Discussion Paper Series

Institute-Wide Joint Research Project
Methodology of Social Sciences: Measuring Phenomena and Values

Why Not Prospect Utilitarianism Instead of Sufficientarianism? A Formal Comparison in Terms of Continuity and Lifeboat Cases

Susumu Cato
University of Tokyo

Hun Chung
Waseda University

August 3, 2023

E-23-003

Institute of Social Science, University of Tokyo
[https://web.iss.u-tokyo.ac.jp/methodology/en/dp/dp/]

Why Not Prospect Utilitarianism Instead of Sufficientarianism?

A Formal Comparison in Terms of Continuity and Lifeboat Cases

Susumu Cato[1] Hun Chung[2]

[Abstract]

In his 2017 paper, "Prospect Utilitarianism: A Better Alternative to Sufficientarianism," Hun Chung proposed a theory of distributive justice called 'Prospect Utilitarianism (PU).' According to Chung, PU retains all the major attractions of sufficientarianism, while avoiding two major problems. The two problems are: (a) sufficientarianism fails to prescribe the right distribution under conditions of scarcity (i.e., 'lifeboat' situations), and (b) sufficientarianism fails to provide continuous ethical evaluations. Recently, Ben Davies (2022) and Lasse Nielsen (2019; 2023) have provided a defense of sufficientarianism from these two charges. This paper aims to provide a comprehensive analysis of the shortcomings in Davies's and to a lesser degree Nielsen's defenses of sufficientarianism. Our paper will highlight that both Davies's and Nielsen's defenses of sufficientarianism stem from fundamental misunderstandings related to the concepts of continuity, welfarism, value satiability, and their interconnectedness with sufficientarianism. In the end, we will argue that Prospect Utilitarianism (PU) is a superior alternative to sufficientarianism.

[Key Words]: Prospect Utilitarianism, Sufficientarianism, Welfarism, Continuity, Lifeboat Cases

[1] Associate Professor, Institute of Social Science, University of Tokyo

[2] Associate Professor, Faculty of Political Science & Economics, Waseda University

1. Introduction: What is Prospect Utilitarianism?

In his 2017 paper, "Prospect Utilitarianism: A Better Alternative to Sufficientarianism," Hun Chung proposed a theory of distributive justice called 'Prospect Utilitarianism (PU).'[3] PU combines the utilitarian social welfare function[4] with individual utility functions characterized in accordance with Kahneman and Tversky (1979)'s prospect theory.[5] The main selling point of PU is that "it has elements that can appeal to utilitarians, sufficientarians, prioritarians, and egalitarians at the same time"; it is, according to Chung, a "total package." (Chung 2017: 1932) To summarize:

(1) PU "recommends a distribution that always maximizes the instances of sufficiency" (Chung 2017: 1932), and, thereby, accommodates sufficientarianism's *positive thesis.*[6]

(2) As a result, PU gives right answers to 'lifeboat' scenarios.[7] (Chung 2017: 1923–1925)

---

[3] More recently, Chung (2022) has shown how PU can be defended from the original position.

[4] Let $N = \{1, \dots, n\}$ be the set of individuals. Let X be the set of social alternatives. For each $i \in N$, let $u_i: X \to R$ be individual i's utility function. Then, for any social alternative $x \in X$, the utilitarian social welfare function $U: X \to R$ is defined as: $U(x) = u_1(x) + \cdots + u_n(x) = \sum_{i=1}^{n} u_i(x)$. Hence, according to utilitarian social welfare function, for any social alternatives $x, y \in X$, x is socially preferred to y iff $U(x) > U(y)$ iff $\sum_{i=1}^{n} u_i(x) > \sum_{i=1}^{n} u_i(y)$. That is, social alternative x is socially preferred to social alternative y iff the total sum of individual utility generated by x is greater than that generated by y.

[5] Kahneman and Tversky propose that each individual's utility is: "(i) defined on deviations from the reference point; (ii) generally concave for gains and commonly convex for losses; (iii) steeper for losses than for gains." (Kahneman and Tversky 1979: 279) From this, Chung defines a reference utility function that meets these three characteristics and defines each individual's utility function as a horizontal translation of the reference utility function so that it generates a utility of 0 at their respective critical sufficiency threshold of resource levels. (Chung 2017: 1922–1923) The reference utility function is designed to measure the extent to which providing a specific benefit to an individual situated at a certain distance from their reference point contributes to the overall value of the entire distribution.

[6] Sufficientarianism's 'positive thesis' claims that it is morally important for people to have enough material resources.

[7] "A 'lifeboat situation' is characterized by the fact that some people will necessarily fall below a critical sufficiency threshold no matter how we distribute the remaining resources." (Chung 2017: 1913)

(3) PU "explains why we are not usually morally disturbed by the relative inequality between the rich and super-rich" (ibid.) and thereby accommodates (partially) sufficientarianism's *negative thesis*[8].

(4) PU always generates *continuous* ethical judgments[9]. (Chung 2017: 1925–27)

(5) Despite being a version of utilitarianism, PU is "immune to a standard objection to utilitarianism – namely, that it may justify vastly unequal distributions." (Chung 2017: 1932)

(6) This is so because "[a]fter maximizing sufficiency, prospect utilitarianism prioritizes the individual who is worse-off in terms of his/her welfare level." (Chung 2017: 1927, Proposition 4) In other words, although PU does not 'directly' aim to equalize welfare, because of the specific way that individual utility functions are characterized in PU, PU has a built-in bias towards equality and is (like prioritarianism) "derivatively egalitarian." (Benbaji 2005: 312; Chung 2017: 1929)

As a version of utilitarianism, PU appeals to utilitarians by default. In addition, (1) and (3) make PU appealing to sufficientarians; (2) and (4) are independent moral desiderata that any principle of distributive ethics should ideally satisfy; and (5) and (6) make PU appealing to both egalitarians and prioritarians.

[8] Sufficientarianism's 'negative thesis' claims that once everybody has enough material resources, whether somebody has more or less material resources than others has no moral significance. According to Chung, PU implies a slightly weaker version of the negative thesis called "the weak negative thesis" according to which "As two people have more and more material resources, the ethical significance of their relative inequality matters less and less." (Chung 2017: 1930)

[9] Intuitively, ethical judgments are continuous if "there are no 'jumps' in the ethical preference order. … it says that two social states that are almost the same, in terms of the welfare levels of society's members must be viewed as almost ethically indifferent." (Roemer 2004: 272) We may think of continuity as formally expressing Aristotle's principle to treat 'like cases alike and different cases differently.'

According to Chung, the two central problems of sufficientarianism including many of its recent (non-headcount) variants[10] that make PU a better alternative is that it fails to meet (2) and (4): that is, "[sufficientarianism] cannot provide right answers to lifeboat situations … [failing (2)] and … it fails to provide continuous ethical evaluations… [failing (4)]." (Chung 2017: 1916) In his 2017 paper, Chung uses these two failures of sufficientarianism as a foil to motivate PU as distributive ethical principle that retains all major attractions of sufficientarianism while avoiding its drawbacks.

In a recent paper entitled, "The Prospects for 'Prospect Utilitarianism'," Ben Davies (2022) attempts to defend sufficientarianism from these two charges. With respect to desideratum (2), Davies (2022: 336–38) argues that recent non-headcount versions of sufficientarianism can indeed give right answers to lifeboat cases by considering "benefit size." With respect to desideratum (4), Davies (2022: 339–41) argues that sufficientarianism can satisfy a more general concept of continuity even if it fails to satisfy the narrower (and, hence, more questionable) concept of continuity (which Davies calls "welfarist continuity") on which Chung's critique of sufficientarianism purported relies. In a related vein, Lasse Nielsen has recently argued that critiquing sufficientarianism on grounds of discontinuity is based on what he calls "the numbers fallacy." According to Nielsen, this fallacy arises from the utilization of numerical examples featuring "empty numbers" that fail to accurately capture the underlying value framework that may possess satiability and/or range properties. (Nielsen 2019: 802–809; Nielsen 2023: 8)

This paper aims to provide a formal analysis of the shortcomings in Davies's and to a lesser degree Nielsen's defenses of sufficientarianism. Our paper will highlight that both

[10] E.g., Crisp (2003), Brown (2005), Casal (2007), Huseby (2010), Hirose (2016), Bossert, Cato, and Kamaga (2022, 2023).

Davies's and Nielsen's defenses of sufficientarianism stem from fundamental misunderstands regarding Chung's Prospect Utilitarianism (PU) and its associated concepts, including continuity, welfarism, value satiability, and their interconnectedness with sufficientarianism. We believe that it is very important to correct these misunderstandings and achieve a clear and precise understanding of these important technical concepts as these concepts are quite integral to any principle of distributive ethics and such endeavors may help protect other axiologists and normative ethical theorists from committing similar errors when producing scholarly work in the ethics of distribution.

In the end, this paper highlights the importance of providing precise "formal" definitions and rigorous treatments for various normative concepts. The significance of this practice is underscored by the fact that certain fundamental mathematical concepts, such as continuity, are often employed in philosophical literature without a precise mathematical definition. Such imprecise usage gives rise to multiple formulations of continuity based on the specific context or ethical framework employed, leading to potential discrepancies in the arguments and conclusions drawn from them. Hence, a precise formal approach to handling various normative concepts is imperative for ensuring clarity, consistency, and robustness in theoretical discussions in distributive ethics.

2. What is Continuity? Continuity *per se* vs. Welfare Continuity

Let us first probe the issue of continuity. Intuitively, we say that ethical judgments are continuous if "there are no 'jumps' in the ethical preference order. … two social states that are almost the same … must be viewed as almost ethically indifferent." (Roemer 2004: 272) We may think of continuity as a formalization of Aristotle's principle that requires to 'treat

like cases alike and different cases differently.' To understand continuity, one must first ask: two social states that are almost the same *with respect to what*? The definition of continuity presumes a space (a domain) over which moral evaluations of ethical preferability are made. If the domain under consideration is a space of resources, then continuity claims that there is a continuous relationship between the distribution of resources and ethical preferability. If the domain under consideration is a space of welfare, then continuity claims that there is a continuous relationship between distribution of welfare and ethical preferability.

Many theories of distributive ethics that work under a welfarist framework are either explicitly or implicitly committed to the view that there is a continuous relationship between the distribution of welfare and ethical preferability. (See Kaplow and Shavell (2001), Roemer (2004), Fleurbaey (2015: 207), Chung (2017), etc.[11]) Davies (2022) calls this "welfarist continuity" and distinguishes it from "continuity *per se*." (Davies 2022: 339)

> Continuity *per se* is an attractive feature of an ethical view if understood as saying that, where is only a slight difference *across all ethically relevant features* between two outcomes, this should not make a big difference to our ethical preferences. But

[11] Among ethical theorists and axiologists, continuity has usually been discussed in the context of determining the value ordering of risky lotteries in accordance with 'expected utility theory.' Intuitively, the continuity axiom of expected utility theory states that for any triple of outcomes x, y, and z, such that x is strictly preferred to y, which, in turn, is strictly preferred to z, there exists some probability $p \in (0, 1)$ that makes the decision-maker indifferent between getting y for sure on the one hand, and playing a gamble (or lottery) that gives x with probability p and z with probability $1 - p$ on the other hand. Temkin (2001) has argued that such a continuity axiom is implausible when x is very good while z is extremely bad ['continuity of extreme cases'], but this is implied by the more plausible continuity axiom that is restricted to adjacent cases, where the differences in the desirability of three outcomes x, y, and z are small, together with 'substitutivity of equivalents' and 'transitivity.' Temkin concludes that, therefore, there must be something wrong with 'transitivity.' [See also Temkin (1996)] Binmore and Voorhoeve (2003), Arrhenius and Rabinowicz (2005), and Stefansson (2022) point out that Temkin (1996; 2001)'s arguments are technically flawed.

> we can clearly embrace a principle of continuity understood in this way, while rejecting *welfarist* continuity, since two outcomes may differ only slightly in welfare, but differ more significantly on some other ethically relevant factor. For instance, egalitarians will strongly prefer a distribution A, which has identical population welfare to distribution B, but where the welfare is much more equally distributed. … However, sufficientarians can plausibly reject welfarist continuity while embracing the less specific principle of continuity I have suggested. (Davies 2022: 340)

In sum, Davies is asserting two claims:

(I) Although continuity in general (i.e., continuity *per se*) is a plausible principle, *welfarist* continuity (i.e., continuity restricted to welfare space) is not. (Davies 2022: 341)

(II) Although sufficiantarianism violates *welfarist* continuity, it can still satisfy a more *general notion of continuity* (i.e., continuity *per se*) defined over *all ethically relevant features*. (Davies 2022: 340).

Both claims are *false*. Let us first examine claim (I) in more detail. According to Davies, one can clearly embrace a principle of continuity defined across all ethically relevant features, while rejecting a principle of continuity restricted to distributions of welfare. This cannot be true because no ethical view can satisfy a general notion of continuity (i.e., continuity *per se*) defined over all ethically relevant features without first satisfying continuity in the space of welfare distributions. The logical mistake that Davies is making here is analogous to accepting a more general statement while rejecting a particular case of that statement: e.g.,

accepting "all *F*s have property *P*" and at the same time rejecting "all *F*s that are also *G*s have property *P*," which is a logical contradiction.

To understand this precisely, we need to first clarify a couple of things. When arguing against welfarist continuity, Davies attacks a view that is not only importantly different from continuity over welfare space, but which is also much stronger (and hence more implausible) than the claim that ethical preferences should be continuous over distributions of welfare. We might call the view Davies mistakenly labels as "welfarist continuity" *Total Population Welfare Continuity* (TPWC).

To understand this formally, let $N = \{1, \dots, n\}$ be a set of n individuals, let $W = \mathbb{R}^n$ be a space of welfare distributions, and let $u, v \in W$ be two distributions of welfare across the n individuals. Let $\succsim$ denote the ethical/moral preference relation such that $u \succsim v$ if and only if the welfare distribution in u is weakly ethically preferred to the welfare distribution in v. Let $\succ$ and $\sim$ be the asymmetric and symmetric parts of $\succsim$. Then, we can define total population welfare continuity as follows:

Total Population Welfare Continuity (TPWC, Davies's 'Welfarist Continuity'): *For any* $u, v \in W$ *and any sequence* $\{v^k\}$ *such that* $\sum_{i \in N} v_i^k \longrightarrow \sum_{i \in N} v_i$*, if* $u \succ v$ *(respectively,* $v \succ u$*), then there exists a* $k^* \in \mathbb{N}$ *such that* $u \succ v^k$ *(respectively,* $v^k \succ u$*) for all* $k > k^*$.

In words, TPWC claims that if two welfare distributions are almost the same in their *total sum of population welfare*, then the two welfare distributions should be almost ethically indifferent in the sense that if one welfare distribution is ethically preferred to another welfare distribution, then the former distribution should remain ethically preferred to any welfare distribution whose total sum of population welfare is sufficiently close to that of the latter

distribution. Contrast this with welfare continuity to which most welfarist theories of distributive justice are committed:

Welfare Continuity: *For any $u, v \in W$ and any sequence $\{v^k\}$ such that $v^k \longrightarrow v$, if $u \succ v$ (respectively, $v \succ u$), then there exists a $k^* \in \mathbb{N}$ such that $u \succ v^k$ (respectively, $v^k \succ u$) for all $k > k^*$.*

In words, welfare continuity claims that if two welfare distributions are almost the same, then the two welfare distributions should be almost ethically indifferent in the sense that if one welfare distribution is ethically preferred to another welfare distribution, then the former distribution should remain ethically preferred to any welfare distribution that is sufficiently close to the latter distribution. Notably, welfare continuity does not say that welfare is our only ethically relevant consideration. (Such a view is called 'welfarism' which we will later define precisely and analyze.) Rather, welfare continuity says that our ethical preference relation is *continuous* in the space of welfare in the sense that if the welfare distribution u is strictly ethically preferred to another welfare distribution v for whatever (welfare or non-welfare) reasons, then any welfare distribution u^k that is *in the neighborhood* of u should also be strictly ethically preferred to v. The point is that, regardless of which features of a distribution we deem ethically relevant, TPWC is a much stronger condition than welfare continuity; in particular, TPWC logically implies welfare continuity, but welfare continuity does not logically imply TPWC.

Proposition 1. *TPWC implies Welfare Continuity, but not vice versa.*

Proof. Suppose the ethical preference relation $\succsim$ satisfies total population welfare continuity. Pick any $u, v \in W$, any sequence $\{v^k\}$ such that $v^k \longrightarrow v$, and suppose $u \succ v$. Since $v^k \longrightarrow v$, we have $\sum_{i \in N} v_i^k \longrightarrow \sum_{i \in N} v_i$. By total population welfare continuity, there exists a $k^* \in \mathbb{N}$ such that $u \succ v^k$ for all $k > k^*$ as desired. (The case where $v \succ u$ can be proved analogously.) Now, consider the maximin criterion according to which $u \succsim v$ if and only if $\min_{i \in N} u \geq \min_{i \in N} v$. We claim that the maximin criterion satisfies welfare continuity but violates total population welfare continuity. To show that the maximin criterion satisfies welfare continuity, pick any $u, v \in W$, any sequence $\{v^k\}$ such that $v^k \longrightarrow v$, and suppose $u \succ v$ by the maximin criterion. By definition, we have $\min_{i \in N} u > \min_{i \in N} v$. Define $\delta = \min_{i \in N} u - \min_{i \in N} v$. Since $v^k \longrightarrow v$, there exists a $k^* \in \mathbb{N}$ such that $\left|\min_{i \in N} v - \min_{i \in N} v^k\right| < \delta$ for all $k > k^*$. Hence, for all $k > k^*$, we have $\min_{i \in N} u > \min_{i \in N} v^k$, and, hence, $u \succ v^k$ for all $k > k^*$ as desired. Now, to show that the maximin criterion does not satisfy total population welfare continuity, pick any $u, v \in W$ such that $\overline{v} > \min_{i \in N} u > \min_{i \in N} v$, where $\overline{v} = \frac{\sum_{i \in N} v_i}{n}$. Define a sequence $\{v^k\}$ such that $v^k = (\overline{v}, \ldots, \overline{v})$ for all $k \in \mathbb{N}$. Then, we have $\sum_{i \in N} v_i^k = n\overline{v} = \sum_{i \in N} v_i$ for all $k \in \mathbb{N}$, and, hence, $\sum_{i \in N} v_i^k \longrightarrow \sum_{i \in N} v_i$. However, there exists no $k^* \in \mathbb{N}$ such that $u \succ v^k$ for all $k > k^*$ because $\min_{i \in N} v^k = \overline{v} > \min_{i \in N} u$ for all $k \in \mathbb{N}$, and, hence, $v^k \succ u$ all $k \in \mathbb{N}$. ■

The example that "egalitarians will strongly prefer a distribution A, which has identical population welfare to distribution B, but where the welfare is much more equally distributed" (Davies 2022: 340) is an example that shows that certain ethical views may plausibly violate TPWC; but it is *not* a counter example to welfare continuity. For instance, suppose there are two individuals and let (w_1, w_2) denote a distribution of welfare to these two individuals. Although welfarist continuity requires a welfare distribution of $(5, 5)$ to be

almost ethically indifferent to a welfare distribution of $(5+\varepsilon, 5)$ for a sufficiently small $\varepsilon > 0$, (unlike TPWC) it does not require the welfare distribution of $(5, 5)$ to be ethically indifferent to a welfare distribution of $(10, 0)$ even though the total population welfare of $(5, 5)$ and $(10, 0)$ are exactly the same. Instead, welfare continuity says that if one believes that $(5, 5)$ is ethically preferred to $(10, 0)$ (for egalitarian or other ethically relevant reasons), then one should also believe that $(5+\varepsilon, 5)$ is ethically preferred to $(10, 0)$ for a sufficiently small $\varepsilon > 0$.

Hence, unlike what Davies imprecisely thinks, egalitarianism does satisfy welfare continuity, even though it violates TPWC. The same holds true for other welfarist ethical theories such as prioritarianism, the maximin criterion (i.e., Rawls's difference principle applied to people's welfare levels), etc. According to Fleurbaey (2015), "[t]he only sensible ranking that this mild condition [viz., welfare continuity] rules out is the leximin, but [even] the leximin can be treated as the limit of a family of continuous social rankings." (Fleurbaey 2015: 207 footnote 4) Hence, it is a mistake to call TPWC 'welfarist' because very few welfarist ethical theories satisfy it. Indeed, as we will now show, the only welfarist ethical theory that satisfies TPWC is utilitarianism in the presence of weak Pareto.

Weak Pareto: *For any $u, v \in W$, if $u_i > v_i$ for all $i \in N$, then $u \succ v$.*

Proposition 2. *An ethical preference relation $\succsim$ is utilitarian if and only if it satisfies weak Pareto and TPWC.*

Proof. We first show that the utilitarian ethical preference relation $\succsim$ implies weak Pareto and total population welfare continuity. (Weak Pareto): Let $\succsim$ be utilitarian. Pick any $u, v \in W$

and suppose $u_i > v_i$ for all $i \in N$. Then, we have $\sum_{i\in N} u_i > \sum_{i\in N} v_i$, which implies that $u \succ v$ as desired. (Total Population Welfare Continuity): Let $\succsim$ be utilitarian. Pick any $u, v \in \mathbb{R}^n$, any sequence $\{v^k\}$ such that $\sum_{i\in N} v_i^k \longrightarrow \sum_{i\in N} v_i$ and suppose $u \succ v$. Since $u \succ v$, we have $\sum_{i\in N} u_i > \sum_{i\in N} v_i$. Define $\delta = [\sum_{i\in N} u_i - \sum_{i\in N} v_i]$. Since $\sum_{i\in N} v_i^k \longrightarrow \sum_{i\in N} v_i$, there exists a $k^* \in \mathbb{N}$ such that $\left|\sum_{i\in N} v_i^k - \sum_{i\in N} v_i\right| < \delta$ for all $k > k^*$. This implies that $\sum_{i\in N} u_i - \sum_{i\in N} {v_i}^k > 0$, which implies $\sum_{i\in N} u_i > \sum_{i\in N} {v_i}^k$, which implies $u \succ v^k$ for all $k > k^*$ as desired.

We now show that any ethical preference relation $\succsim$ satisfying weak Pareto and total population welfare continuity is utilitarian. Pick any $u, v \in W$ such that $\sum_{i\in N} u_i = \sum_{i\in N} v_i$. We need to show that $u \sim v$. For a proof by contradiction, suppose not. Without loss of generality suppose $u \succ v$. Define a sequence $\{v^k\}$ where $v_i^k = u_i + \frac{1}{k}$ for each $k \in \mathbb{N}$. Note that $\sum_{i\in N} v_i^k \longrightarrow \sum_{i\in N} v_i$. By total population welfare continuity, there exists a $k^* \in \mathbb{N}$ such that $u \succ v^k$ for all $k > k^*$. However, by weak Pareto, we have $v^k \succ u$ for all $k \in \mathbb{N}$, a contradiction. Therefore, we must have $u \sim v$. Now, suppose $\sum_{i\in N} u_i > \sum_{i\in N} v_i$. We need to show that $u \succ v$. Consider $u' = (\overline{u}, \dots, \overline{u})$ and $v' = (\overline{v}, \dots, \overline{v})$, where $\overline{u} = \frac{\sum_{i\in N} u_i}{n}$ and $\overline{v} = \frac{\sum_{i\in N} v_i}{n}$. (Note that $\overline{u} > \overline{v}$.) Since $\sum_{i\in N} u_i = \sum_{i\in N} \overline{u}_i$ and $\sum_{i\in N} \overline{v}_i = \sum_{i\in N} v_i$, by our previous result, we must have $u \sim u'$ and $v' \sim v$. Since $\overline{u} > \overline{v}$, by weak Pareto, we must have $u' \succ v'$. Hence, we have $u \sim u' \succ v' \sim v$. By transitivity, we have $u \succ v$ as desired. ■

Now that we have clarified and distinguished welfare continuity from TPWC, let us now examine the argument Davies uses to defend claim (II). According to Davies, *autonomy* is an example that shows how sufficientarians can reject welfarist continuity while accepting a more general principle of continuity:

> However, sufficientarians can plausibly reject welfarist continuity while embracing the less specific principle of continuity I have suggested. For instance, Shields (2018: 44–81) suggests that one candidate for a sufficientarian threshold is autonomy, and that we have especially weighty reasons to secure for people a level of autonomy sufficient to develop and pursue their own view of the good. … If welfarist continuity is an axiom, this seems to rule out by fiat any view that is interested in autonomy except insofar as it impacts welfare. Take, for example, *On Liberty* (Mill 1859/2003), where Mill says we should prefer that people have significant autonomy over their own lives, because a life you choose for yourself is better than one chosen for you. Mill's view is welfarist: the betterness of the life is understood in terms of its welfare or prudential value. But we might instead think that self-chosen lives are *non-prudentially* better even if they contain less welfare, simply because they are lives which individuals choose for themselves. On this view, it is possible to have two identical distributions of welfare, one paternalistically imposed and the other autonomously chosen, and prefer the latter for non-welfarist reasons. (Davies 2022: 340–41)

To appropriately respond to Davies, we would need to define the concept of *general continuity* defined over *all ethically relevant features*. Let $W = \mathbb{R}^n$ denote a space of welfare distributions. Let $X = X_1 \times \cdots \times X_m = \mathbb{R}^{nm}$ denote a space of distributions of *all ethically relevant non-welfare features*. For instance, X_1 might be the space of resource distributions, X_2 might be the set of distributions of autonomy, X_3 might be the set of distributions of desert, and so on. Let our ethical preference relation $\succsim$ be defined over $W \times X$, that is, it is

defined over all ethically relevant features (including both welfare and non-welfare features.) From this, we may define general continuity (or what Davies calls "continuity *per se*") over all ethically relevant features as follows:

General Continuity (over *all ethically relevant features*): *For any* $(u, x), (v, y) \in W \times X = \mathbb{R}^n \times \mathbb{R}^{nm}$ *and any sequence* $\{(v^k, y^k)\}$ *such that* $(v^k, y^k) \longrightarrow (v, y)$, *if* $(u, x) \succ (v, y)$ *(respectively,* $(v, y) \succ (u, x)$*), then there exists a* $k^* \in \mathbb{N}$ *such that* $(u, x) \succ (v^k, y^k)$ *(respectively,* $(v^k, y^k) \succ (u, x)$*) for all* $k > k^*$.

General continuity formally expresses Davies's intuitive thought that "[if there] is only a slight difference *across all ethically relevant features* between two outcomes, this should not make a big difference to our ethical preferences." (Davies 2022: 340) From this, we would need to extend our definition of welfare continuity to apply to a space of all ethically relevant features (i.e., $W \times X$):

Welfare Continuity* (extended to *all ethically relevant features*): *For any* $(u, x), (v, y) \in W \times X$ *and any sequence* $\{(v^k, y)\}$ *such that* $(v^k, y) \longrightarrow (v, y)$, *if* $(u, x) \succ (v, y)$ *(respectively,* $(v, y) \succ (u, x)$*), then there exists a* $k^* \in \mathbb{N}$ *such that* $(u, x) \succ (v^k, y)$ *(respectively,* $(v^k, y) \succ (u, x)$*) for all* $k > k^*$.

Welfare continuity* states that (given that welfare is one among many ethically relevant considerations) if one distribution of all ethically relevant features is ethically preferred to another distribution of all ethically relevant features, then the former distribution should remain ethically preferred to any other distribution that closely approximates the latter in

terms of welfare distribution while maintaining identical non-welfare features. Defined in this way, general continuity defined over all ethically relevant features necessarily implies welfare continuity.

Proposition 3. *General Continuity defined over all ethically relevant features implies Welfare Continuity*.*

Proof. Suppose that an ethical preference relation $\succsim$ on $W \times X$ satisfies general continuity. Pick any $(u, x), (v, y) \in W \times X$, any sequence $\{(v^k, y)\}$ such that $(v^k, y) \longrightarrow (v, y)$, and suppose $(u, x) \succ (v, y)$. Define a new sequence $\{(\hat{v}^k, \hat{y}^k)\}$ such that $\hat{v}^k = v^k$ and $\hat{y}^k = y$ for all $k \in \mathbb{N}$. Then, we have $(\hat{v}^k, \hat{y}^k) \longrightarrow (v, y)$. By general continuity, there exists a $k^* \in \mathbb{N}$ such that $(u, x) \succ (\hat{v}^k, \hat{y}^k) = (v^k, y)$ for all $k > k^*$. (The proof for the case where $(v, y) \succ (u, x)$ is analogous.) Hence, our ethical preference relation $\succsim$ on $W \times X$ satisfies welfare continuity*. ■

In other words, if an ethical preference relation is continuous over all ethically relevant features (where welfare is one such ethically relevant feature), such an ethical preference relation restricted to welfare space must also be continuous over that space. Otherwise, if there is discontinuity at some point in welfare space, then there will be a discontinuity in the space of all ethically relevant features at that point where welfare discontinuity occurs. This is the fundamental point that Davies misunderstands.

3. Welfarism and Welfare Continuity

Recall that Davies's argument was that it is possible to have two identical distributions of welfare, but ethically prefer one distribution over the other on grounds that one distribution has been autonomously chosen while the other has been paternalistically imposed. This is a valid point. But it is a mistake to think that this violates welfare continuity. Here, Davies seems to mistakenly assume that if one is committed to welfare continuity, then this forces one to regard any two situations that contain the same distributions of welfare to be ethically indifferent (regardless of how the two situations differ in any other ethically relevant features.) Such a view is *welfarism* (*not welfare continuity*), which may be formally defined as follows:[12]

Welfarism: *There exists a relation* $\succsim^*$ *on* W *such that for any* $(u,x),(v,y) \in W \times X$, $u \succsim^* v \iff (u,x) \succsim (v,y)$.

According to welfarism, there is an ethical preference relation on the space of welfare distributions such that one welfare distribution u is ethically preferred to another welfare distribution v if and only if (u,x) is ethically preferred to (v,y) for all non-welfare features x and y; essentially, it says that all non-welfare information is irrelevant, and the only ethically relevant consideration is welfare. By contrast, welfare continuity says that if some situation is all things considered ethically preferred to another situation, then the former situation should remain ethically preferred to the latter situation even if we ever so slightly

[12] Note that our definition is compatible with the definition of welfarism, which has been used by many welfare economists since the work of Sen (1979). A concise argument on this topic can be found in Blackorby, Bossert, and Donaldson (2005). Our definition is particularly similar to that of Kaplow and Shavell (2001).

change the distribution of welfare in either the former or the latter situation. Welfarism is not equivalent to welfare continuity; in fact, the two concepts are logically independent.

Proposition 4. *Welfarism and Welfare Continuity* are logically independent on* $W \times X$.

Proof. First, we show that Welfarism does not imply Welfare Continuity*. Let $\underline{u} > 0$ and for any $(u, x) \in W \times X$, let $f(u, x) = \#\{i \in N \mid u_i \geq \underline{u}\}$. Define an ethical preference relation $\succsim$ on $W \times X$ as follows: for all $(u, x), (v, y) \in W \times X$, $(u, x) \succsim (v, y)$ if and only if $f(u, x) \geq f(v, y)$. We can confirm that $\succsim$ satisfies Welfarism. To see this, we can define a relation $\succsim^*$ on W by letting: $u \succsim^* v$ if and only if $\#\{i \in N \mid u_i \geq \underline{u}\}. \geq \#\{i \in N \mid v_i \geq \underline{u}\}$. Note that for all $(u, x), (v, y) \in W \times X$, $u \succsim^* v$ if and only if $(u, x) \succsim (v, y)$. Thus, Welfarism holds. Now, let $u = (\underline{u}, 0, \dots, 0)$ and $v = (\underline{u}, \dots, \underline{u})$. Then, we have for all $x, y \in X$, $f(u, x) = 1$ and $f(v, y) = n$. Hence, we have $(v, y) \succ (u, x)$. Consider the sequence $\{(v^k, y)\}$, where $v^k = (\underline{u} - \frac{1}{k}, \dots, \underline{u} - \frac{1}{k})$ for all $k \in \mathbb{N}$. Then, we have $(v^k, y) \longrightarrow (v, y)$. However, we have $(u, x) \succ (v^k, y)$ for all $k \in \mathbb{N}$. Therefore, the ethical preference relation $\succsim$ violates Welfare Continuity*. Second, we show that Welfare Continuity* does not imply Welfarism. Now, define another ethical preference relation $\succsim'$ such that for all $(u, x), (v, y) \in W \times X$, $(u, x) \succsim' (v, y)$ if and only if $\sum_{i \in N} \sum_{\ell=1}^{m} x_{\ell i} \geq \sum_{i \in N} \sum_{\ell=1}^{m} y_{\ell i}$. Suppose $(u, x) \succ' (v, y)$. Pick any sequence $\{(v^k, y)\}$ such that $(v^k, y) \longrightarrow (v, y)$. Then, we have $(u, x) \succ' (v^k, y)$ for all $k \in \mathbb{N}$. Hence, the ethical preference relation $\succsim'$ satisfies Welfare Continuity*. However, $\succsim'$ violates Welfarism because we have $(w, x) \succ' (w, y)$ for all $w \in W$. ■

Again, Davies claims: "If welfarist continuity is an axiom, this seems to rule out by fiat any view that is interested in autonomy except insofar as it impacts welfare." (Davies

2022: 341) This is incorrect: "ruling out by fiat any view that is interested in autonomy except insofar as it impacts welfare" is a logical implication of *welfarism* (which states that welfare is the only ethically relevant consideration), but it is *not* a logical implication of *welfarist continuity* (which states that our ethical judgments must be continuous over the space of welfare.) Even if one acknowledges the existence of other ethically relevant non-welfare considerations, it remains possible for ethical judgments to demonstrate continuity within the realm of welfare.

Going back to Davies's example of autonomy, consider two identical distributions of welfare, where one is generated autonomously, while the other is imposed paternalistically. Since the two distributions are identical in terms of welfare, welfarism implies that the two distributions should be ethically indifferent. By contrast, welfare continuity says that if the autonomously generated distribution is ethically preferred to the paternalistically imposed distribution for non-welfare reasons (despite the fact that the two distributions contain identical distributions of welfare), then the autonomously generated distribution *should remain ethically preferred* to the paternalistically imposed distribution even if the autonomously generated distribution contained *slightly less welfare* or the paternalistically imposed distribution contained *slightly more welfar*e. The two concepts are clearly different, but it is not easy to recognize the theoretical difference between the two concepts if one conflates two different questions that should be kept separate. The two questions are:

Q1: Should welfare be *the only* relevant consideration in our ethical evaluations?

Q2: Should our ethical evaluations be *continuous*?

Welfarism says ‘yes’ to the first question. Utilitarians say ‘yes’ to both questions. However, the answers to the two questions, as we have seen, are independent; so, a ‘yes’ or ‘no’ answer to the first question does not imply either a ‘yes’ or ‘no’ answer to the second question, and vice versa. Davies wants to say ‘yes’ to the second question and ‘no’ to the first question (i.e., there are morally relevant considerations other than welfare.) If so, then Davies’s criticism is a criticism against welfarism, *not* continuity. As a matter of fact, we can now see (after defining welfarism and welfare continuity precisely) that Davies’s purported counter examples and criticisms are all targeting welfarism, not continuity.

We believe that this is the root of the Davies’s apparent contradiction. That is, what really bothers Davies is ‘welfarism’ to which all versions of utilitarianism are committed. At the same time, Davies wants to defend sufficientarianism. However, Chung (2017) does not criticize sufficientarianism on grounds of welfarism; he criticizes sufficientarianism on grounds of generating discontinuity. Hence, to defend sufficientarianism from Chung’s criticism, Davies had to find a way to defend sufficientarianism from the charge of discontinuity, which is essentially an impossible task. Hence, he artificially splits the notion of continuity and argues that he is rejecting not continuity *per se*, but rather a much narrower view of continuity, which he calls ‘welfarist’ continuity, on which Chung purportedly bases his criticisms of sufficientarianism. Davies then claims (without argument) that although sufficientarianism fails to satisfy this narrow version of (welfarist) continuity (which is okay since welfarist continuity is implausible), it can still satisfy a more general principle of continuity defined over all ethically relevant features. We will now explain why such a strategy cannot succeed.

4. Sufficientarianism and Continuity

As we have seen, welfare continuity, when precisely defined over all ethically relevant features, is different from welfarism, and is, therefore, not implausible for reasons that somebody might find welfarism implausible. But even if it were implausible for other valid reasons, Davies's argumentative strategy would still not work because as long as sufficientarianism presumes the existence of *some* critical sufficiency threshold of *some* ethically relevant feature (be it autonomy, welfare, reasonable contentment, or whatever) over which the ethical value of the distribution suddenly 'jumps,' it will necessarily be discontinuous in whatever domain that our working notion of continuity (be it welfare continuity, general continuity, etc.) is defined over.

To see this, assume that there are two individuals (i.e., $N = \{1, 2\}$) and that welfare and autonomy are our two main ethically relevant considerations. Let $W = \mathbb{R}^2$ be a space of welfare distributions and let $A = \mathbb{R}^2$ be a space of distributions of autonomy. Suppose $\underline{a} > 0$ is the level of autonomy "sufficient to develop and pursue [one's] own view of the good" (Davies 2022: 340) and suppose our sufficientarian theory claims that we have "especially weighty reasons to secure" each individual an autonomy level of $\underline{a}$. Suppose our ethical preference relation ranks each distribution of autonomy and welfare *lexicographically*: specifically, it first looks at the levels of autonomy the two individuals enjoy and ranks a distribution with more individuals above the autonomy threshold $\underline{a}$ over any distribution with less individuals above the autonomy threshold $\underline{a}$; next, when the number of individuals above the critical autonomy threshold $\underline{a}$ are the same, then it applies utilitarianism and prefers the distribution that yields a larger total sum of individual well-being levels in their welfare distributions.

Suppose $u > v$. Then, we have $\big((u, \underline{a}), (u, \underline{a})\big) \succ \big((v, \underline{a}), (v, \underline{a})\big)$. Now, consider a sequence $\left\{\left((u^k, \underline{a}^k), (u^k, \underline{a}^k)\right)\right\}$ where $u^k \longrightarrow u$ and $\underline{a}^k = \underline{a} - \frac{1}{k}$. Then, we have $\left\{\left((u^k, \underline{a}^k), (u^k, \underline{a}^k)\right)\right\} \longrightarrow \big((u, \underline{a}), (u, \underline{a})\big)$. However, we have $\big((v, \underline{a}), (v, \underline{a})\big) \succ \left((u^k, \underline{a}^k), (u^k, \underline{a}^k)\right)$ for all $k \in \mathbb{N}$. Therefore, our ethical preference relation that incorporates sufficientarian considerations about individual autonomy is *discontinuous* from the perspective of (*not welfare* continuity, but) *general* continuity defined over *all ethically relevant considerations*, which, in our working example, are welfare and autonomy. (Note that such an ethical preference relation will satisfy welfare continuity.) This argument can be easily generalized to any number of ethically relevant considerations (besides welfare and autonomy.)

For this purpose, let us consider a generalized version of headcount sufficientarianism. Consider any continuous and monotonic function $h\colon \mathbb{R}^{1+m} \to \mathbb{R}$ and some critical level $\theta^* \in \mathbb{R}$ such that for each $i \in N$, $h(u_i^*, x_i^*) = \theta^*$ for some $(u_i^*, x_i^*) \in \mathbb{R}^{1+m}$. Then, according to generalized headcount sufficientarianism, for any $(u, x), (v, y) \in \mathbb{R}^{(1+m)n}$,

$$(u, x) \succsim (v, y) \Leftrightarrow \#\{i \in N \mid h(u_i, x_i) \geq \theta^*\} \geq \#\{i \in N \mid h(v_i, y_i) \geq \theta^*\},$$

where (u_i, x_i) is a vector that consists of individual i's welfare level and the distribution of all ethically relevant non-welfare features. By monotonicity, we mean that $h(u_i, x_i) > h(v_i, y_i)$ if every coordinate in (u_i, x_i) is larger than that in (v_i, y_i) (i.e., if $(u_i, x_i) \gg (v_i, y_i)$ holds). Here, $h(u_i, x_i)$ works as a general aggregation measure for sufficiency. That is, if its value for individual i is not lower than θ^*, then this individual is considered to have met their

critical sufficiency threshold and has enough. Our notion of generalized headcount sufficientarianism generalizes sufficientarianism by allowing us to regard sufficientarianism based on autonomy as a special case. From this, we are able to prove the following general result that essentially shows that expanding the number of ethically relevant considerations does not save sufficientarianism from the problem of discontinuity.

Proposition 5. *Generalized headcount sufficientarianism violates General Continuity defined over all ethically relevant considerations.*

Proof. For each $i \in N$, let (u_i^*, x_i^*) be such that $h(u_i^*, x_i^*) = \theta^*$. Let $\mathbf{1}_s$ be the s-dimentional vector composed of s ones. Let (u^*, x^*) be a profile such that all individuals $i \in N$ obtain (u_i^*, x_i^*). Also, let (v, y) denote a profile such that the first individual obtains (u_1^*, x_1^*) and all of other individuals $i \in N \setminus \{1\}$ obtains $(u_i^*, x_i^*) - \mathbf{1}_{(1+m)}$. Note that $n = \#\{i \in N \mid h(u_i^*, x_i^*) \geq \theta^*\} > \#\{i \in N \mid h(v_i, y_i) \geq \theta^*\} = 1$, and, hence, $(u^*, x^*) \succ (v, y)$. Now, take the following sequence $\{(u, x)^k\}$of profiles:$(u, x)^k = (u^*, x^*) - \frac{1}{k}\mathbf{1}_{n(1+n)}$. This sequence converges to (u^*, x^*). However, $(v, y) \succ (u, x)^k$ holds for all $k > 0$. Therefore, this violates General Continuity defined over all ethically relevant considerations. ■

In contrast, utilitarianism satisfies this stronger notion of general continuity defined over all ethically relevant considerations.

Proposition 6. *Utilitarianism satisfies General Continuity defined over all ethically relevant considerations.*

Proof. Let the ethical preference relation $\gtrsim$ defined on $W \times X$ be utilitarian: that is, for any $(u, x), (v, y) \in W \times X$, $(u, x) \gtrsim (v, y)$ if and only if $\sum_{i \in N} u_i \geq \sum_{i \in N} v_i$. Pick any $(u, x), (v, y) \in W \times X$ and suppose $(u, x) \succ (v, y)$. Let $\delta = (\sum_{i \in N} u_i - \sum_{i \in N} v_i)$ and pick any sequence $\{(v^k, y^k)\}$ such that $(v^k, y^k) \longrightarrow (v, y)$. Since $(v^k, y^k) \longrightarrow (v, y)$, there exists a $k^* \in \mathbb{N}$ such that for all $i \in N$, $\left|v_i^k - v_i\right| < \frac{\delta}{n}$ for all $k > k^*$. Then, by the triangle inequality, we have for all $k > k^*$:

$$\left|\sum_{i \in N} v_i^k - \sum_{i \in N} v_i\right| = \left|\left(v_1^k + \cdots + v_n^k\right) - (v_1 + \cdots + v_n)\right|$$

$$= \left|\left(v_1^k - v_1\right) + \cdots + (v_n^k - v_n)\right|$$

$$\leq \left|v_1^k - v_1\right| + \cdots + |v_n^k - v_n| < \frac{\delta}{n} + \cdots + \frac{\delta}{n} = \delta.$$

This implies that for all $k > k^*$, $\sum_{i \in N} u_i > \sum_{i \in N} v_i^k$, and, hence, $(u, x) \succ (v^k, y^k)$, as desired. ■

As Chung's PU is utilitarian, it follows from Proposition 6 that it satisfies general continuity defined over all ethically relevant considerations.

Corollary of Proposition 6. *Prospect Utilitarianism satisfies General Continuity defined over all ethically relevant considerations.*

Proposition 5 shows that Davies was simply wrong. Continuity is not a concept that can be arbitrarily stretched to be made consistent with sufficientarianism. If one is a sufficientarian, then one cannot avoid discontinuous ethical judgments defined over whatever domain of ethically relevant considerations; discontinuity is simply a cost that sufficientarians must stomach.

5. Satiability and Continuity

To this, some sufficientarians may acknowledge the discontinuity inherent in sufficientarianism but dismiss its severity. Recently, Lasse Nielsen (2023) has contended that the three standard objections raised against sufficientarianism, namely the problem of indifference[13], the problem of outweighing priority[14], and the problem of discontinuity[15], can all be viewed as specific manifestations of a more broader and general criticism termed "the threshold abruptness objection." (Nielsen, 2023: 3–6) According to Nielsen, all instances of the threshold abruptness objection (including the objection of discontinuity) are based on the construction of numerical counterexamples, whose intuitive force draws from "empty numbers" that lack any substantive moral content, and, therefore, falls prey to what he calls "the numbers fallacy." (Nielsen 2023: 12; Nielsen 2019: 812)

The reason why the numbers used in these numerical examples, according to Nielsen, are "empty" is that they "falsely … assume the Archimedean properties of real numbers in their underlying value-theory." (Nielsen 2019: 812) This assumption suggests that the underlying value being represented is *insatiable*, implying that regardless of the existing level of value, it is in principle always possible to introduce an additional unit of the underlying

[13] This is the objection that claims that sufficientarianism "implies indifference about inequality between the superrich and people who are barely above the threshold." (Nielsen 2023: 3)

[14] This is the objection that claims that sufficientarianism "implies that we should allow marginal benefits given to people below the threshold to outweigh very large benefits given to people above the threshold." (Nielsen 2023: 4)

[15] This is the objection that claims that sufficientarianism "fails to take into consideration the moral significance of 'being almost there'" such that "any state below the cut-off point will be significanly deficient, and thus there cannot be any (even almost) ethical indifference." (Nielsen 2023: 5)

value's currency to further augment its magnitude. However, Nielsen argues that various fundamental values such as autonomy (Raz, 1986), capability (Nussbaum, 2000), and reasonable contentment (Huseby, 2010), among others, exhibit the capacity for satiation or may exhibit range properties. In the latter case, these values are characterized by a binary nature wherein individuals either possess them or do not, and once obtained, further accumulation is impossible. (Nielsen, 2023: 8, 15) Consequently, if sufficientarians were to establish their framework of sufficientarianism by establishing the relevant thresholds based on such satiable foundational values, the apparent implausibility of the threshold abruptness objection could be readily explained away. Nielsen refers to this resultant form of sufficientarianism as "value-satiability sufficientarianism."

> Value-satiability sufficientarianism assumes that the relevant value is (or relevant values are) satiable and that distributive justice is fulfilled if and only if the distribution is such that the [sic] everyone is sated in regard to the relevant value(s). In other words, satiable-value sufficientarianism identifies the threshold as the point above which any person will not become better-off in terms of the relevant value by having more of whatever can be allocated to her. Here, the upper limit is founded on that it is impossible for any human person to be relevantly better-off than this in terms of justice-relevant values. (Nielsen 2019: 806–807)

The thought seems to be that if we accept that many important foundational normative values (such as autonomy, capability, freedom, and so on) can be sated, then the fact that a certain principle of distributive ethics, [such as sufficientarianism, whose critical sufficiency thresholds are defined on the basis of such satiable foundational values (Nielsen 2019,

2023)], displays discontinuity may not be that problematic. In other words, value satiability may *explain* why our ethical judgements can plausibly be discontinuous.

Davies essentially makes the same point when he distinguishes between 'ethical attractiveness' and the 'demands of justice,' and argues, following Nielsen, that "the values that are relevant to justice are *satiable*; once the relevant value has been sated, people can become better off, but not in a way that is relevant to justice." (Davies 2022: 341) From this, Davies argues that

> What justice demands – and sufficientarianism is a theory of *justice* – may be different from the 'moral value' of a total distribution. Thus, even if small differences in welfare necessarily implied only small differences in the ethical preferability of an outcome … such a relationship may not hold between welfare and the demands of justice. (Davies 2022: 341)

What Davies is saying is that if demands of justice is *satiable* (because they are based on satiable foundational values), then this implies that demands of justice display *discontinuity*. This is incorrect. Continuity says that small differences in our ethical data should not generate big differences in our ethical evaluations: however, continuity does *not* say that big differences in our ethical data should always generate big differences in our ethical evaluations (which is the inverse statement of continuity, and inverse statements are generally not logically equivalent to their original statements.)

For the sake of argument, let us accept Nielsen's point that many foundational values pertinent to justice, or relevant to defining sufficientarian thresholds, exhibit the property of satiation. Would this make our evaluations of justice *discontinuous*? Not necessarily. Rather,

value-foundational satiability merely implies that our evaluations of justice may not be *strictly monotonic* across the whole domain.

Let us illustrate this with a simple example. Suppose that a person's autonomy is satiable: that is, autonomy increases with a person's resource levels up to a point after which the person's autonomy is completely fulfilled and sated. Suppose that individual i's autonomy function α_i, which measures the degree of autonomy individual i enjoys, is defined as follows:

$$\alpha_i(x) = \begin{cases} x & \text{for } x < \underline{x} \\ \underline{x} & \text{for } x \geq \underline{x} \end{cases}$$

where $x \in \mathbb{R}$ is the amount of resources i has. So, individual i's autonomy increases linearly in the amount of resources up to $\underline{x}$ after which i's autonomy is fully sated and remains constant. Nevertheless, the autonomy function α_i that measures individual i's degree of autonomy *is continuous* in resources.[16] Contrast this with another autonomy function β_i:

$$\beta_i(x) = \begin{cases} x & \text{for } x < \underline{x} \\ \underline{x} + 100 & \text{for } x \geq \underline{x} \end{cases}$$

Here, individual i's autonomy increases linearly by the amount of resources that i receives up to *right before* $\underline{x}$; but *at* $\underline{x}$, individual i's autonomy *abruptly jumps* to $\underline{x} + 100$ and stays constant thereafter. Clearly, the autonomy function β_i is *discontinuous* at $\underline{x}$.

[16] **Proof.** Since both a linear function and a constant function is continuous, all we need to show is that $\alpha_i(x)$ is continuous at $\underline{x}$. Pick any $\varepsilon > 0$, and let $\delta = \varepsilon$. Then, for all $x \in \mathbb{R}$ such that $|x - \underline{x}| < \delta$, we have $|\alpha_i(x) - \alpha_i(\underline{x})| < \varepsilon$. ■

Note that both autonomy functions α_i and β_i are consistent with Nielsen's value-foundational satiability. Specifically, both autonomy functions "identif[y] the threshold [namely, $\underline{x}$] as the point above which any person will not become better-off in terms of the relevant value [in this case, autonomy] by having more of whatever can be allocated to her [in this case, resources]." (Nielsen 2019: 806–807) But, unlike β_i, which is *discontinuous* at the critical sufficiency threshold $\underline{x}$, α_i is *continuous* at the critical sufficiency threshold $\underline{x}$.

What these examples show is that endorsing value-foundational satiability does not necessarily commit us to endorse discontinuity. We used autonomy as an example, but the same thing can be said to any other satiable foundation value such as freedom, capability, reasonable contentment, and so on. Hence, even if one successfully builds sufficientarianism or any theory of justice on the basis of value-foundational satiability as Nielsen proposes, such fact does not explain away why the demands of justice may be discontinuous. As a consequence, it is a mistake to dismiss the problem of discontinuity raised against sufficientarianism merely as an instance of *the numbers fallacy*.

Then, the question remains. If discontinuity is indeed a *cost*, and if PU can avoid the cost while retaining all the main attractions of sufficientarianism, why not accept PU instead of sufficientarianism?

6. Sufficientarianism and Lifeboat Cases

Let us now move on to lifeboat cases, which were originally introduced by Frankfurt (1987: 30)[17]. In a typical lifeboat case, we are faced with a choice of two options:

[17] See Chung (2016) for a critical discussion of Frankfurt's seminal work on sufficientarianism.

- ***Some Survive***: Save some people by allocating resources unequally, or
- ***All Die***: Allocate resources equally and let everyone die.

(Telic) Egalitarianism claims that (undeserved) inequality is bad in itself. Prioritarianism claims that benefiting people matters more the worse off these people are. (Parfit 1997: 213) Both egalitarianism and prioritarianism imply an equal distribution of resources. In so far as Crisp (2003)'s and Huseby (2010)'s recent non-headcount versions of sufficientarianism endorse prioritarianism below the critical sufficiency threshold[18], they also imply an equal distribution of resources below the critical sufficiency threshold; see also Brown (2005), Casal (2007), Hirose (2016), and Bossert, Cato, and Kamaga (2022, 2023) for related accounts of sufficientarianism. Since everybody in a lifeboat scenario is clearly below the critical sufficiency threshold (however this may be defined), egalitarianism, prioritarianism, and Crisp's and Huseby's non-headcount sufficientarianisms all imply that we should choose *All Die* instead of *Some Survive*. Insofar as we agree that choosing *Some Survive* is the right answer,[19] lifeboat cases pose a problem for all these views. This is the gist of Chung (2017)'s original argument. Of course, Chung makes it clear that

[18] According to Crisp, "absolute priority is to be given to benefits to those below the threshold at which compassion enters. Below the threshold, benefitting people matters more the worse off those people are, the more of those people there are, and the greater the size of the benefit in question. Above the threshold, or in cases concerning only trivial benefits below the threshold, no priority is to be given." (Crisp 2003: 758) Similarly, Huseby claims that "First, individuals below the maximal sufficiency threshold should have absolute priority over individuals above this threshold. … Between the minimal and maximal sufficiency thresholds, I propose that we should apply a constrained and inverse form of prioritarianism. … Second, strong priority should be given to those below the minimal sufficiency threshold. By strong priority, I intend something less than absolute priority, but something more than straightforward weighted aggregation." (Huseby 2010: 184–185)

[19] We understand that some sufficientarians would not even consider that *Some Survive* is the right answer to lifeboat situation. For example, if there are two individuals and 10 units of resources, these sufficiantarians may think that (5, 5) is ethically acceptable or even ethically preferrable even when

> This is not to deny that there are ways for both egalitarians and prioritarians to [choose Some Survive] in such scenario. For instance, as long as the egalitarian and the prioritarian do not give *absolute* priority to the worst off, it could be possible for them to recommend [Some Survive] by allowing inegalitarian or anti-prioritarian distributions in exceptional cases in which sufficiently larges gains for a sufficient number of people who aren't worst-off will outweigh a gain to the worst-off. However, the point is that both egalitarianism and prioritarianism will be able to give right answers to lifeboat situations only allowing exceptions, not as a matter of principle. Note that the same criticism would apply to both Crisp's and Huseby's versions of sufficientarianism, as both versions endorse prioritarianism below the critical sufficiency threshold. (Chung 2017: 1914)

Davies argues that, unlike Chung's accusations, the two recent *non-headcount* versions of sufficientarianism by Crisp (2003) and Huseby (2010) are able to provide right answers to lifeboat cases by considering *benefit size* and that "the idea of allowing benefit size to outweigh the importance of benefits going to the worst off is not an *ad hoc* exception, but rather a principle that applies in all cases." (Davies 2022: 337)

We would first like to point out that there are quite a few places where both Crisp and Huseby seem to acknowledge that their non-headcount versions of sufficientarianism cannot provide right answers to lifeboat cases even if we consider benefit size. For instance, Crisp explains that "[o]ne possible problem with [his] view is … that the view will prefer the

---

the sufficiency threshold is 10. However, this is not Davies's stance: unlike these sufficientarians, Davies agrees that *Some Survive* is the right answer to lifeboat situations, and, instead, tries to argue that sufficientarianism can prescribe *Some Survive* in lifeboat situations by considering benefit size.

smallest nontrivial benefit to any number of individuals below the threshold to any benefit, no matter how large, to any number of individuals above the threshold." (Crisp 2003: 758) Similarly, Huseby explains that a potential problem with his view is that it may allow "giving a small benefit to a person who is below the sufficiency threshold" at the expense of giving "a much larger (or extremely much larger) benefit to many (or extremely many) people above the sufficiency threshold." (Huseby 2010: 186) Huseby calls this general problem (of which giving wrong answers to lifeboat scenarios is a special case) "the problem of *waste*." (Huseby 2010: 186) (This is essentially what Nielsen called the problem of outweighing priority.)

Unlike what Davies thinks, neither Crisp nor Huseby thought that considerations of benefit size will help their non-headcount versions of sufficientarianism provide right answers to lifeboat cases. Instead, Crisp simply bites the bullet and argues that the problem of waste, "may not be as implausible as it seems once we give proper recognition to the fact that the threshold is the point at which compassion no longer applies." (Crisp 2003: 758) Similarly, instead of saying that there is a way to overcome the problem of waste by considering benefit size, Huseby tries to dilute the severity of the objection by pointing out that other well-known distributive principles all face the same problem: "Egalitarianism, absolute prioritarianism, and the difference principle all demand waste also in situations *where everyone is insufficiently well off*." (Huseby 2010: 187 emphasis his) His response is not that there is a way to solve the problem by considering benefit size; rather, he finds such a problem a *cost* of sufficientarianism that he is willing to bear:

> In my view, sufficiency is, all things considered, a more defensible principle than its main alternatives. The principle is grounded in the concern for the badly off. This

concern has as its cost the problem of waste. I find this cost acceptable. (Huseby 2010: 187)

In sum, both Crips and Huseby acknowledged that their non-headcount versions of sufficientarianism may fail to provide right answers to lifeboat cases and they also did not think that considerations of benefit size can be used to overcome such shortcomings.

But maybe both Crisp and Huseby were not perfectly aware of the true potential of their theoretical framework in coping with lifeboat cases. Hence, let us be maximally charitable and examine what formal requirements are necessary to construct a version of sufficientarianism that can provide right answers to lifeboat cases in the specific way that Davies envisions by considering benefit size.

The key is to define sufficientarianism's social welfare function in a way that accounts for the (huge) benefit that accrues to somebody meeting their sufficiency threshold. We may do this by assuming two critical values of transferable resources (or utility): (a) a critical level $\alpha > 0$, and (b) a critical sufficiency threshold $\theta \geq 0$, where $\alpha \geq \theta$. Now, take any continuous, concave, and strictly increasing function g: $\mathbb{R} \rightarrow \mathbb{R}$, and define the sufficientarian social welfare function (SWF) as follows:

$$S(x) = \sum_{i \in N : x_i < \theta} [g(x_i) - g(\alpha)],$$

where $x = (x_1, \ldots, x_n)$ denotes any distribution and x_i denotes the amount of resources that individual i receives in distribution x.[20] Then, for any two distributions $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$, our sufficientarian theory will strictly ethically prefer x to y if and only if:

[20] This is similar to the sufficientarian SWF introduced by Bossert, Cato, and Kamaga (2023) under the name of generalized critical-level sufficientarianism. In a variable-population setting, these authors consider a Paretian sufficientarianism, where a critical level is different from the threshold. By contrast, the sufficientarian SWF proposed in this paper does *not* satisfy the Pareto principle because all changes in utilities/well-beings above the threshold do not increase the sufficientarian

$$S(x) > S(y) \Leftrightarrow \sum_{i \in N: x_i < \theta} [g(x_i) - g(\alpha)] > \sum_{i \in N: y_i < \theta} [g(y_i) - g(\alpha)].$$

Since g is concave, our sufficientarian SWF has prioritarian tendencies when everybody is below the sufficiency threshold and is unable to reach it. Let us now decompose our sufficientarian SWF into two components:

$$S(x) = \underbrace{\sum_{i \in N: x_i < \theta} [g(x_i) - g(\theta)]}_{\text{First Component}} + \underbrace{\sum_{i \in N: x_i < \theta} [g(\theta) - g(\alpha)]}_{\text{Second Component}}$$

If we use $m < |N| = n$ to denote the number of those below the critical sufficiency threshold θ, this can be written as follows:

$$S(x) = \underbrace{\sum_{i \in N: x_i < \theta} [g(x_i) - g(\theta)]}_{\text{Shortfall Component}} + \underbrace{m[g(\theta) - g(\alpha)]}_{\text{Headcount Component}}.$$

Then, how should we interpret these two terms? Note that the first term (i.e., the Shortfall Component) is negative unless the set of those below the critical sufficiency threshold θ is empty, while the second term (i.e., the Headcount Component) is negative unless the critical level α coincides with the critical sufficiency threshold θ or m is zero. Together, the value of our sufficientarian SWF $S(x)$ cannot be positive.

Note that for each individual $i \in N$, $[g(x_i) - g(\theta)]$ represents the disvalue of i falling short of meeting their sufficiency threshold θ. The disvalue of this gap $[g(x_i) - g(\theta)]$ decreases as x_i approaches the critical sufficiency threshold θ and becomes zero when x_i

SWF and are not taken into account. This implies that our sufficientarian SWF $S(x)$ is congruent with Nielsen's value satiability assumption and satisfies sufficientarianism's negative thesis.

reaches or exceeds θ. The first term of our sufficientarian SWF sums the disvalues of such a gap across all individuals who are below the sufficiency threshold θ: i.e., $\sum_{i \in N: u_i < \theta} [g(x_i) - g(\theta)]$. Therefore, we can say that the first term (i.e., the Shortfall Component) of our sufficientarian SWF represents the overall negative impact on society caused by the cumulative gap of individuals falling below the critical sufficiency threshold θ in a specific distribution. As noted above, we can see that our sufficientarian SWF has a prioritarian tendency when everybody is below their sufficiency threshold. This is because, unless the second term changes, the overall disvalue of the sufficientarian SWF can most effectively be reduced by distributing any available resources or utility to the worse off.

Then, what about the second term, $m[g(\theta) - g(\alpha)]$, the Headcount Component? When somebody who was previously below the sufficiency threshold θ successfully reaches or surpasses it, the total number of individuals who are below the sufficiency threshold θ decreases from m to $m - 1$. As a result, the value of the sufficientarian SWF increases by an increment of $[g(\alpha) - g(\theta)]$ (which is the difference between the value of the critical level α and the value of the critical sufficiency threshold θ) for each additional person reaching their critical sufficiency threshold θ. We might think of $[g(\alpha) - g(\theta)]$ as representing the *size* of the *moral benefit* of letting somebody reach their critical sufficiency threshold. Or, to put differently, we might think of $[g(\theta) - g(\alpha)]$ as representing the negative moral value of somebody being below the critical sufficiency threshold θ, and $m[g(\theta) - g(\alpha)]$ as representing the cumulative moral disvalue of having m individuals below the critical sufficiency threshold θ.

For instance, suppose $x_i < \theta$, that is, individual i is below the sufficiency threshold, and suppose we move individual i from x_i to x_i' (where $x_i < x_i'$.) If $x_i < x_i' < \theta$, then the increased moral value of the distribution becomes: $g(x_i') - g(x_i)$. However, if $x_i < \theta < x_i'$,

the increased moral value of the distribution becomes: $[g(x_i') - g(x_i)] + [g(\alpha) - g(\theta)]$. That is, there is an additional moral value that gets added to the distribution by allowing somebody to not just reduce their gap while still falling short of the critical sufficiency threshold θ, but allowing that person to successfully reach the critical sufficiency threshold θ.

We can see that our sufficientarian SWF $S(x)$ is designed to find a distribution that minimizes the negative moral disvalue of the distribution. This is in accordance with Nielsen (2023)'s recent proposal:

> This invites the idea that sufficientarianism should be interpreted negatively in the sense of focusing on elimination of deficiency rather than on securing enough of some given currency. Thus, we arrive at the following generic principle:
>
> (S) A distribution is just if, and only if, no one suffers deficiencies from a justice-relevant threshold. (Nielsen 2023: 17)

Our sufficientarian SWF $S(x)$ is also in line with value-satiability sufficientarianism as it assumes that the moral value that accrues to the distribution by giving more to a given individual is sated once that individual reaches their sufficiency threshold θ. Following Nielsen, let us call our formalization of sufficientarianism, *sufficentarianism S*: *Sufficientarianism S* generates a moral order over the set of distributions in accordance with the sufficientarian SWF $S(x)$. Would *sufficientarianism S* be able to now provide right answers to lifeboat situations?

Proposition 7. *In any lifeboat situation, Sufficientarianism S always maximizes the number of instances meeting their critical sufficiency threshold θ as long as the critical level α is sufficiently large.*

Proof. Consider any lifeboat situation where the available resources are $k\theta$, where $k < |N| = n$. So, we have just enough resources to allow $k < n$ people meet their critical sufficiency threshold θ. Let $x^k = \left(x_1^k, \dots, x_n^k\right)$ be any distribution of the transferable goods that makes k (the maximum number of) individuals meet their critical sufficiency thresholds θ, and, without loss of generality, rearrange the individuals so that $x_i^k = \theta$ for $i = 1, \dots, k$ and $x_j^k = 0$ for $j = k+1, \dots, n$, i.e.,

$$x^k = \left(\underbrace{\theta, \theta, \theta, \dots, \theta}_{(k \text{ individuals})}, \underbrace{0, 0, 0, \dots, 0}_{(n-k \text{ individuals})}\right).$$

We claim that distribution x^k maximizes our sufficientarian SWF. To show this, we will show that that there is no way to increase the value of our sufficientarian SWF by moving to an alternate distribution. Obviously, moving to any other distribution that makes a different set of k individuals meet their critical sufficiency thresholds θ will generate the same value for our sufficientarian SWF, and, hence, will not increase its value. Now, consider moving to another distribution $x^a = (x_1^a, \dots, x_n^a)$ under which $a < k$ individuals meet their critical sufficiency thresholds θ. Without loss of generality, let the individuals from 1 to a be the individuals who meet their sufficiency threshold θ in distribution x^a. Then, for individuals $i \in \{a+1, a+2 \dots, k\}$, we have $x_i^k = \theta > x_i^a$. Since g is increasing and concave, the best distribution that would generate the highest value for the sufficientarian SWF, while having $a < k$ individuals meet their critical sufficiency thresholds θ, would give $x_i^a = \theta$ for $i = 1, \dots, a$ and give $x_j^a = \frac{k-a}{n-a}\theta < \theta$ for $j = a+1, \dots, n$, i.e.,

$$x^a = \left(\underbrace{\theta, \dots, \theta}_{(a \text{ individuals})}, \underbrace{\frac{k-a}{n-a}\theta, \dots, \frac{k-a}{n-a}\theta}_{(n-a \text{ individuals})} \right).$$

According to our sufficientarian SWF, for individuals $i \in \{a+1, a+2 \dots, k\}$, moving from $x_i^k = \theta$ to $x_i^a = \frac{k-a}{n-a}\theta$ generates the following total social loss:

$$-\underbrace{(k-a)}_{\text{a total of } k-a \text{ individuals}} \left\{ \underbrace{\left[g\left(x_i^k = \theta\right) - g\left(x_i^a = \frac{k-a}{n-a}\theta\right) \right]}_{i's \text{ welfare loss}} + \underbrace{[g(\alpha) - g(\theta)]}_{\text{moral disvalue of failing to meet threshold } \theta} \right\} \cdots (1)$$

For individuals $j = k+1, k+2, \dots, n$, moving from $x_j^k = 0$ to $x_j^a = \frac{k-a}{n-a}\theta$ generates the following total social gain:

$$(n-k) \underbrace{\left[g\left(x_j^a = \frac{k-a}{n-a}\theta\right) - g\left(x_j^k = 0\right) \right]}_{j's \text{ welfare gain } (j=k+1,\dots,n.)} \cdots (2)$$

A move from distribution x^k to distribution x^{k-1} would *not* be morally preferable, if and only if,

$$(1) + (2) \leq 0$$

$$\Leftrightarrow (n-k)\left[g\left(\frac{k-a}{n-a}\theta\right) - g(0) \right] + (k-a) g\left(\frac{k-a}{n-a}\theta\right) \leq (k-a) g(\alpha)$$

$$\Leftrightarrow \frac{n-a}{k-a} g\left(\frac{k-a}{n-a}\theta\right) - \frac{n-k}{k-a} g(0) \leq g(\alpha)$$

which is true whenever the critical level α is sufficiently high. Hence, in any lifeboat situation, moving from a distribution that makes the maximum number of individuals meet their critical sufficiency thresholds θ to another distribution that makes a lesser number individuals meet their critical sufficiency thresholds θ will never increase the value of the sufficientarian SWF as long as the critical level α is sufficiently high. ■

Proposition 7 implies that *sufficentarianism S* can provide right answers to lifeboat situations by recommending *Some Survive* instead of *All Die* as long as it *assumes* that the 'size' of the moral benefit (i.e., $[g(\alpha) - g(\theta)]$) that accrues to the distribution *in addition to* the gain in individual welfare of making somebody reach their critical sufficiency threshold θ is sufficiently large.[21] Defined in this way, we can see that there is a sense in which *sufficentarianism S* begs the very question it tries to answer: specifically, it presupposes that the ethical significance of ensuring individuals reach their critical sufficiency threshold θ is substantial enough to demonstrate that it will endorse a distribution that maximizes the number of individuals attaining the said threshold. In this context, the determination of what qualifies as assigning an appropriately high moral significance to meeting the critical sufficiency threshold θ is defined *endogenously* by its capacity to yield morally correct resolutions in lifeboat scenarios.

Note that *the size* of the moral benefit (i.e., $[g(\alpha) - g(\theta)]$) of reaching the critical sufficiency threshold in our *sufficentarianism S* is measured by the difference between the value of the critical level α (i.e., $g(\alpha)$) and the value of the critical sufficiency threshold θ (i.e., $g(\theta)$). To give right answers to lifeboat situations, the critical level α must sufficiently greater than the critical sufficiency threshold θ. However, unless the critical level α is identical to the critical sufficiency threshold θ, the moral ordering generated by

[21] Interestingly, the pure headcount approach can be seen as a limit of our sufficientarian theory. More precisely, sufficientarianism *S* approaches to headcount sufficientarianism as $[g(\alpha) - g(\theta)]$ approaches to infinity. It is noteworthy that our sufficientarianism can be regarded as a hybrid theory because the first term corresponds to prioritarianism below the threshold, and the second term essentially corresponds to the headcount approach to sufficientarianism. In this sense, $[g(\alpha) - g(\theta)]$ is a "weight" on headcount sufficientarianism. This offers an intuitive reason why headcount sufficientarianism can be obtained as a limit.

sufficentarianism S will necessarily display discontinuity at the critical sufficiency threshold θ.

Proposition 8. *The moral ordering of Sufficientarianism S is discontinuous at the critical sufficiency threshold θ whenever the critical level α exceeds the critical sufficiency threshold θ, i.e., whenever $\alpha > \theta$.*

Proof. According to our sufficientarian SWF, a distribution x is strictly morally preferred to another distribution y if and only if

$$S(x) = \sum_{i \in N: x_i < \theta} [g(x_i) - g(\alpha)] > \sum_{i \in N: y_i < \theta} [g(y_i) - g(\alpha)] = S(y).$$

If $\alpha = \theta$ (i.e., if the critical level α is equal to the critical sufficiency threshold θ), then, for any distribution $x = (x_1, \dots, x_n)$, our sufficientarian SWF can be re-written as:

$$S(x) = \sum_{i \in N} \min\{g(x_i) - g(\theta), 0\}.$$

We note that both $g(x_i) - g(\theta)$ and $h(x) = 0$ (i.e., a constant function) is continuous. Therefore, $\min\{g(x_i) - g(\theta), 0\}$ is also continuous. Moreover, the sum of continuous functions yields a continuous function. We thus conclude that the moral ordering induced by our sufficientarian SWF, S, is continuous when $\alpha = \theta$.

Now, suppose $\alpha > \theta$. Consider a distribution $x = (\theta, \dots, \theta)$ where everybody meets their critical sufficiency threshold θ. Note that $S(x) = 0$. Now, consider a sequence of distributions $\{x^k\} = \left(\theta - \frac{1}{k}, \dots, \theta - \frac{1}{k}\right)$ for each $k \in \mathbb{N}$. Note that $x^k \longrightarrow x$ (i.e., x^k converges to x.) Then, we have:

$$\lim_{k \to \infty} S(x^k) = \lim_{k \to \infty} \left\{ n \left[g\left(\theta - \frac{1}{k}\right) - g(\theta) \right] + n[g(\theta) - g(\alpha)] \right\}$$

$$= [g(\theta) - g(\theta)] + n[g(\theta) - g(\alpha)]$$

$$= n[g(\theta) - g(\alpha)] \neq 0 = S(x).$$

Hence, our sufficientarian SWF is discontinuous at $x = (\theta, \dots, \theta)$. As an illustration, consider a two-person case and consider two distributions: $u = (u_1, u_2) = (\theta, \theta)$ and $v = (v_1, v_2) = (\theta, 0)$ and suppose $g(\alpha) > 2g(\theta) - g(0)$. Then, since $S(u) = 0 > g(0) - g(\alpha) = S(v)$, we have $u \succ v$. Now, consider a sequence $\{u^k\} = \left(u_1^k, u_2^k\right) = (\theta - \frac{1}{k}, \theta - \frac{1}{k})$. Note that $u^k \longrightarrow u$.

Then, from $g(\alpha) > 2g(\theta) - g(0)$, we have, for all $k \in \mathbb{N}$:

$$g(\alpha) > 2g(\theta) - g(0)$$

$$\Rightarrow\ g(0) - g(\alpha) > 2[g(\theta) - g(\alpha)] > 2\left[g\left(\theta - \frac{1}{k}\right) - g(\alpha)\right]$$

$$\Rightarrow g(0) - g(\alpha) > 2\left[g\left(\theta - \frac{1}{k}\right) - g(\alpha)\right]$$

$$\Rightarrow S(v) > S(u^k)$$

and, hence, $v \succ u^k$. Hence, there exists no $k^* \in \mathbb{N}$ such that $u^k \succ v$ for all $k > k^*$. So, the moral ordering of sufficientarian *S* is discontinuous at the critical sufficiency threshold θ. ■

Hence, even if we are able to formulate sufficientarianism (negatively) to provide right answers to lifeboat situations by considering 'benefit size' as Davies proposes, we can only do so by completely giving up continuity. Thus, the question remains: if offering appropriate solutions to lifeboat situations is indeed a desired outcome and if generating discontinuous ethical judgments is indeed a cost, and if PU can circumvent such costs while preserving the key merits of sufficientarianism, why not embrace PU instead of sufficientarianism?

7. Concluding Remarks

In this paper, we have argued that both Davies's and Niesen's respective defenses of sufficientariarnism are not entirely successful. Defined precisely, continuity is a much more plausible normative requirement than what people normally think. Even if a discontinuous ethical theory can be defended all things considered, this does not change the fact that discontinuity is a theoretical cost. Also, providing right answers to lifeboat situations is a desideratum for any principle of distributive ethics. We have explored how sufficientarianism can formally provide right answers to lifeboat situations by assuming two critical values: (a) a critical sufficiency threshold, and (b) a critical level that is sufficiently larger than the said critical sufficiency threshold. However, doing so necessarily made sufficientarianism discontinuous at the critical sufficiency threshold. So, we will ask one last time: if violation of continuity is indeed a *cost*, and if PU can avoid the cost while retaining all the main attractions of sufficientarianism including providing right answers to lifeboat situations, why not accept PU instead of sufficientarianism?

References

- Adler, Matthew D. (2018). Prioritarianism: Room for Desert?. *Utilitas* 30: 172–97.
- Arrhenius, Gustaf, and Wlodek Rabinowicz. (2005). Millian Superiorities. *Utilitas* 17: 127–46.
- Benbaji, Yitzhak. (2005). The Doctrine of Sufficiency: A Defence. *Utilitas* 17: 310–332.
- Binmore, Ken, and Alex Voorhoeve. (2003). Defending Transitivity against Zeno's paradox. *Philosophy & Public Affairs* 3: 272–79.
- Blackorby, Charles, Walter Bossert, and David Donaldson. (2005). *Population Issues in Social Choice Theory, Welfare Economics, and Ethics*. Cambridge: Cambridge University Press
- Bossert, Walter, Susumu Cato, and Kohei Kamaga. (2022). Critical-level Sufficientarianism. *Journal of Political Philosophy* 30: 434–61.

- Bossert, Walter, Susumu Cato, and Kohei Kamaga. (2023). Thresholds, critical levels, and generalized sufficientarian principles. *Economic Theory 75*: 1099–1139.
- Brown, Campbell. (2005). Priority or Sufficiency… or Both? *Economics & Philosophy* 21: 199–220.
- Casal, Paula. (2007). Why Sufficiency is Not Enough. *Ethics* 117: 296–326.
- Chung, Hun. (2016). Is Harry Frankfurt's 'Doctrine of Sufficiency' Sufficient? *Organon F* 23 (1): 50-71
- Chung, Hun. (2017). Prospect Utilitarianism: A Better Alternative to Sufficientarianism. *Philosophical Studies* 174: 1911–33
- Chung, Hun. (2022). Prospect Utilitarianism and the Original Position. *Journal of the American Philosophical Association*, 1–35. doi:10.1017/apa.2022.31
- Davies, Ben. (2022). The Prospects for 'Prospect Utilitarianism.' *Utilitas* 34: 33–43.
- Fleurbaey, Marc. (2015). Equality Versus Priority: How Relevant is the Distinction?. *Economics & Philosophy* 31: 203–17.
- Hirose, Iwao. (2016). Axiological sufficientarianism. In: C. Fourie and A. Rid. *What is Enough? Sufficiency, Justice, and Health*. Oxford: Oxford University Press. pp. 51–68.
- Huseby, Robert. (2010). Sufficiency: Restated and defended. *Journal of Political Philosophy* 18: 178–97.
- Kahneman, Daniel, and Amos Tversky. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica* 47: 263–92.
- Kaplow, Louis, and Steven Shavell (2001). Any Non-welfarist Method of Policy Assessment Violates the Pareto Principle. *Journal of Political Economy* 109: 281–86.
- Nielsen, Lasse. (2019). Sufficiency and Satiable Values. *Journal of Applied Philosophy* 36: 800–16.
- Nielsen, Lasse. (2023). The Numbers Fallacy: Rescuing Sufficientarianism from Arithmeticism. *Inquiry* 1–22. https://doi.org/10.1080/0020174X.2023.2186950
- Nussbaum, Martha. (2000). *Women and Human Development*. Cambridge: Cambridge University Press.
- Parfit, Derek. (1997). Equality and Priority. *Ratio* 10: 202–21.
- Raz, Joseph. (1986). *The Morality of Freedom*. Oxford: Clarendon Press.
- Roemer, John. (2004). Eclectic Distributional Ethics. *Politics, Philosophy and Economics* 3: 267–81.

- Sen, Amartya. (1979). Personal Utilities and Public Judgements: Or What's Wrong with Welfare Economics? *Economic Journal* 89: 537–58.
- Stefánsson, H. Orri. (2022). Continuity and Catastrophic Risk. *Economics & Philosophy* 38: 266–74.
- Temkin, Larry. S. (1996). A Continuum Argument for Intransitivity. *Philosophy & Public Affairs*,25: 175–210.
- Temkin, Larry S. (2001). Inequality: a complex, individualistic, and comparative notion. *Philosophical Issues* 11: 327–53.