



HAL
open science

The Acceptability of Accountability

John Bone, Paolo Crosetto, John Hey, Carmen Pasca

► **To cite this version:**

John Bone, Paolo Crosetto, John Hey, Carmen Pasca. The Acceptability of Accountability. Constitutional Political Economy, 2021, 32 (4), pp.476-501. 10.1007/s10602-021-09331-z . hal-03720678

HAL Id: hal-03720678

<https://hal.inrae.fr/hal-03720678v1>

Submitted on 11 Oct 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



The Acceptability of Accountability

John Bone¹ · Paolo Crosetto² · John Hey¹ · Carmen Pasca³

Accepted: 19 April 2021 / Published online: 12 May 2021
© The Author(s) 2021

Abstract

This paper reports on an experimental test of the acceptability of the Principle of Accountability. This is a principle of social justice, and states, “individuals should be rewarded for factors under their control [...], but not for factors outside their control” (Cappelen and Tungodden (2009)). We specifically ask for acceptability of the *principle* underlying it, rather than for particular rewards in particular instances. We carry out the test with both an Internal and an External Dictator, conducting a laboratory experiment with a total of 240 subjects. We find that there is broad, but not overwhelming support for the Principle. When the Principle is internally inconsistent no clear preference emerges, which is not surprising.

Keywords Chance · Choice · Compensation · Luck · The principle of accountability · Responsibility

JEL Classifications D31 · D63 · C91

We are grateful to Noemi Pace and Bertil Tungodden for comments on this paper, and are particularly grateful to an anonymous referee, whose comments and references to the literature have significantly improved the paper.

More detail can be found in the Appendix (which is not intended for publication). This and other material are available through the paper’s website.

✉ John Bone
john.bone@york.ac.uk

Paolo Crosetto
paolo.crosetto@inrae.fr

John Hey
john.hey@york.ac.uk

Carmen Pasca
carmen.mariana.pasca@gmail.com

¹ University of York, York, North Yorkshire, UK

² INRAE, the French National Research Institute for Agriculture, Food and Environment, Grenoble, France

³ CMH - International Hospitality Management School, Paris, France

1 Introduction

This paper considers *both* luck *and* effort, and their incorporation into principles of social justice. In particular, we investigate the support for the idea that Society should compensate an individual for low income due to bad luck, but not for low income due to lack of effort. This idea is inspired by the Principle of Accountability.¹

The following simple model provides a framework for articulating these ideas. It assumes that an individual's exogenously and randomly determined luck is either *bad* or *good*, and that the chosen effort level is either *none* or *some*. For all individuals, the raw, pre-redistribution *income* is a function of effort and luck. A numerical example of a possible situation before Society has imposed redistribution is shown in Table 1.

People with bad luck have lower incomes than those with good luck (given any effort level), and those who make no effort have lower incomes than those who make some effort (given any luck). The question we are investigating in this paper is whether Society should compensate people for bad luck and/or for lack of effort. Society can do this through the distribution of some social dividend.

Different liberal approaches to social redistribution,² such as the Principle of Accountability, or liberal egalitarianism (Cappelen & Tungodden, 2006) or luck egalitarianism (Cohen, 1989), or, more loosely speaking, meritocracy, appear to provide a solution to the choice of the distribution rules that Society should apply: *respecting the effort decisions of individuals, yet treating equally individuals who are subject to different levels of luck*. This implies compensating individuals for bad luck, but not compensating them for choosing not to provide effort. Unfortunately, this Principle of Accountability is not generally implementable; as demonstrated in a different context by Cappelen and Tungodden (2006), building on Bossert and Fleurbaey (1996), there are several situations in which this simple meritocratic principle cannot be implemented, making compromises necessary.

This paper reports on an experiment eliciting preferences for redistribution principles related to the Principle of Accountability, and the compromises that subjects prefer to make when the Principle cannot be fully implemented.

The focus of this paper is on the elicitation of preferences on social compensation principles; our aim being to assess the extent of the support for the Principle of Accountability. Some of the experimental literature *indirectly* infers subjects'

¹ Others use the terms *Meritocracy*, *Liberal Egalitarianism* or *Luck Egalitarianism*. Cohen (1989) is seminal, although drawing critically on Dworkin (1981). Arneson has been prominent in the elaboration and development of this general principle of Liberal Egalitarianism, such as in Arneson (2000) which defends a prioritarian version. Our interpretation of the expression 'The Principle of Accountability' is similar to that of Cappelen and Tungodden (2006), who describe it as a 'compromise' between Strict Egalitarianism: ("For any distribution of talent and effort, the redistribution mechanism assigns the same post-tax income to everyone") and Libertarianism ("For any distribution of talent and effort, the redistribution mechanism assigns to each person a post-tax income equal to her pre-tax income"). Here they are using 'talent' in the same sense as we are using 'Luck'.

² We should note that different terminologies usually reflect differences in nuances, rather than differences in substance.

preferences for redistribution from subjects' choices. While that is a valid experimental methodology, we think that *directly* eliciting preferences over principles yields more easily generalisable results, and does not require the use of econometric techniques to identify the preferences from the noisy choices. Moreover, since the attractive Principle of Accountability is not in general applicable, in a choice-based design the experimenter has to choose between avoiding conflicting situations, and hence being unable to observe which compromises subjects are willing to make in these situations, or expose the subjects to those situations, and hence adding substantial amounts of noise to the choice data. We hence develop an experiment to elicit *directly* preferences over principles, and *directly* observe which compromises subjects are prepared to make in the face of impossibilities.

Given its intuitive appeal, we expected many of our subjects to agree with the Principle of Accountability. Our design allows us to produce exploratory results on the compromises that subjects are willing to make when faced with the inapplicability of the Principle to specific situations.

The paper is organised as follows. In Sect. 2, we explore the implication of compensation rules, focusing in particular on the non-generalisability of the Principle of Accountability solution. Related literature is discussed in Sect. 3. The experiment that we conducted is described in Sect. 4. Section 5 presents our results, and Sect. 6 concludes.

2 Compensating for income differences

We consider the following framework. Assume there is a social fund, of value X , to be divided between the individuals in Society. Each individual's *dividend* (D) from this fund is a supplement to (if positive) or a diminution of (if negative) their income (Y), giving their final total *payment* ($P = Y + D$). Therefore, these dividends provide a means of compensating for income differences such as those in Table 1. The following examples all assume a Society of four individuals, as in our experiment. Corresponding examples may easily be constructed for Societies of any size.

Dividends can be used to achieve different social redistribution goals.

One such goal is to *compensate fully for all income differences*, be they due to luck or effort, thus achieving (*final payments equal for all*). Two examples of how this principle plays out in different configuration of incomes, luck and effort choices in the Society are demonstrated in Table 2. Table 2A shows how full compensation can be reached with a social fund of 40 ($X = 40$). The result is Equal Payments of 15 (the total income divided by 4). Table 2B illustrates, by comparison with Table 2A, that the dividends and payments entailed by *equal payments for all* depend on, *inter alia*, the number of individuals in each cell of the matrix.

The diametrically opposite principle is *no compensation* at all for any income differences; this can be characterised as *equal dividends for all*, as can be seen from the example in Table 3A.

The Principle of Accountability, which suggests compensating only for those income differences due to luck can then be characterised as a hybrid principle (a 'compromise' to quote Cappelen & Tungodden, 2006), comprising (PA) *Equal*

Table 1 Pre-distribution incomes in four different states of the world

		Effort	
		None	Some
Luck	Bad	0	6
	Good	4	10

Payments for Equal Effort, and (PB) *Equal Dividends for Equal Luck*. Table 3B illustrates. At each level of luck, the income differential due to effort, here 6, is preserved in the payment differential. All ex-post income differences due to luck have been compensated, while ex-post inequalities due to effort have not. Table 3C again illustrates that the dividends and payments implied by (PA) and (PB) depend on the distribution of individuals across the four income positions.

Given its intuitive appeal, we expect most of our subjects to agree with the idea of Principle of Accountability. In this case, we expect them to prefer *Equal Payments for Equal Effort* and *Equal Dividends for Equal Luck*, if there is no conflict between these two principles.

Unfortunately, the Principle of Accountability is not always implementable in practice, as Fleurbaey (2008) and many others have pointed out. In Tables 1, 2, 3, the income differential due to effort is assumed to be the same, here 6, at each level of luck. Without this assumption, however, (PA) and (PB) are generally incompatible. This is illustrated in Table 4, where the income at good luck and low effort is 8, instead of 4 as previously assumed.

In Table 4A, dividends are equal at each level of luck. Given this, payments can be equal at only one level of effort. Thus equal payments for those choosing low effort, as shown here, entail unequal payments for those choosing high effort.

In Table 4B, payments are equal at each level of effort. Given this, dividends can be equal at only one level of luck. Equal dividends for those with bad luck, as shown here, entail unequal dividends for those with good luck, and thus a payment differential greater (in this case) than the income differential.

Generally, given such income values, (PA) and (PB) can be satisfied for any three of the four rows and columns, but not for all four.³ An advocate of (PA) and (PB) needs to decide which of these should have priority in the event that they are incompatible, as in Tables 4A, B. Note, however, that this assumes that all four income positions are occupied. As Table 4C illustrates, with (at least) one position unoccupied it is, after all, possible to satisfy (PA) and (PB) for all individuals in this Society.

A formal proof of the lack of generality of the Principle of Accountability is given in Appendix 2.⁴

³ This is an instance of a variety of related impossibilities investigated by Fleurbaey (1994, 1995a, 1995b), Bossert (1995), Bossert and Fleurbaey (1996), and Cappelen and Tungodden (2006), traceable to a problem initially identified by Pazner and Schmeidler (1974).

⁴ We expand on the problems of applying the Principle of Accountability in general, and how one can assess the hybrid compromises that arise in different situations, in Appendix 3.

Table 2 Examples of full compensation for all income differences (Equal Payments)

		<i>Effort</i>			
		A		B	
<i>X=40</i>		<i>Low</i>	<i>high</i>	<i>low</i>	<i>HIGH</i>
<i>Bad</i>		<i>n=1</i>	<i>n=1</i>	<i>n=1</i>	<i>n=0</i>
		<i>Y 0</i>	<i>Y 6</i>	<i>Y 0</i>	
		<i>D 15</i>	<i>D 9</i>	<i>D 16</i>	
		<i>P 15</i>	<i>P 15</i>	<i>P 16</i>	
<i>Good</i>		<i>n=1</i>	<i>n=1</i>	<i>n=1</i>	<i>n=2</i>
		<i>Y 4</i>	<i>Y 10</i>	<i>Y 4</i>	<i>Y 10</i>
		<i>D 11</i>	<i>D 5</i>	<i>D 12</i>	<i>D 6</i>
		<i>P 15</i>	<i>P 15</i>	<i>P 16</i>	<i>P 16</i>
		A		B	

Table 3 Examples of no compensation (A) and compensation using the Principle of Accountability (B and C)

			<i>Effort</i>					
			A		B		C	
<i>X=40</i>			<i>Low</i>	<i>High</i>	<i>Low</i>	<i>High</i>	<i>Low</i>	<i>High</i>
<i>Luck</i>	<i>Bad</i>		<i>n=1</i>	<i>n=1</i>	<i>n=1</i>	<i>n=1</i>	<i>n=1</i>	<i>n=2</i>
			<i>Y 0</i>	<i>Y 6</i>	<i>Y 0</i>	<i>Y 6</i>	<i>Y 0</i>	<i>Y 6</i>
			<i>D 10</i>	<i>D 10</i>	<i>D 12</i>	<i>D 12</i>	<i>D 11</i>	<i>D 11</i>
			<i>P 10</i>	<i>P 16</i>	<i>P 12</i>	<i>P 18</i>	<i>P 11</i>	<i>P 17</i>
<i>Good</i>		<i>n=1</i>	<i>n=1</i>	<i>n=1</i>	<i>n=1</i>	<i>n=0</i>	<i>n=1</i>	
		<i>Y 4</i>	<i>Y 10</i>	<i>Y 4</i>	<i>Y 10</i>		<i>Y 10</i>	
		<i>D 10</i>	<i>D 10</i>	<i>D 8</i>	<i>D 8</i>		<i>D 7</i>	
		<i>P 14</i>	<i>P 20</i>	<i>P 12</i>	<i>P 18</i>		<i>P 17</i>	

Our design allows us to produce exploratory results on the choices of subjects when there is a conflict between the preferred principles (PA) and (PB).

3 Related literature

Our experiment was designed to assess the support for the Principle of Accountability. As much of the literature makes clear, we need to specify who is expressing support. Is it a member of the Society, or an interested outside observer? Many experiments have two treatments to capture these two points of view: an *Internal Dictator* treatment and an *External Dictator* treatment. In the former, we need also to specify what the Internal Dictator knows about his⁵ position in Society. The specification could vary from trying to reproduce Rawls’ Veil of Ignorance (Rawls, 1971), where the Internal Dictator does not know his future position in Society,⁶ to a situation

⁵ For ‘his’ read ‘his or hers’ throughout, and similarly for ‘he’.

⁶ This is what we do, following Rawls as much as possible.

Table 4 Examples of inapplicability of the Principle of Accountability

		<i>effort</i>					
<i>X = 40</i>		<i>low</i>	<i>high</i>	<i>Low</i>	<i>High</i>	<i>low</i>	<i>High</i>
<i>luck</i>	<i>bad</i>	<i>n = 1</i>	<i>n = 1</i>	<i>n = 1</i>	<i>n = 1</i>	<i>n = 0</i>	<i>n = 1</i>
		<i>Y 0</i>	<i>Y 6</i>	<i>Y 0</i>	<i>Y 6</i>	<i>Y 6</i>	<i>Y 6</i>
		<i>D 14</i>	<i>D 14</i>	<i>D 13</i>	<i>D 13</i>	<i>D 13</i>	<i>D 13</i>
		<i>P 14</i>	<i>P 20</i>	<i>P 13</i>	<i>P 19</i>	<i>P 19</i>	<i>P 19</i>
	<i>good</i>	<i>n = 1</i>	<i>n = 1</i>	<i>n = 1</i>	<i>n = 1</i>	<i>n = 2</i>	<i>n = 1</i>
		<i>Y 8</i>	<i>Y 10</i>	<i>Y 8</i>	<i>Y 10</i>	<i>Y 8</i>	<i>Y 10</i>
		<i>D 6</i>	<i>D 6</i>	<i>D 5</i>	<i>D 9</i>	<i>D 9</i>	<i>D 9</i>
		<i>P 14</i>	<i>P 16</i>	<i>P 13</i>	<i>P 19</i>	<i>P 17</i>	<i>P 19</i>
		<i>A</i>		<i>B</i>		<i>C</i>	
		PA violated		PA fulfilled		PA fulfilled	
		PB fulfilled		PB violated		PB fulfilled	

where the Internal Dictator *does* know his position.⁷ In this latter case, the experiment essentially determines the ‘other-regarding preferences’ of the subjects.

Our experiment considers a case in which the income of an individual is a function of a choice (endogenous) variable and an exogenous variable, but we start our discussion here with the simpler one-dimensional case. Many of the previous experimental contributions to the literature have considered just that—dividend rules in one dimension – most commonly with exogenous risky income. Experiments using this latter approach, sometimes with both an Internal and External Dictator treatment, include Bukszar and Knetsch (1997), Cabrales et al.,(2010), Frignani and Ponti (2012), Herne and Suojanen (2004), Johannesson and Gerdtham (1995), Schildberg-Horisch (2010) and Sutter and Weck-Hanneman (2003). Some of these studies use a Dictator game from the game-theoretic literature, but all of them get one of the subjects, either pre-specified, or chosen randomly after all have expressed their preferences, to play the Dictator.

A recent strand of the literature has taken the debate into two dimensions: the two dimensions that we are considering—one exogenous and the other endogenous to the individual. Prominent in this literature are Cappelen et al.(2007, 2010 and 2013) and Konow (2000). Once again, they use a one-shot Dictator game, and the two dimensions are the amount of work that the subjects do, and the rate of pay they get for the work (and, in Cappelen et al.(2010), their productivity). The rate of pay is exogenous while the amount of work is endogenous. We adopt a similar approach (though we omit Cappelen et al.(2010)’s productivity). However, our study differs from theirs in that we ask our subjects to state general *principles* of distribution relevant to the Principle of Accountancy, while they ask either for different principles or for preferred distributions in particular contexts and try to infer the underlying general

⁷ See Voigt (2013) for a discussion of ways of implementing the Veil of Ignorance.

principles from these particular cases.⁸ In addition, our study is different from theirs in the way we try to create the Veil of Ignorance with an Internal Dictator.

There are also important experimental papers by Mollerstrom and associates (Mollerstrom et al., 2015 and Cappelen et al., 2019), which have some features in common with ours. Mollerstrom et al. (2015) focus on two kinds of luck: option luck and brute luck. Brute luck is outside the control of the individual while option luck is partially controllable through an insurance decision that, in terms of individual responsibility, has a similar role to effort in our experiment. For each subject-pair, income inequality can result from varying combinations of brute and option luck for the two subjects individually. An external dictator has to decide *ex ante*, for each of those different possibilities, whether or not incomes are to be equalised between that subject-pair. Cappelen et al. (2019) extends the setup to one in which the external dictator is unable to tell with certainty the role of brute luck versus option luck; interestingly, subjects were significantly more egalitarian in this version. In their requirement of *ex ante* decision-making, covering exhaustively all possible unequal situations, this approach comes closer to ours than do many experiments in this field. Even so, in this setting preferences over distributional principles can be only inferred statistically, rather than directly elicited. Also, as already discussed, a primary reason for our interest in explicit preferences over principles lies in the possibility of such principles being incompatible.

A recent experiment by Becker (2013) focuses specifically on the luck/effort trade off, though her prime interest is in the effect of the *type* of luck on people's principles (she considers wage luck, endowment luck and payment luck – essentially framing differences). She concludes that “Our results suggest that normative beliefs about which share should be taken are driven mostly by the factor that subjects' can be held responsible for, namely effort, whereas actual behavior sometimes deviates from this norm dependent on which type of luck is introduced.” The frame is important, and part of the Principle of Accountability is supported.

Bjerk (2016) also ran an experiment with two dimensions – effectively luck and effort, and considered the effect of the type of Dictator on stated preferences. He was interested in the re-distribution rules that subjects favoured. Each player had to propose a “redistribution rule” which was specified by p , a number between zero and one, “where p represents what fraction of each participant's task earnings will be taken from each participant and split equally across all participants”. Clearly if p is zero, then no re-distribution occurs; this equivalent in our context, to Equal Dividends for all. If p is one, this is equivalent to our Equal Payments for all. However, there is no value of p for which we get (PA) Equal Dividends for Equal Effort and (PB) Equal Payment for Equal Luck. Hence, his principles do not correspond to ours. Deffains et al. (2016), Espinosa et al. (2020) and Gerber et al. (2014) also investigate the Bjerk (2016) re-distribution principles. Once again, we are investigating different principles of social justice. One advantage of their re-distribution

⁸ This methodology might be sensitive to the adoption of particular functional forms (for the social preference functional) and to the particular problems posed to the subjects.

schemes is that they are always implementable, unlike ours, where compromise might be necessary.

A seminal paper is that of Konow (2000). He “differentiates *discretionary* from *exogenous* variables. A discretionary variable is one that affects production and that the individual can influence, which in this experiment is the subject’s input or letters produced. An exogenous variable is one that the person cannot reasonably influence but that may have an impact on output.” These are equivalent to our effort and luck respectively.

4 The experimental design

One feature crucially distinguishes our experiment from some of the previous related experiments: we did not ask subjects to take decisions for *particular ex ante* income distributions; instead, we asked them to select a *compensation principle* that could be applied to *any ex ante* income distribution. We realise that this essentially forces subjects to have a consistent compensation principle; the contrary procedure, of trying to infer an underlying principle from decisions for particular distributions, might well lead to imprecise or undefined inferred underlying principles, especially for the cases in which the principles cannot be applied.

This methodology of eliciting principles has been used elsewhere, most notably in game theory experiments, where principles are called strategies, and where the statement of a strategy is simple. In other contexts, where decisions are dependent on many variables, principles/strategies are usually not elicited because of the complexity of stating them. In our experiment, principles can be simply stated: either *Equal Payments* or *Equal Dividends*. The difficulty for subjects in our experiment was not in understanding what they are being asked to do, but rather in appreciating the *consequences* of their chosen principle.

Our experiment had two *dimensions*, luck and effort, each of them having just two values: luck, good or bad; effort, some or none. Thus, there were four possibilities: (good luck, some effort), (bad luck, some effort), (good luck, no effort) and (bad luck, no effort). As before, we represent these possibilities with the cells of a two-by-two matrix in which the rows are bad luck and good luck and the columns are no effort and some effort, and the cell entries are the *ex-ante* income levels. These cell entries were *ex ante* random (between given and known lower and upper bounds). There was a social dividend of a fixed amount that could be distributed by a rule that would be determined by the stated principles of one of the subjects over these four cells.

We had two different treatments: (1) an *Internal Dictator* treatment in which the rule would be that of a randomly-selected member of Society after preferences were elicited; (2) an *External Dictator* treatment in which groups of five subjects expressed their preferences, after which one of them was randomly chosen to be outside the Society consisting of the other four, and his principles imposed on the other four, while the chosen dictator would get a flat payment. The External Dictator treatment captures the idea of an external person who is not part of Society but has preferences about its organisation. The Internal Dictator treatment was designed to

reproduce as far as possible Rawls' Veil of Ignorance, in which subjects do not know *ex ante* their position in Society: as far as luck is concerned, it is clear that subjects did not know *ex ante* whether they would have good luck or bad; as far as effort is concerned, we tried to make their future decision (as to whether to make an effort or not) as uncertain as possible, by making the initial incomes uncertain.⁹

Subjects were asked to state their preferred *compensation rules/principles*. We deliberately restricted the set of rules from which they could choose. To be specific, we asked them to state, separately, for each row and for each column of the matrix, whether they preferred Equal Dividends (no compensation) or Equal Payments (full compensation) in that row or column.

Unfortunately, as we have already noted, some combinations of principles are incoherent, and cannot be jointly implemented. To accommodate this, if a subject wished to state mutually-exclusive principles,¹⁰ then he was also asked to state which one of these conditions was droppable,¹¹ if not all conditions could be simultaneously implemented.

Subjects were told that, after all had stated their preferred compensation rules, each of them would be randomly assigned bad or good luck (if appropriate), and that each of them would be asked (if appropriate) whether they wished to 'make an effort' or not. 'Making an effort' was implemented by them staying in the laboratory for an extra 20 min and completing a simple questionnaire; this we call Part 2. 'Leave' and 'Stay' below refer to their decision, and correspond to the 'no effort' and 'some effort' of the theory.

We realised that stating their preferred rules for the full matrix might be a difficult task. Therefore, we simplified the problem for the subjects by building up to the full matrix in stages, with four simple scenarios, in which subjects were faced with just a row or just a column of the matrix:

Scenario 1a: All members Leave after Part 1; members individually have Bad or Good luck;

Scenario 1b: All members Stay for Part 2; members individually have Bad or Good luck;

Scenario 2a: All members have Bad luck; members individually choose to Leave or Stay;

Scenario 2b: All members have Good luck; members individually choose to Leave or Stay;

Scenario 3: Full matrix.

We asked them in each scenario whether they preferred Equal Dividends (no compensation) or Equal Payments (full compensation). In all the four simple

⁹ We appreciate that we cannot exclude the possibility that a subject took a decision to (not) provide effort before their incomes were known, and perhaps were innately optimistic (pessimistic) and assumed that they would have good (bad) luck.

¹⁰ For example, "Equal Payments in both columns and Equal Dividends in both rows".

¹¹ By which is meant that, in the case of conflict, it could be dropped, that is, not implemented.

scenarios, the decisions, and their implications, were clear and simple, since just one dimension was varied at each time, the other being fixed. The scenarios were presented to subjects in the fixed order 1a, 1b, 2a, 2b and 3. While this might have created order effects, we chose the fixed order as we needed the full matrix to come at the end, and we saw no particular reason for order effects to have an effect on the elicitation. We deliberately included the simple scenarios to assess support for the two elements of the Principle of Accountability individually.

Subjects knew that, at the end of this part of the experiment, one of the above five scenarios would be selected at random¹² and played out; subjects would be paid according to the dictator's choices, their individual luck and effort level, and the *ex-ante* payments that would be determined randomly. The payout differed according to scenario. In Scenario 1b, for example, all members would (with common knowledge) be required to Stay for Part 2 and the payout would simply assign each member either Bad or Good luck, followed by the application of the compensation rules, that is, either Equal Payments or Equal Dividends, that had been chosen by the Dictator for Scenario 1b. In Scenario 2a, by contrast, all members would (with common knowledge) be assigned Bad luck, and each would then decide whether or not to Stay for Part 2, again followed by the application of the compensation rules that had been chosen for Scenario 2a. The payout in Scenario 3 was as already described: independently and privately, each member was assigned either Bad luck or Good, and then decided whether to Stay for Part 2.

In terms of practical detail, the experimental sessions were organised as follows: preceding the decision stage (Stage 1) was a (silent) PowerPoint show explaining the task, auto-timed and running simultaneously on each subject's computer screen, and lasting 15 min. Each subject also had a hardcopy complete handout of that show, to which they could refer at any time during Part 1.¹³ They could ask questions at any stage; few did.

Stage 1—the essence of the experiment—took 20 min. Subjects considered the five scenarios in sequence, starting with Scenario 1a and ending with Scenario 3. In each case, our experimental software¹⁴ allowed each subject, before making a choice, to explore, through simulation, the effects of different compensation rules, against different hypothetical positions of the different members of society and intermediate income values, all controllable by the subject. Appendix 1 contains screenshots, for Scenarios 1a and 3. For each elementary scenario, subjects could explore the effects of their choices for two minutes, after which they had 30 s to submit their choices. Exploration could continue up to the moment of submission. For the more complex full matrix scenario (Scenario 3), subjects had eight minutes of simulation followed by a two-minute period for submission (during which exploration could continue). The countdown clock can be seen in the screenshots, above the large 'Submit your

¹² The probability was 50% that it would be the full scenario and 12.5% that it would be one of the four elementary scenarios.

¹³ This can be found in the Appendix (not intended for publication) and will be made available on the paper's website in due course.

¹⁴ This was written in Python and is available on the paper's website.

choice' button that was greyed out until the submission period. The exploration phase was meant to allow subjects to understand fully the implications of their decisions for a variety of different situations.

For the full scenario, if at any time in the simulation the subject had a mixed compensation rule (a combination of Equal Payments somewhere and Equal Dividends elsewhere) and a saturated profile (that is, all four cells of the matrix were filled), then the program prompted him to select one of the four elements (in the graphic, one of the two rows or one of the two columns) as *droppable*, before he could continue. This meant that in case of inapplicability of the chosen rules, the rules would not be applied for that particular row or column. It then revealed the hypothetical implications of this for the dividends and payments, as for example in the scenario 3 screenshot in Appendix 1, where the droppable element is identified by the check box. A similar prompt appeared if he tried to submit as his choice a mixed compensation rule without having selected a droppable element.

In the event of non-submission within the allotted time, (known) default rules were in place. Within each Society, the same default rule (either Equal Dividends or Equal Payments) applied to all subjects in their choices for all four elements, and in all five scenarios.¹⁵ To control for default-bias, we gave half of our societies a default of Equal Dividends, and the other half Equal Payments. This produced a 2 by 2 treatment-structure as set out in Table 5, which also records the number of subjects in each treatment.

Once the elicitation stage was completed, one member of the Society was randomly chosen as a dictator, and then all members were anonymously informed of who was the dictator, which scenario was going to be payoff-relevant, the ex-ante income values, and the rules chosen by the Dictator. This was followed by the actual payout of the drawn scenario. The final payout used a software interface similar to that in Stage 1. First, the software generated random incomes in each of the four cells. Then, in Scenarios 1a and 1b, each subject individually clicked a button to reveal his luck, while in Scenarios 2a and 2b each subject individually registered his effort decision by clicking one of two buttons, labelled Leave and Stay respectively. In Scenario 3, each of these events occurred in turn. In all cases, the onscreen display then revealed the positions of each of the four members of that Society, and the consequent incomes, dividends and payments.

Therefore, the overall sequence of events in Part 1 was as follows:

Stage 1 Each member chooses compensation rules for each of the five possible scenarios;

¹⁵ The default rule had an additional role in the compound scenario. Each of the four elements constrains dividends either along a row or down a column of Table 7. In the event that, after the payout, all members of the society were located on one or other of the two diagonals, none of these constraints would be binding, leaving dividends undetermined. After considering several contingencies for this event, including randomisation, we decided that the given default rule should apply. This had an obvious downside, however. For example, a subject was unable to impose Equal Payments, on such a profile, given a default of Equal Dividends.

Stage 2 One member is randomly selected as Dictator (and in the External Dictator treatments takes no further part);

Stage 3 Each member is informed:

- which (numbered¹⁶) member is selected as Dictator for his Society;
- which of the five scenarios applies to his Society;
- the income values relevant to that scenario;
- the compensation rule, as chosen by the Dictator;

Stage 4 The scenario is then played out, determining for each subject whether he has Bad or Good luck, whether or not he Stays for Part 2, and also his final payment.

Part 1 was then complete. Those subjects staying for Part 2 remained in their places, and completed the simple task. The others left, collecting payment on the way out.

There was a total of 240 subjects; subjects were predominately students at the University of York; they spent on average 25 min in the laboratory if choosing not to stay for part 2, and 45 min if they stayed; since about half subjects chose or were forced to stay, on average subjects spent about 35 min in the laboratory, and earned on average £17.42.

The essence of the design was producing an experimental setup in which subjects could express their preferences for (PA) *Equal Payments for Equal Effort* and (PB) *Equal Dividends for Equal Luck* in various contexts: subjects were asked to state their *preferred principle* in different contexts. The ‘full context’ is where both Luck and Effort play role – this we call Scenario 3. As this is quite a complicated context, we built up to it via a series of simplified contexts, in each of which there was *either* Luck *or* Effort, but not both (these are Scenarios 1a, 1b, 2a and 2b).¹⁷ To shed light on whether the expressed principles differ depending upon whether they are those of a disinterested spectator or an interested member of Society, we had two treatments – which we call the External and Internal Dictator treatments. Whether choices in the two treatments might differ depends upon how the subjects viewed the two treatments, and whether they regarded them to be essentially the same, as far as expressing their preferences was concerned.

We conclude this section with a brief illustration of the experimental software.

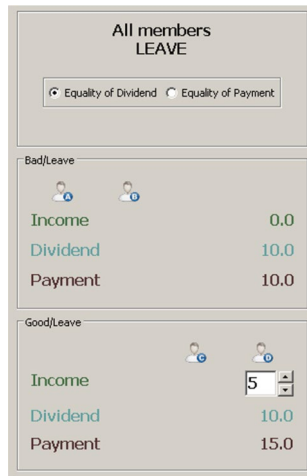
For scenario 1a the interface was as follows:

¹⁶ Numbers were used to preserve anonymity. No subjects knew at any stage the identity of the other members of his Society.

¹⁷ We note that the decision whether to provide effort or not varied from scenario to scenario, but it was clear in each scenario what was exogenous and what was not; so subjects were expressing their preferences given any exogenous conditions. As such, their stated preferences are valid.

Table 5 The 2×2 treatment structure, with subject numbers

		Default dividend rule	
		Equal dividends	Equal payments
Dictator type	Internal	<i>Int-Div</i> (n = 60)	<i>Int-Pay</i> (n = 60)
	External	<i>Ext-Div</i> (n = 59)	<i>Ext-Pay</i> (n = 60)



The only decision that the subject was asked to take was whether he/she thought that one should apply ‘Equality of Dividend’ or ‘Equality of Payment’. As will be seen, in the figure above the former is ticked, implying that the effect of bad luck (a loss of income of 5 ECU) is not compensated and so the payments differ by the same amount as the income. Note, as should be clear from the figure that 2 of the 4 members have bad luck and 2 have good. The subjects were encouraged to vary the Good/Leave income by using the drop down box, and hence explore the implications. In this case, changing the Good/Leave income would change everyone’s payment but not the gap between the good and bad luck subjects. The interface also allowed subjects to experiment with changing the number of subjects in each box. In this case, keeping the Good/Leave income fixed at 5, but having 1 subject in bad leave and 3 in good/leave would increase the income and payment of the one subject moved to good/leave, but, with ‘Equality of Dividend’ ticked, not change the gaps between income and payment. However, if ‘Equality of Dividend’ was ticked, then all would receive a payment of 55/4 instead of 50/4.

5 Results

For each of the four treatments we had 60 subjects in the laboratory. However, in the (External Dictator, Equal Dividends default) treatment there was a computer problem for one subject, which led to the loss of that subject's data but not the subject's participation. We did not collect demographics.

Table 5 shows that we had four treatments in a 2 by 2 structure: Internal/External Dictator and Equal Payments/Equal Dividends Default Rule. The data allows us to rule out any effect of the Default Rule effect—see Appendix 4 for details. Hence, in the remainder of the paper, we pool the data of the two default rules, and consider only the two treatments: *Internal* and *External Dictator*.

5.1 The four elementary scenarios

Table 6 records for each elementary scenario the number of subjects choosing Equal Payments for All.

Equal Payments is the Principle of Accountability (PoA) rule for Scenarios 1a and 1b, while Equal Dividends is for 2a and 2b. Therefore, in each case a large majority of subjects chose the Principle of Accountability rule. For Scenario 1a (all members leave) the proportion choosing Equal Payments is significantly (χ^2 p -value = 0.021) greater with an Internal than with an External Dictator. Otherwise, the differences between the two treatments are not significant.

Table 7 shows the results when aggregating choices over the four simple scenarios and in the full matrix compound scenario. In a perfect world, both columns should show the same preferences; in the case of simple scenarios, this preference has been elicited piecewise, one element at a time; in the compound scenario, it has been elicited jointly. A crucial difference between the two is that in the simple scenarios subjects could not be and were not exposed to the impossibility of some of their choices; in the compound scenario, they were. We discuss here only the simple scenarios.

The Principle of Accountability rule implies Equal Payments in Scenarios 1a and 1b, and Equal Dividends in Scenarios 2a and 2b. We could define a *Somewhat Principle of Accountability* rule ('Somewhat PoA') as differing from this in exactly one element. A still broader class of what might be called *orthodox* responses comprises these together with responses with either Equal Dividends for each level of luck, or Equal Payments for each level of effort. As can be seen from Table 7, these five categories constitute some 63% to 70% of all responses in the simple scenarios – a significant majority.

Figure 1 shows consistency of choice between scenarios 1a and 1b, and/or between 2a and 2b. The majority of subjects (67%) showed consistency, with a relatively small fraction changing their reported preferences.

5.2 The compound scenario

Table 7 reports the results for the compound scenario. Recall that with the exception of Equal Dividends or Equal Payments everywhere, each regime in the compound scenario is contingent on at least one income position being unoccupied, the subject specifying which one of the four elements is droppable otherwise.

As with the elementary scenarios, orthodox responses (that is, non- ‘other’) account for a significant majority of subjects in each treatment. However, PoA responses are inevitably lower, not significantly so (χ^2 p -values 0.0076 and 0.0057 for the Internal and External treatments respectively), in the compound scenario than in the elementary scenarios, and are less attractive than each of Equal Dividends or Equal Payments everywhere. It looks as if the increased complexity of the full-matrix scenario made simpler rules (equal dividends everywhere, that is no compensations, or equal payments everywhere, that is full compensation) more attractive than the more difficult (and not universally applicable) PoA rule of compensating for luck but not for effort.

Figure 2 shows the consistency between the elementary and compound scenarios. One striking feature is that no subject switched from Equal Payments in the simple scenarios to Equal Dividends in the compound, and no subject switched from Equal Dividends in the simple scenarios to Equal Payments in the compound. This might reflect the wide perceived gap between left-wing (full compensation) vs right-wing (no compensation) principles. However, there is a considerable amount of switching elsewhere. Table 8 shows the percentage of subjects who are consistent with their choice of principle in the simple and compound scenarios. For example, 46.7% of all subjects in the Internal Dictator treatment preferred Equal Payments in both the simple and the compound scenarios.

Summarising, we find support (albeit limited) for the Principle of Accountability. Only about 20% of subjects support its pristine form – fully compensating for exogenous luck while not compensating for endogenous effort choices. This percentage decreases in the compound scenario, where subjects have to face the full complexity of the consequences of their choices. A further 20% support principles that can be described as ‘somewhat PoA’, i.e. that keep the PoA core idea but do not apply it in specific situations (e.g., if all subjects have bad luck, or if all subjects exert no effort). About 15% of subjects show right-wing libertarian views, choosing support for no compensation for either luck no effort, and a symmetric 15% show left-wing views, choosing full compensation for both. About a third of the subjects do not express any coherent principle.

5.3 Compromises in the face of the inapplicability of the chosen principles

A secondary purpose of the experiment was to see how those subjects attracted to the Principle of Accountability would respond to its infeasibility in a saturated profile, and in particular, which rules they would select as droppable in the event that they chose the (contingent) Principle of Accountability in the compound scenarios.

Table 6 Percentage of subjects choosing Equal Payments in the elementary scenarios

	Internal dictator	External dictator	All
Scenario 1a: all members leave	64.2	53.8	59.0
Scenario 1b: all members stay	63.3	62.2	62.8
Scenario 2a: all bad luck	36.7	38.7	37.6
Scenario 2b: all good luck	36.7	39.5	38.0

Table 7 Subjects' responses over all scenarios (rounded percentages)

Principle	Internal dictator		External dictator	
	Simple scenarios	Compound scenario	Simple scenarios	Compound scenario
PoA (Principle of Accountability)	21.7	12.5	16.8	8.4
Somewhat PoA	26.7	20.8	20.2	22.7
Equal payments	12.5	17.5	15.1	21.0
Equal dividends	15.8	13.3	12.6	12.6
Other	23.3	35.8	35.3	35.3

The number of subjects we can identify as being attracted to PoA, on the basis of choosing PoA or Somewhat PoA in the compound scenario is 37 in the Internal Dictator treatment, and 32 in the External Dictator Treatment. Table 9 sets out the data on droppable rules among these subjects.

There is a clear and significant difference between the Internal and External treatments for the droppable element. Interestingly, in the External Dictator treatment the responses were not significantly different from a uniform distribution (χ^2 p -value=0.267), while in the Internal Dictator treatment, they *were* significantly different from a uniform distribution (χ^2 p -value=0.010) with a strong preference for dropping Equal Payments for Leave and Equal Dividends for Bad Luck. The External Dictator results seem consistent with our principal hypotheses,¹⁸ while the Internal Dictator results suggest that many of the subjects were planning to Stay or were hoping for Good Luck (in other words, misinterpreting the implementation of the Internal Dictator treatment).

5.4 Extent of exploration of the consequences

As noted above, our experimental software allowed subjects to explore in detail the consequences of their choices. They could vary, for each choice of rules for the different rows and columns, the position of the players and the income

¹⁸ These are: (PA) *Equal Payments for Equal Effort*, and (PB) *Equal Dividends for Equal Luck*.

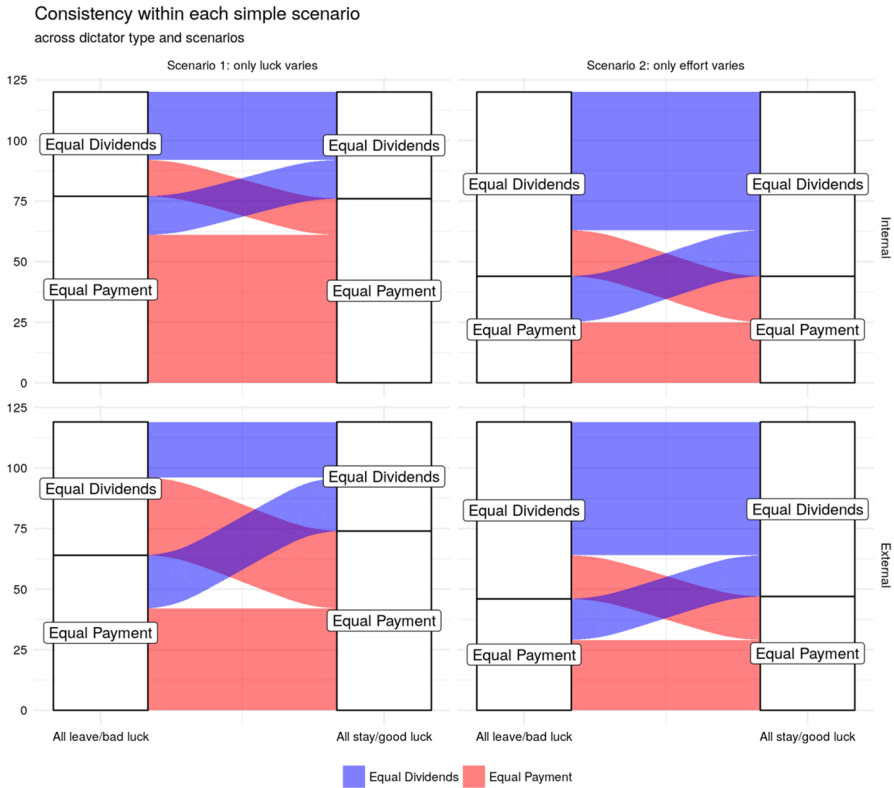


Fig. 1 Consistency within each simple scenario

occurred in different positions. In particular, they had four ways of interacting with the software: they could (a) change the allocation rules; (b) change the position of the players, (c) change the income attached to different choices and (d) use a button that would randomly change the position and the incomes. We tracked each action for both the four simple and the compound scenario. Table 10 details the data. (Table 10)

Subjects explored the consequences of their choices extensively. On average, subjects had between 35 and 45 interactions for each of the 2-min simple scenarios, and 85 interactions for the 8-min compound scenario. There was large heterogeneity, as testified by the large standard deviations, and especially so for the random action that shuffled all positions and incomes. This makes sense as the random exploration allows them to understand the implications only if repeated several times. The random button was the most heterogeneous in its use, with some subjects relying mostly on it and others not using it at all.

The extent of the overall exploration shows no correlation with the chosen principle, neither in the simple (kruskal-wallis, p -value=0.610) nor in the compound scenario (kruskal-wallis, p -value=0.671); the same is true when

Consistency across simple and complex scenarios

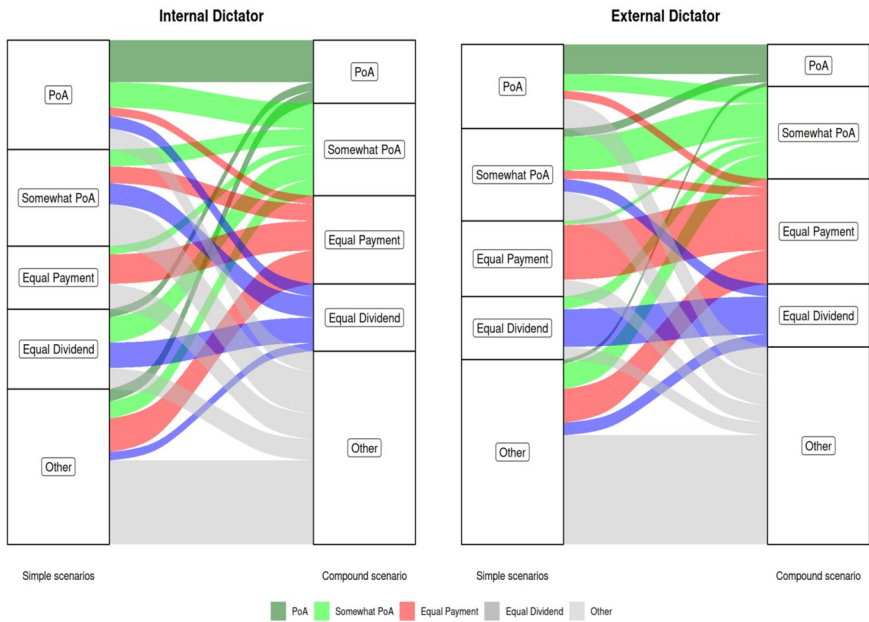


Fig. 2 Consistency across simple and complex scenarios

Table 8 Within-subject consistency across simple and compound scenarios, by principle (rounded percentages)

	Internal dictator	External dictator	All
PoA	38.5	35.0	37.0
Somewhat PoA	15.6	37.5	25.0
Equal Payments	46.7	72.2	60.6
Equal Dividends	31.6	60.0	44.1
Other	46.4	51.4	51.4

Table 9 Distribution of the “droppable” constraint chosen by subjects having chosen PoA or Somewhat PoA in the compound scenario (rounded percentage)

Dropped constraint	Internal dictator	External dictator
Equal dividends for bad luck	37.5	21.6
Equal dividends for good luck	12.5	37.8
Equal payments for leave	40.0	27.0
Equal payments for stay	10.0	13.5

disaggregating the exploration by type, both for the simple (kruskal-wallis, all p -values ≥ 0.142) and compound scenario (kruskal-wallis, all p -values ≥ 0.396). There is further no significant difference in exploration across dictator treatments (t test, p -value = 0.149). No particular axiom comes ‘intuitively’

Table 10 Mean (standard deviation) of each exploratory action, and overall, by scenario

	Income	Position	Random	Rule	Overall
Scenario 1a	14.67 (15.1)	17.4 (11.29)	16.96 (23.77)	8.92 (6.15)	45.86 (25.13)
Scenario 1b	15.5 (14.32)	13.75 (9.85)	21.18 (27.47)	6.62 (5.37)	39.63 (26.59)
Scenario 2a	12.5 (11.97)	13.09 (10.26)	23.51 (24.43)	5.96 (4.12)	35.85 (22.61)
Scenario 2b	14.15 (14.95)	13.42 (16.97)	23.88 (23.13)	6.44 (6.15)	36.83 (29.64)
Compound Scenario	23.72 (20.83)	20.16 (24.83)	47.16 (60.8)	25.15 (19.53)	85.94 (61.72)

and there is no special relation between the need to explore and the choice of more orthodox or appealing social redistribution axioms.

6 Discussion

In this experiment, we directly elicited individual preferences over compensation principles. Even if we think that this direct elicitation mechanism can help reveal social redistribution preferences better than the more popular method of inferring them from given choices, the design has limitations.

First, it can be a daunting task for the subjects. Our priority in designing the structure, the program/interface and the instructions was to make the experiment as simple and intelligible as possible. However, the task might have been challenging for subjects. They had to first understand the nature of the problem and then identify a best response to it. We feel that most of our subjects were able to clear the first of these hurdles. However, even given full comprehension of the problem, comparing regimes is a complex task, especially in the compound scenario, as our discussion above shows. Furthermore, optimal choices could vary substantially across subjects depending, for example on their self-interested motivation, attitude to risk, effort/payment preferences, and beliefs about the corresponding preferences of other members and about the unknown income values. For all of these reasons, the disparate results reported in Sect. 5 concerning the droppable principle were perhaps to be expected.

Subjects' overall responses are clearly not arbitrary, however, and cannot be said to be driven by miscomprehension of the task; we know from our background data recording every keystroke and mouse click, that most subjects, in all treatments, made full use of the simulation program to explore the implications of different regimes. We found that there was no correlation between the extent of the exploration and the submitted set of rules.

A majority of subjects submitted what we call orthodox choices, and most subjects are consistent when moving from the simple to the full scenario. Nonetheless, the design must have been difficult to navigate for some subjects since about a third of them did not submit orthodox principles. This large amount of noise can indeed be a consequence of our design. On the other hand, such a high level of noise would be a problem for *any* design. Ours has the advantage that noisy subjects signal themselves out through their choices, while traditional choice-based experimental approaches have no way of dropping the noisy subjects and this could bias their estimations.

It is difficult to compare our results with those in other papers, since they do not seek ‘orthodox responses’. There is, however, quite clearly much noise in subjects’ responses in other experiments, as shown in their results. For example, large variation can be seen in Becker’s Fig. 1; Konow’s Fig. 2 likewise shows considerable variation.

Second, we might have been forced to oversimplify the scenarios. We suspect that one difficulty was that the binary alternatives presented to the subjects (either Equal Payments or Equal Dividends – dictated to us by the Principle of Accountability) were too constraining, and that some subjects might have preferred partial compensation in certain contexts. In this respect, we should note that the evidence from other experiments (most notably those of Cappelen et al. (2007, 2010 and 2013)), indicates that subjects trade-off the various possibilities in a more continuous way. However, these other studies were not attempts to *test* the Principle of Accountability, nor the other principles discussed in this paper, but were more concerned with *estimating* utility functions – and usually in contexts in which *either* luck *or* effort were manipulated, and not *both*. Our work provides an important complement to theirs, both methodologically and theoretically. We provide a novel, incentivised design to directly elicit general principles, rather than inferring them from a limited set of choices; moreover, the uncertainty over payoffs allows us to have Internal Dictators that were unaware of their future position in Society – a more appropriate implementation of the Rawlsian Veil of Ignorance from a theoretical point of view¹⁹; and we directly test a situation in which both luck and effort vary, and put the subjects face to face with the impossibility of following the Principle of Accountability in a world with people in different positions.

As regards our principal interest, the Principle of Accountability receives modal support in the elementary scenarios. This is in line with other studies that have estimated social redistributive preferences, as for instance Almås et al. (2020). The distribution of preferred compensation strategies is also similar to theirs. With about 15% of the sample being strict Egalitarian (Equal Payments everywhere) and a further 15% Libertarian (Equal Dividends everywhere), our results sit in between the results that Almås et al. (2020) report for Norway and the USA. The fact that similar results stem from widely different methods (following common practice, their results are indirectly estimated from choices of external Dictators only) seems to be an indication of the robustness of the findings.

It is clear that the majority of our subjects agree that Society should compensate citizens for bad luck, and that it should *not* compensate subjects for lack of effort.

When these two principles conflict, subjects do not display a clear preference over second-best solutions, as choices do not reveal clear patterns for which constraint to be dropped in case of conflict. The practical impossibility of following the Principle of Accountability might be too difficult a setting to comprehend and act upon, or might be a scenario that our subjects – like most other people – never contemplated and are ill equipped to answer. While this last remark might seem discouraging, there is hope. New applications of our direct and incentivised elicitation

¹⁹ See Voigt (2013) for a discussion of ways of implementing the Veil of Ignorance.

method, coupled with simulation exercises, could be run on only those following the Principle of Accountability, stressing their preference’s lack of coherence in certain situations and asking which condition(s) should be dropped in these cases. Given the widespread acceptance of the Principle of Accountability in Western democracies, and the never-ending tensions that pervade them, we believe the results of such an experiment would be important.

The implications for constitutional economics are relatively clear: a significant majority of our subjects indicated approval of compensating citizens for bad luck, and a somewhat smaller majority approved *not* compensating citizens for lack of effort. These are in line with most Western practices and agree with survey evidence. The situation becomes less clear when luck and effort are combined; here there seems no consensus for a clear set of principles. It would seem that the preferences depend upon the context – and there was no context in our experiment. This suggests that we should run a similar experiment with the effort decision put into some kind of context.

Appendix 1: Screenshots from Stage 1

Scenario 3

For all members who LEAVE <input type="radio"/> Equality of Dividend <input checked="" type="radio"/> Equality of Payment <input type="checkbox"/> Droppable		For all members who STAY <input checked="" type="radio"/> Equality of Dividend <input type="radio"/> Equality of Payment <input checked="" type="checkbox"/> Droppable	
For all members who have BAD luck <input type="radio"/> Equality of Dividend <input type="checkbox"/> Droppable <input checked="" type="radio"/> Equality of Payment	Bad/Leave Income: 0.0 Dividend: 14.5 Payment: 14.5	Bad/Stay Income: 8 Dividend: 6.5 Payment: 14.5	
For all members who have GOOD luck <input checked="" type="radio"/> Equality of Dividend <input type="checkbox"/> Droppable <input type="radio"/> Equality of Payment	Good/Leave Income: 5 Dividend: 9.5 Payment: 14.5	Good/Stay Income: 10.0 Dividend: 9.5 Payment: 19.5	

Control panel

<input checked="" type="radio"/> Bad/Leave <input type="radio"/> Bad/Stay <input type="radio"/> Good/Leave <input type="radio"/> Good/Stay	<input type="radio"/> Bad/Leave <input checked="" type="radio"/> Bad/Stay <input type="radio"/> Good/Leave <input type="radio"/> Good/Stay	<input type="radio"/> Bad/Leave <input type="radio"/> Bad/Stay <input checked="" type="radio"/> Good/Leave <input type="radio"/> Good/Stay	<input type="radio"/> Bad/Leave <input type="radio"/> Bad/Stay <input type="radio"/> Good/Leave <input type="radio"/> Good/Stay	5:00 RePosition Submit your choice
---	---	---	--	--

Appendix 2: Showing that it is impossible to satisfy Income Compensation for Luck in general

It requires that the following five equations have a solution:

$Y_{BL} + D_{BL} = Y_{GL} + D_{GL}$ (If they Leave they get the same payment whether they have Bad or Good luck).

$Y_{BS} + D_{BS} = Y_{GS} + D_{GS}$ (If they Stay they get the same payment whether they have Bad or Good luck).

$D_{BL} = D_{BS}$ (If they have Bad luck they get the same dividend whether they Leave or Stay).

$D_{GL} = D_{GS}$ (If they have Good luck they get the same dividend whether they Leave or Stay).

$D_{BL} + D_{BS} + D_{GL} + D_{GS} = 40$ (the dividends add up to the total amount to distribute).

We have five equations in four unknowns which are mutually inconsistent unless it is the case that $Y_{GL} - Y_{BL} = Y_{GS} - Y_{BS}$ (which is possible but unlikely).

Appendix 3: Evaluating an incomplete Principle of Accountability

Our experiment was designed to assess the support for the Principle of Accountability. This implies Equal Payments at each level of effort, and Equal Dividends at each level of luck. However, the Principle of Accountability as we have already noted in Sect. 2, may be incoherent.

This incoherence happens when all four income positions are occupied; that is, when there is what can be called a saturated profile. Suppose instead, as in Table 4C, for example, that there are no members with bad luck who choose low effort; that is the profile is *not* saturated. Then the rule of Equal Dividends for all those with bad luck is non-binding, as is the rule of Equal Payments for those with low effort. In effect, the Principle of Accountability requires here only that dividends be equal for those with good luck, and that payments be equal for those with high effort, which is arithmetically unproblematic. Therefore, a contingent PoA regime is feasible: Equal Payments at each level of effort, and Equal Dividends at each level of luck, but on condition that the profile is not saturated.

Given saturation, however, the follower of the PoA faces in effect a trade-off. This could feasibly take the form of an incomplete PoA regime: Equal Payments at each level of effort, and Equal Dividends at one level of luck (with unequal dividends at the other level of luck being whatever the arithmetic entails, as in Table 4b); or, alternatively, Equal Dividends at each level of luck, and Equal Payments at one level of effort (with unequal payments at the other being whatever the arithmetic entails, as in Table 4a).

Evaluating these lesser versions of the Principle of Accountability is far from straightforward. We could refer to Fleurbaey (2008) who approaches the question of the appropriate choice of the principles of distributive justice by applying the idea of *non-envy*; this is his interpretation of fairness. In Fleurbaey (2008),²⁰ an outcome

²⁰ We should note that there are several other ways of applying 'no envy', and that choosing between them is not an easy task. An alternative interpretation, leading to a 'proportional concept of egalitarian-equivalence' can be found in Almås et al. (2011).

is fair if no one envies anyone else *ex post*: no individual would prefer anyone else's final position²¹ to his own.

In some simple situations, the implications of non-envy are clear. Suppose that the only dimension is luck, with all individuals constrained to the same level of effort. Then straightforwardly envy-freeness requires Equal Payments for all. Now suppose instead that the only dimension is effort, with all individuals having the same level of luck. The implications of envy-freeness here are less straightforward, but can be reasoned as follows. If all individuals know that an Equal Dividends regime is in place, then each knows that his dividend will be $X/4$, and each makes his effort decision according to whether the additional payment is worth the additional effort. The outcome now must be envy-free, since a low-effort chooser could have achieved exactly the same outcome as a high-effort chooser, simply by choosing differently, and *vice-versa*. An Equal Payments regime *could* produce an envy-free outcome, for example if each individual chose low effort, but not necessarily.²²

However, when both dimensions are in play, things are more complicated, crucially because members do not know *ex ante* how many members will be in each cell of the matrix. An incomplete PoA regime, as described above, is generally not envy-free. Less obviously, neither a contingent PoA nor an incomplete PoA ensures envy-freeness at any level of luck. Suppose, for example, that bad luck dividends are to be equal. If it so happens that all members have bad luck, then the bad luck dividend will be $X/4$. (In addition, if all members know this then the outcome will plausibly be envy-free, as described above for Equal Dividends.) However, given also that payments are to be equalised at some level of effort, the bad luck dividend will be greater the more members there are, at that level of effort, with good luck rather than bad, their incremental incomes being in effect redistributed through increased dividends for other members. Therefore, a member with bad luck, when deciding his effort level, cannot know the bad luck dividend, and thus his payment at each effort level, without knowing how many members with good luck there are at each effort level for which payments are to be equalised. Of course, he does know the difference between his two relevant payments, since he knows that his dividend, whatever it is, will be unaffected by his choice. In addition, if his preferences happen to be linear in the payment then this will suffice for him to know whether high effort is *ex post* preferable to low effort, each together with its payment. But given instead non-linear preferences, he might make an *ex ante* rational choice of (say) high effort, but *ex post* regret this, correspondingly envying any member with bad luck who chose low effort.

It is at least doubtful, therefore, that with regard to fairness a contingent or an incomplete PoA regime is 'second-best' optimal. Partly this is because it is not obvious how to rank outcomes that are not fully envy-free, for example in terms of the

²¹ In the present context, 'final position' is best understood as the individual's own effort and payment.

²² There may be some individual for whom the additional effort is worthwhile, even though giving him in effect only a $1/4$ share of the additional income. But this produces an outcome in which that individual envies all those on low effort and the same payment.

degree of envy-freeness.²³ In addition, on any such measure, it is not clear that a contingent/incomplete Principle of Accountability will necessarily be superior to Equal Dividends everywhere or Equal Payments everywhere.

Appendix 4 default rule tests

In the experiment, we ran each treatment (internal and external dictator) under two different default conditions, to check for the existence of default bias in our experiment. In the *Equal payment default* condition, at the start of the simulations all choices are set to *equal payment*. This means that a subject not willing to do any effort can sit the experiment out and will submit an *equal payment* rule on both luck and effort dimensions. The *Equal dividends default* condition operates in a similar way, with the default set to *equal dividends*. This allows us to test whether there is an effect of a default (Table 11).

We test the default effect separately for each of the two treatments, and for each of the scenarios. The table reports p-values for χ^2 tests applied to raw choice data. Across the board, the default conditions have no significant effect.

These tests were done at the level of the *scenario*. This means that, in simple scenarios, an observation is a vector of two choices, one for each possible outcome of the variable of interest (luck or effort). We can run tests for sub-scenario variables. Results of the χ^2 tests are reported in the table. Again, no significance across the board. The only marginally significant value is the third choice of Scenario 3; this is significant at 10%. Nonetheless, these p-values were not corrected for multiple testing. Since the third choice in itself is likely not meaningful when taken in isolation, we are not worried by this little bit of significant evidence in a sea of insignificant correlations (Table 12).

We can conclude that there is no effect of the given default in our experiment. This is in line with qualitative data that tell us that subjects made extensive use of the simulator to inspect the results of different rules and the impact of different parameters, thereby watering down any default effect that might have existed.

In the paper, we therefore pool the data of the two conditions, and consider only two treatments: *Internal* and *External dictator*.

²³ Fleurbaey (2008) discusses on page 50 what might be meant by this.

Table 11 Tests of the effect of the default rule at scenario level

	Scenario 1	Scenario 2	Scenario 3
Internal dictator	0.3529	0.3749	0.8508
External dictator	0.4241	0.4247	0.2577

Table 12 Tests of the effect of the default rule at sub-scenario level

	Sc1a	Sc1b	Sc2a	Sc2b	Sc3a	Sc3b	Sc3c	Sc3d
Int	1	0.1848	0.3436	0.8498	1	0.5762	0.1972	0.2679
Ext	0.4119	0.2879	0.2131	0.7634	0.6449	1	0.05338	0.771

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Alesina, A., & Angeletos, G.-M. (2005). Fairness and redistribution. *American Economic Review*, *95*, 960–980
- Almås, I., Cappelen, A. W., Lind, J. T., Sørensen, E. Ø., & Tungodden, B. (2011). Measuring unfair (in) equality. *Journal of Public Economics*, *95*, 488–499
- Almås, I., Cappelen, A. W., & Tungodden, B. (2020). Cutthroat capitalism versus cuddly socialism: Are americans more meritocratic and efficiency-seeking than scandinavians? *Journal of Political Economy*, *128*, 1753–1788
- Arneson, R. (2000). Luck egalitarianism and prioritarianism. *Ethics*, *110*, 339–349
- Becker, A. (2013). Accountability and the fairness bias: the effects of effort vs. luck. *Social Choice and Welfare*, *41*, 685–699
- Bjerk, D. (2016). In front of and behind the veil of ignorance: an analysis of motivations for redistribution. *Social Choice and Welfare*, *47*, 791–824
- Bossert, W. (1995). Redistribution mechanisms based on individual characteristics. *Mathematical Social Sciences*, *29*, 1–17
- Bossert, W., & Fleurbaey, M. (1996). Redistribution and compensation. *Social Choice and Welfare*, *13*, 343
- Bukszar, E., & Knetsch, J. L. (1997). Fragile redistribution choices behind the veil of ignorance. *Journal of Risk and Uncertainty*, *14*, 63–74
- Cabrales, A., Miniaci, R., Piovesan, M., & Ponti, G. (2010). Social preferences and strategic uncertainty: An experiment on markets and contracts. *American Economic Review*, *100*, 2261–2278
- Cappelen, A. W., Hole, A. D., Sørensen, E. Ø., & Tungodden, B. (2007). Pluralism of fairness ideals: An experimental approach. *American Economic Review*, *97*, 818–827
- Cappelen, A. W., Konow, J., Sørensen, E. Ø., & Tungodden, B. (2013). Just luck: An experimental study of risk taking and fairness. *American Economic Review*, *103*, 1398–1413

- Cappelen, A. W., Mollerstrom, J., Reme, B. A., & Tungodden, B. (2019). A Meritocratic Origin of Egalitarian Behavior. *Research Institute of Industrial Economics*, Working Paper Series 1277
- Cappelen, A. W., Sorensen, E. Ø., & Tungodden, B. (2010). Responsibility for what? Fairness and individual responsibility. *European Economic Review*, *54*, 429–441
- Cappelen, A. W., & Tungodden, B. (2006). A liberal egalitarian paradox. *Economics and Philosophy*, *22*, 393–408
- Cettolin, E., & Reidl, A. (2017). Justice under uncertainty. *Management Science*, *63*, 3739–3759
- Cohen, G. A. (1989). On the currency of egalitarian justice. *Ethics*, *99*, 906–944
- Deffains, B., Espinosa, R., & Thöni, . (2016). Political self-serving bias and redistribution. *Journal of Public Economics*, *134*, 67–74
- Dworkin, R. (1981). What is equality? Part 2: Equality of resources. *Philosophy and Public Affairs*, *10*, 283–345
- Espinosa, R., Deffains, B., & Thöni, C. (2020). Debiasing preferences over redistribution: An experiment. *Social Choice and Welfare*, *55*, 823–843
- Fleurbaey, M. (1994). On fair compensation. *Theory and Decision*, *36*, 277–307
- Fleurbaey, M. (1995). Three solutions for the compensation problem. *Journal of Economic Theory*, *65*, 505–521
- Fleurbaey, M. (1995a). The Requisites of Equal Opportunity. In *Social Choice, Welfare and Ethics*, W. A. Barnett, H. Moulin, M. Salles, & N. Schofield (Eds.), Cambridge University Press, Cambridge
- Fleurbaey, M. (2008). *Fairness*. Oxford University Press.
- Frignani, N., & Ponti, G. (2012). Risk versus social preferences under the veil of ignorance. *Economics Letters*, *116*, 143–146
- Gerber, A., Nicklisch, A., & Voigt, S. (2014). The role of ignorance in the emergence of redistribution. *Journal of Economic Behavior and Organization*, *163*, 239–261
- Herne, K., & Suojanen, M. (2004). “The role of information in choices over income distributions. *Journal of Conflict Resolution*, *48*, 173–192
- Johannesson, M., & Gerdtham, U.-G. (1995). A pilot test of using the veil of ignorance approach to estimate a social welfare function for income. *Applied Economics Letters*, *2*, 400–402
- Konow, J. (2000). Fair shares: Accountability and cognitive dissonance in allocation decisions. *American Economic Review*, *90*, 1072–1091
- Mollerstrom, J., Reme, B. A., & Sørensen, E. O. (2015). Luck, choice and responsibility—An experimental study of fairness views. *Journal of Public Economics*, *131*:C, 33–40
- Pazner, E., & Schmeidler, D. (1974). A difficulty in the concept of fairness. *Review of Economic Studies*, *41*, 441–443
- Rawls, J. (1971). *A Theory of Justice*. Harvard University Press.
- Schildberg-Horisch, H. (2010). Is the veil of ignorance only a concept about risk? An experiment. *Journal of Public Economics*, *94*, 1062–1066
- Sutter, M., & Weck-Hanneman, H. (2003). Taxation and the veil of ignorance: A real effort experiment on the laffer curve. *Public Choice*, *115*, 217–240
- Voigt, S. (2013). Veilonomics: On the Use and Utility of Veils in Constitutional Political Economy. working paper

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.