

The geometry of consumer preference aggregation^{*}

Fedor Sandomirskiy (Caltech)[†]

Philip Ushchev (ECARES, Université libre de Bruxelles)

November 9, 2022

the latest version: [link](#)

Abstract

This paper revisits a classical question in economics: how individual preferences and incomes shape aggregate behavior. We develop a method that reduces the hard problem of aggregation to simply computing a weighted average. The method applies to populations with homothetic preferences. The key idea is to handle aggregation in the space of logarithmic price indices.

We demonstrate the power of this method by (i) characterizing classes of preferences invariant with respect to aggregation, i.e., such that any population of heterogeneous consumers with preferences from the class behaves as if it were a single aggregate consumer from the same class; (ii) characterizing classes of aggregate preferences generated by popular preference domains such as linear or Leontief; (iii) describing indecomposable preferences, i.e., those that do not correspond to aggregate behavior of any non-trivial population; (iv) representing any preference as an aggregation of indecomposable ones.

We discuss connections and applications of our findings to stochastic discrete choice, information design, welfare analysis and gains from trade estimation, pseudo-market mechanisms, and preference identification.

^{*}The paper has benefited from our discussions with (in alphabetic order) Anna Bogomolnaia, Federico Echenique, Peter Caradonna, Tommaso Denti, Ben Golub, Bram De Rock, Peter Klibanoff, Alexey Kushnir, Arthur Lewbel, Kiminori Matsuyama, Hervé Moulin, Joseph Root, Roberto Serrano, Denis Shishkin, Ali Shourideh, Omer Tamuz, Aleh Tsyvinski, Xavier Vives, Richard Van Weelden, Adam Wierman, Kai Hao Yang, and comments by seminar participants at CMU, Northwestern, and University of Pittsburgh.

[†]Fedor thanks the Linde Institute at Caltech and National Science Foundation (grant CNS 1518941) for their support.

1 Introduction

Although economic decisions are made by individuals, it is often the aggregate behavior that matters. We examine the relation between individual and collective behavior in the context of consumer choice. Whenever one makes assumptions about incomes and preferences of individual consumers, it is important to understand the restrictions these assumptions impose on the aggregate demand. The reverse question is also relevant, especially for welfare analysis: if we know the aggregate demand, what can be said about the individual characteristics of consumers which could give rise to this aggregate demand? Since (Sonnenschein, 1973), the profession has been quite pessimistic regarding the extent to which the above questions can be answered. The following quote from (Kreps, 2020) provides an accurate summary of the “anything goes” consensus:

“So what can we say about aggregate demand based on the hypothesis that individuals are preference/utility maximizers? Unless we are able to make strong assumptions about the distribution of preferences or income throughout the economy (e.g., everyone has the same preferences) there is little we can say.”

Our paper argues that there is rich enough setting where aggregation questions are already non-trivial and more than “little” can be said. We develop a method for aggregate demand analysis applicable to populations of consumers with homothetic preferences. For example, what can be said about aggregate demand if all goods are perfect substitutes at the individual level, so that consumers have linear (but not identical) preferences? Or, can a CES aggregate demand with complements be obtained by aggregating Leontief preferences? Our method delivers precise and exhaustive answers to these questions among many others and has immediate applications to robust welfare analysis and the algorithmic complexity of markets.

The key idea of our method is strikingly simple: instead of looking at direct utilities of individual consumers, we look at logarithms of their price indices (logarithmic price indices in what follows).¹ We find that the allegedly hard problem of demand aggregation boils down to simply computing a weighted average of logarithmic price indices, with weights equal to relative incomes. This weighted average corresponds to a homothetic preference of the aggregate consumer whose demand coincides with the demand of the original population for all prices.²

¹For a given preference, the price index is a function of prices equal to the minimal budget that an agent with this preference needs to achieve the unit utility level.

²In contrast to the textbook definition of a representative consumer by Gorman (1961), the preference of the aggregate consumer is allowed to depend on income distribution. As a result, the aggregate consumer is well-defined for all populations with homothetic preferences (Eisenberg, 1961) while Gorman’s representative fails to exist for such populations unless all preferences are identical.

The problem of demand aggregation reduces to the problem of preference aggregation: understanding how the preference of the aggregate consumer depends on preferences and incomes of the underlying population. Our method provides insights into economics and geometry of this dependence.

- *Aggregation-invariant classes of preferences.* Consider a class of preferences, e.g., linear, Leontief, or any other subset of homothetic preferences. We refer to such classes as domains. Let us call a domain invariant with respect to aggregation if the aggregate preference always belongs to this domain. For example, the domain of all homothetic preferences, a domain containing just one preference, and the domain of Cobb-Douglas preferences are invariant. Invariant domains are important as in such domains the aggregate behavior is as simple as the behavior of a single agent.

We characterize all invariant domains by the convexity property of the set of associated logarithmic price indices. This characterization allows us to construct simple parametric invariant domains and describe the minimal invariant domains containing popular ones.

- *Characterization of possible aggregate behaviors for a given class of preferences.* Suppose we know a domain to which individual consumers' preferences belong. What can we say about the aggregate demand of such a population? This question boils down to understanding what preferences can be obtained by aggregation of individual preferences from this domain. We call the set of all such aggregate preferences the domain completion. The notion of completion is closely related to invariance: the completion of a domain is the minimal invariant collection of preferences containing the domain. The characterization of the invariance implies that the completion of a domain can be found by computing the convex hull of the set of logarithmic price indices.

We describe the completion explicitly for the domains of linear and Leontief preferences.³ A viable conjecture would be that the completion of linear preferences gives all preferences exhibiting substitutability among goods. We show that this guess is correct in the case of two goods only. For Leontief preferences, the completion turns out to be a proper subset of preferences with complementarity even for two goods.

- *Decomposition of preferences.* Consider the inverse to the problem of aggregation: given a preference from a particular domain, represent it as an aggregation of preferences from the same domain. There is always a trivial representation since we can take a population

³Surprisingly, this problem happens to be connected to several branches of economics and mathematics such as additive random utility models (ARUM), completely monotone functions, the Stieltjes transform, and even complex analysis.

where each agent has the same given preference. We call those preferences that can only be represented by themselves, indecomposable.

Geometrically, indecomposable preferences correspond to extreme points in the space of logarithmic price indices. As any point of a convex set can be represented as a convex combination of extreme points, indecomposable preferences play the role of elementary building blocks: any preference can be represented as an aggregation of indecomposable ones.

For example, linear and Leontief preferences are indecomposable in the domain of all homothetic preferences. We show that the set of indecomposable preferences is much bigger and contains all Leontief preferences on linear composite goods. In particular, aggregation of linear and Leontief preferences together does not give the whole domain of homothetic preferences. We also explore indecomposability in the domains of preferences with substitutability or complementarity.

We illustrate how our approach to aggregation can be applied in several economic environments.

- *Preference identification.* Given a domain of individual preferences, we ask whether observing the price dependence of market demand is enough to identify the distribution of preferences and income over the population. We relate the possibility of identification with the geometric simplex property of the domain meaning that there is a unique way to represent each preference as an aggregation of indecomposable ones. Examples of domains where identification is possible include Leontief and linear preferences over two goods. In settings similar to ours, identification has been previously obtained either for small populations, e.g., two-agent households as in (Chiappori, 1988) or under the assumption of preferences “orthogonality” (Chiappori and Ekeland, 2009). Our results do not restrict the size of populations and allow agents to have preferences that are close.
- *Robust welfare analysis.* An analyst observing aggregate behavior aims to estimate the population’s welfare which depends on individual preferences and incomes. As the same aggregate behavior can be compatible with different populations it can also be compatible with a range of values for the welfare functional. We show that this range can be computed by solving an auxiliary Bayesian persuasion problem. As a corollary, we obtain that the welfare of the aggregate consumer can be used as a proxy for the population’s welfare only for a narrow class of welfare functionals. This conclusion suggests a possible explanation for unexpectedly low gains from trade as measured in recent quantitative literature relying on a representative consumer as a proxy (Arkolakis et al., 2012, 2019).
- *Domain complexity, Fisher markets, and bidding languages for pseudo-market mechanisms.* For invariant domains, the aggregate behavior is as simple as that of a single agent. Since the

completion of a domain is the minimal invariant domain containing it, the completion reflects the complexity of aggregate behavior.

We formalize this intuition in application to Fisher markets: simple exchange economies where consumers with fixed incomes face a fixed supply of goods. Such markets are essential for the pseudo-market (or competitive) approach to fair allocation of resources (Moulin, 2019; Pycia, 2022) and serve as a benchmark model for equilibrium computation in algorithmic economics (Nisan et al., 2007). Computing an equilibrium of a Fisher market turns out to be a challenging problem even in a seemingly innocent case of linear preferences thus limiting the applicability of pseudo-market mechanisms. We explore the origin of the complexity and demonstrate that computing equilibria can be hard even in small parametric domains if their completion is large. We show how to construct domains with small completion and describe an algorithm making use of this smallness. The choice of a domain is interpreted as bidding language design.

The methodological importance of the link between aggregation and weighted averages is that it brings new tools — the most important of which are convexification and extreme-point techniques of Choquet theory — to consumer demand literature. These tools are increasingly popular in other branches of economic theory such as information economics and mechanism design (Kleiner et al., 2021); our paper demonstrates their power for the analysis of aggregate demand. We also uncover a connection between aggregation and modern literature in convex geometry on the geometric mean of convex sets (Milman and Rotem, 2017; Böröczky et al., 2012). This connection not only enables geometric tools in our economic problem but also leads to new insights about the geometric mean of convex sets suggested by the economic interpretation.

1.1 Related literature

As suggested by the quote from (Kreps, 2020), the existing results on demand aggregation have fallen into one of the two extremes. One extreme stems from the classical general equilibrium literature dealing with economies where agents have general convex preferences and earn money by trading their endowments. This literature concludes that the aggregate demand inherits no properties of individual behavior (Sonnenschein, 1973; Mantel, 1974; Debreu, 1974; Chiappori and Ekeland, 1999). The opposite extreme is given by the representative-agent literature aiming to replace the population with a single rational agent whose preferences are independent of the income distribution (Gorman, 1953, 1961). The independence requirement is so restrictive that it is fair to say that the representative almost never exists except for degenerate cases, e.g., the whole population having identical homothetic preferences. The profession has been divided on how seriously

the non-existence should be taken. Applied researchers often postulate the existence of a representative consumer (e.g., [Chamley, 1986](#); [Rogoff, 1990](#)) but this approach is criticized as lacking microfoundations (e.g., [Kirman, 1992](#); [Carroll, 2000](#)).

As demonstrated by [Jackson and Yariv \(2019\)](#), weakening Gorman’s requirements while maintaining income independence does not alter substantially the non-existence conclusion. We escape this conclusion by allowing the aggregate consumer to depend on the income distribution. The existence of such an aggregate consumer was pointed out by [Eisenberg and Gale \(1959\)](#) for populations with linear preferences and, by [Eisenberg \(1961\)](#), in the whole domain of homothetic preferences; see a survey by [Shafer and Sonnenschein \(1982\)](#).⁴ In the modern theoretical literature, this insight has gone largely unnoticed with the exception of algorithmic economics and fair allocation mechanisms; see Section 6.3. The converse statement to Eisenberg’s result was obtained by [Jerison \(1984\)](#) who showed that homotheticity is necessary for the existence.

The existence of income-dependent aggregate consumers confirms the common wisdom that the unrestricted endogeneity of incomes is more important for the negative Sonnenschein-Mantel-Debreu results than the generality of preferences ([Mantel, 1976](#); [Varian, 1984](#); [Jerison, 1994](#); [Hildenbrand, 2014](#)). The empirical literature has been using income-dependent aggregates since ([Jorgenson et al., 1982](#); [Deaton and Muellbauer, 1980](#); [Lewbel, 1989](#)).

Our results on robust welfare analysis are close in spirit to those of [Kang and Vasserman \(2022\)](#). The two papers address different aspects of robustness and are complementary. [Kang and Vasserman \(2022\)](#) derive robust predictions when the number of distinct observations of the aggregate demand is not enough to pin it down but assume that the aggregate demand itself is a sufficient statistic for welfare. By contrast, we develop the robust approach if the same aggregate demand is compatible with a range of welfare levels corresponding to different populations generating the same observed behavior. Curiously, both papers rely on auxiliary Bayesian persuasion problems of different forms and origins.

2 Preliminaries

This section is about notation and basic concepts from consumer demand theory.

Notation. We use \mathbb{R} for the set of all real numbers, \mathbb{R}_+ and \mathbb{R}_- for non-negative/non-positive ones, and \mathbb{R}_{++} and \mathbb{R}_{--} for strictly positive/negative ones. Ratios of the form $t/0$ with $t \geq 0$ are

⁴[Eisenberg and Gale \(1959\)](#) were motivated by the question of probabilistic forecast aggregation and introduced an auxiliary exchange economy of bets, a “prediction market” in modern terms. A closely related idea for belief aggregation is known in the financial-market literature as Negishi approach; see ([Jouini and Napp, 2007](#); [Calvet et al., 2001](#)) and references therein.

assumed to be equal $+\infty$.

Bold font is used for vectors, e.g., $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$. For a pair of vectors of the same dimension, we write $\mathbf{x} \geq \mathbf{y}$ if the inequality holds component-wise, i.e., $x_i \geq y_i$ for all i . The scalar product of $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ is denoted by $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{j=1}^n x_j \cdot y_j$.

For subsets of \mathbb{R}^n , multiplication by a scalar and summation are defined element-wise: $\alpha \cdot X = \{\alpha \cdot \mathbf{x} : \mathbf{x} \in X\}$ and $X + Y = \{\mathbf{x} + \mathbf{y} : \mathbf{x} \in X, \mathbf{y} \in Y\}$ (the Minkowski sum of sets). The standard $(n-1)$ -dimensional simplex is denoted by $\Delta_{n-1} = \{\mathbf{x} \in \mathbb{R}_+^n : x_1 + \dots + x_n = 1\}$.

The gradient of a function $f = f(\mathbf{x})$ is the vector of its partial derivatives $\nabla f = \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} \right)$.

Preferences and demand. Consider a consumer who is endowed with a budget $b \in \mathbb{R}_{++}$ and has a preference \succsim over vectors $\mathbf{x} \in \mathbb{R}_+^n$ interpreted as bundles of $n \geq 1$ divisible goods. We assume that preferences satisfy the following standard requirements:

- **homotheticity:** $\mathbf{x}' \succsim \mathbf{x}$ implies $\alpha \cdot \mathbf{x}' \succsim \alpha \cdot \mathbf{x}$ for any $\alpha \geq 0$
- **convexity:** for any $\mathbf{x}, \mathbf{x}' \in \mathbb{R}_+^n$ between which the consumer is indifferent, $\lambda \mathbf{x} + (1-\lambda)\mathbf{x}' \succsim \mathbf{x}$ for any $\lambda \in [0, 1]$
- **monotonicity:** if $\mathbf{x}, \mathbf{x}' \in \mathbb{R}_+^n$ and $\mathbf{x}' \geq \mathbf{x}$, then $\mathbf{x}' \succsim \mathbf{x}$
- **continuity:** for any $\mathbf{x}, \mathbf{x}' \in \mathbb{R}_+^n$ and a convergent sequence $\mathbf{z}^{(l)}$, $l = 1, 2, 3, \dots$, such that $\mathbf{x}' \succ \mathbf{z}^{(l)} \succ \mathbf{x}$, we have $\mathbf{x}' \succ \lim_{l \rightarrow \infty} \mathbf{z}^{(l)} \succ \mathbf{x}$
- **non-degeneracy:** there exist $\mathbf{x}, \mathbf{x}' \in \mathbb{R}_{++}^n$ such that the consumer is not indifferent between them.

For brevity, we will refer to all preferences satisfying these assumptions as homothetic preferences.

Given a vector of prices $\mathbf{p} \in \mathbb{R}_{++}^n$, the budget set of the consumer is the set of affordable bundles: $\{\mathbf{x} \in \mathbb{R}_+^n : \langle \mathbf{p}, \mathbf{x} \rangle \leq b\}$. The demand of the consumer consists of her most preferred bundles from the budget set

$$D(\mathbf{p}, b) = \underset{\mathbf{x} \in \mathbb{R}_+^n : \langle \mathbf{p}, \mathbf{x} \rangle \leq b}{\operatorname{argmax}} \succsim.$$

The demand is a non-empty closed convex subset satisfying homogeneity with respect to budgets: $D(\mathbf{p}, b) = b \cdot D(\mathbf{p}, 1)$. It is a singleton (one-element set) for almost all \mathbf{p} which allows us to think of the demand as a single-valued function of \mathbf{p} defined almost everywhere.

Representations of preferences. We will use several ways to represent homothetic preferences.

Any homothetic preference \succsim can be represented by a utility function $u = u_{\succsim}(\mathbf{x})$ so that $u(\mathbf{x}) \geq u(\mathbf{x}')$ if and only if $\mathbf{x} \succsim \mathbf{x}'$. This utility function can be selected to be continuous, non-decreasing,

concave, homogeneous ($u(\alpha \cdot \mathbf{x}) = \alpha \cdot u(\mathbf{x})$ for all bundles \mathbf{x} and $\alpha \geq 0$), non-negative, and not identically zero. Utility functions satisfying all these requirements are called homogeneous in what follows. Any homogeneous utility function defines a homothetic preference and each homothetic preference pins down a unique homogeneous utility function up to a multiplicative factor.

A homothetic preference is determined by its upper contour set $\{\mathbf{x} \in \mathbb{R}_+^n : u(\mathbf{x}) \geq 1\}$. A set $X \subset \mathbb{R}^n$ is called upward-closed if $\mathbf{x} \in X$ implies that all vectors $\mathbf{x}' \in \mathbb{R}^n$ such that $\mathbf{x}' \geq \mathbf{x}$ also belong to X . The upper contour set is a closed convex subset of \mathbb{R}_+^n that does not contain 0, and is upward-closed. Any set with these properties corresponds to a homothetic preference. Hence, we can use such sets as another representation for homothetic preferences keeping in mind that, for a given preference, the set is defined up to a homothetic transformation inheriting the freedom in the choice of the multiplicative factor in the utility function.

A dual representation of preferences through price indices will be the most convenient for our analysis. For a consumer with a preference \succsim , the price index $P = P_{\succsim}(\mathbf{p})$ is defined by

$$P(\mathbf{p}) = \min_{\mathbf{x} \in \mathbb{R}_+^n : u(\mathbf{x}) \geq 1} \langle \mathbf{p}, \mathbf{x} \rangle. \quad (1)$$

The price index is the minimal budget that the consumer needs to achieve the unitary utility level, i.e., $P(\mathbf{p})$ is consumer's individual estimate of the "cost of living" given the prices. Geometrically, P is equal to the support function of the upper contour set $\{\mathbf{x} \in \mathbb{R}_+^n : u(\mathbf{x}) \geq 1\}$ up to a sign. A price index $P: \mathbb{R}_+^n \rightarrow \mathbb{R}$ is continuous, non-decreasing, concave, homogeneous, non-negative, and not identically equal to zero. Conversely, any function with these properties is a price index for some homothetic preference. We get yet another way of representing homothetic preferences. Similarly to utility functions, the price index for a given preference is defined up to a multiplicative factor.

For homothetic preferences, Roy's identity becomes

$$D(\mathbf{p}, b) = b \cdot \nabla \ln(P(\mathbf{p})), \quad (2)$$

i.e., the demand is proportional to the gradient of the logarithm of the price index (*logarithmic price index* in what follows). The identity holds for those prices $\mathbf{p} \in \mathbb{R}_{++}^n$ where P is differentiable which coincides with the set of \mathbf{p} where the demand is a singleton. As any concave function is differentiable almost everywhere, this set has full measure.

Consider the budget share function $\mathbf{s} = \mathbf{s}_{\succsim}(\mathbf{p})$ whose i -th component $s_i(\mathbf{p})$ is the fraction of the budget that the consumer spends on good $i = 1, \dots, n$ given the prices, i.e.,

$$s_i(\mathbf{p}) = p_i \cdot \frac{D_i(\mathbf{p}, b)}{b} = p_i \cdot D_i(\mathbf{p}, 1). \quad (3)$$

We treat \mathbf{s} as a single-valued vector function taking values in the standard simplex Δ_{n-1} and defined on the set of $\mathbf{p} \in \mathbb{R}_{++}^n$ of full measure where the demand is a singleton. By (2), budget shares can

be computed as the elasticities of the price index with respect to prices

$$s_i(\mathbf{p}) = p_i \cdot \frac{\partial \ln(P(\mathbf{p}))}{\partial p_i} = \frac{\partial \ln(P(\mathbf{p}))}{\partial \ln(p_i)}. \quad (4)$$

For two goods, preferences can be represented via budget share functions using the following characterization.⁵ For any homothetic preference \succsim over \mathbb{R}_+^2 , the budget share of the first good takes the form

$$s_1(p_1, p_2) = \frac{1}{1 + Q\left(\frac{p_1}{p_2}\right) / \frac{p_1}{p_2}} \quad (5)$$

for some non-decreasing non-negative function $Q: \mathbb{R}_{++} \rightarrow \mathbb{R}_+ \cup \{+\infty\}$. Moreover, for any such function Q , there is a unique homothetic preference; see Lemma 5 in Appendix B.6.

By plugging a function Q with an infinite number of jumps in (5), we see that, rather counter-intuitively, $s_1(p_1, p_2)$ may change monotonicity infinitely many times as p_1 increases, i.e., the consumer starts spending more on the first good as its price goes up, then less, then more again, and so on; an explicit example of such preferences can be found in Section 5.1.

Substitutes and complements. The two important subdomains of homothetic preferences are free from non-monotone behavior of budget shares described above.

A preference \succsim is said to exhibit substitutability among the goods if the budget share $s_i(\mathbf{p})$ is a non-decreasing function of p_j for each pair of goods $i \neq j$. For differentiable budget share functions,

$$\frac{\partial s_i(\mathbf{p})}{\partial p_j} > 0 \quad \text{for all } i \neq j. \quad (6)$$

The intuition is that whenever the price of a good increases, the consumer starts spending more on other goods since this good can be substituted. The canonical example is given by linear preferences \succsim that correspond to utility functions

$$u(\mathbf{x}) = \langle \mathbf{v}, \mathbf{x} \rangle$$

for some vector of values $\mathbf{v} \in \mathbb{R}_+^n \setminus \{0\}$. An elementary computation gives the price index and formula (2) provides budget shares

$$P(\mathbf{p}) = \min_{i=1, \dots, n} \frac{p_i}{v_i} \quad \text{and} \quad s_i(\mathbf{p}) = \begin{cases} 1, & \text{if } \frac{v_i}{p_i} > \frac{v_j}{p_j} \text{ for all } j \neq i, \\ 0, & \text{otherwise} \end{cases}. \quad (7)$$

As we see, under linear preferences, the consumer spends her whole budget on the good with the highest value-to-price ratio.

⁵To the best of our knowledge, this characterization has not appeared in the literature.

A preference \succsim exhibits complementarity among goods if $s_i(\mathbf{p})$ is a non-increasing function of p_j for each pair of goods $i \neq j$. Each of the complementary goods is essential for consumer's satisfaction and so, when one good becomes more expensive, more money is spent on it and less on other goods. The standard example is given by the Leontief preferences which correspond to the following utility function

$$u(\mathbf{x}) = \min_{i=1, \dots, n} \frac{x_i}{v_i}$$

for some vector of values $\mathbf{v} \in \mathbb{R}_+^n \setminus \{0\}$. Note that the utility function has the same functional form as the price index for linear preferences. By duality, the price index for Leontief preferences is linear

$$P(\mathbf{p}) = \langle \mathbf{v}, \mathbf{p} \rangle \quad \text{and} \quad s_i(\mathbf{p}) = \frac{v_i \cdot p_i}{\langle \mathbf{v}, \mathbf{p} \rangle}. \quad (8)$$

The intersection of the domains of preferences exhibiting substitutability and complementarity consists of those preferences \succsim for which budget shares are constant, i.e., there is a fixed vector $\mathbf{a} \in \Delta_{n-1}$ such that $\mathbf{s}(\mathbf{p}) = \mathbf{a}$ for any \mathbf{p} . Budget shares determine the price index by (2)

$$P(\mathbf{p}) = \prod_{i=1}^n p_i^{a_i} \quad (9)$$

and the corresponding preference is given by the Cobb-Douglas utility function

$$u(\mathbf{x}) = \prod_{i=1}^n x_i^{a_i}. \quad (10)$$

Leontief, Cobb-Douglas, and linear preferences are contained as limit cases in a widely used parametric family of preferences with constant elasticity of substitution (CES). A preference \succsim is a CES preference with elasticity of substitution $\sigma \in \mathbb{R}_{++} \setminus \{1\}$ if the corresponding utility function has the form

$$u(\mathbf{x}) = \left(\sum_{i=1}^n (a_i \cdot x_i)^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}} \quad (11)$$

for some vector $\mathbf{a} \in \mathbb{R}_{++}^n$. The corresponding price indices and budget shares are given by

$$P(\mathbf{p}) = \left(\sum_{i=1}^n \left(\frac{p_i}{a_i} \right)^{1-\sigma} \right)^{\frac{1}{1-\sigma}} \quad \text{and} \quad s_i(\mathbf{p}) = \frac{\left(\frac{p_i}{a_i} \right)^{1-\sigma}}{\sum_{j=1}^n \left(\frac{p_j}{a_j} \right)^{1-\sigma}}. \quad (12)$$

CES preferences exhibit substitutability for $\sigma > 1$ and complementarity for $\sigma \in (0, 1)$. Leontief, Cobb-Douglas, and linear preferences are the limiting cases as σ goes, respectively, to 0, 1, and $+\infty$. The limits are taken with respect to the topology that we discuss next.

Topology on preferences. Convergence of preferences, closed and open sets, and the Borel structure are understood with respect to the following metric. We define the distance between preferences \succsim and \succsim' with price indices P and P' by

$$d(\succsim, \succsim') = \sup_{\mathbf{p} \in \Delta_{n-1} \cap \mathbb{R}_{++}^n} \left| \frac{(\ln P(\mathbf{p}) - \ln P(\mathbf{e})) - (\ln P'(\mathbf{p}) - \ln P'(\mathbf{e}))}{(1 + \max_i |\ln p_i|)^2} \right|, \quad (13)$$

where $\mathbf{e} = (1, \dots, 1)$. The main advantage of this way to introduce the distance is that it makes the set of all homothetic preferences a compact metric space. In particular, the distance between any pair of preferences is finite and bounded by 2. See Appendix A for the intuition behind the definition and the detailed discussion.

3 Preference aggregation

Consider $m \geq 1$ consumers $k = 1, \dots, m$. Consumer k has a positive budget $b_k \in \mathbb{R}_{++}$ and a homothetic preference \succsim_k over bundles of $n \geq 1$ divisible goods as in Section 2. For any vector of prices \mathbf{p} , this population generates the market demand equal to the sum of individual demands $D_1(\mathbf{p}, b_1) + \dots + D_m(\mathbf{p}, b_m)$. To study the market demand, we aim to replace the population of m consumers with a single aggregate consumer generating the same demand. The following definition plays the central role in this methodology.

Definition 1. A preference $\succsim_{\text{aggregate}}$ is the aggregate preference for a population of consumers with preferences $\succsim_1, \dots, \succsim_m$ and budgets b_1, \dots, b_m if

$$D_{\text{aggregate}} \left(\mathbf{p}, \sum_{k=1}^m b_k \right) = D_1(\mathbf{p}, b_1) + \dots + D_m(\mathbf{p}, b_m) \quad (14)$$

for any price vector $\mathbf{p} \in \mathbb{R}_{++}^n$. A consumer with preference $\succsim_{\text{aggregate}}$ is referred to as the aggregate consumer.

In other words, the market demand generated by the population of consumers coincides with the demand of the aggregate consumer endowed with the total budget. We stress that the aggregate consumer is selected for a given collection of budgets b_1, \dots, b_m of individual consumers, and so, for a different distribution of incomes over the population, we may end up with a different aggregate consumer. This is an important distinction between Definition 1 and the approach of Gorman (1961) who insist on independence of the aggregate preference on the income distribution which can be achieved in knife-edge cases only.

Example 1. Consider $m = n$ consumers with degenerate linear preferences: consumer i 's utility is $u_i(\mathbf{x}) = x_i$, i.e., she is interested in the good i only. Hence, no matter what the prices are,

consumer i spends her total budget b_i on good i . This observation helps to guess the aggregate consumer without any computations. Indeed, the aggregate consumer spends the amount b_i out of her total budget $b_1 + \dots + b_n$ on good i independently of prices. In other words, the budget share of each good i for the aggregate consumer is price-independent and equal to $s_{\text{aggregate},i}(\mathbf{p}) = b_i/(b_1 + \dots + b_n)$. Hence, the aggregate consumer must have the Cobb-Douglas preferences (10) with $a_i = b_i/(b_1 + \dots + b_n)$. One can verify this guess directly by checking that the demand identity (14) holds. Alternatively, the result can be deduced immediately from Theorem 1 below and explicit formulas for price indices of Cobb-Douglas and linear preferences.

The existence of an aggregate preference was established by [Eisenberg \(1961\)](#) for any population of consumers with homothetic preferences. Denote by B the total budget of the population and by β_k the relative fraction of consumer k 's budget

$$B = \sum_{k=1}^m b_k, \quad \beta_k = \frac{b_k}{B}. \quad (15)$$

[Eisenberg \(1961\)](#) showed that the aggregate preference corresponds to the utility function obtained as the solution to the following optimization problem

$$u_{\text{aggregate}}(\mathbf{x}) = \max \left\{ \prod_{k=1}^m \left(\frac{u_k(\mathbf{x}_k)}{\beta_k} \right)^{\beta_k} : \mathbf{x}_k \in \mathbb{R}_+^n, \quad k = 1, \dots, m, \quad \sum_{k=1}^m \mathbf{x}_k = \mathbf{x} \right\}. \quad (16)$$

In other words, the utility for the aggregate preference at a bundle \mathbf{x} is equal to the maximal weighted Nash social welfare where the maximum is taken over all possible allocations of \mathbf{x} over the consumers and consumer's weight is equal to her relative budget.⁶ The optimization problem (16) is called the Eisenberg-Gale problem as it is similar to a problem studied by [Eisenberg and Gale \(1959\)](#) in the context of probabilistic forecast aggregation.

To determine the utility of an aggregate consumer, one needs to solve the Eisenberg-Gale problem (16) for each $\mathbf{x} \in \mathbb{R}_+^n$. Except for special cases such as Cobb-Douglas preferences, it does not admit an explicit solution and is not easy to work with both analytically and computationally; see Section 6.3.

We observe that the question of describing the aggregate consumer substantially simplifies if we use the dual representation of preferences via price indices.

Theorem 1. *Consider a population of consumers with homothetic preferences $\succsim_1, \dots, \succsim_m$ and budgets b_1, \dots, b_m . The preference of the aggregate consumer is described by the price index $P_{\text{aggregate}}$ satisfying*

$$\ln P_{\text{aggregate}}(\mathbf{p}) = \sum_{k=1}^m \beta_k \cdot \ln P_k(\mathbf{p}), \quad (17)$$

⁶The welfare function equal to the product of consumer's utilities is dubbed the Nash social welfare or the Nash product as this welfare function naturally arises in the context of axiomatic bargaining studied by [Nash \(1950\)](#).

where the weights β_k are given by (15).

Hence, preference aggregation is equivalent to taking convex combinations of individual logarithmic price indices. The simplicity of this operation will allow us to describe domains invariant with respect to aggregation (Section 4) and to study decomposition of a given preference as an aggregation of elementary ones (Section 5).

The identity (17) becomes almost immediate if we take into account the relation (2) between the demand and the gradient of a price index: $D(\mathbf{p}, b) = b \cdot \nabla \ln(P(\mathbf{p}))$. The definition of the aggregate consumer implies the equality

$$B \cdot \nabla \ln P_{\text{aggregate}}(\mathbf{p}) = \sum_{k=1}^m b_k \cdot \nabla \ln P_k(\mathbf{p}) \quad (18)$$

which must hold at all points of differentiability of the price indices. As any concave function is differentiable almost everywhere with respect to the Lebesgue measure, the equality (18) holds on the set of full measure and can be integrated resulting in the identity (17). Integration constants get absorbed by the price indices since they are defined up to multiplicative factors.

In Appendix B.1, we prove Theorem 1 using an approach similar to the one used by Eisenberg (1961) and not relying on formula (2). This alternative proof clarifies that Theorem 1 is the dual to Eisenberg's result.

3.1 Connection to the geometric mean of convex sets

Theorem 1 links preferences aggregation and recent attempts to define the geometric mean of convex sets; see a survey by Milman and Rotem (2017). The following concept was proposed by Böröczky et al. (2012): the weighted geometric mean of convex sets X and Y with weights $(\lambda, 1-\lambda)$, $\lambda \in [0, 1]$, is defined as the convex set $Z = X^\lambda \otimes Y^{1-\lambda}$ such that the logarithm of its support function P_Z is equal to the convex combination of logarithms of support functions of X and Y :

$$\ln P_{X^\lambda \otimes Y^{1-\lambda}} = \lambda \cdot \ln P_X + (1 - \lambda) \cdot \ln P_Y. \quad (19)$$

In other words, the weighted geometric mean of convex sets is defined by taking the usual weighted geometric mean in the space of support functions.^{7,8} The weighted geometric mean extends to any number of convex sets in a straightforward manner.

⁷Böröczky et al. (2012) refer to $X^\lambda \otimes Y^{1-\lambda}$ as the logarithmic sum of convex sets to distinguish it from other definitions of the geometric mean. Since we do not consider other definitions, we will call $X^\lambda \otimes Y^{1-\lambda}$ the weighted geometric mean.

⁸Defining algebraic operations on convex sets through the standard algebraic operations on their support functions is a standard approach. For example, the Minkowski addition of convex sets corresponds to pointwise summation of their support functions.

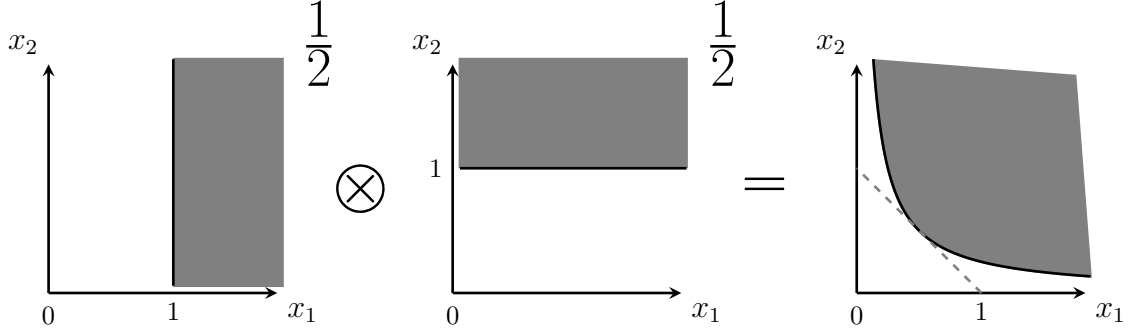


Figure 1: **Geometry:** the set bounded by the hyperbola is the geometric mean of the two orthogonal halfspaces. **Economics:** an aggregation of the two extreme linear preferences where each consumer cares only about her own good gives a Cobb-Douglas preference.

To see the connection between preference aggregation and the geometric mean, recall that the price index is the support function of the upper contour set. We obtain the equivalent version of Theorem 1.

Corollary 1. *An upper contour set of the aggregate consumer's preferences $\{u_{\text{aggregate}}(\mathbf{x}) \geq 1\}$ is the weighted geometric mean of individual upper contour sets with budget-proportional weights:*

$$\{u_{\text{aggregate}}(\mathbf{x}) \geq 1\} = \{u_1(\mathbf{x}) \geq 1\}^{\beta_1} \otimes \{u_2(\mathbf{x}) \geq 1\}^{\beta_2} \otimes \dots \otimes \{u_m(\mathbf{x}) \geq 1\}^{\beta_k}. \quad (20)$$

In Example 1, we saw that Cobb-Douglas preferences over n goods originate as an aggregation of n extreme linear preferences. Figure 1 illustrates the corresponding identity of convex sets for $n = 2$ and equal budgets.

Corollary 1 highlights a peculiar property of the class of convex sets that can be obtained as upper contour sets of homothetic preferences. From formula (19), it is not evident that the logarithmic mean is well-defined, i.e., that we can always find a convex set whose support function is equal to $P_{X \wedge Y^{1-\lambda}}$. A byproduct of Corollary 1 is that the weighted geometric mean is well-defined within the class of all convex subsets of \mathbb{R}_+^n that do not contain zero and are upward-closed. Indeed, any such set is an upper contour set of some homothetic preference and the weighted geometric mean is an upper contour set of the aggregate consumer's preference. Contrast this observation with the case of bounded convex sets which mathematical literature has mostly focused on. The weighted geometric mean is well-defined for bounded convex sets containing zero; however, sets that do not contain zero are problematic as the support function can be negative and the definition of the weighted geometric mean requires ad hoc modifications.

4 Invariant domains

In Section 3, we saw that a population of consumers with homothetic preferences can be replaced with a single aggregate consumer. Here we study domains of homothetic preferences invariant with respect to aggregation: if each consumer's preference belongs to the domain, so does the aggregate preference. Tools developed in the previous section reduce invariance to convexity of the set of logarithmic price indices in a functional space and yield a flexible procedure for constructing invariant domains.

Definition 2. *A domain \mathcal{D} of homothetic preferences over \mathbb{R}_+^n is invariant with respect to aggregation if for any $m \geq 2$ and any population of m consumers with preferences $\succsim_k \in \mathcal{D}$ and budgets $b_k \in \mathbb{R}_{++}$, $k = 1, \dots, m$, the aggregate preference $\succsim_{\text{aggregate}}$ also belongs to \mathcal{D} .*

The set of all homothetic preferences or a domain containing just one preference $\mathcal{D} = \{\succsim\}$ are elementary examples of invariant domains.

Note that it is enough to check the condition of invariance for populations of $m = 2$ consumers. Indeed, aggregation for a population of $m > 2$ consumers reduces to aggregation for pairs by adding consumers one by one sequentially. Hence, if the outcome of aggregation belongs to the domain for any pair, the outcome will belong to this domain for any population.

With a domain \mathcal{D} , we associate the set $\mathcal{L}_{\mathcal{D}}$ of all logarithmic price indices of preferences from \mathcal{D}

$$\mathcal{L}_{\mathcal{D}} = \left\{ f : \mathbb{R}_{++}^n \rightarrow \mathbb{R} \quad : \quad f = \ln P_{\succsim}, \quad \succsim \in \mathcal{D} \right\}.$$

The set $\mathcal{L}_{\mathcal{D}}$ inherits the freedom in the choice of price indices: if $f \in \mathcal{L}_{\mathcal{D}}$, then $f + \text{const} \in \mathcal{L}_{\mathcal{D}}$ and corresponds to the same preference. The following result is a direct corollary of Theorem 1.

Corollary 2. *A domain \mathcal{D} is invariant with respect to aggregation if and only if the set of logarithmic price indices $\mathcal{L}_{\mathcal{D}}$ is a convex set of functions on \mathbb{R}_{++}^n .*

In other words, \mathcal{D} is invariant whenever, for any pair of preferences $\succsim', \succsim'' \in \mathcal{D}$ with price indices P' and P'' and $\lambda \in (0, 1)$, the preference \succsim corresponding to the price index P defined by

$$\ln P = \lambda \cdot \ln P' + (1 - \lambda) \cdot \ln P'' \tag{21}$$

also belongs to \mathcal{D} .

For example, the domain of Cobb-Douglas preferences (10) satisfies the requirement (21) and, hence, is invariant. The domains of preferences exhibiting substitutability or complementarity are also invariant. Indeed, budget shares can be obtained by differentiating logarithmic price indices (4) and so the monotonicity conditions defining these domains are preserved under convex combinations.

Corollary 2 not only characterizes invariant domains in geometric terms but also gives a handy tool to construct invariant domains containing a given one. Suppose \mathcal{D} is not invariant. How to complete it to an invariant domain? Of course, \mathcal{D} is contained in the domain of all homothetic preferences which is invariant. To exclude such a trivial answer we require the completion be minimal with respect to set inclusion.

Definition 3. *For a domain \mathcal{D} , its completion $\mathcal{D}^{\text{complete}}$ is the minimal closed domain that is invariant with respect to aggregation and contains \mathcal{D} .*

Closure is defined with respect to the metric structure (13) on preferences. The closedness assumption helps to get a tractable answer for infinite domains and, essentially, means that $\mathcal{D}^{\text{complete}}$ is enriched by aggregate preferences of non-atomic populations with preferences from \mathcal{D} .

The completion $\mathcal{D}^{\text{complete}}$ exists since it can be obtained as the intersection of all closed invariant domains containing \mathcal{D} and there is at least one such domain, namely, the domain of all homothetic preferences. Corollary 2 implies a geometric characterization of $\mathcal{D}^{\text{complete}}$.

Corollary 3. *For any subdomain \mathcal{D} of homothetic preferences, its completion $\mathcal{D}^{\text{complete}}$ is equal to the set of all preferences corresponding to logarithmic price indices from the closed convex hull of $\mathcal{L}_{\mathcal{D}}$:*

$$\mathcal{D}^{\text{complete}} = \left\{ \succsim : \ln(P_{\succsim}) \in \text{conv}[\mathcal{L}_{\mathcal{D}}] \right\},$$

where $\text{conv}[X]$ denotes the minimal closed convex set containing X .

This corollary assumes that the choice of the topology on preferences is aligned with that on logarithmic price indices. This requirement is satisfied by the topology from Appendix A.

Note that $\text{conv}[X]$ can be obtained as the closure of the set of all convex combinations of finite collections of elements from X . For finite subdomains $\mathcal{D} = \{\succsim_1, \dots, \succsim_q\}$, looking at combinations of at most $q = |\mathcal{D}|$ elements is enough and, hence, Corollary 3 is especially easy to apply. For such \mathcal{D} , the completion $\mathcal{D}^{\text{complete}}$ consists of all preferences \succsim with price indices of the form $\ln P(\mathbf{p}) = \sum_{k=1}^q t_k \cdot \ln P_k(\mathbf{p})$ with $\mathbf{t} \in \Delta_{q-1}$. Reinterpreting Example 1, we conclude that Cobb-Douglas preferences over n goods is the completion of $\mathcal{D} = \{\succsim_1, \dots, \succsim_n\}$ where \succsim_i corresponds to the utility function $u_i(\mathbf{x}) = x_i$.

To compute the completion for infinite subdomains \mathcal{D} , we need to take closure of the set of preferences \succsim corresponding to all finite convex combinations of logarithmic price indices

$$\ln P(\mathbf{p}) = \sum_{k=1}^q t_k \cdot \ln P_k(\mathbf{p}),$$

where $q \geq 1$, a vector $\mathbf{t} \in \Delta_q$, and P_k represents some preference \succsim_k from \mathcal{D} . It is convenient to think about this convex combination as a result of integration with respect to the atomic distribution

μ placing weight t_k on preference \succsim_k :

$$\ln P(\mathbf{p}) = \int_{\mathcal{D}} \ln P_{\succsim'}(\mathbf{p}) \, d\mu(\succsim'). \quad (22)$$

It turns out that taking closure is equivalent to allowing arbitrary probability measures μ in (22), not necessarily atomic. For parametric domains such as linear or Leontief preferences discussed below, the integral in (22) can be seen as the integral over the space of parameters and, hence, passing to an arbitrary μ is straightforward. In Appendix A, we explain how to define (22) for any domain \mathcal{D} and measure μ .

Theorem 2. *The completion of a domain \mathcal{D} consists of all preferences \succsim such that their price index P can be represented as*

$$\ln P(\mathbf{p}) = \int_{\overline{\mathcal{D}}} \ln P_{\succsim'}(\mathbf{p}) \, d\mu(\succsim') \quad (23)$$

with some Borel probability measure μ supported on the closure of $\overline{\mathcal{D}}$ of \mathcal{D} .

With general μ , representation (23) can be interpreted as the result of preference aggregation where non-atomic populations are allowed and μ plays the role of preference distribution over the population. In what follows, we refer to (23) as continuous aggregation.

A generalization of Theorem 2 is proved in Appendix B.2 together with Theorem 3 formulated in the next section. Both results rely on Choquet theory which studies extreme points of compact convex sets in topological vector spaces (Phelps, 2001). Application of this theory requires careful choice of a topology and a measurable structure. For the proof to work, it is crucial that the sets of preferences and logarithmic price indices endowed with the distance (13) are compact and admit an isometric embedding in a Banach space.

4.1 ARUM and completion of linear preferences

Consider the domain \mathcal{D} of all linear preferences. Our goal is to characterize its completion $\mathcal{D}^{\text{complete}}$. This problem turns out to be tightly related to stochastic discrete choice theory. The intuition behind this connection is as follows. By Theorem 2, finding the completion boils down to taking average of logarithmic price indices with respect to some measure μ . This average can be thought as expectation over random preferences of a single decision maker and budget shares can be interpreted as probabilities of choosing one of n possible alternatives.

In the additive random utility model (ARUM), there is a single decision maker choosing between one of n alternatives. Her utility for alternative i is equal to $w_i + \varepsilon_i$, where w_i is a deterministic component and ε_i is a stochastic shock. The vector $\mathbf{w} \in \mathbb{R}^n$ and the joint distribution of shocks $(\varepsilon_1, \dots, \varepsilon_n) \in \mathbb{R}^n$ are given. For each realization of the shocks, the agent selects the alternative with

the highest utility. Hence, the expected utility of the decision maker and the probability that she chooses alternative i are equal to⁹

$$U(\mathbf{w}) = \mathbb{E} \left[\max_{i=1, \dots, n} (w_i + \varepsilon_i) \right] \quad \text{and} \quad S_i(\mathbf{w}) = \mathbb{P}[\{w_i + \varepsilon_i > w_j + \varepsilon_j \ \forall j \neq i\}],$$

where \mathbb{E} and \mathbb{P} denote the expectation and the probability with respect to the shock distribution.

Proposition 1. *A preference \succsim with price index P belongs to the completion of the domain of all linear preferences over n goods if and only if there is a distribution of shocks such that*

$$U(\mathbf{w}) = -\ln(P(e^{-w_1}, \dots, e^{-w_n})) \quad (24)$$

is the expected utility in ARUM with deterministic utilities $\mathbf{w} \in \mathbb{R}^n$.

Taking the gradient on both sides of (24) gives a version of the statement for budget shares:¹⁰ \succsim is in the completion of linear preferences whenever $\mathbf{s}(e^{-w_1}, \dots, e^{-w_n})$ is the vector of choice probabilities for some additive random utility model, i.e., there exists a distribution of shocks such that

$$s_i(e^{-w_1}, \dots, e^{-w_n}) = \mathbb{P}[\{w_i + \varepsilon_i > w_j + \varepsilon_j \ \forall j \neq i\}] \quad (25)$$

for all $i = 1, \dots, n$ and Lebesgue almost all $\mathbf{w} \in \mathbb{R}^n$.

Substituting formula (7) for price index of linear preferences into Theorem 2, we see that the completion of linear preferences consists of all preferences \succsim whose price indices P can be represented as

$$\ln P(\mathbf{p}) = \int_{\mathbb{R}_+^n} \ln \left(\min_{i=1, \dots, n} \frac{p_i}{v_i} \right) d\mu(\mathbf{v}) \quad (26)$$

for some measure μ such that the integral converges. To get (24), it is enough to change variables in (26). Denote $\varepsilon_i = \ln v_i$ and interpret $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)$ as a random vector by assuming that \mathbf{v} is sampled from distribution μ . Plugging in $p_i = e^{-w_i}$, we get $-\ln \left(\min_{i=1, \dots, n} \frac{p_i}{v_i} \right) = \max_{i=1, \dots, n} (w_i + \varepsilon_i)$ and, hence, (26) is equivalent to

$$-\ln(P(e^{-w_1}, \dots, e^{-w_n})) = \mathbb{E} \left[\max_{i=1, \dots, n} (w_i + \varepsilon_i) \right].$$

⁹The formula for the choice probabilities holds for those \mathbf{w} for which the probability of a tie $w_i + \varepsilon_i = w_j + \varepsilon_j$ is zero. This is the case for Lebesgue almost all \mathbf{w} no matter what the distribution of the shocks is.

¹⁰The fact that the choice probabilities $S_i(\mathbf{w})$ can be obtained as partial derivatives of decision maker's utility is known as the Williams–Daly–Zachary theorem and its classic version requires regularity assumptions on the distribution of shocks (McFadden, 1981). The possibility to drop all the assumptions and get the conclusion for Lebesgue almost all \mathbf{w} is a recent result (Sørensen and Fosgerau, 2022). The connection between ARUM and an aggregation of linear preferences make this result a corollary of general formula (4) expressing budget shares as the gradient of logarithmic price indices for almost all prices.

As the right-hand side has the form of the expected utility in ARUM, we obtain Proposition 1.

The class of vector-functions that can arise as choice probabilities $\mathbf{S}(\mathbf{w})$ for some ARUM is well-studied in the discrete choice theory. We need the following necessary condition applicable to smooth vector-functions. For any ARUM with n alternatives and any subset of distinct alternatives i, j_1, j_2, \dots, j_q with $q \leq n - 1$, the following inequality holds

$$\frac{\partial^q S_i(\mathbf{w})}{\partial w_{j_1} \partial w_{j_2} \dots \partial w_{j_q}} \cdot (-1)^q \leq 0$$

at any \mathbf{w} where \mathbf{S} is q times differentiable. Taking into account the connection between budget shares and choice probabilities (25) and the identity $\frac{\partial}{\partial w_i} = -\frac{\partial}{\partial \ln p_i}$ for $p_i = e^{-w_i}$, we obtain the following corollary of Proposition 1.

Corollary 4. *If a preference \succsim belongs to the completion of the domain of all linear preferences, then its budget shares satisfy the following inequalities*

$$\frac{\partial s_i(\mathbf{p})}{\partial \ln p_{j_1} \partial \ln p_{j_2} \dots \partial \ln p_{j_q}} \geq 0 \quad (27)$$

for any distinct goods i, j_1, j_2, \dots, j_q with $q \leq n - 1$ at any price vector $\mathbf{p} \in \mathbb{R}_{++}^n$ where \mathbf{s} is differentiable q times.

For $q = 1$, the condition (27) becomes the substitutability condition (6). In other words, any preference \succsim from the completion of linear preferences exhibits substitutability among goods. This conclusion is not surprising as linear preferences exhibit substitutability and aggregation respects this property.

One could expect that any preference exhibiting substitutability is in the completion of linear preferences. However, for $n \geq 3$ goods, the condition (27) gives extra restrictions on top of substitutability by restricting the second-order derivatives.

Corollary 5. *For $n \geq 3$ goods, the completion of the domain of all linear preferences is a proper subset of the domain of preferences exhibiting substitutability.*

The corollary tells nothing about the case of $n = 2$ goods which turns out to be an exception.

Proposition 2. *For $n = 2$ goods, the completion of the domain of all linear preferences coincides with the set of all preferences exhibiting substitutability.*

This result follows from an explicit construction. Given \succsim such that the budget share of the first good $s_1(p_1, p_2) = \frac{\partial \ln P(p_1, p_2)}{\partial \ln p_1}$ is non-decreasing in p_2 , we need to find a distribution μ of value vectors \mathbf{v} so that the continuous aggregation (26) of linear preferences leads to \succsim .

To guess an explicit formula for such μ , take the partial derivative $\frac{\partial}{\partial \ln p_1}$ on both sides of (26):

$$s_1(p_1, p_2) = \mu \left(\left\{ \frac{v_1}{v_2} \geq \frac{p_1}{p_2} \right\} \right). \quad (28)$$

The derivative exists and the identity holds for Lebesgue almost all (p_1, p_2) . The ratio $\text{MRS} = v_1/v_2$ is the marginal rate of substitution for the corresponding linear preference. We conclude that $1 - s_1(\cdot, 1)$ must be the cumulative distribution function of MRS and the monotonicity of s_1 makes this possible. Choosing any such distribution μ and adding atoms of the weight $1 - \lim_{p_1 \rightarrow +0} s_1(p_1, 1)$ at $\mathbf{v} = (0, 1)$ and of the weight $\lim_{p_1 \rightarrow \infty} s_1(p_1, 1)$ at $\mathbf{v} = (1, 0)$ completes the construction.

Note that we pinned down the distribution of the $\text{MRS} = v_1/v_2$ but not the magnitude of \mathbf{v} . As \mathbf{v} and $\lambda \cdot \mathbf{v}$ with $\lambda > 0$ correspond to the same linear preference, the distribution of preferences over the population is determined uniquely and we are free to choose any normalization of \mathbf{v} so that the integral in (26) converges, e.g., we can assume that μ is supported on $v_1 + v_2 = 1$.

Corollary 6. *Any preference over two goods exhibiting substitutability can be represented as a continuous aggregation of linear preferences (26). The distribution of linear preferences over the population corresponding to \succsim is pinned down uniquely and admits an explicit formula: the cumulative distribution function of the marginal rate of substitution $\text{MRS} = v_1/v_2$ equals $1 - s_1(\cdot, 1)$.*

Example 2 (translog preferences and Benford's law). To illustrate Corollary 6, let us show how a family of consumers with linear preferences over two goods can generate translog preference, a popular class of homothetic preferences obtained as a perturbation of Cobb-Douglas preferences in the space of price indices (Diewert (1974), p. 139). A preference \succsim is translog if its logarithmic price index has the following form

$$\ln P(p_1, p_2) = \begin{cases} \ln p_1, & \ln \left(\frac{p_1}{p_2} \right) < -\frac{1-\alpha}{\beta} \\ \alpha \ln p_1 + (1-\alpha) \ln p_2 - \frac{\beta}{2} \left(\ln \frac{p_1}{p_2} \right)^2, & -\frac{1-\alpha}{\beta} \leq \ln \left(\frac{p_1}{p_2} \right) \leq \frac{\alpha}{\beta} \\ \ln p_2, & \ln \left(\frac{p_1}{p_2} \right) > \frac{\alpha}{\beta} \end{cases},$$

where $\alpha \in (0, 1)$ and $\beta > 0$.

By elementary computations and formula (28), we obtain that a distribution of value vectors $\mathbf{v} = (v_1, v_2)$ aggregates up to the translog preference if and only if the logarithm of the marginal rate of substitution $\text{MRS} = v_1/v_2$ is distributed uniformly:

$$\ln \text{MRS} \sim \mathbb{U} \left(\left[-\frac{1-\alpha}{\beta}, \frac{\alpha}{\beta} \right] \right),$$

where $\mathbb{U}([c, d])$ denotes the uniform distribution supported on $[c, d]$. Curiously enough, this is equivalent to MRS following the so-called Benford law of digit bias (Benford, 1938).

Consider a particular case of translog preference \succsim with $\alpha = \beta = 1/2$. We obtain that \succsim can be represented as aggregation over the continuous population of consumers distributed uniformly in $[-1, 1]$ so that consumer $\xi \in [-1, 1]$ has utility $u(x_1, x_2) = e^{\frac{\xi}{2}} \cdot x_1 + e^{-\frac{\xi}{2}} \cdot x_2$. The corresponding identity (26) takes the following form:

$$\ln P(p_1, p_2) = \int_{-1}^1 \ln \left(\min \left\{ p_1 \cdot e^{-\frac{\xi}{2}}, p_2 \cdot e^{\frac{\xi}{2}} \right\} \right) d\xi.$$

4.2 Complete monotonicity and the completion of Leontief preferences

We saw that, for $n = 2$ goods, the completion of all linear preferences is the whole domain of homothetic preferences with substitutability. By contrast, the completion of all Leontief preferences turns out to be substantially narrower than the domain of all preferences with complementarity, even for $n = 2$.

By Theorem 2 and formula (8) for price indices of Leontief preferences, the completion of Leontief preferences over $n \geq 2$ goods is the set of all preferences \succsim with price index P of the following form:

$$\ln P(\mathbf{p}) = \int_{\mathbb{R}_+^n} \ln \langle \mathbf{v}, \mathbf{p} \rangle d\mu(\mathbf{v}) \quad (29)$$

for some probability measure μ on \mathbb{R}_+^n such that the integral converges.

Note that $\ln \langle \mathbf{v}, \mathbf{p} \rangle$ is an infinitely smooth function of $\mathbf{p} \in \mathbb{R}_{++}^n$. Exchanging integration and differentiation in (29), we conclude that the left-hand side must also be infinitely smooth. Thus the completion of Leontief preferences cannot contain preferences with price indices and budget shares that are not infinitely smooth.

Corollary 7. *For any number $n \geq 2$ of goods, the completion of the domain of Leontief preferences is a proper subset of the domain of preferences exhibiting complementarity.*

The following example provides a concrete preference over two complements that is outside of the completion of Leontief preferences.

Example 3 (A preference over two complements outside of the completion of Leontief). We aim to find a preference \succsim over $n = 2$ complements such that its price index P is not infinitely smooth. It is enough to find a preference such that the budget share of the first good $s_1(p_1, p_2) = \frac{\partial \ln P(p_1, p_2)}{\partial \ln p_1}$ has a discontinuous derivative. The existence of such \succsim follows from the characterization of budget shares (5). We will describe \succsim explicitly.

The idea is to combine two distinct preferences exhibiting complementarity so that the consumer's preference alternates between the two depending on prices. Consider \succsim corresponding to the following utility function:

$$u(x_1, x_2) = \min \{ \sqrt{x_1 \cdot x_2}, x_1 \}. \quad (30)$$

A consumer with this preference behaves as if she switches between Cobb-Douglas and Leontief preferences at $p_1 = p_2$. A simple computation gives the budget share:

$$s_1(p_1, p_2) = \begin{cases} \frac{1}{2}, & \frac{p_1}{p_2} < 1 \\ \frac{p_1}{p_1 + p_2}, & 1 \leq \frac{p_1}{p_2} \end{cases}.$$

As we see, s_1 is decreasing in p_2 and, hence, \succsim exhibits complementarity. However, this preference cannot be obtained as a continuous aggregation of Leontief preferences (29) since the budget share has discontinuous derivative.¹¹

The condition that a preference \succsim belongs to the completion of Leontief preferences is substantially more restrictive than the requirement of smoothness of the price index. An infinitely smooth function $f = f(\lambda)$, $\lambda \in \mathbb{R}_{++}$, is called completely monotone if¹²

$$(-1)^k \cdot \frac{d^k}{d\lambda^k} f \geq 0 \quad \text{for all } k = 0, 1, 2, \dots$$

Complete monotonicity provides a necessary condition for \succsim to be in the completion of Leontief preferences.

Proposition 3. *If a preference \succsim belongs to the invariant closure of Leontief preferences, then the price-normalized budget share $s_i(\mathbf{p})/p_i$ of a good i is a completely monotone function of p_i for each good $i = 1, \dots, n$.*

This proposition follows from the integral representation (29) of the price index P and the possibility to exchange differentiation with respect to p_i and integration. The derivatives of the integrand in (29) can be computed explicitly

$$\frac{\partial^{k+1}}{\partial^k p_i} \ln \langle \mathbf{v}, \mathbf{p} \rangle = (-1)^k \frac{v_i^k \cdot k!}{(\langle \mathbf{v}, \mathbf{p} \rangle)^{k+1}}.$$

Hence,

$$(-1)^k \frac{\partial^{k+1}}{\partial^k p_i} \ln P(\mathbf{p}) = k! \int_{\mathbb{R}_+^n} \frac{v_i^k}{(\langle \mathbf{v}, \mathbf{p} \rangle)^{k+1}} d\mu(\mathbf{v}) \geq 0.$$

Since $s(\mathbf{p})/p_i = \frac{\partial}{\partial p_i} \ln P(\mathbf{p})$, we conclude that $s_i(\mathbf{p})/p_i$ is a completely monotone functions of p_i .

¹¹For complements, budget shares can have discontinuous derivatives but are necessarily continuous themselves. This is a simple corollary of (5). By contrast, budget shares for substitutes can be discontinuous, e.g., for linear preferences budget shares are step functions.

¹²In economics, completely monotone functions arise as the dependence of decision maker's payoff on the discount factor λ . Indeed, by Bernstein's theorem a function f is completely monotone if and only if $f(\lambda) = \int_{\mathbb{R}_+} e^{-\lambda t} d\nu(t)$ for some positive measure ν , i.e., f is the expected utility of a risk-neutral decision maker with geometric discounting for a stream of payoff ν .

For $n = 2$ goods, we are able to provide a simple criterion for a preference \succsim to be in the completion of Leontief preferences. This criterion suggests that the necessary condition of complete monotonicity established in Proposition 4.2 is almost sufficient.

A function $f = f(\lambda)$, $\lambda \in \mathbb{R}_{++}$, is called a Stieltjes function if it can be represented as

$$f(\lambda) = \int_{\mathbb{R}_+} \frac{1}{\lambda + z} d\nu(z) \quad (31)$$

for some positive measure ν on \mathbb{R}_+ .¹³ The Stieltjes functions are exactly those completely monotone functions that themselves can be obtained as the Laplace transform of a completely monotone density.¹⁴

Proposition 4. *For $n = 2$ goods, a preference \succsim belongs to the completion of Leontief preferences if and only if the price-normalized budget share of the first good $s_1(p_1, 1)/p_1$ is a Stieltjes function of the price p_1 .*

If \succsim is in the completion, it is easy to see that it is a Stieltjes function. By (29), for any such preference we have

$$\frac{\partial}{\partial p_1} \ln P(p_1, 1) = \int_{\mathbb{R}_+^2} \frac{v_1}{v_1 p_1 + v_2} d\mu(v_1, v_2) = \int_{\mathbb{R}_+} \frac{1}{p_1 + z} d\nu(z),$$

where ν is the distribution of $\frac{v_2}{v_1}$. Since $\frac{s_1(p_1, 1)}{p_1} = \frac{\partial}{\partial p_1} \ln P(p_1, 1)$, we get

$$\frac{s_1(\lambda, 1)}{\lambda} = \int_{\mathbb{R}_+} \frac{1}{\lambda + z} d\nu(z) \quad (32)$$

and conclude that $s_1(\lambda, 1)/\lambda$ is a Stieltjes function for any \succsim from the completion.

The right-hand side of (32) is the Stieltjes transform of ν . The Stieltjes transform is invertible. Hence, if \succsim belongs to the completion of Leontief preferences, the budget shares determine the distribution ν satisfying (32) uniquely. As a result, the distribution of Leontief preferences over the population that leads to \succsim is pinned down uniquely. Namely, the continuous aggregation of Leontief preferences (29) with distribution μ of (v_1, v_2) such that the ratio v_2/v_1 is distributed according to ν gives \succsim . Leontief preferences with the same ratio coincide and so the distribution of preferences corresponding to \succsim is indeed unique.

The Stieltjes transform can be inverted explicitly using tools from complex analysis. Before describing the tools, we give an example obtained with their help.

¹³These functions are omnipresent in various branches of mathematics such as probability theory, spectral theory, continued fractions, and potential theory (Schilling et al., 2012, Chapter 2).

¹⁴The integral operator on the right-hand side of (31) is known as the Stieltjes transform. It equals the Laplace transform applied to ν twice: $f(\lambda) = \int_{\mathbb{R}_+} e^{-\lambda t} \left(\int_{\mathbb{R}_+} e^{-tz} d\nu(z) \right) dt$. By Bernstein's theorem, the set of completely monotone functions is the set of all functions equal to the Laplace transform of positive measures. Hence, Stieltjes functions are those completely monotone functions that are Laplace transforms of completely monotone ones.

Example 4 (CES with complements as an aggregation of Leontief preferences). We show that any CES preference \succsim over $n = 2$ complements (11) belongs to the completion of Leontief preferences.

First, consider a particular case with the elasticity of substitution $\sigma = \frac{1}{2}$ and weights $\mathbf{a} = (1, 1)$; the corresponding utility function is the harmonic mean. The utility and the budget share of the first good are as follows:

$$u(x_1, x_2) = \left(\frac{1}{x_1} + \frac{1}{x_2} \right)^{-1} \quad \text{and} \quad s_1(p_1, p_2) = \frac{\sqrt{p_1}}{\sqrt{p_1} + \sqrt{p_2}}.$$

By Proposition 4, finding a probability distribution ν on \mathbb{R}_+ such that the identity (32) holds is enough to show that \succsim is in the completion of Leontief preferences. We end up with the following equation:

$$\frac{1}{\lambda + \sqrt{\lambda}} = \int_{\mathbb{R}_+} \frac{1}{\lambda + z} d\nu(z).$$

One can check that ν with a density φ given by

$$\varphi(z) = \frac{1}{\pi} \frac{1}{\sqrt{z}(1+z)} \quad (33)$$

is a solution, hence, \succsim is indeed in the completion. By taking any distribution μ of $\mathbf{v} = (v_1, v_2)$ such that v_2/v_1 is ν -distributed (e.g., v_1 equals 1 identically and v_2 has distribution ν), we represent \succsim via a continuous aggregation of Leontief preferences (29).

The above analysis extends to any CES preference over two complements. The utility function and the budget share have the form

$$u(x_1, x_2) = \left((a_1 \cdot x_1)^{\frac{\sigma-1}{\sigma}} + (a_2 \cdot x_2)^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}} \quad \text{and} \quad s_1(p_1, p_2) = \frac{(p_1)^{1-\sigma}}{(p_1)^{1-\sigma} + \left(\frac{a_1}{a_2} p_2 \right)^{1-\sigma}},$$

where $\sigma \in (0, 1)$. The corresponding distribution ν of v_2/v_1 has to solve the equation

$$\frac{1}{\lambda + \left(\frac{a_1}{a_2} \right)^{1-\sigma} \cdot \lambda^\sigma} = \int_{\mathbb{R}_+} \frac{1}{\lambda + z} d\nu(z).$$

One can check that ν with density

$$\varphi(z) = \frac{\sin(\pi\sigma)}{\pi} \left(\frac{1}{\left(\frac{a_2}{a_1} \right)^{1-\sigma} \cdot z^{2-\sigma} + z \cdot \cos(\pi\sigma) + \left(\frac{a_1}{a_2} \right)^{1-\sigma} \cdot z^\sigma} \right) \quad (34)$$

is a solution. Formula (33) is a particular case of (34) for $\sigma = 1/2$ and $a_1 = a_2$.

Formulas (33) and (34) were derived using the following observation from complex analysis. For any distribution ν on \mathbb{R}_+ , its Stieltjes transform is defined not only for $\lambda \in \mathbb{R}_{++}$ but also for all

complex values of $\lambda \in \mathbb{C} \setminus \mathbb{R}_-$, where \mathbb{C} denotes the complex plane. Moreover, the function is analytic on $\mathbb{C} \setminus \mathbb{R}_-$. The values of this analytic continuation above and below the “cut” over the negative reals can be used to reconstruct ν . The answer is given by the Stieltjes-Perron formula: if f is the Stieltjes transform of a measure ν with density φ , then

$$\varphi(z) = \frac{1}{2\pi i} \cdot \lim_{\varepsilon \rightarrow 0} (f(-z + i\varepsilon) - f(-z - i\varepsilon)), \quad (35)$$

where i is the imaginary unit and ε tends to zero from above.

Combining the Stieltjes-Perron formula and Proposition 4, we get the following corollary.

Corollary 8. *If a preference \succsim over $n = 2$ goods belongs to the completion of Leontief preferences, then the budget share of the first good $s_1(p_1, 1)$ as a function of its price p_1 admits an analytic continuation to complex prices $p_1 \in \mathbb{C} \setminus \mathbb{R}_+$. The function φ given by (35) for*

$$f(\lambda) = \frac{s_1(\lambda, 1)}{\lambda}$$

is the density of the unique distribution of v_2/v_1 such that the continuous aggregation of Leontief preferences (29) gives \succsim .

Note that the analytic continuation is unique if exists. Hence, Corollary 8 can be used to check whether a given preference belongs to the completion of Leontief preferences. First we check whether the budget share admits analytic continuation. If it does, we compute a candidate for the distribution ν via the Stieltjes-Perron formula. Finally, we check that what we got is a probability distribution and the budget share can be obtained as its Stieltjes transform. A preference passes the test if and only if it is in the completion. Example 4 illustrated this approach.

5 Indecomposable preferences

In this section, we study those preferences that cannot be represented as an aggregation of distinct preferences withing a given domain. We call such preferences indecomposable. They play the role of elementary building blocks as any preference can be represented as an aggregation of indecomposable ones.

We already saw an example of such a representation in Section 4, when represented any preference over two substitutes as a continuous aggregation of linear preferences. In contrast to the discussion of Sections 3 and 4 where we started from specifying “elementary” preferences and asked what can be obtained by aggregating them, now we start from a given domain and aim to identify these elementary preferences.

Definition 4. For a given domain \mathcal{D} , a preference \succsim from \mathcal{D} is indecomposable if it cannot be represented as an aggregation of two distinct preferences \succsim' and \succsim'' from \mathcal{D} . The set of all indecomposable preferences from \mathcal{D} is denoted by $\mathcal{D}^{\text{indec}}$.

Recall that a point x from a subset X of a linear space is called an extreme point of X if it cannot be represented as $\alpha x' + (1 - \alpha)x''$ with $\alpha \in (0, 1)$ and distinct¹⁵ $x', x'' \in X$. The set of all extreme points of X is denoted by X^{extrem} . Theorem 1 implies the following corollary.

Corollary 9. A preference \succsim is indecomposable in \mathcal{D} if and only if the corresponding logarithmic price index $\ln P$ is an extreme point of the set of logarithmic price indices

$$\mathcal{L}_{\mathcal{D}} = \{f = \ln(P_{\succsim'}) : \succsim' \in \mathcal{D}\}.$$

The Choquet theorem states that, if X is a compact convex subset of a locally convex topological vector space, then any point $x \in X$ can be obtained as the average of its extreme points $x' \in X^{\text{extrem}}$ with respect to some Borel probability measure $\mu = \mu_x$ supported on X^{extrem} :

$$x = \int_{X^{\text{extrem}}} x' d\mu(x'); \quad (36)$$

see (Phelps, 2001). Using the Choquet theorem, we obtain the following result demonstrating that indecomposable preferences can indeed be seen as elementary building blocks.

Theorem 3. If \mathcal{D} is a closed domain invariant with respect to aggregation, then any preference \succsim from \mathcal{D} can be obtained as a continuous aggregation of indecomposable preferences from \mathcal{D} , i.e., there exists a Borel measure μ supported on $\mathcal{D}^{\text{indec}}$ such that the price index $P = P_{\succsim}$ can be represented as follows

$$\ln P(\mathbf{p}) = \int_{\mathcal{D}^{\text{indec}}} \ln P_{\succsim'}(\mathbf{p}) d\mu(\succsim') \quad (37)$$

for any vector of prices $\mathbf{p} \in \mathbb{R}_{++}^n$.

As in Theorem 2, the integral (37) is formally defined in Appendix A. Both theorems are proved in Appendix B.2. The essence of the proof is checking that the topological assumptions of the Choquet theorem are satisfied.

Representation (37) is especially useful if the set of indecomposable preferences is small relative to the whole domain \mathcal{D} . We will see that this is the case for substitutes but not the case for complements and the full domain.

¹⁵Usually, one assumes that X is convex but we do not make this assumption.

5.1 Indecomposability in the full domain

Let \mathcal{D} be the domain of all homothetic preferences. It is easy to guess some indecomposable preferences from \mathcal{D} : for example, linear and Leontief preferences are indecomposable. It turns out that there are many more. Let us call \succsim a Leontief preference over linear composite goods if it corresponds to a utility function of the form

$$u(\mathbf{x}) = \min_{\mathbf{a} \in A} \{\chi_{\mathbf{a}}(\mathbf{x})\}, \quad (38)$$

where A is a finite or countably infinite subset of \mathbb{R}_+^n and each $\mathbf{a} \in A$ defines a linear composite good $\chi_{\mathbf{a}}(\mathbf{x})$ by

$$\chi_{\mathbf{a}}(\mathbf{x}) = \sum_{i=1}^n a_i x_i.$$

The interpretation is that an agent treats the collection of bundles $\mathbf{a} \in A$ as perfect complements. Leontief and linear preferences are particular cases. For Leontief preferences, the bundles are, in fact, single goods and so each $\mathbf{a} \in A$ has only one non-zero coordinate. Linear preferences correspond to a single bundle \mathbf{a} , i.e., $A = \{\mathbf{a}\}$. Geometrically, Leontief preferences over linear composite goods are exactly those preferences that have upper contour sets with piecewise linear boundary.

Proposition 5. *For any number of goods n , Leontief preferences over linear composite goods (38) are indecomposable in the domain of all homothetic preferences.*

The proposition implies that linear and Leontief preferences are indeed indecomposable. Another immediate corollary is that aggregation of linear and Leontief preferences together is far from giving the full domain. Any preference of the form (38) is indecomposable and, hence, cannot be represented as an aggregation of linear or Leontief preferences unless it is linear or Leontief itself. For example, one can take a preference \succsim corresponding to

$$u(\mathbf{x}) = \min\{x_1 + 2 \cdot x_2, 2 \cdot x_1 + x_2\}.$$

The corollary can be strengthened. Budget shares for \succsim are not monotone, i.e., \succsim exhibits neither substitutability nor complementarity. Since \succsim is indecomposable, we conclude that not every preference can be represented as an aggregation of preferences exhibiting substitutability and preferences exhibiting complementarity.

We see that the full domain has a lot of indecomposable preferences. To formalize this observation, note that piecewise linear concave functions are dense in the set of all concave functions. Accordingly, indecomposable preferences are dense in the full domain \mathcal{D} and extreme points of the set $\mathcal{L}_{\mathcal{D}}$ of logarithmic price indices are dense in this set.¹⁶

¹⁶The existence of non-trivial convex sets with dense extreme points highlights that finite-dimensional intuition

The main insight behind Proposition 5 is as follows. We know that describing indecomposable preferences in a domain \mathcal{D} boils down to finding extreme points of the set of logarithmic price indices $\mathcal{L}_{\mathcal{D}}$. Finite-dimensional linear programming intuition suggests that natural candidates for extreme points of a convex set are those points where the maximal number of constraints defining the set are active. Leontief preference over linear composite goods are those preferences \succeq for which the concavity constraint on the price index P is active almost everywhere.

The formal proof of Proposition 5 is contained in Appendix B.3; we sketch the argument here. A utility function u is of the form (38) if and only if the corresponding price index is also piecewise linear: $P = \min_{c \in C} \sum_{i=1}^n c_i \cdot p_i$ for finite or countable $C \subset \mathbb{R}_+^n$. To demonstrate indecomposability, we need to show that if $P = (P_1)^\alpha (P_2)^{1-\alpha}$, then P_1 and P_2 are proportional to each other (and thus to P). By strict concavity of the function $h(\mathbf{t}) = t_1^\alpha \cdot t_2^{1-\alpha}$ on rays not passing through the origin, P cannot be linear in regions where P_1 and P_2 are not proportional. Hence, P_1 and P_2 must be proportional in each of the linearity regions of P . As these regions cover the whole space, P_1 and P_2 are proportional everywhere implying that \succeq is indecomposable.

In Appendix B.3 we also explore how close Proposition 5 is to characterizing all indecomposable preferences. We show that if there is a neighborhood of a point where the concavity constraint on P is inactive, then a preference can be decomposed (Proposition 13). The idea is that we can find small perturbation $\psi = \psi(\mathbf{p})$ vanishing outside of this neighborhood and such that $P_1 = P \cdot (1 + \psi)$ and $P_2 = P / (1 + \psi)$ are valid logarithmic price indices. Since $\ln P = 1/2 \cdot \ln P_1 + 1/2 \cdot \ln P_2$, the preference corresponding to P can indeed be decomposed.

Intuitively, a concave function is either piecewise linear or there is a neighborhood where it is strictly concave and, hence, Propositions 5 and 13 seem to cover all possible cases. However, there is a family of pathological examples not captured by this intuition, e.g., concave functions whose second derivative is a continuous measure supported on a Cantor set. Proposition 14 formulated and proved in the appendix shows that such pathological preferences are also indecomposable.

Proposition 5 has implications for the geometric mean of convex sets. Consider the collection \mathcal{X} of all closed convex subsets X of \mathbb{R}_+^n that do not contain zero and are upward-closed, i.e., all those that can be obtained as upper contour sets of homothetic preferences. We call a set $X \in \mathcal{X}$ indecomposable if it cannot be represented as the geometric mean $X_1^\lambda \otimes X_2^{1-\lambda}$ with distinct X_1 and X_2 from \mathcal{X} and $\lambda \in (0, 1)$. Proposition 5 implies that convex polytopes (with possibly infinite number of faces) are indecomposable. There is mathematical literature inspired by Gale (1954) and studying a similar concept of indecomposability where instead of taking weighted geometric mean can be misleading in infinite-dimensional convex geometry (Poulsen, 1959). In economic literature, such sets have appeared in the context of the n -good monopolist problem with $n \geq 2$, where mechanisms can be identified with convex functions on $[0, 1]^n$ such that their gradients also belong to $[0, 1]^n$ (Manelli and Vincent, 2007).

means, one takes convex combinations with respect to the Minkowski addition.¹⁷ In contrast to our setting, planar sets indecomposable in the sense of Gale form a simple parametric family (Gale, 1954; Silverman, 1973). However, in the dimension 3 and higher, Gale’s indecomposability behaves similarly to ours: indecomposable sets are dense in all convex sets and one can derive some necessary and some sufficient conditions of indecomposability that almost match each other but yet no criterion is known; see, e.g., (Sallee, 1972).

5.2 The domain of substitutes and the simplex property

Consider the domain \mathcal{D}_S of all homothetic preferences over n substitutes. Linear preferences belong to \mathcal{D}_S and are indecomposable since they are indecomposable even in the larger domain of all homothetic preferences by Proposition 5. For two goods, there are no other indecomposable preferences in \mathcal{D}_S .

Proposition 6. *For $n = 2$ goods, a preference \succsim is indecomposable in the domain \mathcal{D}_S of homothetic preferences with substitutability if and only if \succsim is linear.*

From Corollary 6, we know that any preference over $n = 2$ goods exhibiting substitutability can be obtained by aggregating linear preferences. Hence, any non-linear preference can be decomposed and we get Proposition 6.

Corollary 6 provides an explicit Choquet decomposition (37) for \mathcal{D}_S . Moreover, the corollary states that the decomposition is unique in the sense that the distribution of linear preferences over the population is pinned down uniquely. This phenomenon has the following geometric interpretation.

Consider a collection of d points in a finite-dimensional linear space such that no subset of $k \leq d$ points belongs to a $(k - 2)$ -dimensional linear subspace. The convex hull of such a collection is called a simplex. A simplex has the property that any point has the unique decomposition as a convex combination of extreme points. The uniqueness of the decomposition characterizes simplices among all other closed convex subsets of a finite-dimensional space. In the infinite-dimensional space, this property can be used to define a simplex, namely, a compact convex set is called a simplex if each point can be uniquely represented as the average of the extreme points, i.e., the measure in the Choquet integral (36) is uniquely defined (Phelps, 2001). Accordingly, we say that a closed domain of preferences is a simplex domain if there is a unique way to represent any preference as an aggregation of indecomposable ones, i.e., the measure μ in (37) is unique.

Corollary 10. *For two goods, the domain of homothetic preferences exhibiting substitutability is a simplex domain.*

¹⁷Gale (1954) calls such sets irreducible.

For $n \geq 3$ goods, there are other indecomposable preferences in \mathcal{D}_S except for linear ones. Indeed, by Corollary 5, aggregation of linear preferences does not give the whole domain \mathcal{D}_S . Since any preference can be represented as an aggregation of indecomposable ones, we conclude that there must be other indecomposable preferences. Describing them explicitly and checking whether \mathcal{D}_S is a simplex domain for $n \geq 3$ remains an open question.

5.3 The domain of complements

Let us discuss the domain \mathcal{D}_C of homothetic preferences exhibiting complementarity. Leontief preferences are indecomposable in \mathcal{D}_C since they are indecomposable in the full domain. By Corollary 7, aggregation of Leontief preferences does not give the whole \mathcal{D}_C even for $n = 2$ goods and, hence, there must be other indecomposable preferences. It turns out that indecomposable preferences are dense in \mathcal{D}_C and their structure resembles the one for the full domain. We call a preference \succeq a Leontief preference over Cobb-Douglas composite goods if it corresponds to a utility function

$$u(\mathbf{x}) = \min_{\mathbf{a} \in A} \{\chi_{\mathbf{a}}(\mathbf{x})\}, \quad (39)$$

where A is finite or countably infinite subset of $\mathbb{R}_{++} \times \Delta_{n-1}$ and each $\mathbf{a} = (a_0, a_1, \dots, a_n) \in A$ defines a Cobb-Douglas composite good $\chi_{\mathbf{a}}(\mathbf{x})$ by

$$\chi_{\mathbf{a}}(\mathbf{x}) \equiv a_0 \cdot \prod_{i=1}^n x_i^{a_i}.$$

Cobb-Douglas and Leontief preferences are particular cases of (39) corresponding, respectively, to a singleton $A = \{\mathbf{a}\}$ and to $A = \{(a_0^1, \mathbf{e}_1), \dots, (a_0^n, \mathbf{e}_n)\}$ where \mathbf{e}_i is the i 'th basis vector. We call a Leontief preference \succeq over Cobb-Douglas composite goods non-trivial if the set A contains at least two vectors \mathbf{a} and \mathbf{a}' with $(a_1, \dots, a_n) \neq (a'_1, \dots, a'_n)$. Equivalently, \succeq is non-trivial if it is not a standard Cobb-Douglas preference.

Proposition 7. *For $n = 2$ goods, non-trivial Leontief preferences over Cobb-Douglas composite goods are indecomposable in the domain of homothetic preferences with complementarity.*

The requirement of non-triviality is needed as standard Cobb-Douglas preferences can be decomposed as an aggregation of \succeq_i corresponding to $u_i(\mathbf{x}) = x_i$; see Example 1. Note that \succeq_i — which can be seen as either extreme linear or extreme Cobb-Douglas preference — is indecomposable in \mathcal{D}_C since it is indecomposable even in the full domain by Proposition 5.

Proposition 7 is proved in Appendix B.4. The idea is similar to Proposition 5 dealing with indecomposability in the full domain: indecomposable preferences correspond to price indices P with maximal number of active constraints. In contrast to Proposition 5 where the concavity of

P was the only constraint that matters, now we have the new monotonicity constraint on the budget share. Leontief preferences over Cobb-Douglas composite goods are obtained if the space is partitioned into regions where the monotonicity constraint is active (budget shares are constant) or the concavity constraint is active (the price index is linear). The former regions correspond to hyperbolic parts of the upper contour sets and the latter regions, to cusps.

6 Applications

This section illustrates how the geometric approach to preference aggregation can be used in various economic contexts.

6.1 Preference identification and simplex domains

Market demand reflects individual preferences but information loss is unavoidable. For example, aggregate behavior does not allow to distinguish populations where a pair of agents swapped their preferences and incomes or where a pair of agents with identical preferences is replaced with one agent with the joint income. We can still ask whether market demand determines the distribution of preferences over population, i.e., whether, by looking at the aggregate behavior, it is possible to determine what fraction of the population's income corresponds to agents with preferences of a particular kind.

Consider a population of consumers with homothetic preferences from some domain \mathcal{D} . An analyst knows neither the income distribution nor the size of the population and observes market demand generated by this population for any vector of prices. For any subset of preferences $\mathcal{D}' \subset \mathcal{D}$, the analyst aims to identify what fraction of the total income corresponds to agents in \mathcal{D}' .

In general, identification is impossible. For example, if \mathcal{D} is the domain of Cobb-Douglas preferences, the aggregate demand corresponding to $u_{\text{aggregate}}(x_1, x_2) = x_1^{1/3} \cdot x_2^{2/3}$ can be generated by a population where each agent has the same preference $\succeq_{\text{aggregate}}$ or, alternatively, by the population where 1/3 of the total income is earned by agents with preference $u_1(\mathbf{x}) = x_1$ and 2/3 by those with $u_2(\mathbf{x}) = x_2$; see Example 1.

The domain \mathcal{D} of linear preferences over $n = 2$ goods is an exception. A linear preference over two goods is determined by its marginal rate of substitution $\text{MRS} = v_1/v_2$. By Corollary 6, the fraction of income corresponding to consumers with MRS above a certain threshold α is equal to the fraction of income spent by the population on the first good for prices $p_1 = \alpha \cdot p_2$, i.e.,

$$\mu \left(\text{MRS} \geq \frac{p_1}{p_2} \right) = s_{\text{aggregate},1}(\mathbf{p}) = \frac{p_1 \cdot D_{\text{aggregate},1}(\mathbf{p}, B)}{p_1 \cdot D_{\text{aggregate},1}(\mathbf{p}, B) + p_2 \cdot D_{\text{aggregate},2}(\mathbf{p}, B)}.$$

Hence, even a few observations of market demand $D_{\text{aggregate}}$ at non-collinear price vectors can give a good understanding of the preference distribution over the population.

More generally, the distribution of preferences from a domain \mathcal{D} can be identified if any preference \succeq obtained by aggregation of preferences from \mathcal{D} cannot be decomposed over \mathcal{D} in a different way. A geometric interpretation of this property relies on the notion of simplex domains from Section 5.2 and is contained in the following corollary.

Corollary 11. *If the completion $\mathcal{D}^{\text{complete}}$ of \mathcal{D} is a simplex domain and \mathcal{D} consists of indecomposable preferences, then the distribution of preferences over the population can be identified from price dependence of market demand.*

Apart from linear preferences over two goods, there are many other domains satisfying the requirements of Corollary 11. One can take \mathcal{D} given by any finite collection of preferences $\{\succeq_1, \dots, \succeq_q\}$ none of which can be obtained as an aggregation of the others. For example, if q equals the number of goods n and each \succeq_k corresponds to $u(\mathbf{x}) = x_k$, then the income fraction of consumers with preference \succeq_k is equal to the budget share $s_{\text{aggregate},k}(\mathbf{p})$ at any price \mathbf{p} ; see also Example 1. We stress that just one observation of aggregate behavior at any particular vector of prices \mathbf{p} turns out to be enough to determine the distribution of preferences. The origin of this phenomenon is not the orthogonality of preferences but the fact that the dimension of the domain of preferences does not exceed the dimension of the consumption space. To illustrate this point, note that if $\succeq_1, \dots, \succeq_q$ are Cobb-Douglas preference with vectors of parameters $\mathbf{a}_1, \dots, \mathbf{a}_q$ that are linearly independent (possible only if $q \leq n$), then one observation of market demand also gives a linear system enough for identification.

Another domain \mathcal{D} satisfying conditions of Corollary 11 is the domain of Leontief preferences over two goods. By Corollary 8, any preference from its completion $\mathcal{D}^{\text{complete}}$ can be uniquely decomposed over Leontief preferences. Hence, $\mathcal{D}^{\text{complete}}$ is a simplex domain and Leontief preferences are indecomposable in it. We conclude that, in theory, the distribution of Leontief preferences can be identified. A peculiarity is that the Stieltjes-Perron inversion formula underlying Corollary 8 requires continuation of demand to complex prices. Therefore, it guarantees identification but gives no practical recipe of reconstructing the distribution of preferences for an analyst who observes demand for real prices only. Instead, the analyst can use real-inversion techniques for the Stieltjes transform, e.g., (Widder, 1938; Love and Byrne, 1980).

6.2 Robust welfare analysis via information design

Understanding how the population's welfare changes as a function of prices is crucial for designing government market interventions. Consider an analyst who observes market demand as a function

of prices and, based on this information, aims to estimate a certain aggregate measure of individual well-being.

The empirical literature uses representative consumer’s utility as an aggregate welfare measure. Hence, there is a one-to-one mapping between the market demand and the welfare functional. As pointed out by [Arkolakis et al. \(2012\)](#) and, more recently, by [Arkolakis et al. \(2019\)](#), this approach leads to surprisingly low gains from trade.

The assumption that market demand is a sufficient statistic for welfare is hardwired in the approach taken by the empirical literature. However, the same market demand — hence, the same aggregate preferences — can be generated by different populations of consumers. As a result, the same aggregate behavior may be compatible with a range of aggregate welfare levels. We illustrate how one can combine Theorem 1 with ideas from information design to compute the range of values of the welfare functional (or any other functional depending on individual preferences) compatible with given aggregate behavior, being fully agnostic about the specific decomposition of the market demand into individual demands. We call this a robust approach to welfare analysis.

As market demand pins down the aggregate consumer’s preference, we can assume that the aggregate preference $\succsim_{\text{aggregate}}$ is given. In addition, the analyst is given the total income B in the economy and a measure of individual “welfare” $w = w(\succsim, b)$ for a consumer with preference \succsim and income b . Note that w may also depend on other parameters — e.g., prices before and after a market intervention — but such dependence does not affect our analysis and hence omitted. The goal is to find the range of values of the welfare functional

$$W = \sum_j w(\succsim_j, b_j) \quad (40)$$

over all finite populations of consumers $j = 1, 2, \dots$ with preferences $\succsim_1, \succsim_2, \dots$ and incomes b_1, b_2, \dots such that the individual preferences aggregate up to $\succsim_{\text{aggregate}}$ and incomes sum up to $B = \sum_j b_j$.

Let us focus on computing the maximal value of W . We first analyze the case where w depends on income linearly

$$w(\succsim, b) = b \cdot h(\succsim)$$

and then discuss general w . Let us represent preferences by price indices and rewrite the problem with the help of Theorem 1. Denote by \mathcal{L} the set of all logarithmic price indices of homothetic preferences and by $\beta_j = b_j/B$, the relative income. We obtain that the maximal value of (40) is equal to the maximal value of

$$B \cdot \sum_j \beta_j \cdot h(\succsim_j),$$

where the maximum is taken over all possible ways to represent $\ln P_{\text{aggregate}}$ as a finite convex combination $\ln P_{\text{aggregate}} = \sum_j \beta_j \ln P_j$ with $\ln P_j$ from \mathcal{L} .

Similar optimization problems are well-known in Bayesian persuasion, a benchmark model for a situation where an informed party decides what information to reveal to an uninformed one and has an objective depending on induced beliefs (Kamenica and Gentzkow, 2011). Mathematically, persuasion boils down to solving the following optimization problem. We are given a set of states Ω , a prior belief $\mu \in \Delta(\Omega)$ where $\Delta(\Omega)$ denotes the simplex of probability distributions over Ω , and an objective function g defined on $\Delta(\Omega)$. The goal is to maximize

$$\sum_j \beta_j \cdot g(\mu_j)$$

over all possible ways to represent the prior μ as a finite convex combination $\mu = \sum_j \beta_j \cdot \mu_j$ with $\mu_j \in \Delta(\Omega)$.

Similarity between the two problems must be apparent: the set \mathcal{L} of logarithmic price indices plays the role of $\Delta(\Omega)$, the logarithmic price index of the aggregate preference corresponds to the prior μ , and $h = h(\succsim)$ considered as a function on \mathcal{L} is the analog of informed party's objective g .

The persuasion problem has an elegant geometric solution. For a function f on a convex subset X of a linear space, its concavification $\text{cav}_X[f]$ is the smallest concave functions on X larger than f . The optimal value of the persuasion problem is $\text{cav}_{\Delta(\Omega)}[g](\mu)$ (Kamenica and Gentzkow, 2011). Inspired by this result, we obtain a similar answer for welfare maximization.

Proposition 8. *For $w(\succsim, b) = b \cdot h(\succsim)$, the maximal welfare (40) compatible with an aggregate preference $\succsim_{\text{aggregate}}$ and income B is given by*

$$B \cdot \text{cav}_{\mathcal{L}}[h](\succsim_{\text{aggregate}}). \quad (41)$$

The function $h(\succsim)$ in (41) is treated as a functional on the space of logarithmic price indices. Hence, the concavification is over the infinite-dimensional functional space. Although the proposition may look abstract, it has straightforward economic implications and can be used for numeric simulations via finite-dimensional approximations.

Applying Proposition 8 to $\tilde{h} = (-1) \cdot h$, we get a version of the result for the minimal welfare: the minimal welfare equals $B \cdot \text{vex}_{\mathcal{L}}[h](\succsim_{\text{aggregate}})$ where vex denotes convexification $\text{vex}_X[f] = -\text{cav}_X[-f]$. Thus

$$W \in \left[B \cdot \text{vex}_{\mathcal{L}}[h](\succsim_{\text{aggregate}}), B \cdot \text{cav}_{\mathcal{L}}[h](\succsim_{\text{aggregate}}) \right] \quad (42)$$

is the range of possible values that the welfare can take for a given aggregate behavior.

The use of the aggregate agents's welfare $B \cdot h(\succsim_{\text{aggregate}})$ as a proxy of populations welfare is justified if the interval (42) is, in fact, a singleton, i.e., the convexification coincides with the concavification. The two coincide only for affine functions.

Corollary 12. *Aggregate consumer's welfare is a sufficient statistic for the population's welfare if the measure of individual welfare $h(\succsim)$ is an affine functional of the logarithmic price index $\ln(P_{\succsim})$. If it is not affine, there is an aggregate preference $\succsim_{\text{aggregate}}$ and two populations with the same total budget whose preferences aggregate up to $\succsim_{\text{aggregate}}$ but welfare levels are different.*

For example, $h(\succsim) = \ln P_{\succsim}(\mathbf{p}') - \ln P_{\succsim}(\mathbf{p})$ for some fixed price vectors \mathbf{p} and \mathbf{p}' is an affine functional. Normalizing, letting \mathbf{p}' go to \mathbf{p} , and taking into account the relation between budget shares and logarithmic price indices (4), one concludes that the budget share of any of the goods $h(\succsim) = s_i(\mathbf{p})$ is also affine. However, arguably the most natural choice of w equal to the normalized indirect utility is not affine. Assuming that the indirect utility is normalized to 1 at the unit budget and some price vector \mathbf{p}' , we get $w(\succsim, b) = b \cdot P_{\succsim}(\mathbf{p}')/P_{\succsim}(\mathbf{p}) = b \cdot \exp(\ln P_{\succsim}(\mathbf{p}') - \ln P_{\succsim}(\mathbf{p}))$ and, hence, the dependence of h on the logarithmic price index is exponential.

Let us discuss extensions and important particular cases of Proposition 8. In the maximization problem, one can easily replace the assumption of linearity of w in b with concavity and $w(\succsim, 0) = 0$. For concave w , splitting a consumer with preference \succsim and income b into a large number l of clones with the same preference and incomes b/l can only increase welfare and does not affect the aggregate behavior. Hence, only the behavior of w around zero plays matters: the maximal welfare for concave w is equal to the maximal welfare for linear $\tilde{w} = b \cdot h$ with

$$h(\succsim) = \lim_{b \rightarrow 0} \frac{\partial w(\succsim, b)}{\partial b}.$$

One can similarly handle the case of convex dependence on b in the minimization problem. To get bounds for general w , one can squeeze it between a concave upper bound and a convex lower bound (without loss of generality, both can be taken to be linear).

The analysis extends to the case where the analyst additionally knows that individual preferences are not arbitrary but come from a certain subdomain \mathcal{D} of homothetic preferences. For domains \mathcal{D} such that the set of logarithmic price indices $\mathcal{L}_{\mathcal{D}}$ corresponding to \mathcal{D} is convex (equivalently, \mathcal{D} is invariant with respect to aggregation) one just need to replace \mathcal{L} with $\mathcal{L}_{\mathcal{D}}$ in the above treatment.

A domain restriction can result in a tractability gain. If \mathcal{D} is generated by a finite collection of preferences as in Example 1, finding concavification in (41) becomes a finite-dimensional problem. Domain restriction also helps to establish the formal equivalence between welfare maximization and persuasion beyond similarity.

We saw that welfare maximization with w linear in b is similar to Bayesian persuasion. The difference is that in Bayesian persuasion, the concavification takes place over a simplex $\Delta(\Omega)$ while the set $\mathcal{L}_{\mathcal{D}}$ of logarithmic price indices is not necessarily a simplex. Recall that a convex set is a simplex if the decomposition over the extreme points is unique; we call \mathcal{D} a simplex domain if the corresponding set of logarithmic price indices $\mathcal{L}_{\mathcal{D}}$ is a simplex (Section 5.2).

For simplex domains, welfare maximization is equivalent to a persuasion problem. We will exemplify the equivalence for the domain \mathcal{D}_S of all preferences exhibiting substitutability over $n = 2$ goods. Any preference $\succsim \in \mathcal{D}_S$ can be represented as an aggregation of linear preferences and this representation is unique (Corollary 6). Linear preferences over two goods form a one-parametric family with the marginal rate of substitution $\text{MRS} = v_1/v_2 \in \mathbb{R}_+ \cup \{+\infty\}$ as a parameter. A preference $\succsim \in \mathcal{D}_S$ defines a unique distribution μ of MRS by formula (28): the cumulative distribution function is equal to $1 - s_1(\cdot, 1)$. The function $h(\succsim)$ can equivalently be thought as a function of μ . Thus the welfare maximization problem takes the following form. We are given $\mu_{\text{aggregate}} \in \Delta(\mathbb{R}_+ \cup \{+\infty\})$ and a functional $h = h(\mu)$. The goal is to maximize

$$B \cdot \sum_j \beta_j \cdot h(\mu_j)$$

over all possible ways to represent the prior μ as a finite convex combination $\mu = \sum_j \beta_j \cdot \mu_j$ with $\mu_j \in \Delta(\mathbb{R}_+ \cup \{+\infty\})$.

We conclude that, for two substitutes, welfare maximization is equivalent to persuasion with the set of states $\Omega = \mathbb{R}_+ \cup \{+\infty\}$, the cumulative distribution function of the prior μ equal to $1 - s_1(\cdot, 1)$, and the objective $h = h(\mu)$. Persuasion problems with a continual state space are not easy to solve analytically unless some further assumptions are made. For example, if μ is finitely supported, i.e., there is a finite number of preference “types” in the population, then the support can be taken as the new set of states reducing the problem to the well-understood case of persuasion with finite number of states. If μ has infinite support, tractability can be gained by imposing assumptions on the objective h . Tractable cases include h depending on μ through the mean value of a given function φ , i.e., $h = h(\int \varphi(z) d\mu(z))$ as in (Dworczak and Martini, 2019; Arieli et al., 2019; Kleiner et al., 2021) or h depending on μ through a quantile of φ as in (Yang and Zentefis, 2022).

6.3 Fisher markets, fair division, complexity, and bidding languages

Consider a population of consumers with budgets b_1, \dots, b_m and homothetic preferences $\succsim_1, \dots, \succsim_m$ over n goods. Let us augment this setting by adding a bundle $\mathbf{x} \in \mathbb{R}_{++}^n$ interpreted as fixed total supply of the goods. This economy is known in algorithmic economics literature as the Fisher market¹⁸ and is by far the most studied economy from computational perspective (Nisan et al., 2007, Chapters 5 and 6). Since the classic works of Varian (1974) and Hylland and Zeckhauser (1979), Fisher markets and their modifications are used for fair allocation of private goods without monetary transfers giving rise to a famous mechanism known as the competitive equilibrium with equal incomes (CEEI) or the pseudo-market mechanism (Moulin, 2019; Pycia, 2022).

¹⁸Named after Irving Fisher who introduced a hydraulic method for equilibrium price computation; see (Brainard and Scarf, 2005).

A collection of bundles $\mathbf{x}_1, \dots, \mathbf{x}_m$ and a price vector \mathbf{p} form a competitive equilibrium (CE) of the Fisher market with preferences $\succsim_1, \dots, \succsim_m$, budgets b_1, \dots, b_m , and total supply \mathbf{x} if

$$\mathbf{x}_k \in D_k(\mathbf{p}, b_k), \quad \text{for each consumer } k, \quad \text{and} \quad \sum_{k=1}^m \mathbf{x}_k = \mathbf{x}, \quad (43)$$

i.e., each consumer buys the most preferred bundle within her budget and the market clears. We pinpoint that money have no intrinsic value and the Fisher market is equivalent to an exchange economy where each agent k is endowed with the fraction $\beta_k = b_k/B$ of \mathbf{x} where B is the total budget.

One can think of a CE as an allocation mechanism: agents report their preferences, the mechanism computes an equilibrium and allocates each agent her bundle \mathbf{x}_k . In this interpretation, budgets b_k represent agents' entitlement to the goods in the bundle \mathbf{x} . The case of equal entitlements $b_1 = \dots = b_m$ (CEEI) is especially important. In this case, each agent selects her best bundle from the same budget set and, hence, the resulting allocation is envy-free in the sense that $\mathbf{x}_k \succsim_k \mathbf{x}_l$ for any pair of agents k and l . Since any CE is Pareto optimal by the first welfare theorem, CEEI gives a simple recipe to combine strong fairness and efficiency guarantees. CEEI and its variants have been applied to rent division (Goldman and Procaccia, 2015), chores allocation (Bogomolnaia et al., 2017), course allocation (Budish et al., 2017; Kornbluth and Kushnir, 2021; Soumalias et al., 2022), cloud computing (Devanur et al., 2018), school choice (Ashlagi and Shi, 2016; He et al., 2018), and other problems (Echenique et al., 2021).

Despite its attractive properties, popularity of CEEI remains limited as computing its outcome is a challenging problem. It is known that an equilibrium allocation $\mathbf{x}_1, \dots, \mathbf{x}_m$ can be obtained via maximizing the Nash social welfare

$$\prod_{k=1}^m \left(\frac{u_k(\mathbf{x}_k)}{\beta_k} \right)^{\beta_k} \quad (44)$$

over all bundles $\mathbf{x}_1, \dots, \mathbf{x}_m$ such that $\sum_{k=1}^m \mathbf{x}_k = \mathbf{x}$. This result tightly related to the existence of an aggregate consumer was established by Eisenberg and Gale (1959) for linear preferences but holds for all homothetic preferences; see (Shafer and Sonnenschein, 1982). Although the Eisenberg-Gale problem is convex, computing its solutions is not an easy task unless n or m are small. Even in the benchmark case of linear preferences, algorithms with good theoretical performance have required more than a decade of research and dozens of papers using cutting-edge techniques; see, e.g., (Devanur et al., 2002; Orlin, 2010; Végh, 2012). Developing algorithms with good performance in practice is critical for large-scale applications of Fisher markets — e.g., to fair recommender systems (Gao and Kroer, 2022) and Internet ad markets (Conitzer et al., 2022) — but despite the recent progress this problem is yet to be solved.

We examine the question of finding a CE from preference aggregation perspective. This per-

spective sheds light on why computing a CE can be challenging in seemingly innocent domains such as linear preferences and helps to identify domains where computing a CE is easy.

To find a CE, it is enough to compute the vector of equilibrium prices \mathbf{p} . Once we know \mathbf{p} , each agent is allocated her demanded bundle \mathbf{x}_k at these prices.¹⁹ Thus the essence of computing a CE is finding a vector of prices \mathbf{p} such that the market demand matches the supply. In other words, we need to find \mathbf{p} such that *the aggregate consumer's demand contains \mathbf{x}* . This simple observation combined with our insights about the structure of aggregate preferences has many implications.

Finding a CE for a population of consumers boils down to finding a CE for one aggregate consumer and we know that aggregation is easier to handle in the space of logarithmic price indices. Recall that the demand is proportional to the gradient of the logarithmic price index (2) and, hence, \mathbf{p} is an equilibrium price vector if and only if²⁰

$$\mathbf{x} = B \cdot \nabla \ln P_{\text{aggregate}}(\mathbf{p}),$$

where B is the total budget. Interpreting this identity as the first order condition and taking into account concavity of $\ln P_{\text{aggregate}}$, we conclude that \mathbf{p} is a vector of equilibrium prices whenever

$$\mathbf{p} \quad \text{is the global maximum of} \quad \langle \mathbf{x}, \mathbf{p} \rangle - B \cdot \ln P_{\text{aggregate}}(\mathbf{p}). \quad (45)$$

Combining this result with Theorem 1, we get the following proposition.

Proposition 9. *A vector \mathbf{p} is a vector of equilibrium prices for a population of consumers with homothetic preferences $\succeq_1, \dots, \succeq_m$, budgets b_1, \dots, b_m , and total supply \mathbf{x} if and only if \mathbf{p} is the global maximum of*

$$\langle \mathbf{x}, \mathbf{p} \rangle - \sum_{k=1}^m b_k \cdot \ln P_k(\mathbf{p}). \quad (46)$$

This optimization problem is convex. Its particular case for linear preferences has been known and obtained as the Lagrange dual to the Eisenberg-Gale optimization problem (Cole et al., 2017; Devanur et al., 2016; Shmyrev, 2009). Our approach explains the preference-aggregation origin of this dual and provides the generalization to all homothetic preferences almost without any computations.

The optimization problems (45) and (46) indicate that to find a CE for a population of consumers with preferences from a certain domain \mathcal{D} we must be able to find the market equilibrium for any preference that can be obtained as an aggregation of preferences from \mathcal{D} . In other words, the

¹⁹If utilities are not strictly concave, $D_k(\mathbf{p}, b)$ may not be a singleton, e.g., in the domain of linear preferences. Even in this case, once equilibrium prices are known, choosing bundles \mathbf{x}_k from each agent's demand so that $\sum_k \mathbf{x}_k = \mathbf{x}$ is a simple problem which boils down to a maximum flow computation (Devanur et al., 2002; Brânzei and Sandmirskiy, 2019).

²⁰For price indices that are not smooth the gradient is to be replaced with the subdifferential.

complexity of finding a CE is determined not by the domain \mathcal{D} of individual preferences itself but rather by its completion $\mathcal{D}^{\text{complete}}$. We illustrate this point for the domain of linear preferences over two goods.

By Proposition 2, aggregation of linear preferences over $n = 2$ goods (domain \mathcal{D}) gives all preferences with substitutability ($\mathcal{D}^{\text{complete}}$). We will show that any algorithm computing an approximate CE for preferences from \mathcal{D} can be used to compute an approximate CE for $\mathcal{D}^{\text{complete}}$. Hence, finding CE for \mathcal{D} cannot be easy if it is hard for $\mathcal{D}^{\text{complete}}$. Let us call \mathbf{p} an ε -equilibrium price vector if there are $\mathbf{x}_k \in D_k(\mathbf{p}, b_k)$, $j = 1, \dots, m$ such that

$$\langle \mathbf{p}, \mathbf{e} \rangle \leq \varepsilon \cdot B, \quad \text{where} \quad e_i = \left| x_i - \sum_{k=1}^m x_{k,i} \right|,$$

i.e., the excess demand is relatively small compared to the total budget.

Proposition 10. *Let \mathcal{D} be the domain of linear preferences over two goods and assume we have an algorithm computing an ε -equilibrium price vector for any population of agents with preferences from \mathcal{D} . Then a 3ε -equilibrium price vector for a population of m agents with preferences from $\mathcal{D}^{\text{complete}}$ can be computed by applying the algorithm as a black box to an auxiliary population with preferences from \mathcal{D} and the number of agents of the order of m/ε .*

The idea is to approximate preferences from $\mathcal{D}^{\text{complete}}$ by the aggregate preference of linear consumers so that the budget shares differ by at most ε . Such approximation can be constructed via Corollary 6 and requires of the order of $1/\varepsilon$ auxiliary linear consumers. As we show in Appendix B.5, if budget shares in two populations differ by at most ε , then ε -equilibrium price vector for one population is an $(1 + n)\varepsilon$ -equilibrium price vector for the other. Hence, an ε -equilibrium price vector for the approximating population of linear consumers gives a 3ε -equilibrium price vector for the original population.

The example of linear preferences demonstrates that even a simple parametric domain — if the choice of parameters is not aligned with aggregation — can have a large non-parametric completion. As a result, the simplicity of a parametric domain does not carry over to the aggregate behavior thus complicating computation of a CE. To preserve the simplicity of a parametric domain, the choice of parameters is to be aligned with aggregation. Motivated by this concern, we consider computing a CE in parametric domains invariant with respect to aggregation.

Fix a finite family of “elementary” preferences $\succsim_1, \dots, \succsim_q$ and consider the domain $\mathcal{D} = \mathcal{D}^{\text{complete}}$ of all preferences that can be obtained by aggregating the elementary preferences. We will call such invariant domains finitely-generated. Cobb-Douglas preferences are an example of a finitely generated domain; see Example 1. By Theorem 1, a finitely generated \mathcal{D} consists of all preferences \succsim

whose price index P can be represented as $\ln P = \sum_{l=1}^q t_l \ln P_l$ and, hence, the vector of coefficients $\mathbf{t} \in \Delta_{q-1}$ provides a parameterization of \mathcal{D} .

Proposition 11. *Consider a finitely-generated invariant domain $\mathcal{D} = \{\succsim_1, \dots, \succsim_q\}^{\text{complete}}$ and fix $\varepsilon \geq 0$. Assume we have access to an algorithm finding an ε -equilibrium vector of prices for $m = 1$ agent and using at most T operations. Then an ε -equilibrium price vector for a population of $m \geq 1$ agents can be computed in time of the order of $m \cdot q + T$.*

The proof is straightforward. If preferences of individual agents are represented by $\mathbf{t}_1, \dots, \mathbf{t}_m$ and β_1, \dots, β_m are relative incomes, then, by Theorem 1, the aggregate consumer corresponds to $\mathbf{t} = \sum_{k=1}^m \beta_k \cdot \mathbf{t}_k$. Computing \mathbf{t} requires of the order of $m \cdot q$ operations. Applying the one-agent algorithm to the aggregate agent, we get an ε -equilibrium vector of prices for the original population in T operations.

The linear growth of running time with the number of agents m and absence of large hidden constants suggests that finitely-generated domains can be a natural candidate for scalable fair division mechanisms. Note that the best running time for linear preferences achieved by [Orlin \(2010\)](#) and [Végh \(2012\)](#) grows as m^4 .

In economic design, the choice of a preference domain corresponds to the choice of a bidding language, i.e., the information about the true — possibly substantially more complicated — preferences that the participants can report to a mechanism. Our results indicate that the advantages of bidding languages corresponding to finitely-generated invariant domains.

Finitely-generated invariant domains offer enough flexibility to the designer. For example, apart from Cobb-Douglas preferences, one can consider domains generated by a finite collection of linear preferences. By adding preferences exhibiting complementarity among certain subsets — e.g., pairs — of goods to the collection of elementary preferences, we can allow agents to express both substitutability and complementarity patterns while keeping the domain narrow. Of course, the use of such domains and bidding languages in practice requires additional experimental evaluation as in ([Budish and Kessler, 2022](#)).

References

- C. D. Aliprantis and K. C. Border. *Infinite Dimensional Analysis: A Hitchhiker's Guide*. Springer-Verlag Berlin and Heidelberg GmbH & Company KG, 2013.
- I. Arieli, Y. Babichenko, R. Smorodinsky, and T. Yamashita. Optimal persuasion via bi-pooling. *Available at SSRN 3511516*, 2019.

- C. Arkolakis, A. Costinot, and A. Rodríguez-Clare. New trade models, same old gains? *American Economic Review*, 102(1):94–130, 2012.
- C. Arkolakis, A. Costinot, D. Donaldson, and A. Rodríguez-Clare. The elusive pro-competitive effects of trade. *The Review of Economic Studies*, 86(1):46–80, 2019.
- I. Ashlagi and P. Shi. Optimal allocation without money: An engineering approach. *Management Science*, 62(4):1078–1097, 2016.
- F. Benford. The law of anomalous numbers. *Proceedings of the American philosophical society*, 78(4):551–572, 1938.
- A. Bogomolnaia, H. Moulin, F. Sandomirskiy, and E. Yanovskaya. Competitive division of a mixed manna. *Econometrica*, 85(6):1847–1871, 2017.
- K. J. Böröczky, E. Lutwak, D. Yang, and G. Zhang. The log-Brunn–Minkowski inequality. *Advances in Mathematics*, 231(3-4):1974–1997, 2012.
- W. C. Brainard and H. E. Scarf. How to compute equilibrium prices in 1891. *American Journal of Economics and Sociology*, 64(1):57–83, 2005.
- S. Brânzei and F. Sandomirskiy. Algorithms for competitive division of chores. *arXiv preprint arXiv:1907.01766*, 2019.
- E. Budish and J. B. Kessler. Can market participants report their preferences accurately (enough)? *Management Science*, 68(2):1107–1130, 2022.
- E. Budish, G. P. Cachon, J. B. Kessler, and A. Othman. Course match: A large-scale implementation of approximate competitive equilibrium from equal incomes for combinatorial allocation. *Operations Research*, 65(2):314–336, 2017.
- L. Calvet, J.-M. Grandmont, I. Lemaire, et al. Aggregation of heterogeneous beliefs and asset pricing in complete financial markets. Technical report, 2001.
- C. D. Carroll. Requiem for the representative consumer? Aggregate implications of microeconomic consumption behavior. *American Economic Review*, 90(2):110–115, 2000.
- C. Chamley. Optimal taxation of capital income in general equilibrium with infinite lives. *Econometrica: Journal of the Econometric Society*, pages 607–622, 1986.
- P.-A. Chiappori. Rational household labor supply. *Econometrica: Journal of the Econometric Society*, pages 63–90, 1988.

- P.-A. Chiappori and I. Ekeland. Aggregation and market demand: an exterior differential calculus viewpoint. *Econometrica*, 67(6):1435–1457, 1999.
- P.-A. Chiappori and I. Ekeland. The microeconomics of efficient group behavior: Identification 1. *Econometrica*, 77(3):763–799, 2009.
- R. Cole, N. Devanur, V. Gkatzelis, K. Jain, T. Mai, V. V. Vazirani, and S. Yazdanbod. Convex program duality, Fisher markets, and Nash social welfare. In *Proceedings of the 2017 ACM Conference on Economics and Computation*, pages 459–460, 2017.
- V. Conitzer, C. Kroer, D. Panigrahi, O. Schrijvers, N. E. Stier-Moses, E. Sodomka, and C. A. Wilkens. Pacing equilibrium in first price auction markets. *Management Science*, 2022.
- A. Deaton and J. Muellbauer. An almost ideal demand system. *The American economic review*, 70(3):312–326, 1980.
- G. Debreu. Excess demand functions. *Journal of mathematical economics*, 1(1):15–21, 1974.
- N. R. Devanur, C. H. Papadimitriou, A. Saberi, and V. V. Vazirani. Market equilibrium via a primal-dual-type algorithm. In *The 43rd Annual IEEE Symposium on Foundations of Computer Science, 2002. Proceedings.*, pages 389–395. IEEE, 2002.
- N. R. Devanur, K. Jain, T. Mai, V. V. Vazirani, and S. Yazdanbod. New convex programs for Fisher’s market model and its generalizations. *arXiv preprint arXiv:1603.01257*, 2016.
- N. R. Devanur, J. Garg, R. Mehta, V. V. Vaziranb, and S. Yazdanbod. A new class of combinatorial markets with covering constraints: Algorithms and applications. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 2311–2325. SIAM, 2018.
- W. E. Diewert. Applications of duality theory. In *Frontiers of Quantitative Economics. Volume II*, ed. M.D. Intriligator and D.A. Kendrick., pages 106–199. North Holland Publishing Company, 1974.
- P. Dworczak and G. Martini. The simple economics of optimal persuasion. *Journal of Political Economy*, 127(5):1993–2048, 2019.
- F. Echenique, A. Miralles, and J. Zhang. Constrained pseudo-market equilibrium. *American Economic Review*, 111(11):3699–3732, 2021.
- E. Eisenberg. Aggregation of utility functions. *Management Science*, 7(4):337–350, 1961.
- E. Eisenberg and D. Gale. Consensus of subjective probabilities: The pari-mutuel method. *The Annals of Mathematical Statistics*, 30(1):165–168, 1959.

- D. Gale. Irreducible convex sets. *Proc. Intern. Congr. Math., Amsterdam*, 2:217–218, 1954.
- Y. Gao and C. Kroer. Infinite-dimensional Fisher markets and tractable fair division. *Operations Research*, 2022.
- J. Goldman and A. D. Procaccia. Spliddit: Unleashing fair division algorithms. *ACM SIGecom Exchanges*, 13(2):41–46, 2015.
- W. M. Gorman. Community preference fields. *Econometrica: journal of the Econometric Society*, pages 63–80, 1953.
- W. M. Gorman. On a class of preference fields. *Metroeconomica*, 13(2):53–56, 1961.
- Y. He, A. Miralles, M. Pycia, and J. Yan. A pseudo-market approach to allocation with priorities. *American Economic Journal: Microeconomics*, 10(3):272–314, 2018.
- W. Hildenbrand. Market demand. In *Market Demand*. Princeton University Press, 2014.
- A. Hylland and R. Zeckhauser. The efficient allocation of individuals to positions. *Journal of Political economy*, 87(2):293–314, 1979.
- M. O. Jackson and L. Yariv. The non-existence of representative agents. *Available at SSRN 2684776*, 2019.
- M. Jerison. Aggregation and pairwise aggregation of demand when the distribution of income is fixed. *Journal of economic Theory*, 33(1):1–31, 1984.
- M. Jerison. Optimal income distribution rules and representative consumers. *The Review of Economic Studies*, 61(4):739–771, 1994.
- D. W. Jorgenson, L. J. Lau, and T. M. Stoker. The transcendental logarithmic model of aggregate consumer behavior. *Advances in econometrics*, 1:97–238, 1982.
- E. Jouini and C. Napp. Consensus consumer and intertemporal asset pricing with heterogeneous beliefs. *The Review of Economic Studies*, 74(4):1149–1174, 2007.
- E. Kamenica and M. Gentzkow. Bayesian persuasion. *American Economic Review*, 101(6):2590–2615, 2011.
- Z. Y. Kang and S. Vasserman. Robust bounds for welfare analysis. Technical report, National Bureau of Economic Research, 2022.
- A. P. Kirman. Whom or what does the representative individual represent? *Journal of economic perspectives*, 6(2):117–136, 1992.

- A. Kleiner, B. Moldovanu, and P. Strack. Extreme points and majorization: Economic applications. *Econometrica*, 89(4):1557–1593, 2021.
- D. Kornbluth and A. Kushnir. Undergraduate course allocation through pseudo-markets. *Available at SSRN 3901146*, 2021.
- D. M. Kreps. *A course in microeconomic theory*. Princeton university press, 2020.
- A. Lewbel. Exact aggregation and a representative consumer. *The Quarterly Journal of Economics*, 104(3):621–633, 1989.
- E. Love and A. Byrne. Real inversion theorems for generalized stieltjes transforms. *Journal of the London Mathematical Society*, 2(2):285–306, 1980.
- A. M. Manelli and D. R. Vincent. Multidimensional mechanism design: Revenue maximization and the multiple-good monopoly. *Journal of Economic theory*, 137(1):153–185, 2007.
- R. R. Mantel. On the characterization of aggregate excess demand. *Journal of economic theory*, 7(3):348–353, 1974.
- R. R. Mantel. Homothetic preferences and community excess demand functions. *Journal of Economic Theory*, 12(2):197–201, 1976.
- D. McFadden. Econometric models of probabilistic choice. *Structural analysis of discrete data with econometric applications*, 198272, 1981.
- V. Milman and L. Rotem. Non-standard constructions in convex geometry: geometric means of convex bodies. In *Convexity and concentration*, pages 361–390. Springer, 2017.
- H. Moulin. Fair division in the internet age. *Annual Review of Economics*, 11:407–441, 2019.
- J. F. Nash. The bargaining problem. *Econometrica: Journal of the econometric society*, pages 155–162, 1950.
- N. Nisan, T. Roughgarden, E. Tardos, and V. V. Vazirani. *Algorithmic Game Theory*. Cambridge University Press, 2007.
- J. B. Orlin. Improved algorithms for computing Fisher’s market clearing prices: Computing Fisher’s market clearing prices. In *Proceedings of the forty-second ACM symposium on Theory of computing*, pages 291–300, 2010.
- R. R. Phelps. *Lectures on Choquet’s theorem*. Springer, 2001.

- E. T. Poulsen. Convex sets with dense extreme points. *The American Mathematical Monthly*, 66(7):577–578, 1959.
- M. Pycia. Pseudomarkets. In *Online and Matching-Based Market Design* Eds. Federico Echenique, Nicole Immorlica, and Vijay V. Vazirani. Cambridge University Press, 2022.
- K. Rogoff. Equilibrium political budget cycles. *American Economic Review*, (80):21–36, 1990.
- G. Salée. Minkowski decomposition of convex sets. *Israel Journal of Mathematics*, 12(3):266–276, 1972.
- R. L. Schilling, R. Song, and Z. Vondracek. Bernstein functions. In *Bernstein functions*. de Gruyter, 2012.
- W. Shafer and H. Sonnenschein. Market demand and excess demand functions. *Handbook of mathematical economics*, 2:671–693, 1982.
- V. I. Shmyrev. An algorithm for finding equilibrium in the linear exchange model with fixed budgets. *Journal of Applied and Industrial Mathematics*, 3(4):505–518, 2009.
- R. Silverman. Decomposition of plane convex sets. I. *Pacific Journal of Mathematics*, 47(2):521–530, 1973.
- H. Sonnenschein. Do Walras’ identity and continuity characterize the class of community excess demand functions? *Journal of economic theory*, 6(4):345–354, 1973.
- J. R.-V. Sørensen and M. Fosgerau. How McFadden met Rockafellar and learned to do more with less. *Journal of Mathematical Economics*, 100:102629, 2022.
- E. Soumalias, B. Zamanlooy, J. Weissteiner, and S. Seuken. Machine learning-powered course allocation. *arXiv preprint arXiv:2210.00954*, 2022.
- H. R. Varian. Equity, envy, and efficiency. *Journal of Economic Theory*, 9(1):63–91, 1974.
- H. R. Varian. Social indifference curves and aggregate demand. *The Quarterly Journal of Economics*, 99(3):403–414, 1984.
- L. A. Végh. Strongly polynomial algorithm for a class of minimum-cost flow problems with separable convex objectives. In *Proceedings of the forty-fourth annual ACM symposium on Theory of computing*, pages 27–40, 2012.
- D. Widder. The Stieltjes transform. *Transactions of the American Mathematical Society*, 43(1):7–60, 1938.

A Topology on preferences and integration

There are two high-level reasons why we need a topology on preferences. The topology is necessary to define closure of preference domains as in our discussion of completion but, most importantly, it is needed to formalize integration over preferences and to apply Choquet theory (Phelps, 2001). Recall that Choquet theory deals with compact convex subsets of locally convex topological vector spaces. Our goal is to identify the domain of all homothetic preferences with a compact convex subset of a Banach space (complete normed and, hence, locally convex vector space).

We represent a homothetic preference \succsim by its logarithmic price index $\ln P$. We call two functions f and g equivalent if $f - g = \text{const}$. Since the price index is defined up to a multiplicative factor, each preference corresponds to the class of equivalent logarithmic price indices.

Let \mathcal{L} be the set of classes of equivalent continuous functions f on \mathbb{R}_{++}^n that can be obtained as logarithmic price indices of homothetic preferences. The set \mathcal{L} is in one-to-one correspondence with the domain of homothetic preferences. Hence, to define a topology and integration for preferences, it is enough to define them for \mathcal{L} . We first introduce a metric structure. To motivate the definition of a distance, we need some estimates on price indices.

Lemma 1. *For any price index P , the following inequality holds*

$$|\ln P(\mathbf{p}) - \ln P(\mathbf{p}')| \leq \max_i |\ln p_i - \ln p'_i| \quad (47)$$

for any pair of price vectors \mathbf{p} and \mathbf{p}' from \mathbb{R}_{++}^n .

In other words, logarithmic price indices are 1-Lipshitz functions of logarithms of prices.

Proof. We need to show that

$$\min_i \frac{p_i}{p'_i} \leq \frac{P(\mathbf{p})}{P(\mathbf{p}')} \leq \max_i \frac{p_i}{p'_i}.$$

It is enough to demonstrate the upper bound and the lower bound will follow by flipping the roles of \mathbf{p} and \mathbf{p}' .

Recall that $P(\mathbf{p})$ is the minimal budget that the agent needs to achieve the unit level of utility for prices \mathbf{p} . Given prices \mathbf{p} and \mathbf{p}' , define $\mathbf{p}'' = \max_i \frac{p_i}{p'_i} \cdot \mathbf{p}'$. The price of each good under \mathbf{p}'' is higher than for \mathbf{p} and, hence, the agent needs at least as much money to achieve the same welfare level. Thus

$$P(\mathbf{p}) \leq P(\mathbf{p}'') = \max_i \frac{p_i}{p'_i} \cdot P(\mathbf{p}'),$$

where we used the homogeneity of the price index. Dividing both sides by $P(\mathbf{p}')$, we obtain the desired inequality and complete the proof. \square

Denote by \mathbf{e} the vector of all ones $\mathbf{e} = (1, \dots, 1)$. By the lemma, we see that any price index satisfies the following estimate

$$\left| \frac{\ln P(\mathbf{p}) - \ln P(\mathbf{e})}{1 + \max_i |\ln p_i|} \right| \leq 1 \quad (48)$$

for any vector of prices. The normalization in (48) suggest how to define a distance so that the set of logarithmic price indices has a bounded diameter.

We define the distance between preferences \succsim and \succsim' or, equivalently, between the corresponding logarithmic price indices $f = \ln P$ and $f' = \ln P'$ as follows:

$$d(\succsim, \succsim') = d(f, f') = \sup_{\mathbf{p} \in \Delta_{n-1} \cap \mathbb{R}_{++}^n} \left| \frac{(\ln P(\mathbf{p}) - \ln P(\mathbf{e})) - (\ln P'(\mathbf{p}) - \ln P'(\mathbf{e}))}{(1 + \max_i |\ln p_i|)^2} \right|. \quad (49)$$

The denominator in (49) is squared so that

$$\frac{\ln P(\mathbf{p}) - \ln P(\mathbf{e})}{(1 + \max_i |\ln p_i|)^2} \rightarrow 0 \quad \text{as } \mathbf{p} \text{ approaches the boundary of } \Delta_{n-1}. \quad (50)$$

Hence, the supremum is always attained at an interior point of the simplex and so can be replaced with the maximum.

Note that the ratio in (49) does not depend on the choice of a logarithmic price index from the class of equivalent ones. On the other hand, the distance between any two distinct preferences or, equivalently, between two non-equivalent logarithmic price indices is non-zero as the values of logarithmic price indices on \mathbb{R}_{++}^n are determined by their values on the interior of the simplex $\Delta_{n-1} \cap \mathbb{R}_{++}^n = \{\mathbf{p} \in \mathbb{R}_{++}^n : \sum_i x_i = 1\}$ since $P(\alpha \cdot \mathbf{p}) = \alpha \cdot P(\mathbf{p})$.

The metric structure on preferences allows one to define convergence and closed sets. A closure of a domain \mathcal{D} of preference consists of all limit points of \mathcal{D} , i.e., of all the preferences \succsim such that there exists a sequence of preferences $\succsim^{(l)} \in \mathcal{D}$ with $d(\succsim, \succsim^{(l)}) \rightarrow 0$ as $l \rightarrow \infty$. A closed domain is a domain that coincides with its closure.

Open sets are complements of closed ones and so the metric defines a topology. Once the topology is defined, one constructs the Borel measurable structure in the standard way (Aliprantis and Border, 2013, Section 4.4). Hence, we can write integrals of the form

$$\int_{\mathcal{D}} G(\succsim) d\mu(\succsim) = \int_{\mathcal{L}} G(f) d\mu(f)$$

formally where G is a Borel-measurable function and μ is a Borel measure (as usual, we identify functions and measures on preferences and on logarithmic price indices). In all our examples, the integrated function G is continuous and, hence, measurable.

By (48), the diameter of \mathcal{L} does not exceed 2. Hence, \mathcal{L} is a bounded convex set. To fit the assumptions of the Choquet theory, we need to show that \mathcal{L} is compact and can be thought as a subset of a Banach space. We achieve both goals by constructing an isometric compact embedding of \mathcal{L} into a Banach space.

Consider the Banach space $\mathcal{C}(\Delta_{n-1})$ of all continuous functions on the simplex endowed with the standard sup-norm $\|h\| = \sup_{\mathbf{p} \in \Delta_{n-1}} |h(\mathbf{p})|$.

Lemma 2. *Let E be a map that maps a logarithmic price index $f = \ln P$ to a function $E[f]$ on Δ_{n-1} given by*

$$E[f](\mathbf{p}) = \begin{cases} \frac{\ln P(\mathbf{p}) - \ln P(\mathbf{e})}{(1 + \max_i |\ln p_i|)^2}, & \mathbf{p} \in \Delta_{n-1} \cap \mathbb{R}_{++}^n \\ 0, & \text{otherwise} \end{cases}. \quad (51)$$

Then E is an isometric embedding of the set \mathcal{L} of logarithmic price indices in the Banach space of continuous functions $\mathcal{C}(\Delta_{n-1})$ and the image $E[\mathcal{L}]$ is a compact convex set.

Proof. The function $E[f]$ is continuous in the interior of the simplex by the continuity of price indices and it is continuous on the boundary by (50). Hence, $E[f]$ belongs to $\mathcal{C}(\Delta_{n-1})$. By the definition of the distance (49) and the norm in $\mathcal{C}(\Delta_{n-1})$, we get $d(f, f') = \|E[f] - E[f']\|$. Hence, E preserves the distance and, in particular, $f \neq f'$ implies $E[f] \neq E[f']$. Thus E is an isometric embedding of \mathcal{L} in $\mathcal{C}(\Delta_{n-1})$.

The diameter of the image $E[\mathcal{L}]$ of \mathcal{L} does not exceed 2 by (48). Hence, $E[\mathcal{L}]$ is a bounded subset of $\mathcal{C}(\Delta_{n-1})$. By Lemma (47), functions from $E[\mathcal{L}]$ are uniformly equicontinuous. Applying the Arcellà-Ascoli theorem, we conclude that the closure of $E[\mathcal{L}]$ is compact.²¹

It remains to show that $E[\mathcal{L}]$ is closed and convex. The set \mathcal{L} is convex by Theorem 1 and E maps convex combinations to convex combinations, hence $E[\mathcal{L}]$ is convex. To show that it is closed, consider a sequence of functions $h^{(l)} \in E[\mathcal{L}]$ converging to some h and show that the limit belongs to $E[\mathcal{L}]$. Convergence in $\|\cdot\|$ implies pointwise convergence and hence h is equal to zero at the boundary of the simplex. At any \mathbf{p} from the interior, we obtain that the sequence of price indices $P^{(l)}(\mathbf{p})/P^{(l)}(\mathbf{e})$ corresponding to $h^{(l)}$ converges to $g(\mathbf{p}) = \exp((1 + \max_i |\ln p_i|)^2 \cdot h(\mathbf{p}))$. As concavity is preserved under pointwise limits, g is a non-negative concave function on $\Delta_{n-1} \cap \mathbb{R}_{++}^n$ and, hence, there is a preference with price index $P = g$. Therefore, $h = E[\ln P]$ and so $E[\mathcal{L}]$ is closed. \square

By Lemma 2, one can think of \mathcal{L} and the set of all homothetic preferences as a closed convex subset of $\mathcal{C}(\Delta_{n-1})$ and thus can use Choquet theory.

²¹Consider a subset \mathcal{T} of the set $\mathcal{C}(X)$ of continuous functions on a compact set X with sup-norm. Arcellà-Ascoli theorem states that the closure $\overline{\mathcal{T}}$ of \mathcal{T} is compact in $\mathcal{C}(X)$ if \mathcal{T} is bounded and functions from \mathcal{T} are uniformly equicontinuous.

B Proofs

B.1 Proof of Theorem 1

Proof. By the result of [Eisenberg \(1961\)](#), we know that the aggregate consumer exists and her preference corresponds to the following utility function

$$u_{\text{aggregate}}(\mathbf{x}) = \max \left\{ \prod_{k=1}^m \left(\frac{u_k(\mathbf{x}_k)}{\beta_k} \right)^{\beta_k} : \mathbf{x}_k \in \mathbb{R}_+^n, \quad k = 1, \dots, m, \quad \sum_{k=1}^m \mathbf{x}_k = \mathbf{x} \right\}. \quad (52)$$

Our goal is to compute the corresponding price index $P_{\text{aggregate}}(\mathbf{p})$ and check that it satisfies the identity

$$\ln(P_{\text{aggregate}}(\mathbf{p})) = \sum_{k=1}^m \beta_k \cdot \ln(P_k(\mathbf{p})). \quad (53)$$

As an intermediate step, we compute the indirect utility of the aggregate consumer. Recall that the indirect utility for a consumer with preference \succeq represented by a utility function u is given by

$$v(\mathbf{p}, b) = \max \{ u(\mathbf{x}) : \mathbf{x} \in \mathbb{R}_+^n, \langle \mathbf{x}, \mathbf{p} \rangle \leq b \}$$

and is related to the price index by

$$v(\mathbf{p}, b) = \frac{b}{P(\mathbf{p})}. \quad (54)$$

For the aggregate consumer, we get

$$v_{\text{aggregate}}(\mathbf{p}, b) = \max \left\{ \prod_{k=1}^m \left(\frac{u_k(\mathbf{x}_k)}{\beta_k} \right)^{\beta_k} : \mathbf{x}_k \in \mathbb{R}_+^n, \quad k = 1, \dots, m, \quad \left\langle \mathbf{p}, \sum_{k=1}^m \mathbf{x}_k \right\rangle \leq b \right\}. \quad (55)$$

Plug in $b = 1$ and consider an optimal collection of bundles \mathbf{x}_k , $k = 1, \dots, m$ in (55). Denote their prices $\langle \mathbf{p}, \mathbf{x}_k \rangle$ by α_k . Our goal is to show that $\alpha_k = \beta_k$. The argument is along the lines of [Eisenberg \(1961\)](#). Rescale each bundle \mathbf{x}_k to make its price equal to β_k . We obtain a new collection of bundles $\mathbf{x}'_k = \frac{\beta_k}{\alpha_k} \cdot \mathbf{x}_k$ which also satisfies the aggregate budget constraint $\langle \mathbf{p}, \sum_{k=1}^m \mathbf{x}'_k \rangle \leq 1$. By the optimality of \mathbf{x}_k , the product of utilities $\prod_k (u_k(\mathbf{x}_k))^{\beta_k}$ is at least as big as $\prod_k (u_k(\mathbf{x}'_k))^{\beta_k}$. By homogeneity of utilities, this inequality on the products can be rewritten as follows:

$$1 \leq \prod_{k=1}^m \left(\frac{\alpha_k}{\beta_k} \right)^{\beta_k}.$$

Taking logarithm, we get an equivalent inequality

$$0 \leq \sum_{k=1}^m \beta_k \cdot \ln \frac{\alpha_k}{\beta_k}. \quad (56)$$

The concavity of the logarithm implies an upper bound on the right-hand side

$$\sum_{k=1}^m \beta_k \cdot \ln \frac{\alpha_k}{\beta_k} \leq \ln \left(\sum_k \beta_k \cdot \frac{\alpha_k}{\beta_k} \right) = \ln \left(\sum_k \alpha_k \right) \leq \ln(1) = 0. \quad (57)$$

Inequalities (56) and (57) can only be compatible if they are, in fact, equalities. As the logarithm is strictly concave, the equality between the first two expressions in (57) implies that the ratio $\frac{\alpha_k}{\beta_k}$ is a constant independent of k . Since the average value of the logarithms is zero, this constant equals one. We conclude that $\alpha_k = \beta_k$.

We proved that $\langle \mathbf{p}, \mathbf{x}_k \rangle = \beta_k$, for any optimal collection of bundles \mathbf{x}_k , $k = 1, \dots, m$, from the optimization problem (55). In particular, the inequality $\langle \mathbf{p}, \mathbf{x}_k \rangle \leq \beta_k$ always holds at the optimum. Therefore, we can replace the budget constraint of the aggregate consumer $\langle \mathbf{p}, \sum_{k=1}^m \mathbf{x}_k \rangle \leq 1$ with a stronger requirement of individual budget constraints $\langle \mathbf{p}, \mathbf{x}_k \rangle \leq \beta_k$, $k = 1, \dots, m$, and this modification will not alter the value:

$$v_{\text{aggregate}}(\mathbf{p}, B) = \max \left\{ \prod_{k=1}^m \left(\frac{u_k(\mathbf{x}_k)}{\beta_k} \right)^{\beta_k} : \mathbf{x}_k \in \mathbb{R}_+^n, \quad \langle \mathbf{p}, \mathbf{x}_k \rangle \leq \beta_k, \quad k = 1, \dots, m \right\}.$$

The maximization of the product reduces to maximizing each term $u_k(\mathbf{x}_k)$ separately over the corresponding budget set $\langle \mathbf{p}, \mathbf{x}_k \rangle \leq \beta_k$, which gives the indirect utility of consumer k :

$$v_r(\mathbf{p}, B) = \prod_{k=1}^m \left(\frac{\max \left\{ u_k(\mathbf{x}_k) : \mathbf{x}_k \in \mathbb{R}_+^n, \quad \mathbf{p} \cdot \mathbf{x}_k \leq \beta_k \right\}}{\beta_k} \right)^{\beta_k} = \prod_{k=1}^m \left(\frac{v_k(\mathbf{p}, \beta_k)}{\beta_k} \right)^{\beta_k}.$$

Expressing each indirect utility through price indices via formula (54), we end up with the following equality: $P_{\text{aggregate}}(\mathbf{p}) = \prod_{k=1}^m (P_k(\mathbf{p}))^{\beta_k}$. Taking logarithm of both sides, we obtain identity (53) completing the proof. \square

B.2 Generalizations and proofs of Theorems 2 and 3

We first prove Theorem 3 and then formulate and prove a general result containing both Theorem 2 and Theorem 3 as particular cases.

Recall that $\mathcal{D}^{\text{complete}}$ is the completion of a domain \mathcal{D} , i.e., the minimal closed domain invariant with respect to aggregation and containing \mathcal{D} . The set of indecomposable preferences in \mathcal{D} is denoted by $\mathcal{D}^{\text{indec}}$ and contains all preferences $\succsim \in \mathcal{D}$ that cannot be represented as an aggregation of two distinct preferences $\succsim', \succsim'' \in \mathcal{D}$.

For the reader's convenience, we repeat the statement of Theorem 3.

Theorem. *If \mathcal{D} is a closed domain such that $\mathcal{D} = \mathcal{D}^{\text{complete}}$, then a preference \succsim belongs to \mathcal{D} if and only if there exists a Borel probability measure μ supported on $\mathcal{D}^{\text{indec}}$ such that the price index*

$P = P_{\succsim}$ can be represented as follows

$$\ln P(\mathbf{p}) = \int_{\mathcal{D}^{\text{indec}}} \ln P_{\succsim'}(\mathbf{p}) d\mu(\succsim') \quad (58)$$

for any vector of prices $\mathbf{p} \in \mathbb{R}_{++}^n$.

The requirement that μ is supported on $\mathcal{D}^{\text{indec}}$ means that the complement of this set of preferences has μ -measure zero. The topology on preferences and logarithmic price indices is described in Appendix A. The Borel structure is defined by this topology.

We will need the Choquet theorem formulated below. Consider a (not necessarily convex) subset X of a linear space. A point $x \in X$ is an extreme point of X if it cannot be represented as $\alpha x' + (1 - \alpha)x''$ with $\alpha \in (0, 1)$ and distinct $x', x'' \in X$. All extreme points of X are denoted by X^{extrem} .

Theorem (Choquet's theorem; see Phelps (2001), Section 3). *If X is a metrizable compact convex subset of a locally convex space, then a point x belongs to X if and only if there is a Borel probability measure μ on X supported on X^{extrem} such that*

$$x = \int_{X^{\text{extrem}}} x' d\mu(x'). \quad (59)$$

In our application, X will be a subset of the Banach space of continuous functions with the sup-norm. A Banach space is a complete separable normed space. Each such space is locally convex and metrizable via the metric induced by the norm.

The identity (59) is to be understood in the weak sense, i.e., for any continuous linear functional F

$$F[x] = \int_{X^{\text{extrem}}} F[x'] d\mu(x').$$

Proof of Theorem 3. A homothetic preference \succsim is represented by a family of equivalent logarithmic price indices which differ by a constant. Let \mathcal{L} be the set of all classes of equivalent logarithmic price indices corresponding to homothetic preferences. Denote by $\mathcal{L}_{\mathcal{D}}$ the subset of \mathcal{L} corresponding to the domain \mathcal{D} . The set $\mathcal{L}_{\mathcal{D}}$ is closed and convex. Indeed, convexity follows from invariance of \mathcal{D} by Corollary 3 and closedness of \mathcal{D} is inherited by $\mathcal{L}_{\mathcal{D}}$ as the topologies on preferences and logarithmic price indices are aligned.

By Lemma 2, the set \mathcal{L} admits an affine isometric compact embedding E in the Banach space $\mathcal{C}(\Delta_{n-1})$ of continuous functions on the simplex Δ_{n-1} with the sup-norm. Since $\mathcal{L}_{\mathcal{D}}$ is a closed convex subset of \mathcal{L} , the embedding $E[\mathcal{L}_{\mathcal{D}}]$ is a compact convex subset of $\mathcal{C}(\Delta_{n-1})$.

Applying the Choquet theorem to $X = E[\mathcal{L}_{\mathcal{D}}]$, we conclude that a logarithmic price index $\ln P$ belongs to $\mathcal{L}_{\mathcal{D}}$ if and only if there is a measure μ supported on X^{extrem} such that $x = E[\ln P]$ is given by the integral of the form (59).

As there is a natural bijection between \mathcal{D} and X , we can assume that μ is a measure on \mathcal{D} . By Theorem 1, a preference \succsim' is indecomposable in $\mathcal{D}^{\text{indec}}$ if and only if its logarithmic price index $\ln P_{\succsim'}$ cannot be represented as a convex combination of two non-equivalent price indices from \mathcal{D} (Corollary 9). Hence, \succsim' belongs to $\mathcal{D}^{\text{indec}}$ if and only if $E[\ln P_{\succsim'}]$ is in X^{extrem} . We obtain that \succsim with a price index P is contained in the completion $\mathcal{D}^{\text{complete}}$ if and only if

$$E[\ln P] = \int_{\mathcal{D}^{\text{indec}}} E[\ln P_{\succsim'}] d\mu(\succsim') \quad (60)$$

for some μ supported on $\mathcal{D}^{\text{indec}}$. To get the desired pointwise identity (58), it remains to apply an appropriate linear functional on both sides.

Let $F_{\mathbf{p}}$ be the functional on $\mathcal{C}(\Delta_{n-1})$ evaluating a function at some $\mathbf{p} \in \Delta_{n-1}$. This functional is continuous and the family of such functionals with $\mathbf{p} \in \Delta_{n-1} \cap \mathbb{R}_{++}^n$ separates points, i.e., if two functions are not equal, there is a functional taking different values on them. Hence, (60) is equivalent to the following identity

$$F_{\mathbf{p}}[E[\ln P]] = \int_{\mathcal{D}^{\text{indec}}} F_{\mathbf{p}}[E[\ln P_{\succsim'}]] d\mu(\succsim')$$

for all $\mathbf{p} \in \Delta_{n-1} \cap \mathbb{R}_{++}^n$. Plugging in the explicit form (51) of the embedding E , we conclude that $\succsim \in \mathcal{D}$ if and only if its logarithmic price index \succsim can be represented as follows

$$\frac{\ln P(\mathbf{p}) - \ln P(\mathbf{e})}{(1 + \max_i |\ln p_i|)^2} = \int_{\mathcal{D}^{\text{indec}}} \frac{\ln P_{\succsim'}(\mathbf{p}) - \ln P_{\succsim'}(\mathbf{e})}{(1 + \max_i |\ln p_i|)^2} d\mu(\succsim')$$

for all $\mathbf{p} \in \Delta_{n-1} \cap \mathbb{R}_{++}^n$. Multiplying both sides by the denominator, we get

$$\ln P(\mathbf{p}) = \text{const} + \int_{\mathcal{D}^{\text{indec}}} \ln P_{\succsim'}(\mathbf{p}) d\mu(\succsim') \quad (61)$$

for $\mathbf{p} \in \Delta_{n-1} \cap \mathbb{R}_{++}^n$. Since $P(\alpha \cdot \mathbf{p}) = \alpha \cdot P(\mathbf{p})$, the identity extends to \mathbb{R}_{++}^n . As price indices that differ by a constant correspond to the same preference, the constant in (61) can be absorbed by $\ln P$. This completes the proof of Theorem 3. \square

With the help of Theorem 3, we can prove that, without any assumptions on the domain \mathcal{D} , the completion $\mathcal{D}^{\text{complete}}$ is obtained by continuous aggregation of preferences from $\overline{\mathcal{D}}^{\text{indec}}$, i.e., of indecomposable preferences from the closure of \mathcal{D} . This result extends both Theorem 3 and Theorem 2.

Theorem 4. *For any domain \mathcal{D} , a preference \succsim belongs to its completion $\mathcal{D}^{\text{complete}}$ if and only if the price index P of \succsim admits the following representation*

$$\ln P(\mathbf{p}) = \int_{\overline{\mathcal{D}}^{\text{indec}}} \ln P_{\succsim'}(\mathbf{p}) d\mu(\succsim'), \quad \mathbf{p} \in \mathbb{R}_{++}^n, \quad (62)$$

for some Borel probability measure μ supported on $\overline{\mathcal{D}}^{\text{indec}}$.

Theorem 3 corresponds to closed invariant domains \mathcal{D} and Theorem 2 is a corollary since $\overline{\mathcal{D}}^{\text{indec}}$ is a subset of $\overline{\mathcal{D}}$.

Our proof of Theorem 4 relies on already proved Theorem 3 and on Milman's converse to the Krein-Milman theorem. Recall that $\text{conv}[Z]$ denotes the closed convex hull of a set Z .

Proposition 12 (Milman; see Phelps (2001), Proposition 1.5). *If X is a compact convex subset of a locally convex space and $X = \text{conv}[Z]$, then extreme points X^{extrem} are contained in the closure \overline{Z} .*

From the definition of extreme points it is immediate that if $Z \subset X$, then any extreme point of X contained in Z is an extreme point of Z . Hence, the conclusion of Proposition 12 can be strengthened as $X^{\text{extrem}} \subset \overline{Z}^{\text{extrem}}$.

Since the completion corresponds to taking closed convex hull (Corollary 3) and indecomposable preferences correspond to extreme points, Proposition 12 implies the following corollary.

Corollary 13. *For any preference domain \mathcal{D} , all indecomposable preferences of its completion are contained in indecomposable preferences of its closure, i.e., $(\mathcal{D}^{\text{complete}})^{\text{indec}} \subset \overline{\mathcal{D}}^{\text{indec}}$.*

With this corollary, Theorem 4 follows from Theorem 3 almost immediately.

Proof of Theorem 4. Apply Theorem 3 to the closed invariant domain $\mathcal{D}' = \mathcal{D}^{\text{complete}}$. We get that \succsim is in $\mathcal{D}^{\text{complete}}$ if and only if there is a measure supported $(\mathcal{D}^{\text{complete}})^{\text{indec}}$ such that

$$\ln P(\mathbf{p}) = \int_{(\mathcal{D}^{\text{complete}})^{\text{indec}}} \ln P_{\succsim'}(\mathbf{p}) \, d\mu(\succsim').$$

By Corollary 13, $(\mathcal{D}^{\text{complete}})^{\text{indec}} \subset \overline{\mathcal{D}}^{\text{indec}}$ which completes the proof. \square

B.3 Indecomposability in the full domain and proof of Proposition 5

We will need the following simple lemma.

Lemma 3. *Consider a function $h(\mathbf{t}) = t_1^\alpha \cdot t_2^{1-\alpha}$ where $\alpha \in (0, 1)$ and $\mathbf{t} \in \mathbb{R}_{++}^2$. If $\mathbf{t} \neq \text{const} \cdot \mathbf{t}'$, then*

$$h(\lambda \mathbf{t} + (1 - \lambda) \mathbf{t}') > \lambda h(\mathbf{t}) + (1 - \lambda) h(\mathbf{t}')$$

for any $\lambda \in (0, 1)$.

Proof. The result follows from strict concavity of $g(\lambda) = h(\lambda \mathbf{t} + (1 - \lambda) \mathbf{t}')$. To demonstrate strict concavity, it is enough to show that the second derivative $g''(\lambda) < 0$. After a linear change of variable, this requirement boils down to negativity of the second derivative of $\gamma^\alpha (1 + \gamma)^{1-\alpha}$ with respect to γ . We omit the elementary computation. \square

With the help of this lemma, we prove Proposition 5.

Proof of Proposition 5. A utility function u is of the form (38) if and only if the corresponding price index is also piecewise linear:

$$P = \min_{c \in C} \left(\sum_{j=1}^n c_j p_j \right), \quad (63)$$

where $C \subset \mathbb{R}_+^n$ is finite or countable. We need to show that preferences with such price indices are indecomposable. Towards a contradiction, assume that P of the form (63) can be represented as

$$\ln P = \alpha \ln P_1 + (1 - \alpha) \ln P_2,$$

where P_1 and P_2 are price indices representing distinct homothetic preferences and $\alpha \in (0, 1)$. Hence, P_1 and P_2 are not proportional to each other, i.e., the ratio $P_1/P_2 \neq \text{const}$. By continuity of price indices, this means that there is a linearity region of P where $P_1/P_2 \neq \text{const}$. Therefore, we can find $\mathbf{p}, \mathbf{p}' \in \mathbb{R}_{++}^n$ from the same linearity region of P such that

$$\frac{P_1(\mathbf{p})}{P_2(\mathbf{p})} \neq \frac{P_1(\mathbf{p}')}{P_2(\mathbf{p}')} \quad (64)$$

By homogeneity of price indices, we can assume that \mathbf{p} and \mathbf{p}' are normalized so that $P(\mathbf{p}) = P(\mathbf{p}') = 1$. Since \mathbf{p} and \mathbf{p}' belong to the same linearity region, the value of P at the mid-point $\mathbf{p}'' = (\mathbf{p} + \mathbf{p}')/2$ is also equal to 1. Therefore,

$$1 = P(\mathbf{p}'') = P_1(\mathbf{p}'')^\alpha P_2(\mathbf{p}'')^{1-\alpha} \geq \left(\frac{1}{2} P_1(\mathbf{p}) + \frac{1}{2} P_1(\mathbf{p}') \right)^\alpha \left(\frac{1}{2} P_2(\mathbf{p}) + \frac{1}{2} P_2(\mathbf{p}') \right)^{1-\alpha},$$

where we used concavity of P_1 and P_2 . By Lemma 3, the right-hand side admits the following lower bound

$$\left(\frac{1}{2} P_1(\mathbf{p}) + \frac{1}{2} P_1(\mathbf{p}') \right)^\alpha \left(\frac{1}{2} P_2(\mathbf{p}) + \frac{1}{2} P_2(\mathbf{p}') \right)^{1-\alpha} > \frac{1}{2} P_1(\mathbf{p})^\alpha P_2(\mathbf{p})^{1-\alpha} + \frac{1}{2} P_1(\mathbf{p}')^\alpha P_2(\mathbf{p}')^{1-\alpha}.$$

The right-hand side can be rewritten as

$$\frac{1}{2} P_1(\mathbf{p})^\alpha P_2(\mathbf{p})^{1-\alpha} + \frac{1}{2} P_1(\mathbf{p}')^\alpha P_2(\mathbf{p}')^{1-\alpha} = \frac{1}{2} P(\mathbf{p}) + \frac{1}{2} P(\mathbf{p}') = 1.$$

We end up with a contradictory inequality $1 > 1$. Therefore, P cannot be represented as a convex combination (64) and we conclude that the corresponding preference is indecomposable. \square

Let us explore whether there are other indecomposable preferences in the full domain. For simplicity, we focus on the case of $n = 2$ goods. As opposed to preferences with piecewise linear u and P considered in Proposition 5, we examine preferences with price index P that is strictly concave in a neighborhood of a certain point.

We say that a function of one variable $h = h(t)$ is strictly concave in the neighborhood of $t = t_0$ if there is $\varepsilon > 0$ and $\delta > 0$ such that the second derivative of $h''(t) < -\delta$ for almost all t in the ε -neighborhood $[t_0 - \varepsilon, t_0 + \varepsilon]$ of t_0 . We note that the second derivative exists almost everywhere for any concave function by Alexandrov's theorem.

Proposition 13. *Consider a preference \succsim over two goods with price index P . If there is a point $\mathbf{p}_0 \in \mathbb{R}_{++}^2$ and a direction $\mathbf{r} \in \mathbb{R}^2 \setminus \{0\}$ such that $g(t) = P(\mathbf{p}_0 + t \cdot \mathbf{r})$ is strictly concave in the neighborhood of $t = 0$, then \succsim is not indecomposable.*

Proof. Since $P(\alpha \cdot \mathbf{p}) = \alpha \cdot P(\mathbf{p})$, the values of the price index on the line $p_2 = 1$ determine its values everywhere by $P(p_1, p_2) = p_2 \cdot P(p_1/p_2, 1)$. Accordingly, the condition from the statement is equivalent to the existence of t_0 such that $g(t) = P(t, 1)$ is strictly concave in the neighborhood of t_0 .

Let us show that if $g(t) = P(t, 1)$ is strictly concave in the neighborhood of t_0 , then the preference \succsim is an aggregation of some distinct \succsim_1 and \succsim_2 . By strict concavity $g'' < -\delta$ on $[t_0 - \varepsilon, t_0 + \varepsilon]$ for some $\varepsilon, \delta > 0$. Let $\varphi(z)$ be a smooth function on \mathbb{R} not equal to zero identically and vanishing outside of the interval $[-1, 1]$ together with all its derivatives. For example, one can take

$$\varphi(z) = \exp\left(-\frac{1}{1-z^2}\right)$$

for $z \in (-1, 1)$ and zero outside. Define

$$g_1(t) \equiv (1 + \gamma \cdot \varphi(\varepsilon(t - t_0))) g(t) \quad \text{and} \quad g_2(t) \equiv \frac{1}{1 + \gamma \cdot \varphi(\varepsilon(t - t_0))} g(t)$$

for some constant $\gamma > 0$. Note that $g_1 = g_2 = g$ outside the ε -neighborhood of t_0 . The second derivatives of g_1 and g_2 continuously depends on γ and, for $\gamma = 0$, the derivatives are bounded from above by $-\delta$ in the ε -neighborhood of t_0 . Hence, for small enough $\gamma > 0$, the second derivative is non-positive, i.e., both g_1 and g_2 are concave.

Define $P_1(\mathbf{p}) = p_2 \cdot g_1(p_1/p_2)$ and $P_2(\mathbf{p}) = p_2 \cdot g_2(p_1/p_2)$. These are non-negative homogeneous concave functions that are not proportional to each other. Hence, P_1 and P_2 are price indices of some distinct preferences \succsim_1 and \succsim_2 . By the construction,

$$\ln P = \frac{1}{2} \ln P_1 + \frac{1}{2} \ln P_2$$

and thus \succsim is the aggregate preference for a pair of consumers with preferences \succsim_1 and \succsim_2 and equal incomes. \square

It may seem that Propositions 5 and 13 cover all possible preferences in the case of two goods: intuitively, a price index P is either piecewise linear or there is a point in the neighborhood of

which P is strictly concave. However, there are pathological examples not captured by the two propositions.

Any concave function f on \mathbb{R}_+ can be represented as

$$f(t) = f(0) - \int_0^t \left(\int_0^s d\nu(q) \right) ds$$

for some positive measure ν on \mathbb{R}_+ . This ν is uniquely defined distributional second derivative of f . Abusing the notation, we will write $\nu = f''$. Note that the classic pointwise second derivative (where exists) equals the density of the absolutely continuous component of ν .

Propositions 5 and 13 address the cases where the second derivative of $P(t, 1)$ is either an atomic measure with nowhere dense set of atoms or has an absolutely continuous component with a strictly negative density on a certain small interval.

Recall that ν is called singular if there is a set of zero Lebesgue measure such that its complement has ν -measure zero. For example, atomic measures with discrete set of atoms are singular, but there are other singular measures such as non-atomic measures supported on a Cantor set or atomic measures with everywhere dense set of atoms.

Proposition 14. *If \succsim is a preference over two goods with price index P such that the second distributional derivative of $g(t) = P(t, 1)$ is singular, then \succsim is indecomposable in the full domain.*

We see that the set of indecomposable preferences is broader than suggested by Proposition 5. Note that, in the particular case of two goods, Proposition 5 is a direct corollary of Proposition 14.

Proof. It is enough to show that if \succsim is an aggregation of two distinct preferences \succsim_1 and \succsim_2 , then the second distributional derivative of g has a non-zero absolutely continuous component. In other words, we need to show that the classic derivative $g'' \neq 0$ on a set of positive Lebesgue measure.

Let P_1 and P_2 be price indices of \succsim_1 and \succsim_2 . Since the preferences are distinct, $P_1 \neq \text{const} \cdot P_2$. By the assumption, $P = P_1^\alpha \cdot P_2^{1-\alpha}$ with some $\alpha \in (0, 1)$. Without loss of generality, we can assume that $\alpha = 1/2$. Indeed, if $\alpha \neq \frac{1}{2}$, one can define new price indices $P'_1 = P_1^{\alpha-\varepsilon} \cdot P_2^{1-\alpha+\varepsilon}$ and $P'_2 = P_1^{\alpha+\varepsilon} \cdot P_2^{1-\alpha-\varepsilon}$ for some $\varepsilon < \min\{\alpha, 1-\alpha\}$ so that $P = \sqrt{P'_1 \cdot P'_2}$.

Hence, $g = \sqrt{g_1 \cdot g_2}$ where $g_1(t) = P_1(t, 1)$ and $g_2(t) = P_2(t, 1)$ are non-negative concave functions not proportional to each other. Computing the classic second derivative of g , we obtain

$$g'' = \frac{g_1'' \cdot g_2 + g_2'' \cdot g_1}{2\sqrt{g_1 \cdot g_2}} - \frac{(g_1' \cdot g_2 - g_2' \cdot g_1)^2}{4(g_1 \cdot g_2)^{3/2}}.$$

Both terms are non-positive. The numerator in the second term can be rewritten as follows

$$(g_1' \cdot g_2 - g_2' \cdot g_1)^2 = \left((g_2)^2 \cdot \left(\frac{g_1}{g_2} \right)' \right)^2$$

Since the ratio g_1/g_2 is non-constant, its derivative $(g_1/g_2)'$ is non-zero on a set of positive measure. Thus the distributional derivative g'' contains a non-zero absolutely continuous component. We conclude that preferences such that g'' has no absolutely continuous component are indecomposable. \square

The approach from Proposition 14 extends to $n > 2$ goods. It can be used to show that if the distributional second derivative of

$$g(t) = P(p_1, \dots, p_{i-1}, t, p_{i+1}, \dots, p_n)$$

is singular for any $i = 1, \dots, n$ and any fixed

$$p_{-i} = (p_1, \dots, p_{i-1}, p_{i+1}, \dots, p_n) \in \mathbb{R}_{++}^{n-1},$$

then the corresponding preference is indecomposable. This result extends Proposition 5.

B.4 Proof of Proposition 7

Proof. Consider a Leontief preference \succsim over Cobb-Douglas composite goods. It corresponds to a utility function

$$u(\mathbf{x}) = \min_{\mathbf{a} \in A} \left\{ a_0 \cdot \prod_{i=1}^n x_i^{a_i} \right\},$$

where A is finite or countably infinite subset of $\mathbb{R}_{++} \times \Delta_{n-1}$. Assume that \succsim is non-trivial, i.e., there are $\mathbf{a}, \mathbf{a}' \in A$ such that $(a_1, \dots, a_n) \neq (a'_1, \dots, a'_n)$. Intersection of convex sets corresponds to convexification of the maximum of their support functions (Aliprantis and Border, 2013, Theorem 7.56). Since price indices are support functions of upper contour sets up to a sign, the price index corresponding to \succsim takes the following form

$$P(\mathbf{p}) = \text{cav} \left[\max_{\mathbf{a} \in A} \left\{ \frac{1}{a_0} \cdot \prod_{i=1}^n p_i^{a_i} \right\} \right],$$

where cav denotes concavification.

Let us focus on the case of $n = 2$ goods. In this case, P has a particularly simple structure. The positive orthant \mathbb{R}_{++}^2 is partitioned into a finite or countably infinite number of cones of two types: (I) cones where P is linear (II) cones where P coincides with the Cobb-Douglas price index $1/a_0 \cdot \prod_{i=1}^n p_i^{a_i}$ for some $\mathbf{a} \in A$. The cones of type (I) and (II) interlace and derivatives of P change continuously. Note that there must be at least one cone of type (I) as otherwise P would be a price index of a standard Cobb-Douglas preferences which is ruled out by the non-triviality assumption.

Let us show that such a preference \succsim over two goods is indecomposable. Towards a contradiction, assume that

$$\ln P = \alpha \cdot \ln P_1 + (1 - \alpha) \ln P_2, \quad \alpha \in (0, 1), \quad (65)$$

where P_1 and P_2 corresponds to two distinct preferences \succsim_1 and \succsim_2 exhibiting complementarity. As in the proof of Proposition 5, one shows that in each linearity region of P , the price indices P_1 and P_2 are proportional to each other. In other words, $P_1 = \text{const} \cdot P_2$ in each cone of type (I), where the constant can depend on the cone.

Recall that by (4), the partial derivative of a logarithmic price index with respect to $\ln p_i$ is the budget shares of good i . Denote the budget shares for \succsim , \succsim_1 , and \succsim_2 by s_i , $s_{1,i}$, and $s_{2,i}$, respectively. Since P_1 and P_2 are proportional in cones of type (I), we obtain that $s_i = s_{1,i} = s_{2,i}$ there.

Consider cones of type (II). In these cones, s_i is constant since budget shares are constant for Cobb-Douglas preferences. Taking the partial derivative on both sides of (65), we get

$$s_i(p_1, p_2) = \alpha \cdot s_{1,i}(p_1, p_2) + (1 - \alpha)s_{2,i}(p_1, p_2).$$

The budget shares depend only on the ratio of prices. By complementarity, $s_{1,i}$ and $s_{2,i}$ must be non-increasing functions of the ratio p_{3-i}/p_i . Note that if s_i is constant and one of $s_{1,i}$ or $s_{2,i}$ increases, the other must decrease violating the monotonicity requirement. Hence, in all the cones of type (II), s_i , $s_{1,i}$ and $s_{2,i}$ are constant. These constants must be all equal. Indeed, suppose that $s_i = c$, $s_{1,i} = c_1$ and $s_{2,i} = c_2$ in some cone of type (II) with $c_1 \neq c_2$. Since s_i is continuous and coincides with $s_{1,i}$ and $s_{2,i}$ in the neighboring cone of type (I), $s_{1,i}$ and $s_{2,i}$ are discontinuous on the boundary between the two cones and at least one of these discontinuities necessarily violates the monotonicity requirement. Thus in cones of type (II), $s_i = s_{1,i} = s_{2,i}$.

We conclude that $s_i = s_{1,i} = s_{2,i}$ everywhere, i.e., partial derivatives of $\ln P$, $\ln P_1$, and $\ln P_2$ coincide. Hence, $\ln P_1 = \ln P_2 + \text{const}$, i.e., P_1 and P_2 are proportional. Thus $\succsim_1 = \succsim_2$ contradicting the assumption that the two preferences are distinct. We conclude that, for $n = 2$, any non-trivial Leontief preference over Cobb-Douglas composite goods is indecomposable. \square

B.5 Proof of Proposition 10

The proof relies on the following lemma showing that if preferences in two populations have budget shares that are close, then ε -equilibrium price vectors are close as well.

Lemma 4. *Consider two populations of m consumers with the same budgets b_1, \dots, b_m but different preferences over n goods: $\succsim_1, \dots, \succsim_m$ in the first population and $\succsim'_1, \dots, \succsim'_m$ in the second one. Assume that the budget shares $s_{k,i}(\mathbf{p})$ and $s'_{k,i}(\mathbf{p})$ differ by at most some $\delta > 0$ for any consumer k , good i , and price \mathbf{p} . Then any ε -equilibrium price vector for one population is an $(\varepsilon + n\delta)$ -equilibrium price vector for the other.*

Proof of Lemma 4. The demand of an agent k for a good i can be expressed through budget shares

as follows:

$$D_{k,i}(\mathbf{p}, b_k) = s_{k,i}(\mathbf{p}) \cdot \frac{b_k}{p_i}. \quad (66)$$

Let $\mathbf{x}_1 + \dots + \mathbf{x}_m$ and $\mathbf{x}'_1 + \dots + \mathbf{x}'_m$ be market demands of the two populations from the statement of the lemma at some vector of prices \mathbf{p} . By (66) and the assumption that budget shares differ by at most δ ,

$$\sum_{i=1}^n p_i \cdot \left| \sum_{k=1}^m x_{k,i} - \sum_{k=1}^m x'_{k,i} \right| \leq \sum_{i=1}^n \sum_{k=1}^m b_k \cdot \max_{i,k} |s_{k,i}(\mathbf{p}) - s'_{k,i}(\mathbf{p})| \leq n \cdot B \cdot \delta.$$

Hence, if \mathbf{p} is an ε -equilibrium price vector for $\succsim'_1, \dots, \succsim'_m$, it is an $(\varepsilon + n\delta)$ -equilibrium price vector for $\succsim_1, \dots, \succsim_m$. \square

To prove the proposition, it remains to show that any preference \succsim over two substitutes can be approximated by the aggregate preference $\succsim' = \succsim_{\text{aggregate}}$ of an auxiliary population with linear preferences so that the budget shares differ by at most ε at any vector of prices and the number of auxiliary agents is of the order of $1/\varepsilon$.

Proof of Proposition 10. Since $s_{\succsim,1} + s_{\succsim,2} = 1$ and budget shares depend on the ratio of prices only, it is enough to ensure that $|s_{\succsim,1}(p_1, 1) - s_{\text{aggregate},1}(p_1, 1)| \leq \varepsilon$ for any $p_1 \in \mathbb{R}_{++}$. As \succsim exhibits substitutability, $s_{\succsim,1}(\cdot, 1)$ is a non-increasing function with values in $[0, 1]$. For any such function f , there is a piecewise-constant function f_ε with at most $1/\varepsilon + 1$ jumps such that the two functions differ by at most ε ; indeed, one can take $f_\varepsilon = \varepsilon \cdot [f/\varepsilon]$, where $[t]$ denotes the integer part of a real number t . By Corollary 6, any piecewise-constant non-decreasing function with values in $[0, 1]$ is equal to $s_{\text{aggregate},1}(\cdot, 1)$ for a population of linear consumers with the number of consumers equal to the number of jumps, marginal rates of substitution given by positions of the jumps, and budgets determined by jumps' magnitude. We conclude that $1/\varepsilon + 1$ linear consumers are enough to approximate budget shares of any preference exhibiting substitutability with precision ε . Combined with Lemma 4, this observation completes the proof. \square

Note that constructing the approximation in a computationally efficient way requires solving the equation $s_{\succsim,1}(p_1, 1) = \varepsilon \cdot l$ multiple times for various \succsim and l . Provided that there is an oracle computing budget shares, one can use binary search for this task.

B.6 Characterization of budget shares for two goods

For a preference \succsim over $n = 2$ goods, consider the budget shares $s_1(\mathbf{p})$ and $s_2(\mathbf{p})$ of these goods. We explore what functions one can get as budget shares. Since $s_1 + s_2 = 1$, we can focus on the budget share s_1 of the first good. As $s_1(\alpha \cdot \mathbf{p}) = s_1(\mathbf{p})$, the budget share can be seen as a function

of one variable $z = p_1/p_2$. Our goal is to characterize functions $h = h(z)$ such that $h = s_1$ for some homothetic preference \succsim .

Lemma 5. *A function $h : \mathbb{R}_{++} \rightarrow \mathbb{R}$ is the budget share of the first good associated with some homothetic preference \succsim over two goods (i.e., $h(p_1/p_2) = s_1(\mathbf{p})$, $\mathbf{p} \in \mathbb{R}_{++}^2$) if and only if*

$$h(z) = \frac{z}{z + Q(z)}, \quad (67)$$

where $Q : \mathbb{R}_{++} \rightarrow \mathbb{R}_+ \cup \{+\infty\}$ is a non-negative non-decreasing function.

Leontief and linear preferences correspond to the two extreme cases: a constant $Q = \text{const}$ and an infinite step function

$$Q = \begin{cases} 0, & z \leq \alpha \\ +\infty & z > \alpha \end{cases},$$

respectively.

Note that for any non-increasing function h with values in $[0, 1]$, the function

$$Q(z) = \frac{1}{h(z)} - 1$$

satisfies the requirement of Lemma 5.

Corollary 14. *Any non-increasing function h with values in $[0, 1]$ is the budget share of the first good for some preference \succsim exhibiting substitutability among the two goods.*

Proof of Lemma 5. Consider a homothetic preference \succsim for $n = 2$ goods. The budget share $s_1(\mathbf{p})$ of the first good satisfies $s_1(\alpha \cdot \mathbf{p}) = \alpha \cdot s_1(\mathbf{p})$ and, hence, $s_1(\mathbf{p}) = s_1(z, 1)$, where $z = p_1/p_2$.

Let us show that, for any \succsim , the function $h(z) = s_1(z, 1)$ admits the representation (67). In other words, we need to show that

$$Q(z) = \frac{z}{s_1(z, 1)} - z$$

is non-negative and non-decreasing. Expressing the budget share through the logarithmic price index by (4) and denoting $\pi(z) = P(z, 1)$, we get

$$Q(z) = \frac{\pi(z)}{\pi'(z)} - z. \quad (68)$$

Note that π is a non-negative non-decreasing concave function of z . Hence, Q is non-negative as well. To show that Q is non-decreasing, let us differentiate both sides of (68). We get

$$Q'(z) = -\frac{\pi(z)}{(\pi'(z))^2} \cdot \pi''(z). \quad (69)$$

We see that $Q' \geq 0$ and so Q is non-decreasing.²² We conclude that $h = s_1(z, 1)$ admits the representation (67).

To prove the converse, consider h of the form (67) with non-negative non-decreasing Q and construct the corresponding preference. The identity (68) suggests how to define the price index. We get

$$\frac{\pi'(z)}{\pi(z)} = \frac{1}{z + Q(z)},$$

where $\pi(z) = P(z, 1)$. Integrating this identity, we obtain

$$\pi(z) = \exp \left(\int_1^z \frac{1}{w + Q(w)} dw \right).$$

By the construction, π is non-negative and non-decreasing. Since the identity (69) is hardwired in the definition of π and the function Q is non-decreasing, we conclude that π'' is non-positive. Hence, π is concave. Define P by

$$P(p_1, p_2) = p_2 \cdot \pi(p_1/p_2).$$

This function is homogeneous, non-negative, non-decreasing, and concave. Thus P is a price index corresponding to some homothetic preference. By the construction, $s_1(z, 1) = h(z)$ completing the proof. \square

²²If π' is not differentiable, the identity (69) is to be understood in the sense of distributional derivatives: π'' is a non-negative measure, and the right-hand side is a measure having density $-\pi(z)/(\pi'(z))^2$ with respect to π'' .