Pareto optimality and welfare functions

Robert J. Carroll

Last revised: 6 May 2026

---

1 Motivation

Most political-economy modeling concerns settings in which several individuals — voters, coalition members, legislators, citizens, regulators — have conflicting preferences over the alternatives at issue. A reform proposal benefits some constituents and disadvantages others. A budget allocation across districts pleases the recipients of any particular tranche and disappoints those who got less. A coalition negotiating a policy faces the structural fact that no single policy can be each member's first-best. In each case, the basic conceptual question is what we can say about "efficient" or "socially desirable" alternatives without first having to decide whose preferences count for how much. The answer that the welfare-economics literature has organized itself around is the concept of *Pareto optimality*: an alternative is Pareto optimal if no available alternative is strictly preferred by at least one individual and weakly preferred by all of them. The criterion does not require interpersonal comparisons of utility, does not single out any individual's preferences as more important than another's, and is the most uncontroversial efficiency notion in the welfare-and-aggregation literature.

Three structural insights organize this handout. First, Pareto optimality is the right efficiency concept precisely because it is the strongest claim one can make about a set of alternatives without imposing an interpersonal weighting (§2–§3). Second, under standard convexity and concavity conditions, every Pareto-optimal alternative *is* the maximizer of some weighted-sum welfare function $\sum_i \alpha_i u_i$ for some non-negative weights $\alpha = (\alpha_1, \ldots, \alpha_n)$ (§4); the converse, that every weighted-sum maximizer is Pareto optimal, holds with appropriate sign conditions on the weights. The two directions together give a complete optimization-theoretic characterization of the Pareto set. Third, the welfare-function representation is the structural bridge to a substantial substantive literature on aggregation: utilitarianism is the special case of equal weights; Bergson–Samuelson welfare functions generalize beyond linear aggregation; bargaining theory (the next handout) defines bargaining problems on Pareto sets and characterizes which Pareto point gets selected; social-choice theory (a forthcoming cluster) addresses what happens when the weights themselves are contested.

This handout is the first in a two-handout cluster on welfare and bargaining. It leans on the convex-sets-and-concave-functions handout (#21) for the separating hyperplane theorem and the structural reading of concavity, and on the static-optimization handout (#22) for Weierstrass-existence. The bargaining handout that follows will use the Pareto set developed here as its primitive — a bargaining problem is a Pareto set together with a disagreement point Pareto-dominated by at least one alternative — and will characterize the Nash and Kalai–Smorodinsky solutions.

2 Pareto dominance and Pareto optimality

Consider a coalition of legislators allocating a fixed budget across districts, each member of the coalition representing a different district and preferring a larger share for her own. A reform commission considering policy proposals against the status quo: each commissioner prefers the proposal she favors, and they disagree. A two-candidate election in which both candidates choose policy positions, each preferring positions closer to her ideal point. In each setting, several individuals with conflicting preferences face a common alternative space, and the structural question is which alternatives are *efficient* in the sense that no alternative is unambiguously better. The formalization is the Pareto-dominance relation, defined on the alternative space using individuals' utility functions.

Definition 1. Let $N = \{1, 2, \ldots, n\}$ be a finite set of individuals, A a set of alternatives, and $u_i : A \to \mathbb{R}$ each individual i's utility function.

- Alternative y *Pareto dominates* x, written $y \succ_P x$, if $u_i(y) \geq u_i(x)$ for every $i \in N$ and $u_j(y) > u_j(x)$ for at least one $j \in N$.
- Alternative y *weakly Pareto dominates* x, written $y \succsim_P x$, if $u_i(y) \geq u_i(x)$ for every $i \in N$.
- An alternative $x \in A$ is *Pareto optimal* (or *Pareto efficient*) if there is no alternative $y \in A$ that Pareto dominates x. The set of Pareto optimal alternatives is the *Pareto set* (or *Pareto frontier*), denoted $\mathrm{P}(A)$.

A few features of the definition are worth flagging. First, Pareto dominance is a binary relation on A that is *reflexive-free* (no alternative Pareto dominates itself, by the strict-inequality clause), *transitive*, and *antisymmetric*; it is a strict partial order on A in the sense of the order-theory handout (#4). Second, the definition is invariant under any monotone transformation of each individual's utility function: if u_i is replaced by $\phi_i \circ u_i$ for a strictly increasing ϕ_i, the Pareto-dominance relation is unchanged. The Pareto criterion is therefore an *ordinal* concept, not a cardinal one — it does not depend on the choice of utility scale, and it does not require interpersonal comparisons of utility. Third, in general, $\mathrm{P}(A)$ is not a singleton; it is typically a substantial subset of A, and the choice of which Pareto optimum to pick *is* where political conflict and bargaining show up.

Example 2 (Two-voter spatial-loss example)**.** Let $A = \mathbb{R}^2$ (a two-dimensional policy space) and let two voters have ideal points $\hat{x}_1, \hat{x}_2 \in \mathbb{R}^2$ with quadratic-loss utilities $u_i(x) = -\|x - \hat{x}_i\|^2$. The Pareto set $\mathrm{P}(A)$ is exactly the line segment between $\hat{x}_1$ and $\hat{x}_2$: every point on the segment is Pareto optimal (any movement toward one ideal point hurts the other voter), and every point off the segment is Pareto-dominated by its orthogonal projection onto the segment (which moves both voters closer to their ideal points or leaves one indifferent and helps the other). With elliptical level curves — corresponding to $u_i(x) = -(x - \hat{x}_i)^\top M_i (x - \hat{x}_i)$ for positive definite M_i — the Pareto set is a curve (not a straight line) connecting the two ideal points; the standard political-economy terminology is *contract curve*. The Pareto set is the substantive object on which the bargaining handout will define bargaining problems.

Example 3 (Coalition allocation across districts)**.** A coalition of n legislators must allocate a fixed budget B across n districts: $A = \{(x_1, \ldots, x_n) : x_i \geq 0, \sum_i x_i = B\}$, the simplex of allocations. Each legislator i values her district's allocation, with utility $u_i(x) = \phi_i(x_i)$ for some increasing ϕ_i. The Pareto set is exactly A itself: every allocation is Pareto optimal, since giving more to any district means taking from another. The substantive content of the example is that the Pareto criterion alone says nothing about which allocation is socially desirable; that question is downstream of efficiency, and is the territory of social-choice theory and bargaining theory.

3 Existence of Pareto optima

Substantive political-economy questions about Pareto optimality usually presume that Pareto-optimal alternatives *exist*; the model would otherwise be vacuous. The structural fact that delivers existence under standard hypotheses is also the structural fact that licenses optimization more generally: under compactness of the alternative space and continuity of the individual utilities, a maximum of any continuous social welfare function exists, and any such maximum is Pareto optimal. Existence of Pareto optima then follows from existence of welfare-maximizers, by way of a clever sufficient condition.

Theorem 4 (Sufficient condition via positive-weight maximization)**.** *Let A be any non-empty set, $u_1, \ldots, u_n : A \to \mathbb{R}$ utility functions, and $\alpha_1, \ldots, \alpha_n > 0$ strictly positive weights. If $x^* \in A$ solves*

$$\max_{y \in A} \sum_{i=1}^{n} \alpha_i u_i(y),$$

then x^* is Pareto optimal.

Proof. Suppose x^* is the welfare-maximizer but is Pareto dominated by some $y \in A$, so $u_i(y) \geq u_i(x^*)$ for all i and $u_j(y) > u_j(x^*)$ for some j. Then $\alpha_i u_i(y) \geq \alpha_i u_i(x^*)$ for all i (since each $\alpha_i > 0$), with strict inequality at $i = j$. Summing, $\sum_i \alpha_i u_i(y) > \sum_i \alpha_i u_i(x^*)$, contradicting the assumption that x^* is the maximizer. □

The intuition is that strictly positive weights ensure no individual's utility gain is "invisible" to the welfare function: any Pareto-improving alternative must be visible as a welfare improvement, so any welfare maximizer is necessarily Pareto optimal. With this in hand, existence of Pareto optima follows under the standard Weierstrass conditions.

Corollary 5 (Existence under compactness and continuity)**.** *Suppose $A \subseteq \mathbb{R}^d$ is compact and each $u_i : A \to \mathbb{R}$ is continuous. Then $\mathrm{P}(A)$ is non-empty.*

Proof. The function $f(y) = \sum_{i=1}^{n} u_i(y)$ is continuous on A as a sum of continuous functions. By Weierstrass (handout #22, leveraging compactness from #7), f attains its maximum on A at some $x^* \in A$. By Theorem 4 (with all $\alpha_i = 1$), x^* is Pareto optimal. □

The corollary is more powerful than it looks: existence of *any* continuous social welfare function with positive weights buys existence of a Pareto optimum, and the most uncontroversial choice of welfare function (the unweighted utilitarian sum) is sufficient. The substantive reading is that the modeler does not need to pin down the "right" weights to argue that the model has a Pareto-optimal solution; she only needs the standard regularity conditions and the structural insight of Theorem 4.

A natural strengthening allows weakly positive weights when utilities are sufficiently well-behaved.

Theorem 6 (Sufficient condition under quasi-concavity)**.** *Suppose $A \subseteq \mathbb{R}^d$ is convex and each u_i is strictly quasi-concave on A. If there exist weights $\alpha_1, \ldots, \alpha_n \geq 0$ (not all zero) such that x^* solves $\max_{y \in A} \sum_i \alpha_i u_i(y)$, then x^* is Pareto optimal.*

Proof sketch. Suppose x^* is Pareto dominated by y, with $u_i(y) \geq u_i(x^*)$ for all i. Let $w = \frac{1}{2}x^* + \frac{1}{2}y \in A$ by convexity. Strict quasi-concavity of u_i implies $u_i(w) > \min\{u_i(x^*), u_i(y)\} = u_i(x^*)$ for each i. Then $\sum_i \alpha_i u_i(w) > \sum_i \alpha_i u_i(x^*)$ (using non-negativity of weights and the strict-positivity of at least one α_j), contradicting the maximization. □

The strengthening is useful in practice because it allows the modeler to put zero weight on some individuals' utilities while still concluding that a maximizer is Pareto optimal — an important degree of freedom for boundary alternatives in bargaining models.

4 Welfare-function characterization with concavity

Theorems 4 and 6 are sufficient conditions: they tell us *from* weighted-sum maximization *to* Pareto optimality. The structurally more interesting question is the converse: given a Pareto-optimal alternative x^*, can we always exhibit a weighted-sum welfare function that x^* maximizes? Under convexity of the alternative space and concavity of the individual utilities, the answer is yes, and the result is the structural payoff of the welfare-economics framework. Every Pareto optimum is the maximizer of *some* weighted-sum utilitarianism — the weights may not be the modeler's preferred ones, but they exist. The proof technique is the separating hyperplane theorem from the convex-sets handout (#21), applied to two convex sets in utility space.

The setup that makes the separation argument work is the *set of utility imputations* — the set of utility profiles attainable from the alternative space, together with all profiles weakly dominated by them.

Definition 7. The *set of utility imputations* associated with A and utilities $u_1, \dots, u_n$ is

$$\mathcal{U} = \big\{\boldsymbol{z} \in \mathbb{R}^n : \exists\, x \in A \text{ such that } (u_1(x), \dots, u_n(x)) \geq \boldsymbol{z} \text{ componentwise}\big\}.$$

Geometrically, $\mathcal{U}$ is the set of all utility profiles attainable at some alternative x, plus everything to the southwest — every profile that some alternative weakly dominates. The set is the natural object on which the welfare-function characterization lives because the welfare function is itself a function on $\mathbb{R}^n$ (the dot product $\alpha \cdot \boldsymbol{z}$ for utility profile $\boldsymbol{z}$ and weight vector α). Under convexity and concavity, $\mathcal{U}$ inherits convex structure.

Lemma 8 (Convexity of imputations)**.** *Suppose $A \subseteq \mathbb{R}^d$ is convex and each u_i is concave on A. Then $\mathcal{U} \subseteq \mathbb{R}^n$ is convex.*

Proof sketch. Take $\boldsymbol{z}, \boldsymbol{z}' \in \mathcal{U}$ with witnesses $x, x' \in A$ such that $(u_1(x), \dots, u_n(x)) \geq \boldsymbol{z}$ and $(u_1(x'), \dots, u_n(x')) \geq \boldsymbol{z}'$. For $\alpha \in [0,1]$, set $x_\alpha = \alpha x + (1-\alpha)x' \in A$ (by convexity of A). Concavity of each u_i gives $u_i(x_\alpha) \geq \alpha u_i(x) + (1-\alpha)u_i(x') \geq \alpha z_i + (1-\alpha)z_i'$. So $\alpha\boldsymbol{z} + (1-\alpha)\boldsymbol{z}' \in \mathcal{U}$. □

The separating-hyperplane argument now proceeds in standard form. Fix a Pareto optimum x^* with utility profile $\boldsymbol{z}^* = (u_1(x^*), \dots, u_n(x^*))$. The set of utility profiles strictly dominating $\boldsymbol{z}^*$ is $\mathcal{V} = \{\boldsymbol{z} \in \mathbb{R}^n : \boldsymbol{z} > \boldsymbol{z}^*\}$, a non-empty open convex set. Pareto optimality of x^* implies $\mathcal{U} \cap \mathcal{V} = \emptyset$ (no alternative attains a utility profile strictly dominating $\boldsymbol{z}^*$). The separating hyperplane theorem (#21) then gives a linear function separating $\mathcal{U}$ and $\mathcal{V}$, and the gradient of that linear function is a vector of weights showing x^* to be the welfare-maximizer.

Theorem 9 (Welfare-function characterization)**.** *Suppose $A \subseteq \mathbb{R}^d$ is convex and each u_i is concave on A. If $x^* \in A$ is Pareto optimal, then there exist weights $\alpha_1, \dots, \alpha_n \geq 0$ (not all zero) such that*

$$x^* \in \arg\max_{y \in A} \sum_{i=1}^{n} \alpha_i u_i(y).$$

Proof sketch. With $\mathcal{U}$, $\boldsymbol{z}^*$, and $\mathcal{V}$ as above, the separating hyperplane theorem (#21) gives a non-zero $\alpha \in \mathbb{R}^n$ and $c \in \mathbb{R}$ such that $\alpha \cdot \boldsymbol{z} \le c \le \alpha \cdot \boldsymbol{w}$ for all $\boldsymbol{z} \in \mathcal{U}$ and all $\boldsymbol{w} \in \mathcal{V}$. We claim $\alpha \cdot \boldsymbol{z}^* = c$ and $\alpha \in \mathbb{R}^n_+$, i.e. each $\alpha_i \ge 0$. For the first: $\boldsymbol{z}^* \in \mathcal{U}$, so $\alpha \cdot \boldsymbol{z}^* \le c$; if strict, take $\boldsymbol{w} = \boldsymbol{z}^* + \epsilon \mathbf{1} \in \mathcal{V}$ for small $\epsilon > 0$ and observe $\alpha \cdot \boldsymbol{w} = \alpha \cdot \boldsymbol{z}^* + \epsilon \sum_i \alpha_i$, which can be made less than c for small ϵ if $\alpha \cdot \boldsymbol{z}^* < c$, a contradiction. For the second: if $\alpha_i < 0$ for some i, pick $\boldsymbol{w} = \boldsymbol{z}^* + \epsilon \mathbf{1} + \beta \boldsymbol{e}_i$ for small $\epsilon > 0$ and large $\beta > 0$; then $\boldsymbol{w} \in \mathcal{V}$ but $\alpha \cdot \boldsymbol{w}$ can be made strictly less than $c = \alpha \cdot \boldsymbol{z}^*$, again a contradiction. With both claims established, $\alpha \cdot \boldsymbol{z} \le \alpha \cdot \boldsymbol{z}^*$ for all $\boldsymbol{z} \in \mathcal{U}$, hence $\sum_i \alpha_i u_i(x) \le \sum_i \alpha_i u_i(x^*)$ for all $x \in A$. □

Combining Theorem 9 (necessary direction) with Theorem 6 (sufficient direction) gives a clean characterization under appropriate concavity.

Corollary 10 (Full characterization under strict concavity)**.** *Suppose $A \subseteq \mathbb{R}^d$ is convex and each u_i is strictly concave on A. Then $x^* \in A$ is Pareto optimal if and only if there exist weights $\alpha_1, \dots, \alpha_n \ge 0$ (not all zero) such that $x^* \in \arg\max_{y \in A} \sum_i \alpha_i u_i(y)$.*

A subtle but important caveat: the necessary direction of Theorem 9 only gives non-negative weights, not strictly positive ones. Boundary Pareto optima — alternatives at extreme points of the contract curve, for instance — may require some $\alpha_i = 0$, corresponding to "ignoring" some individuals' utility. The standard counter-example illustrates the point.

Example 11 (Why the necessary direction cannot be strengthened to strictly positive weights)**.** Let $N = \{1, 2\}$, $A = [0, 1]$, $u_1(x) = -x^2$, $u_2(x) = -(1-x)^2$. The Pareto set is A itself (every point in $[0, 1]$ is Pareto optimal: any move toward 0 helps voter 1 and hurts voter 2, and any move toward 1 does the reverse). Consider the boundary point $x^* = 1$: it maximizes voter 2's utility but minimizes voter 1's. There is no strictly positive weight pair (α_1, α_2) with $\alpha_1 > 0$ such that $x = 1$ maximizes $\alpha_1 u_1(x) + \alpha_2 u_2(x) = -\alpha_1 x^2 - \alpha_2 (1-x)^2$ on $[0, 1]$ — the FOC gives $-2\alpha_1 x + 2\alpha_2(1-x) = 0$, hence the maximizer is $x = \alpha_2/(\alpha_1 + \alpha_2) < 1$ unless $\alpha_1 = 0$. Only with $\alpha_1 = 0$ does $x^* = 1$ become a maximizer of the welfare function. The boundary Pareto optimum genuinely needs the zero weight, and the corresponding "social welfare" literally ignores voter 1.

5 Welfare functions, utilitarianism, and aggregation

The welfare-function representation of the Pareto set, theoretically clean as it is, raises an immediate substantive question for political-economy modeling: what *are* the weights $\alpha_1, \dots, \alpha_n$, where do they come from, and what does it mean for a social-choice problem that any vector of non-negative weights gives rise to a different Pareto-optimal alternative? These questions are not answered by the optimization theory of §4; they are answered, in different ways, by the substantive sub-fields the welfare framework is designed to feed into.

The simplest case is *utilitarianism*, the special case of the welfare function with equal weights:

$$W_{\text{util}}(\boldsymbol{z}) = \sum_{i=1}^{n} z_i.$$

Utilitarianism is the canonical reference welfare function in welfare economics, deeply rooted in Bentham and refined by Mill, Sidgwick, and the twentieth-century welfare economists (Pigou, Bergson, Samuelson). Theorem 4 with $\alpha_i = 1$ for all i says that the utilitarian maximizer is always

Pareto optimal — a central pre-Pareto observation. The substantive critique of utilitarianism, going back at least to Rawls (1971), is that the equal-weights treatment of individual utilities silently presupposes interpersonal comparability of utility (utilities are commensurable in such a way that adding them is meaningful) and is invariant under "transferring" utility from one individual to another (a transfer that leaves the sum invariant has zero welfare effect, regardless of which individual gives or receives). The Rawlsian alternative replaces the sum with the minimum, $W_{\text{Rawls}}(\boldsymbol{z}) = \min_i z_i$ (the maximin or "leximin" criterion), which gives lexicographic priority to the worst-off individual.

The general framework of *Bergson–Samuelson welfare functions* (Bergson, 1938; Samuelson, 1947) allows arbitrary increasing functions of the individual utilities, $W(\boldsymbol{z}) = F(z_1, \ldots, z_n)$ for F strictly increasing in each argument, weakening the linear-aggregation structure. Theorem 9's necessary direction can be extended to this more general framework: under mild regularity conditions, every Pareto-optimal alternative is the maximizer of some Bergson–Samuelson welfare function, with the linear-aggregation case as a special instance.[1]

For the purposes of bargaining theory (the next handout), the welfare-function characterization plays a different role. A bargaining problem is a Pareto set together with players whose conflicting preferences are exactly the source of weight-disagreement: each player would like the welfare function to put a higher weight on her own utility. Bargaining solutions (Nash, Kalai–Smorodinsky) are particular ways of selecting a Pareto-optimal alternative from the bargaining problem, equivalent (under regularity) to particular choices of weights in the welfare-function representation.

Example 12 (Utilitarian budget allocation)**.** Continuing Example 3: the coalition allocates a fixed budget B across n districts with utilities $u_i(x) = \phi_i(x_i)$ for strictly increasing concave ϕ_i. The utilitarian welfare-maximizer solves $\max \sum_i \phi_i(x_i)$ subject to $\sum_i x_i = B$, $x_i \geq 0$. The FOC gives $\phi_i'(x_i) = \lambda$ for every i (Lagrangian condition with multiplier λ): the utilitarian allocation equates marginal utilities across districts. With identical ϕ_i, the utilitarian allocation gives every district B/n; with heterogeneous ϕ_i, the allocations are unequal in proportion to inverse marginal-utility. The political-economy reading is that utilitarianism as a normative criterion delivers equal allocations only under the strong assumption of identical (or symmetric) preferences across districts. With realistically heterogeneous preferences, the utilitarian benchmark is itself non-egalitarian.

6 Beyond concavity

Theorem 9 relies on convexity of the alternative space and concavity of the individual utilities. When these hypotheses fail, the characterization can fail too: there are Pareto-optimal alternatives that are not maximizers of *any* non-negative-weighted-sum welfare function. The structural reason is that the separating hyperplane argument requires both sets ($\mathcal{U}$ and $\mathcal{V}$) to be convex, and concavity of

[1] The substantive welfare-economics literature has repeatedly returned to the question of what welfare functions *should* be used when interpersonal utility comparisons are at stake. The Bergson–Samuelson framework finesses the question by treating the welfare function as a primitive and asking only what its mathematical properties imply for the Pareto-optimal alternatives it selects. Arrow (1951) reframes the question as one of preference aggregation, with the famous impossibility result that no aggregation rule satisfying minimal axioms (universal domain, Pareto, independence of irrelevant alternatives, non-dictatorship) yields a coherent social ordering. Sen (1970) pushes further, showing that even the Pareto criterion can conflict with minimal liberal rights (the Sen paradox of "the impossibility of a Paretian liberal"). The forthcoming social-choice cluster develops this line of inquiry; for the present handout, the operative observation is that Pareto optimality and welfare-function maximization are coextensive at the optimization-theoretic level, and the substantive content of welfare economics — which welfare function, which weights, why — is downstream of the structural characterization in Theorem 9.

utilities is what gives $\mathcal{U}$ its convex structure. Without concavity, the imputation set may have "dents" that contain Pareto-optimal alternatives in their interior — alternatives that no linear functional separates from the strictly-dominating cone.

The standard diagnostic is the following. Without concavity, there can be Pareto optima at which the utility-imputation set is "locally non-convex"; a non-linear welfare function can still pick them out (since non-linear level sets can wrap around dents), but no linear welfare function can. The corresponding Pareto optima are not utilitarian / weighted-sum maximizers under any choice of weights.

Example 13 (Pareto optimum that is not a weighted-sum maximizer)**.** Let $A = [0, 1]$, $u_1(x) = x$, $u_2(x) = 1 - x^2$ (strictly concave) until $x = 1/\sqrt{2}$, then redefine u_2 to wrap upward in a non-concave manner so that the imputation set develops a non-convex region. (The full diagnostic is left as an exercise; Duggan (2017) works through an explicit case.) The point of the example is structural: the welfare-function characterization is a feature of the convex case, not a universal one, and the convex case is the workhorse precisely because it is the one in which the optimization-theoretic and welfare-economic readings line up.

The substantive lessons. First, convexity-and-concavity assumptions are not bookkeeping — they are what makes the welfare-economics framework analytically tractable, and their failure is the entry point for more delicate analysis (matching theory, mechanism design, models with indivisibilities). Second, in political-economy modeling, the standard regularity conditions are usually defensible: spatial-loss preferences are concave on convex policy spaces; concave preferences over divisible budgets are the rule; even discrete-choice models often admit continuous welfare-function representations via probabilistic-choice structure. The pathologies are real but not pervasive, and where they do bite, the welfare-theoretic literature has developed a substantial body of generalizations.[2]

7 What's next

This handout is the first of a two-handout cluster on welfare and bargaining. The bargaining handout that follows uses the Pareto set developed here as its primitive: a bargaining problem is a Pareto set together with a disagreement point Pareto-dominated by at least one alternative, with players whose preferences over the Pareto set conflict. The Nash bargaining solution and the Kalai–Smorodinsky solution are characterized axiomatically and then connected back to the welfare-function representation of §4 (each solution corresponds to a particular way of selecting weights).

Two further strands extend the cluster.

[2]The more general characterization theory — which welfare functions characterize Pareto optima beyond the linear-aggregation case — is the topic of a substantial mathematical-welfare-economics literature. Negishi (1960) introduced the technique of using welfare-function maximization to characterize competitive equilibria (the "Negishi" or "planner's problem" approach), with Pareto optimality as the bridge: every Pareto optimum is a Negishi-welfare maximizer for some weights, and every competitive equilibrium is a Pareto optimum (the first welfare theorem) — so equilibria are welfare maximizers. Mas-Colell (1985) provides the canonical infinite-dimensional generalization, treating commodity spaces of arbitrary dimension. The non-concave case has substantive applications in matching theory and mechanism design: matching outcomes are Pareto optima on a non-convex alternative space (the discrete set of matchings) and require non-utilitarian welfare-theoretic tools (the Roth–Sotomayor stability framework, the Shapley–Shubik assignment-game framework). The bargaining handout that follows stays in the convex case — the standard Nash and Kalai–Smorodinsky framework presumes a convex bargaining set — which is sufficient for the political-economy applications under consideration.

Social choice theory (a forthcoming cluster) addresses the question of where the welfare weights come from when individuals' preferences themselves are the primitive. Arrow's impossibility theorem shows that no preference-aggregation rule satisfying minimal axioms produces a coherent social ordering; Sen's results sharpen the trade-offs between Pareto efficiency and individual rights; the Gibbard–Satterthwaite theorem extends to strategic manipulation. The welfare-function framework of this handout is, at the structural level, a shortcut around Arrow's problem: by stipulating a welfare function, the modeler bypasses the aggregation question. Social-choice theory asks whether the bypass is justified.

Mechanism design (forthcoming game-theory cluster, eventually) addresses the practical problem of selecting Pareto optima when players have private information and strategic incentives. The revelation principle, the Vickrey–Clarke–Groves mechanism, and the broader literature on optimal auctions and contract design build on the welfare-function characterization of this handout: a designer asks which Pareto optimum she would like to implement, and the question becomes whether incentive-compatibility and individual-rationality constraints allow her to implement it.

For graduate-level treatments at this handout's level: Mas-Colell, Whinston, and Green (1995, Ch. 16) is the standard micro-theoretic reference on welfare economics with the Pareto characterization at center; Austen-Smith and Banks (1999, 2005) treat the political-economy applications with the political-science audience in mind; Duggan (2017) (the source of the structural sequence in this handout) treats the optimization-theoretic side in detail with explicit attention to the convex / non-convex distinction.

8 Exercises

Exercise 14. *Two-voter contract curve.* Two voters in $\mathbb{R}$ have ideal points $\hat{x}_1 = 0, \hat{x}_2 = 1$ and quadratic-loss utilities $u_i(x) = -(x - \hat{x}_i)^2$. (a) Show that the Pareto set is $[0, 1]$. (b) Find, for each $\alpha \in (0, 1)$, the unique $x^*(\alpha) \in [0, 1]$ that maximizes $\alpha u_1(x) + (1 - \alpha)u_2(x)$. (c) Verify that as α ranges over $(0, 1)$, $x^*(\alpha)$ traces out the interior of the Pareto set, and identify which weights pick out the boundary points $x = 0$ and $x = 1$.

Exercise 15. *Budget allocation.* Two legislators allocate $B = 10$ across two districts, with utilities $u_i(x_1, x_2) = \log(x_i + 1)$ (each cares only about her own district). (a) Show that the Pareto set is the budget line $\{(x_1, x_2) : x_1 + x_2 = 10, x_1, x_2 \geq 0\}$. (b) Find the utilitarian-welfare maximizer (equal weights). (c) Find the welfare-maximizer with weights $(\alpha_1, \alpha_2) = (1, 2)$. (d) Discuss in two sentences how the heterogeneity of weights translates into the asymmetry of the resulting allocation.

Exercise 16. *Existence under compactness.* Let $A = [0, 1]^2$ and $u_1(x_1, x_2) = x_1$, $u_2(x_1, x_2) = x_2$. (a) Verify the Weierstrass conditions and conclude that a Pareto optimum exists. (b) Describe the Pareto set explicitly. (c) Identify (without solving) the utilitarian welfare-maximizer.

Exercise 17. *Convexity of the imputation set.* Let $A = \mathbb{R}$ and $u_1(x) = -(x-1)^2$, $u_2(x) = -(x+1)^2$. (a) Compute the set of utility imputations $\mathcal{U} \subseteq \mathbb{R}^2$. (b) Verify that $\mathcal{U}$ is convex (consistent with Lemma 8). (c) Find the Pareto set in A and the corresponding boundary of $\mathcal{U}$.

Exercise 18. *Separating hyperplane in action.* For the setting of Exercise 17, fix the Pareto optimum $x^* = 0$ with utility profile $\boldsymbol{z}^* = (-1, -1)$. (a) Identify the strictly-dominating set $\mathcal{V} \subseteq \mathbb{R}^2$. (b) Find a non-zero vector of non-negative weights $\alpha = (\alpha_1, \alpha_2)$ such that $\alpha \cdot \boldsymbol{z} \leq \alpha \cdot \boldsymbol{z}^*$ for every $\boldsymbol{z} \in \mathcal{U}$ and $\alpha \cdot \boldsymbol{w} \geq \alpha \cdot \boldsymbol{z}^*$ for every $\boldsymbol{w} \in \mathcal{V}$. (c) Verify that x^* is the maximizer of $\alpha_1 u_1(x) + \alpha_2 u_2(x)$ on $\mathbb{R}$.

Exercise 19. *Rawlsian (maximin) welfare.* Consider Example 12 with $n = 2$, $\phi_1(y) = \log(y+1)$, $\phi_2(y) = \log(y+1)$ (identical preferences), $B = 10$. (a) Find the utilitarian welfare-maximizer. (b) Find the Rawlsian (maximin) welfare-maximizer, $\max_{x \in A} \min_i \phi_i(x_i)$. (c) Verify that both are Pareto optimal but typically not the same alternative; identify when they coincide. (d) Discuss in two sentences which welfare criterion the political-economy literature on income inequality more closely resembles.

Exercise 20. *Bergson–Samuelson welfare.* Let $W : \mathbb{R}^n \to \mathbb{R}$ be a strictly increasing continuous Bergson–Samuelson welfare function (i.e., $W(\boldsymbol{z}) > W(\boldsymbol{z}')$ whenever $\boldsymbol{z} > \boldsymbol{z}'$ componentwise). Show that any maximizer of $W \circ (u_1, \ldots, u_n)$ on A is Pareto optimal, mirroring Theorem 4 but for non-linear W. The exercise illustrates that the welfare-function framework is more general than the linear-utilitarian special case.

Exercise 21. *Reform vs. status quo as Pareto-improvement.* Three commissioners must accept or reject a reform proposal r relative to a status quo q. The commissioners have utilities u_i over the alternative $\{r, q\}$. (a) Define what it means for r to Pareto-dominate q in this setting. (b) Argue that any reform that strictly Pareto-improves on the status quo is implementable by unanimity. (c) Discuss in two or three sentences why most actually-proposed reforms in legislative settings do not Pareto-dominate the status quo, and what that implies for the structure of legislative bargaining.

Exercise 22. *Coalition vs. unanimity.* Three voters $\{1, 2, 3\}$ with utilities u_1, u_2, u_3 on $A = \mathbb{R}$ have ideal points $\hat{x}_1 = 0$, $\hat{x}_2 = 1$, $\hat{x}_3 = 3$ and quadratic-loss utilities. (a) Identify the Pareto set. (b) Show that any majority coalition (any two of the three voters) can find a movement away from voter 3's ideal point that the coalition prefers. (c) Discuss in two sentences how this exercise illustrates the gap between Pareto efficiency (a unanimity-based criterion) and majority-based political bargaining (which can move the outcome away from a Pareto optimum, on a strict reading).

Exercise 23. *What Pareto optimality buys, conceptually.* In two or three sentences, explain to a political-science colleague why Pareto optimality is the most uncontroversial welfare criterion — and why, despite this, most substantive social-choice and welfare-economics analysis goes *beyond* Pareto optimality (to weighted-sum welfare functions, Bergson–Samuelson welfare, social-choice axioms, etc.). The exercise asks the reader to articulate the conceptual position of Pareto optimality in the welfare-economics framework: its strength (uncontroversiality) is also its weakness (non-uniqueness in selecting an alternative).

References

Arrow, Kenneth J. (1951). *Social Choice and Individual Values.* New York: Wiley.

Austen-Smith, David and Jeffrey S. Banks (1999). *Positive Political Theory I: Collective Preference.* Ann Arbor: University of Michigan Press.

– (2005). *Positive Political Theory II: Strategy and Structure.* Ann Arbor: University of Michigan Press.

Bergson, Abram (1938). “A Reformulation of Certain Aspects of Welfare Economics”. In: *Quarterly Journal of Economics* 52.2, pp. 310–334.

Duggan, John (2017). “Notes on Optimization and Pareto Optimality”. Lecture notes, University of Rochester.

Mas-Colell, Andreu (1985). *The Theory of General Economic Equilibrium: A Differentiable Approach.* Cambridge: Cambridge University Press.

Mas-Colell, Andreu, Michael D. Whinston, and Jerry R. Green (1995). *Microeconomic Theory*. New York: Oxford University Press.

Negishi, Takashi (1960). "Welfare Economics and Existence of an Equilibrium for a Competitive Economy". In: *Metroeconomica* 12.2–3, pp. 92–97.

Rawls, John (1971). *A Theory of Justice*. Cambridge, MA: Harvard University Press.

Samuelson, Paul A. (1947). *Foundations of Economic Analysis*. Cambridge, MA: Harvard University Press.

Sen, Amartya K. (1970). *Collective Choice and Social Welfare*. San Francisco: Holden-Day.