880, USA, email: ola.mahmoud@unisg.ch

‡We are grateful to Simone Cerreia-Vioglio, Urs Fischbacher, Itzhak Gilboa, Lisa Goldberg, Georg Nöldeke, Marciano Siniscalchi, and two anonymous referees for providing valuable feedback on our work. We also thank numerous seminar and conference audiences for their comments, and the Basic Research Fund of the University of St. Gallen for financial support.

1 Introduction

Diversification is one of the cornerstones of decision making in economics and finance. In its essence, it conveys the idea of choosing variety over similarity. Informally, one might say that the goal behind introducing variety through diversification is the reduction of risk or uncertainty, and so one might identify a diversifying decision maker with a risk averse one. This is indeed the case in the expected utility theory (EUT) of von Neumann and Morgenstern (1944), where risk aversion and preference for diversification are exactly captured by the concavity of the utility function which the decision maker is maximizing. However, this equivalence fails to hold in more general models of choice, as shown by De Giorgi and Mahmoud (2016).

In the context of portfolio construction, standard economic theory postulates that an investor should optimize amongst various choice alternatives by maximizing portfolio return while minimizing portfolio risk, given by the return variance (Markowitz 1952). In practice, however, these traditional optimization approaches to choice are plagued by technical difficulties.[1] Experimental work in the decades after the emergence of the classical theories of von Neumann and Morgenstern (1944) and Markowitz (1952) has shown that economic agents in reality systematically violate the traditional diversification assumption when choosing among risky gambles. Indeed, seminal psychological and behavioral economics research by Tversky and Kahneman (1981) (see also Simon (1955) and Simon (1979)) suggests that the portfolio construction task may be too complex for decision makers to perform. Consequently, investors adopt various types of simplified diversification paradigms in practice.

One of the most widely applied such simple rules of choice is the so-called *naive diversification* heuristic. It stipulates that an economic agent allocates equal weights among a given choice set, independent of the individual characteristics of the underlying choice alternatives. In the context of portfolio construction, this rule is often referred to as the *equal-weighted* or $1/n$ strategy. This naive diversification paradigm goes as far back as the Talmud, where the relevant passage states that "*it is advisable for one that he should divide his money in three parts, one of which he shall invest in real estate, one of which in business, and the third part to remain always in his hands*" (Duchin and Levy 2009). It is documented that even Harry Markowitz used the simple $1/n$ heuristic when he made his own retirement investments. He justifies his choice on psychological grounds: "My intention was to minimize my future regret. So I split my contributions fifty-fifty between bonds and equities" (Gigerenzer 2010).

[1] These difficulties are stemming from the instability of the optimization problem with respect to the available data. As is the case with any economic model, the true parameters are unknown and need to be estimated, hence resulting in uncertainty and estimation error. For a discussion of the problems arising in implementing mean-variance optimal portfolios, see for example Hodges and Brealy (1978), Best and Grauer (1991), Michaud (1998), and Litterman (2003).

1.1 Towards choice-theoretic foundations

The word *naive* inherently implies a lack of sophistication. Indeed, naive diversification is widely viewed as an anomaly linked to irrational behavior that does not assure sensible or coherent decision making. In its essence, the naive diversification paradigm is considered a simple and practical rule of thumb with no economic foundation guaranteeing its optimality. Moreover, despite the large experimental and empirical evidence of the prevalence and outperformance of naive diversification, a formal descriptive choice theoretic or economic model does not seem to exist.

With the purpose of filling this gap, this paper provides a mathematically and economically sound choice theoretic formalization of the naive approach to diversification of decision makers and investors. To this end, we axiomatize naive diversification by framing it as a choice theoretic preference for equality over inequality, which has a utility representation, and derive its relationship to the classical diversification paradigm. The crux of our choice theoretic axiomatization of the naive diversification heuristic lies in the idea that equality is preferred over inequality, a concept that is simultaneously simple and complex, as put by Sen (1973): "*At one level, it is the simplest of all ideas and has moved people with an immediate appeal hardly matched by any other concept. At another level, however, it is an exceedingly complex notion which makes statements of inequality highly problematic, and it has been, therefore, the subject of much research by philosophers, statisticians, political theorists, sociologists and economists.*" We complement this line of research from a decision theoretic perspective by using the mathematical concept of *majorization* to describe a preference relation which exhibits *preference for naive diversification*. Historically, majorization has been used to describe inequality orderings in the economic context of inequality of income, as developed by both Lorenz (1905) and Dalton (1920).[2]

The goal of our choice-theoretic approach is threefold. First, our main objective is to develop an axiomatic system that precisely captures widely observed regularities of behavior. We thus provide a formal descriptive model of what is considered to be an anomalous yet strongly prevalent paradigm such as naive diversification. Second, this axiomatic descriptive model enables us to gain novel insights into the nature of the preferences and the utility of the naive diversifier. In particular, by relating it to other known axiomatized behavioral paradigms, we show that preferences for naive diversification are equivalent to convex preferences that additionally exhibit an indifference among the choice alternatives, which is formalized via a notion of permutation invariance. We also show preferences for naive diversification arise when naive diversifiers treat assets as being conditionally independent and identically distributed, which implies that they exhibit a level of correlation neglect.

Finally, one may use the axioms underlying naive diversification to test the behavioral drivers of this choice heuristic in reality. For example, one of our axioms, that of permuta-

[2] We refer the reader to Marshall, Olkin, and Arnold (2011) for a comprehensive self-contained account of the theory and applications of majorization.

tion invariance, implies that the given alternatives are considered in some way symmetric or equivalent by the naive decision maker. This is an axiom that can be directly tested in, say, an experimental setting by relating it to Laplace's principle of indifference and varying the amount of information available for each of the choice alternatives.

1.2 Synopsis

The remainder of the paper is structured as follows. Section 2 discusses some principles related to naive diversification and provides an overview of the evidence of both naive diversification and correlation neglect in the real world. Section 3 sets up the choice theoretic framework in the Anscombe-Aumann setting and provides the necessary background on majorization and doubly stochastic matrices, both of which are fundamental concepts in our development. Section 4 presents an axiomatic formalization of naive diversification preferences and derives its relationship to the traditional (convex) diversification axiom. We then show that the notion of permutation invariance lies at the core of our definition and that a preference relation exhibits preference for naive diversification if and only if it is convex and permutation invariant. In Section 5, we provide necessary and sufficient conditions on the utility functions that give rise to preferences for naive diversification. Section A considers two potentially useful applications of our formalism, namely comparison of levels of naive diversification and rebalancing of allocation to equality.

2 Background

2.1 Related principles

Naive diversification implies a preference of equality over inequality in the choice weights. One of the earliest, closely related hypotheses concerning decisions under subjective uncertainty is the *principle of insufficient reason*, also called the *principle of indifference*. It is generally attributed to Bernoulli (1738) and invoked by Bayes (1763) in his development of the binomial theorem. The principle states that in situations where there is no logical or empirical reason to favor any one of a set of mutually exclusive events or choices over any other, one should assign them all equal probability. In Bayesian probability, this is the simplest non-informative prior.

Outside the choice theoretic framework, the notion of preference of equality over inequality dominates several prominent problems in economic theory. Early in the twentieth century, economists became interested in measuring inequality of incomes or wealth. More specifically, it became desirable to determine how income or wealth distributions might be compared in order to say that one distribution was more equal than another. The first discussion of this kind was provided by Lorenz (1905). He suggested a graphical manner in which to compare inequality in finite populations in terms of nested curves. If total wealth is uniformly distributed, the so-called *Lorenz curve* is a straight line. With an unequal distribution, the curves will always begin and end in the same points as with an equal distribution, but they will be bent in the middle. The rule of interpretation, as he puts it, is: as the bow is bent, concentration increases. Later, Dalton (1920) described the closely related *principle of transfers*. Under the theoretical proposition of a positive functional relationship between income and economic welfare, stating that economic welfare increases at an exponentially decreasing rate with increased income, Dalton concludes that maximum social welfare is achievable only when all incomes are equal. Following a suggestion by Pigou (1912), he proposed the condition that a transfer of income from a richer to a poorer person, so long as that transfer does not reverse the ranking of the two, will result in greater equity. Such an operation, involving the shifting of wealth from one individual to a relatively poorer individual, is known as the *Pigou-Dalton transfer* and has also been labeled as a *Robin Hood transfer*. The seminal ideas of Lorenz (1905) and Dalton (1920) will be referenced frequently throughout our development of naive diversification preferences, as the mathematical framework upon which we rely coincides with theoretical formalizations of the Lorenz curve and the Dalton transfer.

2.2 Experimental and empirical evidence of naive diversification

Academics and practitioners have long studied the occurrence of naive diversification, along with its downside and potential benefits. Some of the first academic demonstrations of naive diversification as a choice heuristic were made by Simonson (1990) in marketing in the context of consumption decisions by individuals, and by Read and Loewenstein (1995) in the context of experimental psychology. In the context of economic and financial decision making, empirical evidence suggests behavior which is consistent with naive diversification. For instance, Benartzi and Thaler (2001) turned to study whether the effect manifests itself among investors making decisions in the context of defined contribution saving plans. Their experimental evidence suggests that some people spread their contributions evenly across the investment options irrespective of the particular mix of options. The authors point out that while naive diversification can produce a “reasonable portfolio”, it affects the resulting asset allocation and can be costly. In particular, people might choose a portfolio that is not on the efficient frontier, or they might pick the wrong point along the frontier. Moreover, it does not assure sensible or coherent decision making. Subsequently, Huberman and Jiang (2006) find that participants tend to invest in only a small number of the funds offered to them, and that they tend to allocate their contributions evenly across the funds that they use, with this tendency weakening with the number of funds used. More recently, Baltussen and Post (2011) find strong evidence for what they coin as *irrational* behavior. Their subjects follow a conditional naive diversification heuristic as they exclude the assets with an unattractive marginal distribution and divide the available funds equally between the remaining, attractive assets. This strategy is applied even if it leads to allocations that are dominated in terms of first-order stochastic dominance – hence the term irrational.

Irrationality has been since then frequently used to describe naive diversification behavior. In Fernandes (2013), the naive diversification bias of Benartzi and Thaler (2001) was replicated across different samples using a within-participant manipulation of portfolio options. It was found that the more investors use intuitive judgments, the more likely they are to display the naive diversification bias.

In the context of portfolio construction, naive diversification has enjoyed a revival during the last few years because of its simplicity on one hand and the empirical evidence on the other hand suggesting superior performance compared to traditional diversification schemes. In addition to the relative outperformance, the empirical stability of the naive $1/n$ diversification rule has made it particularly attractive in practice, as — unlike Markowitz's risk minimization strategies — it does not rely on unknown correlation parameters that need to be estimated from data. Moreover, its outperformance has been investigated and a range of reasons have been proposed for why naive diversification may outperform other diversification paradigms. The most widely documented of these is the so-called small-cap-effect within the universe of equities. This theory stipulates that stocks with smaller market capitalization tend to ourperform larger stocks, and by construction, naive diversification gives more exposure to smaller cap stocks compared to capitalization weighting. Empirical support for the superior performance of equal weighted portfolios relative to capitalization weighting include Lessard (1976), Roll (1981), Ohlson and Rosenberg (1982), Breen, Glosten, and Jagannathan (1989), Grinblatt and Titman (1989), Korajczyk and Sadka (2004), Hamza, Kortas, L'Her, and Roberge (2007) and Pae and Sabbaghi (2010). Furthermore, DeMiguel, Garlappi, and Uppal (2007) show the strong performance relative to optimized portfolios. Duchin and Levy (2009) provide a comparison of naive and Markowitz diversification and show that an equally weighted portfolio may often be substantially closer to the true mean variance optimality than an optimized portfolio. On the other hand, Tu and Zhou (2011) propose a combination of naive and sophisticated strategies, including Markowitz optimization, as a way to improve performance, and conclude that the combined rules not only have a significant impact in improving the sophisticated strategies, but also outperform the naive rule in most scenarios.

2.3 Correlation neglect

Typically, financial decision makers are faced with not only an analysis of risk and return profiles of their assets, but also the correlations across different asset returns. It can however be a challenging task to work with joint distributions of multiple random variables. Even though a decision maker could in principle adequately analyze the choice variables' co-movement, he may fail to account for correlation in the decision making process.

Correlation neglect is a cognitive bias by which individuals treat choice options as if they are independent. This phenomenon has been recently explored in different contexts in the behavioral economics and bounded rationality literature. It was first documented experimentally by Kroll, Levy, and Rapoport (1988), whose experiment participants were asked to allocate an endowment between assets, where only the correlation between assets was varied between participants (from -0.8 to 0.8). They found that allocation was not affected by the treatment. Ortoleva and Snowberg (2015) analyze the effect of correlation neglect on the polarisation of beliefs. DeMarzo, Vayanos, and Zwiebel (2003) study how it affects the diffusion of information in social networks. Glaeser and Sunstein (2009) and Levy and Razin (2015) explore the implications for group decision making in political applications. Recent experimental evidence in Eyster and Weizsäcker (2011) shows how correlation neglect biases choices in an investment portfolio decision problem. Moreover, the experiment of Kallir and Sonsino (2009) found that subjects neglect correlations in their allocation decisions, even if it could be shown that they generally noticed the structure of or the changes in co-movement.

In Section 4 we derive a general result that formalizes the link between preference for naive diversification and correlation neglect.

3 Theoretical setup

3.1 Preference relation

We adopt the generalized Anscombe-Aumann choice theoretic setup presented by Cerreia-Vioglio, Maccheroni, Marinacci, and Montrucchio (2011), where S is a set of states of the world, Σ is an algebra of subsets of S and X is the set of consequences, which is assumed to be a convex subset of a vector space, such as the set of lotteries on a set of prizes. We denote by $\mathcal{F}$ the set of simple acts, i.e., functions $f : S \to X$ that are Σ-measurable and with finitely many values. As usual, we identify X with the set of constant acts in $\mathcal{F}$, i.e., $x \in X$ is identified with the constant act x such that $x(s) = x$ for all $s \in S$. Moreover, for $\alpha \in [0,1]$ and $f, g \in \mathcal{F}$, the act $\alpha\, f + (1-\alpha)\, g$ is defined by $(\alpha\, f + (1-\alpha)\, g)(s) = \alpha\, f(s) + (1-\alpha)\, g(s)$ for all $s \in S$.

The decision maker's preferences on $\mathcal{F}$ are modeled by a binary relation $\succsim$, which induces an indifference relation $\sim$ on $\mathcal{F}$ defined by $f \sim g \Leftrightarrow (f \succsim g) \wedge (g \succsim f)$ and a strict preference relation $\succ$ on $\mathcal{F}$ defined by $f \succ g \Leftrightarrow f \succsim g \wedge \neg(f \sim g)$. The preference relation $\succsim$ is a weak order, i.e., satisfies the following properties:

(i) *Non-triviality*: $f, g \in \mathcal{F}$ exist such that $f \succ g$.

(ii) *Completeness*: For all $f, g \in \mathcal{F}$, $f \succsim g \vee g \succsim f$.

(iii) *Transitivity*: For all $f, g, h \in \mathcal{F}$, $f \succsim g \wedge g \succsim h \Rightarrow f \succsim h$.

Moreover, emulating the majority of frameworks of economic theory, we assume that the preference relation $\succsim$ is *monotone*.

(iv) *Monotonicity*: For all $f, g \in \mathcal{F}$ with $f(s) \succsim g(s)$ for all $s \in S$ we have $f \succsim g$.

Finally, we impose the following two standard additional assumptions:

(v) *Risk independence*: For $x, y, z \in X$ and $\alpha \in (0,1)$, $x \sim y \Longrightarrow \alpha\, x + (1-\alpha)\, z \sim \alpha\, y + (1-\alpha)\, z$.

(vi) *Continuity*: For $f, g, h \in \mathcal{F}$, the sets $\{\alpha \in [0,1] \,:\, \alpha\, f + (1-\alpha)\, g \succsim h\}$ and $\{\alpha \in [0,1] \,:\, h \succsim \alpha\, f + (1-\alpha)\, g\}$ are closed.

In the remainder of the paper a preference relation $\succsim$ is assumed to satisfy properties (i)-(vi). It is well-known (Herstein and Milnor 1953, Fishburn 1970) that properties (i)-(vi) imply the existence of a non-constant affine function $u : X \to \mathbb{R}$ such that

$$x \succsim y \Longleftrightarrow u(x) \geq u(y).$$

Note that for $f \in \mathcal{F}$ and u as above, $u(f)$ is an element of the set $B_0(\Sigma)$ of real-valued Σ-measurable simple functions. The dual space of $B_0(\Sigma)$ is the set $ba(\Sigma)$ of all bounded finitely additive measures on (S, Σ) and Δ denotes the set of all probabilities in $ba(\Sigma)$.

3.2 Choice weights and majorization

We use the theory of majorization from linear algebra to measure the variability of weights when diversifying across a set of n possible choices. Majorization, which was formally introduced by Hardy, Littlewood, and Pólya (1934), captures the idea that the components of a weight vector $\boldsymbol{\alpha} \in \mathbb{R}^n$ are less spread out or more nearly equal than the components of a vector $\boldsymbol{\beta} \in \mathbb{R}^n$. For any $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n) \in \mathbb{R}^n$, let

$$\alpha_{(1)} \geq \cdots \geq \alpha_{(n)}$$

denote the components of $\boldsymbol{\alpha}$ in *decreasing* order, and let

$$\boldsymbol{\alpha}_{\downarrow} = (\alpha_{(1)}, \ldots, \alpha_{(n)})$$

denote the decreasing rearrangement of $\boldsymbol{\alpha}$. The weight vector with i-th component equal to 1 and all other components equal to 0 is denoted by $\boldsymbol{e}_i$, and the vector with all components equal to 1 is denoted by $\boldsymbol{e}$. We restrict our attention to non-negative weights which sum to one, that is, $\boldsymbol{\alpha} \in \mathbb{S}^n = \left\{\boldsymbol{v} = (v_1, \ldots, v_n) \in \mathbb{R}^n_+ \mid \sum_{i=1}^n v_i = 1\right\}$. This means that the decision maker is assumed to use his full capital and is not taking "inverse" positions such as shorting in financial economics. Moreover, we will sometimes refers to the set

$$\mathbb{S}^n_{\downarrow} = \left\{\boldsymbol{v}_{\downarrow} = (v_{(1)}, \ldots, v_{(n)}) \in \mathbb{R}^n_+ \mid \sum_{i=1}^n v_{(i)} = 1\right\}.$$

We now define the notion of majorization:

Definition 1 (Majorization)**.** *For* $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n) \in \mathbb{R}^n$ *and* $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_n) \in \mathbb{R}^n$*,* $\boldsymbol{\beta}$ *is said to* (weakly) majorize $\boldsymbol{\alpha}$ *(or, equivalently,* $\boldsymbol{\alpha}$ *is* majorized by $\boldsymbol{\beta}$*), denoted by* $\boldsymbol{\beta} \geq_m \boldsymbol{\alpha}$*, if*

$$\sum_{i=1}^{n} \alpha_i = \sum_{i=1}^{n} \beta_i$$

and for all $k = 1, \ldots, n-1$*,*

$$\sum_{i=1}^{k} \alpha_{(i)} \leq \sum_{i=1}^{k} \beta_{(i)} \ .$$

Majorization is a preorder on the weight vectors in $\mathbb{S}^n$ and a partial order on $\mathbb{S}^n_{\downarrow}$. It is trivial but important to note that all vectors in $\mathbb{S}^n_{\downarrow}$ majorize the uniform vector $\boldsymbol{u}_n = (\frac{1}{n}, \ldots, \frac{1}{n})$, since the uniform vector is the vector with minimal differences between its components.

A key mathematical result in the study of majorization and inequality measurement is a theorem due to Hardy, Littlewood, and Pólya (1929). It roughly states that a vector $\boldsymbol{\alpha}$ is majorized by a vector $\boldsymbol{\beta}$ if and only if $\boldsymbol{\alpha}$ is an averaging of $\boldsymbol{\beta}$. This "averaging" operation is formalized via doubly stochastic matrices.[3] A square matrix P is said to be stochastic if its elements are all non-negative and all rows sum to one. If, in addition to being stochastic, all columns sum to one, the matrix is said to be doubly stochastic. A formal definition follows.

Definition 2 (Doubly stochastic matrix)**.** *An* $n \times n$ *matrix* $P = (p_{ij})$ *is* doubly stochastic *if* $p_{ij} \geq 0$ *for* $i, j = 1, \ldots, n$*,* $\boldsymbol{e}P = \boldsymbol{e}$ *and* $P\boldsymbol{e}' = \boldsymbol{e}'$*. We denote by* $\mathbb{D}_n$ *the set of* $n \times n$ *doubly stochastic matrices.*

Theorem 1 (Hardy, Littlewood, and Pólya (1929))**.** *For* $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{R}^n$*,* $\boldsymbol{\alpha}$ *is majorized by* $\boldsymbol{\beta}$ *if and only if* $\boldsymbol{\alpha} = \boldsymbol{\beta}P$ *for some doubly stochastic matrix* P*.*[4]

An obvious example of a doubly stochastic matrix is the $n \times n$ matrix in which every entry is $1/n$, which we shall denote by P_n. Other simple examples are given by the $n \times n$ identity matrix I_n and by permutation matrices: a square matrix Π is said to be a permutation matrix if each row and column has a single unit entry with all other entries being zero. There are $n!$ such matrices of size $n \times n$ each of which is obtained by interchanging rows or columns of the identity matrix. The set $\mathbb{D}_n$ of doubly stochastic matrices is convex and permutation matrices constitute its extreme points.

Use of a special type of doubly stochastic matrix, the so-called T-transform, will be made in this paper.

[3] A note on terminology: the term "stochastic matrix" goes back to the large role that they play in the theory of discrete Markov chains. Doubly stochastic matrices are also sometimes called "Schur transformations" or "bistochastic".

[4] We refer the reader to Schmeidler (1979) for several economic interpretations of Theorem 1, including decisions under uncertainty and welfare economics.

Definition 3 (T-transform). *A (elementary) T*-transform *is a matrix that has the form $T = \lambda I + (1-\lambda)\Pi$, where $\lambda \in [0,1]$ and Π is a permutation matrix that interchanges exactly two coordinates. For $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n) \in \mathbb{S}^n$, $\boldsymbol{\alpha} T$ thus has the form*

$$\boldsymbol{\alpha} T = (\alpha_1, \ldots, \alpha_{j-1}, \lambda\alpha_j + (1-\lambda)\alpha_k, \alpha_{j+1}, \ldots, \alpha_{k-1}, \lambda\alpha_k + (1-\lambda)\alpha_j, \alpha_{k+1}, \ldots, \alpha_n) \ ,$$

where we assume that the j-th and k-th coordinates of $\boldsymbol{\alpha}$ are averaged.

The importance of T-transforms can be seen from the following result, which is essential in the proof of Theorem 1 and which we shall utilize in some of the proofs of this article.

Proposition 1 (Muirhead (1903); Hardy, Littlewood, and Pólya (1934)). *If $\boldsymbol{\alpha} \in \mathbb{R}^n$ is majorized by $\boldsymbol{\beta} \in \mathbb{R}^n$, then $\boldsymbol{\alpha}$ can be derived from $\boldsymbol{\beta}$ by successive applications of a finite number of T-transforms.*

4 Naive diversification preferences

4.1 Classical diversification

An economic agent who chooses to diversify is traditionally understood to prefer variety over similarity. Axiomatically, preference for diversification is formalized as follows; see Dekel (1989).

Definition 4 (Preference for diversification). *A preference relation $\succsim$ exhibits* preference for diversification *if for any $f_1, \ldots, f_n \in \mathcal{F}$ and $\alpha_1, \ldots, \alpha_n \in [0,1]$ for which $\sum_{i=1}^n \alpha_i = 1$,*

$$f_1 \sim \cdots \sim f_n \Longrightarrow \sum_{i=1}^n \alpha_i \, f_i \succsim f_j \quad \textit{for all } j = 1, \ldots, n.$$

This definition states that an individual will want to diversify among a collection of choices all of which are ranked equivalently. The most common example of such diversification is within the universe of asset markets, where an investor faces a choice amongst risky assets.

The related notion of convexity of preferences inherently relates to the classic ideal of diversification, as introduced by Bernoulli (1738). By combining two choices, the decision maker is ensured under convexity that he is never "worse off" than the least preferred of these two choices.

Definition 5 (Convex preferences). *A preference relation $\succsim$ on $\mathcal{F}$ is* convex *if for all $f, g \in \mathcal{F}$ and $\alpha \in (0,1)$,*

$$f \sim g \Longrightarrow \alpha \, f + (1-\alpha) \, g \succsim f.$$

Indeed, a preference relation is convex if and only if it exhibits preference for diversification. Therefore, preference relations that exhibit preference for diversification coincide with uncertainty averse preferences, as pointed out by Schmeidler (1989). Moreover, it is well-known that a preference relation that is represented by a concave utility function is convex, and that a preference relation is convex if and only if its utility representation is quasi-concave. Variations on this classical definition of diversification exist in the literature (see, for example, Chateauneuf and Tallon (2002) and Chateauneuf and Lakhnati (2007)). We refer to De Giorgi and Mahmoud (2016) for a recent analysis of the classical definitions of diversification in the theory of choice.

4.2 Naive diversification

We now present an axiomatic formalization of the notion of naive diversification in terms of preference of equal decision weights over unequal decision weights.

Definition 6 (Preference for naive diversification). *A preference relation $\succsim$ exhibits* preference for naive diversification *if for $n \in \mathbb{N}$, and $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n) \in \mathbb{S}^n$, $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_n) \in \mathbb{S}^n$ it follows that:*

$$\boldsymbol{\alpha} \leq_m \boldsymbol{\beta} \quad \Longrightarrow \quad \sum_{i=1}^{n} \alpha_i\, f_i \succsim \sum_{i=1}^{n} \beta_i\, f_i \text{ for all } f_1, \ldots, f_n \in \mathcal{F} \text{ with } f_1 \sim \cdots \sim f_n.$$

A preference relation $\succsim$ exhibits preference for weak naive diversification *if for $n \in \mathbb{N}$ and $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n) \in \mathbb{S}^n$ it follows that:*

$$\frac{1}{n} \sum_{i=1}^{n} f_i \succsim \sum_{i=1}^{n} \alpha_i\, f_i \text{ for all } f_1, \ldots, f_n \in \mathcal{F} \text{ with } f_1 \sim \cdots \sim f_n.$$

This definition states that a preference relation $\succsim$ exhibits preference for naive diversification if, for alternatives that are equally ranked, an allocation to these alternatives is preferred to any alternative weight allocation that majorizes it. In other words, weight allocations that are closer to equality are always more preferred; see Ibragimov (2009).

We now derive some initial properties of a preference relation $\succsim$ that exhibits preference for naive diversification:

(1) On naive versus weak naive diversification. Definition 6 implies that $\frac{1}{n}\sum_{i=1}^{n} f_i \succsim \sum_{i=1}^{n} \alpha_i f_i$ for any $\boldsymbol{\alpha} \in \mathbb{S}^n$ and $f_1 \sim \cdots \sim f_n$, because any $\boldsymbol{\alpha} \in \mathbb{S}^n$ majorizes the equal-weighted decision vector $\boldsymbol{u}_n = \left(\frac{1}{n}, \ldots, \frac{1}{n}\right)$. It follows that the equal-weighted decision vector $\boldsymbol{u}_n$ is the most preferred choice allocation when $\succsim$ exhibits naive diversification preferences. This means that preference for naive diversification implies preference for weak naive diversification. However, the converse does not necessarily hold.

(2) On naive diversification and number of alternatives. In general, we have

$$\frac{1}{n} \sum_{i=1}^{n} f_i \succsim \frac{1}{n-1} \sum_{i=1}^{n-1} f_i \succeq \frac{1}{2}\,(f_1 + f_2) \succsim f_1,$$

for all $n \in \mathbb{N}$ and $f_1, \ldots, f_n \in \mathcal{F}$ such that $f_1 \sim \cdots \sim f_n$. This ordering entails the informal diversification paradigm that *more is better*, as analyzed by Elton and Gruber (1977), since an equal weighted allocation to n choices is more preferred to an equal weighted allocation to m choices if and only if $n \geq m$.

(3) On indifference under naive diversification. Note that choice weights under naive diversification preferences are equivalent whenever their ordered vectors coincide. Moreover, whenever a collection of choices are pairwise equally ranked, a convex combination of each of these must be equally ranked. The following formalization of these observations is hence an immediate consequence of Definition 6.

Lemma 1. *Let* $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n) \in \mathbb{S}^n$, $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_n) \in \mathbb{S}^n$, $f_1, \ldots, f_n \in \mathcal{F}$ *with* $f_1 \sim \cdots \sim f_n$*, and* $g_1, \ldots, g_n \in \mathcal{F}$ *with* $g_1 \sim \cdots \sim g_n$*, such that* $f_i \sim g_i$ *for* $i = 1, \ldots, n$*. Suppose that* $\succsim$ *exhibits preference for naive diversification. Then*

(i) $\sum_{i=1}^n \alpha_i\, f_i \sim \sum_{i=1}^n \beta_i\, g_i$ *if* $\sum_{i=1}^k \alpha_{(i)} = \sum_{i=1}^k \beta_{(i)}$ *for all* $k = 1, \ldots, n$*;*

(ii) $\sum_{i=1}^n \alpha_i\, f_i \sim \sum_{i=1}^n \alpha_i\, g_i$*.*

(4) On naive diversification and convex preferences. An agent whose preferences are convex chooses to diversify by taking a convex combination over individual choices without specifying a preference ordering over choice weights. So the classical notion of diversification does not necessarily imply preferences for naive diversification. The converse holds however: suppose that $\succsim$ exhibits preferences for naive diversification and let $f_1, \ldots, f_n \in \mathcal{F}$ with $f_1 \sim \cdots \sim f_n$. Then, for $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n) \in \mathbb{S}^n$, we have $\sum_{i=1}^n \alpha_i f_i \succsim f_j$ for all $j = 1, \ldots, n$, since the components of the choice vector $\boldsymbol{\alpha}$ are more nearly equal than those of $\boldsymbol{e}_j$, i.e., any $\boldsymbol{\alpha} \in \mathbb{S}^n$ is majorized by $\boldsymbol{e}_j$. This proves the following result.

Proposition 2. *Naive diversification preferences are convex, or, equivalently, exhibit preferences for diversification.*

4.3 Permutation invariant preferences

The notion of permutation invariance lies at the core of the definition of naive diversification. Permutation invariance captures the idea that the underlying characteristics of the individual choices are irrelevant in the decision making process. In other words, the economic agent is indifferent towards a permutation of the components of choice vectors. We formalize such permutation invariant preferences through permutation matrices. For a permutation matrix Π and choice vector $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n) \in \mathbb{S}^n$, we shall write $\boldsymbol{\alpha}\Pi$ for the vector whose components have been shuffled using Π and whose i-th component we denote by $(\boldsymbol{\alpha}\Pi)_i$. When ordering the components of $\boldsymbol{\alpha}\Pi$ in decreasing order, we denote its i-th ordered component by $(\boldsymbol{\alpha}\Pi)_{(i)}$.

Definition 7 (Permutation invariant preferences)**.** *A preference relation* $\succsim$ *on* $\mathcal{F}$ *is* permutation invariant *if for all* $\boldsymbol{f} = (f_1, \ldots, f_n) \in \mathcal{F}^n$ *with* $f_1 \sim \cdots \sim f_n$*, and* $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n) \in \mathbb{S}^n$*,*

$$\boldsymbol{\alpha} \cdot \boldsymbol{f} \sim (\boldsymbol{\alpha}\Pi) \cdot \boldsymbol{f} \ ,$$

where Π *is a permutation matrix.*

The following lemma shows that naive diversification preferences are permutation invariant.

Lemma 2. *Naive diversification preferences are permutation invariant.*

Proof. For all $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n) \in \mathbb{S}^n$, we have $\boldsymbol{\alpha}_\downarrow = (\boldsymbol{\alpha}\Pi)_\downarrow$. Therefore, $\sum_{i=1}^k \alpha_{(i)} = \sum_{i=1}^k (\boldsymbol{\alpha}\Pi)_{(i)}$ for all $k = 1, \ldots, n$. By Lemma 1, this implies that $\boldsymbol{\alpha} \cdot \boldsymbol{f} \sim (\boldsymbol{\alpha}\Pi) \cdot \boldsymbol{f}$. □

The significance of permutation invariance manifests itself in its implication for classical diversification. Indeed, imposing permutation invariance on convex preferences yields preferences for naive diversification (Proposition 4). We start by showing the weaker result.

Proposition 3. *A preference relation $\succsim$ that is permutation invariant and convex exhibits preference for weak naive diversification.*

Proof. Because any $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n) \in \mathbb{S}^n$ majorizes the vector $\mathbf{u}_n$, then, according to Proposition 1, $\mathbf{u}_n$ can be derived from $\boldsymbol{\alpha}$ by successive applications of a finite number of T-transforms, i.e.,

$$\mathbf{u}_n = \boldsymbol{\alpha} T_1 T_2 \cdots T_k$$

where $T_1, T_2, \cdots T_k$ are T-transforms. For $f_1, \ldots, f_n \in \mathcal{F}$, we have:

$$\frac{1}{n} \sum_{i=1}^n f_i = \mathbf{u}_n \cdot \boldsymbol{f} = (\boldsymbol{\alpha} T_1 \cdots T_k) \cdot \boldsymbol{f}.$$

We prove that $(\boldsymbol{\alpha} T_1 \cdots T_k) \cdot \boldsymbol{f} \succsim \boldsymbol{\alpha} \cdot \boldsymbol{f}$ by mathematical induction. First of all, we show that $(\boldsymbol{\alpha} T) \cdot \boldsymbol{f} \succsim \boldsymbol{\alpha} \cdot \boldsymbol{f}$ when T is T-transform and $\succsim$ is permutation invariant and convex. Indeed,

$$(\boldsymbol{\alpha} T) \cdot \boldsymbol{f} = [\boldsymbol{\alpha}(\lambda\, I + (1-\lambda) Q)] \cdot \boldsymbol{f} = \lambda\, \boldsymbol{\alpha} \cdot \boldsymbol{f} + (1-\alpha)\, (\boldsymbol{\alpha} Q) \cdot \boldsymbol{f}$$

where Q is a permutation matrix. Because $\succsim$ is permutation invariant, then $(\boldsymbol{\alpha} Q) \cdot \boldsymbol{f} \sim \boldsymbol{\alpha} \cdot \boldsymbol{f}$. Finally, because $\succsim$ is convex, then

$$\lambda\, \boldsymbol{\alpha} \cdot \boldsymbol{f} + (1-\lambda)\, (\boldsymbol{\alpha} Q) \cdot \boldsymbol{f} \succsim \boldsymbol{\alpha} \cdot \boldsymbol{f}.$$

It follow that:

$$(\boldsymbol{\alpha} T) \cdot \boldsymbol{f} \succsim \boldsymbol{\alpha} \cdot \boldsymbol{f}.$$

Now suppose that $(\boldsymbol{\alpha} T_1 \cdots T_{k-1}) \cdot \boldsymbol{f} \succsim \boldsymbol{\alpha} \cdot \boldsymbol{f}$. Let $\tilde{\boldsymbol{\alpha}} = \boldsymbol{\alpha} T_1 \cdots T_{k-1}$. It follows that:

$$(\boldsymbol{\alpha} T_1 \cdots T_k) \cdot \boldsymbol{f} = (\tilde{\boldsymbol{\alpha}} T_k) \cdot \boldsymbol{f} \succsim \tilde{\boldsymbol{\alpha}} \cdot \boldsymbol{f} = (\boldsymbol{\alpha} T_1 \cdots T_{k-1}) \cdot \boldsymbol{f} \succsim \boldsymbol{\alpha} \cdot \boldsymbol{f}.$$

Therefore,

$$(\boldsymbol{\alpha} T_1 \cdots T_k) \cdot \boldsymbol{f} \succsim \boldsymbol{\alpha} \cdot \boldsymbol{f}.$$

This proves the statement of the proposition. □

We recall that T-transforms (Definition 3) are averaging operators between two components of the original weight vector. This averaging operator is always weakly preferred under permutation invariant and convex preferences. The proof of Proposition 3 shows that repeated averaging of two components of a weight vector reaches its limit at the equal-weighted decision vector $\boldsymbol{u}_n$. Therefore, Proposition 3 can be viewed as a corollary to Muirhead's result (Proposition 1).

Another seminal result tangentially related to Proposition 3 appeared in Samuelson (1967), where the first formal proof of the following, at the time seemingly well-understood, diversification paradigm is given: "*putting a fixed total of wealth equally into independently, identically distributed investments will leave the mean gain unchanged and will minimize the variance.*" One may hence think of the conditions of having non-negative, independent and identically distributed random variables in Theorem 1 of Samuelson (1967) being replaced by the permutation invariance condition in Proposition 3 to yield an equal weighted allocation as optimal.[5]

We next derive the stronger statement, which gives naive diversification under permutation invariance and convexity.

Proposition 4. *A preference relation $\succsim$ that is permutation invariant and convex exhibits preference for naive diversification.*

Proof. Suppose that $\succsim$ is permutation invariant and convex. We have to show that $\boldsymbol{\alpha}\cdot\boldsymbol{f} \succsim \boldsymbol{\beta}\cdot\boldsymbol{f}$ for all $\boldsymbol{f} \in \mathcal{F}^n$ when $\boldsymbol{\beta} \geq_m \boldsymbol{\alpha}$. If $\boldsymbol{\beta} \geq_m \boldsymbol{\alpha}$, then $\boldsymbol{\alpha}$ can be derived from $\boldsymbol{\beta}$ by successive applications of a finite number of T-transforms. By applying the same argument as in the proof of Proposition 3, we have $\boldsymbol{\alpha}\cdot\boldsymbol{f} \succsim \boldsymbol{\beta}\cdot\boldsymbol{f}$. Therefore, $\succsim$ exhibits preference for naive diversification. □

Combining Proposition 2 and Lemma 2 with Proposition 4 yields the following equivalence of preferences.

Theorem 2. *A monotonic and continuous preference relation $\succsim$ exhibits preference for naive diversification if and only if it is convex and permutation invariant.*

4.4 A geometric characterization

In this subsection, we give a geometric characterization of convex preferences that are permutation invariant. The characterization relies on classical results from convex analysis and linear algebra, which we briefly recall first.

A set which is the convex hull of finitely many points is called a *polytope*. Fix an allocation vector $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n) \in \mathbb{S}^n$. The convex hull of all vectors in $\mathbb{S}^n$ obtained by permutations of the coordinates α_i of $\boldsymbol{\alpha}$ is called the *permutation polytope* $\mathcal{K}_{\boldsymbol{\alpha}}$ of the vector $\boldsymbol{\alpha}$:

$$\mathcal{K}_{\boldsymbol{\alpha}} = \mathrm{conv}\{\boldsymbol{\alpha}\Pi : \Pi \text{ permutation matrix }\}.$$

Another polytope of relevance in this discussion is the *Birkhoff polytope* $\mathcal{B}_n$, which is the convex hull of the set of all permutation matrices of dimension n. The Birkhoff-von-Neumann Theorem (Birkhoff 1946) states that every doubly stochastic real matrix is in

[5]See Hadar and Russell (1969), Hadar and Russell (1971), Tesfatsion (1976) and Li and Wong (1999) for generalizations of Samuelson's classical result.

fact a convex combination of permutation matrices of the same order. The permutation matrices are then precisely the extreme points of the set of doubly stochastic matrices.

We now reformulate the decision making problem from choice amongst objects in $\mathcal{F}$ to an allocation problem to a given selection of objects $f_1, \ldots, f_n \in \mathcal{F}$. That is, faced with n different objects, a decision maker must decide on an allocation vector in $\mathbb{S}_n$. Permutation invariance implies indifference amongst all possible permutations of allocation vectors. The decision maker's preference relation thus reduces to the majorization preorder $\leq_m$ on $\mathbb{S}^n$. For a given allocation vector $\boldsymbol{\alpha} \in \mathbb{S}^n$, consider the contour set

$$C(\boldsymbol{\alpha}) = \{\boldsymbol{\beta} \in \mathbb{S}^n : \boldsymbol{\beta} \leq_m \boldsymbol{\alpha}\},$$

which is the set of all antecedents of $\boldsymbol{\alpha}$ in the majorization preordering $\leq_m$. This set is in fact the permutation polytope of the allocation vector $\boldsymbol{\alpha}$ (Rado 1952), and is thus generated as the convex hull of points obtained by permuting the components of $\boldsymbol{\alpha}$. This means that indifference curves associated with permutation invariance are in fact the vertices of the permutation polytope $K_{\boldsymbol{\alpha}} = C(\boldsymbol{\alpha})$. Consequently, if $\boldsymbol{\beta} \leq_m \boldsymbol{\alpha}$, so that by Theorem 1, $\boldsymbol{\beta} = \boldsymbol{\alpha} P$ for some doubly stochastic matrix P, then there exist constants $c_i \geq 0$ with $\sum c_i = 1$, such that

$$\boldsymbol{\beta} = \boldsymbol{\alpha}\left(\sum c_i \Pi_i\right) = \sum c_i(\boldsymbol{\alpha}\Pi_i),$$

where the Π_i are permutation matrices. This means, as was noted by Rado (1952), that $\boldsymbol{\beta}$ lies in the convex hull of the orbit of $\boldsymbol{\alpha}$ under the group of permutation matrices. Figure 1 illustrates indifference curves and associated contour sets for the cases $n = 2$ and $n = 3$.

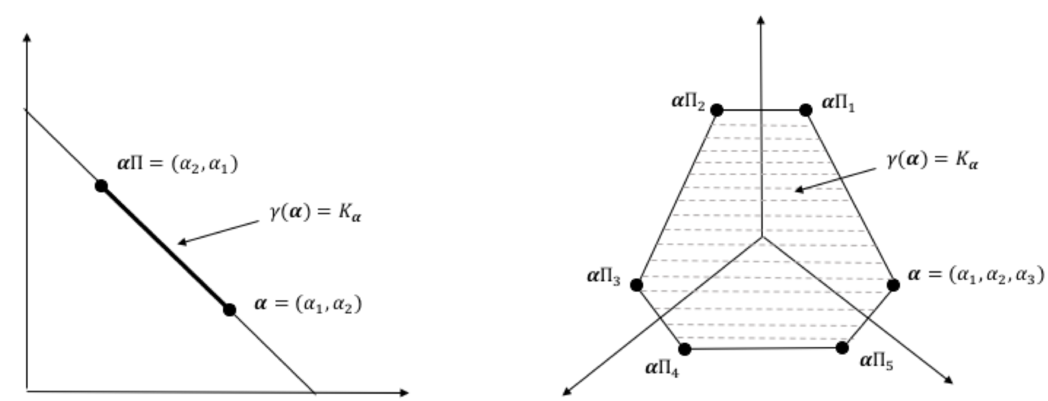

Figure 1: Indifference curves and associated contour sets for the allocation to $n = 2$ choice options (left) and $n = 3$ choice options (right).

5 Representation

We now derive necessary and sufficient conditions on a utility function U such that the corresponding preference relation $\succsim$, with $f \succsim g \iff U(f) \geq U(g)$, exhibits preference for naive diversification. In particular, we show that naive diversification preferences arise when decision makers treat choice alternatives as being mixtures of conditionally independent and identically distributed random variables, with correlation neglect as a special case.

Our main result so far states that a preference relation exhibits preference for naive diversification if and only if it is convex and permutation invariant. Cerreia-Vioglio, Maccheroni, Marinacci, and Montrucchio (2011) provide a characterization for a general class of preferences that are non-trivial, complete, transitive, monotone, risk independent, continuous and convex, which are known as *uncertainty averse preferences*. This means that naive diversification preferences constitute a subclass of uncertainty averse preferences that are additionally permutation invariant. We thus build our derivation on the representation results for uncertainty averse preferences of Cerreia-Vioglio, Maccheroni, Marinacci, and Montrucchio (2011).

A preference relation $\succsim$ on $\mathcal{F}$ is uncertainty averse if and only if its representation takes the form

$$U(f) = \inf_{\mathbb{Q} \in \Delta} G\left(\mathbb{E}_{\mathbb{Q}}\left[u(f)\right], \mathbb{Q}\right),$$

where $u : X \to \mathbb{R}$ is non-constant and affine, and $G : u(X) \times \Delta \to (-\infty, \infty]$, called the uncertainty aversion index, is linearly continuous, quasi-convex, increasing in the first variable with $\inf_{\mathbb{Q} \in \Delta} G(t, \mathbb{Q}) = t$ for $t \in u(X)$.

Under this representation, decision makers consider all possible probabilities $\mathbb{Q}$ and the associated expected utilities. They then summarize all these evaluations by taking their minimum. The function G can be interpreted as an index of uncertainty aversion; higher degrees of uncertainty aversion correspond to pointwise smaller indices G. The quasi-convexity of G and the cautious attitude reflected by the minimum derive from the convexity of preferences, or, equivalently, from preferences for traditional diversification. Uncertainty aversion is hence closely related to convexity of preferences. Under this formalization, convexity reflects a basic negative attitude of decision makers towards the presence of uncertainty in their choices.

Now, we assume that decision makers exclusively form convex combinations of non-constant acts from an infinite sequence $\mathbf{f} = (f_1, f_2, \dots)^T$ in $\mathcal{F}$. This means that choice alternatives are elements of the convex hull $\text{conv}\{f_1, f_2, \dots\}$ of $\{f_1, f_2, \dots\}$. This assumption hold for example when the set of consequences is a convex subset of a vector space with countable basis.[6]

The following definition will play a central role in our main representation result:

[6]With some abuse of notation we consider the infinite sequence $\mathbf{f} = (f_1, f_2, \dots)^T$ as a vector of acts with values in X^∞.

Definition 8 (Exchangeability). *Let* $\mathbb{Q} \in \Delta$. *An infinite sequence* $\mathbf{w}^T = (w_1, w_2, \dots)$ *of elements in* $B_0(\Sigma)$ *is said to be* $\mathbb{Q}$-exchangeable *if and only if* $\mathbf{w}$ *has the same distribution under* $\mathbb{Q}$ *as* $\Pi\,\mathbf{w}$ *for any permutation matrix* $\Pi \in \mathbb{R}^\infty \times \mathbb{R}^\infty$ *that only permutes a finite number of elements of* $\mathbf{w}$.

A well-known result on exchangeable sequences is de Finetti's theorem, which states that an infinite sequence is exchangeable if and only if it corresponds to a mixture of independent and identically distributed sequences (Aldous 1985). Formally, we first define random measures as follows:

Definition 9 (Random measure). *The function* $\nu : S \times \mathcal{B}(\mathbb{R}) \to \mathbb{R}_+$, *where* $\mathcal{B}(\mathbb{R})$ *is the Borel* σ*-algebra on* $\mathbb{R}$, *is a* random measure *if* $\nu(s, \cdot)$ *is a probability measure on* $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ *for all* $S \in S$ *and* $\nu(\cdot, A)$ *is a random variable on* (S, Σ) *for all* $A \in \mathcal{B}(\mathbb{R})$.

The following result holds:

Lemma 3 (de Finetti). *An infinite sequence* $\mathbf{w} = (w_1, w_2, \dots)$ *of elements in* $B_0(\Sigma)$ *is* $\mathbb{Q}$*-exchangeable if and only if a random measure* ν *exists such that:*

(1) $w_1, w_2, \dots$ *are conditionally independent given* $\mathcal{G}$, *i.e.,*

$$\mathbb{Q}\left[w_i \in A_i,\, 1 \le i \le n | \mathcal{G}\right] = \prod_{i=1}^{n} \mathbb{Q}\left[w_i \in A_i | \mathcal{G}\right], \quad A_1, \dots, A_n \in \mathcal{B}(\mathbb{R}), \quad n \ge 1;$$

and

(2) the conditional distribution of x_i *given* $\mathcal{G}$ *is* ν, *i.e.,*

$$\mathbb{Q}\left[w_i \in A_i | \mathcal{G}\right] = \nu(\cdot, A_i), \quad A_i \in \mathcal{B}(\mathbb{R}), i = 1, 2, \dots$$

where $\mathcal{G}$ *is the* σ*-algebra generated by the family of random variables* $(\nu(\cdot, A))_{A \in \mathcal{B}(\mathbb{R})}$.

Following the result of Lemma 3, we say that an exchangeable infinite sequence of elements in $B_0(\Sigma)$ is a mixture of i.i.d. sequences dictated by a random measure ν. An immediate implication of Lemma 3 is that any two terms of an exchangeable infinite sequence have zero conditional correlation:

Corollary 1 (Exchangeability and correlation neglect). *Let* $\mathbf{w} = (w_1, w_2, \dots)$ *be a* $\mathbb{Q}$*-exchangeable infinite sequence in* $B_0(\Sigma)$ *dictated by the random measure* ν. *It follows that:*

$$\rho_{\mathbb{Q}}(x_i, x_j | \mathcal{G}) = 0$$

where $\mathcal{G}$ *is the* σ*-algebra generated by the family of random variables* $(\nu(\cdot, A))_{A \in \mathcal{B}(\mathbb{R})}$.

Note that the (unconditional) correlation of any two terms in an exchangeable infinite sequence does not have to be zero, as illustrated in the following example.

To see the link between exchangeability and correlation neglect, consider the following example. Let $m, s \in B_0(\Sigma)$ and $\mathbf{w} = (w_1, w_2, \dots)^T$ be a sequence of independent and identically distributed random variables in $B_0(\Sigma)$ under probability measure $\mathbb{Q} \in \Delta$. Define

$v_i = m + s\,w_i$ for $i = 1, 2, \dots$. It follows that $\mathbf{v} = (v_1, v_2, \dots)$ is $\mathbb{Q}$-exchangeable. Indeed, conditioning on m and s, $v_1, v_2, \dots$ are independent and identically distributed. Clearly, conditioning on m and s, the correlation between any two random variables v_i and v_j for $i \neq j$ is equal to zero. However, (unconditionally) v_i and v_j for $i \neq j$ are not independent and in general also not identically distributed.

We now present our main representation result.

Theorem 3 (Representation of naive-diversification preferences)**.** *Let $\mathbf{f} = (f_1, f_2, \dots)^T$ be an infinite sequence of non-constant acts in $\mathcal{F}$ from which the decision maker forms convex combinations. Then a preference relation on conv$\{f_1, f_2, \dots\}$ exhibits preference for naive diversification if and only if its utility representation is given by*

$$U(f) = \inf_{\mathbb{Q}\in\Delta} G_e\left(\mathbb{E}_{\mathbb{Q}}\left[u(f)\right], \mathbb{Q}\right),$$

where $u : X \to \mathbb{R}$ is affine and $G_e : u(X) \times \Delta \to (-\infty, \infty]$ is an index of uncertainty aversion with $G_e(\mathbb{Q}, \cdot) = \infty$ for $\mathbb{Q} \in \Delta \setminus \Delta_e$ and $\Delta_e \subset \Delta$ is the set of probability measures $\mathbb{Q}$ on (S, Σ) such that $(u(f_1), u(f_2), \dots)^T$ is a mixture of i.i.d. sequences dictated by some random measure ν under $\mathbb{Q}$.

Proof. One direction directly follows from Lemma 3. Let $\tilde{\mathbf{f}} = (f_{i_1}, \dots, f_{i_n})^T$ and $u(\tilde{\mathbf{f}}) = (u(f_{i_1}), \dots, u(f_{i_n}))^T$ where $i_j \geq 1$ for all $j = 1, \dots, n$ and $i_j \neq i_k$ for $j \neq k$. Because for any $\boldsymbol{\alpha} \in \mathbb{S}^n$ and any $n \times n$ permutation matrix Π we have:

$$u(\boldsymbol{\alpha}\Pi \cdot \tilde{\mathbf{f}}) = \boldsymbol{\alpha}\Pi \cdot u(\tilde{\mathbf{f}}) = \boldsymbol{\alpha} \cdot \Pi u(\tilde{\mathbf{f}})$$

then according to Lemma 3, $u(\boldsymbol{\alpha}\Pi \cdot \tilde{\mathbf{f}})$ as the same distribution as $u(\boldsymbol{\alpha} \cdot \tilde{\mathbf{f}})$ under any probability measure $\mathbb{Q} \in \Delta_e$. Therefore, for

$$U(f) = \inf_{\mathbb{Q}\in\Delta} G_e\left(\mathbb{E}_{\mathbb{Q}}\left[u(f)\right], \mathbb{Q}\right)$$

we have:

$$U(\boldsymbol{\alpha}\Pi \cdot \tilde{\mathbf{f}}) = U(\boldsymbol{\alpha} \cdot \tilde{\mathbf{f}})$$

for any $\boldsymbol{\alpha} \in \mathbb{S}^n$ and any $\tilde{\mathbf{f}}$. It follows that the preference relation represented by U is convex and permutation invariant and thus exhibits preference for naive diversification.

The other direction works as follows. Suppose that $\succsim$ is convex and permutation invariant. For any $\boldsymbol{\alpha} \in \mathbb{S}^n$ and any $\tilde{\mathbf{f}} = (f_{i_1}, \dots, f_{i_n})^T$ and $u(\tilde{\mathbf{f}}) = (u(f_{i_1}), \dots, u(f_{i_n}))^T$ where $i_j \geq 1$ for all $j = 1, \dots, n$ and $i_j \neq i_k$ for $j \neq k$, we have:

$$\boldsymbol{\alpha} \cdot \tilde{\mathbf{f}} \sim (\boldsymbol{\alpha}\Pi) \cdot \tilde{\mathbf{f}}$$

For any $n \times n$ permutation matrix Π. As this must also apply to constant acts, in this case we have,

$$u(\boldsymbol{\alpha} \cdot \tilde{\mathbf{f}}) = u\left((\boldsymbol{\alpha}\Pi) \cdot \tilde{\mathbf{f}}\right) \Leftrightarrow \boldsymbol{\alpha} \cdot u(\tilde{\mathbf{f}}) = (\boldsymbol{\alpha}\Pi) \cdot u(\tilde{\mathbf{f}}) = \boldsymbol{\alpha} \cdot (\Pi u(\tilde{\mathbf{f}})) \Leftrightarrow \boldsymbol{\alpha} \cdot \left(u(\tilde{\mathbf{f}}) - \Pi u(\tilde{\mathbf{f}})\right) = 0.$$

for any $n \times n$ permutation matrix Π and any $\boldsymbol{\alpha} \in \mathbb{S}^n$. It follows that

$$u(\tilde{\mathbf{f}}) = \Pi u(\tilde{\mathbf{f}})$$

for any $n \times n$ permutation matrix Π and any $n \geq 1$. For general acts, the equality is in distribution and this is violated if $(u(f_1), u(f_2), \dots)$ is not exchangeable. Therefore, in the representation of $\succsim$ we can limit the set of measures in Δ to those under which $(u(f_1), u(f_2), \dots)$ is exchangeable. □

The utility of naive diversification preferences thus represents uncertainty averse preferences with the additional requirement that decision makers only consider beliefs that imply some weak form of independence. Therefore, naive diversification is closely related to correlation neglect. However, the latter is a much stronger condition under which naive diversification arises.

We end this Section with a simple example illustrating the connection between naive diversification and correlation neglect. Consider three non-degenerate normally and identically distributed random variables x_1, x_2, and x_3 representing payoffs to assets, where x_3 is independent of x_1 and x_2, but with x_1 and x_2 perfectly negatively correlated, that is, $\rho(x_1, x_2) = -1$. For any risk-averse investor with preferences represented by the relation $\succsim$,

$$\frac{1}{2}(x_1 + x_2) \succsim \frac{1}{3}(x_1 + x_2 + x_3).$$

This is because allocating equally to perfect-negatively correlated choices is risk-free, whereas $\frac{1}{3}(x_1 + x_2 + x_3)$ is not, although both have the same mean. However, under our formalization of preferences for naive diversification, the distribution on the right weakly dominates the one on the left. In other words, naive diversifiers ignore correlations among assets and this may lead to indifference to mean-preserving spreads and thus to a preference for second-degree stochastically dominated alternatives.

6 Concluding Remarks

In this paper, we provided mathematically and economically sound choice theoretic foundations for the naive approach to diversification. In particular, we axiomatized naive diversification by defining it as a preference for equality over inequality, and showed that the notion of permutation invariance lies at the core of naive diversification. Moreover, we derived necessary and sufficient conditions on the utility functions that give rise to preferences for naive diversification by showing that naive diversification preferences arise when decision makers only consider beliefs that imply some weak form of independence, which is closely related to correlation neglect.

The theory of majorization underlying the formalization of naive diversification preferences and their representation is a rich theory that lends itself to wider extensions going beyond the axiomatization and representation results of this paper. Appendices A.2 and A.3 give an overview of two potentially useful extensions of our theory, namely comparison of levels of naive diversification and rebalancing of allocation to equality.

We conclude by briefly discussing the relationship between our axiomatic system and observed behavior in reality, followed by sketching choice theoretic extensions of our work.

6.1 Testing the reality of naive diversification

Even though desirability for diversification is a cornerstone of a broad range of portfolio choice models, the precise formal definition differs from model to model. Analogously, the way in which the notion of diversification is interpreted and implemented in the real world varies greatly. Traditional diversification paradigms are consistently violated in practice. Indeed, empirical evidence suggests that economic agents often choose diversification schemes other than those implied by Markowitz's portfolio theory or expected utility theory. Diversification heuristics thus span a vast range, and naive diversification, in particular, has been widely documented both empirically and experimentally.

However, despite the growing literature pointing to the common existence of naive diversification in practice, experimental research investigating the *behavioral drivers* of diversifiers remains rather limited. Our axiomatization can help empirical and experimental economists test diversification preferences, and their underlying drivers, of economic agents in the real world. In particular, we can now look for the main parameters driving the decision process of naive diversifiers. One such parameter or heuristic implied by our axiomatization is that of permutation invariance. In practice, it is arguably rather rare that a diversifier would know so little about the given assets to be essentially indifferent among them. Despite this, naive diversification continues to be applied by both experienced professionals and regular people. By varying the amount of information available to subjects in an experimental setting, one may be able to deduce whether the indifference axiom applies in general or whether it is information dependent, as implied by Laplace's principle of indifference. Another insight gained through our axiomatization was that of consistency with traditional convex diversification and concave expected utility maximization. In particular, consider that a risk averse investor would in theory be expected to diversify in the traditional convex sense. Hence, the level of risk aversion may be yet another parameter driving naive diversification, and this again can be directly tested.

6.2 Choice-theoretic generalizations

Comparing allocations among different numbers of choices. Our discussion of naive diversification throughout has focused on a fixed number of choice alternatives n. Suppose that an economic agent is faced with an allocation among either $\boldsymbol{f} = (f_1, \dots, f_n)$ or $\boldsymbol{g} = (g_1, \dots, g_m)$, where $n \neq m$. In Section 3, we showed that an equal allocation among a larger number of alternatives is always more preferred under naive diversification. More generally, however, given unequal choice weights $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_n) \in \mathbb{S}^n$ and $\boldsymbol{\beta} = (\beta_1, \dots, \beta_m) \in \mathbb{S}^m$ and allocations $\boldsymbol{\alpha} \cdot \boldsymbol{f}$ and $\boldsymbol{\beta} \cdot \boldsymbol{g}$, one can cannot infer a preference of one over the other without generalizing the naive diversification axiom. Such an extension has been developed by Marshall, Olkin, and Arnold (2011) in the context of the majorization order on vectors of unequal lengths. In fact, they showed that the components of $\boldsymbol{\alpha}$ are less spread out than the components of $\boldsymbol{\beta}$ if and only if the Lorenz curve $L_{\boldsymbol{\alpha}}$ associated with the vector $\boldsymbol{\alpha}$ is greater or equal than the Lorenz curve $L_{\boldsymbol{\beta}}$ associated with $\boldsymbol{\beta}$ for all values in its domain $[0,1]$, and that this is equivalent to requiring that $1/n \sum_{i=1}^{n} \phi(\alpha_i) \leq 1/m \sum_{i=1}^{m} \phi(\beta_i)$ for all convex functions $\phi : \mathbb{R} \to \mathbb{R}$.

Multidimensional diversification. One may think of naive diversification as being *univariate*, in the sense that a naive diversifier is concerned with only one dimension, namely that of equality of choice weights. Suppose that an economic agent would like to diversify naively, but would also like to reduce variability along a second dimension. Consider for example the dimension of "risk weights" as opposed to "capital weights". This is a commonly applied risk diversification strategy in practice, known under *risk parity*. Parity diversification focuses on allocation of risk, usually defined as volatility, rather than allocation of capital. Here, risk contributions across choice alternatives are equalized (and are in practice typically levered to match market levels of risk). It can be viewed as a middle ground between the naive approach and the minimum risk approach (see for example Maillard, Roncalli, and Teiletche (2010)).

When allocations along more than one dimension are to be compared simultaneously, we move from the linear space of choice vectors to the space of choice *matrices*. Each row of a choice matrix represents a particular attribute or dimension, whereas each column represents the choice weights along that dimension. The generalization of the mathematical formalism of naive diversification is then straightforward. For example, a choice matrix X is more diversified (along some given dimensions) than a choice matrix Y if $X = PY$ for some doubly stochastic matrix P. This definition is part of an established field within linear algebra known as multivariate majorization.

Towards an inequality aversion coefficient. The naive diversification axiom implies that a weight allocation that is closest to the equal weighted vector $\boldsymbol{u}_n$ is always more preferred. This in turn induces the idea of being *averse to inequality*, which we discussed in Section 4. One may formalize this notion, together with a characterization of different levels of inequality aversion as follows.

First, yet another generalization of naive diversification can be obtained by substituting a more general vector $\boldsymbol{d} \in \mathbb{S}^n$ for the equality vector $\boldsymbol{u}_n$. In that case, weight allocations closest to $\boldsymbol{d}$ are preferred. To do this, we need to define the concept of d-stochastic matrix. For $\boldsymbol{d} \in \mathbb{S}^n$, an $n \times n$ matrix $A = (a_{ij})$ is said to be $\boldsymbol{d}$*-stochastic* if (i) $a_{ij} \geq 0$ for all $i, j \leq n$; (ii) $\boldsymbol{d}A = \boldsymbol{d}$; and (iii) $A\boldsymbol{u}_n' = \boldsymbol{u}_n'$. To get an intuition for $\boldsymbol{d}$-stochastic matrices, note that since $\sum_{i=1}^{n} d_i = 1$ by construction, a $\boldsymbol{d}$-stochastic matrix in our setting can be viewed as the transition matrix of a Markov chain. Clearly, when $\boldsymbol{d} = \boldsymbol{u}_n$, a $\boldsymbol{d}$-stochastic matrix is doubly stochastic. One can then say that a preference relation $\succsim$ exhibits preference for *relative naive diversification* if there is a weight allocation $\boldsymbol{d} = (d_1, \ldots, d_n) \in \mathbb{S}^n$ such that for any $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n) \in \mathbb{S}^n$ and $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_n) \in \mathbb{S}^n$,

$$\boldsymbol{\alpha} \leq_m \boldsymbol{\beta} \iff \boldsymbol{\alpha} = \boldsymbol{\beta} A$$

for some $\boldsymbol{d}$-stochastic matrix A. The interpretation here is that an individual with naive diversification preferences relative to some $\boldsymbol{d} \neq \boldsymbol{u}_n$ is *less averse* to inequality than one with naive diversification preferences.

To be then able to compare levels of aversion to inequality within relative naive diversification preferences, we can introduce the *coefficient of inequality aversion*. For naive diversification preferences relative to $\boldsymbol{d} \in \mathbb{S}^n$, the corresponding *inequality aversion coefficient* ε is defined as $\varepsilon = \|\boldsymbol{d} - \boldsymbol{u}_n\|$, where $\|\cdot\|$ is the Euclidean norm taken up-to-permutation. Clearly, this inequality aversion coefficient ε lies within $[0, \infty)$, with $\varepsilon = 0$ for naive diversification preferences, in which case we can say that the decision maker possesses *absolute aversion to inequality*.

A Appendix

A.1 Measures of naive diversification

An evaluation of the optimality of a given choice allocation of a naive diversifier essentially reduces to a measure of inequality of the decision weights of his choice. Measures of inequality arise in various disciplines within economic theory, particularly within the context of wealth and income. Indeed, there is a vast literature on diversity and inequality indices in economics — see classical discussions and surveys by Sen (1973), Szal and Robinson (1977), Dalton (1920), Atkinson (1970), Blackorby and Donaldson (1978), and Krämer (1998). Most of these indices have been developed primarily based on foundations of the concept of social welfare, and hence may not necessarily be applicable to our setting.

Since a measure of inequality strongly depends on the context, we provide an axiomatization that is consistent with our definition of preference for naive diversification, which has a precise mathematical formulation in terms of majorization and Schur-concave functions. Many existing indices measuring allocation optimality or inequality are qualitative in nature focused on ranking with no indication of a quantification of the comparison. We do not only seek a qualitative ranking of choice allocations, but we aim to quantify the distance between two weight allocations. The resulting measure hence indicates how far from optimality a given choice allocation is and allows for comparison of two non-equal choice allocations in terms of their distance.

Let $\succsim$ be a preference relation on $\mathcal{F}$ exhibiting preferences for naive diversification. To derive the qualitative and quantitative properties that are consistent with naive diversification, we fix the optimal choice allocation $\boldsymbol{u}_n = (\frac{1}{n}, \ldots, \frac{1}{n})$ for a given n and look at comparisons with respect to this vector. The following are the minimal requirements that a measure $\mu_n : \mathcal{F} \to \mathbb{R}$ of naive diversification should satisfy:

(A1) Positivity: For all $f \in \mathcal{F}$, $\mu_n(f) \geq 0$.

(A2) Normality: For all $f \in \mathcal{F}$, $\mu_n(f) = 0$ if and only if $f \sim \boldsymbol{u}_n \cdot \boldsymbol{f}$ for some $\boldsymbol{f} = (f_1, \ldots, f_n)$.

(A3) Boundedness: For all $f \in \mathcal{F}$, $\mu_n(f) < \infty$.

(A4) Representation: For all $f, g \in \mathcal{F}$, $f \succsim g$ implies $\mu_n(f) \geq \mu_n(g)$.

(A5) Permutation invariance: For $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n), \boldsymbol{\beta} = (\beta_1, \ldots, \beta_n) \in \mathbb{S}^n$ and $f_1, \ldots, f_n \in \mathcal{F}$, if $\sum_{i=1}^n \alpha_i f_i \succsim \sum_{i=1}^n \beta_i f_i$ and $\boldsymbol{\alpha} \neq \Pi\boldsymbol{\beta}$ for any permutation matrix Π, then $\mu_n(\sum_{i=1}^n \alpha_i f_i) > \mu_n(\sum_{i=1}^n \beta_i f_i)$.

Axioms A1, A2, and A3 essentially ensure that the function μ_n is a well-behaved probability metric (Rachev, Stoyanov, and Fabozzi 2011) and hence an analytically sound measure of the distance between two random quantities. Axiom A4 implies Schur-concavity and thus that the qualitative ranking is preserved. By introducing invariance under permutation (Axiom A5), we require strict Schur-concavity. This distinguishes equivalence, and hence a zero distance from equality, from a strict preference ordering of choice weights, which should give a strictly positive distance.

Some well-known classes of measures from statistics, economics and asset management that satisfy the above axioms include statistical dispersion measures, economic inequality indices, such as the Gini coefficient (Gini 1921), Dalton's measure (Dalton 1920) and Atkinson's measure (Atkinson 1970), and diversification indices such as the Herfindahl-Hirschman Index (Hirschman 1964) and the Simpson diversity index (Simspon 1949).

A.2 Rebalancing to equality

Based on Theorem 1 of Hardy, Littlewood, and Pólya (1929), a doubly stochastic matrix can be thought of as an operation between two weight allocations leading towards greater equality in the weight vector. With this in mind, we define a *rebalancing transform* to be a doubly stochastic matrix. Clearly, rebalancing in this context cannot yield a less diversified allocation. In other words, applying a rebalancing transform to a vector of decision weights is equivalent to averaging the decision weights.

In this Section, we characterize such transforms which start with a suboptimal weight allocation $\sum_{i=1}^n \alpha_i f_i$ and produce equality $\frac{1}{n}\sum_{i=1}^n f_i$ in terms of their implied turnover in practice. Our analysis is focused on the asset allocation problem, where rebalancing is understood in terms of buying and selling positions. However, this discussion can be generalized to characterize transforms in the context of reallocation of wealth, such as Dalton's principle of transfers.

Starting from an allocation $\boldsymbol{\alpha} \in \mathbb{S}^n$, there are, in general, more than one possible transforms that rebalance $\boldsymbol{\alpha}$ to $\boldsymbol{u}_n$ or, more generally, to an allocation $\boldsymbol{\beta} \in \mathbb{S}^n$ that is closer to equality. Given two weight allocations $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{S}^n$ with $\boldsymbol{\alpha}$ majorized by $\boldsymbol{\beta}$, the set

$$\Omega_{\boldsymbol{\alpha} \leq_m \boldsymbol{\beta}} = \{P \in \mathbb{D}_n \mid \boldsymbol{\alpha} = \boldsymbol{\beta} P\}$$

is referred to as the *rebalancing polytope* of the order $\boldsymbol{\alpha} \leq_m \boldsymbol{\beta}$.[7] The set $\Omega_{\boldsymbol{\alpha} \leq_m \boldsymbol{\beta}}$ is nonempty, compact and convex. In the case that the components of $\boldsymbol{\beta}$ are simply a rearrangement of

[7]Within the linear algebra literature, this set is referred to as the "majorization polytope". As pointed out by Marshall, Olkin, and Arnold (2011), very little is known about this polytope.

the components of $\boldsymbol{\alpha}$, then $\Omega_{\boldsymbol{\alpha}\leq_m\boldsymbol{\beta}}$ contains one unique permutation matrix. In general, however, $\Omega_{\boldsymbol{\alpha}\leq_m\boldsymbol{\beta}}$ contains more than one element.

Now, for $\boldsymbol{\lambda} \in \mathbb{S}^n$, we have $\boldsymbol{u}_n \leq_m \boldsymbol{\lambda}$, and so our focus henceforth is the set

$$\Omega_{n,\boldsymbol{\lambda}} := \Omega_{\boldsymbol{u}_n\leq_m\boldsymbol{\lambda}} = \{P \in \mathbb{D}_n \mid \boldsymbol{u}_n = \boldsymbol{\lambda}P\} \ .$$

It contains all rebalancing transformations that lead to an equal allocation. In particular, it includes the matrix P_n with all entries equal to $1/n$.

We are interested in rebalancing a weight allocation towards equality in practice. However, it is not clear how or why one would choose one transform in a given polytope $\Omega_{n,\boldsymbol{\lambda}}$ over another. We provide a precise distinction in terms of *turnover*. In the context of asset allocation, the particular rebalancing transform applied to rebalance one weight allocation to another has an interpretation in terms of the fraction of assets bought and sold and, consequently, in terms of the implied transaction costs.

Definition 10 (Turnover). *For $\boldsymbol{\lambda} \in \mathbb{S}^n$, the* turnover vector $\boldsymbol{\tau}(\boldsymbol{\lambda})$ *corresponding to rebalancing $\boldsymbol{\lambda}$ to equality $\boldsymbol{u}_n$ is given by $\boldsymbol{\tau}(\boldsymbol{\lambda}) = \boldsymbol{\lambda} - \boldsymbol{u}_n$, and the resulting* turnover $\tau(\boldsymbol{\lambda})$ *is defined by $\tau(\boldsymbol{\lambda}) = \frac{1}{2}\sum_{i=1}^n |\tau_i|$, where τ_i are the components of the turnover vector $\boldsymbol{\tau}(\boldsymbol{\lambda})$.*

The turnover is intuitively equal to the portion of the total decision weights that would have to be redistributed by taking from weights exceeding $1/n$ and assigning these portions to weights that are less than $1/n$. The turnover hence always lies between 0 and 1. Graphically, it can be represented as the longest vertical distance between the Lorenz curve associated with a choice vector, and the diagonal line representing perfect equality. Note the similarities between Definition 10 and the *Hoover Index* (Hoover 1936), a measure of income metrics which is also known as the Robin Hood Index, as uniformity is achieved in a population by taking from the richer half and giving to the poorer half.

Lemma 4. *Let $\boldsymbol{\lambda} \in \mathbb{S}^n$ and $\Omega_{n,\boldsymbol{\lambda}} = \{P \in \mathbb{D}_n \mid \boldsymbol{u}_n = \boldsymbol{\lambda}P\}$. Then for all $P \in \Omega_{n,\boldsymbol{\lambda}}$,*

$$\boldsymbol{\lambda}(I_n - P) = \boldsymbol{\tau}(\boldsymbol{\lambda}) \ .$$

Proof. The equation follows by definition, as $\boldsymbol{\lambda}(I_n - P) = \boldsymbol{\lambda} - \boldsymbol{\lambda}P = \boldsymbol{\lambda} - \boldsymbol{u}_n = \boldsymbol{\tau}(\boldsymbol{\lambda})$. □

Based on Definition 10, every transformation $P \in \Omega_{n,\boldsymbol{\lambda}}$ applied to $\boldsymbol{\lambda}$ theoretically yields the same turnover. However, there is a subtle difference. In practice, some rebalancing transformations imply a higher practical turnover than the theoretical turnover of Definition 10. This is because more assets are bought or sold than is theoretically needed to obtain equality. In simple cases where there are only 2 or 3 possible choices, choosing a transformation that minimizes turnover is straightforward. However, for larger collections, the choice of the optimal rebalancing transformation may not be obvious.

We refer to the actual turnover induced in practice as the *practical turnover*.

Definition 11 (Practical turnover). *Let $\boldsymbol{\lambda} \in \mathbb{S}^n$. For $P \in \Omega_{n,\boldsymbol{\lambda}}$, the* practical turnover *is given by $\tilde{\tau}_P(\boldsymbol{\lambda}) = \tau(\boldsymbol{\lambda})\,\|P - I_n\|$, where $\|\cdot\|$ is the Frobenius norm taken up-to-permutation.*[8]

[8]For a $m \times n$ matrix $A = (a_{ij})$, the Frobenious norm is defined as $\|A\| = \sqrt{\sum_{i=1}^m \sum_{j=1}^{\prime} n|a_{ij}|^2}$.

The practical turnover is thus determined in terms of the distance of the corresponding rebalancing transform from the identity transform (up-to-permutation). The idea is that the closer one is to the identity transform, the smaller the changes that are applied to the entries of the choice vector.

Proposition 5. *Let* $\boldsymbol{\lambda} \neq \boldsymbol{u}_n \in \mathbb{S}^n$*. For* $P \in \Omega_{n,\boldsymbol{\lambda}} = \{P \in \mathbb{D}_n \mid \boldsymbol{u}_n = \boldsymbol{\lambda}P\}$*, denote by* $\tilde{\tau}(\boldsymbol{\lambda}) = \{\tilde{\tau}_P(\boldsymbol{\lambda}) \mid P \in \Omega_{n,\boldsymbol{\lambda}}\}$ *the set of all possible practical turnovers. Then*

$$\inf\left(\tilde{\tau}(\boldsymbol{\lambda})\right) = \tau(\boldsymbol{\lambda}) \ .$$

In other words, the smallest possible practical turnover is the theoretical turnover.

Proof. We will show that $\|P - I_n\| \geq 1$ for all $P \in \Omega_{n,\boldsymbol{\lambda}}$. Note that we obtain the smallest possible norm if all rows of P and I_n coincide up to permutation, except for two rows, say i and j. In other words, all entries of $\boldsymbol{\lambda}$ and $\boldsymbol{u}_n$ coincide (up to permutation) apart from the i-th and j-th entries that need to be averaged out to give $1/n$ each. Because P is a doubly stochastic matrix, the entries of both rows i and j must be some $a \in (0,1)$ and $1-a$. Consequently, $\|P - I_n\| = \sqrt{2a^2 + 2(1-a)^2}$ and its minimum is reached at $a = 1/2$, implying that the smallest possible norm is equal to $\|P - I_n\| = \sqrt{4(1/2)^2} = 1$. □

To characterize the rebalancing transform that would yield the theoretical turnover, and thus by Proposition 5 the smallest possible practical turnover, we use the notion of *T-transform* (Definition 3). Recall that in the economic context of equalizing wealth or income, T-transforms are also known as *Dalton* or *Robin Hood transfers* and are interpreted as the operation of shifting income or wealth from one individual to a relatively poorer individual. The following observation follows directly from the proof of Proposition 5.

Corollary 2. *Suppose one can transform* $\boldsymbol{\lambda} \in \mathbb{S}^n$ *to equality* $\boldsymbol{u}_n$ *directly through a single* T*-transform, i.e.* $T \in \Omega_{n,\boldsymbol{\lambda}}$*. Then* $\|T - I_n\| = 1$*.*

Also recall that according to Hardy, Littlewood, and Pólya (1934) (Proposition 1), if a vector $\boldsymbol{\alpha} \in \mathbb{S}^n$ is majorized by another vector $\boldsymbol{\beta} \in \mathbb{S}^n$, then $\boldsymbol{\alpha}$ can be derived from $\boldsymbol{\beta}$ by successive applications of at most $n-1$ such T-transforms. Therefore, every rebalancing polytope $\Omega_{n,\boldsymbol{\lambda}}$ contains (not necessarily unique) products of T-transforms. In Example **??**, $P(0,0)$ is itself a T-transform. Such successive applications of T-transforms do indeed produce the least possible turnover, that is the theoretical turnover. The following is an immediate consequence of the proof of Proposition 5 and the proof of Lemma 2, p.47 of Hardy, Littlewood, and Pólya (1934).

Proposition 6. *Let* $\boldsymbol{\lambda} \neq \boldsymbol{u}_n \in \mathbb{S}^n$*. Then*

$$\inf\left(\tilde{\tau}(\boldsymbol{\lambda})\right) = \tilde{\tau}_Q(\boldsymbol{\lambda}) \ ,$$

where $Q \in \Omega_{n,\boldsymbol{\lambda}}$ *is a product of at most* $n-1$ T*-transforms.*

Corollary 3. *For* $\boldsymbol{\lambda} \neq \boldsymbol{u}_n \in \mathbb{S}^n$ *and the rebalancing polytope* $\Omega_{n,\boldsymbol{\lambda}}$*, the minimum distance from identity* I_n *of any rebalancing transform* $P \in \Omega_{n,\boldsymbol{\lambda}}$ *is a product of* T*-transforms.*[9]

[9] Based on a private correspondence with the authors of Marshall, Olkin, and Arnold (2011), the problem of characterizing the closest element to an identity matrix within a given polytope has not been tackled in linear algebra. Our characterization through T-transforms can hence be of interest to mathematicians and economists working with inequalities and the theory of majorization in general.

References

Aldous, D. J. (1985): "Exchangeability and related topics," in *École d'Été de Probabilités de Saint-Flour XIII — 1983*, ed. by P. L. Hennequin, pp. 1–198, Berlin, Heidelberg. Springer Berlin Heidelberg.

Atkinson, A. B. (1970): "On the Measurement of Inequality," *Journal of Economic Theory*, 2, 244–263.

Baltussen, G., and T. Post (2011): "Irrational Diversification: An Examination of the Portfolio Construction Decision," *Journal of Financial and Quantitative Analysis*, 46(5).

Bayes, T. (1763): "An Essay Towards Solving a Problem in the Doctrine of Chances," *Philosophical Transactions of the Royal Society of London*, 53, 370–418.

Benartzi, S., and R. Thaler (2001): "Naive Diversification Strategies in Defined Contribution Saving Plans," *American Economic Review*, 91, 79–98.

Bernoulli, D. (1738): "Exposition of a New Theory on the Measurement of Risk," *Econometrica*, 22, 23–36.

Best, M. J., and R. R. Grauer (1991): "On the Sensitivity of Mean-Variance-Efficient Portfolios to Changes on Asset Means: Some Analytical and Computational Results," *The Review of Financial Studies*, 4, 315–342.

Birkhoff, G. (1946): "Three Observations on Linear Algebra," *Univ. Nac. Tacuman Rev. Ser. A*, 5, 147–151.

Blackorby, C., and D. Donaldson (1978): "Measures of Relative Equality and their Meaning in Terms of Social Welfare," *Journal of Economic Theory*, 18, 59–79.

Breen, W., L. R. Glosten, and R. Jagannathan (1989): "Economic Significance of Predictable Variations in Stock Index Returns," *The Journal of Finance*, 44, 1177–1189.

Cerreia-Vioglio, S., F. Maccheroni, M. Marinacci, and L. Montrucchio (2011): "Risk Measures: Rationality and Diversification," *Mathematical Finance*, 21(4), 743–774.

Chateauneuf, A., and G. Lakhnati (2007): "From Sure to Strong Diversification," *Economic Theory*, 32, 511–522.

Chateauneuf, A., and J.-M. Tallon (2002): "Diversification, Convex Preferences and Non-empty Core in the Choquet Expected Utility Model," *Economic Theory*, 19, 509–523.

Dalton, H. (1920): "On the Measurement of Inequality of Incomes," *The Economic Journal*, 30(119), 348–361.

De Giorgi, E. G., and O. Mahmoud (2016): "Diversification Preferences in the Theory of Choice," *Decisions in Economics and Finance*, 39(2), 143–174.

Dekel, E. (1989): "Asset Demands Without the Independence Axiom," *Econometrica*, 57, 163–169.

DeMarzo, P. M., D. Vayanos, and J. Zwiebel (2003): "Persuasion Bias, Social Influence and Unidimensional Opinions," *Quarterly Journal of Economics*, 118, 909–968.

DeMiguel, V., L. Garlappi, and R. Uppal (2007): "Optimal Versus Naive Diversification: How Inefficient is the $1/n$ Portfolio Strategy?," *The Review of Financial Studies*, 22, 1915–1953.

Duchin, R., and H. Levy (2009): "Markowitz Versus the Talmudic Portfolio Diversification Strategies," *Journal of Portfolio Management*, 35, 71–74.

Elton, E. J., and M. J. Gruber (1977): "Risk Reduction and Portfolio Size: An Analytic Solution," *Journal of Business*, 50, 415–437.

Eyster, E., and G. Weizsäcker (2011): "Correlation Neglect in Financial Decision Making," Working paper.

Fernandes, D. (2013): "The 1/N Rule Revisited: Heterogeneity in the Naive Diversification Bias," *International Journal of Marketing Research*, 30(3), 310–313.

Fishburn, P. C. (1970): *Utility Theory For Decision Making*. John Wiley & Sons.

Gigerenzer, G. (2010): *Rationality for Mortals: How People Cope with Uncertainty*. Oxford University Press.

Gini, C. (1921): "Measurement of Inequality of Income," *The Economic Journal*, 31(121), 124–126.

Glaeser, E., and C. R. Sunstein (2009): "Extremism in Social Learning," *Journal of Legal Analysis*, 1(1).

Grinblatt, M., and S. Titman (1989): "Mutual Fund Perfomance: An Analysis of Quarterly Portfolio Holdings," *The Journal of Business*, 62, 393–416.

Hadar, J., and W. R. Russell (1969): "Rules for Ordering Uncertain Prospects," *American Economic Review*, 59, 25–34.

——— (1971): "Stochastic Dominance and Diversification," *Journal of Economic Theory*, 3, 288–305.

Hamza, O., M. Kortas, J.-F. L'Her, and M. Roberge (2007): "International Equity Indices: Exploring Alternatives to Market-Cap Weighting," *Journal of Investing*, 16, 103–118.

Hardy, G. H., J. E. Littlewood, and G. Pólya (1929): "Some Simple Inequalities Satisfied by Convex Functions," *Messenger of Mathematics*, 58, 145–152.

——— (1934): *Inequalities.* Cambridge University Press.

Herstein, I. N., and J. Milnor (1953): "An Axiomatic Approach to Measurable Utility," *Econometrica*, 21(2), 291–297.

Hirschman, A. O. (1964): "The Paternity of an Index," *The American Economic Review*, 54(5).

Hodges, S. D., and R. A. Brealy (1978): "Portfolio Selection in a Dynamic and Uncertain World," in *Modern Developments in Investment Management*, ed. by J. H. Lorie, and R. A. Brealy. Dryden Press.

Hoover, E. M. (1936): "The Measurement of Industrial Localization," *Review of Economics and Statistics*, 18, 162–171.

Huberman, G., and W. Jiang (2006): "Offering vs. Choice in 401(k) Plans: Equity Exposure and Number of Funds," *Journal of Finance*, 61, 763–801.

Ibragimov, R. (2009): "Portfolio diversification and value at risk under thick-tailedness," *Quantitative Finance*, 9(5), 565–580.

Kallir, I., and D. Sonsino (2009): "The Neglect of Correlation in Allocation Decisions," *Southern Economic Journal*, 75(4), 1045–1066.

Korajczyk, R. A., and R. Sadka (2004): "Are Momentum Profits Robust to Trading Costs?," *The Journal of Finance*, 59, 1039–1082.

Krämer, W. (1998): "Measurement of Inequality," in *Handbook of Applied Economic Statistics*, ed. by A. Ullah, and D. E. A. Giles, pp. 39–61. Marcel Dekker, New York.

Kroll, Y., H. Levy, and A. Rapoport (1988): "Experimental Tests of the Separation Theorem and the Capital Asset Pricing Model," *American Economic Review*, 78(3), 500–519.

Lessard, D. R. (1976): "World, Country and Industry Relationships in Equity Returns," *Financial Analysts Journal*, 32, 32–41.

Levy, G., and R. Razin (2015): "Correlation Neglect, Voting Behavior and Information Aggregation," *American Economic Review*, 105, 1634–1645.

Li, C.-K., and W.-K. Wong (1999): "Extension of Stochastic Dominance Theory to Random Variables," *RAIRO Operations Research*, 33, 509–524.

Litterman, R. (2003): *Modern Investment Management: An Equilibrium Approach.* Wiley, New York.

Lorenz, M. O. (1905): "Methods of Measuring Concentration of Wealth," *Journal of the American Statistical Association*, 9, 209–219.

Maillard, S., T. Roncalli, and J. Teiletche (2010): "The Properties of Equally Weighted Risk Contribution Portfolios," *Journal of Portfolio Management.*

Markowitz, H. M. (1952): "Portfolio Selection," *Journal of Finance*, 7, 77–91.

Marshall, A. W., I. Olkin, and B. C. Arnold (2011): *Inequalities: Theory of Majorization and its Applications.* Springer.

Michaud, R. O. (1998): *Efficient Asset Management.* Harvard Business School Press.

Muirhead, R. F. (1903): "Some Methods Applicable to Identities and Inequalities of Symmetric Algebraic Functions of n Letters," *Proceedings of the Edinburgh Mathematical Society*, 21, 144–157.

Ohlson, J., and B. Rosenberg (1982): "Systematic Risk of the CRSP Equal-Weighted Common Stock Index: A History Estimated by Stochastic-Parameter Regression," *The Journal of Business*, 55, 121–145.

Ortoleva, P., and E. Snowberg (2015): "Overconfidence in Political Economy," *American Economic Review*, 105, 504–535.

Pae, Y., and N. Sabbaghi (2010): "Why Do Equally Weighted Portfolios Outeprform Value Weighted Portfolios?," Working Paper, Lewis University College of Business and Illinois Institute of Technology.

Pigou, A. C. (1912): *Wealth and Welfare.* Macmillan, New York.

Rachev, S. T., S. V. Stoyanov, and F. J. Fabozzi (2011): *A Probability Metrics Approach to Financial Risk Measures.* Wiley-Blackwell.

Rado, R. (1952): "An Inequality," *Journal of the London Mathematical Society*, 27, 1–6.

Read, D., and G. Loewenstein (1995): "Diversification Bias: Explaining the Discrepancy in Variety Seeking Between Combined and Separated Choices," *Journal of Experimental Psychology: Applied*, 1(1), 34–49.

Roll, R. (1981): "A Possible Explanation of the Small Firm Effect," *The Journal of Finance*, 36, 879–888.

Samuelson, P. (1967): "General Proof that Diversification Pays," *Journal of Financial and Quantitative Analysis*, 2, 1–13.

Schmeidler, D. (1979): "A Bibliographical Note on a Theorem of Hardy, Littlewood, and Polya," *Journal of Economic Theory*, 20, 125–128.

Schmeidler, D. (1989): "Subjective Probability and Expected Utility Without Additivity," *Econometrica*, 57(3), 571–587.

Sen, A. (1973): *On Economic Inequality*. Clarendon Press Oxford.

Simon, H. A. (1955): "A Behavioral Model of Rational Choice," *Quarterly Journal of Economics*, 69(1), 99–118.

——— (1979): "Rational Decision Making in Business Organizations," *American Economic Review*, 69(4), 493–513.

Simonson, I. (1990): "The Effect of Purchase Quantity and Timing on Variety-Seeking Behavior," *Journal of Marketing Research*, 27, 150–162.

Simspon, E. H. (1949): "Measurement of Diversity," *Nature*, 163, 688.

Szal, R., and S. Robinson (1977): "Measuring Income Inequality," in *Income Distribution and Growth in Less-Developed Countries*, ed. by C. R. Frank, and R. C. Webbs, pp. 491–533. Brookings Institute Washington.

Tesfatsion, L. (1976): "Stochastic Dominance and Maximization of Expected Utility," *Review of Economic Studies*, 43, 301–315.

Tu, J., and G. Zhou (2011): "Markowitz meets Talmud: A Combination of Sophisticated and Naive Diversifcation Strategies," *Journal of Financial Economics*, 99, 204–215.

Tversky, A., and D. Kahneman (1981): "The Framing of Decisions and the Psychology of Choice," *Science*, 211, 453–458.

von Neumann, J., and O. Morgenstern (1944): *Theory of Games and Economic Behavior*. Princeton University Press.