Uncertainty, Equality, Fraternity

Rush T. Stewart

MCMP, LMU Munich

April 21, 2021

Abstract

Epistemic states of uncertainty play important roles in ethical and political theorizing. Theories that appeal to a "veil of ignorance," for example, analyze *fairness* or *impartiality* in terms of certain states of ignorance. It is important, then, to scrutinize proposed conceptions of ignorance and explore promising alternatives in such contexts. Here, I study Lerner's probabilistic egalitarian theorem in the setting of imprecise probabilities. Lerner's theorem assumes that a social planner tasked with distributing income to individuals in a population is "completely ignorant" about which utility functions belong to which individuals. Lerner models this ignorance with a certain uniform probability distribution, and shows that, under certain further assumptions, income should be equally distributed. Much of the criticism of the relevance of Lerner's result centers on the representation of ignorance involved. Imprecise probabilities provide a general framework for reasoning about various forms of uncertainty including, in particular, ignorance. To what extent can Lerner's conclusion be maintained in this setting?

Keywords. Distributive justice; egalitarianism; imprecise probabilities; Lerner's theorem; uncertainty; utilitarianism; veil of ignorance

1 Introduction

The veil of ignorance is an important theoretical construction for moral and political philosophy. For Rawls, just arrangements of social institutions and general principles that govern them are those that members of a society would choose under a hypothetical state of ignorance about the citizens' individual social and economic standing, their abilities, etc. (Rawls, 1971). The idea is that such ignorance insulates choice from bias and unfair tailoring of policies to special interests. Similarly, in making his case for utilitarianism, Harsanyi assumes that an agent choosing between social systems without knowing his or her position in them is making a *moral* value judgment. On the question of the just distribution of income, he writes, "a value judgment on the distribution of income would show the required impersonality to the highest degree if the person who made this judgment had to choose a particular income distribution in *complete ignorance* of what his own relative position (and the position of those near to his heart) would be within the system chosen" (Harsanyi, 1953, pp. 434-435,

emphasis mine). "This," Harsanyi reflects later, "is my own version of the concept of the 'original position'" (Harsanyi, 1975, p. 598).

Prior to either Harsanyi's or Rawls's invocation of some sort of original position, Abba Lerner appealed to something of a conceptual forerunner of the veil of ignorance in the setting of an interesting theorem concerning distributive justice (1946).[1] Suppose that there is a fixed sum of money to be distributed to individuals in a population. Importantly, Lerner stipulates, the distribution must be chosen under ignorance as to who has which utility function. Given Lerner's particular formulation of ignorance and assumptions about individual and social welfare, the *equal distribution* maximizes *expected* social welfare. If the goal is to maximize expected social welfare, then Lerner provides an argument for an egalitarian distribution of income.

Ignorance in the intended setting, according to Lerner, amounts to it being equally likely, for any two people in the population, that one person possesses a given utility function as that the other possesses it. Ignorance, in other words, is represented as all matchings of utility functions to individuals being *equally probable*. Harsanyi conceives of ignorance along similar lines. Continuing the passage quoted earlier, Harsanyi claims that an agent would choose an income distribution "in complete ignorance" of his own relative position in society "if he had *exactly the same chance* of obtaining the first position (corresponding to the highest income) or the second or the third, etc., up to the last position (corresponding to the lowest income) available within that scheme" (Harsanyi, 1953, p. 435, emphasis mine). This way of thinking about ignorance has been regarded with a good deal of skepticism in general—being in essence an application of the principle of indifference—and, as McManus et al. note, in the context of Lerner's theorem in particular: "Of particular interest in the past has been Lerner's assumption of complete ignorance about which individual has which utility function, and his jump from that assumption to the hypothesis that each possible matching is equally likely" (1972, p. 494). About the relevant probabilities, Graaff, for example, writes, "There is no justification for assuming them equal. From absolute ignorance we can derive nothing but absolute ignorance" (1957, p. 100 fn). Similarly, Little, remarking on Lerner's theorem, claims, "From complete ignorance nothing but complete ignorance can follow" (1950, p. 59).

How should we model ignorance in Lerner's intended setting? Here, I consider Lerner's theorem in the setting of imprecise probabilities (IP). By the lights of many, IP allows for a more general and compelling representation of uncertainty and of "complete ignorance" in particular than standard probability models do (e.g., Keynes, 1921; Good, 1952; Levi, 1974; Walley, 1991; Kaplan, 1996; Joyce, 2011; Hájek and Smithson, 2012; Weatherson, 2015; Stewart and Ojea Quintana, 2018; Hill, 2019). So it is interesting and natural, it would seem, to consider arguments that appeal to some version of the veil of ignorance or related ideas—as in Lerner's theorem—in this more general setting. This allows us to perform a robustness check of sorts. Lerner's particular representation of ignorance, the equi-probability assumption, constitutes a sticking point in the appreciation of his theorem, as we have just seen. To what extent does his conclusion depend on this assumption? Some robustness analysis has been performed, with theorems that establish the equal distribution of income as optimal according to the maximin rule for decisions under ignorance (Theorem 2). But

---

[1]Paul Samuelson recasts Lerner's theorem as an argument appealing more explicitly to a sort of veil of ignorance (Samuelson, 1964, pp. 175–176). There are, of course, antecedent uses of related ideas. For some of this historical background, see, for example, (Freeman, 2019, § 1).

this rule is extremely conservative, moves decidedly away from the maximization of expected utility, and is susceptible to certain intuitive counterexamples (e.g., Luce and Raiffa, 1957, pp. 279–280).

After rehearsing the basics of the framework and stating a general version of Lerner's theorem due to Sen in Section 2, I introduce and motivate IP as a framework for representing and reasoning about uncertainty in Sections 3 and 4. The primary conceptual move is to employ IP theory to model ignorance in the context of ethical arguments that exploit ignorance, e.g., veil of ignorance arguments or, in the case at hand, Lerner's theorem.[2] A crucial issue here is that there are several candidate decision criteria for IP, each of which can claim to generalize expected utility maximization. As a result, the admissibility of the egalitarian distribution is a more subtle issue than in the standard setting. In the IP setting, we find that, from the assumption of total ignorance, we do not automatically arrive at egalitarianism, partially vindicating those skeptical of Lerner's "jump" from complete ignorance to equal probability. It depends, in part, on the IP generalization of expected utility that we adopt. But, *pace* Little and Graaff, it is not true that we arrive at *nothing* from total ignorance. Even for the most extreme form of ignorance in the IP setting, what we arrive at depends, again, on the decision rule that we adopt. The admissibility of the equal distribution is fairly robust for IP decision rules under complete ignorance (Observation 1). Moreover, certain properties of social welfare that admit ethical motivation entail the *unique* admissibility of the equal distribution under complete ignorance for some decision criteria (Observation 2). On the other hand, certain intuitively very unjust distributions are likewise admissible for some of the decision rules considered (Example 1). To the extent one finds the implied decision theory compelling, this calls into question the inference from a distribution's choiceworthiness under ignorance to its status as fair or just.

2 (Sen's Version of) Lerner's Theorem

In presenting Lerner's theorem, I follow Sen's more general version (1973). Sen's version of the theorem generalizes away from the assumption that social welfare is additive or even separable, which already responds to certain objections to Lerner's original formulation (Friedman, 1947). We assume that there is a group N of individuals, $i = 1, \ldots, n$. We also assume that there is a collection of n utility functions, $U^1, \ldots, U^n$.[3] It is unknown—and this is the crucial point—which utility function is associated with which individual. Let y be any income vector $(y_1, \ldots, y_n)$, with y_i denoting the income of individual i. Let z be a vector of equal incomes, $z_i = z_j$ for all $i, j \in N$. A function of n arguments is called *symmetric* if the value of the function is constant under permutations of the arguments. For income vectors, we say

---

[2]For related uses of IP theory, see (e.g., Levi, 1977; Gajdos and Kandil, 2008). Debates about appropriate epistemic states (e.g., Buchak, 2017; Stefánsson, 2019) and decision-theoretic principles (e.g., Kurtulmuş, 2012; Liang, 2017; Gustafsson, 2018) behind the veil remain active. Stefánsson (2019), for example, considers ambiguity averse preferences behind the veil, and finds such preferences support a form of egalitarianism. Liang (2017), to take another example, employs cumulative prospect theory and finds an optimal form of inequality. In contrast, the present paper considers a range of (purportedly normative) IP choice rules, focuses on Lerner's theorem rather than Rawls's theory, makes no central appeal to objective probabilities, and seems to reach relatively more equivocal conclusions about the extent to which a form of egalitarianism is or is not vindicated.

[3]Sen notes that the proofs to follow can be extended to cover larger, finite collections of utility functions. Following Sen, I will use "utility" and "welfare" interchangeably for individuals.

$y = (y_1, \ldots, y_n) \geq y' = (y'_1, \ldots, y'_n)$ if $y_i \geq y'_i$ for all $i \in N$. A function f is *increasing* if $y \geq y'$ implies $f(y) \geq f(y')$. A function $f : \mathcal{S} \to \mathbb{R}$ defined on a convex subset of a real vector space is called *concave* if for all $x, y \in \mathcal{S}$ and all $\lambda \in [0,1]$, $f(\lambda x + (1 - \lambda)y) \geq \lambda f(x) + (1 - \lambda)f(y)$. Standardly, concavity of individual utility is taken to reflect the rate at which the marginal utility of income decreases and, in the context of decisions under uncertainty, risk aversion. Concavity of a social welfare function, on the other hand, reflects a form of inequality aversion. The assumptions for Sen's generalization of Lerner's theorem are as follows.

(A.1) *(Total Income Fixity). There is a fixed sum y^* to be distributed:* $\sum_{i=1}^{n} y_i = y^*$.

(A.2) *(Concavity of the Group Welfare Function). Social welfare W, an increasing and symmetric function of individual utilities, is concave.*

(A.3) *(Concavity of Individual Welfare Functions). Individual welfare functions are concave.*

(A.4) *(Equi-probability). For each j, $p_i^j = p_m^j$, for all $i, m \in N$.*

Here, p_i^j is the probability that agent i has utility function U^j. One way to think about assumption A.4 is that a social planner tasked with deciding the distribution of income uses the uniform probability distribution over the possible matchings of utility functions to individuals.

For any income distribution y, let $\tilde{y}$ be a permutation of y such that $\tilde{y}^j$ is the income of the individual with the jth utility function. There are $n!$ ways of assigning n utility functions to n individuals. For each such assignment k, there is a particular permutation vector $\tilde{y}(k)$ reflecting a particular assignment of utility functions to individuals (or, maybe more carefully, a matching of incomes with utility functions). So for any income vector y, the $n!$ possible social welfare values are given by $F(\tilde{y}(k))$, $k = 1, \ldots, n!$, where F is the reduced or compound function $F(\tilde{y}(k)^1, \ldots, \tilde{y}(k)^n) = W(U^1(\tilde{y}(k)^1), \ldots, U^n(\tilde{y}(k)^n))$ for all y and all k (exploiting the symmetry of W to order profiles by the utility function index). A social planner facing the income distribution problem would know, for any given assignment k, that $\tilde{y}^j$ is the income going to the person with utility function U^j, but not know which individual in the population possesses that utility function. In general, the expectation of social welfare is

$$E_p(y) = \sum_{k=1}^{n!} F(\tilde{y}(k))p(k) \tag{1}$$

for some probability p. Under assumption A.4, expected social welfare is given by

$$E(y) = \frac{1}{n!} \sum_{k=1}^{n!} F(\tilde{y}(k)). \tag{2}$$

We can now state (Sen's version of) Lerner's theorem.

Theorem 1. *Given (A.1), (A.2), (A.3), and (A.4), expected social welfare is maximized by z, the equal distribution of income.*

As mentioned above, many find A.4 to be unmotivated as a representation of ignorance about who has which utility function. One attempt to respond to these concerns is to eschew equating ignorance with any probabilistic judgment whatsoever and to appeal to some decision theoretic alternative to expected utility theory. In particular, versions of a theorem to the effect that the equal distribution is a maximin strategy have been proved, demonstrating that Lerner's conclusion is not as dependent on A.4 as it might initially appear (Sen, 1969; McManus et al., 1972; Sen, 1973). The maximin policy is to maximize the minimum possible level of social welfare, where the possibilities are the assignments of utility functions to individuals. Dispensing with A.4, Sen makes two further assumptions (1973, Theorem 2).[4]

(A.5) *(Shared Set of Welfare Functions). For any individual i and any utility function j, it is possible that i has j.*

(A.6) *(Bounded Individual Utility Functions). Each function U^j is bounded from below.*

Theorem 2. *Given* (A.1), (A.2), (A.3), (A.5), *and* (A.6), *the equal distribution z is the maximin strategy for social welfare.*

As McManus et al. put it, "the worrisome hypothesis of equal probability is not necessary for Lerner's conclusion" (1972, p. 494).

However, maximin is itself of questionable normative status. For example, that o_2 and not o_1 is the maximin solution in the following decision problem gives some pause about the general appeal of the rule for decisions under uncertainty (Cf. Luce and Raiffa, 1957, pp. 279–280).

Table 1: A Problem with Maximin

	s_1	s_2	s_3	$\cdots$	s_{100}
o_1	10^{-7}	10^{60}	10^{60}	$\ldots$	10^{60}
o_2	10^{-6}	10^{-6}	10^{-6}	$\ldots$	10^{-6}

Option o_1's payoff from state 2 on can be increased arbitrarily and option o_2's payoff can be decreased to any positive number (provided o_1's payoff in s_1 is still lower) without altering maximin's verdict. Notice, however, that the force of this objection depends heavily on measurability assumptions about utilities. Still, with no probabilistic assessments of the states whatsoever, it is, at the very least, extremely difficult to evaluate such choices.

3 Complete Ignorance

The theory of imprecise probabilities has at its disposal the representation of a form of ignorance more severe than any that admits representation when it is assumed that probability judgments are numerically precise. There are a number of IP frameworks, but I will

---

[4]As Sen observes, assumption A.6 is stronger than he needs. The same is true for the generalizations below.

work with arbitrary *sets* of probability functions. In his review of Walley's treatise on IP, Wasserman takes note of this "nice feature": "A nice feature about [IP] is that there exists a transformation-invariant expression of ignorance, a holy grail in Bayesian statistics. To represent ignorance, we use the set $\mathbb{P}$ of all probabilities" (Wasserman, 1993, p. 701, notation changed for consistency).[5] Many advocates of IP have found the set of all probabilities to be an eminently natural representation of complete ignorance. Isaac Levi, for example, calls such a state "probabilistically ignorant in the extreme sense" (1977). Since *no* probability distribution is ruled out, $\mathbb{P}$ reflects no information about the relevant possibilities.

The primary concern about this proposal for representing ignorance has to do with what is sometimes called *belief inertia* (Levi, 1980; Vallinder, 2018; Bradley, 2019). After his flattering remarks, Wasserman continues, "It is tempting to conclude that this solves the problem of finding an objective prior for Bayesian inference. But a vacuous prior gives a vacuous posterior, no matter how much data we obtain. So we cannot represent ignorance after all, at least not in standard statistical problems [...] What a shame that we cannot drink from the grail" (1993, p. 701). But belief inertia isn't the slightest problem for using this representation of ignorance to carry out an exercise of making hypothetical choices behind the veil of ignorance.[6] There is no concern to learn or update our "initial" ignorance in this context.

Let $\mathbb{P}_i^j$ be the *set* of probabilities assigned to individual i having utility U^j. I propose to consider replacing Lerner's equi-probability assumption A.4 with the following assumption.

(A.7) *For each j, $\mathbb{P}_i^j = [0, 1]$ for all $i \in N$.*

One way to think about assumption A.7 is that a social planner responsible for the choice of income distribution uses the same *set* of probability distributions over the possible utility functions—the set of *all distributions*—for each individual. A.7 immediately implies that $\mathbb{P}_i^j = \mathbb{P}_m^j$ for all $i, m \in N$, in analogy to the equi-probability assumption A.4. A.7 also implies A.5 since, for any utility function j and individual i, $\mathbb{P}_i^j$ includes positive values (even 1) for the possibility that j belongs to i, which is not the case when it is not *possible* that i has j. On the one hand, A.7 is a very weak assumption, too weak, one might think, to derive something like Lerner's egalitarian conclusion from. On the other hand, A.7 would seem to be far less open to the objections leveled against the equi-probability assumption, A.4, as a representation of ignorance. However, Observations 1 and 2 below show that, together with the other assumptions discussed, interesting content remains.

We have assumed a fixed, finite set of utility functions—in fact, of the same cardinality as the population (but see footnote 3). But there could be greater uncertainty about the appropriate set of possible utility functions. A larger set of possible individual utility functions would imply a larger set of possible assignments of utility functions to individuals. So, we

---

[5]The relevant notion of transformation invariance amounts to the possibility of reparameterizing or re-describing the space of possibilities without thereby having to alter the probabilities assigned to events, an issue with which so-called *objectivist* methods struggle.

[6]Levi thinks that his concept of *confirmational commitment* can help to cope with the problem of belief inertia (1980, §13.2, §13.3). A confirmational commitment is a function from states of full belief or evidence to sets of probabilities and can be rationally revised under some circumstances, including in the case of ignorance. Such a theory of the revision of confirmational commitments would entail that being in a state of complete ignorance at one point does not preclude an agent from making more informative probability judgments in the future.

could be confronted with a state of what Leve calls "modal ignorance" that is more severe than we have assumed (1977). (A smaller set of possible assignments of utility functions to individuals is also possible—if the social planner knows that one particular utility function belongs to a particular individual, for example—and would be a less severe state of modal ignorance than we are assuming.) In such a state of greater modal ignorance, we could still face probabilistic ignorance regarding the relative likelihoods of the various assignments. While I will not pursue this issue here, extending Lerner's theorem to significantly greater states of modal ignorance may provide for an even more severe and convincing robustness check on Lerner's egalitarian argument.[7]

4 Choice under IP

Part of the debate between Harsanyi and Rawls concerns, not only the form of ignorance faced, but the appropriate decision rule to use behind the veil of ignorance. Where Harsanyi advocates the maximization of expected utility, making certain assumptions of equi-probability, Rawls advocates a non-probabilistic rule, maximin, and denies the legitimacy of informative probabilistic judgments about social standing.[8] The mere introduction of IP forms of ignorance—which are in a sense probabilistic without necessarily being as precise as those Rawls objects to—does not settle the issue of the appropriate decision criterion with which to pair such representations, even if we retain a general commitment to the spirit of expected utility.

There are a number of candidates considered in the literature, each with a claim to generalize expected utility maximization. Since this is a matter of some controversy, and the subsequent analysis depends on it, let's briefly review the primary proposals. Let Y be a given set of options (for simplicity, I supress reference to a state space here). One of the more restrictive IP decision rules embodies some of the pessimism of maximin (e.g., Gilboa and Schmeidler, 1989). Γ-maximin restricts choice to the set

$$\left\{ x \in Y : \inf_{p \in \mathbb{P}} EU_p(x) \geq \inf_{p \in \mathbb{P}} EU_p(y) \text{ for all } y \in Y \right\},$$

where $EU_p(x)$ is the p-expected U-utility of x. (In the context of Lerner's theorem, we drop the social welfare function W from our notation and write $E_p(x)$ as in Equation 1.) Another of the more restrictive IP decision rules is optimistic where Γ-maximin is pessimistic. The Γ-maximax options of Y are given by the set

$$\left\{ x \in Y : \sup_{p \in \mathbb{P}} EU_p(x) \geq \sup_{p \in \mathbb{P}} EU_p(y) \text{ for all } y \in Y \right\}.$$

E-admissibility, a more liberal rule, was proposed by Isaac Levi (1974). E-admissible options maximize expected utility relative to *some* $p \in \mathbb{P}$. That is, the E-admissible options are those in the set

$$\left\{ x \in Y : \exists p \in \mathbb{P} \ \forall y \in Y \ EU_p(x) \geq EU_p(y) \right\}.$$

---

[7]Cf. Thistle (2007, fn. 9) on this topic. Thistle demonstrates how Lerner's theorem can be generalized in various other ways.

[8]The maximin principle Ralws favors differs from the one that is the subject of Theorem 2. Rawls's rule seeks to maximize the lot of the least well off individual in society. Put another way, social welfare, which we aim to maximize, is identified with the status of the worst off individual.

More liberal still, Maximality (Walley, 1991) enjoins us to choose options undominated in expectation, restricting choice to the set

$$\left\{ x \in Y : \forall y \in Y \ \exists p \in \mathbb{P} \ EU_p(x) \geq EU_p(y) \right\}.$$

Interval Dominance appeals to a weaker notion of avoiding dominated options. Admissible options according to this rule are those in the set

$$\left\{ x \in Y : \sup_{p \in \mathbb{P}} EU_p(x) \geq \inf_{p \in \mathbb{P}} EU_p(y) \text{ for all } y \in Y \right\}.$$

Γ-maximin, Γ-maximax, E-admissibility, Maximality, and Interval Dominance all generalize expected utility maximization. When $\mathbb{P} = \{p\}$, all rules amount to maximizing expected utility with respect to p. For other discussions of this rule set, see for example (Troffaes, 2007; Chandler, 2014; Huntley et al., 2014). Troffaes establishes that these decision rules stand in certain logical relations to each other generally, as depicted in Figure 1 (2007, Theorem 1). An arrow represents the fact that admissibility according to one rule implies admissibility according to the rule the arrow points to.

Figure 1: Relations among IP Decision Rules
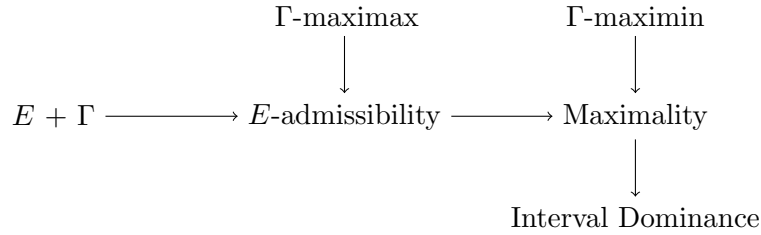
Since E-admissibility is a rather permissive decision rule—Maximality and Interval Dominance even more so—some have considered lexicographic choice procedures that would narrow the class of admissible options further. One candidate often considered in this context is E-admissibility + Γ-maximin (abbreviated $E + \Gamma$ in Figure 1) (Levi, 1986; Seidenfeld et al., 2012). This rule first restricts choice to E-admissible options, and then applies Γ-maximin to the set of E-admissible options as a tie-breaking procedure. What are the logical relations between this rule and those considered above? While the admissibility of an option according to E-admissibility + Γ-maximin clearly implies its E-admissibility, it is not true that it implies the option's admissibility according to Γ-maximin (e.g., Seidenfeld, 2004, Example 1). The following example demonstrates, among other things, that admissibility according to Γ-maximax does not imply admissibility according to E-admissibility + Γ-maximin.

Example 1. *Let $N = \{1,2\}$ and $y^* = \$100$. Let $U^1(x) = x$ and $U^2(x) = \sqrt{x}$. Both of these utility functions are concave (on their domains). Define $W(u^1, u^2) = u^1 + u^2$, where u^j is the utility of the person with the jth utility function. That W is increasing, symmetric, and concave is easily verified. For each j, let $\mathbb{P}_i^j = [0,1]$ for all $i \in N$. Consider the distribution $y = (99.75, 0.25)$. For some $p \in \mathbb{P}$, $E_p(y) \approx 10.24$, and for another $p' \in \mathbb{P}$, $E_{p'}(y) = 100.25$. For the equal distribution z, $E_p(z) \approx 57.07$ for all $p \in \mathbb{P}$.* △

Two comments on Example 1 are in order. First, distribution y guarantees that z is not a Γ-maximax option nor uniquely E-admissible.[9] The point about Γ-maximax suffices to establish the last claim of Observation 1 below. The point about the E-admissibility of y motivates the search for suitable assumptions that allow us to be more discriminating. Some such assumptions figure into Observation 2 below. Second, as mentioned just above, while y is a Γ-maximax option, it is not admissible according to E-admissibility + Γ-maximin since z will eliminate it at the second (Γ-maximin) stage. This helps us to further locate E-admissibility + Γ-maximin with respect to its logical relations to the other decision criteria.

We can now state our first observation.

Observation 1. *Given* (A.1), (A.2), (A.3), (A.6), *and* (A.7), *the equal distribution z is admissible according to*

1. *Γ-maximin*

2. *E-admissibility*

3. *E-admissibility + Γ-maximin*

4. *Maximality*

5. *Interval Dominance.*

However, z is not generally a Γ-maximax distribution.

Condition A.2 assumes that social welfare is concave. This is unnecessarily strong for some of the subclaims of Observation 1. For the claim about Γ-maximin, for instance, it suffices to assume that social welfare is *quasi-concave*.[10] But I would like to consider strengthening concavity or quasi-concavity to their *strict* versions for two reasons. First, such a strengthening has been thought to have important *ethical* ramifications for social welfare precisely because of its egalitarian nature. So it is already of theoretical interest here. Second, one might regard Observation 1 as less than a fully persuasive rational case for equal distribution. That z is admissible does not at all require its ultimate selection. Many other, decidedly non-egalitarian distributions may likewise be admissible. This is precisely the case in Example 1, where both the equal distribution and the lopsided $(99.75, 0.25)$ distribution are E-admissible. As a result, the ethical upshot of Observation 1 is at most that z is *permissible*, not that it is ethically *mandatory*. As we will see, the strengthening has further interesting consequences in this respect.

Let's begin with the ethical motivation for strict concavity/quasi-concavity. The following famous example is due to Diamond (1967), and was originally intended as a criticism of Harsanyi's utilitarianism.

---

[9]Even more skewed allocations are optimal for other sets of utility functions. For U^1 and $U^{2'}$ defined by $U^{2'}(x) = x/2$, for example, it is optimal to allocate everything to the individual with U^1.

[10]A function $f : \mathcal{S} \to \mathbb{R}$ defined on a convex subset of a real vector space is called *quasi-concave* if for all $x, y \in \mathcal{S}$ and all $\lambda \in [0, 1]$, $f(\lambda x + (1 - \lambda)y) \geq \min\{f(x), f(y)\}$. And f is *strictly quasi-concave* if for all $x \neq y$ and $\lambda \in (0, 1)$, $f(\lambda x + (1 - \lambda)y) > \min\{f(x), f(y)\}$.

Example 2. *Let $N = \{1, 2\}$ and suppose that we have to allocate a single unit of an indivisible good. Consider the following two policies. The first policy allocates the good to individual 1 for sure. The second policy is to randomize between the allocations $(1, 0)$ and $(0, 1)$, each with equal probability of being selected. Suppose that $U^i(1) = 1, U^i(0) = 0$. If the interests of individuals 1 and 2 are given equal weight, then society will be indifferent between the two policies if the group welfare function is linear in individual utilities.[11] In defense of social indifference, one might say that under either policy, a distribution results allocating the good to just one individual. In criticism of social indifference, one might say that the second policy gives individual 2 a "fair shake" while the first policy does not (1967, p. 766). The crucial distinction here is about concern for outcomes alone versus concern for procedure.* △

On the basis of the sort of fairness consideration displayed by the second reaction to Diamond's example—which might be viewed as a concern for equality of opportunity in one sense—Epstein and Segal construct a theory of social choice that relaxes expected utility at the social level and imposes a strict preference for randomization in scenarios like the one in Example 2 (Epstein and Segal, 1992). Their social preferences are strictly quasi-concave. (In particular, their theory entails that the $.50 - .50$ randomization is preferred to randomizing with any other probabilities in Example 2.)[12] We must note again that the force of Example 2 depends on both measurability and (strong) comparability assumptions about utilities.

In our setting, concavity and quasi-concavity reflect a social aversion to inequality in the distribution of utilities. Their strict versions reflect a stronger form of such aversion and rule out more options as inadmissible for social choice, as the following observation attests. However, most of the IP rules under consideration fail to secure the unique admissibility of the equal distribution in general—that is, for many sets of utility functions about which a social planner could be uncertain, z will not be uniquely admissible according to most of the rules under consideration.

Observation 2. *Given (A.1), (A.3), (A.6), (A.7) and modifying (A.2) so that social welfare is **strictly** quasi-concave (resp. strictly concave), the equal distribution z is **uniquely** admissible according to Γ-maximin (resp. E-admissibility + Γ-maximin). However, z generally fails to be uniquely admissible according to Γ-maximax, E-admissibility, Maximality, and Interval Dominance.*

We see clearly here how, in general, the case for Lerner's egalitarianism depends delicately on the IP generalization of expected utility that we adopt. One strategy, not pursued here, is to find ways to discriminate among the various IP decision criteria in terms of normative attractiveness. Seidenfeld (2004) and Troffaes (2007), for example, both discuss this possibility, each finding an advantage for E-admissibility. Levi (1980) defends the E-admissibility of an option as a necessary condition for its rational selection, but thinks further tie-breaking crite-

---

[11]Let δ_i, $i = 1, 2$ be the (degenerate) lotteries allocating the good to 1 and 2, respectively, with certainty. Von Neumann and Morgenstern's independence axiom implies that $\delta_1 = \lambda\delta_1 + (1 - \lambda)\delta_1 \sim \lambda\delta_2 + (1 - \lambda)\delta_1$ for $\lambda \in [0, 1]$, since $\delta_1 \sim \delta_2$.

[12]As Sen points out in his debate with Harsanyi, to take Example 2 as motivation for violating independence in *every* case is hasty (1977, pp. 297–298). At most, it motivates thinking that independence should sometimes be violated, not that it should always be. Imposing *strict* concavity on social welfare, then, requires further justification. I won't pursue that issue here. I introduce the strict versions for the sake of the argument: some people have advocated them and they allow us to establish Lerner's conclusion for certain decision rules.

ria like a lexicographic application of Γ-maximin as in the rule E-admissibility $+$ Γ-maximin are matters of individual discretion.

5 Discussion

Many ethical and political theories make crucial appeals to epistemic states of uncertainty and ignorance. Appropriate representation of such states, then, is an important issue in these contexts. This essay is a case study in applying the theory of imprecise probabilities in the general area of veil-of-ignorance-type arguments. Taking for granted the philosophical premise of the relevance of ignorance for evaluating issues impartially or fairly, we can perform a conceptual robustness check on Lerner's argument for the egalitarian income distribution by considering other, more plausible representations of "complete ignorance." There are two broad types of reactions one might have to the foregoing analysis.

First, one might take the present study as good news for egalitarianism. Observation 1 shows that Lerner's conclusion regarding the optimality of the egalitarian income distribution is, in a particular sense, robust to variation between certain important representations of ignorance. We need not assume "the worrisome hypothesis of equal probability" to reach the admissibility of the equal distribution of income, but neither need we move to totally non-probabilistic decision rules and representations of ignorance either as Theorem 2 does. Even the extremely weak A.7 suffices for the admissibility of the equal distribution for most rules. However, for some IP choice rules, admissibility is quite a weak notion. In general, it is much weaker than optimality according to a weak order, for example. It is worth noting, as Observation 2 does, that the *unique* admissibility of the equal distribution can be secured without moving away from the assumption of complete ignorance. This conclusion now depends both on ethical assumptions about social welfare and on the IP decision rule that we adopt. But perhaps the ethical assumptions and decision rules that secure uniqueness can be motivated.[13]

Second, one might take the present study as bad news for Lerner's argument. While Observation 1 shows that the equal distribution is admissible, it is not uniquely admissible in general. In Example 1, for instance, *severely* skewed income distributions are also admissible according to several of the generalizations of expected utility under consideration, and for other sets of utility functions, completely skewed distributions will be admissible. Pre-theoretic fairness-based considerations to reject a completely skewed income distribution may militate against an analysis of fairness in terms of admissibility under ignorance. That is, such observations may well incline some to reject the general decision-making under IP-ignorance approach to evaluating issues of distributive justice—or even the more general philosophical premise mentioned at the beginning of this section that gives ignorance a distinguished role

---

[13]Such justifications for IP decision rules have important consequences for ethics more generally, outside of veil of ignorance settings. For example, the problem of *cluelessness* for consequentialist moral theories is essentially that we are usually very uncertain about the knock on effects and longterm consequences of our actions, making the moral assessment of actions much more difficult than is often appreciated. Mogensen (2020) investigates the issue from the perspective of IP, arguing that his favored rule—Maximality—is unable to support the positions of longtermists and effective altruists unless the depth of our uncertainty is downplayed. As explained above, Maximality is a particularly permissive rule, and, as Observation 2 shows, less permissive rules can lead to decidedly less equivocal conclusions—perhaps even to more ethically attractive ones—even in the face of complete ignorance.

in evaluating issues of impartiality and fairness. And while Observation 2 demonstrates that the unique admissibility of the equal distribution can be secured under certain assumptions about social welfare for certain IP decision rules, for most of the rules under consideration, the equal distribution is not uniquely admissible even relative to the stronger egalitarian assumptions about social welfare made in the observation. The vindication of Lerner's argument, on such views, awaits vindication of some set of assumptions that imply that the equal distribution is uniquely admissible.

This second sort of reaction, in objecting to the admissibility of completely and severely skewed distributions, makes appeal to a mild procedure-independent notion of fairness. By contrast, *pure* procedural justice makes no such appeal, defining any distribution resulting from a fair procedure as just—like any holdings resulting from a fair gamble. Still other notions of procedural justice do appeal to independent standards of fairness (e.g., Rawls, 1971, §14). If the IP statement of Lerner's problem is the more appropriate or compelling representation of choice under complete ignorance, one consequence of the critique voiced in the second reaction would seem to be either that unfair distributions can result from this type of fair procedure—compare well-conducted legal trials that reach a wrong verdict—or that choice under ignorance is not a fair procedure.[14]

Appendix

Since W is symmetric, we can set $F(\tilde{y}) = W(U^1(\tilde{y}^1), ..., U^n(\tilde{y}^n))$ for any given permutation $\tilde{y}$ of y. Properties of individual and social welfare, then, imply certain properties of the reduced function F. For example, in the proof of Observation 2 below, we use the fact that, when W is assumed to be increasing and strictly quasi-concave and the U^j are concave, F is itself strictly quasi-concave. I sketch the argument for this particular claim here because the reasoning can be extended to analogous propositions appealed to below.

So suppose that W is increasing and strictly quasi-concave and that individual utility is concave. The first inequality below uses the facts that the U^j are concave and W is increasing, the second, the fact that W is strictly quasi-concave.

$$
\begin{aligned}
F(\lambda(\tilde{y}^1, ..., \tilde{y}^n) + (1 - \lambda)(\tilde{x}^1, ..., \tilde{x}^n)) &= F(\lambda\tilde{y}^1 + (1 - \lambda)\tilde{x}^1, ..., \lambda\tilde{y}^n + (1 - \lambda)\tilde{x}^n) \\
&= W(U^1(\lambda\tilde{y}^1 + (1 - \lambda)\tilde{x}^1), ..., U^n(\lambda\tilde{y}^n + (1 - \lambda)\tilde{x}^n)) \\
&\geq W(\lambda U^1(\tilde{y}^1) + (1 - \lambda)U^1(\tilde{x}^1), ..., \lambda U^1(\tilde{y}^n) + (1 - \lambda)U^1(\tilde{x}^n)) \\
&= W(\lambda(U^1(\tilde{y}^1), ..., U^n(\tilde{y}^n)) + (1 - \lambda)(U^1(\tilde{x}^1), ..., U^n(\tilde{x}^n))) \\
&> \min\{W(U^1(\tilde{y}^1), ..., U^n(\tilde{y}^n)), W(U^1(\tilde{x}^1), ..., U^n(\tilde{x}^n))\} \\
&= \min\{F(\tilde{y}^1, ..., \tilde{y}^n), F(\tilde{x}^1, ..., \tilde{x}^n)\}
\end{aligned}
$$

This establishes that F is strictly quasi-concave.

Proof of Observation 1

Proof. We prove the items in order.

1. The equal distribution z is an average of $\tilde{y}(k)$, $k = 1, ..., n!$. By the quasi-concavity of social welfare (implied by the stronger A.2),

$$F(z) \geq \min_k F(\tilde{y}(k)). \tag{3}$$

Since z is invariant under permutations, we obtain

$$\min_k F(\tilde{z}(k)) \geq \min_k F(\tilde{y}(k)). \tag{4}$$

By A.7, for any y, there is a $\underline{p} \in \mathbb{P}$ such that $\underline{p}(\arg\min F(\tilde{y}(k))) = 1$.[15] Hence, $E_{\underline{p}}(y) = \min_k F(\tilde{y}(k))$. And for any other $p \in \mathbb{P}$, $E_p(y)$ is a convex combination with terms at least as great as $\min_k F(\tilde{y}(k))$. So,

$$\inf_{p \in \mathbb{P}} E_p(z) \geq \inf_{p \in \mathbb{P}} E_p(y). \tag{5}$$

It follows that z is a Γ-maximin solution.[16]

2. That z is E-admissible follows from Theorem 1. In particular, A.7 requires that an equi-probability distribution as dictated by A.4 is in $\mathbb{P}$, guaranteeing the E-admissibility of z.

3. Immediate from the preceding two claims (Observation 1.1 and 1.2).

4. Any distribution that is E-admissible is of course undominated in expectation (Figure 1). Another way to see the Maximality claim in our setting is that, by A.7, for every distribution y, there is a $p \in \mathbb{P}$ such that $p(\arg\min F(\tilde{y}(k)) = 1$. Then, $E_p(z) = F(z) \geq E_p(y) = \min_k F(\tilde{y}(k))$ by (3). So, z is admissible according to Maximality.

5. Again, this follows immediately from the foregoing and Figure 1, but we can show the claim more directly. Since $\sup_{p \in \mathbb{P}} E_p(z) \geq \inf_{p \in \mathbb{P}} E_p(z)$, it follows immediately from (5) that, for all y,
$$\sup_{p \in \mathbb{P}} E_p(z) \geq \inf_{p \in \mathbb{P}} E_p(y).$$

Finally, Example 1 suffices to establish that z is not generally a Γ-maximax solution.

$\square$

---

[15]Cf. McManus et al. (1972, p. 495, fn. 8).

[16]A general relationship between Γ-maximin and maximin options obtains under the assumption of complete ignorance in the IP sense (e.g., Berger, 1985, p. 216).

13

Proof of Observation 2

Proof. The assumptions that individual utilities are concave (A.3) and that W is increasing and strictly quasi-concave (modified A.2) imply that F is a strictly quasi-concave function. This, in turn, allows us to promote the inequality in (3) to a strict one for any $y \neq z$.

$$F(z) > \min_k F(\tilde{y}(k))$$

Accordingly, (4) can be strengthened to

$$\min_k F(\tilde{z}(k)) > \min_k F(\tilde{y}(k)),$$

and (5) to

$$\inf_{p \in \mathbb{P}} E_p(z) > \inf_{p \in \mathbb{P}} E_p(y).$$

This establishes the unique admissibility of z according to Γ-maximin.

Assuming now that social welfare is strictly *concave* (which implies that it is strictly quasi-concave), we have that z is E-admissible by Observation 1.2. The unique admissibility of z according to E-admissibility $+$ Γ-maximin follows immediately from this fact and z's unique admissibility according to Γ-maximin, established just above.

For the claim about the rules for which z is *not* generally uniquely admissible, by Figure 1, it suffices to show that, under the stated assumptions, z is not uniquely admissible according to Γ-maximax for some set of utility functions consistent with the assumptions. From this, it follows that z is not uniquely admissible according to the other rules downstream in the figure. To that end, consider again Example 1. Recall $n = 2$, $U^1(x) = x$ and $U^2(x) = \sqrt{x}$. Both U^1 and U^2 are concave. But now suppose that $W(u^1, u^2) = \sqrt{u^1 + u^2}$. As defined, W is an increasing (for these utility functions—we could define W piecewise to preserve its being increasing for negative numbers), symmetric, and *strictly* concave function of individual welfare levels. To see this, note that addition is an increasing and concave function of individual welfare levels and that the square root operation is an increasing and *strictly* concave function. By an argument analogous to the one at the beginning of the Appendix, the composition of those functions, W, is then an increasing and strictly concave function of individual welfare levels. Moreover, W is invariant under permutations of its arguments and thus symmetric. Now, consider again the distribution $y = (99.75, 0.25)$. Since $\mathbb{P}^1_1 = [0, 1]$, $E_p(y) = \sqrt{99.75 + 0.25^{1/2}} \approx 10.02$ for some $p \in \mathbb{P}$. However, for all $p \in \mathbb{P}$, $E_p(z) = \sqrt{50 + 50^{1/2}} \approx 7.55$. Hence, it is not the case that $\sup_{p \in \mathbb{P}} E_p(z) \geq \sup_{p \in \mathbb{P}} E_p(x)$ for all possible distributions x. In other words, z is not admissible according to Γ-maximax— but (99.75, 0.25) is—and, consequently, not uniquely admissible according to E-admissibility, Maximality, or Interval Dominance. $\qquad\square$

References

Berger, J. O. (1985). *Statistical Decision Theory and Bayesian Analysis* (2 ed.). New York: Springer-Verlag.

Bradley, S. (2019). Imprecise Probabilities. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2019 ed.). Metaphysics Research Lab, Stanford University.

Buchak, L. (2017). Taking risks behind the veil of ignorance. *Ethics 127*(3), 610–644.

Chandler, J. (2014). Subjective probabilities need not be sharp. *Erkenntnis 79*(6), 1273–1286.

Diamond, P. A. (1967). Cardinal welfare, individualistic ethics, and interpersonal comparison of utility: Comment. *The Journal of Political Economy 75*(5), 765–766.

Epstein, L. G. and U. Segal (1992). Quadratic social welfare functions. *Journal of Political Economy 100*(4), 691–712.

Freeman, S. (2019). Original Position. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2019 ed.). Metaphysics Research Lab, Stanford University.

Friedman, M. (1947). Lerner on the economics of control. *Journal of Political Economy 55*(5), 405–416.

Gajdos, T. and F. Kandil (2008). The ignorant observer. *Social Choice and Welfare 31*(2), 193–232.

Gilboa, I. and D. Schmeidler (1989). Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics 18*(2), 141–153.

Good, I. J. (1952). Rational decisions. *Journal of the Royal Statistical Society. Series B (Methodological)*, 107–114.

Graaff, J. d. V. (1967, originally published 1957). *Theoretical Welfare Economics*, Volume 446. Cambridge University Press.

Gustafsson, J. E. (2018). The difference principle would not be chosen behind the veil of ignorance. *The Journal of Philosophy 115*(11), 588–604.

Hájek, A. and M. Smithson (2012). Rationality and indeterminate probabilities. *Synthese 187*(1), 33–48.

Harsanyi, J. C. (1953). Cardinal utility in welfare economics and in the theory of risk-taking. *Journal of Political Economy 61*(5), 434–435.

Harsanyi, J. C. (1975). Can the maximin principle serve as a basis for morality? a critique of John Rawls's theory. *American Political Science Review 69*(2), 594–606.

Hill, B. (2019). Confidence in beliefs and rational decision making. *Economics & Philosophy 35*(2), 223–258.

Huntley, N., R. Hable, and M. C. Troffaes (2014). Decision making. In *Introduction to Imprecise Probabilities*, Wiley Series in Probability and Statistics, Augustin, thomas and coolen, frank p.a., and de cooman, gert and troffaes, matthias c.m. 8, pp. 190–206. Wiley.

Joyce, J. M. (2011). The development of subjective bayesianism. In D. M. Gabbay and J. H. Woods (Eds.), *Inductive Logic*, Volume 10, pp. 415–475.

Kaplan, M. (1996). *Decision Theory as Philosophy*. Cambridge University Press.

Keynes, J. M. (1921). *A treatise on probability*. Courier Dover Publications.

Kurtulmuş, A. F. (2012). Uncertainty behind the veil of ignorance. *Utilitas 24*(1), 41–62.

Lerner, A. P. (1946). *The Economics of Control*. New York: The MacMillan Company.

Levi, I. (1974). On indeterminate probabilities. *The Journal of Philosophy 71*(13), 391–418.

Levi, I. (1977). Four types of ignorance. *Social Research 44*(4), 745–756.

Levi, I. (1980). *The Enterprise of Knowledge*. MIT Press, Cambridge, MA.

Levi, I. (1986). The paradoxes of Allais and Ellsberg. *Economics and Philosophy 2*(1), 23–53.

Liang, C.-Y. (2017). Optimal inequality behind the veil of ignorance. *Theory and Decision 83*(3), 431–455.

Little, I. M. D. (2002, originally published in 1950). *A Critique of Welfare Economics*. Oxford University Press.

Luce, R. D. and H. Raiffa (1957). *Games and decisions: Introduction and critical survey*. Courier Dover Publications.

McManus, M., G. M. Walton, and R. B. Coffman (1972). Distributional equality and aggregate utility: Further comment. *The American Economic Review 62*(3), 489–496.

Mogensen, A. L. (2020). Maximal cluelessness. *The Philosophical Quarterly 71*(1), 141–162.

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

Samuelson, P. A. (1964). AP Lerner at sixty. *The Review of Economic Studies 31*(3), 169–178.

Seidenfeld, T. (2004). A contrast between two decision rules for use with (convex) sets of probabilities: Γ-maximin versus E-admissibilty. *Synthese 140*(1-2), 69–88.

Seidenfeld, T., M. J. Schervish, and J. B. Kadane (2012). Forecasting with imprecise probabilities. *International Journal of Approximate Reasoning 53*(8), 1248–1261.

Sen, A. (1969). Planner's preferences: Optimality, distribution and social welfare. In J. Margolis and H. Guitton (Eds.), *Public Economics*. London: MacMillan.

Sen, A. (1973). On ignorance and equal distribution. *The American Economic Review 63*(5), 1022–1024.

Sen, A. (1977). Non-linear social welfare functions: A reply to Professor Harsanyi. In R. E. Butts and J. Hintikka (Eds.), *Foundational Problems in the Special Sciences*, pp. 297–302. Springer.

Stefánsson, H. O. (2019). Ambiguity aversion behind the veil of ignorance. *Synthese doi.org/10.1007/s11229-019-02455-8*, 1–24.

Stewart, R. T. and I. Ojea Quintana (2018). Probabilistic opinion pooling with imprecise probabilities. *Journal of Philosophical Logic 47*(1), 17–45.

Thistle, P. D. (2007). Generalized probabilistic egalitarianism. In *Equity*, pp. 7–32. Emerald Group Publishing Limited.

Troffaes, M. C. (2007). Decision making under uncertainty using imprecise probabilities. *International Journal of Approximate Reasoning 45*(1), 17–29.

Vallinder, A. (2018). Imprecise Bayesianism and global belief inertia. *The British Journal for the Philosophy of Science 69*(4), 1205–1230.

Walley, P. (1991). *Statistical reasoning with imprecise probabilities*. Chapman and Hall London.

Wasserman, L. (1993). Review: Statistical reasoning with imprecise probabilities by Peter Walley. *Journal of the American Statistical Association 88*(422), pp. 700–702.

Weatherson, B. (2015). For Bayesians, rational modesty requires imprecision. *Ergo, an Open Access Journal of Philosophy 2*.