CrossMark

Optimal inequality behind the veil of ignorance

Che-Yuan Liang[1] (ORCID)

Abstract In Rawls' (A theory of justice. Harvard University Press, Cambridge, 1971) influential social contract approach to distributive justice, the fair income distribution is the one that an individual would choose behind a veil of ignorance. Harsanyi (J Polit Econ 61:434–435, 1953, J Polit Econ 63:309–332, 1955, Am Polit Sci Rev 69:594–606, 1975) treated this situation as a decision under risk and arrived at utilitarianism using expected utility theory. This paper investigates the implications of applying cumulative prospect theory instead, which better describes behavior under risk. I find that the specific type of inequality in bottom-heavy right-skewed income distributions, which includes the log-normal income distribution, could be perceived as desirable. This optimal inequality result contrasts the implications of other social welfare criteria.

Keywords Veil of ignorance · Prospect theory · Social welfare function · Income inequality

1 Introduction

How to distribute income fairly is a question that has been discussed across different disciplines of social science and philosophy. Harsanyi (1953, 1955, 1975) and Rawls

✉ Che-Yuan Liang
che-yuan.liang@nek.uu.se

1   Department of Economics, Institute for Housing and Urban Research, Uppsala Center for Fiscal Studies, Uppsala University, P.O. Box 514, 75120 Uppsala, Sweden

(1971) offered two of the most influential theories of distributive justice, both using the popular social contract approach. A central idea is that the normative question can be transformed to the *descriptive* question of what income distribution an individual would choose in a hypothetical original position before knowing her identity in the society.[1] Behind such a veil of ignorance, the decision maker becomes an impartial observer, internalizing the interests of *each* member of the society appropriately, as she has to account for potentially becoming each of them. Under the impartiality constraint which seems to be a prerequisite for justice, social preferences could be derived from individual preferences. Whereas Rawls' argued that the impartial observer would choose distributions according to the maximin principle, Harsanyi favored utilitarianism.

There are, however, many difficulties with deriving a theory of justice using the original position. Some issues concern what assumptions to make about the impartial observer's preferences (e.g., for income and risk) and the exact nature of the decision problem (e.g., whether probabilities of possible income levels are known).[2] Motivating the exact framing of the original position seems to be a normative task in itself that could be as difficult as the question of fair income distribution. There are also more general methodological concerns such as whether the social contract approach appropriately captures impartiality and whether justice requires more than impartiality.[3]

While these are interesting issues that have received a lot of attention, there is also a literature that stays agnostic about them and that focuses on answering the descriptive question about the desired outcome behind the veil of ignorance by asking individuals in surveys (e.g, Frohlich et al. 1987; Bosmans and Schokkaert 2004; Herne and Suojanen 2004; Johansson-Stenman et al. 2002; Traub et al. 2005; Amiel et al. 2009). The conclusions of these studies are not clear-cut and depend crucially on the framing of the original position.[4] A general pattern is that both the maximin and utilitarian social welfare functions perform poorly in explaining the survey responses.

A parsimonious way to characterize the original position is to think of income distributions as lotteries of birth because the impartial observer *randomly* becomes somebody in her chosen distribution.[5] Harsanyi embraced the lottery interpretation of the original position and used von Neumann and Morgenstern's (1944) theory of

---

[1] The original position therefore seems to violate Hume's law (1739) asserting that normative conclusions cannot follow from descriptive premises.

[2] Both Rawls and Harsanyi argued that some restrictions *should* be made on the impartial observer and the decision problem, although they did not agree about the restrictions.

[3] For instance, Moreno-Ternero and Roemer (2008) argued that justice requires prioritarianism and they show that the veil of ignorance is not compatible with prioritarianism when different individuals have different preferences.

[4] Among other things, the conclusions depend on factors such as whether the scenario is about risk or uncertainty, whether the respondents should consider themselves to be external observers or involved in the realized distributions, the thickness of the veil, individual background characteristics, and the exact rules of negotiation in the case of groups agreeing on a principle of justice.

[5] Randomness is sometimes perceived to be *unfair ex post* because it is beyond individuals' control. However, because the impartial observer chooses and accepts the randomness *ex ante*, it makes the chosen income distribution impartial and arguably *fair*.

decision under risk applying expected utility theory to the problem.[6] Since Harsanyi's seminal work, there is plenty of new empirical evidence that expected utility theory provides a poor description of individual behavior under risk.[7] To cope with the deficiencies of expected utility theory, Kahneman and Tversky (1979) developed prospect theory and later modified it to cumulative prospect theory (Tversky and Kahneman 1992). There are, by now, many empirical studies in support of these theories.[8] In this paper, I explore the consequences of applying the cumulative version of prospect theory[9] *given* the lottery interpretation of the original position. The result represents an impartially perceived social preference for income distributions based on realistic and empirically verified individual human preferences.[10]

The exercise corresponds to asking individuals about their preferences in a rudimental original position without imposing normative concerns, but without involving the complications of actually surveying individuals. Realistic prospect-theory type of human risk preferences have been used before for individual welfare evaluations by Günther and Maier (2008) and Jänttietal et al (2014). While they evaluated the effects of real income *changes*, I evaluate static hypothetical distributions. Furthermore, the original position intends to connect individual choice to *social* welfare.

I study the problem of distributing a fixed amount of income in a population once. The individual preference for risk-free income is assumed to have the same functional form across individuals.[11] It is a decision under risk because the frequencies of different income levels are known. Production and efficiency concerns are ignored. I start out by investigating the simplest two-level income distribution for analytical tractability and to pin down the intuition before moving on to multi-level income distributions using numerical methods. Unlike original prospect theory, cumulative prospect theory that I use can handle such distributions in an uncontroversial way.[12] Furthermore, nobody has successfully developed such a foundation for original prospect theory which also violates stochastic dominance. It may therefore be argued that a rational individual should not follow original prospect theory. These deficiencies are solved with cumulative prospect theory. When referring to prospect theory, I will typically refer to its cumulative version.

---

[6] To be precise, Harsanyi used von Neuman–Morgernstern individual preferences and the principle of insufficient reasons to weight the outcomes, which results in the expected utility formula.

[7] See, e.g., Kahneman and Tversky (1979), Fishburn and Kochenberger (1979), Hershey and Schoemaker (1980), Payne et al. (1981), Wehrung (1989), Tversky and Kahneman (1992), Camerer and Ho (1994), and Wu and Gonzalez (1996).

[8] See Kahneman and Tversky (2000) and the references in footnote 7.

[9] Empirically, some data fits original prospect theory better (e.g., Camerer and Ho 1994), while other data fits cumulative prospect theory better (e.g., Fennema and Wakker 1997).

[10] This contrasts Rawls impartial observer who is stripped of most human preferences, e.g., risk preferences, as Rawls argued that the impartial observer should disregard the frequencies of different income levels.

[11] A major modern discussion on the veil of ignorance concerns how to deal with heterogeneous preferences for risk-free income (see, e.g., Karni 1998; Mongin 2001). The simple setting here abstracts from issues of interpersonal comparability of preferences.

[12] Wakker and Tversky (1993) provided a straightforward extension of cumulative prospect theory for continuous distributions. Rieger and Wang (2008) suggested a way to extend original prospect theory to handle multi-outcome and continuous distributions that Kothiyal et al. (2011) find "unsatisfactory".

In my simple setting, social welfare functions in the previous literature all favor complete equality. These criteria are typically based on diminishing marginal utility, direct positional concerns for the lower end of income distributions, or are directly inversely related to income inequality.[13]

As shown by Wakker and Tversky (1993), the preference foundation of prospect theory replaces the independence assumption of expected utility theory with reference and positional dependence (often called sign and rank dependence). Prospect theory reference dependence has three distinguishing features: First, income carries utility relative to a reference level. Second, not only gains but also losses exhibit decreasing marginal sensitivity. Third, losses carry more disutility than gains carry utility. Prospect theory positional dependence is implemented through non-linear probability weights where probabilities of the relatively largest gains and losses are overweighted compared to probabilities of other gains and losses.

While reference and positional dependence are realistic psychological features of human decision utility, it is unclear whether such relative comparisons should enter welfare evaluations. At an individual level, this type of behavioral patterns may sometimes seem irrational if the reference point can be manipulated or if it is unstable over time. Nevertheless, there is evidence that experienced well-being also depend on relative comparisons (Kahneman et al. 1997).[14] Furthermore, the happiness research literature provides evidence that income relative to some measure of average income in a country matters for expressed well-being (see, e.g., Easterlin 1974; Frey and Stutzer 2002). Excluding these elements of satisfaction from individual welfare seems highly paternalistic. The individual preference foundation translates into a social preference foundation in the original position, where individual relative comparisons translate into social relative comparisons.

An issue in the original position concerns how to choose the reference income. I first explore the effects of using the mean income as the reference income. This is the income level all individuals would have in the even income distribution with complete equality which is the distribution favored by other approaches to justice such as egalitarianism and prioritarianism. The mean-income reference setting also corresponds to a thought experiment where individuals in a complete equality world evaluate the attractiveness of lottery-based income redistribution. Another interpretation is that complete equality is the default to be implemented if the impartial observer does not choose another distribution.

As an alternative, I also develop and use "representative aggregation of reference incomes" which takes the mean of the social welfare *evaluations* of each of the individuals in the realized income distribution using the realized income of each individual as her reference income. With this method, we do not need to ask individuals to assume a hypothetical counterfactual reference income and we come even closer to captur-

---

[13] They include, besides the utilitarian and maximin social welfare functions, e.g., the Cobb–Douglas social welfare function, the quadratic social welfare function (Epstein and Segal 1992), Atkinson's social welfare function (Atkinson 1970), Gini, entropy, and Boulding's principle (Boulding 1962). Of course, when production is introduced into the problem, inequality may be tolerated because there is usually an efficiency-equity trade-off.

[14] Bernheim and Rangel (2009) proposed an approach to connect individual choice based on decision utility, experienced utility, or remembered utility with individual welfare.

ing realistic individual preferences. This setting corresponds to a thought experiment where we have individuals in realized distributions evaluate the attractiveness of redistributing income in lotteries where everybody has the same probability of switching income position with everybody else. The best society is the one with the best impartial self-evaluations.

A prospect theory impartial observer has two reasons to prefer an uneven income distribution. First, incurring small losses with a high probability to afford large gains with a low probability could be attractive because the large gains are overweighted. Second, incurring large losses with a low probability to afford small gains with a high probability could be attractive because the large losses have low marginal disutility. In a two-level income world, this leads to two possible types of optimal uneven income distributions.[15] The first type is a bottom-heavy right-skewed superstar distribution where few individuals have very high income and many individuals have low income. The second type is a top-heavy left-skewed scapegoat distribution where few individuals have very low income and many individuals have high income.

Whether inequality is perceived as desirable depends on the exact parameterization of prospect theory. I show that the superstar distribution is optimal under some assumptions. Furthermore, the superstar type of inequality is more desirable than complete equality when using a reasonably chosen prospect theory parameterization for two-level income distributions and log-normal income distributions which many countries have (Gibrat 1931; Aitchison and Brown 1957; Battistin et al. 2007). The intuition is that these income distributions resemble fair odds lotteries that people do buy. Such distributions contain the American dream with an ex ante opportunity to become a superstar creating a strong psychological possibility effect.

From a normative standpoint, the evaluation of optimal inequality depends on whether we should give any weight to (dis)satisfaction derived from social comparisons between individuals in a society. It also depends on whether the type of procedural justice in the original position appropriately embodies impartiality and whether impartiality has intrinsic value. In this regard, it may be argued that optimal inequality is an unpleasant and unacceptable result that indicates that the original position needs to be modified or rejected.

Our results are, however, also descriptively interesting. They imply that if complete equality is imposed at a certain point in time (let us say, for fairness reasons), such a distribution would not prevail over time if we allow individuals to redistribute income by participating in lotteries. Instead, they would voluntarily and jointly opt for the optimal-inequality type of distribution.

The next section presents the model used. Section 3 presents the income distributions investigated. Section 4 reports some analytical results. Section 5 reports some numerical results. The final section concludes and further discusses the implications of the results.

---

[15] To avoid overcomplicating phrasing, I sometimes write that an income distribution is optimal, desirable, or social-welfare maximizing when meaning that it is preferred by the impartial observer.

2 Model

The problem at hand concerns how to evaluate different income distributions once. It is a purely static problem and income can be thought of as life-time income, resources, endowment, wealth, or consumption goods. Assume that each risk-free income level x carries a (decision) utility for individuals according to the (basic) utility function $u(x)$. Individuals are identical and they all have the same utility function. By normalizing the population to one, the frequencies in the income distribution can be interpreted as population shares $p(x)$.

The original position transforms society's choice of the optimal income distribution into an individual impartial observer's choice of the optimal lottery interpreting the frequencies described by the function $p(.)$ as probabilities of different lottery outcomes. The lottery interpretation is attractive because it forces the social welfare evaluation to account for the outcome of *each* individual in the income distribution, in the same manner as an individual's preference evaluation of a lottery where she accounts for each of the different lottery outcomes she could end up with.

Choice patterns for lottery distributions have been extensively studied theoretically and empirically before. It is, therefore, possible to apply a calibrated model of decision under risk that relies on the insights of this literature to investigate the problem without the need to ask individuals about their hypothetical preferences in the original position. This circumvents the issues of how to appropriately frame the original position to remove normative elements and to obtain truthful answers of behavior in a hypothetical scenario.

I apply a general model for evaluation of income and lottery distributions that encompasses expected utility theory (Von Neumann and Morgenstern 1944) and cumulative prospect theory. I work with the cumulative version of prospect theory (Tversky and Kahneman 1992) because it can handle multi-level income distributions.[16] Furthermore, it is generally believed to be theoretically sounder because it has a preference foundation (Wakker and Tversky 1993) and does not violate stochastic dominance.

The impartial observer attaches weights $w(x; p(.))$ to each income level in a way that may depend on the entire income distribution. We are interested in evaluating income distributions with a fixed total and mean income x_{m}. Any income distribution can be obtained by starting out from an even income distribution where everyone has income x_{m} and then transferring income from some individuals to others. Any uneven income distribution corresponds to a mean-preserving spread of the even income distribution.

The optimal income distribution is then the distribution that maximizes the weighted average of the utility attached to each income level according to:

$$\max_{p(.)} U(p(.)) = \sum_{x} w(x; p(.)) u(x)$$

---

[16] For two-outcome prospects with both a gain and a loss, cumulative prospect theory collapses to original prospect theory. Cumulative prospect theory is typically used in empirical applications involving multi-outcome prospects (e.g., Barberis and Huang 2008; Barberis and Xiong 2009 provide applications on stock market returns).

$$\text{s.t.} \sum_x p(x) = 1 \text{ and } \sum_x p(x)\, x = x_{\mathrm{m}}. \tag{1}$$

From the society's perspective, the objective function U represents the social welfare of an income distribution. From the impartial observer's perspective, U represents the perceived decision utility of a lottery distribution.

The criterion in Eq. (1) can be further specified by choosing functional forms for $u(.)$ and $w(.)$. For an expected utility impartial observer, the utility function is concave ($u''(x) < 0$), which reflects risk aversion. Furthermore, the probability weight is linear in the probability ($w = p(x)$). Such a decision utility leads to the utilitarian social welfare function. These features are implied by the independence axiom where income and associated probabilities add welfare without depending on the rest of the income distribution.

For a prospect theory impartial observer, the utility function depends on the reference income x_0. It is concave for gains ($u''(x > x_0) < 0$) and convex for losses ($u''(x < x_0) > 0$). Furthermore, it exhibits loss aversion ($u'(x_0 + a) < u'(x_0 - a), a > 0$). Reference dependence reflects the fact that people evaluate income relative to an anchoring point. Diminishing sensitivity reflects the fact that accumulated losses are perceived as better than many small losses.

The probability weights are position dependent and depend on the entire income distribution. It is typically defined trough the concept of capacities (Choquet 1955) which is a cumulative weighting function $W(p)$. Following Tversky and Kahneman (1992), we can express the probability weights according to:

$$w(x; p(.)) = \begin{cases} W(p(x_i \geq x)) - W(p(x_i > x)) & x \geq x_0 \\ W(p(x_i \leq x)) - W(p(x_i < x)) & x < x_0, \end{cases} \tag{2}$$

where subindex i indexes income levels. The probability weighting function contains two pieces and may be discontinuous at the reference income. The *cumulative* weighting function fulfills subcertainty ($W(p) + W(1-p) < 1$) and subproportionality $\left(\frac{1 - W((1-p)q)}{1 - W(1-p)} < \frac{1 - W((1-p)qr)}{1 - W((1-p)r)}, p, q \in (0,1) \right)$. These properties result in the overweighting of probabilities of the relatively largest gains and losses and the underweighting of probabilities of other gains and losses. Because large gains and losses usually occur with low probabilities in applications, this is often interpreted as the overweighting of low probabilities and the underweighting of high probabilities.[17]

Prospect theory weights embody an indirect positional dependence. Features of the income distribution contribute indirectly to social welfare through the welfare contribution of each income level. This dependence reflects that probabilities of income levels tend to be categorized as impossible, possible, probable, and certain.

---

[17] In most applications, such as the ones here, overweighting can be expressed in terms of the income level (Wakker and Tversky 1993) or in terms of the probability of the income level (Tversky and Kahneman 1992). In cases where the two descriptions differ, the former description is more accurate for cumulative prospect theory. Original prospect theory, which uses the latter description, faces issues with potential violation of continuity and invariance to coordination system change as discussed by Rieger and Wang (2008).
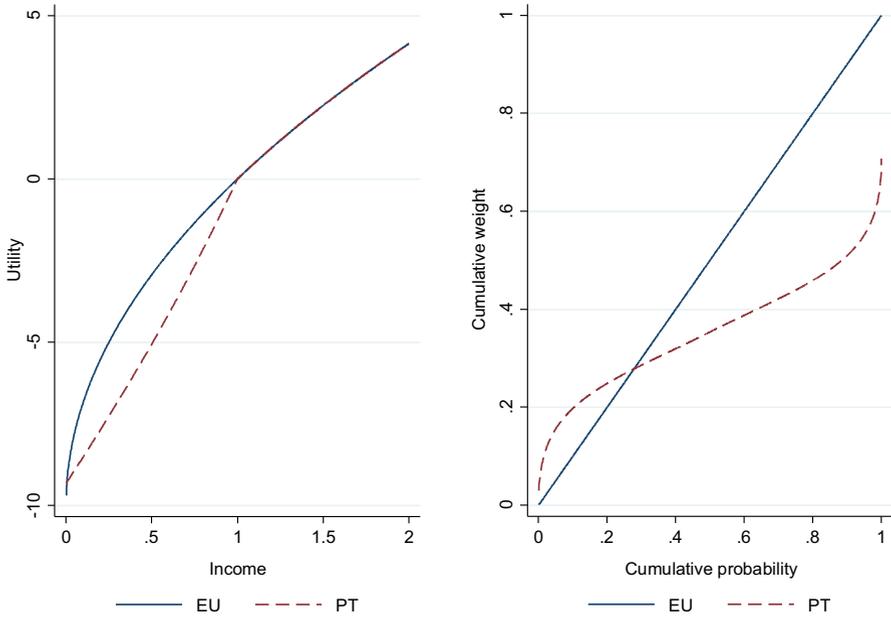
Fig. 1 Individual utility and probability weighting functions

Altogether, prospect theory produces a fourfold pattern of risk attitudes: risk aversion for small gains and large losses and risk seeking for large gains and small losses. The theory can explain, e.g., why some people buy both lottery tickets and insurance.

Given these properties, the utility and probability weighting functions can be parameterized in different ways. In the numerical exercises, I use the standard constant relative risk-aversion (CRRA) utility function. The utility functions for expected utility theory (EU) and prospect theory (PT) are:

$$u_{\text{EU}}(x) = \tilde{x}_{\text{EU}}^{\alpha} - \bar{x}, \quad \tilde{x}_{\text{EU}} = x, \quad \bar{x} = x_0^{\alpha}, \tag{3}$$

$$u_{\text{PT}}(x) = \begin{cases} \tilde{x}_{\text{PT}}^{\alpha} - \bar{x}, & x \geq x_0 \\ -\lambda \left[\tilde{x}_{\text{PT}}^{\alpha} - \bar{x} \right] & x < x_0 \end{cases}, \quad \tilde{x}_{\text{PT}} = |x - x_0| + x_0, \quad \bar{x} = x_0^{\alpha}, \tag{4}$$

where $0 < \alpha < 1$, and $\lambda > 1$. Augmenting the standard functions by $\bar{x}$ normalizes utility to zero at the reference income and restricts the utility of income levels above the reference income to be the same in expected utility theory and prospect theory. $\tilde{x}_{\text{EU}}$ and $\tilde{x}_{\text{PT}}$ accounts reference independence and dependence, respectively.

In Eq. (3), risk aversion decreases when α increases. In Eq. (4), marginal sensitivity of gains and losses increases when α increases. λ measures loss aversion. Marginal utility is also discontinuous at the reference point. The utility functions in Eqs. (3) and (4) are illustrated in the left panel of Fig. 1 in which x_0 is normalized to one.

For the cumulative weighting function, I use the following commonly used parameterizations in the numerical exercises:

$$w_{\mathrm{EU}}(x) = w_{\mathrm{EU}}(p(x)) = W_{\mathrm{EU}}(p) = p(x), \tag{5}$$

$$w_{\mathrm{PT}}(x; p(.)) : W_{\mathrm{PT}}(p) = \frac{p^{\gamma}}{(p^{\gamma} + (1-p)^{\gamma})^{1/\gamma}}. \tag{6}$$

Equation (5) just shows that with a linear cumulative weighting function, we obtain the linear expected utility probability weighting function. γ reflects the degree of overweighting and collapses to linear weights when $\gamma = 1$.

The probability weighting functions described by Eqs. (2), (5) and (6) are illustrated in the right panel of Fig. 1, where I plot the resulting cumulative weight against the cumulative probability with income levels ranked in ascending order. The derivatives of the graphs provide the weight given to each income level depending on its position in the distribution. For prospect theory, we plot the special case with a win probability of 50%.

Neilson and Stowe (2002) showed that given the functional form, only high values of α (> 0.5) can accommodate some gambling on unlikely gains. Furthermore, given high values of α, only low values of γ (< 0.3) can accommodate the Allais paradox. I set $\alpha = 0.5$ and $\gamma = 0.3$ which can accommodate both empirically observed phenomenon.[18] Most result patterns are insensitive to quite large variations of the two parameters. In particular, all patterns are preserved when increasing α and when decreasing γ. I set $\lambda = 2.25$ like estimated in Tversky and Kahneman (1992).

It could be argued from the point of view of individual well-being that risk aversion toward lotteries is rational because marginal income is typically spent on basic goods fulfilling essential needs at low income levels. Furthermore, equal concern of each potential future self seems to support the independence axiom. This is the type of argument used to motivate expected utility theory as a normative theory. On the other hand, reference points could be manipulated and could be unstable over time. Nevertheless, individual satisfaction including experienced and expressed well-being does depend on relative comparisons and excluding such satisfaction from individual welfare seems highly paternalistic. The individual preference foundation translates into a social preference foundation in the original position. Individual relative comparisons translate into social relative comparisons. The welfare of an individual depends on its position in the income distribution and the reference position.

There is, however, no dynamic aspect in the original position, alleviating unstable reference points over time. However, a reference income needs to be chosen. I explore the effects of using the mean income in the population as the reference income. This is also the income everyone has under complete equality, which is the optimal income distribution when using other social welfare criteria. However, it could be criticized

---

[18] In the studies by Tversky and Kahneman (1992), Camerer and Ho (1994), and Wu and Gonzalez (1996), varies between 0.32 and 0.88, and varies between 0.56 and 0.74. As Neilson and Stowe (2002) showed, the combinations of parameter values in these studies cannot account for both the mentioned phenomenon at the same time, unlike the intended design of prospect theory in Kahneman and Tversky (1979). The alternative parameterization in Prelec (1998) suffers from the same issues.

for being a choice that already embodies a normative statement—that some representative individual is the standard. Besides also elaborating with the median income as the reference income, I suggest and use a procedure to overcome the arbitrariness in choosing a *specific* reference income. The procedure takes the mean of the (hypothetical) social welfare evaluations (behind the veil of ignorance) of all individuals in the realized distribution, using the realized income of each individual as her reference income. This is formally defined in Definition 1.

Definition 1 Representative aggregation of reference incomes:

$$U\left(p\left(.\right)\right) = \sum_{x_0} p\left(x_0\right) U\left(p\left(.\right)|x_0\right). \tag{7}$$

$U\left(p\left(.\right)|x_0\right)$ is the objective function of an individual behind the veil of ignorance calculated using Eq. (1), given her reference income x_0. We now take the expectation over a distribution of reference incomes, using the actual distribution $p\left(.\right)$. This corresponds to letting each individual in the evaluated income distribution evaluating the income distribution behind the veil of ignorance *given* her ex post *realized* reference income, and then averaging over all individuals' *evaluations*. The procedure is representative by giving each individual's evaluation the same weight in the aggregation.

In the same way as expected utility theory is egalitarian in giving each individual's utility the same weight, my representative aggregation is egalitarian in giving each individual's impartial perception of the income distribution the same weight. This impartial perception requires of each individual to internalize the utility of other individuals. In this internalization, the individual is *not* hypothetically stripped of her reference income and satisfaction derived from social comparisons. We therefore come even closer to capturing realistic individual preferences. The averaging can be normatively motivated using the same independence argument as used in expected utility theory.[19]

Using representative aggregation of reference incomes with a prospect theory utility function transforms the decision problem in Eq. (1) to:

$$\max_{p(.)} U\left(p\left(.\right)\right) = \sum_{x_0} \sum_{x} p\left(x_0\right) w\left(x; p\left(.\right)|x_0\right) u\left(x|x_0\right). \tag{8}$$

Note that for an expected utility impartial observer, the reference income does not affect the social welfare evaluation.

3 Income distributions

To explore the effects of different components of prospect theory and to illustrate the basic intuition, I start with the simplest problem, where income can take two different

---

[19] It is possible to argue that the weight given to different welfare evaluations should not be equal. However, whereas human satisfaction depends on social comparisons of income, there is no evidence that human satisfaction depends on social comparisons of evaluations.
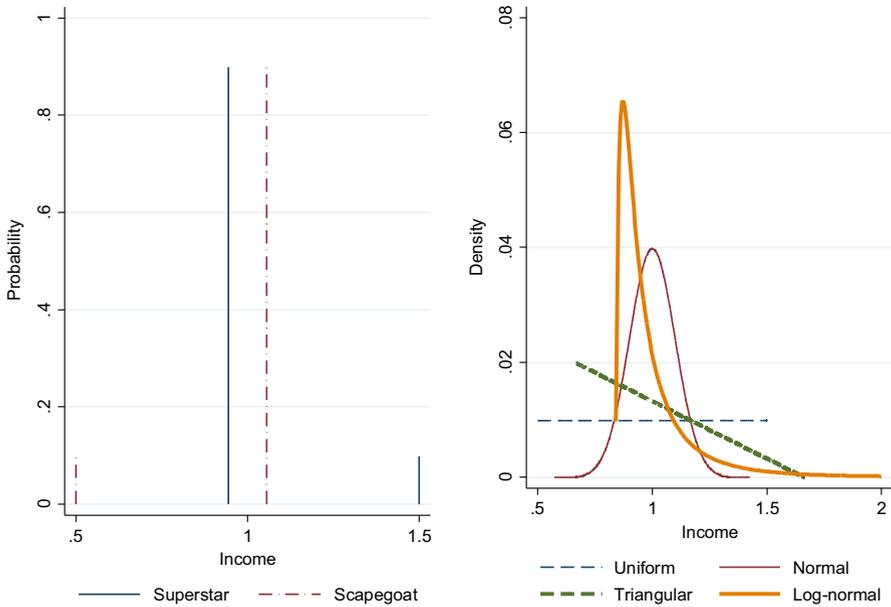
Fig. 2 Two-level and multi-level income distributions with a mean income of one

levels. Because of the revenue neutrality constraint, there are three possible types of income distributions. One type of income distribution is the bottom-heavy right-skewed distribution where a majority of individuals have less than the mean income and a minority of individuals have much more than the mean income. I call this type of distribution "the superstar distribution". Ex ante, before its realization, it embodies the American dream providing the impartial observer the opportunity to take a fair-odds long-shot gamble on becoming a superstar.

Another type of distribution is the top-heavy left-skewed distribution where a minority of individuals have much less than the mean income and a majority of individuals have more than the mean income. I call this type of distribution "the scapegoat distribution". Ex ante, it provides the impartial observer the possibility to take a fair-odds "safe bet" on not becoming the scapegoat. The two different types of distributions are displayed in the left panel of Fig. 2 where mean income is normalized to one. Furthermore, there is also the type of distribution where half of the individuals have less than the mean income and half of the individuals have more than the mean income.

In the numerical section, I also investigate some multi-level discrete approximations of continuous income distributions. I investigate some symmetric income distributions and some asymmetric superstar distributions because the two-level analysis will indicate that superstar distributions are particularly promising. The investigated income distributions include the uniform, normal, triangular, and log-normal income distributions. They are displayed in the right panel of Fig. 2. The log-normal distribution is of particular interest because the income distributions of most countries have this shape (Gibrat 1931; Aitchison and Brown 1957; Battistin et al. 2007).

The impartial observer can, of course, choose from any positive income distribution that preserves the mean and not only the income distributions investigated here. The optimal income distribution may be one that is not investigated here. Because of the difficulties with functional form and parameterization discussed in the last section, the exact welfare numbers (certainty equivalents in the numerical exercises) should not be taken too seriously.

For two-level income distributions, when the mean income in the population is the reference income, the decision problem in Eq. (1) is reduced to:[20]

$$\max_{x_g, p_g} w(p_g) u(x_m + x_g) + w(1 - p_g) u \left(x_m - \frac{p_g}{1 - p_g} x_g \right). \tag{9}$$

x_g is the size of the gains relative to the reference income $x_0 = x_m$, p_g is the gain probability, $x_g \geq 0$, and $w(x_g, p(.)) = w(p_g) = W(p_g)$. With a finite population, $0 < \underline{p}_g < p_g < 1 - \underline{p}_g < 1$. $\underline{p}_g$ is the lower bound corresponding to one individual with more than the mean income. For some results, I assume that the number of individuals in the population is large and that $\underline{p}_g$ is close to zero.

When using representative aggregation of reference incomes, the decision problem in Eq. (1) for a prospect theory impartial observer instead becomes:[21]

$$\max_{x_g, p_g} (1 - p_g) w(p_g) u_{PT} \left(x_0 + \frac{1}{1 - p_g} x_g \right)$$
$$+ p_g w(1 - p_g) u_{PT} \left(x_0 - \frac{1}{1 - p_g} x_g \right). \tag{10}$$

4 Analytical results

Let us start with the expected utility optimum for two-level income distributions. The decision problem is formulated in Eq. (9). Because of diminishing marginal utility, spread cannot be desirable. This classical equality result is stated in Proposition 1. All proofs are presented in the Appendix. The equality solution provides a utility of zero. The result can be extended to the case allowing for continuous income distributions for concave utility functions. See, e.g., Mas-Collel et al. (1995), who show that concavity implies preferences against mean-preserving spreads.

Proposition 1 *For two-level income distributions, complete equality is optimal when using expected utility theory, i.e., $x_g^* = 0$.*

The overweighting of probabilities of the relatively largest gains in prospect theory creates a possibility for an uneven income distribution to be optimal by accumulating

---

[20] For the two-level case the constraint in Eq. (1) becomes $p_g x_g + p_l x_l = 0$, where l indexes the loss state. Making use of $p_l = (1 - p_g)$ gives $x_l = -p_g x_g/(1 - p_g)$ leading to the argument in $u(.)$ for the loss state in Eq. (9).

[21] The first term represents $(1 - p_g)$ individuals with the low income reference-level that see a p_g chance, given the weight $w(p_g)$, of getting $(x_g + x_m) - (x_m + x_l) = x_g/(1 - p_g)$ in addition their reference level, after making use of the in footnote 20. The second term is similarly derived.

gains (incomes above the mean income) among a few individuals at the expense of smaller losses (incomes below the mean income) for a larger number of individuals. With a concave utility function, such an outcome is optimal, both when the mean income is the reference income and when using representative aggregation of reference incomes, according to Proposition 2.

Proposition 2 *For two-level income distributions, when using a prospect theory probability weighting function and a weakly concave utility function, we have that*:

(a) *A superstar distribution with $0 < x_g^* \leq \frac{1-p_g^*}{p_g^*} x_m$ and $p_g^* < 0.5$ is optimal. The crucial condition for this result is:*

$$\frac{w(p_g)}{p_g} > \frac{w(1-p_g)}{1-p_g}, \tag{11}$$

which is implied by the prospect theory probability weighting function.

(b) *With a linear utility function, a superstar distribution with $x_g^* = \frac{1-p_g^*}{p_g^*} x_m$ and $p_g^* < 0.5$ is optimal.*

The superstar distribution in Proposition 2a contains at least one superstar and at most half the population as superstars, with much more than the mean income, supported by all other individuals having less than the mean income. The results depend on the parameterization. The upper bound on gains occurs where the individuals with less than the mean income have no income, and the superstars have all income. With a linear utility function, no factor works against spread, and the optimal income of superstars approach its upper bound.

The diminishing marginal sensitivity in losses in prospect theory creates another possibility for an uneven income distribution to be optimal by accumulating losses among a few individuals to allow smaller gains for a larger number of individuals. With a linear probability weighting function, such an outcome could be optimal when the mean income is the reference income, but not when using representative aggregation of reference incomes, according to Proposition 3.

Proposition 3 *For two-level income distributions, when using a prospect theory utility function and a linear probability weighting function, we have that:*

(a) *When the mean income is the reference income,* either *of the following two conditions is sufficient for a scapegoat distribution with $x_g^* > 0$ and $p_g^* \to 1 - \underline{p}_g$ to be optimal:*

$$\lim_{x_g \to 0} \left[u'_{PT}(x_m + x_g) - u'_{PT}\left(x_m - \frac{1-\underline{p}_g}{\underline{p}_g} x_g \right) \right] > 0 \, or \tag{12}$$

$$(1-\underline{p}_g)\left[u_{PT}\left(\frac{1}{1-\underline{p}_g} x_m \right) - u_{PT}(x_m) \right]$$
$$> \underline{p}_g \left[u_{PT}(x_m) - u_{PT}(0) \right]. \tag{13}$$

(b) *If x is unconstrained and $\underline{p}_g \to 0^+$ in (a), the following weaker condition is sufficient for a scapegoat distribution: there is an x such that $u'_{PT}(x > x_m) > \lim_{z \to -\infty} u'_{PT}(z)$.*

(c) *Complete equality is optimal when applying representative aggregation of reference incomes.*

The scapegoat distribution in Proposition 3a contains one individual scapegoat with much less than the mean income, sacrificed so that all other individuals can have more than the mean income. Whether such a distribution is optimal depends on the parameterization. The first sufficient condition in Eq. (12) requires the marginal utility of gains at the mean income to be larger than the marginal disutility of losses that are larger than the gains at the mean income. Whether the condition holds depends on three factors. Loss aversion works against the condition because it leads to the marginal utility being greater for losses of the same size as gains. The number of individuals and the degree of diminishing sensitivity in losses work in favor of the condition because the losses are larger than the gains and because marginal disutility decreases with losses.

The second sufficient condition in Eq. (13) requires that the gain utility of a small gain weighted by the number of individuals getting the gain is greater than the loss utility of one individual losing all her income. Again, loss aversion works against the condition, whereas diminishing sensitivity in losses works in favor of it. The purpose of having this second condition is to show that we do not require a condition involving the marginal utility at the mean income.

When allowing for income to be unconstrained, the condition required for a scapegoat distribution to be optimal becomes weaker in Proposition 3b. We then only require the marginal disutility at the (possibly hypothetical) worst-off loss to be small enough in comparison with the marginal utility at an arbitrary gain. This is fulfilled, e.g., if the marginal disutility of losses converges to zero or if the marginal utility of gains is infinite for the first dollar, an Inada condition often assumed on utility functions. When interpreting the input in the utility function as income or resources, the lower zero bound is natural. If the amount of resources to distribute is large, the lower constraint may, however, be relatively very low, and the condition assuming unboundedness may be a good approximation. A lower bound greater than zero may be desirable for other reasons, e.g., if there are basic goods without which individuals experience extreme disutility.[22]

The optimization problem becomes much more complicated when combining prospect theory utility and probability weighting functions. In general, the solution depends on the exact parameterization of the functions. In Proposition 4, I state two sufficient conditions for a superstar distribution to increase social welfare compared to complete equality.

---

[22] Prospect theory has not been empirically tested in situations involving extremely low income levels where individuals cannot afford basic goods. The acceptability of the conclusion of one suffering individual to let the others thrive (a tiny bit more) may be considered to be normatively unacceptable. It would be simple to reframe our decision problem as distributing a fixed amount of non-basic goods.

Proposition 4 *For two-level income distributions, when using prospect theory utility and probability weighting functions, we have that:*

(a) *When the mean income is the reference income or when using representative aggregation of reference incomes,*

$$\lim_{p_g \to 0^+} w(p_g) > 0 \tag{14}$$

and $\underline{p}_g \to 0^+$ *are sufficient for a superstar distribution with* $x_g^* > 0$ *and* $p_g^* \to \underline{p}_g$ *to be preferred to complete equality, i.e., there is an* x_g *such that* $U_{PT}(x_g > 0, p_g \to \underline{p}_g) > 0 = U_{PT}(x_g = 0)$.

(b) *When using representative aggregation of reference incomes, the following condition is sufficient for a superstar distribution with* $x_g^* > 0$ *and* $p_g^* < 0.5$:

$$\frac{w(p_g)}{p_g} \frac{1 - p_g}{w(1 - p_g)} > \lim_{x_g \to 0^+} \frac{u'_{PT}(x_m - x_g)}{u'_{PT}(x_m + x_g)}. \tag{15}$$

The superstar distributions in Proposition 4a and 4b are not necessarily the optimal income distribution, but they are preferred to complete equality. Because $w(p) = 0$ when $p = 0$, the condition in Eq. (14) requires the probability weighting function to be discontinuous at $p = 0$. The uneven income distribution is better because increasing the probability of gains from zero results in a positive discontinuous increase in social welfare due to positive probability weights on the positive gain utility, whereas decreasing the accompanying losses much less results in a continuous decrease in the negative loss utility. The exact optimal gains and gain probability depend, however, on the parameterization. The condition is not very strong. In the original version of prospect theory in Kahneman and Tversky (1979), the authors had this type of probability weighting function in mind, which reflects that even a small probability is categorized as a possibility given considerable weight.

The left-hand side of the condition in Eq. (15) reminds of Eq. (11). It is the degree of overweighting of small probabilities of large gains multiplied by the inverse of the degree of underweighting of large probabilities of small losses. As discussed in connection to Proposition 2, this factor is greater than one. The right hand side of Eq. (15) is the quotient between the marginal disutility of losses and the marginal utility of gains around the reference income. Hence, the condition requires the departure from linear probability weights to be larger than the degree of loss aversion.

Thus far, we have dealt with two-level income distributions. Real income distributions, however, certainly allow for and often have more income levels. Can the results be extended to such income distributions? The optimization problem increases in dimensionality by twice the additional number of income levels, increasing the difficulty of obtaining analytical solutions. Nevertheless, the shapes of the prospect theory utility and probability weighting functions still have similar impacts. A three-level income distribution can be created from a two-level income distribution by applying a mean-preserving spread on one of the income levels in a two-level income distribution. In Proposition 5, I claim that such a spread can increase social welfare when the mean

income is the reference income. The argument in Proposition 5 can be iterated to show that much more complicated multi-income level distributions can be preferred to the optimal three-level income distribution. The results depend, however, crucially on the parameterization.

Proposition 5 *When using prospect theory utility and probability weighting functions with the mean income as the reference income, three-level income distributions can be preferred to two-level income distributions.*

As already discussed, a crucial component of prospect theory is the selection of a reference point. We have so far taken this reference point to be the mean income or used representative aggregation of reference incomes. Another option is the median income. Unlike the mean, the median can change discontinuously when altering the income distribution continuously. The preference for an income distribution increases when the reference income decreases, leaving room for manipulation of the reference point. A systematic way to increase the preference for an income distribution is presented in Proposition 6.

Proposition 6 *When the median income is used as the reference income, decreasing the income levels in the lower end of the income distribution including the median income is desirable if it decreases the distance between the median income and every lower income level.*

Note that if we decrease the income levels according to Proposition 6, keeping the mean income fixed, the decrease in income at low income levels provides income that can be redistributed to individuals at high income levels, which additionally increases the preference for the income distribution.

5 Numerical results

The analytical complexity involved in obtaining results when combining prospect theory utility and probability weighting functions can be avoided by using numerical methods. Multi-level income distributions can also easily be investigated numerically. Furthermore, we can quantify the welfare effects. On the down side, some results are driven by the parameterization.

When reporting the social welfare of an income distribution, I report certainty equivalents rather than social welfare. The certainty equivalent is the additional percent of income that the individuals need when each of them have the mean income to reach the social welfare of an income distribution. In constructing the certainty equivalent of social welfare, I use the expected utility social welfare formula to enable straightforward comparison of certainty equivalents independent of which theory is used to calculate social welfare. The numerical exercises are performed using 1,000 individuals spaced at income levels with the same probability density mass between them. At this coarseness, the results are insensitive, but computational time is very sensitive, to varying coarseness.

The certainty equivalents of different two-level superstar distributions are reported in Table 1. I vary the size of income gains relative to the mean income in percent

Table 1 Certainty equivalents of some two-level superstar distributions

Gains	Gain probability	EU	PT mean	PT representative
0.5	1	−0.000006	0.05	0.26
5	1	−0.0006	0.54	2.55
50	1	−0.05	4.91	24.36
5	0.1	−0.00006	0.41	1.61
5	1	−0.0006	0.54	2.55
5	10	−0.007	0.35	3.31

Gain probability is expressed in percent and other numbers are expressed in percent of the mean income

and the gain probability in percent (keeping the mean income constant in all income distributions). I report the certainty equivalents in percent of the mean income for an expected utility (EU) impartial observer, and for prospect theory (PT) impartial observers when the mean income (mean) is the reference income and when using representative aggregation of reference incomes (representative).

We observe that the superstar distributions provide negative social welfare for the expected utility impartial observer. The social welfare loss relative to complete equality increases with the size of the income gains and the gain probability. However, the income distributions provide positive social welfare for the prospect theory impartial observers and are hence preferred to complete equality for them. The social welfare gain is much larger when using representative aggregation of reference incomes than when the mean income is the reference income. It increases with the size of the income gains. It also increases with the gain probability, albeit only up to a gain probability of 1% when the mean income is the reference income.

The size of the social welfare gains for the prospect theory impartial observers is large, up to a certainty equivalent of 24.36% of the mean income when using representative aggregation of reference incomes. The magnitude of the effects is larger for the prospect theory impartial observers than for the expected utility impartial observer. This is because the gain probabilities are small and hence carry small weight for the expected utility impartial observer, whereas the prospect theory impartial observers overweight those probabilities. The patterns found are in line with Propositions 1, 2, and 4.

The certainty equivalents of different two-level scapegoat distributions are reported in Table 2, which is similarly organized as Table 1. The expected utility impartial observer again prefers income distributions that are the closest to complete equality. Unlike superstar distributions, social welfare is also negative for prospect theory impartial observers. The social welfare loss increases with the size of the income losses and the loss probability. Furthermore, the size of the social welfare losses is much larger for the prospect theory impartial observers. This reflects the impact of loss aversion and the overweighting of large losses.

However, as the income losses or loss probability increase, the additional social welfare loss is less in relative terms for the prospect theory impartial observers than for the expected utility impartial observer. For instance, increasing the income losses

Table 2 Certainty equivalents of some two-level scapegoat distributions

Losses	Loss probability	EU	PT mean	PT representative
0.5	1	−0.000006	−0.13	−0.13
5	1	−0.0006	−1.30	−1.29
50	1	−0.09	−11.54	−11.38
5	0.1	−0.00006	−0.94	−0.94
5	1	−0.0006	−1.30	−1.29
5	10	−0.007	−1.39	−1.06

Loss probability is expressed in percent and other numbers are expressed in percent of the mean income

Table 3 Certainty equivalents of some symmetric multi-level income distributions

Spread	Uniform		
	EU	PT mean	PT representative
1	−0.0002	−0.09	−0.10
10	−0.02	−0.89	−0.95
100	−2.18	−7.99	−8.31
Variance	Normal		
	EU	PT mean	PT representative
1	−0.002	−0.50	−0.53
10	−0.02	−1.55	−1.63
100	−0.25	−4.66	−4.89

Spread is top income minus bottom income in percent of the mean income, variance is the distribution variance in percent of the mean income, and other numbers are expressed in percent of the mean income

ten times (from, e.g., 5 to 50) increases the certainty equivalent of the social welfare loss more than 100 times (from −0.0006 to −0.09) for the expected utility impartial observer, but less than ten times (from around −1.3 to around −11.5) for the prospect theory impartial observers.

The certainty equivalents of the symmetric uniform and normal multi-level income distributions are reported in Table 3. We observe that social welfare is negative for both distributions independent of the theory applied. It also decreases as spread and variance increase. More inequality is, therefore, in general, also undesirable for the prospect theory impartial observers. Like for the scapegoat distributions, the social welfare loss is larger for the prospect theory impartial observers because of loss aversion. The additional relative negative effect of additional spread on social welfare is, however, again relatively smaller for the prospect theory impartial observers.

Because of the desirability of two-level superstar distributions, I also investigate some asymmetric multi-level superstar distributions. The certainty equivalents of some triangular and log-normal distributions are reported in Table 4. They are all negative for the triangular distributions. The pattern is very similar to the one of uniform distributions. A difference is that the social welfare loss is smaller for a triangular

Table 4 Certainty equivalents of some asymmetric multi-level superstar distributions

Spread	Triangular		
	EU	PT mean	PT representative
1	−0.0001	−0.02	−0.03
10	−0.01	−0.22	−0.30
100	−1.35	−2.42	−3.02
Variance	Log-normal		
	EU	PT mean	PT representative
1	−0.01	1.65	1.52
10	−0.09	4.76	4.35
100	−0.77	12.19	11.09

Spread is top income minus bottom income in percent of average income, variance is the distribution variance in percent of average income, and other numbers are expressed in percent of the mean income

distribution with the same spread. Furthermore, the social welfare loss for the prospect theory impartial observers relative to that of the expected utility impartial observer is smaller for the triangular distribution (for a spread of 100 when the mean income is the reference income, we have the certainty equivalent comparison −2.42 versus −1.35) than for the uniform distribution (for a spread of 100 when the mean income is the reference income, we have the certainty equivalent comparison −7.99 versus −2.18). This is the effect of prospect theory impartial observers liking superstar distributions, although not enough to make them preferring triangular distributions to complete equality.

The log-normal distribution is also right-skewed like the triangular distribution. However, the skewness is greater. This skewness manages to turn social welfare positive for the prospect theory impartial observers. Using the mean income as the reference income or representative aggregation of reference incomes has small effects on the results. Furthermore, social welfare increases as variance increases. The social welfare gains are large with a certainty equivalent of at most 12% of the mean income. However, they are not as large as in the most preferred two-level superstar distribution which had a social welfare gain corresponding to a certainty equivalent of 24% of the mean income (see Table 1).

6 Concluding discussion

Harsanyi (1953, 1955, 1975) and Rawls (1971) offered two of the most influential theories of distributive justice. Both used the popular social contract approach starting from an original position where the impartial observer does not know her identity in the society. Under such a veil of ignorance, her choice of income distribution could be considered the fair distribution. This paper asked how an impartial observer applying cumulative prospect theory would choose. Applying prospect theory is appealing

because it better describes behavior under risk than expected utility theory, which Harsanyi's impartial observer uses.

I found that the perceived desirability of different income distributions depends on the parameterization of prospect theory. Two properties of prospect theory work in the direction of increasing the preference for uneven income distributions: the overweighting of the relatively largest gains and diminishing sensitivity in losses. For a reasonably chosen parameterization, prospect theory impartial observers are in general *more* inequality averse than an expected utility impartial observer. However, inequality could be perceived as desirable when it comes to a *specific type* of income distribution that is bottom-heavy and right-skewed where few superstars have very high income and many individuals have low income, such as the log-normal income distribution.

Another alternative approach to applying prospect theory in the original position that treats the decision problem in the original position as a descriptive question is to ask real individuals or groups about what they would prefer or could agree on in this position. A problem could be that individuals may not fully interpret the original position the way they are supposed to, that they may not know how they actually would choose in the original position, or that they may not truthfully reveal how they would choose (e.g., by instead revealing how they think they should choose). Given the parsimonious framing of the original position as a pure decision under risk, the survey approach would also arrive at the pattern predicted by prospect theory.

The lottery interpretation of the original position is a way to capture impartiality by forcing individuals to put themselves in other individuals' positions. This paper spelled out the implications. For proponents of the idea that the procedural justice in the original position appropriately embodies impartiality and that impartiality is central for fairness, some types of inequality must be viewed as socially desirable. Disregarding this conclusion, the optimal inequality results imply that individuals would voluntarily join lotteries that lead to the optimal-inequality type of distribution when imposing complete equality.

From the point of view of the welfare foundation, the social desirability of optimal inequality depends on whether we should give any weight to (dis)satisfaction derived from social comparisons between individuals in a society. On the one hand, such satisfaction is not inherent to the practical usefulness of income. On the other hand, the perception of the resources of others does substantially affect even experienced well-being.

It is possible to argue that optimal inequality is an unpleasant and unacceptable result indicating that the original position needs to be modified or rejected. An issue is what type of risk preferences the impartial observer should have, if any, in the original position. It could be argued that some risk preferences such as the one implied by prospect theory are inappropriate in the original position. This line of argument, however, amounts to rejecting that the original position transforms the normative question into a descriptive question; it really implies that the difficult normative question (about the fair income distribution) is replaced by another (equally difficult?) normative question (about what risk preferences an impartial observer in the original position should have). If resorting to this argument, it is possible to argue that expected utility theory

or some other theory of decision under risk is a better normative theory than prospect theory and that the best normative theory should be applied in the original position.

Another issue is whether the problem in the original position is one about decision under risk. Rawls (1971) argues that it is a decision under uncertainty where the impartial observer should disregard the frequencies of different income levels. In comparing distributions with known frequencies, it seems difficult to argue why they should be disregarded. Even then, it is unclear which theory of behavior under uncertainty is the appropriate one, if not a prospect theory for uncertainty. Along this line of reasoning, it is possible to argue that risk and uncertainty preferences are irrelevant or should be ignored. But with such a thick veil, it seems hard to say anything about the social welfare of income distributions.

Appendix

Proof of Proposition 1 We have the following derivative of the objective function in Eq. (9):

$$\frac{dU_{EU}}{dx_g} = p_g \left[u'_{EU}(x_m + x_g) - u'_{EU}\left(x_m - \frac{p_g}{1 - p_g} x_g \right) \right]$$

$$\leq 0 \text{ since } u'(x > x_m) < u'(x < x_m).$$

Hence, increasing spread from $x_g = 0$ decreases expected utility.   □

Proof of Proposition 2

(a) We have the following derivatives of the objective functions in Eqs. (9) and (10) respectively:

$$\frac{dU}{dx_g} = w(p_g)u'_{EU}(x_m + x_g)$$

$$- \frac{p_g}{1 - p_g} w(1 - p_g)u'_{EU}\left(x_m - \frac{p_g}{1 - p_g} x_g \right),$$

$$\frac{dU}{dx_g} = w(p_g)u'_{EU}\left(x_m + \frac{1}{1 - p_g} x_g \right)$$

$$- \frac{p_g}{1 - p_g} w(1 - p_g)u'_{EU}\left(x_m - \frac{p_g}{1 - p_g} x_g \right).$$

For $p_g < p_g^\# < 0.5$, where $w(p_g^\#) = p_g^\#$, subproportionality implies overweighting and $w(p_g) > p_g$, which together with subcertainty implies $w(1 - p_g) <$

$1 - p_\mathrm{g}$. The two inequalities imply Eq. (11), which gives $\frac{\mathrm{d}U}{\mathrm{d}x_\mathrm{g}}(x_\mathrm{g} = 0) > 0$ and $x_\mathrm{g}^* > 0$. $p_\mathrm{g} < p_\mathrm{g}^\#$ is a sufficient condition, but not a necessary condition. $p_\mathrm{g}^* < 0.5$ is, however, a necessary condition.

(b) We have the following derivative of the objective function in Eq. (9):

$$\frac{\mathrm{d}U}{\mathrm{d}x_\mathrm{g}} = w(p_\mathrm{g}) - \frac{p_\mathrm{g}}{1 - p_\mathrm{g}} w(1 - p_\mathrm{g}).$$

The same argument as in (a) leads to Eq. (11) which now implies $\frac{\mathrm{d}U}{\mathrm{d}x_\mathrm{g}} > 0$. We want to increase x_g^* until its maximum. Because of the lower bound of x, $x_\mathrm{m} - \frac{p_\mathrm{g}}{1 - p_\mathrm{g}} x_\mathrm{g} > 0$, we get $x_\mathrm{g}^* = \frac{1 - p_\mathrm{g}^*}{p_\mathrm{g}^*} x_\mathrm{m}$. □

Proof of Proposition 3

(a) We have the following derivatives of the objective function in Eq. (9):

$$\frac{\mathrm{d}^2 U}{\mathrm{d}p_\mathrm{g}^2} = \frac{x_\mathrm{g}^2}{(1 - p_\mathrm{g})^3} u_\mathrm{PT}'' \left(x_\mathrm{m} - \frac{p_\mathrm{g}}{1 - p_\mathrm{g}} x_\mathrm{g} \right),$$

$$\frac{\mathrm{d}U}{\mathrm{d}x_\mathrm{g}} = p_\mathrm{g} u_\mathrm{PT}'(x_\mathrm{m} + x_\mathrm{g}) - p_\mathrm{g} u_\mathrm{PT}' \left(x_\mathrm{m} - \frac{p_\mathrm{g}}{1 - p_\mathrm{g}} x_\mathrm{g} \right).$$

Because $\frac{\mathrm{d}^2 U}{\mathrm{d}p_\mathrm{g}^2} > 0$, U has an interior minimum in p_g, and it must be either that $p_\mathrm{g}^* \to \underline{p}_\mathrm{g}$ or $p_\mathrm{g}^* \to 1 - \underline{p}_\mathrm{g}$. For $p_\mathrm{g}^* \to \underline{p}_\mathrm{g}$, we have $u_\mathrm{PT}'(x_\mathrm{m} + x_\mathrm{g}) < u_\mathrm{PT}'(x_\mathrm{m} - ax_\mathrm{g})$ with $a < 1$ because of loss aversion, which implies $\frac{\mathrm{d}U}{\mathrm{d}x_\mathrm{g}} < 0$, giving $x_\mathrm{g}^* = 0$ as the optimum. For $p_\mathrm{g}^* \to 1 - \underline{p}_\mathrm{g}$, Eq. (12) implies $\frac{\mathrm{d}U}{\mathrm{d}x_\mathrm{g}}(x_\mathrm{g} = 0) > 0$ and $x_\mathrm{g}^* > 0$. If $\frac{\mathrm{d}U}{\mathrm{d}x_\mathrm{g}}$ is strictly decreasing in x_g, we want to increase x_g up until $\frac{\mathrm{d}U}{\mathrm{d}x_\mathrm{g}} = 0$, or until its upper bound $x_\mathrm{g} = \frac{p_\mathrm{g}}{1 - \underline{p}_\mathrm{g}} x_\mathrm{m}$. Equation (13) is sufficient for $x_\mathrm{g}^* > 0$, because it implies $U \left(x_\mathrm{g} = \frac{p_\mathrm{g}}{1 - \underline{p}_\mathrm{g}} x_\mathrm{m}, p_\mathrm{g} = 1 - \underline{p}_\mathrm{g} \right) > 0 = U(x_\mathrm{g} = 0)$.

(b) In (a), there is an x such that $u_\mathrm{PT}'(x > x_\mathrm{m}) > \lim_{z \to -\infty} u_\mathrm{PT}'(z)$ implies $\lim_{x_\mathrm{g} \to 0^+} u_\mathrm{PT}'(x_\mathrm{m} + x_\mathrm{g}) > \lim_{z \to -\infty} u_\mathrm{PT}'(z)$ and $\frac{\mathrm{d}U}{\mathrm{d}x} > 0$. This gives $x_\mathrm{g}^* > 0$ when $p_\mathrm{g}^* \to 1^-$. We want to increase x_g infinitely, i.e., $x_\mathrm{g}^* = \infty$.

(c) We have the following derivative of the objective function in Eq. (10):

$$\frac{\mathrm{d}U}{\mathrm{d}x_\mathrm{g}} = p_\mathrm{g} \left[u_\mathrm{PT}' \left(x_0 + \frac{1}{1 - p_\mathrm{g}} x_\mathrm{g} \right) - u_\mathrm{PT}' \left(x_0 - \frac{1}{1 - p_\mathrm{g}} x_\mathrm{g} \right) \right].$$

We have $u_\mathrm{PT}'(x_\mathrm{m} + z) < u_\mathrm{PT}'(x_\mathrm{m} - z)$ for all z, because of loss aversion. Hence, $\frac{\mathrm{d}U}{\mathrm{d}x_\mathrm{g}} < 0$. □

Proof of Proposition 4

(a) In both cases, Eq. (14) implies $\lim_{p_g \to 0^+} U_{PT}(x_g > 0, p_g) = \lim_{p_g \to 0^+} w(p_g) u_{PT}(x_m + x_g) > 0 = U_{PT}(x_g = 0)$.

(b) We have the following derivatives of the objective function in Eq. (10):

$$\frac{dU_{PT}}{dx_g} = w(p_g) u'_{PT}\left(x_m + \frac{1}{1 - p_g} x_g\right)$$
$$- \frac{p_g}{1 - p_g} w(1 - p_g) u'_{PT}\left(x_m - \frac{1}{1 - p_g} x_g\right).$$

Equation (15) implies $\frac{dU_{PT}}{dx_g}(x_g = 0) > 0$ and hence $x_g^* > 0$. $p_g^* < 0.5$ is needed for $\frac{w(p_g)}{p_g} \frac{1 - p_g}{w(1 - p_g)} > 1$, see the proof of Proposition 2a. □

Proof of Proposition 5 Start out from the optimal two-level income distribution. Can we improve on it by dividing one of the income levels into two levels using a mean-preserving spread? Start with the income level below the mean income. Assume the mean-preserving spread increases income by a small $z < \frac{p_g}{1 - p_g} x_g$ (so that the new income levels are still below the mean income) in half the cases and $-z$ in the other half. Then, the convexity of the prospect theory utility function implies that such a spread increases social welfare if linear probability weights or prospect theory probability weights that are close enough to linear probability weights are used. Can the income level above the mean income be improved by such a spread? Assume the same type of mean-preserving spread with $z < x_g$. Then, if the probability weighting function is convex around p, which it may be in prospect theory, the spread increases social welfare if a linear utility function or a prospect theory utility function that is close enough to the linear function is used. We have thus established at least two situations where a three-level income distribution created from a two-level income distribution can be preferred to the optimal two-level income distribution. □

Proof of Proposition 6 Obviously, decreasing the distance between the median income and every lower income levels decreases the loss utility of those lower income levels. The decrease in the median income also increases the gain utility of the higher income levels. □

References

Aitchison, J., & Brown, J. (1957). *The lognormal distribution*. Cambridge: Cambridge University Press.

Amiel, Y., Cowell, F., & Gaertner, W. (2009). To be or not to be involved: A questionnaire-experimental view on Harsanyi's utilitarian ethics. *Social Choice and Welfare*, *32*, 299–316.

Atkinson, A. (1970). On the measurement of inequality. *Journal of Economic Theory*, *2*, 244–263.

Battistin, E., Blundell, R., & Lewbel, A. (2007). *Why is consumption more log normal than income?*. Manuscript: Gibrat's law revisited.

Barberis, N., & Huang, M. (2008). Stocks as lotteries: The implications of probability weighting for security prices. *American Economic Review*, *98*, 2066–2100.

Barberis, N., & Xiong, W. (2009). What drives the disposition effect? An analysis of a longstanding preference-based explanation. *Journal of Finance*, *64*, 751–784.

Bernheim, D., & Rangel, A. (2009). Beyond revealed preference: Choice-theoretic foundations for behavioral welfare economics. *Quarterly Journal of Economics*, *124*, 51–104.

Bosmans, K., & Schokkaert, E. (2004). Social welfare, the veil of ignorance and purely individual risk: An empirical examination. *Research on Economic Inequality*, *11*, 8–114.

Boulding, K. (1962). Social justice in social dynamics. In R. Brandt (Ed.), *Social justice*. Englewood Cliffs: Prentice Hall.

Camerer, C., & Ho, T. (1994). Violations of the betweenness axiom and nonlinearity in probability. *Journal of Risk and Uncertainty*, *8*, 167–196.

Choquet, G. (1955). Theory of capacities. *Annales de L'Institut Fourier*, *5*, 131–295.

Easterlin, R. (1974). Does economic growth improve the human lot? Some empirical evidence. In P. David & M. Reder (Eds.), *Nations and households in economic growth: Essays in honor of Moses Abramowitz*. New York: Academic Press.

Frey, B., & Stutzer, A. (2002). What can economists learn from happiness research? *Journal of Economic Literature*, *40*, 402–435.

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

Fennema, H., & Wakker, P. (1997). Original and new prospect theory: A discussion and empirical differences. *Journal of Behavioral Decision Making*, *10*, 53–64.

Fishburn, P., & Kochenberger, G. (1979). Two-piece von Neumann–Morgenstern utility functions. *Decision Science*, *10*, 503–518.

Frohlich, N., Oppenheimer, J., & Eavey, C. (1987). Choices of principles of distributive justice in experimental groups. *American Journal of Political Science*, *31*, 606–636.

Gibrat, R. (1931). *Les Inegalites Economiques*. Paris: Librairie du Recueil Sirey.

Günther, I., & Maier, J. (2008). *Poverty, vulnerability and loss aversion*. Manuscript, University of Göttingen.

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

Harsanyi, J. (1955). Cardinal welfare, individualistic ethics and interpersonal comparisons of utility. *Journal of Political Economy*, *63*, 309–332.

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

Herne, K., & Suojanen, M. (2004). The role of information in choices over income distributions. *Journal of Conflict Resolution*, *48*, 173–193.

Hershey, J., & Schoemaker, P. (1980). Prospect theory's reflection hypothesis: A critical examination. *Organizational Behavior and Human Decision Processes*, *25*, 395–418.

Hume, D. (1739). *A treatise of human nature* (2nd ed.). Oxford: Oxford University Press.

Johansson-Stenman, O., Carlsson, F., & Gupta, G. (2002). Measuring future grandparents' preferences for equality and relative standing. *Economic Journal*, *112*, 362–383.

Jäntti, M., Kanbur, R., Nyyssölä, M., & Pirttilä, J. (2014). Poverty and welfare measurement on the basis of prospect theory. *Review of Income and Wealth, 60*, 182–205.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*, 263–291.

Kahneman, D., & Tversky, A. (2000). *Choices, values, and frames*. Cambridge: Cambridge University Press.

Kahneman, D., Wakker, P., & Sarin, R. (1997). Back to Bentham? Explorations of experienced utility. *Quarterly Journal of Economics*, *112*, 375–406.

Karni, E. (1998). Impartiality: Definition and representation. *Econometrica*, *66*, 1405–1415.

Kothiyal, A., Spinu, V., & Wakker, P. (2011). Prospect theory for continuous distributions: A preference foundation. *Journal of Risk and Uncertainty*, *42*, 195–210.

Mas-Collel, A., Whinston, M., & Green, J. (1995). *Microeconomic theory*. Oxford: Oxford University Press.

Mongin, P. (2001). The impartial observer theorem of social ethics. *Economics and Philosophy*, *17*, 147–179.

Moreno-Ternero, J., & Roemer, J. (2008). The veil of ignorance violates priority. *Economics and Philosophy*, *24*, 233–257.

Neilson, W., & Stowe, J. (2002). A further examination of cumulative prospect theory parameterizations. *Journal of Risk and Uncertainty*, *24*, 31–46.

Payne, J., Laughhunn, D., & Crum, R. (1981). Aspiration level effects in risky behavior. *Management Science*, *27*, 953–958.

Prelec, D. (1998). The probability weighting function. *Econometrica*, *66*, 497–527.

Rawls, J. (1971). *A theory of justice*. Cambridge: Harvard University Press.

Rieger, M., & Wang, M. (2008). Prospect theory for continuous distributions. *Journal of Risk and Uncertainty*, *36*, 83–102.

Traub, S., Seidl, C., Schmidt, U., & Levati, M. (2005). Friedman, Harsanyi, Rawls, Boulding—or somebody else? An experimental investigation of distributive justice. *Social Choice and Welfare*, *24*, 283–309.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*, 297–323.

Von Neumann, J., & Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton: Princeton University Press.

Wakker, P., & Tversky, A. (1993). An axiomatization of cumulative prospect theory. *Journal of Risk and Uncertainty*, *7*, 147–176.

Wehrung, D. (1989). Risk taking over gains and losses: A study of oil executives. *Annals of Operations Research*, *19*, 115–139.

Wu, G., & Gonzalez, R. (1996). Curvature of the probability weighting function. *Management Science*, *42*, 1676–1690.