E-companion: "A Unified Framework for Analyzing and Optimizing a Class of Convex Fairness Measures"

EC.1. Properties of the Deviation-Based Fairness Measures in Table 1

EC.1.1. Equivalence

In this section, we investigate the equivalence between fairness measures in Table 1. We say two fairness measures ϕ^1 and ϕ^2 are equivalent if there exists $\beta > 0$ such that $\phi^1(\boldsymbol{u}) = \beta\phi^2(\boldsymbol{u})$ for all $\boldsymbol{u} \in \mathbb{R}^N$. Hence, replacing ϕ^2 by ϕ^1 as a fairness criterion in the fairness-promoting optimization models, e.g., models (2) and (3), essentially scales the weight on the fairness criterion γ or the upper bound η by a factor of $\beta > 0$ (i.e., from ϕ^2 to $\beta\phi^2$). It is easy verify that all of these measures are equivalent when $N = 2$. However, in Proposition EC.1, we show that only some of these measures are equivalent when $N \geq 3$,

PROPOSITION EC.1. *Only the following equivalence relationships between fairness measures shown in Table 1 hold: (a) for any $N \geq 3$, (i) and (iii) are equivalent; (b) for any $N \geq 3$, (vi) and (vii) are equivalent; (c) when $N = 3$, (i), (ii), and (iii) are equivalent; (d) when $N = 3$, (iv), (vi), and (vii) are equivalent. The remaining pairs of fairness measures are not equivalent.*

Proof. We first prove (a)–(d). Without loss of generality, we assume that $\boldsymbol{u}$ is sorted in ascending order, i.e., $u_1 \leq u_2 \leq \cdots \leq u_N$.

(a) Note that the maximum pairwise difference (iii) is equal to $u_N - u_1$, which is the same as the range (i).

(b) We claim that $\max_{i\in[N]} \sum_{j=1}^N |u_i - u_j| = N \max_{i\in[N]} |u_i - \overline{u}|$. To prove this claim, note that we can write (vii) as

$$\max_{i\in[N]} \sum_{j=1}^{N} |u_i - u_j| = \max\left\{ \sum_{i=2}^{N} u_i - (N-1)u_1,\, (N-1)u_N - \sum_{i=1}^{N-1} u_i \right\} =: \max\{C_1, C_2\},$$

where C_1 and C_2 represent the first and second expressions in the $\max$ operator, respectively. Consider the case when $C_1 \leq C_2$. This implies that $2\sum_{j=2}^{N-1} u_i \leq (N-2)(u_N + u_1)$. Adding $2(u_1 + u_N)$ on both sides of the inequality results in $2\overline{u} \leq u_1 + u_N$, implying that $\overline{u} - u_1 \leq u_N - \overline{u}$. Hence, we have

$$\max_{i\in[N]} \sum_{j=1}^{N} |u_i - u_j| = (N-1)u_N - \sum_{i=1}^{N-1} u_i = N(u_N - \overline{u}) = N \max_{i\in[N]} |u_i - \overline{u}|.$$

A similar argument holds for the case when $C_1 \geq C_2$.

(c) By (a), it suffices to show that (i) and (ii) are equivalent when $N = 3$. By Mesa et al. (2003), we can write (ii) as

$$\sum_{i=1}^{3} \sum_{j=1}^{3} |u_i - u_j| = \sum_{i=1}^{3} 2(2i-4)u_i = 4(u_3 - u_1),$$

which shows the equivalence between (i) and (ii).

Table EC.1 Examples that some fairness measures are not equivalent

E.g.	Outcome vectors (in $\mathbb{R}^N$)	(i)	(ii)	(iv)	(v)	(vi)	(viii)
A	$\boldsymbol{u}^1=(1,\,2,\,2.5,\ldots,\,2.5,\,4.5)$	3.5	$14+8(N-3)$	4	$\sqrt{6.5}$	2	$9.5+2(N-3)$
	$\boldsymbol{u}^2=(1,\,1,\,2,\ldots,\,2,\,4)$	3	$12+8(N-3)$	4	$\sqrt{6}$	2	$9+2(N-3)$
B	$\boldsymbol{u}^1=(2,\,5,\,5,\ldots,\,5,\,9)$	7	$28+14(N-3)$	/	$\sqrt{25-\frac{1}{N}}$	/	$18+4(N-3)$
	$\boldsymbol{u}^2=(2,\,2,\,4,\ldots,\,4,\,8)$	6	$24+16(N-3)$	/	$\sqrt{24}$	/	$18+4(N-3)$
B'	$\boldsymbol{u}^1=(2,\,5,\,5,\,6,\,9)\in\mathbb{R}^5$	/	60	/	/	/	26
	$\boldsymbol{u}^2=(2,\,2,\,4,\,4,\,8)\in\mathbb{R}^5$	/	56	/	/	/	26
C	$\boldsymbol{u}^1=(2,\,5,\,\frac{16}{3},\ldots,\,\frac{16}{3},\,9)$	7	$28+\frac{44}{3}(N-3)$	/	$\sqrt{\frac{74}{3}}$	/	/
	$\boldsymbol{u}^2=(2,\,2,\,\frac{13}{3},\ldots,\,\frac{13}{3},\,9)$	7	$28+\frac{56}{3}(N-3)$	/	$\sqrt{\frac{98}{3}}$	/	/
D	$\boldsymbol{u}^1=(1,\,2,\,3,\ldots,\,3,\,6)$	/	$20+12(N-3)$	/	$\sqrt{14}$	/	/
	$\boldsymbol{u}^2=3+(0,\,0,\,\frac{\sqrt{21}}{3},\ldots,\,\frac{\sqrt{21}}{3},\,\sqrt{21})$	/	$4\sqrt{21}+\frac{8\sqrt{21}}{3}(N-3)$	/	$\sqrt{14}$	/	/
E	$\boldsymbol{u}^1=(1,\,7,\,7,\ldots,\,7,\,8,\,12)$	/	/	12	/	6	/
	$\boldsymbol{u}^2=(5,\,10,\,10.5,\ldots,\,10.5,\,13,\,14)$	/	/	12	/	5.5	/

(d) By (b), it suffices to show that (iv) and (vi) are equivalent when $N=3$. First, we claim that if $\overline{u}\in[u_2,u_3]$, then $u_3-\overline{u}\geq\overline{u}-u_1$. Indeed, since $\overline{u}\geq u_2$, we have $u_1+u_3\geq 2u_2$. Adding $2(u_1+u_3)$ on both sides of the inequality results in $u_1+u_3\geq 2\overline{u}$, implying that $u_3-\overline{u}\geq\overline{u}-u_1$. Therefore, when $\overline{u}\in[u_2,u_3]$, we have

$$\begin{aligned}\sum_{i=1}^{3}|u_i-\overline{u}|&=(\overline{u}-u_1)+(\overline{u}-u_2)+(u_3-\overline{u})=\frac{1}{3}(u_1+u_2+u_3)-u_1-u_2+u_3\\&=\frac{2}{3}(2u_3-u_1-u_2)=2(u_3-\overline{u})=2\max_{i=1,2,3}|u_i-\overline{u}|.\end{aligned}$$

Similarly, in the case when $\overline{u}\in[u_1,u_2]$, we have $\overline{u}-u_1\geq u_3-\overline{u}$. Thus,

$$\begin{aligned}\sum_{i=1}^{3}|u_i-\overline{u}|&=(\overline{u}-u_1)+(u_2-\overline{u})+(u_3-\overline{u})=-\frac{1}{3}(u_1+u_2+u_3)-u_1+u_2+u_3\\&=\frac{2}{3}(-2u_1+u_2+u_3)=2(\overline{u}-u_1)=2\max_{i=1,2,3}|u_i-\overline{u}|.\end{aligned}$$

This proves the equivalence between (iv) and (vi).

Finally, we show that the remaining pairs of fairness measures are not equivalent. To prove two measures ϕ and $\widetilde{\phi}$ are not equivalent, it suffices to find vectors $\boldsymbol{u}^1$ and $\boldsymbol{u}^2$ such that $\phi(\boldsymbol{u}^1)=\phi(\boldsymbol{u}^2)$ but $\widetilde{\phi}(\boldsymbol{u}^1)\neq\widetilde{\phi}(\boldsymbol{u}^2)$. In Table EC.1, we provide examples showing that the remaining pairs of fairness measures are not equivalent. Specifically, example A shows that the pairs {(iv,i), (iv,ii), (iv,v), (iv,viii), (vi,i), (vi,ii), (vi,v), (vi,viii)} are not equivalent; example B and B' shows that the pairs {(viii,i), (viii,ii), (viii,v)} are not equivalent; example C shows that the pair (i, ii) is not equivalent when $N>3$ and (i,v) is not equivalent for all N; example D shows that the pair (ii, v) is not equivalent; example E shows that the pair (iv, vi) is not equivalent when $N>3$. Note that in example B, fairness measure (ii) at $\boldsymbol{u}^1$ and $\boldsymbol{u}^2$ are equal when $N=5$. Example B' shows that the pair (viii,ii) is not equivalent even when $N=5$. □

Tables EC.2–EC.3 summarize the equivalence of the fairness measures shown in Table 1. The two groups of equivalent fairness measures proved in Proposition EC.1 are highlighted in red and blue with 'Equiv.' representing equivalence in the tables. If a given pair of fairness measures is not equivalent, one of the corresponding counterexamples from A to E is stated (see Table EC.1).

Table EC.2 Equivalence of fairness measures or counterexamples when $N=3$

$N=3$	i	ii	iii	iv	v	vi	vii	viii
i	/	Equiv.	Equiv.	A	C	A	A	B
ii		/	Equiv.	A	C/D	A	A	B
iii			/	A	C	A	A	B
iv				/	A	Equiv.	Equiv.	A
v					/	A	A	B
vi						/	Equiv.	A
vii							/	A
viii								/

Table EC.3 Equivalence of fairness measures or counterexamples when $N>3$

$N>3$	i	ii	iii	iv	v	vi	vii	viii
i	/	C	Equiv.	A	C	A	A	B
ii		/	C	A	D	A	A	B, B'
iii			/	A	C	A	A	B
iv				/	A	E	E	A
v					/	A	A	B
vi						/	Equiv.	A
vii							/	A
viii								/

EC.1.2. Axioms

In this section, we show that the fairness measures in Table 1 are convex fairness measures (see Section 5).

PROPOSITION EC.2. *The measures (i)–(viii) in Table 1 are convex fairness measures, i.e., satisfying Axioms C, N, S, SCV, TI, PH, and CV.*

Proof. It is easy to verify that measures (i)–(viii) satisfy Axioms C, N, S, TI, and PH. Next, note that if ϕ is convex and symmetric, then ϕ is Schur convex (Marshall et al. 2011). Thus, it suffices to show that measures (i)–(viii) are convex. In the following, we assume that $\{\boldsymbol{u}^1,\boldsymbol{u}^2\}\subseteq\mathbb{R}^N$ and $\lambda\in[0,1]$. For measure (i), note that u_i is a linear function in $\boldsymbol{u}$. Since a maximum (resp. minimum) of linear functions is convex (resp. concave), measure (i) is also convex. For measure (ii), we have

$$\begin{aligned}\phi(\lambda\boldsymbol{u}^1+(1-\lambda)\boldsymbol{u}^2)&=\sum_{i=1}^N\sum_{j=1}^N\Big|\big[\lambda u_i^1+(1-\lambda)u_i^2\big]-\big[\lambda u_j^1+(1-\lambda)u_j^2\big]\Big|\\&\le\lambda\sum_{i=1}^N\sum_{j=1}^N|u_i^1-u_j^1|+(1-\lambda)\sum_{i=1}^N\sum_{j=1}^N|u_i^2-u_j^2|\\&=\lambda\phi(\boldsymbol{u}^1)+(1-\lambda)\phi(\boldsymbol{u}^2).\end{aligned}$$

For measure (iii), following a similar argument in (ii), we have

$$\begin{aligned}\phi(\lambda\boldsymbol{u}^1+(1-\lambda)\boldsymbol{u}^2)&=\max_{i\in[N]}\max_{j\in[N]}\Big|\big[\lambda u_i^1+(1-\lambda)u_i^2\big]-\big[\lambda u_j^1+(1-\lambda)u_j^2\big]\Big|\\&\le\max_{i\in[N]}\max_{j\in[N]}\Big\{\lambda|u_i^1-u_j^1|+(1-\lambda)|u_i^2-u_j^2|\Big\}\\&\le\lambda\max_{i\in[N]}\max_{j\in[N]}|u_i^1-u_j^1|+(1-\lambda)\max_{i\in[N]}\max_{j\in[N]}|u_i^2-u_j^2|\\&=\lambda\phi(\boldsymbol{u}^1)+(1-\lambda)\phi(\boldsymbol{u}^2).\end{aligned}$$

For measure (iv), note that $\overline{u} = \lambda\overline{u}^1 + (1-\lambda)\overline{u}^2$, where $\overline{u}^k = (1/N)\sum_{i=1}^N u_i^k$ for $k \in \{1,2\}$. Convexity follows from a similar argument for measure (ii). For measure (v), we have

$$\begin{aligned}\phi(\boldsymbol{u}) = \sqrt{\sum_{i=1}^N (u_i - \overline{u})^2} = \sqrt{\sum_{i=1}^N \left(u_i - \frac{1}{N}\mathbf{1}^\top \boldsymbol{u}\right)^2} &= \left\|\boldsymbol{u} - \frac{1}{N}\mathbf{1}\mathbf{1}^\top \boldsymbol{u}\right\|_2 \\ &= \left\|\left(I - \frac{1}{N}\mathbf{1}\mathbf{1}^\top\right)\boldsymbol{u}\right\|_2 =: \|A\boldsymbol{u}\|_2,\end{aligned}$$

where $I \in \mathbb{R}^{N\times N}$ is the identity matrix. Since ℓ_2 norm is convex, it follows that ϕ is also convex (Bertsekas 2015). For measure (vi), one can easily verify its convexity by following the same logic used to verify the convexity of (iv). Similarly, one can verify the convexity of (vii) and (viii) by following a similar argument as in measures (ii) and (iii). $\square$

EC.2. Examples Related to Order-Based Fairness Measures

EXAMPLE EC.1 (FAIR RESOURCE ALLOCATION). Consider the problem of allocating R resources fairly to N individuals, where we use $i \in [N]$ to denote each individual. Let x_i be the number of resources allocated to i. The impact on i is measured as $u_i = a_i x_i$, where a_i may represent the efficiency per unit resource allocated to i. Moreover, there is a limit $K \le R$ on the number of resources allocated to each individual. Let $N = 6$ and $a_i = i$, for all $i \in [6]$. Suppose we use the order-based fairness measure $\nu_{\boldsymbol{w}}(\boldsymbol{u})$ to ensure fair allocations with $w_i = 2(2i-7)$, for $i \in [6]$. Then, our fair resource allocation optimization problem can be stated as

$$\underset{\boldsymbol{x},\boldsymbol{u}}{\text{minimize}} \quad -10u_{(1)} - 6u_{(2)} - 2u_{(3)} + 2u_{(4)} + 6u_{(5)} + 10u_{(6)} \tag{EC.1a}$$

$$\text{subject to} \quad u_1 = x_1,\, u_2 = 2x_2,\, \ldots, u_6 = 6x_6,\, \textstyle\sum_{i=1}^6 x_i = R,\, 0 \le x_i \le K,\, \forall i \in [6]. \tag{EC.1b}$$

Clearly, the optimization problem (EC.1) prioritizes allocating resources to less advantaged individuals (i.e., those with lower efficiency). To illustrate, we solve (EC.1) numerically (see Section 6) with $R = 25$. Figure EC.1 shows the optimal allocation decisions with different values of K. It is clear that more resources are allocated to less advantaged individuals. Even when K decreases from 10 to 7, i.e, when a smaller upper bound is imposed on x_i, more resources are allocated to the less advantaged individuals with a priority to individual 2, followed by 3 to 6. (Note that the resources allocated to individual 1 decrease because of the decrease in the imposed upper bound K on x_i).

EXAMPLE EC.2 (RANGE). Since $\max_{i\in[N]} u_i - \min_{i\in[N]} = -u_{(1)} + u_{(N)}$, it follows that the range is an order-based fairness measure with weight vector $\boldsymbol{w}' := (-1, 0, \ldots, 0, 1)^\top \in \mathbb{R}^N$. By Theorem 2, we have that Axioms N, SCV, PH, and PA with $\boldsymbol{w} = \boldsymbol{w}'$ characterize the range. Specifically, Axiom PA with $\boldsymbol{w} = \boldsymbol{w}'$ implies (a) there exists constant $C > 0$ such that $\phi(\boldsymbol{u} + \varepsilon \boldsymbol{e}_1) = \phi(\boldsymbol{u}) - \varepsilon C$ for all $\varepsilon \in [0, u_2 - u_1]$ and

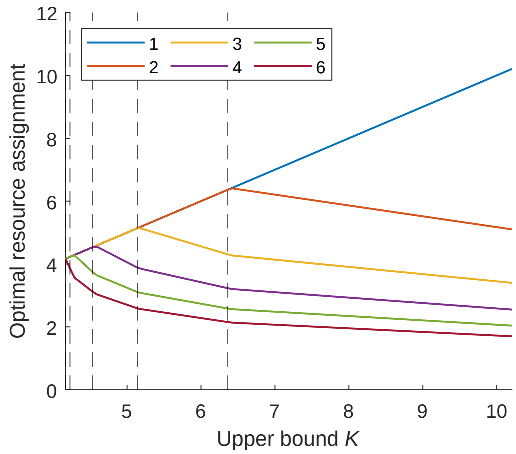

Figure EC.1 Optimal allocation decisions (each line corresponds to an individual)

$\phi(\boldsymbol{u}+\varepsilon\boldsymbol{e}_N)=\phi(\boldsymbol{u})+\varepsilon C$ for all $\varepsilon\geq 0$; (b) for all $j\in[2,N-1]$, $\phi(\boldsymbol{u}+\varepsilon\boldsymbol{e}_j)=\phi(\boldsymbol{u})$ for all $\varepsilon\in[0,u_{j+1}-u_j]$. This indicates that the fairness measure's value decreases (resp. increases) by a constant if the smallest (resp. largest) entry of $\boldsymbol{u}$ increases by a small amount ε, and it remains unchanged otherwise. This reflects the nature of the range that its value depends only on the value of the smallest ($u_{(1)}$) and largest ($u_{(N)}$) entries of $\boldsymbol{u}$.

EXAMPLE EC.3 (GINI DEVIATION). We can rewrite Gini deviation as $\sum_{i=1}^N\sum_{j=1}^N|u_i-u_j|=\sum_{i=1}^N 2(2i-1-N)u_{(i)}=\sum_{i=1}^N w_i' u_{(i)}$, where $w_i'=2(2i-1-N)$ (Mesa et al. 2003). It is easy to verify that $\boldsymbol{w}'\in\mathcal{W}$ (see Definition 1), i.e., Gini deviation is an order-based fairness measure with weight vector $\boldsymbol{w}'$. By Theorem 2, we have that Axioms N, SCV, PH, and PA with $\boldsymbol{w}=\boldsymbol{w}'$ characterize Gini deviation. Specifically, Axiom PA with $\boldsymbol{w}=\boldsymbol{w}'$ implies that there exists constant $C>0$ such that $\phi(\boldsymbol{u}+\varepsilon\boldsymbol{e}_j)=\phi(\boldsymbol{u})+\varepsilon C(2j-N-1)$ for any $\boldsymbol{u}\in\mathbb{R}^N_{\uparrow}$, $j\in[N]$, and $\varepsilon\in[0,u_{j+1}-u_j]$. Thus, if we increase the jth smallest entry $u_{(j)}$ by a small amount ε, the change in fairness measure's value is proportional to ε and scales linearly in j. In particular, if we increase $u_{(j)}$ for any $j\in[\lfloor(N-1)/2\rfloor]$, the fairness measure's value decreases since $2j-N+1<0$, and the decrease is the most pronounced when we increase the smallest entry ($u_{(1)}$).

EC.3. Comparison between the Class of Order-Based Fairness Measures and OWA Operators

Our proposed class of order-based fairness measures differs from OWA operators in the following aspects. First, the OWA operator (and order-median function) takes the form $F_{\widetilde{\boldsymbol{w}}}(\boldsymbol{u})=\sum_{i=1}^N\widetilde{w}_i u_{(i)}$, where the weight vector $\widetilde{\boldsymbol{w}}\in\mathbb{R}^N$ satisfies $\widetilde{w}_i\in[0,1]$ and $\sum_{i=1}^N\widetilde{w}_i=1$. In contrast, the weight vector $\boldsymbol{w}\in\mathbb{R}^N_{\uparrow}$ in our order-based fairness measure satisfies $\sum_{i=1}^N w_i=0$ with $w_1<0$ and $w_N>0$ (see Definition 1). Second, as detailed in Yager (1988), the OWA operator is a way to aggregate different values of u_i, which was designed for multi-criteria decision-making problems and not for measuring unfairness or inequality.

Indeed, if $\widetilde{\boldsymbol{w}} \in \mathbb{R}^N_\uparrow$, which is a common assumption in the literature (see, e.g., Blanco et al. 2016, Nickel and Puerto 2006, Rodríguez-Chía et al. 2000), then we can write $F_{\widetilde{\boldsymbol{w}}}$ as $F_{\widetilde{\boldsymbol{w}}}(\boldsymbol{u}) = f(\boldsymbol{u}) + \nu_{\boldsymbol{w}}(\boldsymbol{u})$, where $f(\boldsymbol{u}) = N^{-1}\sum_{j=1}^N u_j$ and $\boldsymbol{w} = \widetilde{\boldsymbol{w}} - N^{-1}\mathbf{1}$. Here, $\nu_{\boldsymbol{w}}$ is an order-based fairness measure when $\widetilde{\boldsymbol{w}} \neq N^{-1}\mathbf{1}$. Thus, $F_{\boldsymbol{w}}$ (OWA) can be viewed as a special case of the general problem defined in (2) with a specific choice of the inefficiency measure $f(\boldsymbol{u}) = N^{-1}\sum_{j=1}^N u_j$ (i.e., mean of $\boldsymbol{u}$) and the fairness measure $\nu_{\boldsymbol{w}}(\boldsymbol{u})$. In contrast, our framework allows adopting any classical inefficiency measure $f(\boldsymbol{u})$ in (2).

EC.4. Relative Convex Fairness Measures

In this section, building on our analyses of the convex fairness measure proposed in Section 5, we introduce its relative counterpart and study its properties. We relegate all proofs to EC.5. Note that while absolute measures are characterized by translation invariance (Axiom TI), relative measures are characterized by the following scale invariance axiom (Chakravarty 1999, Mussard and Mornet 2019).

AXIOM SI (SCALE INVARIANCE). $\phi(\alpha\boldsymbol{u}) = \phi(\boldsymbol{u})$ for any $\alpha > 0$.

Axiom SI guarantees that if all u_i are multiplied by a positive scalar, then the value of the *relative* fairness measure will not change (i.e., the inequality across the population would not change). This implies that a change in the measurement unit of the vector $\boldsymbol{u}$ (e.g., from US dollar to British pound if $\boldsymbol{u}$ represents costs) will not affect the fairness measure, i.e., the relative fairness measure is unitless.

We are now ready to introduce the relative counterpart of the class of convex fairness measures, which we call relative convex fairness measures.

DEFINITION EC.1. Let $\nu : \mathbb{R}^N_+ \to \mathbb{R}$ be a convex fairness measure and $g : \mathbb{R}^N_+ \to \mathbb{R}$ be a continuous function. A fairness measure $\rho : \mathbb{R}^N_+ \to \mathbb{R}$ is a relative convex fairness measure if

$$\rho(\boldsymbol{u}) = \frac{\nu(\boldsymbol{u})}{g(\boldsymbol{u})} \tag{EC.2}$$

and ρ satisfies Axioms S, SCV, SI, and the following Axiom NR. Here, we adopt the conventions $0/0 = 0$ and $a/0 = \infty$ for any $a > 0$.

AXIOM NR (NORMALIZATION-RELATIVE). $\phi(\boldsymbol{u}) \in [0,1]$ for any $\boldsymbol{u} \in \mathbb{R}^N_+$ and $\phi(\boldsymbol{u}) = 0$ if and only if $\boldsymbol{u} = \alpha\mathbf{1}$ for some $\alpha \in \mathbb{R}_+$.

We make a few remarks in order. First, we consider non-negative $\boldsymbol{u}$, which is common in many practical applications and consistent with prior studies (e.g., $\boldsymbol{u}$ could be patient waiting time, distances from demand nodes to open facilities, income, etc; see Ahmadi-Javid et al. 2017, Chakravarty 1999, Marynissen and Demeulemeester 2019, Mussard and Mornet 2019, Pinedo 2016). Second, Axiom NR is the relative counterpart of the normalization axiom discussed in Section 3. Specifically, we require that the relative convex fairness ρ is normalized such that $\rho(\boldsymbol{u}) \in [0,1]$ for any $\boldsymbol{u} \in \mathbb{R}^N_+$, which is consistent with the literature on relative fairness measures (Donaldson and Weymark 1980, Mehran 1976). In particular, with a suitable choice of g, we have that $\rho = 0$ represents perfect equality while $\rho = 1$ represents perfect inequality (see Proposition EC.3 and Corollary EC.1). Finally, since both ν and g are continuous, ρ is continuous on $\{\boldsymbol{u} \in \mathbb{R}_+^N \mid g(\boldsymbol{u}) \neq 0\}$ by definition in (EC.2). For example, if $g(\boldsymbol{u}) = C\mathbf{1}^\top \boldsymbol{u}$ for some constant $C > 0$, then $\rho(\boldsymbol{u})$ is continuous on $\mathbb{R}_+^N$ except at $\boldsymbol{u} = \mathbf{0}$ (see Theorem EC.2).

It is clear from the definition of ρ in (EC.2) that one should carefully choose the normalization function g such that ρ satisfies the desired set of axioms (i.e., NR, S, SCV, and SI). Hence, we next derive conditions and provide guidelines for choosing the normalization function g. First, in Theorem EC.1, we provide conditions on g such that ρ satisfies Axioms NR, S, and SI.

THEOREM EC.1. *Let $\nu : \mathbb{R}_+^N \to \mathbb{R}$ be a convex fairness measure and $g : \mathbb{R}_+^N \to \mathbb{R}$ be a continuous function. Then, the relative convex fairness measure $\rho = \nu/g$ satisfies Axioms NR, S, and SI if and only if g is symmetric and positive homogeneous with $g \geq \nu$.*

Theorem EC.1 provides sufficient and necessary conditions on g for which the relative convex fairness measure ρ satisfies Axioms NR, S, and SI. Specifically, g belongs to the class of continuous and positive homogeneous functions with $g \geq \nu$. Note that although the convex fairness measure ν is Schur convex, the relative counterpart might not. Thus, in addition to Theorem EC.1, we need to impose conditions on the normalization function g to ensure that ρ is Schur convex. Theorem EC.2 provides a sufficient condition on g such that ρ is Schur convex, and hence, satisfies Definition EC.1.

THEOREM EC.2. *Let $\nu : \mathbb{R}_+^N \to \mathbb{R}$ be a convex fairness measure and $g : \mathbb{R}_+^N \to \mathbb{R}$ be a continuous function. If $g(\boldsymbol{u}) = CN\overline{u}$ with $\overline{u} = N^{-1}\sum_{i=1}^N u_i$ for any $C \geq \sup_{\boldsymbol{w}\in\mathcal{W}_\nu} \|\boldsymbol{w}_+\|_\infty$, where $\boldsymbol{w}_+ = \max\{\boldsymbol{w}, \mathbf{0}\}$ with the maximum taken component-wisely, then the relative convex fairness measure $\rho = \nu/g$ satisfies Axioms NR, S, SCV, and SI.*

Let us provide an intuitive justification for choosing $g(\boldsymbol{u}) = CN\overline{u}$ as suggested by Theorem EC.2. Let $\boldsymbol{u} \in \mathbb{R}_+^N$ with $\mathbf{1}^\top \boldsymbol{u} = \gamma > 0$, i.e., the total impact on the N subjects is γ. Suppose that we can re-distribute the γ units of impact among the subjects. Then, $\boldsymbol{u}_{\min} = \gamma/N \cdot \mathbf{1}$ minimizes ν with $\nu_{\min} = \nu(\boldsymbol{u}_{\min}) = 0$, and $\boldsymbol{u}_{\max} = (0, \ldots, 0, \gamma)^\top$ maximizes ν with $\nu_{\max} = \nu(\boldsymbol{u}_{\max}) = \gamma \sup_{\boldsymbol{w}\in\mathcal{W}_\nu} \|\boldsymbol{w}_+\|_\infty$. Thus, an intuitive way to normalize $\nu(\boldsymbol{u})$ such that $0 \leq \rho(\boldsymbol{u}) \leq 1$ is as follows:

$$\frac{\nu(\boldsymbol{u}) - \nu_{\min}}{\nu_{\max} - \nu_{\min}} = \frac{\nu(\boldsymbol{u})}{\gamma \sup_{\boldsymbol{w}\in\mathcal{W}_\nu} \|\boldsymbol{w}_+\|_\infty} = \frac{\nu(\boldsymbol{u})}{\mathbf{1}^\top \boldsymbol{u} \cdot \sup_{\boldsymbol{w}\in\mathcal{W}_\nu} \|\boldsymbol{w}_+\|_\infty} = \frac{\nu(\boldsymbol{u})}{(N \sup_{\boldsymbol{w}\in\mathcal{W}_\nu} \|\boldsymbol{w}_+\|_\infty) \cdot \overline{u}}. \tag{EC.3}$$

The denominator in (EC.3) takes the same form of g suggested in Theorem EC.2. Indeed, it is also common in the existing literature to normalize a fairness measure by a function of the mean $\overline{u}$ of $\boldsymbol{u}$ (Chakravarty 1999, Mussard and Mornet 2019, Zheng 2007).

Note that there are no general necessary conditions on g for the relative convex fairness measure ρ to be Schur convex. In Example EC.4, we demonstrate that ρ can be Schur convex even for some non-convex and non-concave g.

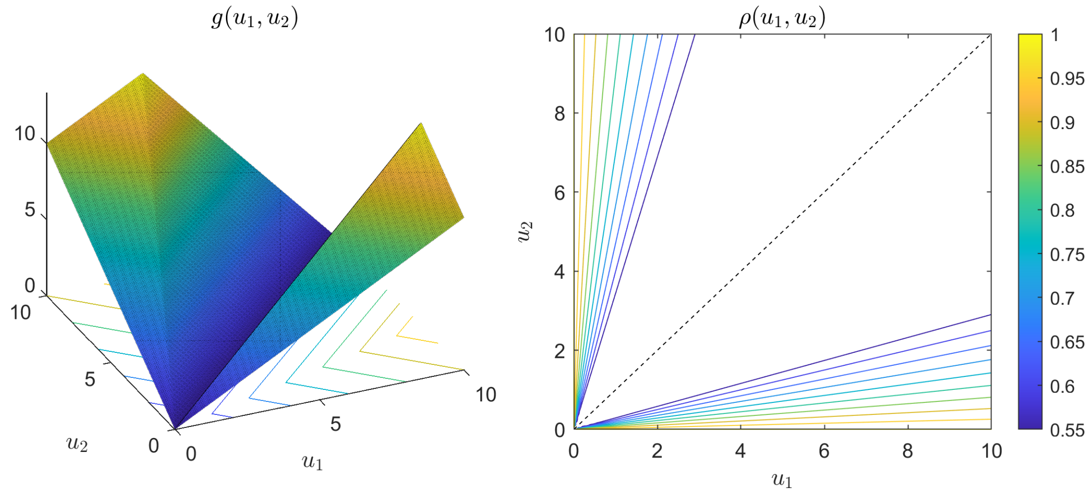

Figure EC.2 Example of the normalization function $g(u_1, u_2) = 2\min\{|u_1 - u_2|, (u_1 + u_2)/2\}$ and the relative convex fairness measure $\rho(u_1, u_2) = |u_1 - u_2|/g(u_1, u_2)$.

EXAMPLE EC.4. Consider the two-dimensional inequality measure $\nu(u_1, u_2) = |u_1 - u_2|$ and the normalization function $g(u_1, u_2) = 2\min\{|u_1 - u_2|, (u_1 + u_2)/2\}$. By definition, g is symmetric and positive homogeneous with $g \geq \nu$, and hence, ρ is symmetric and scale invariant by Theorem EC.1. Note that g is neither convex nor concave, which can be observed in Figure EC.2 (the left plot). However, we can show that the relative convex fairness measure $\rho(u_1, u_2) = \nu(u_1, u_2)/g(u_1, u_2)$ is Schur convex. Indeed, it is straightforward to show that the level sets of ρ, denoted as $\mathcal{L}_\beta = \{(u_1, u_2) \in \mathbb{R}^2_+ \mid \rho(u_1, u_2) \leq \beta\}$ for all $\beta \geq 0$, are convex. Specifically,

$$\mathcal{L}_\beta = \begin{cases} \left\{(u_1, u_2) \in \mathbb{R}^2_+ \mid u_1 = u_2\right\}, & \text{if } \beta \in [0, 1/2), \\ \left\{(u_1, u_2) \in \mathbb{R}^2_+ \,\middle|\, \dfrac{1-\beta}{1+\beta} u_1 \leq u_2 \leq \dfrac{1+\beta}{1-\beta} u_2 \right\}, & \text{if } \beta \in [1/2, 1), \\ \mathbb{R}^2_+, & \text{if } \beta \in [1, \infty). \end{cases}$$

In Figure EC.2 (the right plot), we show the contours of ρ, where the black dotted line represents the level set with $\beta \in [0, 1/2)$. This shows that ρ is quasi-convex, which implies that ρ is Schur convex (see Chapter 3 of Marshall et al. 2011).

Next, in Theorem EC.3, we show that if we restrict the normalization function g to be a function of the mean $\overline{u}$, then $g(\boldsymbol{u})$ can only take the following form $g(\boldsymbol{u}) = CN\overline{u}$, i.e., a necessary condition on g for ρ to be a relative convex fairness measure satisfying Definition EC.1.

THEOREM EC.3. *Let $\nu : \mathbb{R}^N_+ \to \mathbb{R}$ be a convex fairness measure and $g : \mathbb{R}^N_+ \to \mathbb{R}$ be a continuous function. Suppose that $g(\boldsymbol{u}) = \tilde{g}(\overline{u})$ is a function of mean $\overline{u} = N^{-1}\sum_{i=1}^N u_i$ of $\boldsymbol{u}$. Then, the relative convex fairness measure $\rho = \nu/g$ satisfies Axioms NR, S, SCV, and SI if and only if $\tilde{g}(\overline{u}) = CN\overline{u}$ for any $C \geq \sup_{\boldsymbol{w} \in \mathcal{W}_\nu} \|\boldsymbol{w}_+\|_\infty$.*

Theorem EC.3 shows that if the normalization function g is a function of the mean $\overline{u}$, then a linear function in $\overline{u}$ is the only possible choice such that ρ satisfies Axioms NR, S, SCV, and SI. We can also see from (EC.3) that $g(\boldsymbol{u}) = CN\overline{u}$ with $C = \sup_{\boldsymbol{w}\in\mathcal{W}_\nu} \|\boldsymbol{w}_+\|_\infty$ is an intuitive choice.

Observe from (EC.3) that $\rho(\boldsymbol{u}) = 0$ when the convex fairness measure attains its minimum value of zero, i.e., $\nu(\boldsymbol{u}) = \nu(\boldsymbol{u}_{\text{min}}) = 0$, where $\boldsymbol{u}_{\text{min}} \in \arg\min\{\nu(\boldsymbol{u}) \mid \mathbf{1}^\top \boldsymbol{u} = \gamma\}$ for any $\gamma > 0$. Thus, like other relative fairness measures in the literature, $\rho(\boldsymbol{u}) = 0$ implies perfect equality (Axiom NR). On the other hand, $\rho(\boldsymbol{u})$ attains its maximum value of one when the convex fairness measure attains its maximum value, i.e., $\nu(\boldsymbol{u}) = \nu(\boldsymbol{u}_{\text{max}})$, where $\boldsymbol{u}_{\text{max}} \in \arg\max\{\nu(\boldsymbol{u}) \mid \mathbf{1}^\top \boldsymbol{u} = \gamma\}$ for any $\gamma > 0$. It follows that $\rho(\boldsymbol{u}) = 1$ indicates perfect inequality. We formalize the latter observation in Proposition EC.3.

PROPOSITION EC.3. *Let ν be a convex fairness measure and $g(\boldsymbol{u}) = \sup_{\boldsymbol{w}\in\mathcal{W}_\nu} \|\boldsymbol{w}_+\|_\infty N\overline{u}$ with $\overline{u} = N^{-1}\sum_{i=1}^N u_i$. Then, the relative convex fairness measure $\rho(\boldsymbol{u}) = \nu(\boldsymbol{u})/g(\boldsymbol{u}) = 1$ if and only if $\boldsymbol{u} \in \mathcal{U}^\star = \bigcup_{\gamma>0} \mathcal{U}^\star_\gamma$, where $\mathcal{U}^\star_\gamma = \arg\max_{\boldsymbol{u}'\in\mathbb{R}^N_+}\{\nu(\boldsymbol{u}') \mid \mathbf{1}^\top \boldsymbol{u}' = \gamma\}$.*

In Corollary EC.1, we show that under some additional mild assumption on ν, the relative convex fairness measure achieves its maximum $\rho(\boldsymbol{u}) = 1$ if and only if $\boldsymbol{u}$ takes the form $P(0,\ldots,0,\gamma)^\top \in \mathbb{R}^N$ for any $\gamma > 0$ and permutation matrix P.

COROLLARY EC.1. *Let ν be a convex fairness measure and $g(\boldsymbol{u}) = \sup_{\boldsymbol{w}\in\mathcal{W}_\nu} \|\boldsymbol{w}_+\|_\infty N\overline{u}$ with $\overline{u} = N^{-1}\sum_{i=1}^N u_i$. Assume that $\nu(0,\ldots,0,0,\gamma) > \nu(0,\ldots,0,\gamma/2,\gamma/2)$ for any $\gamma > 0$. Then, $\rho(\boldsymbol{u}) = \nu(\boldsymbol{u})/g(\boldsymbol{u}) = 1$ if and only if $\boldsymbol{u} = P(0,\ldots,0,\gamma)^\top$ for any $\gamma > 0$ and permutation matrix P.*

REMARK EC.1. Note that the constant $w_{\text{max}} = \sup_{\boldsymbol{w}\in\mathcal{W}_\nu} \|\boldsymbol{w}_+\|_\infty$ can be computed directly without solving the supremum problem. Specifically, since $\nu(\boldsymbol{u}) = \sup_{\boldsymbol{w}\in\mathcal{W}_\nu} \nu_{\boldsymbol{w}}(\boldsymbol{u})$ (see Theorem 3), we have $\nu(0,\ldots,0,\gamma) = \gamma w_{\text{max}}$, which implies $w_{\text{max}} = \nu(0,\ldots,0,\gamma)/\gamma$ for any $\gamma > 0$.

We close this section with Example EC.5, where we derive the relative counterparts of the deviation-based fairness measures in Table 1 and make connections to existing relative fairness measures.

EXAMPLE EC.5 (RELATIVE COUNTERPARTS OF THE FAIRNESS MEASURES IN TABLE 1).

(i) Recall that measure (i) is order-based with $\boldsymbol{w} = (-1, 0, \ldots, 0, 1)$. Thus, we have $w_{\text{max}} = 1$, and hence,

$$\rho(\boldsymbol{u}) = \frac{\nu(\boldsymbol{u})}{N\overline{u}} = \frac{u_{(N)} - u_{(1)}}{N\overline{u}},$$

which is equivalent to the relative range (Cowell 2011). Note that (iii) is equivalent to (i) (see EC.1). Thus, the relative counterpart of (iii) takes the same form as (i).

(ii) Recall that measure (ii) is order-based with $w_i = 2(2i - N - 1)$ for $i \in [N]$. Thus, we have $w_{\text{max}} = 2(N-1)$, and hence,

$$\rho(\boldsymbol{u}) = \frac{\nu(\boldsymbol{u})}{2N(N-1)\overline{u}} = \frac{\sum_{i\in[N]}\sum_{j\in[N]} |u_i - u_j|}{2N(N-1)\overline{u}},$$

which is equivalent to the Gini index (Gini 1912).

(iv) For measure (iv), we have $w_{\max} = \nu(0,\ldots,0,\gamma)/\gamma = 2[1-(1/N)]$, and hence,

$$\rho(\boldsymbol{u}) = \frac{\nu(\boldsymbol{u})}{2(N-1)\overline{u}} = \frac{\sum_{i\in[N]}|u_i-\overline{u}|}{2(N-1)\overline{u}},$$

which is equivalent to the relative mean absolute deviation, also known as the Hoover index (Hoover 1936).

(v) For measure (v), we have $w_{\max} = \nu(0,\ldots,0,\gamma)/\gamma = \sqrt{1-(1/N)}$, and hence,

$$\rho(\boldsymbol{u}) = \frac{\nu(\boldsymbol{u})}{\sqrt{N}\sqrt{N-1}\overline{u}} = \frac{\sqrt{(N-1)^{-1}\sum_{i\in[N]}(u_i-\overline{u})^2}}{\sqrt{N}\overline{u}},$$

which is equivalent to the relative standard deviation or the coefficient of variation (Cowell 2011).

(vi) For measure (vi), we have $w_{\max} = \nu(0,\ldots,0,\gamma)/\gamma = 1-(1/N)$, and hence,

$$\rho(\boldsymbol{u}) = \frac{\nu(\boldsymbol{u})}{(N-1)\overline{u}} = \frac{\max_{i\in[N]}|u_i-\overline{u}|}{(N-1)\overline{u}}.$$

Note that measure (vii) is equal to N times measure (vi) (see EC.1). Thus, the relative counterpart of (vii) takes the same form as (vi).

(viii) For measure (viii), we have $w_{\max} = \nu(0,\ldots,0,\gamma)/\gamma = N$, and hence,

$$\rho(\boldsymbol{u}) = \frac{\nu(\boldsymbol{u})}{N^2\overline{u}} = \frac{\sum_{i\in[N]}\max_{j\in[N]}|u_i-u_j|}{N^2\overline{u}}.$$

EC.5. Mathematical Proofs

In this section, we present proofs of the theoretical results in the order of their appearance.

EC.5.1. Proof of Proposition 1

(a) *Axiom C.* First, we claim that the sorting operator $S:\mathbb{R}^N\to\mathbb{R}^N$ that maps a vector $\boldsymbol{u}$ to $\boldsymbol{u}^{\uparrow}\in\mathbb{R}^N_{\uparrow}$ with entries in ascending order is continuous. Consider two vectors $\boldsymbol{u}^1$ and $\boldsymbol{u}^2$ with $\|\boldsymbol{u}^1-\boldsymbol{u}^2\|_\infty = \varepsilon$, i.e., $|u^1_i-u^2_i|\le\varepsilon$ for all $i\in[N]$. We claim that $|u^1_{(i)}-u^2_{(i)}|\le\varepsilon$ for all $i\in[N]$. To show this, suppose, on the contrary, that $|u^1_{(i)}-u^2_{(i)}|>\varepsilon$ for some $i\in[N]$. Consider the following two cases. First, if $u^2_{(i)}<u^1_{(i)}-\varepsilon$, define $\mathcal{J}_i^- = \{\pi_2(1),\ldots,\pi_2(i)\}$, where $\pi_2(k)$ is the index such that $\boldsymbol{u}^2_{\pi_2(k)} = \boldsymbol{u}^2_{(k)}$ for $k\in[N]$. For all $j\in\mathcal{J}_i^-$, we have

$$u^1_j \le u^2_j+\varepsilon \le u^2_{(i)}+\varepsilon < u^1_{(i)}, \tag{EC.4}$$

where the first inequality follows from $\varepsilon = \|\boldsymbol{u}^1-\boldsymbol{u}^2\|_\infty$, and the second inequality follows from $j\in\mathcal{J}_i^-$. This contradicts that $u^1_{(i)}$ is the ith smallest entry in $\boldsymbol{u}^1$. Similarly, if $u^2_{(i)}>u^1_{(i)}+\varepsilon$, define $\mathcal{J}_i^+ = \{\pi_2(i),\ldots,\pi_2(N)\}$. Then, following a similar argument in (EC.4), for all $j\in\mathcal{J}_i^+$, we have

$$u^1_j \ge u^2_j-\varepsilon \ge u^2_{(i)}-\varepsilon > u^1_{(i)},$$

which leads to the contradiction that $u^1_{(i)}$ is the ith smallest entry in $\boldsymbol{u}^1$. Therefore,

$$|\nu_{\boldsymbol{w}}(\boldsymbol{u}^1) - \nu_{\boldsymbol{w}}(\boldsymbol{u}^2)| = \left| \sum_{i=1}^N w_i u^1_{(i)} - \sum_{i=1}^N w_i u^2_{(i)} \right| \le \sum_{i=1}^N |w_i|\,|u^1_{(i)} - u^2_{(i)}| \le \|\boldsymbol{w}\|_1 \|\boldsymbol{u}^1 - \boldsymbol{u}^2\|_\infty.$$

This shows that $\nu_{\boldsymbol{w}}$ is continuous.

(b) *Axiom N*. We first show that $\nu_{\boldsymbol{w}}(\boldsymbol{u}) \ge 0$ for any $\boldsymbol{u} \in \mathbb{R}^N$. Letting $u_{(0)} = 0$, we can write

$$\nu_{\boldsymbol{w}}(\boldsymbol{u}) = \sum_{i=1}^N w_i u_{(i)} = \sum_{i=1}^N w_i \sum_{j=1}^i \left[u_{(j)} - u_{(j-1)}\right] = \sum_{j=1}^N \left(\sum_{i=j}^N w_i \right) \left[u_{(j)} - u_{(j-1)}\right].$$

Since $\sum_{i=1}^N w_i = 0$, we have

$$\nu_{\boldsymbol{w}}(\boldsymbol{u}) = \sum_{j=2}^N \left(\sum_{i=j}^N w_i \right) \left[u_{(j)} - u_{(j-1)}\right]. \tag{EC.5}$$

Since $u_{(j)} - u_{(j-1)} \ge 0$, to show that $\nu_{\boldsymbol{w}}(\boldsymbol{u}) \ge 0$, it suffices to show that $\sum_{i=j}^N w_i > 0$ for all $j \in [2,N]_{\mathbb{Z}}$. We show $\sum_{i=j}^N w_i > 0$ by induction. When $j = 2$, since $\sum_{i=1}^N w_i = 0$ and $w_1 < 0$, we have $\sum_{i=2}^N w_i = \sum_{i=1}^N w_i - w_1 > 0$. Next, suppose that $\sum_{i=j-1}^N w_i > 0$. If $w_j \ge 0$, then it is trivial that $\sum_{i=j}^N w_j > 0$ since $0 \le w_j \le \cdots \le w_N$ and $w_N > 0$. If $w_j \le 0$, then $\sum_{i=j}^N w_i = \sum_{i=j-1}^N w_i - w_{j-1} > 0$ by induction hypothesis. This completes the induction step and shows that $\nu_{\boldsymbol{w}}(\boldsymbol{u}) \ge 0$. Finally, we show that $\nu_{\boldsymbol{w}}(\boldsymbol{u}) = 0$ if and only if $\boldsymbol{u} = \alpha\mathbf{1}$ for some $\alpha \in \mathbb{R}$. If $\boldsymbol{u} = \alpha\mathbf{1}$, then it trivial that $\nu_{\boldsymbol{w}}(\boldsymbol{u}) = \alpha \sum_{i=1}^N w_i = 0$. If $\nu_{\boldsymbol{w}}(\boldsymbol{u}) = 0$, by (EC.5) and $\sum_{i=j}^N w_i > 0$, we must have $u_{(j)} - u_{(j-1)} = 0$ for all $j \in [2,N]_{\mathbb{Z}}$, which in turn implies $u_{(1)} = \cdots = u_{(N)}$.

(c) *Axiom S*. Symmetry follows directly from the definition of $\nu_{\boldsymbol{w}}$ that depends only on the order of $\boldsymbol{u}$.

(d) *Axiom SCV*. Letting $w_0 = 0$, we can write

$$\begin{aligned}
\nu_{\boldsymbol{w}}(\boldsymbol{u}) &= \sum_{i=1}^N w_i u_{(i)} = \sum_{i=1}^N \left[\sum_{j=1}^i (w_j - w_{j-1}) \right] u_{(i)} = \sum_{j=1}^N (w_j - w_{j-1}) \left[\sum_{i=j}^N u_{(i)} \right] \\
&= w_1 \left[\sum_{i=1}^N u_{(i)} \right] + \sum_{j=2}^N (w_j - w_{j-1}) \left[\sum_{i=j}^N u_{(i)} \right].
\end{aligned}$$

By definition, $\boldsymbol{u}^1 \preceq \boldsymbol{u}^2$ implies $\sum_{i=1}^N u^1_i = \sum_{i=1}^N u^2_i$ and $\sum_{i=j}^N u^1_i \le \sum_{i=j}^N u^2_i$ for all $j \in [N]$. Since $w_j - w_{j-1} \ge 0$, we have $\nu_{\boldsymbol{w}}(\boldsymbol{u}^1) \le \nu_{\boldsymbol{w}}(\boldsymbol{u}^2)$.

(e) *Axiom TI*. It is straightforward to verify that

$$\nu_{\boldsymbol{w}}(\boldsymbol{u} + \alpha\mathbf{1}) = \sum_{i=1}^N w_i \left[u_{(i)} + \alpha\right] = \sum_{i=1}^N w_i u_{(i)} + \alpha \sum_{i=1}^N w_i = \nu_{\boldsymbol{w}}(\boldsymbol{u}).$$

(f) *Axiom PH*. It is straightforward to verify that

$$\nu_{\boldsymbol{w}}(\alpha\boldsymbol{u}) = \sum_{i=1}^N w_i \left[\alpha u_{(i)}\right] = \alpha \sum_{i=1}^N w_i u_{(i)} = \alpha \nu_{\boldsymbol{w}}(\boldsymbol{u}).$$

This completes the proof. $\square$

EC.5.2. Proof of Theorem 1

From the rearrangement inequality (Marshall et al. 2011), for any $\pi \in \Pi$, we have

$$\sum_{i=1}^{N} w_{\pi(i)} u_i \le \sum_{i=1}^{N} w_{(i)} u_{(i)} = \sum_{i=1}^{N} w_i u_{(i)} = \nu_{\boldsymbol{w}}(\boldsymbol{u}),$$

where the first equality follows from $w_1 \le \cdots \le w_N$. $\square$

EC.5.3. Proof of Theorem 2

Suppose that ν is an order-based fairness measure with weight $\boldsymbol{w}$. Then, ν satisfies Axioms N, SCV, and PH by Proposition 1. In addition, for any $\boldsymbol{u} \in \mathbb{R}^N_{\uparrow}$, $\varepsilon \in [0, u_{j+1} - u_j]$, and $j \in [N]$, since the order of the entries in $\boldsymbol{u} + \varepsilon \boldsymbol{e}_j$ are the same as that of $\boldsymbol{u}$, we also have $\nu(\boldsymbol{u} + \varepsilon \boldsymbol{e}_j) = \nu(\boldsymbol{u}) + \varepsilon w_j$ by definition of the order-based fairness measure. If follows that ν satisfies Axiom PA.

Now, suppose that ν satisfies Axioms N, SCV, PH, and PA. Recall that Axiom SCV implies Axiom S, i.e., ν is symmetric. Thus, without loss of generality, we focus on the function ν in the space $\mathbb{R}^N_{\uparrow}$. We show that ν is an order-based fairness measure in two steps. First, we show that for any $\boldsymbol{u} \in \mathbb{R}^N_{\uparrow}$,

$$\nu(\boldsymbol{u} + \varepsilon \overline{\boldsymbol{u}}^j) = \nu(\boldsymbol{u}) + \varepsilon \nu(\overline{\boldsymbol{u}}^j), \tag{EC.6}$$

for all $j \in [N]$ and $\varepsilon \ge 0$, where $\overline{\boldsymbol{u}}^j = \sum_{i=j}^N \boldsymbol{e}_i$. To show (EC.6), applying Axiom PA iteratively, we obtain that for all $j \in [N]$ and $\boldsymbol{u} \in \mathbb{R}^N_{\uparrow}$,

$$\begin{aligned}
\nu(\boldsymbol{u} + \varepsilon \overline{\boldsymbol{u}}^j) &= \nu\Big(\big[\boldsymbol{u} + \varepsilon \overline{\boldsymbol{u}}^{j+1}\big] + \varepsilon \boldsymbol{e}_j\Big) \\
&= \nu\Big(\boldsymbol{u} + \varepsilon \overline{\boldsymbol{u}}^{j+1}\Big) + \varepsilon w_j \\
&= \nu\Big(\big[\boldsymbol{u} + \varepsilon \overline{\boldsymbol{u}}^{j+2}\big] + \varepsilon \boldsymbol{e}_{j+1}\Big) + \varepsilon w_j \\
&= \nu\Big(\boldsymbol{u} + \varepsilon \overline{\boldsymbol{u}}^{j+2}\Big) + \varepsilon (w_j + w_{j+1}) \\
&= \cdots \\
&= \nu(\boldsymbol{u}) + \varepsilon \sum_{i=j}^{N} w_j.
\end{aligned} \tag{EC.7}$$

In particular, if we set $\boldsymbol{u} = \boldsymbol{0}$ in (EC.7), this gives $\nu(\varepsilon \overline{\boldsymbol{u}}^j) = \varepsilon \sum_{i=j}^N w_j$ for all $j \in [N]$, which implies that $\nu(\overline{\boldsymbol{u}}^j) = \sum_{i=j}^N w_j$ for all $j \in [N]$ by Axiom PH. Replacing $\sum_{i=j}^N w_j$ in (EC.7) by $\nu(\varepsilon \overline{\boldsymbol{u}}^j)$, we obtain the desired equality in (EC.6).

Second, using (EC.6), we show that Axioms N, SCV, PH, and PA characterize order-based fairness measures. For any $\boldsymbol{u} \in \mathbb{R}^N_{\uparrow}$, we can write $\boldsymbol{u}$ as a linear combination of $\{\overline{\boldsymbol{u}}^j\}_{j=1}^N$:

$$\boldsymbol{u} = u_1 \cdot \overline{\boldsymbol{u}}^1 + (u_2 - u_1) \cdot \overline{\boldsymbol{u}}^2 + \cdots + (u_N - u_{N-1}) \cdot \overline{\boldsymbol{u}}^N = u_1 \cdot \overline{\boldsymbol{u}}^1 + \sum_{j=2}^{N} \Big[(u_j - u_{j-1}) \cdot \overline{\boldsymbol{u}}^j\Big].$$

Applying (EC.6) iteratively, we have $\nu(\boldsymbol{u}) = \nu(u_1 \cdot \overline{\boldsymbol{u}}^1) + \sum_{j=2}^N (u_j - u_{j-1})\nu(\overline{\boldsymbol{u}}^j)$. Note that $\nu(u_1 \cdot \overline{\boldsymbol{u}}^1) = \nu(u_1 \cdot \mathbf{1}) = 0$ by Axiom N. Thus, we have $\nu(\boldsymbol{u}) = \sum_{j=2}^N (u_j - u_{j-1})\nu(\overline{\boldsymbol{u}}^j)$. It follows that we can rewrite the function ν as $\nu(\boldsymbol{u}) = \sum_{j=1}^N w'_j u_j$, where $w'_j = \nu(\overline{\boldsymbol{u}}^j) - \nu(\overline{\boldsymbol{u}}^{j+1})$ for $j \in [N-1]$ and $w'_N = \nu(\overline{\boldsymbol{u}}^N)$. Hence, to show that ν is an order-based fairness measure, it suffices to show that $\boldsymbol{w}' = (w'_1, \dots, w'_N)^\top \in \mathcal{W} = \{\boldsymbol{w} \in \mathbb{R}^N \mid \sum_{i=1}^N w_i = 0,\, w_1 \le \cdots \le w_N,\, w_1 < 0,\, w_N > 0\}$.

(a) First, we show that $\sum_{j=1}^N w'_j = 0$. From the definition of w'_j, we have $\sum_{j=1}^N w'_j = \nu(\overline{\boldsymbol{u}}^N) + \sum_{j=1}^{N-1} [\nu(\overline{\boldsymbol{u}}^j) - \nu(\overline{\boldsymbol{u}}^{j+1})] = \nu(\overline{\boldsymbol{u}}^1) = \nu(\mathbf{1}) = 0$ by Axiom N.

(b) Second, we show that $w'_j \ge w'_{j-1}$ for all $j \in [2, N]$. Indeed, note that $(\overline{\boldsymbol{u}}^{j-1} + \overline{\boldsymbol{u}}^{j+1})/2 \preceq \overline{\boldsymbol{u}}^j$ since the vector $(\overline{\boldsymbol{u}}^{j-1} + \overline{\boldsymbol{u}}^{j+1})/2$ is obtained by transferring $1/2$ from the jth to the $(j-1)$th individual in $\overline{\boldsymbol{u}}^j$, where we let $\overline{\boldsymbol{u}}^{N+1} = \mathbf{0}$. Axiom SCV implies that $\nu\big((\overline{\boldsymbol{u}}^{j-1} + \overline{\boldsymbol{u}}^{j+1})/2\big) \le \nu(\overline{\boldsymbol{u}}^j)$. Also, by (EC.6), we have $\nu\big((\overline{\boldsymbol{u}}^{j-1} + \overline{\boldsymbol{u}}^{j+1})/2\big) = \big[\nu(\overline{\boldsymbol{u}}^{j-1}) + \nu(\overline{\boldsymbol{u}}^{j+1})\big]/2$. Therefore, we obtain $\nu(\overline{\boldsymbol{u}}^{j-1}) - \nu(\overline{\boldsymbol{u}}^j) \le \nu(\overline{\boldsymbol{u}}^j) - \nu(\overline{\boldsymbol{u}}^{j+1})$, and thus, $w'_j \ge w'_{j-1}$.

(c) Third, be definition, we have $w'_1 = \nu(\overline{\boldsymbol{u}}^1) - \nu(\overline{\boldsymbol{u}}^2) = -\nu(\overline{\boldsymbol{u}}^2) < 0$, where the inequality follows from Axiom N. Together with (a) and (b), we must have $w'_N > 0$.

Therefore, we have $\nu(\boldsymbol{u}) = \sum_{j=1}^N w'_j u_j$ is an order-based fairness measure. Finally, note that we have $\nu(\overline{\boldsymbol{u}}^j) = \sum_{i=j}^N w_j$ for all $j \in [N]$ from the first step of the proof, which implies $\boldsymbol{w}' = \boldsymbol{w}$. This completes the proof. $\square$

EC.5.4. Proof of Theorem 3

It is easy to verify that (b) implies (c). We first prove that (c) implies (a). Suppose there exists a compact set $\mathcal{W}_\nu \subseteq \{\boldsymbol{w} \in \mathbb{R}^N_\uparrow \mid \mathbf{1}^\top \boldsymbol{w} = 0\}$ such that $\nu(\boldsymbol{u}) = \sup_{\boldsymbol{w} \in \mathcal{W}_\nu} \nu_{\boldsymbol{w}}(\boldsymbol{u})$. We want to verify that $\nu(\boldsymbol{u})$ is a convex fairness measure satisfying Axioms C, N, S, TI, PH, and CV (Definition 2).

(I) *Axiom C*. From the proof of Proposition 1, for any $\{\boldsymbol{u}^1, \boldsymbol{u}^2\} \subset \mathbb{R}^N$, we have $|\nu_{\boldsymbol{w}}(\boldsymbol{u}^1) - \nu_{\boldsymbol{w}}(\boldsymbol{u}^2)| \le \|\boldsymbol{w}\|_1 \|\boldsymbol{u}^1 - \boldsymbol{u}^2\|_\infty$. Hence,

$$
\begin{aligned}
|\nu(\boldsymbol{u}^1) - \nu(\boldsymbol{u}^2)| &= \left| \sup_{\boldsymbol{w} \in \mathcal{W}} \nu_{\boldsymbol{w}}(\boldsymbol{u}^1) - \sup_{\boldsymbol{w} \in \mathcal{W}} \nu_{\boldsymbol{w}}(\boldsymbol{u}^2) \right| \\
&\le \sup_{\boldsymbol{w} \in \mathcal{W}} |\nu_{\boldsymbol{w}}(\boldsymbol{u}^1) - \nu_{\boldsymbol{w}}(\boldsymbol{u}^2)| \le \sup_{\boldsymbol{w} \in \mathcal{W}} \|\boldsymbol{w}\|_1 \cdot \|\boldsymbol{u}^1 - \boldsymbol{u}^2\|_\infty,
\end{aligned}
$$

where $\sup_{\boldsymbol{w} \in \mathcal{W}} \|\boldsymbol{w}\|_1 < \infty$ since $\mathcal{W}$ is compact. Hence, ν is continuous in $\boldsymbol{u}$.

(II) *Axiom N*. Since $\nu_{\boldsymbol{w}}(\boldsymbol{u}) \ge 0$ for any $\boldsymbol{w} \in \mathcal{W}$ by Proposition 1, we have $\nu(\boldsymbol{u}) \ge 0$. Moreover, note that $\nu(\boldsymbol{u}) = 0$ is equivalent to $\nu_{\boldsymbol{w}}(\boldsymbol{u}) = 0$ for all $\boldsymbol{w} \in \mathcal{W}$. By Proposition 1, $\nu_{\boldsymbol{w}}(\boldsymbol{u}) = 0$ if and only if $\boldsymbol{u} = \alpha \mathbf{1}$ for some $\alpha \in \mathbb{R}$.

(III) *Axiom S*. Symmetry holds since each $\nu_{\boldsymbol{w}}(\boldsymbol{u})$ is symmetric by Proposition 1.

(IV) *Axiom CV*. By Theorem 1, $\nu_{\boldsymbol{w}}$ is a maximum of linear functions. It follows that $\nu_{\boldsymbol{w}}$ is convex. Hence, since ν is a supremum of convex functions, ν is also convex.

(V) *Axiom TI.* By Proposition 1, for any $\alpha \in \mathbb{R}$,

$$\nu(\boldsymbol{u}+\alpha\mathbf{1}) = \sup_{\boldsymbol{w}\in\mathcal{W}} \nu_{\boldsymbol{w}}(\boldsymbol{u}+\alpha\mathbf{1}) = \sup_{\boldsymbol{w}\in\mathcal{W}} \nu_{\boldsymbol{w}}(\boldsymbol{u}) = \nu(\boldsymbol{u}).$$

(VI) *Axiom PH.* By Proposition 1, for any $\alpha > 0$,

$$\nu(\alpha\boldsymbol{u}) = \sup_{\boldsymbol{w}\in\mathcal{W}} \nu_{\boldsymbol{w}}(\alpha\boldsymbol{u}) = \alpha \sup_{\boldsymbol{w}\in\mathcal{W}} \nu_{\boldsymbol{w}}(\boldsymbol{u}) = \alpha\nu(\boldsymbol{u}).$$

Next, we prove that (a) implies (b). Note that by definition of convex fairness measures, ν is proper, continuous, and convex. By Fenchel–Moreau theorem, $\nu(\boldsymbol{u}) = \sup_{\boldsymbol{w}\in\mathbb{R}^N}\left\{\boldsymbol{u}^\top\boldsymbol{w} - \nu^*(\boldsymbol{w})\right\}$, where $\nu^*(\boldsymbol{w}) = \sup_{\boldsymbol{u}\in\mathbb{R}^N}\{\boldsymbol{u}^\top\boldsymbol{w} - \nu(\boldsymbol{u})\}$ is the convex conjugate of ν (Bertsekas 2009). We divide the proof into the following four steps.

- *Step 1.* By Axiom PH, we have for any $\alpha > 0$,

$$\begin{aligned}\nu^*(\boldsymbol{w}) &= \sup_{\boldsymbol{u}\in\mathbb{R}^N}\left\{\boldsymbol{u}^\top\boldsymbol{w} - \nu(\boldsymbol{u})\right\}\\ &= \alpha \sup_{\boldsymbol{u}\in\mathbb{R}^N}\left\{\left(\frac{\boldsymbol{u}}{\alpha}\right)^\top\boldsymbol{w} - \nu\left(\frac{\boldsymbol{u}}{\alpha}\right)\right\} = \alpha \sup_{\boldsymbol{u}'\in\mathbb{R}^N}\left\{(\boldsymbol{u}')^\top\boldsymbol{w} - \nu(\boldsymbol{u}')\right\} = \alpha\nu^*(\boldsymbol{w}),\end{aligned}$$

  where we apply a change of variable $\boldsymbol{u}' = \boldsymbol{u}/\alpha$. Thus, $\nu^*(\boldsymbol{w})$ equals 0 if $\boldsymbol{w}\in\mathrm{dom}(\nu^*)$, and ∞ otherwise. Note that $\mathrm{dom}(\nu^*) = \{\boldsymbol{w}\in\mathbb{R}^N \mid \nu^*(\boldsymbol{w}) \le 0\} = \left\{\boldsymbol{w}\in\mathbb{R}^N \mid \boldsymbol{u}^\top\boldsymbol{w} - \nu(\boldsymbol{u}) \le 0,\, \forall \boldsymbol{u}\in\mathbb{R}^N\right\} = \partial\nu(\mathbf{0})$, where $\nu(\mathbf{0})$ is the subdifferential of ν at $\boldsymbol{u} = \mathbf{0}$. Therefore, the set $\mathrm{dom}(\nu^*)$ is closed, bounded, and convex (see Proposition 5.4.2 of Bertsekas 2009). Thus, we have $\nu(\boldsymbol{u}) = \sup_{\boldsymbol{w}\in\mathrm{dom}(\nu^*)}\left\{\boldsymbol{u}^\top\boldsymbol{w}\right\}$.

- *Step 2.* By Axiom TI, we have for any $\widetilde{\boldsymbol{u}}\in\mathbb{R}^N$,

$$\begin{aligned}\nu^*(\boldsymbol{w}) = \sup_{\boldsymbol{u}\in\mathbb{R}^N}\left\{\boldsymbol{u}^\top\boldsymbol{w} - \nu(\boldsymbol{u})\right\} &\ge \sup_{\alpha\in\mathbb{R}}\left\{(\widetilde{\boldsymbol{u}}+\alpha\mathbf{1})^\top\boldsymbol{w} - \nu(\widetilde{\boldsymbol{u}}+\alpha\mathbf{1})\right\}\\ &= \sup_{\alpha\in\mathbb{R}}\left\{\widetilde{\boldsymbol{u}}^\top\boldsymbol{w} - \nu(\widetilde{\boldsymbol{u}}) + \alpha\mathbf{1}^\top\boldsymbol{w}\right\}.\end{aligned}$$

  Therefore, if $\boldsymbol{w}\in\mathrm{dom}(\nu^*)$, we must have $\mathbf{1}^\top\boldsymbol{w} = 0$.

- *Step 3.* By Axiom S, for any permutation matrix $P\in\mathcal{P}$ (the set of all $N\times N$ permutation matrices),

$$\begin{aligned}\nu^*(P\boldsymbol{w}) &= \sup_{\boldsymbol{u}\in\mathbb{R}^N}\left\{\boldsymbol{u}^\top(P\boldsymbol{w}) - \nu(\boldsymbol{u})\right\} = \sup_{\boldsymbol{u}\in\mathbb{R}^N}\left\{(P^\top\boldsymbol{u})^\top\boldsymbol{w} - \nu(\boldsymbol{u})\right\}\\ &= \sup_{\boldsymbol{u}'\in\mathbb{R}^N}\left\{(\boldsymbol{u}')^\top\boldsymbol{w} - \nu(P\boldsymbol{u}')\right\} = \sup_{\boldsymbol{u}'\in\mathbb{R}^N}\left\{(\boldsymbol{u}')^\top\boldsymbol{w} - \nu(\boldsymbol{u}')\right\} = \nu^*(\boldsymbol{w}),\end{aligned}$$

  where we apply a change of variable $\boldsymbol{u}' = P^\top\boldsymbol{u}$. Thus, ν^* is also symmetric. That is, if $\boldsymbol{w}\in\mathrm{dom}(\nu^*)$, then $P\boldsymbol{w}\in\mathrm{dom}(\nu^*)$ for any $P\in\mathcal{P}$. As a result, $\mathrm{dom}(\nu^*) = \bigcup_{P\in\mathcal{P}}\{\widetilde{\boldsymbol{w}}\in\mathbb{R}^N \mid \widetilde{\boldsymbol{w}} = P\boldsymbol{w},\, \boldsymbol{w}\in\mathcal{W}_\nu\}$ with $\mathcal{W}_\nu = \mathrm{dom}(\nu^*)\cap\mathbb{R}^N_\uparrow$ still being compact and convex. Hence, we can write

$$\nu(\boldsymbol{u}) = \sup_{\boldsymbol{w}\in\mathrm{dom}(\nu^*)}\left\{\boldsymbol{u}^\top\boldsymbol{w}\right\} = \sup_{\boldsymbol{w}\in\mathcal{W}_\nu}\sup_{P\in\mathcal{P}}\left\{\boldsymbol{u}^\top(P\boldsymbol{w})\right\} = \sup_{\boldsymbol{w}\in\mathcal{W}_\nu}\nu_{\boldsymbol{w}}(\boldsymbol{u}),$$

  where the last equality follows from Theorem 1.

- *Step 4.* We have shown that if ν satisfies Axioms C, S, CV, TI, PH, then $\nu(\boldsymbol{u}) = \sup_{\boldsymbol{w}\in\mathcal{W}_\nu} \nu_{\boldsymbol{w}}(\boldsymbol{u})$. Therefore, we immediately have $\nu(\boldsymbol{u}) \geq 0$ since $\nu_{\boldsymbol{w}}(\boldsymbol{u}) \geq 0$. Moreover, by Axiom N, since ν equals zero if and only if $\boldsymbol{u} = \alpha\mathbf{1}$ for some $\alpha \in \mathbb{R}$. This implies that $\mathcal{W}_\nu \neq \{\mathbf{0}\}$ (otherwise, $\nu(\boldsymbol{u}) = 0$ for any $\boldsymbol{u} \in \mathbb{R}^N$).

To conclude, if ν is a convex fairness measure, then $\nu(\boldsymbol{u}) = \sup_{\boldsymbol{w}\in\mathcal{W}_\nu} \nu_{\boldsymbol{w}}(\boldsymbol{u})$, where $\mathcal{W}_\nu = \mathrm{dom}(\nu^*) \cap \mathbb{R}^N_\uparrow \subseteq \{\boldsymbol{w} \in \mathbb{R}^N \mid \mathbf{1}^\top \boldsymbol{w} = 0\}$ and $\mathcal{W}_\nu \neq \{\mathbf{0}\}$. This completes the proof. $\square$

EC.5.5. Proof of Proposition 2

Since $\mathcal{W}_\nu$ is compact and convex, we have $\boldsymbol{w} \in \mathrm{conv}(\mathcal{E}(\mathcal{W}_\nu))$. For any $\boldsymbol{w} \in \mathrm{conv}(\mathcal{E}(\mathcal{W}_\nu))$, we have $\boldsymbol{w} = \sum_{j=1}^K \alpha_j \widetilde{\boldsymbol{w}}^j$ for some $K > 0$, $\widetilde{\boldsymbol{w}}^j \in \mathcal{E}(\mathcal{W}_\nu)$, and $\alpha_j \geq 0$ with $\sum_{j=1}^K \alpha_j = 1$. Then,

$$\nu_{\boldsymbol{w}}(\boldsymbol{u}) = \sum_{i=1}^N w_i u_{(i)} = \sum_{i=1}^N \left(\sum_{j=1}^K \alpha_j \widetilde{w}_i^j \right) u_{(i)} = \sum_{j=1}^K \alpha_j \left(\sum_{i=1}^N \widetilde{w}_i^j u_{(i)} \right) = \sum_{j=1}^K \alpha_j \nu_{\widetilde{\boldsymbol{w}}^j}(\boldsymbol{u}).$$

That is, $\nu_{\boldsymbol{w}}(\boldsymbol{u})$ is a convex combination of $\nu_{\widetilde{\boldsymbol{w}}^j}(\boldsymbol{u})$ for $j \in [K]$. As a result, we have $\nu_{\boldsymbol{w}}(\boldsymbol{u}) \leq \max_{j\in[K]} \nu_{\widetilde{\boldsymbol{w}}^j}(\boldsymbol{u})$. Hence, it suffices to consider the supremum over the set of extreme points $\mathcal{E}(\mathcal{W}_\nu)$. Finally, since $\nu_{\mathbf{0}} \equiv 0$ is always dominated by $\nu_{\boldsymbol{w}}$ for any $\boldsymbol{w} \neq \mathbf{0}$, we can consider the supremum over the set $\mathcal{E}(\mathcal{W}_\nu) \setminus \{\mathbf{0}\}$. This completes the proof. $\square$

EC.5.6. Proof of Proposition 3

(a) From Proposition EC.1 (in EC.1), both (i) and (iii) are equivalent to $\max_{i\in[N]} u_i - \min_{i\in[N]} u_i = -u_{(1)} + u_{(N)}$. Therefore, a dual set $\mathcal{W}_\nu$ is given by the singleton $\{(-1, 0, \ldots, 0, 1) \in \mathbb{R}^N\}$.

(b) From Mesa et al. (2003), we can write the Gini deviation (ii) as

$$\sum_{i=1}^N \sum_{j=1}^N |u_i - u_j| = \sum_{i=1}^N 2(2i - 1 - N) u_{(i)} = \sum_{i=1}^N w'_i u_{(i)},$$

where we let $w'_i = 2(2i - 1 - N)$. Note that $w'_1 = 2(1 - N) < 0$, $w'_N = 2(N - 1) > 0$ and w'_i is increasing in i. Moreover,

$$\sum_{i=1}^N w'_i = 2 \sum_{i=1}^N (2i - 1 - N) = 2[N(N+1) - N(N+1)] = 0.$$

Therefore, a dual set $\mathcal{W}_\nu$ is given by the singleton $\{\boldsymbol{w}'\}$.

(c) Consider $\|\boldsymbol{u} - \overline{u}\mathbf{1}\|_p$ for some $p \in [1, \infty]$. Let q be such that $1/p + 1/q = 1$, where we let $q = \infty$ if $p = 1$, and $q = 1$ if $p = \infty$. Since ℓ_p norm is the dual of ℓ_q norm,

$$\|\boldsymbol{u} - \overline{u}\mathbf{1}\|_p = \sup_{\boldsymbol{w}' : \|\boldsymbol{w}'\|_q \leq 1} (\boldsymbol{u} - \overline{u}\mathbf{1})^\top \boldsymbol{w}' = \sup_{\boldsymbol{w}' : \|\boldsymbol{w}'\|_q \leq 1} \boldsymbol{u}^\top (\boldsymbol{w}' - \overline{w}'\mathbf{1}), \tag{EC.8}$$

where $\overline{w}' = (1/N)\mathbf{1}^\top \boldsymbol{w}'$. Note that (iv)–(vi) can be written as $\|\boldsymbol{u} - \overline{u}\mathbf{1}\|_1$, $\|\boldsymbol{u} - \overline{u}\mathbf{1}\|_2$, and $\|\boldsymbol{u} - \overline{u}\mathbf{1}\|_\infty$ respectively. Thus, from (EC.8), the desired dual set is

$$\begin{aligned}\mathcal{W}_\nu &= \mathcal{S}^N \cap \left\{ \boldsymbol{w} \in \mathbb{R}^N \;\middle|\; \boldsymbol{w} = \boldsymbol{w}' - \overline{w}'\mathbf{1},\, \overline{w}' = \frac{1}{N}\sum_{j=1}^N w_j',\, \|\boldsymbol{w}'\|_q \le 1 \right\} \\ &= \left\{ \boldsymbol{w} \in \mathbb{R}^N \;\middle|\; \boldsymbol{w} = \boldsymbol{w}' - \overline{w}'\mathbf{1},\, \overline{w}' = \frac{1}{N}\sum_{j=1}^N w_j',\, \|\boldsymbol{w}'\|_q \le 1,\, \boldsymbol{w}' \in \mathbb{R}^N_\uparrow \right\},\end{aligned}$$

where the last equality follows from the facts that $\mathbf{1}^\top \boldsymbol{w} = \mathbf{1}^\top \boldsymbol{w}' - \overline{w}'(\mathbf{1}^\top \mathbf{1}) = 0$, and $\boldsymbol{w} \in \mathbb{R}^N_\uparrow$ if and only if $\boldsymbol{w}' \in \mathbb{R}^N_\uparrow$.

(d) From Proposition EC.1 (in EC.1), (vii) is equivalent to $N\|\boldsymbol{u} - \overline{u}\mathbf{1}\|_\infty$. The same argument in (EC.8) shows that

$$N\|\boldsymbol{u} - \overline{u}\mathbf{1}\|_p = \sup_{\boldsymbol{w}' : \|\boldsymbol{w}'\|_q \le 1} (\boldsymbol{u} - \overline{u}\mathbf{1})^\top (N\boldsymbol{w}') = \sup_{\boldsymbol{w}' : \|\boldsymbol{w}'\|_q \le N} \boldsymbol{u}^\top (\boldsymbol{w}' - \overline{w}'\mathbf{1}),$$

which gives the desired dual set $\mathcal{W}_\nu$.

(e) Finally, for (viii), let $k = k(\boldsymbol{u})$ be the number of entries in $\boldsymbol{u}$ that are closer to $u_{(1)}$, i.e., $u_{(i)} - u_{(1)} < u_{(N)} - u_{(i)}$ for $i \in [k]$. Then,

$$\begin{aligned}\sum_{i=1}^N \max_{j \in [N]} |u_i - u_j| &= \sum_{i=1}^k \left[u_{(N)} - u_{(i)}\right] + \sum_{i=k+1}^N \left[u_{(i)} - u_{(1)}\right] \\ &= \left[-(N-k) - 1\right] u_{(1)} + \sum_{i=2}^k (-1)u_{(i)} + \sum_{i=k+1}^{N-1} u_{(i)} + (k+1)u_{(N)}.\end{aligned}$$

Note that k takes value in $[N-1]$ only, it suffices to consider $\boldsymbol{w}^k$ with entries $w_1^k = -(N-k) - 1$, $w_2^k = \cdots = w_k^k = -1$, $w_{k+1}^k = \cdots = w_{N-1}^k = 1$, and $w_N^k = k+1$. Moreover, it is easy to verify that if there are k entries in $\boldsymbol{u}$ that are closer to $u_{(1)}$, then $\nu_{\boldsymbol{w}^k}(\boldsymbol{u}) \ge \nu_{\boldsymbol{w}^h}(\boldsymbol{u})$ for $h \ne k$. Indeed, if $h < k$, then

$$\begin{aligned}&\nu_{\boldsymbol{w}^k}(\boldsymbol{u}) - \nu_{\boldsymbol{w}^h}(\boldsymbol{u}) \\ &= \left\{ \sum_{i=1}^k \left[u_{(N)} - u_{(i)}\right] + \sum_{i=k+1}^N \left[u_{(i)} - u_{(1)}\right] \right\} - \left\{ \sum_{i=1}^h \left[u_{(N)} - u_{(i)}\right] + \sum_{i=h+1}^N \left[u_{(i)} - u_{(1)}\right] \right\} \\ &= \sum_{i=h+1}^k \left\{ \left[u_{(N)} - u_{(i)}\right] - \left[u_{(i)} - u_{(1)}\right] \right\} > 0.\end{aligned}$$

Following a similar argument, if $h > k$, we also have $\nu_{\boldsymbol{w}^k}(\boldsymbol{u}) - \nu_{\boldsymbol{w}^h}(\boldsymbol{u}) > 0$. Hence, a dual set $\mathcal{W}_\nu$ is given by $\{\boldsymbol{w}^k\}_{k=1}^N$.

This completes the proof. $\square$

EC.5.7. Proof of Theorem 4

First, if $\mathcal{W}_1 = \beta\mathcal{W}_2$ for some $\beta > 0$, then we have

$$\nu_1(\boldsymbol{u}) = \sup_{\boldsymbol{w}\in\mathcal{W}_1} \nu_{\boldsymbol{w}}(\boldsymbol{u}) = \sup_{\boldsymbol{w}\in\beta\mathcal{W}_2} \nu_{\boldsymbol{w}}(\boldsymbol{u}) = \beta \sup_{\boldsymbol{w}\in\mathcal{W}_2} \nu_{\boldsymbol{w}}(\boldsymbol{u}) = \beta\nu_2(\boldsymbol{u}),$$

implying the equivalence of ν_1 and ν_2. Next, if ν_1 is equivalent to ν_2, then $\nu_1(\boldsymbol{u}) = \beta\nu_2(\boldsymbol{u})$ for some $\beta > 0$. From the proof of Theorem 3, we have

$$\sup_{\boldsymbol{w}\in\mathrm{dom}(\nu_1^*)} \boldsymbol{u}^\top \boldsymbol{w} = \nu_1(\boldsymbol{u}) = \beta\nu_2(\boldsymbol{u}) = \beta \sup_{\boldsymbol{w}\in\mathrm{dom}(\nu_2^*)} \boldsymbol{u}^\top \boldsymbol{w} = \sup_{\boldsymbol{w}\in\beta\,\mathrm{dom}(\nu_2^*)} \boldsymbol{u}^\top \boldsymbol{w}. \tag{EC.9}$$

As a result, (EC.9) implies that the support functions of the convex compact sets $\mathrm{dom}(\nu_1^*)$ and $\beta\,\mathrm{dom}(\nu_2^*)$ are the same. Therefore, we have $\mathrm{dom}(\nu_1^*) = \beta\,\mathrm{dom}(\nu_2^*)$ by Theorem 13.2 of Rockafellar (1970), and thus, $\mathcal{W}_1 = \mathrm{dom}(\nu_1^*) \cap \mathbb{R}^N_\uparrow = \beta\,\mathrm{dom}(\nu_2^*) \cap \mathbb{R}^N_\uparrow = \beta\mathcal{W}_2$. $\square$

EC.5.8. Proof of Theorem 5

From Theorem 1, we can write $\nu(\boldsymbol{u}) = \max_{P\in\mathcal{P}} \boldsymbol{u}^\top(P\boldsymbol{w})$, where $\mathcal{P}$ is the set of all permutation matrices given by

$$\mathcal{P} = \left\{ P \in \mathbb{R}^{N\times N} \;\middle|\; \sum_{i=1}^N P_{ij} = 1,\, \sum_{j=1}^N P_{ij} = 1,\, P_{ij} \in \{0,1\},\, \forall i \in [N],\, j \in [N] \right\}. \tag{EC.10}$$

Note that the objective $\boldsymbol{u}^\top(P\boldsymbol{w})$ is linear in P, and the constraint matrix formed by the assignment constraints in $\mathcal{P}$ is totally unimodular (Martello and Toth 1987). Hence, $\max_{P\in\mathcal{P}} \boldsymbol{u}^\top(P\boldsymbol{w})$ is a linear program in variables P_{ij} and we can take its dual as

$$\max_{P\in\mathcal{P}} \boldsymbol{u}^\top(P\boldsymbol{w}) = \min_{\boldsymbol{\lambda}\in\mathbb{R}^N,\, \boldsymbol{\theta}\in\mathbb{R}^N} \left\{ \mathbf{1}^\top(\boldsymbol{\lambda}+\boldsymbol{\theta}) \;\middle|\; \lambda_i + \theta_j \geq u_i w_j,\, \forall i \in [N],\, j \in [N] \right\}. \tag{EC.11}$$

Combining (EC.11) with the outer minimization over $\boldsymbol{x}$ and $\boldsymbol{u}$ in problem (4), we obtain the reformulation in (5). $\square$

REMARK EC.2. We note that a similar reformulation technique also appears in studies that optimize the ordered-median function (see, e.g., Blanco et al. 2016). However, as pointed out in Section 4, our proposed class of order-based fairness measures is different from the ordered-median functions. In particular, the latter is a combination of the mean (an inefficiency measure) and an order-based fairness measure; see Section EC.3 for a thorough discussion on the differences.

EC.5.9. Proof of Proposition 4

First, note that we can relax the set of permutation matrices $\mathcal{P}$ in (EC.10) as

$$\mathcal{P} = \left\{ P \in \mathbb{R}^{N\times N} \;\middle|\; \sum_{i=1}^{N} P_{ij} = 1,\, \sum_{j=1}^{N} P_{ij} \le 1,\, P_{ij} \in \{0,1\},\, \forall i \in [N],\, j \in [N] \right\}.$$

Indeed, the first constraint requires that every column of P has exactly one entry with value 1 and the second constraint requires that every row of P has at most one entry with value 1. Since P is an N-by-N matrix, this immediately ensures that P has exactly one entry with value 1 in every row and column. As a result, the dual variable λ_i associated with the second constraint is non-negative. Next, for a given $\boldsymbol{u} \in \mathbb{R}^N$, let $P^\star$ be an optimal solution to the primal problem $\max_{P\in\mathcal{P}} \boldsymbol{u}^\top(P\boldsymbol{w})$ and $(\lambda^\star, \theta^\star)$ be an optimal dual solution. Let $\{(i,\pi(i))\}_{i\in[N]}$ be the set of pairs of indices such that $P^\star_{i,\pi(i)} = 1$ (and is 0 otherwise). Without loss of generality, we can assume that there exists $i' \in [N]$ such that $\lambda^\star_{i'} = 0$. Indeed, if $\varepsilon := \min_{i\in[N]} \lambda^\star_i > 0$, then $(\widetilde{\boldsymbol{\lambda}}, \widetilde{\boldsymbol{\theta}})$ defined by $\widetilde{\lambda}_i = \lambda^\star_i - \varepsilon$ and $\widetilde{\theta}_i = \theta^\star_i + \varepsilon$ for all $i \in [N]$ is another optimal solution to (5).

Now, we derive an upper bound on $\theta^\star_i$. By the complementary slackness condition, if $P^\star_{i,\pi(i)} = 1$, then we have $\lambda^\star_i + \theta^\star_{\pi(i)} = u_i w_{\pi(i)}$ (i.e., with a zero slack variable), implying that $\theta^\star_{\pi(i)} = u_i w_{\pi(i)} - \lambda^\star_i$. Since $\lambda^\star_i \ge 0$, we immediately have $\theta^\star_{\pi(i)} \le u_i w_{\pi(i)} \le u_{\max} w_{\pi(i)}$, where $u_{\max} = \max_{i\in[N]} u_i$. Now, we derive an upper bound on λ_i. Using $\theta^\star_{\pi(i)} = u_i w_{\pi(i)} - \lambda^\star_i$ and letting $i = \pi^{-1}(j)$, constraints (5b) implies that $\lambda^\star_i + \left(u_{\pi^{-1}(j)} w_j - \lambda^\star_{\pi^{-1}(j)}\right) \ge u_i w_j$ for all $i \in [N]$ and $j \in [N]$, which is equivalent to $\lambda^\star_i + \left(u_j w_{\pi(j)} - \lambda^\star_j\right) \ge u_i w_{\pi(j)}$ for all $i \in [N]$ and $j \in [N]$. Hence, we have $\lambda^\star_i - \lambda^\star_j \ge (u_i - u_j) w_{\pi(j)}$ for all $i \in [N]$ and $j \in [N]$. Setting $i = i'$, since $\lambda^\star_{i'} = 0$, the inequalities imply that $\lambda^\star_j \le (u_j - u_{i'}) w_{\pi(j)} \le (u_{\max} - u_{\min}) \|\boldsymbol{w}\|_\infty$ for all $j \in [N] \setminus \{i'\}$, where $u_{\min} = \min_{i\in[N]} u_i$. Finally, the lower bound of θ_i follows from the upper bound of λ_i that $\theta^\star_{\pi(i)} = u_i w_{\pi(i)} - \lambda^\star_i \ge u_i w_{\pi(i)} - (u_{\max} - u_{\min}) \|\boldsymbol{w}\|_\infty$.

Note that the above lower and upper bounds on $\boldsymbol{\lambda}$ and $\boldsymbol{\theta}$ are obtained by fixing a vector $\boldsymbol{u} \in \mathbb{R}^N$. Thus, the desired lower bound follows from taking the infimum over all feasible $\boldsymbol{u} \in \mathcal{U} := \{\boldsymbol{u} \in \mathbb{R}^N \mid \boldsymbol{u} = U(\boldsymbol{x}),\, \boldsymbol{x} \in \mathcal{X}\}$ and the desired upper bounds follow from taking the supremum over all $\boldsymbol{u} \in \mathcal{U}$. Thus, we obtain the bounds on $\boldsymbol{\lambda}$

$$0 \le \lambda_i \le \sup_{\boldsymbol{u}\in\mathcal{U}} \left\{u_{\max} - u_{\min}\right\} \cdot \|\boldsymbol{w}\|_\infty \le (U_{\max} - U_{\min}) \|\boldsymbol{w}\|_\infty =: \overline{\lambda},$$

the lower bound on $\boldsymbol{\theta}$

$$\theta_j \ge \inf_{\boldsymbol{u}\in\mathcal{U}} \left\{ u_{\pi^{-1}(j)} w_j - (u_{\max} - u_{\min}) \|\boldsymbol{w}\|_\infty \right\} \ge \min\left\{U_{\max} w_j,\, U_{\min} w_j\right\} - \overline{\lambda},$$

and the upper bound on $\boldsymbol{\theta}$

$$\theta_j \le \sup_{\boldsymbol{u}\in\mathcal{U}} \left\{u_{\max} w_j\right\} \le \max\left\{U_{\max} w_j,\, U_{\min} w_j\right\}.$$

This completes the proof. $\square$

EC.5.10. Proof of Proposition 5

By the LP reformulation of $\nu_{\boldsymbol{w}^k}(\boldsymbol{u})$ in (EC.11), $\delta \geq \nu_{\boldsymbol{w}^k}(\boldsymbol{u})$ if and only if there exist $\boldsymbol{\lambda}^k$ and $\boldsymbol{\theta}^k$ such that $\lambda_i^k + \theta_{i'}^k \geq u_i w_{i'}^k$ for all $i \in [N]$, $i' \in [N]$ and $\delta \geq \mathbf{1}^\top(\boldsymbol{\lambda}^k + \boldsymbol{\theta}^k)$. Therefore, by introducing the variables $\boldsymbol{\lambda}^k$ and $\boldsymbol{\theta}^k$, we can reformulate (9) into (10). $\square$

EC.5.11. Proof of Lemma 1

First, note that

$$\begin{aligned}\sup_{\boldsymbol{w}_1 \in \mathcal{W}_1} \nu_{\boldsymbol{w}_1}(\boldsymbol{u}) - \sup_{\boldsymbol{w}_2 \in \mathcal{W}_2} \nu_{\boldsymbol{w}_2}(\boldsymbol{u}) &= \sup_{\boldsymbol{w}_1 \in \mathcal{W}_1} \inf_{\boldsymbol{w}_2 \in \mathcal{W}_2} \Big\{ \nu_{\boldsymbol{w}_1}(\boldsymbol{u}) - \nu_{\boldsymbol{w}_2}(\boldsymbol{u}) \Big\} \\ &\leq \sup_{\boldsymbol{w}_1 \in \mathcal{W}_1} \inf_{\boldsymbol{w}_2 \in \mathcal{W}_2} \|\boldsymbol{w}_1 - \boldsymbol{w}_2\| \cdot \|\boldsymbol{u}\|_2 \leq d_H(\mathcal{W}_1, \mathcal{W}_2) \cdot \|\boldsymbol{u}\|_2,\end{aligned}$$

where the first inequality follows from Cauchy-Schwarz inequality, and the second inequality follows from the definition of d_H. Next, using the same argument, we have

$$\begin{aligned}\sup_{\boldsymbol{w}_1 \in \mathcal{W}_1} \nu_{\boldsymbol{w}_1}(\boldsymbol{u}) - \sup_{\boldsymbol{w}_2 \in \mathcal{W}_2} \nu_{\boldsymbol{w}_2}(\boldsymbol{u}) &= \inf_{\boldsymbol{w}_2 \in \mathcal{W}_2} \sup_{\boldsymbol{w}_1 \in \mathcal{W}_1} \Big\{ \nu_{\boldsymbol{w}_1}(\boldsymbol{u}) - \nu_{\boldsymbol{w}_2}(\boldsymbol{u}) \Big\} \\ &\geq - \sup_{\boldsymbol{w}_2 \in \mathcal{W}_2} \inf_{\boldsymbol{w}_1 \in \mathcal{W}_1} \|\boldsymbol{w}_1 - \boldsymbol{w}_2\| \cdot \|\boldsymbol{u}\|_2 \geq -d_H(\mathcal{W}_1, \mathcal{W}_2) \cdot \|\boldsymbol{u}\|_2.\end{aligned}$$

Hence, we obtain $\big|\nu_1(\boldsymbol{u}) - \nu_2(\boldsymbol{u})\big| \leq d_H(\mathcal{W}_1, \mathcal{W}_2) \cdot \|\boldsymbol{u}\|_2$. $\square$

EC.5.12. Proof of Theorem 6

First, we prove part (a). Note that

$$\begin{aligned}\upsilon_2^\star - \upsilon_1^\star &= \min_{\boldsymbol{x} \in \mathcal{X}} \Big\{ f(U(\boldsymbol{x})) + \gamma \nu_2(U(\boldsymbol{x})) \Big\} - \min_{\boldsymbol{x} \in \mathcal{X}} \Big\{ f(U(\boldsymbol{x})) + \gamma \nu_1(U(\boldsymbol{x})) \Big\} \\ &\leq \Big[f(U(\boldsymbol{x}_1^\star)) + \gamma \sup_{\boldsymbol{w} \in \mathcal{W}_2} \nu_{\boldsymbol{w}}(U(\boldsymbol{x}_1^\star)) \Big] - \Big[f(U(\boldsymbol{x}_1^\star)) + \gamma \sup_{\boldsymbol{w} \in \mathcal{W}_1} \nu_{\boldsymbol{w}}(U(\boldsymbol{x}_1^\star)) \Big] \qquad \text{(EC.12a)} \\ &\leq \gamma U_{\max} d_H(\mathcal{W}_1, \mathcal{W}_2), \qquad \text{(EC.12b)}\end{aligned}$$

where (EC.12a) follows from $\boldsymbol{x}_1^\star \in \mathcal{X}$ and the optimality of $\boldsymbol{x}_1^\star$, and (EC.12b) follows from Lemma 1. Using the same logic, it is easy to verify that $\upsilon_1^\star - \upsilon_2^\star \leq \gamma U_{\max} d_H(\mathcal{W}_1, \mathcal{W}_2)$. This completes the proof showing that $\big|\upsilon_1^\star - \upsilon_2^\star\big| \leq \gamma U_{\max} d_H(\mathcal{W}_1, \mathcal{W}_2)$.

Next, we proceed to prove part (b). Note that

$$\begin{aligned}\upsilon_2^\star - \upsilon_1^\star &= \min_{\boldsymbol{x} \in \mathcal{X}} \Big\{ f(U(\boldsymbol{x})) + \gamma \nu_2(U(\boldsymbol{x})) \Big\} - \min_{\boldsymbol{x} \in \mathcal{X}} \Big\{ f(U(\boldsymbol{x})) + \gamma \nu_1(U(\boldsymbol{x})) \Big\} \\ &= \Big[f(U(\boldsymbol{x}_2^\star)) + \gamma \nu_{\boldsymbol{w}_2^\star}(U(\boldsymbol{x}_2^\star)) \Big] - \Big[f(U(\boldsymbol{x}_1^\star)) + \gamma \nu_{\boldsymbol{w}_1^\star}(U(\boldsymbol{x}_1^\star)) \Big] \\ &= \bigg\{ \Big[f(U(\boldsymbol{x}_2^\star)) + \gamma \nu_{\boldsymbol{w}_2^\star}(U(\boldsymbol{x}_2^\star)) \Big] - \Big[f(U(\boldsymbol{x}_1^\star)) + \gamma \nu_{\boldsymbol{w}_2^\star}(U(\boldsymbol{x}_1^\star)) \Big] \bigg\} \\ &\quad + \bigg\{ \Big[f(U(\boldsymbol{x}_1^\star)) + \gamma \nu_{\boldsymbol{w}_2^\star}(U(\boldsymbol{x}_1^\star)) \Big] - \Big[f(U(\boldsymbol{x}_1^\star)) + \gamma \nu_{\boldsymbol{w}_1^\star}(U(\boldsymbol{x}_1^\star)) \Big] \bigg\}\end{aligned}$$

$$\leq -\tau_2 \|\boldsymbol{x}_2^\star - \boldsymbol{x}_1^\star\|_2^2 + \gamma \left\{ \sup_{\boldsymbol{w} \in \mathcal{W}_2} \nu_{\boldsymbol{w}}(U(\boldsymbol{x}_1^\star)) - \sup_{\boldsymbol{w} \in \mathcal{W}_1} \nu_{\boldsymbol{w}}(U(\boldsymbol{x}_1^\star)) \right\} \tag{EC.13a}$$

$$\leq -\tau_2 \|\boldsymbol{x}_2^\star - \boldsymbol{x}_1^\star\|_2^2 + \gamma U_{\max} d_H(\mathcal{W}_1, \mathcal{W}_2), \tag{EC.13b}$$

where (EC.13a) follows from (13), $\boldsymbol{w}_2^\star \in \mathcal{W}_2$, and the optimality of $\boldsymbol{w}_1^\star$, and (EC.13b) follows from Lemma 1. Using the same logic, it is easy to verify that $\upsilon_1^\star - \upsilon_2^\star \leq -\tau_1 \|\boldsymbol{x}_1^\star - \boldsymbol{x}_2^\star\|_2^2 + \gamma U_{\max} d_H(\mathcal{W}_1, \mathcal{W}_2)$. Thus, we have

$$0 \leq \left|\upsilon_1^\star - \upsilon_2^\star\right| \leq -\min\{\tau_1, \tau_2\} \|\boldsymbol{x}_1^\star - \boldsymbol{x}_2^\star\|_2^2 + \gamma U_{\max} d_H(\mathcal{W}_1, \mathcal{W}_2),$$

which directly implies the desired inequality (14). □

EC.5.13. Proof of Theorem EC.1

Note that it is straightforward to verify the sufficient part using the definition of ρ. Thus, we focus on the necessary part. Assuming that $\rho = \nu/g$ satisfies Axioms NR, S, and SI, we next show that g is symmetric and positive homogeneous with $g \geq \nu$. For notational simplicity, we define $\mathcal{U}_0 = \{\boldsymbol{u} \mid \boldsymbol{u} = \alpha \mathbf{1},\, \alpha \geq 0\}$ and its complement $\mathcal{U}_0^c = \mathbb{R}_+^N \setminus \mathcal{U}_0$. We divide the proof into the following three steps.

- *Step 1*. First, consider $\boldsymbol{u} \in \mathcal{U}_0^c$. By Axiom NR, we have $\rho(\boldsymbol{u}) > 0$. Also, by definition of convex fairness measures (Axiom N), we have $\nu(\boldsymbol{u}) > 0$. Since $0 < \rho(\boldsymbol{u}) = \nu(\boldsymbol{u})/g(\boldsymbol{u})$, we have $g(\boldsymbol{u}) > 0$. Also, since $\rho(\boldsymbol{u}) \leq 1$ by Axiom NR, it follows that $g(\boldsymbol{u}) \geq \nu(\boldsymbol{u}) \geq 0$ for any $\boldsymbol{u} \in \mathcal{U}_0^c$. Now, consider $\boldsymbol{u} \in \mathcal{U}_0$, i.e., $\boldsymbol{u} = \boldsymbol{u}_\alpha = \alpha \mathbf{1}$ for some $\alpha > 0$. Consider the sequence $\{\boldsymbol{u}_\alpha^k = [(\alpha+1)/k, \alpha, \ldots, \alpha]^\top\}_{k \in \mathbb{N}} \subset \mathbb{R}_+^N$ that converges to $\boldsymbol{u}_\alpha$ as $k \to \infty$. Since $\boldsymbol{u}_\alpha^k \in \mathcal{U}_0^c$, we have $g(\boldsymbol{u}_\alpha^k) \geq \nu(\boldsymbol{u}_\alpha^k)$ for all $k \in \mathbb{N}$. Taking limit $k \to \infty$ on both sides of the inequality, we obtain $g(\boldsymbol{u}_\alpha) \geq \nu(\boldsymbol{u}_\alpha)$ by continuity of g and ν. This shows that $g \geq \nu$.
- *Step 2*. Similar to step 1, consider $\boldsymbol{u} \in \mathcal{U}_0^c$. As shown in step 1, we have $\rho(\boldsymbol{u}) > 0$ and $g(\boldsymbol{u}) > 0$. Note that for any permutation matrix P, we have $g(P\boldsymbol{u}) = \nu(P\boldsymbol{u})/\rho(P\boldsymbol{u}) = \nu(\boldsymbol{u})/\rho(\boldsymbol{u}) = g(\boldsymbol{u})$ since $\rho(P\boldsymbol{u}) = \rho(\boldsymbol{u})$ by Axiom S and $\nu(P\boldsymbol{u}) = \nu(\boldsymbol{u})$ by symmetry of ν. It follows that g is symmetric. Finally, if $\boldsymbol{u} \in \mathcal{U}_0$, we can apply a similar limiting argument in step 1.
- *Step 3*. Similar to step 1, consider $\boldsymbol{u} \in \mathcal{U}_0^c$. As shown in step 1, we have $\rho(\boldsymbol{u}) > 0$ and $g(\boldsymbol{u}) > 0$. Note that for any $\alpha > 0$, we have $g(\alpha\boldsymbol{u}) = \nu(\alpha\boldsymbol{u})/\rho(\alpha\boldsymbol{u}) = \alpha\nu(\boldsymbol{u})/\rho(\boldsymbol{u}) = \alpha g(\boldsymbol{u})$ since $\rho(\alpha\boldsymbol{u}) = \rho(\boldsymbol{u})$ by Axiom SI and $\nu(\alpha\boldsymbol{u}) = \alpha\nu(\boldsymbol{u})$ by positive homogeneity of ν. It follows that g is positive homogeneous. Finally, if $\boldsymbol{u} \in \mathcal{U}_0$, we can apply a similar limiting argument in step 1.

This completes the proof. □

EC.5.14. Proof of Theorem EC.2

First, note that g is symmetric and positive homogeneous. In addition, we claim that $g(\boldsymbol{u}) \geq \nu(\boldsymbol{u})$ for all $\boldsymbol{u} \in \mathbb{R}_+^N$. To prove this claim, consider the following optimization problem:

$$\max_{\boldsymbol{u} \in \mathbb{R}_+^N} \left\{ \nu(\boldsymbol{u}) = \sup_{\boldsymbol{w} \in \mathcal{W}_\nu} \sum_{i=1}^N w_i u_{(i)} \,\middle|\, \mathbf{1}^\top \boldsymbol{u} = \gamma \right\} \tag{EC.14}$$

for any $\gamma > 0$. It is easy to show that $\boldsymbol{u}^\star_\gamma = (0,\dots,0,\gamma)^\top$ is an optimal solution to (EC.14) with objective value $\gamma \sup_{\boldsymbol{w}\in\mathcal{W}_\nu} \|\boldsymbol{w}_+\|_\infty$. Thus, for any $\gamma > 0$ and $\boldsymbol{u} \in \mathbb{R}^N_+$ such that $\mathbf{1}^\top \boldsymbol{u} = \gamma$, we have

$$\nu(\boldsymbol{u}) \le \gamma \sup_{\boldsymbol{w}\in\mathcal{W}_\nu} \|\boldsymbol{w}_+\|_\infty = N\overline{u} \sup_{\boldsymbol{w}\in\mathcal{W}_\nu} \|\boldsymbol{w}_+\|_\infty \le CN\overline{u} = g(\boldsymbol{u}),$$

showing that $g(\boldsymbol{u}) \ge \nu(\boldsymbol{u})$ for all $\boldsymbol{u} \in \mathbb{R}^N_+$. Thus, by Theorem EC.1, ρ satisfies Axioms NR, S, and SI.

Finally, we show that g satisfies Axiom SCV. Note that a function $h : \mathbb{R}^N_+ \to \mathbb{R}$ is Schur convex if and only if h is symmetric and h satisfies the following condition: $h(\boldsymbol{u}^\lambda_{12}) \le h(\boldsymbol{u})$ for any $\boldsymbol{u} \in \mathbb{R}^N_+$ and $\lambda \in (0,1)$, where $\boldsymbol{u}^\lambda_{12} = \lambda \boldsymbol{u} + (1-\lambda)\boldsymbol{u}_{12}$ and $\boldsymbol{u}_{12} = (u_2, u_1, u_3, \dots, u_N)^\top$ (see Theorem 2.4 of Stępniak 2007). Since ρ is symmetric, it suffices to show that $\rho(\boldsymbol{u}^\lambda_{12}) \le \rho(\boldsymbol{u})$ for any $\boldsymbol{u} \in \mathbb{R}^N_+$. Indeed, since $\mathbf{1}^\top \boldsymbol{u}^\lambda_{12} = \mathbf{1}^\top \boldsymbol{u} = \mathbf{1}^\top \boldsymbol{u}_{12}$, letting $\mu = \mathbf{1}^\top \boldsymbol{u}/N$, we have

$$\rho(\boldsymbol{u}^\lambda_{12}) = \frac{\nu(\boldsymbol{u}^\lambda_{12})}{CN\mu} \le \frac{\lambda\nu(\boldsymbol{u}) + (1-\lambda)\nu(\boldsymbol{u}_{12})}{CN\mu} = \frac{\nu(\boldsymbol{u})}{CN\mu} = \frac{\nu(\boldsymbol{u})}{g(\boldsymbol{u})} = \rho(\boldsymbol{u}),$$

where the inequality follows from the convexity of ν, and the second equality follows from $\nu(\boldsymbol{u}_{12}) = \nu(\boldsymbol{u})$ (Axiom S). Hence, ρ is Schur convex. $\square$

EC.5.15. Proof of Theorem EC.3

Note that we have proved the sufficiency part in Theorem EC.2, and hence, we only need to prove the necessity part. From Theorem EC.1, we know that g is symmetric and positive homogeneous with $g \ge \nu$. This implies that $\tilde{g} : \mathbb{R}_+ \to \mathbb{R}$ is also a (one-dimensional) positive homogeneous function. Theorem 2.2.1 of Castillo and Ruiz-Cobo (1992) shows that the only class of solutions of this homogeneous functional equation is of the form $\tilde{g}(\overline{u}) = c\overline{u}$ for some constant c. Finally, by Axiom NR, we have $\rho(\boldsymbol{u}) \le 1$, which implies $\nu(\boldsymbol{u}) \le c\overline{u}$. Recall, from the proof of Theorem EC.2, for any $\gamma > 0$ and $\boldsymbol{u} \in \mathbb{R}^N_+$ such that $\boldsymbol{u}^\top \mathbf{1} = \gamma$, we have $\nu(\boldsymbol{u}) \le \gamma \sup_{\boldsymbol{w}\in\mathcal{W}_\nu} \|\boldsymbol{w}_+\|_\infty = N \sup_{\boldsymbol{w}\in\mathcal{W}_\nu} \|\boldsymbol{w}_+\|_\infty \cdot \overline{u}$. Thus, to ensure that $\rho(\boldsymbol{u}) \le 1$ for all $\boldsymbol{u} \in \mathbb{R}^N_+$, we must have $c \ge N \sup_{\boldsymbol{w}\in\mathcal{W}_\nu} \|\boldsymbol{w}_+\|_\infty$. $\square$

EC.5.16. Proof of Proposition EC.3

Assume that $\boldsymbol{u} \in \mathcal{U}^\star$. Let $\gamma = \mathbf{1}^\top \boldsymbol{u}$. Since $\boldsymbol{u} \in \mathcal{U}^\star$, we have $\boldsymbol{u} \in \arg\max\{\nu(\boldsymbol{u}') \mid \mathbf{1}^\top \boldsymbol{u}' = \gamma\}$, which implies that

$$\nu(\boldsymbol{u}) \ge \nu(0,\dots,0,\gamma) = \sup_{\boldsymbol{w}\in\mathcal{W}_\nu} \|\boldsymbol{w}_+\|_\infty \gamma = \sup_{\boldsymbol{w}\in\mathcal{W}_\nu} \|\boldsymbol{w}_+\|_\infty (\mathbf{1}^\top \boldsymbol{u}) = g(\boldsymbol{u}) > 0.$$

Thus, it follows that $\rho(\boldsymbol{u}) = \nu(\boldsymbol{u})/g(\boldsymbol{u}) \ge 1$. This shows that $\rho(\boldsymbol{u}) = 1$.

Next, assume that $\rho(\boldsymbol{u}) = 1$. From Theorem EC.3, we know that $\rho(\boldsymbol{u}) \le 1$ (Axiom NR). This implies that the set of $\boldsymbol{u} \in \mathbb{R}^N_+$ such that $\rho(\boldsymbol{u}) = 1$ is the set of the maximizers of $\rho(\boldsymbol{u})$, i.e., $\boldsymbol{u} \in \arg\max_{\boldsymbol{u}'\in\mathbb{R}^N_+}\{\rho(\boldsymbol{u}')\}$. Finally, note that

$$\mathop{\arg\max}_{\boldsymbol{u}'\in\mathbb{R}^N_+} \left\{ \rho(\boldsymbol{u}') = \frac{\nu(\boldsymbol{u}')}{g(\boldsymbol{u}')} \right\} = \mathop{\arg\max}_{\boldsymbol{u}'\in\mathbb{R}^N_+} \left\{ \frac{\nu(\boldsymbol{u}')}{\sup_{\boldsymbol{w}\in\mathcal{W}_\nu} \|\boldsymbol{w}_+\|_\infty (\mathbf{1}^\top \boldsymbol{u}')} \right\} \tag{EC.15}$$

$$= \bigcup_{\gamma>0} \arg\max_{\boldsymbol{u}'\in\mathbb{R}^N_+} \{\nu(\boldsymbol{u}') \mid \mathbf{1}^\top \boldsymbol{u}' = \gamma\} \tag{EC.16}$$
$$= \bigcup_{\gamma>0} \mathcal{U}^\star_\gamma = \mathcal{U}^\star.$$

Here, note that the objective function in the maximization problem (EC.15) is a fraction with both nominator and denominator being positive homogeneous. By classical fractional programming results, this is equivalent to maximizing the nominator with a fixed value of the denominator (see Chapter 4.3 of Stancu-Minasian 1997 for details), which leads to (EC.16). □

EC.5.17. Proof of Corollary EC.1

By Proposition EC.3, it suffices to show that $\mathcal{U}^\star_\gamma = \{P(0,\dots,0,\gamma)^\top \mid P\in\mathcal{P}\}$ for all $\gamma>0$. That is, $\boldsymbol{u} = (0,\dots,0,\gamma)^\top$ is the unique maximizer to the optimization problem $\max_{\boldsymbol{u}'\in\mathbb{R}^N_+}\{\nu(\boldsymbol{u}') \mid \mathbf{1}^\top\boldsymbol{u}'=\gamma,\ \boldsymbol{u}'\in\mathbb{R}^N_\uparrow\}$ in which we seek to find the impact vector $\boldsymbol{u}'$ that maximizes the value of the fairness measure ν. Suppose, on the contrary, that there exists $\boldsymbol{u}\in\mathbb{R}^N_\uparrow$ and $i\in[N-1]$ such that $\rho(\boldsymbol{u})=1$ with $u_j>0$ for all $j\in[i,N]_{\mathbb{Z}}$ and $u_j=0$ otherwise, i.e., $\boldsymbol{u}$ does not take the form $(0,\dots,\gamma)^\top$. Let $\gamma=\mathbf{1}^\top\boldsymbol{u}$. Define $\boldsymbol{u}^1=(0,\dots,0,0,\gamma)^\top$ and $\boldsymbol{u}^2=(0,\dots,0,\gamma/2,\gamma/2)^\top$. Consider the following two cases.

- If $u_N\le\gamma/2$, we must have $\boldsymbol{u}\preceq\boldsymbol{u}^2$. By Schur convexity of ν and our assumption that $\nu(\boldsymbol{u}^1)>\nu(\boldsymbol{u}^2)$, we have $\nu(\boldsymbol{u})\le\nu(\boldsymbol{u}^2)<\nu(\boldsymbol{u}^1)$. Since $\mathbf{1}^\top\boldsymbol{u}=\mathbf{1}^\top\boldsymbol{u}^1=\gamma$, we arrive at the contradiction that $1=\rho(\boldsymbol{u})<\rho(\boldsymbol{u}^1)=1$.
- If $u_N>\gamma/2$, we define the vector $\widetilde{\boldsymbol{u}}\in\mathbb{R}^N_\uparrow$ by $\tilde{u}_j=0$ for $j\in[N-2]$, $\tilde{u}_{N-1}=\sum_{j=i}^{N-1}u_j$, and $\tilde{u}_N=u_N$. By construction, we have $\boldsymbol{u}\preceq\widetilde{\boldsymbol{u}}$ and $\widetilde{\boldsymbol{u}}=\lambda\boldsymbol{u}^2+(1-\lambda)\boldsymbol{u}^1$, where $\lambda=2(1-u_N/\gamma)\in(0,1)$. Thus, by Schur convexity of ν, convexity of ν and our assumption that $\nu(\boldsymbol{u}^1)>\nu(\boldsymbol{u}^2)$, we have $\nu(\boldsymbol{u})\le\nu(\widetilde{\boldsymbol{u}})\le\lambda\nu(\boldsymbol{u}^2)+(1-\lambda)\nu(\boldsymbol{u}^1)<\nu(\boldsymbol{u}^1)$. Again, since $\mathbf{1}^\top\boldsymbol{u}=\mathbf{1}^\top\boldsymbol{u}^1=\gamma$, we arrive at the contradiction that $1=\rho(\boldsymbol{u})<\rho(\boldsymbol{u}^1)=1$.

Thus, this shows that $\rho(\boldsymbol{u})=1$ if and only if $\boldsymbol{u}=P(0,\dots,0,\gamma)^\top$ for any $\gamma>0$ and permutation matrix P.
□

EC.6. Solution Approaches with Convex Fairness Measures in Constraint Form

In this section, we propose solution approaches for optimization problem of the form (3) with our proposed convex fairness measures.

EC.6.1. Convex Fairness Measure

Using the dual representation of convex fairness measures (see Section 5), we can reformulate (3) into

$$\min_{\boldsymbol{x},\,\boldsymbol{u}}\left\{f(\boldsymbol{u})\,\middle|\,\nu_{\boldsymbol{w}}(\boldsymbol{u})\le\eta,\ \forall\boldsymbol{w}\in\mathcal{W}_\nu,\ \boldsymbol{u}=U(\boldsymbol{x}),\ \boldsymbol{x}\in\mathcal{X}\right\}. \tag{EC.17}$$

Note that decision-makers typically specify the upper bound η on the convex fairness measure in (EC.17). If they chose a very small η, then (EC.17) could be infeasible.

Let us first consider the case when ν is an order-based fairness measure, i.e., $\mathcal{W}_\nu = \{\boldsymbol{w}\}$ is a singleton. Using the same proof techniques of Theorem 5, we can reformulate (EC.17) as

$$\begin{aligned}
\underset{\boldsymbol{x},\,\boldsymbol{u},\,\boldsymbol{\lambda}\in\mathbb{R}^N,\,\boldsymbol{\theta}\in\mathbb{R}^N}{\text{minimize}} \quad & f(\boldsymbol{u}) && \text{(EC.18a)}\\
\text{subject to} \quad & \mathbf{1}^\top(\boldsymbol{\lambda}+\boldsymbol{\theta}) \le \eta, && \text{(EC.18b)}\\
& \lambda_i + \theta_{i'} \ge u_i w_{i'}, \quad \forall i \in [N],\, i' \in [N], && \text{(EC.18c)}\\
& \boldsymbol{u} = U(\boldsymbol{x}),\, \boldsymbol{x} \in \mathcal{X}. && \text{(EC.18d)}
\end{aligned}$$

Second, when ν is a convex fairness measure, we propose a C&CG algorithm similar to Algorithm 1 to solve (EC.17). Algorithm 1 summarizes the steps of this algorithm. In this C&CG, we solve a master problem and a subproblem at each iteration. Specifically, at iteration j, we solve the following master problem:

$$\min_{\boldsymbol{x},\,\boldsymbol{u}} \left\{ f(\boldsymbol{u}) \,\middle|\, \nu_{\boldsymbol{w}}(\boldsymbol{u}) \le \eta,\, \forall \boldsymbol{w} \in \{\boldsymbol{w}^0, \ldots, \boldsymbol{w}^{j-1}\},\, \boldsymbol{u} = U(\boldsymbol{x}),\, \boldsymbol{x} \in \mathcal{X} \right\}, \tag{EC.19}$$

where $\{\boldsymbol{w}^0, \ldots, \boldsymbol{w}^{j-1}\} \subseteq \mathcal{W}_\nu$. Following the same techniques in the proof Proposition 5, we derive the following equivalent solvable reformulation of (EC.19):

$$\begin{aligned}
\underset{\boldsymbol{x},\,\boldsymbol{u},\,\boldsymbol{\lambda}\in\mathbb{R}^N,\,\boldsymbol{\theta}\in\mathbb{R}^N}{\text{minimize}} \quad & f(\boldsymbol{u}) && \text{(EC.20a)}\\
\text{subject to} \quad & \mathbf{1}^\top(\boldsymbol{\lambda}^k+\boldsymbol{\theta}^k) \le \eta, \quad \forall k \in [0, j-1]_{\mathbb{Z}}, && \text{(EC.20b)}\\
& \lambda_i^k + \theta_{i'}^k \ge u_i w_{i'}^k, \quad \forall i \in [N],\, i' \in [N],\, k \in [0, j-1]_{\mathbb{Z}}, && \text{(EC.20c)}\\
& \boldsymbol{u} = U(\boldsymbol{x}),\, \boldsymbol{x} \in \mathcal{X}. && \text{(EC.20d)}
\end{aligned}$$

Since only a set of weight vectors in $\mathcal{W}_\nu$ is considered, the master problem (EC.20) is a relaxation of the original problem (EC.17), i.e., the feasible region of (EC.17) is a subset of the feasible region of (EC.20). Thus, if (EC.20) is infeasible, then we can conclude that the original problem (EC.17) is also infeasible. Otherwise, with the optimal solution $\boldsymbol{u}^j$ from the master problem, we solve the same subproblem (11) and record the optimal solution $\boldsymbol{w}^j \in \mathcal{W}_\nu$ and optimal value D^j. Note that $D^j = \nu(\boldsymbol{u}^j)$ is the value of the convex fairness measure evaluated at $\boldsymbol{u}^j$. Thus, if $D^j > \eta$, the current solution $(\boldsymbol{x}^j, \boldsymbol{u}^j)$ is infeasible to the original problem (EC.17), and thus we enlarge the set of weight vectors and proceed to the next iteration. Otherwise, we terminate and conclude that $(\boldsymbol{x}^j, \boldsymbol{u}^j)$ is an optimal solution.

EC.6.2. Relative Convex Fairness Measure

Recall that, in Section EC.4, we define a relative convex fairness measure by $\rho(\boldsymbol{u}) = \nu(\boldsymbol{u})/g(\boldsymbol{u})$ for some convex fairness measure ν and continuous normalization function g. Thus, the corresponding fairness-promoting optimization model is

$$\min_{\boldsymbol{x},\,\boldsymbol{u}} \left\{ f(\boldsymbol{u}) \,\middle|\, \rho(\boldsymbol{u}) = \nu_{\boldsymbol{w}}(\boldsymbol{u})/g(\boldsymbol{u}) \le \eta,\, \forall \boldsymbol{w} \in \mathcal{W}_\nu,\, \boldsymbol{u} = U(\boldsymbol{x}),\, \boldsymbol{x} \in \mathcal{X} \right\}. \tag{EC.21}$$

Algorithm 1: A column-and-constraint generation (C&CG) method to solve (EC.17)

Initialization: Set $LB = 0$, $UB = \infty$, $\varepsilon > 0$, $j = 1$, $\boldsymbol{w}^0 \in \mathcal{W}_\nu$.

1. Master problem.

Solve master problem (EC.20) with weights $\{\boldsymbol{w}^0, \ldots, \boldsymbol{w}^{j-1}\}$.

If the master problem (EC.20) is infeasible, terminate and return the infeasibility of (EC.17).

Otherwise, record the optimal solution $(\boldsymbol{x}^j, \boldsymbol{u}^j)$.

2. Subproblem. Solve subproblem (11) for fixed $\boldsymbol{u} = \boldsymbol{u}^j$.

Record the optimal solution $\boldsymbol{w}^j$ and value D^j.

If $D^j \le \eta$, terminate and return the optimal solution $(\boldsymbol{x}^j, \boldsymbol{u}^j)$.

3. Scenario set enlargement.

Update $j \leftarrow j + 1$ and go back to step 1.

---

for some $\eta > 0$. From Theorem EC.1, the normalization function g satisfies $g \ge \nu \ge 0$. In particular, if $g(\boldsymbol{u}) = 0$, we must have $\nu(\boldsymbol{u}) = 0$. Therefore, we can reformulate (EC.21) into

$$\min_{\boldsymbol{x},\boldsymbol{u}} \left\{ f(\boldsymbol{u}) \,\middle|\, \nu_{\boldsymbol{w}}(\boldsymbol{u}) \le \eta g(\boldsymbol{u}),\, \forall \boldsymbol{w} \in \mathcal{W}_\nu,\, \boldsymbol{u} = U(\boldsymbol{x}),\, \boldsymbol{x} \in \mathcal{X} \right\}, \tag{EC.22}$$

which takes the same form of (EC.17). Thus, we can apply similar solution approaches as in EC.6.1 to solve (EC.22).

EC.7. Additional Computational Results for the Fair Facility Location Problem

In this section, we provide additional results comparing the computational performance of our proposed unified reformulations and solution methods and traditional techniques using a set of larger instances of the fair facility location problem introduced in Section 8.1. Specifically, we follow the experimental settings discussed in Section 8.1 to generate five random instances of each of following combinations of problem parameters: (a) $(|I|, |J|) = (60, 40)$ with $p \in \{13, 10, 8\}$, (b) $(|I|, |J|) = (80, 20)$ with $p \in \{7, 5, 4\}$, (c) $(|I|, |J|) = (80, 35)$ with $p \in \{7, 5, 4\}$. These instances reflect realistic scenarios where the number of customer locations is larger than that of potential facility locations, i.e., $|I| > |J|$.

Tables EC.4–EC.5 present the minimum (min), average (avg), and maximum (max) solution time (in seconds) of the generated instances solved using our formulations and traditional ones for the problem with the MAD and Gini deviation, respectively. The observations are similar to those in Section 8.1. First, using our unified reformulation and the proposed C&CG, we can solve all the generated instances significantly faster than the classical reformulation techniques. Second, solution time increases as γ decreases, i.e., when more emphasis is placed on fairness. This suggests that the problem becomes more challenging when seeking

Table EC.4 Solution time (in seconds) over five randomly generated instances with $\gamma \in \{0.3, 0.2\}$ using the absolute deviation from mean (MAD). *Note:* Solution times with '>' indicate that one or more instances cannot be solved within the imposed two-hour time limit.

$\lvert I\rvert = 40$, $\lvert J\rvert = 60$		$\gamma = 0.3$			$\gamma = 0.2$		
		min	avg	max	min	avg	max
$p = 13$	C&CG Method	1	7	11	2	567	2340
	Traditional Reformulation (16)	1	40	63	6	>5802	>7200
		min	avg	max	min	avg	max
$p = 10$	C&CG Method	4	24	50	86	867	3263
	Traditional Reformulation (16)	6	>1490	>7200	405	>4755	>7200
		min	avg	max	min	avg	max
$p = 8$	C&CG Method	4	18	35	15	853	3377
	Traditional Reformulation (16)	4	>1477	>7200	153	>4667	>7200
$\lvert I\rvert = 80$, $\lvert J\rvert = 20$		$\gamma = 0.3$			$\gamma = 0.2$		
		min	avg	max	min	avg	max
$p = 7$	C&CG Method	1	3	8	7	66	162
	Traditional Reformulation (16)	1	5	15	4	>1613	>7200
		min	avg	max	min	avg	max
$p = 5$	C&CG Method	1	2	5	16	30	50
	Traditional Reformulation (16)	1	5	14	17	110	216
		min	avg	max	min	avg	max
$p = 4$	C&CG Method	1	2	5	13	22	34
	Traditional Reformulation (16)	2	4	12	15	114	251
$\lvert I\rvert = 80$, $\lvert J\rvert = 35$		$\gamma = 0.3$			$\gamma = 0.2$		
		min	avg	max	min	avg	max
$p = 7$	C&CG Method	2	9	30	28	756	2055
	Traditional Reformulation (16)	1	14	44	15	>3832	>7200
		min	avg	max	min	avg	max
$p = 5$	C&CG Method	1	5	14	14	551	1458
	Traditional Reformulation (16)	1	45	144	218	>3673	>7200
		min	avg	max	min	avg	max
$p = 4$	C&CG Method	1	9	25	15	654	2064
	Traditional Reformulation (16)	1	11	36	182	>3468	>7200

fairer decisions. Notably, the solution times of traditional approaches are substantially longer than our solution times when γ is smaller. Consider, for example, the largest instances with $(|I|, |J|) = (80, 35)$. We can solve all the instances using our unified reformulations and C&CG method within two hours. Specifically, the average solution using our approach ranges from 9 minutes to 13 minutes for the MAD with $\gamma = 0.2$ and 12 minutes to 62 minutes for the Gini deviation and $\gamma = 0.3$. In contrast, we are unable to solve a large number of these generated instances within two hours using traditional reformulations (16) and (18). For instances that were not solved by traditional formulations within two hours, the relative optimality gap reported by the solver at termination ranges from 2.4% to 15.5%. Third, the gap in solution time between the two approaches generally widens as p decreases, i.e., as the problem becomes more constrained. These results further underscore the computational efficiency of our approach compared with traditional methods.

Table EC.5 Solution time (in seconds) over five randomly generated instances with $\gamma \in \{0.4, 0.3\}$ using the Gini deviation. *Note:* Solution times with '$>$' indicate that one or more instances cannot be solved within the imposed two-hour time limit.

$\lvert I\rvert=40$, $\lvert J\rvert=60$		$\gamma=0.4$			$\gamma=0.3$		
		min	avg	max	min	avg	max
$p=13$	Our Reformulation (19)	39	49	62	79	584	1461
	Traditional Reformulation (18)	4	72	271	53	808	1804
		min	avg	max	min	avg	max
$p=10$	Our Reformulation (19)	39	64	103	81	1569	4149
	Traditional Reformulation (18)	14	137	269	292	$>$3020	$>$7200
		min	avg	max	min	avg	max
$p=8$	Our Reformulation (19)	17	108	207	72	936	1829
	Traditional Reformulation (18)	16	309	820	278	$>$4183	$>$7200
$\lvert I\rvert=80$, $\lvert J\rvert=20$		$\gamma=0.4$			$\gamma=0.3$		
		min	avg	max	min	avg	max
$p=7$	Our Reformulation (19)	59	130	269	795	1800	3557
	Traditional Reformulation	14	189	464	784	$>$2860	$>$7200
		min	avg	max	min	avg	max
$p=5$	Our Reformulation (19)	56	225	742	406	1259	1982
	Traditional Reformulation (18)	38	162	559	1051	1858	2916
		min	avg	max	min	avg	max
$p=4$	Our Reformulation (19)	52	87	134	127	336	745
	Traditional Reformulation (18)	45	113	367	602	1076	2249
$\lvert I\rvert=80$, $\lvert J\rvert=35$		$\gamma=0.4$			$\gamma=0.3$		
		min	avg	max	min	avg	max
$p=7$	Our Reformulation (19)	78	353	1282	1988	3694	5918
	Traditional Reformulation (18)	65	418	1790	5450	$>$6683	$>$7200
		min	avg	max	min	avg	max
$p=5$	Our Reformulation (19)	70	156	319	222	1863	5426
	Traditional Reformulation (18)	44	648	1951	1107	$>$3970	$>$7200
		min	avg	max	min	avg	max
$p=4$	Our Reformulation (19)	70	126	217	143	740	1423
	Traditional Reformulation	41	303	1106	490	$>$4558	$>$7200

EC.8. Comparison with Lan et al. (2010)

In this section, we compare our proposed axiomatic approach for the class of convex fairness measures (see Section 5) with Lan et al. (2010)'s axiomatic approach for a class of fairness measures in resource allocation. Specifically, Lan et al. (2010) considered resource allocation settings where one wants to quantify the degree of fairness associated with a given allocation vector $\boldsymbol{u} \in \mathbb{R}_+^N$ (u_i is the resource allocated to subject i). They construct a class of fairness measures $\{h_N\}_{N\in\mathbb{N}}$ for this setting based on a set of five axioms and a generator function g, where g must be an increasing and continuous function that leads to a well-defined mean function. To facilitate the discussion, we first summarize the axiomatic characterization of Lan et al. (2010)'s class of fairness measures in Theorem EC.4 (we slightly modified Lan et al. (2010)'s notation to align it with our notation for consistency).

THEOREM EC.4 **(Theorems 1 and 2 of Lan et al. (2010)).** *There exists a unique set of fairness measures $\{h_N : \mathbb{R}_+^N \to \mathbb{R}\}_{N\in\mathbb{N}}$ defined using a generator function g such that $\{h_N\}_{N\in\mathbb{N}}$ satisfy the following axioms:*

1. *Continuity: h_N is continuous on $\mathbb{R}_+^N$;*
2. *Homogeneity: $h_N(\alpha\boldsymbol{u}) = h_N(\boldsymbol{u})$ for all $\alpha > 0$ with $|h_1(u)| = 1$ for all $u > 0$;*
3. *Asymptotic Saturation: $\lim_{N\to\infty} h_{N+1}(\mathbf{1})/h_N(\mathbf{1}) = 1$;*
4. *Irrelevance of Partition: for any $\boldsymbol{u} = (\boldsymbol{u}^1, \boldsymbol{u}^2) \in \mathbb{R}^N$ with $\boldsymbol{u}^1 \in \mathbb{R}^{N_1}$ and $\boldsymbol{u}^2 \in \mathbb{R}^{N_2}$,*

$$h_N(\boldsymbol{u}) = h_2(\mathbf{1}^\top\boldsymbol{u}^1,\, \mathbf{1}^\top\boldsymbol{u}^2)\cdot g^{-1}\Big(s_1\cdot g\big(h_{N_1}(\boldsymbol{u}_1)\big) + s_2\cdot g\big(h_{N_2}(\boldsymbol{u}_2)\big)\Big)$$

   for some $\{s_1, s_2\} \subset \mathbb{R}$ with $s_1 + s_2 = 1$, and continuous and strictly monotonic function g such that $m = g^{-1}\Big(s_1\cdot g\big(h_{N_1}(\boldsymbol{u}_1)\big) + s_2\cdot g\big(h_{N_2}(\boldsymbol{u}_2)\big)\Big)$ is a mean function of $\{h_{N_1}(\boldsymbol{u}^1), h_{N_2}(\boldsymbol{u}^2)\}$;
5. *Monotonicity: $h_2(u_1, u_2)$ is monotonically decreasing with $|u_1 - u_2|$.*

As discussed in Lan et al. (2010), from any function $g(\cdot)$ satisfying Axiom 4, Axioms 1–5 generate a unique set $\{h_N\}_{N\in\mathbb{N}}$. Such h_N is a well-defined fairness measure if it also satisfies Axioms 1–5. Hence, to derive a closed-form mathematical expression of h_N that can be used in optimization models, one has to choose a suitable generator function g and specify the values of $\{s_1, s_2\}$ in Axiom 4 such that h_N satisfies Axioms 1–5. Lan et al. (2010) focused on the special case where $g(\cdot)$ is a power function and derived the mathematical expression of h_N as shown in Theorem EC.5.

THEOREM EC.5 **(Theorem 4 of Lan et al. (2010)).** *If the generation function g is chosen as the power function $g(y) = |y|^\beta$ with parameter β and coefficients $s_i = (\mathbf{1}^\top\boldsymbol{u}^i)^\rho/[(\mathbf{1}^\top\boldsymbol{u}^1)^\rho + (\mathbf{1}^\top\boldsymbol{u}^2)^\rho]$ for $\rho \neq 0$ and $i \in \{1, 2\}$, then Axioms 1–5 define a unique family of fairness measures as follows:*

$$h_N(\boldsymbol{u}) = \operatorname{sgn}(1-\beta r)\left[\sum_{i=1}^N \left(\frac{u_i}{\sum_{j=1}^N u_j}\right)^{1-\beta r}\right]^{\frac{1}{\beta}},$$

where $r = (1-\rho)/\beta$ *and* $\operatorname{sgn}(\cdot)$ *is the sign function, i.e.,* $\operatorname{sgn}(a) = 1$ *if* $a > 0$ *and* $\operatorname{sgn}(a) = -1$ *if* $a < 0$*. In particular, if* r *is chosen to* 1*, then*

$$h_N(\boldsymbol{u}) = \operatorname{sgn}(1-\beta)\left[\sum_{i=1}^{N}\left(\frac{u_i}{\sum_{j=1}^{N} u_j}\right)^{1-\beta}\right]^{\frac{1}{\beta}}.$$

We are now ready to discuss the differences between our work and that of Lan et al. (2010). First, we emphasize that our proposed unified framework for convex fairness measures and Lan et al. (2010)'s axiomatic approach to fairness measures consider different classes of fairness measures and different settings, and thus, there is no direct relationship between the two. As mentioned earlier, Lan et al. (2010)'s axiomatic approach is based on the specific context of network resource allocation. Specifically, the outcome vector $\boldsymbol{u} \in \mathbb{R}_+^N$ in Lan et al. (2010) can be interpreted as the non-negative resource allocation decision. In contrast, we propose a unified framework for a class of convex fairness measures suitable for various optimization contexts, and thus, it applies to a broader range of applications. Moreover, while Lan et al. (2010)'s approach is limited to settings where the outcome vector is non-negative, our unified framework for convex fairness measures is applicable to general settings where the entries of the outcome vector can be negative.

Second, Lan et al. (2010) considers fairness in a relative sense, i.e., their class of fairness measures consists of relative measures (more on this below). In contrast, we consider convex fairness measures in an absolute sense and also introduce their relative counterpart. Third, we observe the following differences in the mathematical properties and interpretations between our convex fairness measure ν and the fairness measure h_N:

- The fairness measure h_N may assume negative value depending on the choice of the generator function (see Theorem EC.5). In contrast, the convex fairness measure ν satisfies Axiom N, and thus, is always non-negative. As discussed in Section 3, Axiom N has the following intuitive interpretation: ν equals zero implies perfect fairness or equality.
- A larger value of h_N corresponds to a fairer distribution of outcomes. In contrast, a smaller value of ν means a fairer distribution of outcomes. In particular, while the equal distribution $\boldsymbol{u} = \mathbf{1}$ maximizes h_N, i.e., $h_N(\mathbf{1}) = \max_{\boldsymbol{u} \in \mathbb{R}_+^N} h_N(\boldsymbol{u})$ (see Corollary 2 in Lan et al. 2010), it minimizes ν, i.e., $\nu(\mathbf{1}) = \min_{\boldsymbol{u} \in \mathbb{R}^N} \nu(\boldsymbol{u}) = 0$, both implying perfect fairness. Hence, by construction, the fairness measure h_N is Schur concave (see Theorem 3 in Lan et al. 2010), and our convex fairness measure ν is Schur convex (Axiom SCV).

Finally, Lan et al. (2010) considers a slightly different set of axioms to define their class of relative fairness measures. In the following, we discuss similarities and differences between Lan et al. (2010)'s axioms and some of those we use to define the class of convex fairness measures. (Recall that the class of convex fairness measures is defined by Axioms C, N, S, TI, PH, and CV, and its relative counterpart is defined by Axioms NR, S, SCV, and SI.)

1. *Continuity.* This axiom is the same as Axiom C.
2. *Homogeneity.* This axiom is equivalent to scale invariance, i.e., Axiom SI, used in the definition of the relative convex fairness measure (see Definition EC.1). As discussed in EC.4, this axiom characterizes relative fairness measures and thus implies that Lan et al. (2010)'s class of fairness measures $\{h_N\}_{N\in\mathbb{N}}$ is defined in the relative sense (see, e.g., Theorem EC.5). In contrast, we define the class of convex fairness measures in the absolute sense and also derive its relative counterpart (which satisfies Axiom SI) in EC.4.
3. *Asymptotic Saturation.* As pointed out in Lan et al. (2010), this is a technical condition and not a known axiom for fairness measures. They adopt this condition to guarantee the uniqueness of the set of fairness measures defined in Theorem EC.4.
4. *Irrelevance of Partition.* This is another axiom specific for Lan et al. (2010)'s class of fairness measures. It allows one to compute the value of the fairness measure of a high-dimensional outcome vector (i.e., with a large number of subjects) recursively from that of low-dimensional outcome vectors (i.e., with smaller numbers of subjects). As Chen and Hooker (2023) pointed out, it is still unclear "how one might assess whether the rather abstract axiom of partition is appropriate for a particular practical application."
5. *Monotonicity.* This axiom applies exclusively to the case where there are two subjects, i.e., $N = 2$. Specifically, it indicates that h_2 decreases as the absolute difference $|u_1 - u_2|$ increases. In other words, a higher value of $|u_1 - u_2|$ implies a more unfair distribution of outcomes. Note that when there are two subjects, our convex fairness measure ν can be represented as $c|u_1 - u_2|$, where c is a positive constant (see discussions after Theorem 4). Hence, $\nu : \mathbb{R}^2 \to \mathbb{R}_+$ is monotonoically increasing with $|u_1 - u_2|$. It follows that $\nu(\boldsymbol{u}) = c|u_1 - u_2|$ satisfies the notion of monotonicity.

EC.9. Envy-Based Fairness Measures

In this section, we analyze a general class of envy-based fairness measures and derive conditions under which a subset of them belongs to the class of convex fairness measures. Without loss of generality, we consider u_i as the positive impact on subject $i \in [N]$, i.e., a larger value of u_i is preferred.

We first introduce a general and classical class of envy-based fairness measures widely adopted in the literature (Aleksandrov et al. 2019, Boiney 1995, Feldman and Kirman 1974, Tan et al. 2023). Let $V_i : \mathbb{R} \to \mathbb{R}$ be the utility function of individual $i \in [N]$. That is, $V_i(y) \in \mathbb{R}$ is the utility or value that individual i receives from a given impact $y \in \mathbb{R}$. We define the envy that i has towards j for a given impact vector $\boldsymbol{u} \in \mathbb{R}^N$ as

$$e_{ij}(\boldsymbol{u}) := \big(V_i(u_j) - V_i(u_i)\big)_+ = \begin{cases} 0, & \text{if } V_i(u_i) \geq V_i(u_j), \\ V_i(u_j) - V_i(u_i), & \text{if } V_i(u_i) < V_i(u_j). \end{cases}$$

Then, we define the envy-based fairness measure $E:\mathbb{R}^N \to \mathbb{R}$ as the total envy:

$$E(\boldsymbol{u}) = \sum_{i=1}^{N}\sum_{j=1}^{N} e_{ij}(\boldsymbol{u}) = \sum_{i=1}^{N}\sum_{j=1}^{N}\big(V_i(u_j) - V_i(u_i)\big)_+. \tag{EC.23}$$

In Theorem EC.6, we identify a set of envy-based fairness measures of the form (EC.23) that belong to the class of convex fairness measures.

THEOREM EC.6. *Assume that the utility function $V_i:\mathbb{R}\to\mathbb{R}$ is continuous and non-decreasing for all $i\in[N]$. The total envy $E:\mathbb{R}^N\to\mathbb{R}$ defined in* (EC.23) *is a convex fairness measure if and only if $V_i(y) = cy + V_i(0)$ for some $c>0$ and for all $i\in[N]$, i.e.,*

$$E(\boldsymbol{u}) = c\sum_{i=1}^{N}\sum_{j=1}^{N}\big(u_j - u_i\big)_+ = c\sum_{i=1}^{N}\sum_{j=i+1}^{N}\big|u_j - u_i\big|. \tag{EC.24}$$

Proof. First, suppose that $V_i(y) = cy + V(0)$. We show that $E(\boldsymbol{u})$ defined in (EC.24) is equivalent to the Gini deviation $\phi^{\text{Gini}}(\boldsymbol{u}) = \sum_{i=1}^{N}\sum_{j=1}^{N}\big|u_j - u_i\big|$ and thus is a convex fairness measure. From the definition of $E(\boldsymbol{u})$, we have

$$\begin{aligned}
E(\boldsymbol{u}) = c\sum_{i=1}^{N}\sum_{j=1}^{N}\big(u_j - u_i\big)_+ &= c\sum_{i=1}^{N}\sum_{j=1}^{i-1}\big(u_j - u_i\big)_+ + c\sum_{i=1}^{N}\sum_{j=i+1}^{N}\big(u_j - u_i\big)_+ \\
&= c\sum_{j=1}^{N}\sum_{i=j+1}^{N}\big(u_j - u_i\big)_+ + c\sum_{i=1}^{N}\sum_{j=i+1}^{N}\big(u_j - u_i\big)_+ \\
&= c\sum_{i=1}^{N}\sum_{j=i+1}^{N}\Big[\big(u_i - u_j\big)_+ + \big(u_j - u_i\big)_+\Big] \\
&= c\sum_{i=1}^{N}\sum_{j=i+1}^{N}\big|u_j - u_i\big| \qquad \text{(EC.25)} \\
&= \frac{c}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\big|u_j - u_i\big| = \frac{c}{2}\phi^{\text{Gini}}(\boldsymbol{u}),
\end{aligned}$$

where (EC.25) follows from the equality $(a)_+ + (-a)_+ = |a|$ for any $a\in\mathbb{R}$. It follows that $E(\boldsymbol{u})$ in (EC.24) is a convex fairness measure.

Now, suppose that E defined in (EC.23) is a convex fairness measure. We will show that $V_i(y) = cy + V(0)$ for some $c\in\mathbb{R}$ and for all $i\in[N]$ in two steps.

Step 1. Consider any $(a,b)\in\mathbb{R}^2$. Let $\boldsymbol{u} = (a,b,b,\ldots,b)^\top \in\mathbb{R}^N$ and its permutation $\boldsymbol{u}' = (b,a,b,\ldots,b)^\top\in\mathbb{R}^N$. By Axiom S, we have $E(\boldsymbol{u}) = E(\boldsymbol{u}')$. This implies that

$$(N-1)\cdot\Big(V_1(b) - V_1(a)\Big)_+ + \Big(V_2(a) - V_2(b)\Big)_+ = (N-1)\cdot\Big(V_2(b) - V_2(a)\Big)_+ + \Big(V_1(a) - V_1(b)\Big)_+ \tag{EC.26}$$

for any $(a,b)\in\mathbb{R}^2$. Consider the following three cases.

- Suppose that $V_1(b) - V_1(a) > 0$. Note that if $V_2(a) - V_2(b) > 0$, the left-hand-side of (EC.26) is strictly greater than zero while the right-hand side of (EC.26) equals zero, which is a contradiction. Therefore, we must have $V_2(a) - V_2(b) \le 0$. It then follows from (EC.26) that $V_1(b) - V_1(a) = V_2(b) - V_2(a)$.
- Suppose that $V_1(b) - V_1(a) < 0$. Note that if $V_2(a) - V_2(b) < 0$, the left-hand-side of (EC.26) equals zero while the right-hand side of (EC.26) is strictly greater than zero, which is a contradiction. Therefore, we must have $V_2(a) - V_2(b) \ge 0$. It then follows from (EC.26) that $V_2(a) - V_2(b) = V_1(a) - V_1(b)$.
- Suppose that $V_1(b) - V_1(a) = 0$. It then follows from (EC.26) that $\big(V_2(a) - V_2(b)\big)_+ = (N-1) \cdot \big(V_2(b) - V_2(a)\big)_+$, which implies $V_2(a) - V_2(b) = 0$. Thus, we also have $V_2(a) - V_2(b) = V_1(a) - V_1(b)$.

Combining the three cases, we have that (EC.26) implies $V_1(b) - V_1(a) = V_2(b) - V_2(a)$, or equivalently, $\big(V_1 - V_2\big)(b) = \big(V_1 - V_2\big)(a)$ for any $(a,b) \in \mathbb{R}^2$. This shows that $V_1 - V_2$ is a constant function, i.e., $V_1 - V_2 \equiv \tau_{12}$ for some $\tau_{12} \in \mathbb{R}$. Repeating the same argument, one can show that for any $i \in [N]$ and $j \in [N] \setminus \{i\}$, we have $V_i - V_j \equiv \tau_{ij}$ for some $\tau_{ij} \in \mathbb{R}$. That is, we have $V_i \equiv V + \tau_i$ for some function $V : \mathbb{R} \to \mathbb{R}$ and $\tau_i \in \mathbb{R}$ for all $i \in [N]$. Note that since V_i is non-decreasing by assumption, V is also non-decreasing. We can now write the fairness measure E in (EC.23) as

$$
\begin{aligned}
E(\boldsymbol{u}) &= \sum_{i=1}^{N}\sum_{j=1}^{N}\big(V(u_j) - V(u_i)\big)_+ \\
&= \sum_{i=1}^{N}\sum_{j=1}^{i-1}\big(V(u_j) - V(u_i)\big)_+ + \sum_{i=1}^{N}\sum_{j=i+1}^{N}\big(V(u_j) - V(u_i)\big)_+ && \text{(EC.27)} \\
&= \sum_{i=1}^{N}\sum_{j=i+1}^{N}\big|V(u_i) - V(u_j)\big|, && \text{(EC.28)}
\end{aligned}
$$

where (EC.27) follows from the equality $(a)_+ + (-a)_+ = |a|$ for any $a \in \mathbb{R}$.

Step 2. Consider any $(a,b) \in \mathbb{R}^2$. Let $\boldsymbol{u} = (a, b, b, \ldots, b)^\top \in \mathbb{R}^N$. From Axiom TI, we have $E(\boldsymbol{u} + \alpha \mathbf{1}) = E(\boldsymbol{u})$ for any $\alpha \in \mathbb{R}$. It follows from (EC.28) that

$$
\big|V(a+\alpha) - V(b+\alpha)\big| = \big|V(a) - V(b)\big| \tag{EC.29}
$$

for any $(a,b) \in \mathbb{R}^2$ and $\alpha \in \mathbb{R}$. By setting $\alpha = -b$, (EC.29) implies that

$$
\big|V(a-b) - V(0)\big| = \big|V(a) - V(b)\big| \tag{EC.30}
$$

for any $(a,b) \in \mathbb{R}^2$. Consider the following two cases.

- Suppose that $a \le b$, or equivalently, $a - b \le 0$. Since V is non-decreasing, we have $V(a) \le V(b)$ and $V(a-b) \le V(0)$. Therefore, (EC.30) reduces to $V(0) - V(a-b) = V(b) - V(a)$.
- Suppose that $a > b$, or equivalently, $a - b > 0$. Since V is non-decreasing, we have $V(a) \ge V(b)$ and $V(a-b) \ge V(0)$. Therefore, (EC.30) reduces to $V(a-b) - V(0) = V(a) - V(b)$.

Combining the two cases, we have that (EC.30) implies $V(a-b) = V(a) - V(b) + V(0)$ for any $(a,b) \in \mathbb{R}^2$. Note that this is a one-dimensional generalized Cauchy functional equation. Since V is continuous, it follows from Corollary 2.4.1 of Castillo and Ruiz-Cobo (1992) that $V(y) = cy + V(0)$ for some $c \in \mathbb{R}$, and thus $V_i(y) = V(y) + \tau_i = cy + V(0) + \tau_i$ for all $i \in [N]$. This shows that $V_i(y) = cy + V_i(0)$ for all $i \in [N]$. Finally, since V_i is non-decreasing, we have $c \geq 0$. Plugging $V_i(y) = cy + V_i(0)$ for all $i \in [N]$ in (EC.23), we obtain $E(\boldsymbol{u}) = c\sum_{i=1}^{N}\sum_{j=i+1}^{N} \left|u_j - u_i\right|$ as in (EC.24). By Axiom N, $E(\boldsymbol{u}) = 0$ if and only if $\boldsymbol{u} = \alpha\mathbf{1}$ for any $\alpha \in \mathbb{R}$, which implies that the parameter c should be strictly greater than zero. This completes the proof. $\square$

We remark the following on the utility function V_i. First, the assumption that V_i is non-decreasing is mild. In particular, it is reasonable and common to expect that the utility individual i receives, denoted as $V_i(y)$, is non-decreasing with the (positive) impact y, i.e., $V_i(y_1) \geq V_i(y_2)$ if $y_1 \geq y_2$. Second, several studies have adopted affine utility functions in optimization models of the form (1) (see, e.g., Blanco et al. 2024, Chanta et al. 2011, 2014, Espejo et al. 2009). In such settings, the function U can be interpreted as the utility function, i.e., $\boldsymbol{u} = U(\boldsymbol{x})$ computes the utility of the subjects. Finally, we note that for the envy-based fairness measure, $E(\boldsymbol{u})$, to be a convex fairness measure and satisfy Axiom S—which stipulates that the perceived degree of fairness should not depend on the identity of individuals—the proportional constant c in the utility function $V_i(y) = cy + V_i(0)$ must be identical for all $i \in [N]$. That is, the additional utility or value an individual receives due to a unit increase in y is the same for all individuals.

EC.10. Fairness-Promoting Optimization Models Based on the Rawlsian Principle

In this section, we discuss how our unified framework can be leveraged to analyze fairness-promoting problems incorporating the Rawlsian principle (Chen and Hooker 2023, Karsu and Morton 2015, Shehadeh and Snyder 2023). Let us consider the following Rawlsian-based fairness-promoting optimization model:

$$\max_{\boldsymbol{x},\,\boldsymbol{u}} \left\{ \min_{i\in[N]}\{u_i\} \;\middle|\; \boldsymbol{u} = U(\boldsymbol{x}),\, \boldsymbol{x} \in \mathcal{X} \right\}. \tag{EC.31}$$

Formulation (EC.31) finds $\boldsymbol{x}$ and $\boldsymbol{u}$ that maximize the minimum (positive) outcome across the N subjects. Note that we can equivalently rewrite (EC.31) as follows:

$$\min_{\boldsymbol{x},\,\boldsymbol{u}} \left\{ -\min_{i\in[N]}\{u_i\} \;\middle|\; \boldsymbol{u} = U(\boldsymbol{x}),\, \boldsymbol{x} \in \mathcal{X} \right\} = \min_{\boldsymbol{x},\,\boldsymbol{u}} \left\{ \max_{i\in[N]}\{-u_i\} \;\middle|\; \boldsymbol{u} = U(\boldsymbol{x}),\, \boldsymbol{x} \in \mathcal{X} \right\}. \tag{EC.32}$$

Letting $\overline{u} = N^{-1}(\mathbf{1}^\top \boldsymbol{u})$ be the mean of $\boldsymbol{u}$, we can rewrite the objective function in (EC.32) as

$$\max_{i\in[N]}\{-u_i\} = -\overline{u} + \max_{i\in[N]}\{-u_i + \overline{u}\} = f(\boldsymbol{u}) + \nu(\boldsymbol{u}),$$

where $f(\boldsymbol{u}) = -\overline{u}$ and

$$\nu(\boldsymbol{u}) = \max_{i\in[N]}\{-u_i + \overline{u}\} = \overline{u} - \min_{i\in[N]}\{u_i\}. \tag{EC.33}$$

Note that minimizing $f(\boldsymbol{u}) = -\overline{u}$ is equivalent to maximizing the average outcome. Thus, $f(\boldsymbol{u})$ represents an efficiency measure. It is straightforward to verify that ν in (EC.33) satisfies Axioms C, N, S, TI, PH, and CV. Thus, ν is a convex fairness measure (see Definition 2). Therefore, we can rewrite the objective of the Rawlsian optimization model in the form (EC.32) as the sum of an efficiency measure f and a convex fairness measure ν. Hence, one can use our proposed unified framework for solving and analyzing the Rawlsian optimization problem of the form (EC.31).

References

Ahmadi-Javid A, Seyedi P, Syam SS (2017) A survey of healthcare facility location. *Computers & Operations Research* 79:223–263.

Aleksandrov M, Ge C, Walsh T (2019) Fair division minimizing inequality. *Progress in Artificial Intelligence: 19th EPIA Conference on Artificial Intelligence, EPIA 2019, Vila Real, Portugal, September 3–6, 2019, Proceedings, Part II 19*, 593–605 (Springer).

Bertsekas D (2009) *Convex Optimization Theory* (Athena Scientific).

Bertsekas D (2015) *Convex Optimization Algorithms* (Athena Scientific).

Blanco V, Marín A, Puerto J (2024) Intra-facility equity in discrete and continuous p-facility location problems. *Computers & Operations Research* 162:106487.

Blanco V, Puerto J, Ben-Ali SEH (2016) Continuous multifacility ordered median location problems. *European Journal of Operational Research* 250(1):56–64.

Boiney LG (1995) When efficient is insufficient: Fairness in decisions affecting a group. *Management Science* 41(9):1523–1537.

Castillo E, Ruiz-Cobo MR (1992) *Functional Equations and Modelling in Science and Engineering*, volume 161 (CRC Press).

Chakravarty SR (1999) Measuring inequality: The axiomatic approach. *Handbook of Income Inequality Measurement* 163–186.

Chanta S, Mayorga ME, Kurz ME, McLay LA (2011) The minimum p-envy location problem: A new model for equitable distribution of emergency resources. *IIE Transactions on Healthcare Systems Engineering* 1(2):101–115.

Chanta S, Mayorga ME, McLay LA (2014) The minimum p-envy location problem with requirement on minimum survival rate. *Computers & Industrial Engineering* 74:228–239.

Chen VX, Hooker J (2023) A guide to formulating fairness in an optimization model. *Annals of Operations Research* 1–39.

Cowell F (2011) *Measuring Inequality* (Oxford University Press).

Donaldson D, Weymark JA (1980) A single-parameter generalization of the Gini indices of inequality. *Journal of Economic Theory* 22(1):67–86.

Espejo I, Marin A, Puerto J, Rodríguez-Chía AM (2009) A comparison of formulations and solution methods for the minimum-envy location problem. *Computers & Operations Research* 36(6):1966–1981.

Feldman A, Kirman A (1974) Fairness and envy. *The American Economic Review* 64(6):995–1005.

Gini C (1912) *Variabilità e Mutabilità: Contributo allo Studio delle Distribuzioni e delle Relazioni Statistiche* (Cuppini, Bologna).

Hoover EM (1936) The measurement of industrial localization. *The Review of Economic Statistics* 162–171.

Karsu Ö, Morton A (2015) Inequity averse optimization in operational research. *European Journal of Operational Research* 245(2):343–359.

Lan T, Kao D, Chiang M, Sabharwal A (2010) An axiomatic theory of fairness in network resource allocation. *2010 Proceedings IEEE INFOCOM*, 1–9, URL `http://dx.doi.org/10.1109/INFCOM.2010.5461911`.

Marshall AW, Olkin I, Arnold BC (2011) *Inequalities: Theory of Majorization and its Applications* (Springer).

Martello S, Toth P (1987) Linear assignment problems. *North-Holland Mathematics Studies*, volume 132, 259–282 (Elsevier).

Marynissen J, Demeulemeester E (2019) Literature review on multi-appointment scheduling problems in hospitals. *European Journal of Operational Research* 272(2):407–419.

Mehran F (1976) Linear measures of income inequality. *Econometrica: Journal of the Econometric Society* 805–809.

Mesa JA, Puerto J, Tamir A (2003) Improved algorithms for several network location problems with equality measures. *Discrete Applied Mathematics* 130(3):437–448.

Mussard S, Mornet P (2019) A note on α-Gini measures. *Review of Income and Wealth* 65(3):675–682.

Nickel S, Puerto J (2006) *Location Theory: A Unified Approach* (Springer Science & Business Media).

Pinedo ML (2016) *Scheduling: Theory, Algorithms, and Systems* (Springer).

Rockafellar RT (1970) *Convex Analysis*, volume 18 (Princeton University Press).

Rodríguez-Chía AM, Nickel S, Puerto J, Fernandez FR (2000) A flexible approach to location problems. *Mathematical Methods of Operations Research* 51:69–89.

Shehadeh KS, Snyder LV (2023) Equity in stochastic healthcare facility location. *Uncertainty in Facility Location Problems*, 303–334 (Springer).

Stancu-Minasian IM (1997) *Fractional Programming: Theory, Methods and Applications*, volume 409 (Springer Dordrecht).

Stępniak C (2007) An effective characterization of Schur-convex functions with applications. *Journal of Convex Analysis* 14(1):103–108.

Tan M, Ren Y, Pan R, Wang L, Chen J (2023) Fair and efficient electric vehicle charging scheduling optimization considering the maximum individual waiting time and operating cost. *IEEE Transactions on Vehicular Technology* .

Yager RR (1988) On ordered weighted averaging aggregation operators in multicriteria decisionmaking. *IEEE Transactions on Systems, Man, and Cybernetics* 18(1):183–190.

Zheng B (2007) Unit-consistent decomposable inequality measures. *Economica* 74(293):97–111.