Consensus and Interpersonal Welfare Comparisons*

Niels Boissonnet†
Department of Economics, IMW, and JICE, Bielefeld University

Yves Breitmoser
Department of Economics, IMW, and JICE, Bielefeld University

What welfare comparisons follow when people with different preferences agree about physical resources? We study a rule restricted to bundles their recipients could receive in efficient divisions of a specified resource pool. If both people regard one bundle as at least as good as the other, its recipient must count as no worse off, and as better off if either preference is strict. We characterize all ways of assigning comparable welfare levels that preserve individual preferences and satisfy this rule. The two people's indifference judgments give exact bounds on which bundles can represent equal welfare across them. Every continuous, order-preserving pairing of reference bundles within those bounds is possible. Although some comparisons remain open, the shared comparisons can identify a best facility assignment under symmetric, welfare-increasing social evaluation, without comparing welfare gains. Equal welfare at an efficient division where the bundles can be exchanged also excludes envy.

Key words: Interpersonal welfare comparisons; ordinal preferences; consensus; social choice
JEL codes: D63; D71

1. INTRODUCTION

What can agreement between people with different preferences tell us about their relative welfare? We study a simple rule: if both people regard one physical bundle as at least as good as another, its recipient should count as no worse off. The welfare ranking must be strict when either person strictly prefers that bundle. Applied to every bundle, including agreement in indifference, this rule is inconsistent with different preferences. We restrict it to bundles the people could receive in efficient divisions of a specified pool of resources. This restriction permits a sharp answer. We characterize all ways of assigning comparable welfare levels that preserve each person's preferences and satisfy the rule. These assignments need not agree on who is better off in every case. Yet the comparisons they share can settle some distributive choices without measuring welfare gains.

Consider a regional authority deciding whether a rural town or a city should receive a specialist-care facility. Both want it, so neither placement makes everyone better off. Yet a decision might follow from two comparisons of welfare levels: the city with the facility is better off than the rural town with it, while the rural town without the facility is better off than the city without it. Locating the facility in the city then produces a better treated position and a better untreated position. An authority that treats the towns symmetrically and favors higher welfare at each position prefers this placement. The argument does not compare the city's gain with the rural town's loss. Its missing ingredient is a basis for comparing welfare across recipients.

The distinction is between knowing who gains and knowing what social judgment follows. Even a complete description of each town's preferences cannot say whether a treated city occupies a higher welfare level than a treated rural town. Multiplying one town's utility numbers by a positive constant

* We are grateful for comments from Gerrit Bauch, Jurek Preker, Frank Riedel, and participants at the Economic Theory Lunch Seminar in Bielefeld. Financial support by Verbundinitiative InChange funded by MKW NRW is gratefully acknowledged.

† Corresponding author.

changes none of its choices but can change numerical comparisons with the other town. Simply adding the reported utility numbers would therefore answer a question that the reports have not settled. A welfare principle must specify which cross-person comparisons are meaningful.

One response is to select a complete interpersonal index and a social objective. We ask a prior question: how far can a limited principle take the authority without making those additional choices? An incomplete answer is useful when the unresolved comparisons do not affect the decision. It is also useful when it establishes that the decision *does* require another judgment, rather than a more precise estimate of an individual utility function. Our results identify both possibilities.

This distinction matters for public appraisal. The UK Green Book calls for analysis of distributional effects and describes welfare weights for monetary benefits received by different income groups (HM Treasury, 2026). Such appraisal makes the ethical assumptions behind distributional comparisons explicit. Our question is complementary: which decisions follow before selecting one complete welfare measure or one way of adding up gains? The hospital example also has a concrete policy counterpart. Germany's 2024 hospital reform sought to combine quality requirements for specialized treatment with geographically accessible basic care (Bundesministerium für Gesundheit, 2024). We do not evaluate that reform. It illustrates a conflict in which communities may value the relevant services differently, and establishing the physical consequences does not by itself settle their social evaluation.

Our proposed welfare information comes from a hypothetical sharing problem. The authority specifies the resources the two towns may divide. A division is efficient if neither town can gain without harming the other. Each town's reference set consists of the bundles it receives in some efficient division. Take one bundle from each set and ask both towns which of these same physical alternatives they prefer for themselves. *Efficient Consensus* requires the corresponding ranking of their recipients' welfare. If both weakly prefer one bundle and either preference is strict, the recipient must count as strictly better off. If both are indifferent, the welfare levels must be equal. Individual preferences supply the physical comparisons; the authority adopts the ethical rule connecting them to interpersonal welfare. The compared bundles need not belong to the same joint division.

For example, one town may regard a bundle with more local general care and another with more specialist service as equally good, while the other town strictly prefers the specialist bundle. The rule treats this as agreement sufficient for a strict welfare comparison: neither person disputes the specialist bundle's advantage, and one strictly endorses it. If each instead prefers a different bundle, the rule need not decide. Consensus here is a restriction on a social evaluator, not unanimous approval of the actual policy. The town losing a facility can still oppose a policy that the maintained welfare principles rank higher.

Efficient sharing makes the restriction economically relevant. Its reference bundles reflect both scarcity and the people's different relative valuations of the resources. It also gives consensus a familiar fairness implication. Suppose one person envies the other at an efficient division, and exchanging their bundles is feasible. The envied person must strictly prefer keeping her bundle; otherwise the exchange would improve the division. Both people therefore prefer that bundle, so the rule places its recipient strictly higher. An efficient, exchangeable division that equalizes an admissible welfare scale must consequently be envy-free. The rule does not choose an allocation. It retains this connection between envy and welfare precisely on a domain where mutual envy cannot produce contradictory comparisons.

The central theorem describes exactly how much room the rule leaves. In our environments, each person's efficient reference set is a continuous sequence of increasingly preferred bundles, from none of the resources to the full pool. Fix a bundle on the first sequence. Each person identifies a bundle on the second that she regards as equally good. These are two individual answers to the same physical question, not two measurements of interpersonal utility. If the answers differ, a bundle representing equal welfare across the recipients must lie strictly between them; if they coincide, that answer is required. Every continuous, increasing pairing satisfying these bounds is admissible.

The characterization therefore identifies a class of welfare scales rather than selecting one index. These scales are ordinal: their levels can be compared across people, but differences between the numbers do not measure comparable gains. Some comparisons hold throughout the class; others can be reversed without violating the rule.

The next question concerns the references themselves. One can compare people by giving the same welfare label to the same bundle along a chosen reference sequence and extending the labels through individual indifference. Different sequences may produce different rankings. We show when the two efficient sequences preserve all the weak comparisons that survive every such choice. The key is to find, on each person's indifference curve, the bundles the other person values most and least. Under two conditions ordering relative valuations, the efficient references contain exactly these extremes. One sequence cannot do the same job, and no other pair does. Thus the allocation problem can determine the smallest sufficient reference family. This conclusion requires stronger conditions than the welfare-scale characterization and holds within a fixed sharing institution, not across arbitrary changes in its resource pool.

Incomplete welfare information can still be decisive. Suppose three towns rank 3, 2, 1 from highest to lowest welfare when treated, but 1, 3, 2 when untreated. With one facility, treating town 3 obtains the best treated position but leaves town 2 at the worst untreated position. Treating town 2 avoids that loss at the cost of a worse treated position. Neither placement is best for every symmetric, welfare-increasing evaluation. With two facilities, treating towns 2 and 3 gives the two highest treated positions and the highest untreated position. Every such evaluation prefers that assignment. We construct preferences and efficient references yielding these rankings. The general result compares the ordered lists of treated and untreated welfare levels. It needs no cardinal treatment gains and does not require the stronger conditions used for reference minimality.

These results contribute to a common research program in fair allocation and welfare measurement. No-envy asks whether anyone prefers another person's bundle; egalitarian equivalence asks whether everyone regards her allocation as equivalent to one common physical bundle (Varian, 1974, 1976; Pazner and Schmeidler, 1978; Thomson, 2011). These concepts show how preferences and resources can give equality an economic meaning. Fairness-based welfare measurement develops that idea into interpersonal indices and policy evaluation (Fleurbaey and Maniquet, 2011; Fleurbaey and Schokkaert, 2011; Fleurbaey and Tadenuma, 2014). In particular, Fleurbaey and Maniquet (2017) characterize families of well-being measures from fairness requirements. We build on this foundation by using efficient bilateral sharing to determine the reference sets and characterizing the entire class of scales allowed by consensus on them. Our envy result explains the fairness content of that class; it is not a new general existence theorem for envy-free allocations.

This distinction also separates two uses of fairness. An allocation criterion such as no-envy asks whether a proposed distribution is acceptable. A welfare measure asks how to compare recipients at many distributions, including ones that fail that criterion. Moving from the first question to the second requires consistency across comparisons. Our restriction preserves an envy-based interpretation on efficient bilateral divisions while allowing the resulting scales to evaluate other bundles through each recipient's preferences. This is the connection needed when the actual policy menu contains no envy-free assignment.

The choice of reference affects welfare comparisons, and earlier work gives economic or axiomatic reasons for several choices (Fleurbaey, 2007; Fleurbaey et al., 2003). Maniquet and Moramarco (2026) study welfare measurement relative to reference consumption, requiring that a change to an individually equivalent reference bundle leave measured welfare unchanged. Their approach makes explicit how preferences at the reference enter the measurement problem. We ask a different, complementary question: which reference bundles does efficient bilateral sharing supply, and what welfare comparisons follow from consensus on those bundles?

Decancq and Nys (2021) develop robust incomplete comparisons using indifference contours, study restricted sets of physical alternatives, and provide an ordinal elicitation procedure. Their

analysis explains both the value of reference robustness and the greater decisiveness available from a narrower comparison domain. Our efficient-sharing construction gives that domain a particular economic source. The minimal-reference theorem establishes when the efficient pair preserves exactly the weak comparisons robust to all shared references. Outside its stronger assumptions, efficient references can instead yield additional comparisons; their extra force comes from accepting the restriction, not from new information about preferences.

These contributions lead to two distinct tasks: determining meaningful reference bundles and determining how much welfare information they provide. Efficient sharing addresses the first; the sharp scale characterization addresses the second. We do not claim that the resulting institution is uniquely correct. By retaining every compatible scale, we distinguish what its consensus rule implies from what would follow only after selecting one index.

Consensus and aggregation address different parts of the question. Sprumont (2012) and Sakamoto (2018) use agreement to restrict social evaluation; our rule applies to interpersonal positions associated with efficient shares. Pivato (2013) studies what social conclusions follow from incomplete interpersonal comparisons. We derive such comparisons from preferences and EC before asking what they imply for policies. The policy theorem uses the established anonymous grading and rank-dominance principles (Suppes, 1966; Sen, 1970; Saposnik, 1983): pair every position under one policy with a distinct, no-higher position under its rival. We prove that one favorable pairing is necessary as well as sufficient for agreement across all compatible scales and all anonymous, strongly Pareto evaluations. Consistency and matching results supply the mathematical tools for this step, not a competing answer to the welfare-measurement question.

The conclusion is conditional on explicit choices. The authority must specify the resource pool, accept EC, and, for our policy comparisons, treat recipient names symmetrically. The healthcare example treats towns as welfare subjects; equal treatment of towns is not the same as equal weighting of residents in differently sized towns. Travel, treatment quality and congestion must enter the physical consequences or preferences before evaluation (Mseke et al., 2024; Ramsay et al., 2025). The rule neither identifies these effects nor induces truthful reports. Its contribution is to distinguish decisions supported by the stated welfare information from decisions requiring another ethical judgment. More precise preference answers can resolve uncertainty about the former; they cannot eliminate a disagreement that remains across admissible scales with fully known preferences.

2. CONSENSUS AND EFFICIENT RESOURCE SHARING

The facility conflict separates three tasks: describing what each person receives, comparing welfare across recipients, and evaluating a whole allocation. Each person's preferences rank her own possible bundles. They do not by themselves compare welfare across recipients or select a social objective. We begin with the additional interpersonal structure, explain why an unrestricted consensus rule fails, and construct the restricted domain on which the rule can work.

2.1. Preferences and welfare comparisons

There are two people, $N = \{1, 2\}$, and $L \geq 2$ resources. A personal bundle is $x \in X = [0, 1]^L$, with 0 and 1 denoting the empty and full bundles. Person i's preference $\succsim_i$ is continuous and monotone: $x \geq y$ implies $x \succsim_i y$, and $x \gg y$ implies $x \succ_i y$. The declared sharing institution has feasible set

$$\mathcal{X} \subseteq \{(x_1, x_2) \in X^2 : x_1 + x_2 \leq 1\}$$

and contains $\mathbf{e}_1(x) = (x, 0)$ and $\mathbf{e}_2(x) = (0, x)$ for every $x \in X$. These single-recipient allocations let the authority compare giving a bundle to one person with giving another bundle to the other. The institution may allow hypothetical division of services even when an actual facility is indivisible.

The resource pool describes the institution used for comparison, not necessarily the actual policy menu. A regional authority may evaluate indivisible facilities using hypothetical divisions of service

capacity. It may also protect resources assigned to outsiders. Those choices determine which divisions are available in X. Keeping them explicit prevents a change in entitlements or feasibility from being mistaken for a change in numerical units.

A single-recipient allocation is useful because it isolates the comparison to be made. When we compare $\mathbf{e}_1(x)$ with $\mathbf{e}_2(y)$, the untreated recipient has the common empty bundle in each case. The authority is comparing the two positive welfare positions, not the desirability of a policy that gives both x and y simultaneously. This distinction will matter when the two reference bundles come from different efficient divisions.

An *interpersonal welfare scale* $\mathbf{m} = (m_1, m_2)$ consists of continuous representations $m_i : X \to \mathbb{R}_+$ of the two preferences. It assigns meaning to comparisons such as $m_1(x) \geq m_2(y)$: person 1 at x counts as at least as well off as person 2 at y. Independently increasing transformations preserve individual choices but can reverse this comparison. We impose the common anchor $m_i(0) = 0$. Equal welfare at the empty bundle is an interpersonal premise; using the number zero is only normalization. We do not yet impose a common label at the full bundle. Nor do differences in labels measure comparable welfare gains.

To connect these levels to social evaluation, first compare two allocations that give all resources to a single recipient. If their recipients have the same welfare level and everyone else has level zero, changing the recipient's name should leave the social evaluation unchanged. Giving that recipient a higher welfare level should improve the allocation. Formally, let R be a complete, transitive ordering of welfare profiles in $\mathbb{R}^2_+$. We require *strong Pareto*: a coordinatewise increase with at least one strict increase improves the profile strictly. The name-independence requirement is *single-recipient anonymity*,

$$(a, 0) \sim_{\mathrm{R}} (0, a) \qquad (a \geq 0).$$

Call such an ordering an *aggregation rule*. Together, $\mathbf{m}$ and R represent the social preference

$$\mathbf{x} \succsim \mathbf{y} \quad \Longleftrightarrow \quad (m_1(x_1), m_2(x_2))\, \mathrm{R}\, (m_1(y_1), m_2(y_2)).$$

For single-recipient allocations these assumptions make the social ranking exactly the ranking of the recipients' welfare labels: first move the positive label to the same coordinate by anonymity, then apply strong Pareto. Thus

$$\mathbf{e}_1(x) \succsim \mathbf{e}_2(y) \quad \Longleftrightarrow \quad m_1(x) \geq m_2(y).$$

Online Appendix Lemma O.1 records the same equivalence for strict preference and indifference.

The stronger *anonymity* assumption used for policy comparisons makes every permutation of a welfare profile indifferent. It is not part of the consensus rule. An authority may use the same interpersonal scale but give a named person special priority when evaluating policies; our anonymous policy conclusions would then not apply. Separating measurement from aggregation lets us ask which interpersonal comparisons follow before choosing a complete social objective.

For instance, an authority might give priority to the lowest welfare level, breaking ties by the remaining level, or instead place more weight on improvements at higher levels, subject to strong Pareto. The characterization below does not decide among such distributive judgments. Its question is what numbers may consistently represent the individual *positions* to which an aggregation rule is applied. Single-recipient anonymity is sufficient to interpret the consensus comparisons; full anonymity becomes relevant only when the positions must be rearranged to compare whole policies.

Even on a fixed scale, full anonymity is not the claim that equal physical resources give equal welfare. People can value those resources differently. It says that once interpersonal welfare levels have been specified, changing the recipients' names without changing the list of levels leaves social evaluation unchanged. This distinction is essential in an application with towns of different populations or recipients with different rights. Such features may require different welfare subjects or additional constraints rather than an anonymous comparison of the town-level labels.

2.2. Why the unrestricted rule fails

A natural first proposal applies consensus to any two bundles. If both people weakly prefer x to y, it requires $\mathbf{e}_1(x) \succsim \mathbf{e}_2(y)$, strictly if either preference is strict, and the same requirement after exchanging recipients. Call this *Unrestricted Consensus* (UC).

The problem already appears at identical bundles. Both people are indifferent between x and itself. UC therefore requires a tie between giving x to person 1 and giving it to person 2. Their welfare labels must agree at every x. One utility representation would then have to represent both preferences, excluding the heterogeneity we wish to study.

Proposition 1 (Unrestricted consensus is impossible)**.** *If $\succsim_1 \neq \succsim_2$, no social preference represented by an interpersonal welfare scale and an aggregation rule satisfies* UC.

Proof of Proposition 1. Applying weak UC in both recipient directions to an identical physical bundle x gives $\mathbf{e}_1(x) \sim \mathbf{e}_2(x)$. Lemma O.1 therefore requires $m_1(x) = m_2(x)$ for every x. As both indices represent their respective preferences, $x \succsim_1 y$ if and only if $m_1(x) \geq m_1(y)$, if and only if $m_2(x) \geq m_2(y)$, if and only if $x \succsim_2 y$. This contradicts heterogeneous preferences; neither strict UC nor differentiability is needed. □

The argument uses weak agreement at identical bundles, not a failure of strict agreement. Other consensus principles apply to whole allocations or to unanimous strict comparisons (Sprumont, 2012; Sakamoto, 2018; Brun and Tungodden, 2004). The proposition does not rule out those approaches. It shows why our rule needs an economically specified comparison domain.

The failure is informative about what must be restricted. It does not arise because the people disagree about every bundle: heterogeneous preferences can agree on many comparisons. Rather, imposing equal welfare whenever the same physical bundle is given to either recipient, *at every bundle*, forces one common ordering of the entire resource space. We therefore cannot preserve that unrestricted equality requirement and at the same time represent two different personal orderings.

Restricting the domain also needs a positive rationale. A domain chosen solely to avoid contradictions could exclude precisely the comparisons that matter economically. Efficient sharing instead asks which bundles each person can receive when there are no further opportunities for mutual improvement. It relates the comparison domain to the scarcity problem and to the same preferences the welfare scale must respect. The next construction states the domain before asking what consensus implies on it.

2.3. Reference bundles from efficient sharing

A division is *Pareto efficient* if no feasible division makes one person better off without making the other worse off. Let E be the set of such divisions and $E_i = \{x_i : (x_1, x_2) \in E\}$ its projection for person i. We propose to compare $x \in E_1$ with $y \in E_2$. Each bundle must be obtainable by its recipient in an efficient division; the two bundles need not be jointly feasible or come from the same division.

The two-resource example gives these sets a concrete meaning. Let the rural town value general care s relatively more than specialist services z, while the city has the opposite comparative advantage:

$$u_1(s, z) = 2s + z, \qquad u_2(s, z) = s + 2z.$$

Allow the towns to divide one unit of each resource freely. If the rural town has some specialist service while the city still has general care, a small exchange of these resources can benefit both. Every efficient division therefore allocates general care to the rural town first and specialist services to the city first. As the welfare distribution varies, the rural town's efficient bundles follow the lower and right edges of the unit square; the city's follow the left and upper edges. Figure 1 displays these two reference paths.

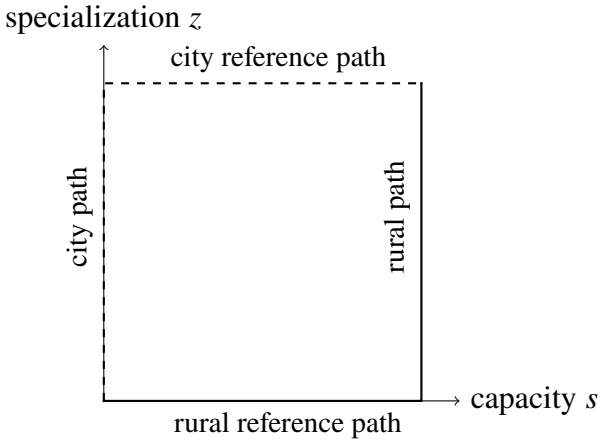


Figure 1. The two reference paths generated by efficient divisions in the health example. The solid path is the rural town's projection of the efficient frontier, whereas the dashed path is the city's projection. Both towns rank each path from 0 to 1.

These are ordered reference sets, not a proposed division. Moving forward on either path improves the physical bundle for both towns. Each person's indifference curve therefore meets either path once. In the general model we require the same continuous progression from empty to full resources, with no repetition or reversal. A *homeomorphism* is the continuous one-to-one parametrization, with continuous inverse, used to express that requirement. The paths need not be right-angled or generated by linear preferences.

Assumption 1 (Regular efficient reference paths)**.** *For each $j \in \{1,2\}$, the projected Pareto set E_j is a continuous path from* 0 *to* 1*, and both agents rank points on that path in the same strict order. Formally, there is a homeomorphism $p_j : [0,1] \to E_j$ with $p_j(0) = 0$, $p_j(1) = 1$, and, for both $i = 1,2$,*

$$t > t' \quad \Longleftrightarrow \quad p_j(t) \succ_i p_j(t').$$

The assumption requires both people to rank each path in the same direction, not a unique joint efficient allocation. Online Appendix O.5 gives sufficient conditions, including comparative advantage and a concave supermodular class. Regularity is a maintained restriction on the sharing problem, not a conclusion from arbitrary preferences.

Efficient divisions are plausible references because they reflect the resources that can be shared and the opportunities for mutually beneficial exchange. Excluding dominated divisions also avoids using a benchmark that wastes such an opportunity. This is a rationale for the domain, not a proof that it is ethically compulsory. The resource pool, protected outsiders' entitlements and permission for hypothetical fractional sharing remain institutional choices. Section 4 asks a separate question: when do these efficient paths preserve all comparisons robust to other shared paths?

Efficient references need not be equal-resource bundles. In the linear example, giving general care to the rural town and specialist services to the city exploits their different relative valuations. Equal division could leave a mutually beneficial exchange available. A reference system tied to efficient sharing therefore responds to heterogeneous preferences rather than suppressing them. The authority still decides how the resource pool is defined; efficiency determines which of its divisions enter the comparison.

The regular-path assumption makes this choice operational. Each path orders a full range of welfare positions from empty to full resources, and both people can compare progress along it. It does not require the people to agree on which of two bundles lying on *different* paths is better. Those cross-path disagreements are precisely what can leave room for different interpersonal welfare rankings. If the preferences or feasible set do not generate regular paths, the present representation cannot simply be applied without checking a new domain.

2.4. The restricted consensus rule

For a bundle on each person's reference path, the authority now asks both people to compare the same two physical alternatives for themselves. Agreement has an explicit interpersonal consequence:

Efficient Consensus (EC). *For $x \in E_1$ and $y \in E_2$, if both agents weakly prefer x to y, then $\mathbf{e}_1(x) \succsim \mathbf{e}_2(y)$, with strict social preference if one agent's preference is strict. If both agents weakly prefer y to x, then $\mathbf{e}_2(y) \succsim \mathbf{e}_1(x)$, again strictly when one individual comparison is strict.*

The reference restriction avoids UC's forced equality at every common physical bundle. Where the two paths do meet, EC still requires equal labels for the same bundle. In particular, both contain 1, so $m_1(1) = m_2(1) > 0$. A common positive rescaling then normalizes both upper labels to one. We use this normalization below. It does not give a cardinal interpretation to differences between labels.

EC can be attractive without settling every ethical question. It refuses to declare two reference recipients equally well off when both people agree that one physical bundle is better, but otherwise allows disagreement to leave room for the authority's judgment. The next section determines exactly how much room remains.

3. WHAT CONSENSUS ESTABLISHES

The first task is to find the entire set of welfare scales that obey EC. A necessary bound would not be enough: unless every permitted comparison can be implemented, we would not know which apparent ambiguities are genuine. The two people's own indifference judgments give both the bound and the construction.

3.1. Two personal answers bound an interpersonal tie

Return to $u_1(s, z) = 2s + z$ and $u_2(s, z) = s + 2z$. Fix the rural-path bundle $x = (1/2, 0)$. How much specialist service alone is equally good? The city answers $(0, 1/4)$, whereas the rural town answers $(0, 1)$. Below the city's answer, both prefer x; above the rural answer, both prefer the city-path bundle. At either answer, one person is indifferent and the other has a strict preference. EC therefore requires any interpersonal tie to lie strictly between the two answers. These quantities are physical equivalents, not interpersonal utility differences.

At the lower answer $(0, 1/4)$, the city is indifferent to the rural bundle, but the rural town strictly prefers its own bundle. EC makes the rural welfare position strictly higher, so this endpoint cannot represent an interpersonal tie. At the upper answer $(0, 1)$, the rural town is indifferent and the city strictly prefers its own bundle. Now the city welfare position must be strictly higher. Between the answers the rural town prefers the rural bundle while the city prefers the city bundle. Consensus itself supplies neither strict direction there.

This example distinguishes the information reported from the judgment constrained. Each person reports an ordinary indifference between two possible bundles for herself. Neither is asked whether she has the same utility as someone else, how much welfare the other receives, or what weight society should attach to her preferences. The authority uses the two reports to constrain a cross-recipient tie because it has accepted EC. Without that ethical step, the individual reports remain only intrapersonal comparisons.

For general preferences, let $\tau_2(x) \in E_2$ be the bundle person 2 regards as equally good as $x \in E_1$. Let $\tau_1(y) \in E_1$ be person 1's equivalent to $y \in E_2$. Both maps are increasing homeomorphisms. Starting from x, the two answers on E_2 are therefore $\tau_2(x)$ and $\tau_1^{-1}(x)$. The inverse is needed because τ_1 runs in the other direction.

An interpersonal scale pairs each $x \in E_1$ with the unique $y \in E_2$ having the same welfare label. Write $\mu : E_1 \to E_2$ for this *social-equivalence map*, so that

$$m_1(x) = m_2(\mu(x)).$$

Continuity, common endpoints and strict ordering along the paths make μ an increasing homeomorphism. EC restricts it to the following band:

$$\mathcal{B}(x) = \begin{cases} \{\tau_2(x)\}, & \tau_2(x) = \tau_1^{-1}(x), \\ \{y \in E_2 : y \text{ lies strictly between } \tau_2(x) \text{ and } \tau_1^{-1}(x)\}, & \tau_2(x) \neq \tau_1^{-1}(x). \end{cases}$$

Coincident personal answers fix the social tie; otherwise the band is open.

A permissible choice must also fit with choices at other welfare levels. If x' is better than x on E_1, their socially equivalent bundles on E_2 must occur in the same order. Choosing a point independently in every band could violate that consistency requirement. The characterization therefore concerns *increasing continuous pairings of the entire paths*, not arbitrary selections from separate intervals. The theorem's sharpness claim is that no further restriction is needed once this order requirement and the bands are respected.

In particular, the authority need not declare which of the two people's personal equivalents is the morally correct one. Except where they coincide, either endpoint would contradict the other person's strict comparison. Nor does EC require the arithmetic midpoint between their numerical labels: such a midpoint changes if the path is relabeled. The economically meaningful requirement is the open interval in the physical path order together with an increasing pairing across welfare levels. Figure 2 shows an admissible tie between the two personal equivalents.

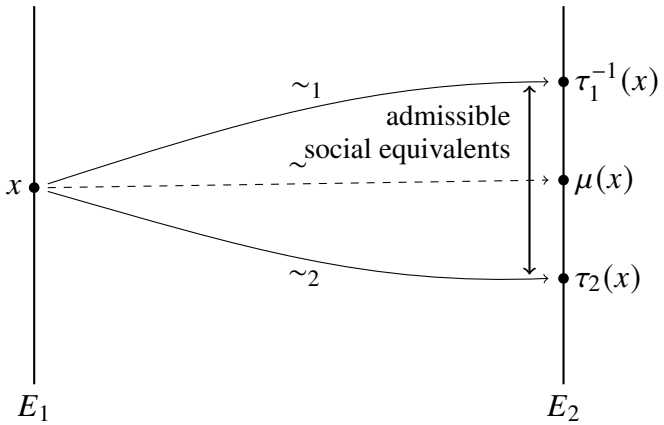


Figure 2. Two personal equivalents and the EC interval. The physical reference paths in Figure 1 are drawn vertically, with more preferred bundles higher. Agent 2's equivalent to x is below agent 1's. An EC social equivalent must lie strictly between them; neither endpoint can be a social tie.

There is another useful description of the same information. Choose one path E_j and give the same welfare label to the same bundle on that path for either person. Extend these labels to every bundle through the recipient's own indifference. A scale $\mathbf{m}^j = (m_1^j, m_2^j)$ constructed in this way is *calibrated to* E_j. The convention defines an interpersonal comparison; individual preferences alone do not require it. Regularity guarantees its existence and uniqueness up to a common increasing relabeling (Online Appendix Proposition O.2).

Use both paths to give a welfare position two ordinal scores:

$$S_k(x) = \big(m_k^1(x), m_k^2(x)\big), \qquad D = \bigcup_{k=1,2} S_k(X).$$

For bundles on their respective reference paths, calibration gives

$$m_1^1(x) \geq m_2^1(y) \iff x \succsim_2 y, \qquad m_1^2(x) \geq m_2^2(y) \iff x \succsim_1 y.$$

The two score comparisons encode the two people's physical judgments. A function on score pairs is *strictly increasing* if it gives a higher label whenever neither coordinate falls and at least one rises. A common function with this property can express the welfare restrictions without choosing one calibration as the governing index.

3.2. The exact class of welfare scales

The theorem gives equivalent descriptions of every admissible scale: a common increasing labeling of attainable score pairs, or a continuous increasing pairing of the efficient paths within the EC band. Its final claim makes the characterization sharp.

Theorem 1 (Efficient Consensus characterizes interpersonal welfare scales)**.** *Let $\succsim$ be represented by $(\mathbf{m}, \mathrm{R})$ with $m_i(0) = 0$ and $m_i(1) = 1$ for both agents, and let Assumption 1 hold. The following are equivalent.*

(a) $\succsim$ satisfies EC.

(b) For some, equivalently every, pair of welfare scales $\mathbf{m}^1, \mathbf{m}^2$ calibrated to E_1, E_2, there is a strictly increasing function $\Psi : D \to \mathbb{R}$ such that

$$m_k(x) = \Psi\big(m_k^1(x), m_k^2(x)\big), \qquad k = 1, 2.$$

(c) The planner's equivalence map satisfies

$$\mu(x) \in \mathcal{B}(x) \qquad \text{for every } x \in E_1.$$

Moreover, every increasing homeomorphism $\mu : E_1 \to E_2$ satisfying the band condition in (c) *is generated by an EC social preference for any given aggregation rule.*

The result concerns an entire class of *ordinal interpersonal scales* for the given preferences and sharing problem. It does not select numerical labels, identify welfare differences, or impose a social welfare function across preference profiles. Any common increasing relabeling preserves interpersonal rankings. More substantively, different admissible pairings may rank the same two recipients differently while satisfying every EC comparison.

The score representation is defined on the attainable set D, not an arbitrary rectangle of independent scores. Nor must the two single-path calibrations themselves satisfy strict EC. One can tie positions that EC strictly ranks when one personal comparison is an indifference. The admissible scales must respect both scores, including these boundary restrictions.

The sufficiency construction explains why the result characterizes scales, not only inequalities. Start with a permissible pairing μ and assign any increasing labels to person 1's reference path. Give each paired bundle $\mu(x)$ on person 2's path the same label as x. For a bundle off a person's own path, use the label of the unique point to which she is indifferent. This construction represents both personal preferences. It has the chosen interpersonal ties and satisfies EC because the increasing pairing keeps every tie within its permitted band.

There are thus two different freedoms. A common increasing change of numerical labels changes no interpersonal ranking at all. Changing the admissible pairing can change a cross-person ranking while leaving the original preferences fixed. Theorem 1 describes both without confusing numerical normalization with genuine normative discretion. The next corollary identifies what does not depend on that second choice: the comparisons common to every permissible pairing.

Nor do the two reference scores supply cardinal units. Their labels can be changed independently along the two paths, provided the same change is used for both people on a given path. The function Ψ adjusts accordingly, so the characterized class is unchanged. Replacing the first score s by s^2, for example, changes no reference ranking on $[0, 1]$, but changes the numerical average $(s + t)/2$. Fixed-weight averaging therefore adds a convention about the labels; its weights cannot be inferred from ordinal indifference answers.

We can now ask which comparisons survive every choice within the class. If one welfare position has no lower score on either path, all admissible scales agree. If the scores cross, the sharpness construction makes either strict ranking possible without changing the aggregation rule.

Corollary 1 (The EC-robust interpersonal preorder)**.** *Fix welfare scales* $\mathbf{m}^1, \mathbf{m}^2$ *calibrated to* E_1, E_2*, and let* S_k *be defined above. Write* $\succeq^{EC}$ *for a comparison holding under every EC social preference. For any single-recipient positions,*

$$\mathbf{e}_k(x) \succeq^{EC} \mathbf{e}_\ell(y) \quad \Longleftrightarrow \quad S_k(x) \geq S_\ell(y) \qquad \textit{coordinatewise}.$$

If one coordinate inequality is strict, every EC social preference has the corresponding strict ranking. If the two evaluation pairs cross, then for every aggregation rule there are EC social preferences with that rule that rank the two positions in either strict direction.

In the linear example, compare the rural recipient at $x = (1/2, 0)$ with a city recipient at $y = (0, 1/5)$. Both people prefer x, so every EC scale ranks the rural position strictly higher. At $y = (0, 2/5)$, the rural town prefers x but the city prefers y. The score pairs cross, and compatible scales can rank either recipient higher. This incompleteness reflects a precise normative freedom, not a failure to recover an available utility number.

More information about the same preferences cannot eliminate that particular ambiguity. If the exact two scores cross, the theorem constructs welfare scales with opposite interpersonal rankings even after every individual comparison is known. By contrast, uncertain or coarsely elicited scores may conceal a dominance relation that additional questions reveal. The distinction between these cases tells an authority whether to improve its information or acknowledge that it needs a further welfare judgment.

The shared comparisons are invariant to the cardinal utility representations used in an example. Replacing u_i by any strictly increasing transformation leaves the personal equivalents, EC bands and admissible interpersonal rankings unchanged. In particular, the useful comparison of two treated recipients is not recovered by subtracting their utilities before and after treatment. The result is useful precisely because it determines some welfare-level rankings without making those gain comparisons meaningful.

The robust-order corollary connects the characterization to restricted-contour comparison. Decancq and Nys (2021) test nested contours on a specified set of physical alternatives. Taking that set to be $E_1 \cup E_2$, their strict empty-intersection test requires both reference comparisons to be strict. EC yields the same direction there and additionally makes a strict comparison when one score is tied and the other improves. The choice of the efficient domain, the full class of implementing scales and the treatment of equality boundaries are distinct parts of our answer.

3.3. What the rule says about envy

No-envy compares bundles using each recipient's own preferences: a division is envy-free when $x_i \succsim_i x_j$ for both people. It does not itself assign interpersonal welfare labels. Efficiency links the two ideas. At a division where the bundles can feasibly be exchanged, envy by person 1 implies that person 2 strictly prefers keeping her bundle; otherwise the exchange would improve the division. Both then prefer person 2's bundle, so EC ranks her welfare higher.

The next proposition gives the full classification, including the indifference cases. Its first implication needs efficiency and feasible exchange, not convexity. Its converse uses the sharpness of the scale characterization.

Proposition 2 (Envy identifies an EC welfare inequality)**.** *Let Assumption 1 hold, let* $(x_1, x_2) \in E$*, and suppose that* $(x_2, x_1) \in \mathcal{X}$*. If person* i *envies person* j*, so that* $x_j \succ_i x_i$*, then*

$$m_j(x_j) > m_i(x_i)$$

under every EC-compatible scale.

If both people strictly prefer their own bundles, EC-compatible scales can rank either recipient higher or give them equal welfare. If one person is indifferent between the bundles and the other strictly prefers her own, every compatible scale ranks the latter recipient higher. If both are indifferent, every compatible scale gives them equal welfare.

Consequently, some EC-compatible scale equalizes welfare at the division if and only if either both people strictly prefer their own bundles or both are indifferent. Such a division is envy-free, but an envy-free efficient division need not admit equal EC welfare levels.

The result gives an economic interpretation of equality within the admissible class. Equality cannot conceal envy at an efficient, exchangeable division. But no-envy alone does not establish equal welfare. To see precisely what remains open, consider the full efficient frontier of the linear example.

On its first part the rural town receives $x_1 = (s, 0)$ and the city receives $x_2 = (1 - s, 1)$. The rural town values these bundles at $2s$ and $3 - 2s$, respectively. It therefore envies the city for $s < 3/4$, is indifferent at $s = 3/4$, and strictly prefers its own bundle for $s > 3/4$. The city strictly prefers its own bundle throughout this part. On the second part, $x_1 = (1, z)$ and $x_2 = (0, 1 - z)$; now the rural town strictly prefers its own bundle throughout, while the city does so exactly when $z < 1/4$. These inequalities identify an interval of envy-free efficient divisions, including two boundary divisions with one indifferent recipient.

Inside that interval both recipients strictly prefer their own bundles. No interpersonal ranking is forced there. At the efficient corner $(1, 0), (0, 1)$, label each reference path's corner $1/2$. The score pairs are

$$S_1((1, 0)) = (1/2, 3/4), \qquad S_2((0, 1)) = (3/4, 1/2).$$

The continuous, strictly increasing functions $\Psi_\alpha(r, t) = \alpha r + (1 - \alpha)t$, $0 < \alpha < 1$, therefore generate admissible scales. Their difference at this division is $1/4 - \alpha/2$: $\alpha < 1/2$ ranks the rural recipient higher, $\alpha > 1/2$ ranks the city higher, and $\alpha = 1/2$ gives equality. These weights describe alternative permitted labelings of the two fixed scores. They are not identified distributive weights or interpersonal marginal utilities.

At the boundary $x_1 = (3/4, 0)$, $x_2 = (1/4, 1)$, the rural town is indifferent between the bundles, while the city strictly prefers its own. The division is envy-free, but the corresponding score pairs,

$$S_1(x_1) = (3/8, 5/8), \qquad S_2(x_2) = (13/16, 5/8),$$

force strictly higher city welfare under every admissible scale. This is why the proposition distinguishes strict own-bundle preferences from weak no-envy. Omitting the equality boundary would turn a correct one-way fairness implication into an incorrect equivalence.

The efficient-domain restriction is important to this interpretation. At the feasible but inefficient division $x_1 = (0, 1)$, $x_2 = (1, 0)$, each person strictly prefers the other's bundle in the linear example. An unrestricted demand that every envier have lower interpersonal welfare would require each recipient's welfare to be below the other's. Swapping the bundles improves both recipients, so efficiency excludes precisely this source of contradiction. The proposed relation between envy and welfare is therefore conditional, not a general way to read interpersonal inequality from any observed envy.

This gives a connection to fairness-based measurement beyond adopting a familiar fairness label (Fleurbaey and Maniquet, 2017). The welfare scales must be usable across bundles, so their local

fairness implications have to fit into a consistent ordering. Efficient sharing identifies a domain on which the envy implication is meaningful; the sharp characterization establishes which complete scales can respect the full consensus restriction there. No-envy by itself would neither specify those scales nor compare the actual scarce-facility assignments.

This is an interpretation of welfare measurement, not an allocation rule. Egalitarian equivalence instead requires a *common physical bundle* to which all recipients are indifferent (Pazner and Schmeidler, 1978; Thomson, 2011); equality under an arbitrary EC scale need not have that form. The restriction to efficient divisions also matters: an unrestricted rule inferring lower welfare from envy would face contradictory requirements whenever two people envy one another. In the facility application, no feasible deterministic assignment is envy-free, yet the common EC comparisons can still be useful.

So far efficiency has supplied the domain and consensus has supplied the interpersonal restrictions. A further question remains. Could a different physical reference sequence overturn comparisons supported by the efficient pair, and when is checking those two paths enough?

4. WHY THESE REFERENCE PATHS?

Efficient sharing gives the references an economic origin. It does not yet establish how much information they retain. To see the distinction, suppose an authority instead measures everyone's welfare along one common increasing sequence of bundles. A person's welfare position is represented by the bundle on that sequence to which she is indifferent. Different sequences can select different representatives and hence reverse an interpersonal ranking. We ask when the two efficient paths retain every comparison that survives this freedom to change the common reference.

This question builds on reference-based welfare measurement and robust contour comparisons (Fleurbaey and Maniquet, 2017; Decancq and Nys, 2021). Their contribution is to identify what interpersonal comparisons can survive alternative physical benchmarks. Our additional question is whether the *allocation problem* supplies a smallest family of benchmarks with that property. The resource pool is held fixed throughout.

4.1. The information a reference family must retain

Let $\mathcal{A}$ be the family of continuous, nondecreasing paths from 0 to 1 on which both preferences are strictly increasing. A common path $Q \in \mathcal{A}$ assigns equal welfare to its same physical bundle, whoever receives it. Write $q_i^Q(x)$ for the unique point on Q indifferent to x for person i, and m_i^Q for any common increasing numerical calibration of these representatives. Then

$$m_i^Q(x) \geq m_j^Q(y) \quad \Longleftrightarrow \quad q_i^Q(x) \succsim_j y. \tag{4.1}$$

Person i supplies a representative of her own position; person j compares that representative with her bundle. This uses only personal preferences over physical alternatives.

For this section strengthen monotonicity to

$$x \geq y, \quad x \neq y \quad \Longrightarrow \quad x \succ_i y \qquad (i = 1, 2). \tag{SM}$$

Every bundle can then be included in an admissible path. Consequently the possible representatives of x exhaust its indifference contour $I_i(x) = \{z : z \sim_i x\}$. The comparison in (4.1) survives all paths exactly when even the representative that person j values *least* is at least as good as y:

$$m_i^Q(x) \geq m_j^Q(y) \text{ for every } Q \in \mathcal{A} \quad \Longleftrightarrow \quad z \succsim_j y \text{ for every } z \in I_i(x). \tag{4.2}$$

The strict version requires every such z to be strictly preferred to y. This is the familiar nested-contour test. Lemma O.3 and Theorem O.1 in the online appendix give the full statements and proofs.

The linear healthcare example makes the information requirement concrete. The rural bundle $x = (1/2, 0)$ lies on $2s + z = 1$. Along this contour, the city's utility $s + 2z$ ranges from $1/2$ at

$(1/2, 0)$ to 2 at $(0, 1)$. A city receiving $(0, 1/5)$, worth $2/5$, lies below even the least city-valued rural representative. Every common path therefore ranks the rural position higher. For a city receiving $(0, 2/5)$, worth $4/5$, some rural representatives lie above its bundle and others below. The comparison depends on the path.

For example, the diagonal path $Q = \{(t, t) : 0 \leq t \leq 1\}$ represents the rural bundle by $(1/3, 1/3)$. The city values this representative at 1, so that calibration ranks the rural position above a city receiving $(0, 2/5)$. The capacity-first path instead represents the rural bundle by itself, worth $1/2$ to the city, and reverses the ranking. Neither calculation misrepresents the rural town's preferences: both reference bundles are equally good for it. The difference is the physical benchmark used to compare the two recipients.

This is why recovering a single reference equivalent is not the same as recovering a reference-robust comparison. A selected path gives one complete interpersonal scale. Requiring a comparison to survive all paths retains less information but removes dependence on that selection. Theorem 2 will not reconcile all the complete scales. It will identify a particularly small, economically generated family whose *common weak comparisons* equal those of the whole family.

The question now changes from one chosen path to a smaller family that preserves the comparisons common to all paths. Such a family must retain both ends of the other person's evaluation range on *every* contour, and for both people. Missing the minimum can make person i's position look no worse than person j's under every selected path even though an omitted path reverses that ranking. Missing the maximum can create the corresponding error in the reverse comparison. Attaining both extrema is also sufficient, because every other representative lies between them in the other person's preference. Theorem O.2 in the online appendix establishes this equivalence. We next ask when the efficient paths attain these extrema.

The finite-family criterion also clarifies the relation to Decancq and Nys (2021). Restricting the physical evaluation set can add comparisons because some potential contrary evaluations are no longer examined. That is a deliberate weakening of robustness, not a discovery of new personal preferences. In our setting the evaluation set has an economic source: the two efficient projections. To establish that no weak comparison is added merely by omitting other paths, we must show that every omitted representative is bounded by representatives already retained. Attainment of the two extrema is exactly this requirement.

The converse matters as much as sufficiency. If a finite family misses an extreme, continuity supplies an opposing welfare position between that missed value and every retained value. All selected paths then support a comparison that a path through the missed extreme rejects. Thus a smaller reference family cannot be declared adequate simply because it happens to work for a few examples. Adequacy is a statement about every pair of welfare positions at the fixed preferences and endowment.

4.2. Efficient sorting and informational sufficiency

Efficiency and contour extrema concern different comparisons. Efficiency asks how a change in one person's bundle affects the resources *left for the other*. Contour extrema ask how the other person values *that same bundle*. Two sorting conditions separate these tasks.

Let $X = \{(x_1, x_2) \in X^2 : x_1 + x_2 \leq 1\}$, so all resources may be shared. Order resources by person 1's relative valuation. For $r < s$, write $y \triangleright_{rs} x$ when $y_r > x_r$, $y_s < x_s$, and all other coordinates are unchanged. Ordinal comparative advantage requires

$$y \triangleright_{rs} x, \quad y \sim_1 x \quad \Longrightarrow \quad 1 - y \succ_2 1 - x. \tag{OCA}$$

A compensated move toward an earlier resource leaves person 1 indifferent and benefits person 2 through the complementary resources. For instance, the rural town is indifferent between $x = (1/4, 1/2)$ and $y = (1/2, 0)$, both worth 1. The city values the complementary bundles at $7/4$ and $5/2$, respectively, so the second division is better.

Now compare x and y *themselves*. The city values them at $5/4$ and $1/2$, preferring the bundle before the rural town's compensated move. Same-bundle single crossing requires this direction along each person's indifference contour:

$$\begin{aligned} y \triangleright_{rs} x, \quad y \sim_1 x \Longrightarrow x \succ_2 y, \\ y \triangleright_{rs} x, \quad y \sim_2 x \Longrightarrow y \succ_1 x. \end{aligned} \tag{OSC}$$

Continuing such compensated moves toward the boundary identifies the contour extrema. OCA alone sorts efficient divisions; OSC supplies the additional information needed for path sufficiency.

These conditions are ordinal. The compensating change is defined by personal indifference, and the conclusion is a strict preference, so changing a utility representation does not alter either requirement. They are also global conditions: they must hold at every bundle and every indicated resource exchange, not merely at the eventual facility allocation. Convexity of a feasible set does not by itself imply them.

The separate conditions prevent an overstatement of efficiency's role. A gain from reallocating complementary resources does not automatically say how the other person ranks the two bundles offered to the first. In nonseparable or nonlinear environments the complementary bundle can change the relevant tradeoff. That is why the main characterization assumes only regular efficient paths, the policy application first uses common paths, and the all-path minimality claim invokes the stronger same-bundle condition explicitly.

The two boundary paths fill resources in opposite orders:

$$\begin{aligned} G^- &= \bigcup_{k=1}^{L} \{1\}^{k-1} \times [0,1] \times \{0\}^{L-k}, \\ G^+ &= \bigcup_{k=1}^{L} \{0\}^{k-1} \times [0,1] \times \{1\}^{L-k}. \end{aligned} \tag{4.3}$$

Figure 3 shows their intersections with an indifference contour. The next result identifies these paths with the efficient projections and, under OSC, establishes both informational sufficiency and uniqueness.

Theorem 2 (Ordinal sorting makes the efficient pair minimally sufficient)**.** *Maintain* (SM)*. Under* (OCA)*, the efficient projections of the full bilateral sharing problem are*

$$E_1 = G^-, \qquad E_2 = G^+.$$

If (OSC) *also holds, these two paths attain, on every indifference contour of either agent, the unique minimum and maximum of the other agent's preference. Consequently:*

(a) the EC-robust interpersonal preorder generated jointly by E_1, E_2 is exactly the all-path robust preorder in Theorem O.1;

(b) no single admissible reference path reproduces that all-path preorder;

(c) if two admissible paths reproduce the all-path preorder, then, up to relabeling, they are E_1, E_2.

Hence the two efficient projections are the unique minimum-cardinality sufficient reference family.

In the linear example, the efficient paths reach $(1/2, 0)$ and $(0, 1)$, the city's unique least and most valued points on the rural contour. The theorem extends that observation to every contour and every resource dimension under the stated conditions. A single path meets a contour only once and cannot retain two distinct extremes. Moreover, a pair of continuous paths that reaches all the required extremes must be the efficient pair itself.

The equality in part (a) concerns rankings that hold *weakly* under every admissible scale. It does not imply that every shared-path calibration ranks those positions strictly. If one score is tied and the other improves, EC requires a strict ranking, but calibration on the tied path remains indifferent. Nor does the theorem select one welfare scale. It identifies the unique smallest sufficient *physical reference family* within the specified sharing problem.

The scientific implication concerns reference selection. Fairness-based welfare measurement already supplies defensible ways to use reference bundles (Fleurbaey and Maniquet, 2011, 2017). Here the efficiency structure identifies the paths, while the contour argument establishes their informational sufficiency. Neither part is redundant. Efficiency without the contour result would provide a plausible but potentially selective domain. A contour criterion without the efficiency result would say which information a reference family must retain without supplying that family from the resource problem.

This also separates minimality from uniqueness of welfare. Exactly two physical paths suffice under the theorem's assumptions, but many increasing pairings between them may remain admissible. The result removes freedom in the smallest reference family for a stated robustness objective; it does not remove every interpersonal judgment inside the EC bands. Its usefulness is to reduce what must be checked while retaining an honest description of what is not identified.

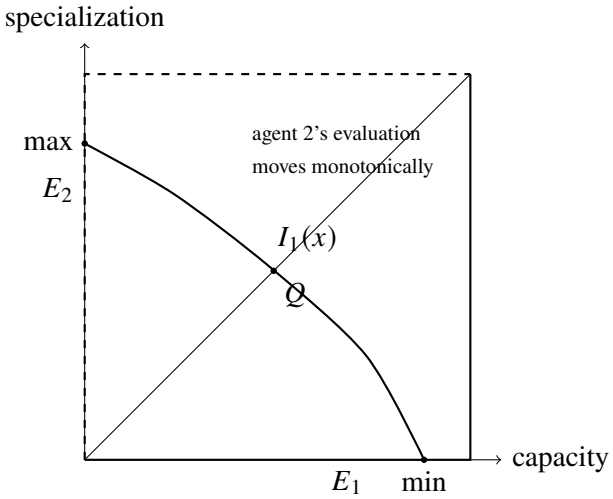


Figure 3. Why the efficient pair is sufficient. Along an indifference contour of agent 1, agent 2's evaluation moves strictly between the two boundary endpoints. The efficient projections E_1, E_2 attain the two extremes, while an interior admissible path Q selects an intermediate point.

The stronger sorting condition is not needed for Theorem 1. Without OSC, efficient references can rank a policy that is not preferred under every shared path. Online Appendix Example O.1 makes this distinction concrete. Its efficient paths rank town 2 higher when treated and town 1 higher when untreated, so placing the facility in town 2 improves both matched positions. Another increasing shared path reverses both comparisons and favors town 1. The preferences and actual allocations are unchanged; only the reference changes.

The example uses smooth preferences that are linear in specialist service but need not be concave in general care. That qualification matters: within this class, concavity together with OCA implies OSC and rules out the reversal (Online Appendix Proposition O.8). The example thus establishes a possible gain in decisiveness, not its prevalence under diminishing marginal valuations. It also shows the normative cost: the extra ranking depends on accepting the efficient references rather than requiring agreement across every shared path.

Finally, none of these results determines which resources belong in the hypothetical pool. Changing that pool changes the comparison problem, not merely its parametrization. We retain this institutional choice explicitly when applying the common welfare comparisons to policy.

5. FROM WELFARE COMPARISONS TO SOCIAL DECISIONS

The preceding results concern welfare positions, such as a particular town receiving a particular bundle. A policy gives one position to *every* town. To evaluate policies we need two further ingredients: bilateral judgments must fit together across the population, and a social evaluation must compare the resulting distributions. Neither requirement is supplied by an individual's preferences alone.

5.1. Combining comparisons and matching positions

Before combining bilateral judgments, consider three welfare positions: town 1 at bundle x, town 2 at y, and town 3 at z. If the reference comparisons required the first to be better than the second, the second no worse than the third, and the third no worse than the first, no welfare scale could satisfy them. These are comparisons of recipients *at specified bundles*, not a ranking of town names. A town's own preference can also link its positions at different bundles.

Let $N = \{1, \dots, n\}$. A connected graph G records the pairs on which the authority imposes EC. Each pair has a specified bilateral sharing problem satisfying the regular-path assumptions. All scales use common empty and full endpoints. Write $\mathcal{M}^{EC}$ for the class of continuous normalized population scales representing the personal preferences and satisfying every pair's EC restriction. For a welfare position (i, x), define

$$(i, x) \succeq^{EC} (j, y) \quad \Longleftrightarrow \quad m_i(x) \geq m_j(y) \quad \text{for every } \mathbf{m} \in \mathcal{M}^{EC}. \tag{5.1}$$

We use this definition only when the class is nonempty.

Let $\succeq^0$ collect direct comparisons from bilateral EC and individual preferences. A *strict cycle* is a finite chain of such comparisons returning to its starting position with at least one strict step. A common scale exists exactly when there is no strict cycle; moreover, the comparisons shared by all common scales are precisely those obtained by chaining direct comparisons. Theorem O.3 in the online appendix proves these assertions, including the continuous representation. This is a setting-specific application of the consistency problem studied by Suzumura (1976); Bossert and Suzumura (2010). The common-reference application below guarantees compatibility directly.

A valid chain must connect the same welfare position at successive steps, or use the recipient's own preferences to connect different bundles. Requiring every comparison to be elicited directly would discard these implications. The consistency theorem tells us both when all the direct judgments fit together and which additional comparisons their chains establish.

The consistency result provides a safeguard before policy evaluation. If the declared bilateral institutions generate a strict cycle, there is no admissible population scale and hence no meaningful universal ranking over such scales. Universal quantification over an empty class would otherwise make every proposed policy comparison vacuously true. An authority facing that conflict must reconsider the collection of reference institutions or the maintained restrictions, rather than interpret inconsistency as evidence in favor of any policy. This is different from the legitimate incompleteness of a nonempty class.

Now return to the two possible facility locations. Suppose the city has the higher welfare level when treated, whereas the rural town has the higher level when untreated. City placement produces a better treated position and a better untreated position. An authority that treats names symmetrically can match these two positions to those under rural placement and prefer city placement. This reasoning compares *levels within treatment states*, not the two towns' gains from treatment.

For many towns, the same logic requires a one-to-one match. One high welfare position cannot be used twice to offset two lower positions under the rival policy. Call a policy comparison *universal* if every EC-compatible scale and every anonymous, strongly Pareto aggregation agree. The Suppes–Sen grading principle supplies the basic anonymous comparison (Suppes, 1966; Sen, 1970; Saposnik, 1983); Pivato (2013) studies aggregation with incomplete interpersonal information. Here we

establish the exact test for the information generated by EC, including the necessity of one matching that works for all compatible scales.

Theorem 3 (Perfect matching characterizes anonymous policy dominance)**.** *Maintain a finite connected comparison graph and assume equivalently that $\mathcal{M}^{EC} \neq \varnothing$ or that the direct bilateral order has no strict cycle. For any two policies $\mathbf{x} = (x_i)_{i\in N}$ and $\mathbf{y} = (y_i)_{i\in N}$, the following are equivalent:*

(a) every anonymous EC social preference ranks $\mathbf{x} \succsim \mathbf{y}$;

(b) there is a permutation π of N such that

$$(i, x_i) \succeq^{EC} (\pi(i), y_{\pi(i)}) \qquad \text{for every } i \in N.$$

Universal strict ranking holds if and only if a permutation in (b) *has a strict matched comparison,*

$$(i, x_i) \succ^{EC} (\pi(i), y_{\pi(i)}) \quad \text{for some } i \in N.$$

Equivalently, (b) *holds and no compatible scale makes the two welfare profiles permutations of one another. If* (b) *holds but no such strict match exists, every anonymous EC social preference is indifferent between the policies.*

The sufficiency argument is immediate once the partners are identified: anonymity permits the permutation, and strong Pareto respects every favorable match. Necessity is the substantive part. If no favorable matching exists, some set of rival positions has too few possible partners. The proof uses this shortage to construct a compatible welfare scale and a symmetric welfare-increasing evaluation that reject the proposed ranking. Thus the test identifies exactly what the maintained principles imply, not merely an easily verified sufficient condition.

Anonymity is an additional ethical restriction. An authority that gives priority to a named town, a protected entitlement, or a different population weight need not accept the same policy ranking. For the application, towns are the welfare recipients; the result does not automatically aggregate the welfare of their residents.

The comparison does not need society to agree on how large one improvement is relative to another. Once a favorable matching exists, every maintained aggregator ranks the resulting profile in the same direction. Evaluators can disagree about equality, priority to the worst off, or the value of improvements at different levels, yet agree on that particular comparison. An unresolved policy pair, however, need not be ethically equivalent. It means that the maintained class contains evaluations with different rankings, so selecting between them requires another restriction.

A favorable match is consequently a certificate of a decision, not a full social ranking of all outcomes. It uses one interpersonal comparison for each matched position. This can be useful when the authority faces a small policy menu or needs to evaluate a proposed transfer. The mathematical certificate alone does not reveal how many preference questions are needed to establish its comparisons, and it does not ensure the respondents have incentives to answer truthfully. Those are separate information and mechanism questions.

5.2. Common efficient references

The comparison is especially simple when every pair uses the same physical paths. Suppose all pairs share the same strictly positive reference endowment, and measure resources in its units. For $i < j$, require population ordinal comparative advantage:

$$y \rhd_{rs} x, \quad y \sim_i x \quad \Longrightarrow \quad 1 - y \succ_j 1 - x. \tag{P-OCA}$$

Maintain strong monotonicity and full bilateral sharing. Applying the efficient-sorting argument to each pair yields

$$E_i^{ij} = E^- = G^-, \qquad E_j^{ij} = E^+ = G^+, \qquad i < j. \tag{CR}$$

The paths arise from the allocation problem, rather than being selected separately for each comparison. Proposition O.3 in the online appendix states the implication. Population OSC is *not* imposed here.

Choose common increasing parametrizations $p^-, p^+ : [0,1] \to E^-, E^+$. For each town, let $m_i^-(x)$ and $m_i^+(x)$ be the parameters of its indifferent bundles on those paths, and write

$$S_i(x) = (m_i^-(x), m_i^+(x)), \qquad D_N = \bigcup_{i \in N} S_i(X).$$

The same physical calibration serves every town. The population counterpart of Theorem 1 says that all compatible scales have the form $m_i(x) = \Psi(S_i(x))$, with one continuous function $\Psi : D_N \to [0,1]$ strictly increasing in the product order and normalized at the endpoints. Such functions exist, so the bilateral requirements are compatible. Theorem O.4 in the online appendix gives the representation.

The common welfare comparisons and the policy test therefore become

$$(i,x) \succeq^{EC} (j,y) \iff S_i(x) \geq S_j(y), \tag{5.2}$$

$$\mathbf{x} \succeq^{U} \mathbf{y} \iff S_i(x_i) \geq S_{\pi(i)}(y_{\pi(i)}) \quad \text{for all } i \text{ and some permutation } \pi. \tag{5.3}$$

All vector inequalities are coordinatewise, and $\succeq^U$ denotes universal policy comparison. A strict matched coordinate gives a strict policy conclusion. These scores locate equivalent physical bundles; they do not measure welfare gains.

5.3. Allocating scarce specialist facilities

There are $K \in \{1, \ldots, n-1\}$ indivisible facilities and one unit of general-care capacity. Every actual policy leaves each town its baseline $a \in (0, 1/n]$, with $1 - na$ units unused. A treated town receives $H_a = (a, 1)$; an untreated town receives $L_a = (a, 0)$. Write p^A for the policy treating the K-town set A.

No distinct policies in this menu Pareto-dominate one another: each transfer benefits its new recipient and harms its former recipient. No deterministic assignment is envy-free either, because every untreated town prefers the treated bundle. The purpose of welfare measurement here is therefore not to find an envy-free allocation. It is to compare feasible policies when that fairness criterion cannot select one.

The hypothetical sharing problem must be specified separately. Protect all outsiders' general-care baselines and the $K - 1$ facilities not involved in a possible transfer. The compared pair can then share one reference facility and the residual general-care capacity:

$$\omega_a = (R_a, 1), \qquad R_a = 1 - (n-2)a. \tag{5.4}$$

The reference facility may be divided fractionally even though actual assignments are indivisible. This institution admits the pair's baselines and unused capacity, but not outsiders' protected resources. It is a declared premise, not an implication of EC.

This distinction matters for interpreting the hospital example. The actual question is where to place complete facilities. The hypothetical question asks what resources two affected towns may use to compare their welfare positions. Allowing fractional specialist service in that hypothetical question does not claim that a facility can be divided in the actual policy. It is comparable to asking someone to evaluate a continuous service level when the available choices are discrete. Whether this is an acceptable benchmark must be defended in the application.

The protection of outsiders also does work. Moving one facility need not reopen every other town's entitlement to a specialist facility. The proposed reference pool keeps those entitlements fixed

while allowing the compared pair to share its own baseline resources and unused capacity. Protecting unused resources as well, or admitting outsiders' facilities, would change the institution. Below we show why allowing more specialist resources into the hypothetical pool can change the welfare comparisons even when every actual allocation stays fixed.

In units of R_a, the actual bundles become

$$L_b = (b, 0), \qquad H_b = (b, 1), \qquad b = a/R_a \leq 1/2. \tag{5.5}$$

Assume the preferences, in these units, satisfy population OCA. Every untreated bundle lies on E^-, and every treated bundle on E^+. Hence their score pairs have the form

$$S_i(L_b) = (\lambda, \ell_i), \qquad S_i(H_b) = (h_i, \eta).$$

All untreated positions share the first coordinate, so ℓ_i alone ranks them. All treated positions share the second, so h_i ranks them. Strict preference for treatment gives $h_i > \lambda$ and $\eta > \ell_i$: every treated position is above every untreated position in both coordinates. The rankings *within* the two groups can nevertheless differ.

For a treated set A, sort its h_i's from lowest to highest and call the list $h(A)^\uparrow$. Likewise let $\ell(A^c)^\uparrow$ list the untreated scores. Since both policies treat K towns, their favorable matches can be checked separately within the two groups. This gives a particularly transparent form of the general matching criterion.

Theorem 4 (Two ordinal rankings determine fixed-K policy comparisons)**.** *Under the preceding common-reference assumptions, for any K-recipient sets A, B, every anonymous EC social preference ranks $p^A \succsim p^B$ if and only if*

$$h(A)^\uparrow \geq h(B)^\uparrow, \qquad \ell(A^c)^\uparrow \geq \ell(B^c)^\uparrow \quad \textit{coordinatewise}. \tag{5.6}$$

The ranking is universally strict if at least one inequality is strict; equality of both sorted lists implies indifference under every such preference.

A policy must therefore improve both distributions: the welfare positions it creates among recipients of the facility and those it leaves among nonrecipients. The criterion is also the intersection of the policy rankings obtained by calibrating everyone on E^- alone and on E^+ alone. The former ties all untreated positions; the latter ties all treated positions. Neither calibration is asserted to satisfy EC by itself. The point is that this fixed-capacity decision need not select between the two benchmarks.

For a single transfer from town j to town i, the test is $h_i \geq h_j$ and $\ell_j \geq \ell_i$. A universally best K-town assignment must therefore select the K highest treated scores *and* the K lowest untreated scores, with ties handled as in Corollary O.3 of the online appendix. The same $2n$ reference answers can be used throughout the policy menu.

Why are the relevant scores levels rather than gains? A high h_i says that town i's treated bundle is equivalent, for that town, to a high point on the capacity-first path. A low ℓ_i says that its untreated bundle is equivalent to a low point on the other path. Subtracting these two labels would combine different physical reference parametrizations. The subtraction is neither needed nor invariant to their relabeling. The assignment test instead compares treated positions with treated positions and untreated positions with untreated positions.

This is an informational advantage for the fixed-capacity problem, not a claim that two answers reveal a town's entire preference ordering. Each answer is used only because every actual outcome is one of two treatment states and lies on one efficient reference path. For a richer policy with different care levels, quality or congestion at every location, more equivalents would generally be needed. The general matching theorem still applies once the relevant welfare comparisons have been established, but the two-ranking reduction would have to be checked for that new menu.

Consider three towns with $a = 1/6$, so $R_a = 5/6$. Their preferences, in physical units, are

$$u_i(s, z) = g_i(s) + z, \qquad g_i(s) = A_i s + \frac{k_i}{2}(s^2 - s),$$
$$A = (2.20, 1.85, 1.50), \qquad k = (-0.55, 0.55, -0.55).$$

These preferences satisfy population OCA at this endowment but not the stronger population OSC. Thus they illustrate the policy theorem without assuming the condition needed for all-path minimality.

Ask each town two personal questions. How much specialist service alone, $(0, z_i)$, equals its untreated bundle? How much general care alone, $(q_i, 0)$, equals its treated bundle? The answers satisfy $z_i = g_i(a)$ and $g_i(q_i) = g_i(a) + 1$:

i	z_i	q_i	treated rank	untreated rank
1	0.405	0.609	3	1
2	0.270	0.717	2	3
3	0.288	0.833	1	2

(rank 1 is highest).

Larger q_i means a higher treated welfare position; larger z_i means a higher untreated position. The rankings are $3 \succ 2 \succ 1$ among treated towns and $1 \succ 3 \succ 2$ among untreated towns. The online appendix verifies the preferences and answers. None of these amounts is a measured interpersonal welfare gain.

With *one* facility, treating town 1 is dominated by treating either town 2 or town 3. The latter two choices remain incomparable. Both leave town 1 untreated. Town 3 supplies the higher treated position, but leaving town 3 untreated supplies the higher remaining untreated position. Each placement wins one comparison, so neither passes both parts of the test.

With *two* facilities, treating towns 2 and 3 is uniquely best. It creates the two highest treated positions and leaves town 1 in the highest untreated position. Against $\{1, 3\}$, match treated town 3 to itself, treated town 2 above treated town 1, and untreated town 1 above untreated town 2. The comparison with $\{1, 2\}$ works similarly. Every admissible anonymous evaluation strictly favors $\{2, 3\}$, although each reallocation harms someone. The same incomplete welfare information thus settles one capacity decision and not another.

The two-facility conclusion can also be understood by reversing the question. Which town should remain untreated? Leaving town 1 untreated creates the highest untreated welfare position. It also leaves the two towns with the highest treated positions to receive facilities. Both considerations favor the same choice. With one facility they conflict: the town with the highest treated position is not the town with the lowest untreated position. The theorem explains the difference without changing the preference information or the reference institution between the two exercises.

Opposite treated and untreated rankings give a useful special case. If higher-indexed towns rank higher when treated and lower when untreated, moving a facility from a lower to a higher index improves both matched positions. Repeating these moves selects the highest K indices, although each transfer remains Pareto-incomparable. Imposing the same-bundle sorting condition on every pair delivers those opposite rankings; Online Appendix Section O.4.2 derives them. The two-ranking theorem itself does not require this additional condition, as the three-town example demonstrates.

Identifying a best assignment need not rank all the remaining policies. With four towns, two facilities, and strict opposite state rankings, the top two towns form a uniquely best assignment. Yet $\{1, 4\}$ and $\{2, 3\}$ are incomparable: the former has the higher best treated position but the lower worst treated position, so neither sorted treated list dominates. This follows directly from Theorem 4. The criterion certifies some decisions rather than completing the entire policy order.

The example also shows how incomplete elicitation can suffice for a particular decision. Compare treating $\{2, 3\}$ with treating $\{1, 3\}$. Town 3 receives the same bundle and can be matched to itself. Suppose ordinal answers place the other relevant physical equivalents in the intervals

$$q_1 \in [0.60, 0.62], \quad q_2 \in [0.71, 0.73], \qquad z_1 \in [0.40, 0.41], \quad z_2 \in [0.26, 0.28].$$

These are hypothetical answer intervals containing the example's exact equivalents, not observations from a survey. Even their least favorable endpoints establish $q_2 > q_1$ and $z_1 > z_2$. The two remaining matches are therefore strict, and the policy ranking follows without locating any of these four indifference points exactly.

This is a decision-focused use of ordinal elicitation, related to the nonparametric approach of Decancq and Nys (2021). Asking whether an actual bundle is preferred to a reference bundle narrows an interval; it does not ask a respondent to supply an interpersonal welfare number. Overlapping intervals might justify further questions. Once exact treated and untreated rankings conflict, however, asking more of the same questions cannot make the policy comparison universal. The distinction separates a practical information problem from an unresolved normative choice.

The argument presumes accurate answers to the stated hypothetical questions. It supplies no incentive-compatibility result and no statistical procedure for noisy reports. Its immediate use is a transparent conditional certificate: anyone accepting the physical consequences, interval information, reference institution and welfare principles can inspect the two matched inequalities and verify the conclusion. An empirical implementation would have to establish those inputs rather than infer them from the theoretical example.

The institutional boundary is substantive. Suppose the actual treated bundle still contains one facility, but the hypothetical pair may share more than one. That bundle now lies inside the reference rectangle rather than on its top edge. Under the stronger sorting conditions, the two reference scores rank different towns' treated positions in opposite directions. No distinct fixed-capacity assignments can then be matched favorably, so neither universally outranks the other (Online Appendix Proposition O.9). Pooling has changed the comparison institution, not the policy's physical effects. Efficient Consensus evaluates welfare within the declared institution; it does not choose that institution.

6. CONCLUSION

Efficient Consensus connects interpersonal welfare measurement to the preferences and resource constraints of an allocation problem. It characterizes a class of ordinal scales rather than selecting one index. Its exact bounds explain both the comparisons shared by all admissible scales and the disagreements they leave open. At efficient, exchangeable bilateral divisions, this information gives envy a robust welfare interpretation.

Under stronger sorting conditions, efficient sharing also supplies the unique smallest family of paths preserving all weak comparisons robust to alternative shared references. Within a fixed institution, the allocation problem can therefore discipline both the welfare restrictions and the physical information needed to apply them. Anonymity and strong Pareto then turn some of these incomplete comparisons into decisive policy rankings, including assignments for which Pareto and no-envy do not select an outcome.

The remaining choices are substantive. The authority must specify the reference institution, identify the welfare recipients, and establish the policies' physical consequences. An unresolved ranking can reflect missing preference information or a genuine limit of the consensus principle. Distinguishing these cases is useful before seeking more precise welfare numbers or imposing a complete social objective.

REFERENCES

Bossert, W. and Suzumura, K. (2010). *Consistency, Choice, and Rationality*. Harvard University Press, Cambridge, MA. doi:10.2307/j.ctvjhzq9s.

Brun, B. C. and Tungodden, B. (2004). Non-welfaristic theories of justice: Is "the intersection approach" a solution to the indexing impasse? *Social Choice and Welfare*, 22(1):49–60. doi:10.1007/s00355-003-0276-0.

Bundesministerium für Gesundheit (2024). Bundestag beschließt krankenhausreform. Press release, 17 October. Accessed 9 November 2026. URL: https://www.bundesgesundheitsministerium.de/presse/pressemitteilungen/bundestag-beschliesst-krankenhausreform-pm-17-10-2024.

Decancq, K. and Nys, A. (2021). Non-parametric well-being comparisons. *European Economic Review*, 133:103666. doi:10.1016/j.euroecorev.2021.103666.

Fleurbaey, M. (2007). Social choice and the indexing dilemma. *Social Choice and Welfare*, 29(4):633–648. doi:10.1007/s00355-007-0248-x.

Fleurbaey, M., Hagnéré, C., and Trannoy, A. (2003). Welfare comparisons with bounded equivalence scales. *Journal of Economic Theory*, 110(2):309–336. doi:10.1016/S0022-0531(03)00035-8.

Fleurbaey, M. and Maniquet, F. (2011). *A Theory of Fairness and Social Welfare*. Cambridge University Press. doi:10.1017/CBO9780511851971.

Fleurbaey, M. and Maniquet, F. (2017). Fairness and well-being measurement. *Mathematical Social Sciences*, 90:119–126. doi:10.1016/j.mathsocsci.2016.11.003.

Fleurbaey, M. and Schokkaert, E. (2011). Equity in health and health care. In Pauly, M. V., McGuire, T. G., and Barros, P. P., editors, *Handbook of Health Economics, Volume 2*, pages 1003–1092. Elsevier. doi:10.1016/B978-0-444-53592-4.00016-5.

Fleurbaey, M. and Tadenuma, K. (2014). Universal social orderings: An integrated theory of policy evaluation, inter-society comparisons, and interpersonal comparisons. *The Review of Economic Studies*, 81(3):1071–1101. doi:10.1093/restud/rdu006.

Hall, P. (1935). On representatives of subsets. *Journal of the London Mathematical Society*, s1-10(1):26–30. doi:10.1112/jlms/s1-10.37.26.

HM Treasury (2026). The Green Book: UK government guidance on appraisal. Technical report, HM Treasury, London. Updated 28 October 2026; accessed 9 November 2026. URL: https://www.gov.uk/government/publications/the-green-book--2/the-green-book-2026.

Maniquet, F. and Moramarco, D. (2026). Well-being measurement with reference consumption. *Economic Theory*, 82(2):437–476. doi:10.1007/s00199-026-01706-3.

Mseke, E. P., Jessup, B., and Barnett, T. (2024). Impact of distance and/or travel time on healthcare service access in rural and remote areas: A scoping review. *Journal of Transport & Health*, 37:101819. doi:10.1016/j.jth.2024.101819.

Pazner, E. and Schmeidler, D. (1978). Egalitarian equivalent allocations: A new concept of economic equity. *Quarterly Journal of Economics*, 92(4):671–687. doi:10.2307/1883182.

Pivato, M. (2013). Social welfare with incomplete ordinal interpersonal comparisons. *Journal of Mathematical Economics*, 49(5):405–417. doi:10.1016/j.jmateco.2013.04.004.

Ramsay, A. I. G., Tomini, S. M., Gandhi, S., Fulop, N. J., and Morris, S. (2025). Centralisation of specialised healthcare services: a scoping review of definitions, types, and impact on outcomes. *Health and Social Care Delivery Research*, 13(43):21–90. doi:10.3310/REMD6648.

Sakamoto, N. (2018). Equity criteria based on the dominance principle and individual preferences: Refinements of the consensus approach. RCNE Discussion Paper Series 5, Research Center for Normative Economics, Institute of Economic Research, Hitotsubashi University. URL: https://econpapers.repec.org/paper/hitrcnedp/5.htm.

Saposnik, R. (1983). On evaluating income distributions: Rank dominance, the Suppes–Sen grading principle of justice, and pareto optimality. *Public Choice*, 40:329–336. doi:10.1007/BF00114529.

Sen, A. K. (1970). *Collective Choice and Social Welfare*. Holden-Day, San Francisco. URL: https://search.worldcat.org/title/140528.

Sprumont, Y. (2012). Resource egalitarianism with a dash of efficiency. *Journal of Economic Theory*, 147(4):1602–1613. doi:10.1016/j.jet.2010.07.009.

Suppes, P. (1966). Some formal models of grading principles. *Synthese*, 16(3–4):284–306. doi:10.1007/BF00485084.

Suzumura, K. (1976). Remarks on the theory of collective choice. *Economica*, 43(172):381–390. doi:10.2307/2553273.

Thomson, W. (2011). Fair allocation rules. In Arrow, K. J., Sen, A. K., and Suzumura, K., editors, *Handbook of Social Choice and Welfare*, volume 2, chapter 21, pages 393–506. Elsevier. doi:10.1016/S0169-7218(10)00021-3.

Varian, H. R. (1974). Equity, envy, and efficiency. *Journal of Economic Theory*, 9(1):63–91. doi:10.1016/0022-0531(74)90075-1.

Varian, H. R. (1976). Two problems in the theory of fairness. *Journal of Public Economics*, 5(3–4):249–260. doi:10.1016/0047-2727(76)90018-9.

A. PROOFS OF THE PRINCIPAL RESULTS

This appendix proves the scale characterization and its robust order, the envy implication, efficient-reference minimality, and the two policy results. The online appendix supplies the supporting path, continuity, and population-compatibility lemmas. Each proof below gives the argument for its stated main result rather than substituting a reference to an omitted proof.

A.1. The admissible scales and their common comparisons

For the theorem's score representation, it is useful to fix one *social welfare level* c and ask where that same label appears on each person's efficient reference path. The unique-intersection property gives bundles $X_1(c) \in E_1$ and $X_2(c) \in E_2$. We now record how the *other person* ranks these two physical reference bundles; that information determines the signs of the EC inequalities at this social tie.

$$a \equiv m_1 \circ X_2, \qquad b \equiv m_2 \circ X_1.$$

Both are increasing homeomorphisms because the reference paths and the corresponding ordinal evaluations rise continuously from the common bottom to the common top. At the cross-person indifference $(X_1(c), X_2(c))$,

$$X_1(c) \succsim_1 X_2(c) \iff c \geq a(c), \qquad X_2(c) \succsim_2 X_1(c) \iff c \geq b(c). \tag{A.1}$$

At social tie level c, Lemma O.2 permits the tie precisely when both towns disagree on which physical bundle is better or are both indifferent. The sign comparison in (A.1) restates this: $c - a(c)$

and $c - b(c)$ must have the same sign, which converts the economic no-forbidden-tie condition into the inequalities used to construct the score representation.

Choose one scale calibrated to each reference path. To locate the social welfare level c in those *two ordinal coordinate systems*, define

$$f = m_1^1 \circ X_1, \qquad g = m_2^2 \circ X_2.$$

These are increasing homeomorphisms. The reference-path calibration conditions imply

$$m_1^2(X_1(c)) = g(a^{-1}(c)), \qquad m_2^1(X_2(c)) = f(b^{-1}(c)).$$

For example, $a^{-1}(c)$ is the social label of the point on E_2 that person 1 finds indifferent to $X_1(c)$, since $m_1(X_2(a^{-1}(c))) = c$. The second calibrated scale gives this physical point the same label for either recipient. Its second score is therefore $g(a^{-1}(c))$. The corresponding calculation on E_1 gives $f(b^{-1}(c))$. These formulas record the two individual equivalences at a fixed social tie; they do not assume that the individual equivalences themselves are social ties.

The path-calibration identities now locate the two welfare positions tied at level c as the following two score pairs:

$$P_c = \big(f(c), g(a^{-1}(c))\big) \quad \text{for agent 1}, \qquad Q_c = \big(f(b^{-1}(c)), g(c)\big) \quad \text{for agent 2}.$$

Proof of Theorem 1. There are three tasks in the theorem's proof. We must show that the score conversion respects EC, that every EC scale admits such a conversion, and that every increasing selection inside the EC bands can be implemented. First note how the reference scores encode physical agreement. Because f and g increase, $P_c \leq Q_{c'}$ means both towns weakly prefer $X_2(c')$ to $X_1(c)$; $Q_c \leq P_{c'}$ reverses the physical comparison; within the same person's positions, both scores follow that person's own ordering.

Suppose a strictly increasing conversion Ψ represents social welfare from the two reference scores; at any social tie level c, it assigns $\Psi(P_c) = \Psi(Q_c)$. If both towns weakly preferred one physical bundle to the other, with one strict preference, the associated score pair would strictly dominate its rival; strict monotonicity of Ψ would then prohibit equal social labels. By Lemma O.2, the social preference satisfies EC.

Conversely, suppose the social preference satisfies EC. Every social welfare level c is attained by a position on each reference path. Define a common function Ψ that gives *both* reference-score pairs corresponding to level c the number c:

$$\Psi(P_c) = \Psi(Q_c) = c.$$

To check that this construction defines one function, first consider two positions of the same person. Their two reference scores follow the same personal ordering. Equal scores therefore imply personal indifference and equal social labels. For positions of different people, replace each bundle by its indifferent point on its recipient's own reference path. This replacement changes neither reference score nor social label. Equal score pairs then mean that both people are indifferent between the two reference bundles, so EC in both directions equates their labels. Hence intersecting score curves never assign two different values to Ψ.

The same argument proves strict monotonicity on D. For positions of one person, a strict coordinatewise improvement is a strict personal improvement. For positions of different people, the two favorable score comparisons encode the two weak physical preferences; one strict coordinate makes at least one personal preference strict. EC therefore requires a strict welfare improvement. These cases cover all pairs in D. In particular, we have not assumed that every point of the surrounding rectangle is attainable or extended Ψ to points outside D. By construction,

$$\Psi\big(m_i^1(x), m_i^2(x)\big) = m_i(x)$$

for every x and i.

Continuity holds on the attainable set without requiring an extension to the whole score rectangle. Let $F_1(c) = P_c$ and $F_2(c) = Q_c$, and write $D_i = F_i([0,1])$. Each F_i is continuous and one-to-one: a higher personal welfare level has strictly higher coordinates on both reference paths. Because the parameter interval is compact and the score space is Hausdorff, F_i is a homeomorphism onto its image. The assigned label on D_i is its continuous inverse, F_i^{-1}.

Both images are compact and hence closed in $D = D_1 \cup D_2$. The well-definedness argument above shows that the inverse labels agree on their intersection. The pasting lemma for these two closed sets therefore makes Ψ continuous on D. In particular, intersections of the score curves introduce neither inconsistent labels nor discontinuities; no assumption that the curves are disjoint is needed. This also explains the "some, equivalently every" quantifier. Suppose another pair of calibrations replaces the labels on reference path E_j by an increasing homeomorphism h_j. Since each calibration gives both people the same labels on that path and extends them through their own indifference, it transforms both people's jth scores by that same h_j. The new attainable score set is the coordinatewise image of D, and

$$\widetilde{\Psi}(r,t) = \Psi\big(h_1^{-1}(r), h_2^{-1}(t)\big)$$

gives exactly the original welfare labels on it. Continuity and strict product monotonicity are preserved. Thus the representation does not depend on a particular numerical parametrization of either physical path.

To recover the EC band, fix a rural-path bundle x. Let $u = \tau_2(x)$ be the *city's* equivalent on E_2, and let $v = \tau_1^{-1}(x)$ be the *rural town's*; the social-equivalence map chooses $y = \mu(x)$. If $u = v$, the towns agree on one equivalent, and tying x to any other city-path bundle would violate EC. Thus $y = u$ is required.

If $u \neq v$, a tie at or beyond either personal endpoint would conflict with a weak physical agreement that includes a strict preference. A point *strictly between* them instead has opposite personal rankings and permits the social tie; the same argument works whichever town's equivalent lies higher. Hence the allowed social equivalents are exactly those in $\mathcal{B}(x)$.

Sharpness requires the reverse construction. Take *any* increasing homeomorphism μ that selects a permitted social equivalent inside each EC band; start with a normalized representation m_1 of the rural town's own preferences. On the city path E_2, assign every bundle $\mu(x)$ the same welfare number as its paired rural bundle x:

$$m_2(\mu(x)) = m_1(x), \qquad x \in E_1,$$

For an arbitrary city bundle z, let $q_2(z) \in E_2$ be its unique indifferent representative. The extension is explicitly

$$m_2(z) = m_1\big(\mu^{-1}(q_2(z))\big).$$

If v_2 is any continuous normalized representation of the city's preferences, then $q_2 = (v_2|_{E_2})^{-1} \circ v_2$. Proposition O.1 makes the restricted map a homeomorphism, so the extension is continuous. It represents the city's ordering because each of q_2, μ^{-1}, and $m_1|_{E_1}$ preserves the relevant strict path order. It also shares the zero and full-resource labels with m_1.

The resulting scale realizes precisely the prescribed ties μ. Every tie lies in its EC band, so the no-forbidden-tie lemma gives EC. Single-recipient comparisons depend only on the two welfare labels, not on how the aggregation rule ranks other profiles. The construction therefore works for any maintained aggregation rule, as claimed.

□

Proof of Corollary 1. If one position scores at least as high as another on both reference paths, Theorem 1(b) ranks it higher under every EC scale, strictly when a score improves. Two positions

belonging to the *same* person always have concordant scores; to prove that *crossed* scores between different people really leave both directions open, move the positions to their personal indifference equivalents on the reference paths. The individuals disagree over those physical bundles, so the relevant city-path point lies strictly within the EC band, and we construct valid social-equivalence maps placing it on either side.

Parameterize the paths from 0 to 1. Let L and U denote the lower and upper EC-band boundaries. Both increase continuously and share endpoints. Let $P = \{t : L(t) = U(t)\}$, where the social equivalent is pinned down. On each connected component J of $\{t : L(t) < U(t)\}$, any interior choice between the boundaries is permitted. Choose a weight $\theta_J \in (0, 1)$ and set

$$M(t) = (1 - \theta_J)L(t) + \theta_J U(t) \quad (t \in J), \qquad M(t) = L(t) = U(t) \quad (t \in P).$$

Each convex combination is continuous and strictly increasing on its component and lies strictly inside the band. At a point of P, including a limit of components, continuity follows by squeezing between the continuous bounds L, U, which agree there. Strict increase also holds across components: a point of P between them has a strictly intermediate image because both bounds increase strictly. The common endpoints are 0, 1, so M is an increasing homeomorphism of the whole interval.

Let t_0 parameterize the disputed first-person position and let y_0 be the second person's path parameter. Crossed scores put y_0 strictly between $L(t_0)$ and $U(t_0)$. The number

$$\theta_0 = \frac{y_0 - L(t_0)}{U(t_0) - L(t_0)}$$

therefore lies strictly between zero and one. Choose the component's weight below θ_0 for one construction and above it for another, retaining weight $1/2$ elsewhere. These maps put the social equivalent on opposite sides of y_0. Sharpness in Theorem 1 implements both strict interpersonal rankings under the same given aggregation rule. The interpolation is a construction in path parameters, not a comparison of cardinal welfare differences. □

A.2. Envy at efficient bilateral divisions

Proof of Proposition 2. First consider envy. Suppose $x_2 \succ_1 x_1$. If person 2 weakly preferred x_1 to x_2, the feasible exchanged division would improve person 1's welfare without harming person 2, contrary to efficiency. Hence $x_2 \succ_2 x_1$. Because $x_i \in E_i$, EC applies to these two bundles and ranks the recipient of x_2 strictly higher. The argument is symmetric when person 2 envies person 1.

Now suppose $x_1 \succ_1 x_2$ and $x_2 \succ_2 x_1$. The calibration identities for bundles on their respective reference paths give

$$m_1^1(x_1) < m_2^1(x_2), \qquad m_1^2(x_1) > m_2^2(x_2).$$

Thus the score pairs cross. Corollary 1 supplies normalized EC-compatible scales $\mathbf{m}^+$ and $\mathbf{m}^-$ with

$$d^+ \equiv m_1^+(x_1) - m_2^+(x_2) > 0, \qquad d^- \equiv m_1^-(x_1) - m_2^-(x_2) < 0.$$

Choose $t = -d^-/(d^+ - d^-) \in (0, 1)$ and set $m_i^0 = t m_i^+ + (1 - t) m_i^-$ for each i. Each m_i^0 is continuous, normalized, and represents the same individual preferences, since both component functions do. Every weak or strict EC comparison is preserved by this common positive combination. Hence $\mathbf{m}^0$ is EC-compatible, and its two welfare labels at the division are equal. The displayed differences are used only to construct an equalizing scale; they have no cardinal welfare interpretation.

If person 1 is indifferent and person 2 strictly prefers her own bundle, both weakly prefer x_2, with one preference strict. EC therefore ranks person 2 strictly higher. The reverse case is symmetric. If both are indifferent, EC in both directions forces equality.

These cases exhaust efficient, swap-feasible divisions: whenever either person envies, the first part applies; otherwise each weakly prefers her own bundle. Equality is impossible in the envy cases and in the cases with one indifference and one strict own-bundle preference. It is possible when both own-bundle preferences are strict, and required when both are indifferent. This proves the final equivalence. □

A.3. The unique smallest sufficient reference family

Proof of Theorem 2. We first use OCA to show that efficient bilateral divisions have the two greedy projections. We next use OSC to locate the other person's minimum and maximum evaluations on every indifference contour. The finite-family theorem then proves sufficiency, while the fact that the two extreme points are distinct proves minimality and uniqueness.

Strong monotonicity rules out wasting resources in an efficient division, so every candidate has form $(x, 1-x)$. If $x \notin G^-$, Lemma O.4 moves person 1 to her unique indifferent bundle $x^- \in G^-$. Each transfer shifts toward an earlier resource while leaving person 1 indifferent. By OCA, person 2's complementary bundle improves strictly. Thus $1-x^- \succ_2 1-x$, and the original division cannot be efficient. On any contour, $1-x^-$ is consequently the *unique best complementary bundle* for person 2.

Conversely, suppose $(x, 1-x)$ with $x \in G^-$ were Pareto-dominated by $(y, 1-y)$. If person 1 were indifferent between x and y, the strict maximum of person 2's complementary valuation on that contour would force $y = x$, contradicting dominance. If person 1 strictly preferred y, continuity and strong monotonicity would allow us to scale y down to $z = \alpha y \leq y$, with $\alpha < 1$ and $z \sim_1 x$. Then person 2 would strictly prefer $1-z$ to $1-y$, which by the proposed dominance is weakly preferred to $1-x$, but $1-x$ is already person 2's best complement on person 1's indifference contour. Contradiction. Therefore $E_1 = G^-$, and complements give $E_2 = G^+$.

Now impose OSC as well. On a person-1 indifference curve, each compensated move toward G^- strictly lowers person 2's evaluation of that same personal bundle. Every point outside G^- admits a finite sequence with at least one strict step, so the unique greedy intersection has strictly lower value than every other point on the contour. Reversing the resource order applies the same condition in the opposite direction and makes the G^+ intersection the unique maximum. The argument includes boundary bundles because the compensated-sorting lemma requires only an unfilled earlier coordinate and a positive later one.

On a person-2 contour, the second implication of OSC reverses the roles of the two paths: sorting toward G^- strictly raises person 1's evaluation, and sorting toward G^+ strictly lowers it. Thus both people's contours have the required unique extrema. At the empty and full contours the extrema coincide at the unique bundle; no strict distinction between them is required there. By Theorem O.2, the efficient paths then preserve every comparison robust to arbitrary shared paths.

When $L \geq 2$, the two greedy paths meet only at 0 and 1; every interior indifference contour therefore has two distinct extreme representatives. One shared path crosses that contour only once and cannot select both, proving that at least two paths are needed.

For uniqueness, suppose another pair Q_1, Q_2 suffices. Parameterize person 1's indifference curves continuously by $c \in (0,1)$. Let $z_k(c)$ be the unique point where Q_k meets the contour, and $g^-(c), g^+(c)$ the corresponding greedy-path intersections. Sufficiency and uniqueness of the extrema force each $z_k(c)$ to equal one of $g^-(c), g^+(c)$, with the two selected points covering both extrema:

$$\{z_1(c), z_2(c)\} = \{g^-(c), g^+(c)\}$$

for every interior c, with $g^-(c) \neq g^+(c)$.

A continuous path Q_1 cannot switch from the first extreme to the second as c varies: the two sets of levels where its intersection equals either extreme are disjoint, closed relative to the connected interval $(0,1)$, and together cover that interval. One set must therefore be the whole interval. Q_2 must trace the other extreme, and continuity fixes the endpoints. Thus $\{Q_1, Q_2\} = \{G^-, G^+\}$. □

A.4. Anonymous policy comparisons

Proof of Theorem 3. The matching theorem's sufficiency is short: anonymity reorders matched recipients and strong Pareto compares their levels; necessity needs a construction. If no *robust* perfect matching exists, Hall's theorem finds rival-policy positions that cannot all get distinct favorable partners. Different EC scales might individually reverse different missing comparisons, so we

combine the witness scales into *one* compatible scale and then choose *one* anonymous rank-sensitive evaluation that favors the rival. We finally check the strict and tied cases.

If a robust perfect matching π exists, every compatible scale satisfies

$$m_i(x_i) \geq m_{\pi(i)}(y_{\pi(i)}) \quad \text{for all } i.$$

Anonymity lets us put the recipients into the matching order, and strong Pareto ranks $\mathbf{x} \gtrsim \mathbf{y}$. This proves sufficiency of a favorable robust perfect matching.

Now suppose no perfect matching pairs every rival-policy position with a different robustly favorable position under $\mathbf{x}$. Form the bipartite graph whose left vertices are the positions under $\mathbf{x}$, whose right vertices are those under $\mathbf{y}$, and whose edges are robustly favorable comparisons. For a set J of right vertices, let $\Gamma(J)$ be the left vertices that dominate at least one position in J. *Hall's marriage theorem* (Hall, 1935) gives a nonempty J with too few possible partners: $|\Gamma(J)| < |J|$. For every winner position $i \notin \Gamma(J)$ and rival position $j \in J$, robust dominance fails, and therefore *some* compatible scale $\mathbf{m}^{ij}$ puts i's position strictly below j's. There are finitely many such failures. List the witness scales as $\mathbf{m}^1, \ldots, \mathbf{m}^K$, and record each welfare position's labels across these scales as a vector:

$$A_i = (m_i^1(x_i), \ldots, m_i^K(x_i)), \qquad B_j = (m_j^1(y_j), \ldots, m_j^K(y_j)).$$

On the finite set $F = \{A_i, B_j : i, j \in N\}$, define the product upper set

$$U = \{z \in F : z \geq B_j \text{ coordinatewise for some } j \in J\}.$$

Every rival vector B_j in J belongs to the coordinatewise upper set U. A proposed-winner vector A_i with $i \notin \Gamma(J)$ does *not*: for every $j \in J$, at least one witness scale puts A_i below B_j in some coordinate; this remains true even if distinct welfare positions yield the same vector. As U is closed upward under coordinatewise dominance, we can extend the finite partial order into a strict total ranking that puts all vectors in U above those outside. Lemma O.6 gives one continuous, strictly increasing function Φ and a cutoff c separating these groups:

$$\Phi(B_j) > c \ (j \in J), \qquad \Phi(A_i) < c \ (i \notin \Gamma(J)).$$

Apply Φ to the vector of witness-scale labels of *every* person's bundle, not just the finite policy positions: $m_i^*(z) = \Phi(m_i^1(z), \ldots, m_i^K(z))$. Lemma O.5 ensures that this single scale is EC-compatible after common affine normalization. In detail, every component represents the same personal order, and every EC-required comparison is favorable in all components, strictly in at least one whenever required. Strict product monotonicity preserves these comparisons. The common lower and upper endpoint values are $\Phi(0, \ldots, 0)$ and $\Phi(1, \ldots, 1)$; their difference is positive, so one affine transformation normalizes every person's scale together.

Normalize the cutoff c in the same way. Write $a_i = m_i^*(x_i)$, $b_i = m_i^*(y_i)$, and use parentheses for their increasing order statistics. At least $|J|$ rival positions lie above the cutoff, whereas at most $|\Gamma(J)| < |J|$ proposed-winner positions do. The rival therefore has strictly higher welfare at rank $n - |J| + 1$:

$$a_{(n-|J|+1)} < c < b_{(n-|J|+1)}.$$

For sufficiently small $\varepsilon > 0$, the anonymous ordering represented by

$$W_\varepsilon(z) = z_{(n-|J|+1)} + \varepsilon \sum_{\ell=1}^{n} z_\ell$$

ranks $\mathbf{y}$ strictly above $\mathbf{x}$. The order statistic is weakly increasing in every coordinate, and the positive sum term makes the ordering strongly Pareto.

To make the reversal explicit, let $\delta = b_{(n-|J|+1)} - a_{(n-|J|+1)} > 0$. All normalized welfare labels lie in $[0, 1]$, so the sum term can favor $\mathbf{x}$ over $\mathbf{y}$ by at most εn. Any $0 < \varepsilon < \delta/n$ leaves the rival's advantage at the chosen rank strictly positive. This produces one admissible scale and one admissible social evaluation that reverse the proposed comparison. A robust perfect matching is therefore necessary for universal dominance.

Finally consider strictness. A robust perfect matching may contain either positions that every compatible scale ties or a comparison that every scale makes *strict*. The population theorem shows that the first case corresponds to chains working both ways, and the second to a chain working in only one direction. One strict matched improvement makes the policy ranking strict under strong Pareto for every compatible scale. If every match is a robust tie, the policies give the same multiset of welfare numbers under every scale and every anonymous rule is indifferent. A perfect matching with a strict edge cannot coexist with a different all-ties matching: chaining the latter would force the supposedly strict comparison to be reversible.

□

A.5. Fixed-capacity facility assignments

Proof of Theorem 4. Every town prefers treatment, so $h_i > \lambda$ and $\eta > \ell_i$. Every treated position therefore strictly dominates every untreated position in both score coordinates, even when the two recipients differ.

Suppose a favorable perfect matching paired a proposed-winner treated position with a rival untreated position. Because both policies have exactly K treated positions, some proposed-winner untreated position would then have to be paired with a rival treated position. That comparison is impossible: both of the untreated position's coordinates are lower. Hence every favorable matching separates into a matching of the treated positions and a matching of the untreated positions.

Within the treated group, the second coordinate is constant. Only the scalar list of h_i's matters. If the proposed winner's sorted list is at least the rival's at every rank, pairing corresponding ranks is favorable. Conversely, suppose the inequality fails at rank k, with both lists written increasingly. The rival has $K - k + 1$ positions at or above its kth value, whereas at most $K - k$ proposed-winner positions can reach that value. Those rival positions cannot all have distinct favorable partners. The same argument applies to the untreated group, whose first coordinate is constant and whose varying list consists of the ℓ_i's.

Thus the two sorted-list inequalities are necessary and sufficient for a favorable matching. A strict rank inequality gives a strict matched comparison. If both lists are equal, matching equal scores makes the two policies permutations of the same welfare levels under every compatible scale. The general policy theorem then gives the stated weak, strict and indifference conclusions. □