WORKING PAPERS IN ECONOMICS

No. 09/22

ALEXANDER W. CAPPELEN
THOMAS DE HAAN
BERTIL TUNGODDEN

Fairness and limited information: Are people Bayesian meritocrats?

Department of Economics
UNIVERSITY OF BERGEN

Fairness and limited information: Are people Bayesian meritocrats?

Alexander W. Cappelen, Thomas de Haan, Bertil Tungodden*

4th October 2022

Abstract

Meritocracy is a prominent fairness view in many societies, but often difficult to apply because there is limited information about the source of inequality. This paper studies theoretically and empirically how limited information affects inequality acceptance. We connect the hitherto unrelated literatures on fairness and belief updating and show that irrationality in belief updating may be as important as differences in fairness views in explaining inequality acceptance. In many economic environments with limited information, signal-neglecting meritocrats act as egalitarians and base-rate neglecting meritocrats act as libertarians. The findings contribute to better understanding of the foundations of inequality acceptance in society.

*Alexander W. Cappelen, NHH Norwegian School of Economics; Thomas de Haan, University of Bergen; Bertil Tungodden, NHH Norwegian School of Economics. We are grateful to Peter Andre, Stefano DellaVigna, Dirk Engelmann, Benjamin Enke, Christine Exley, Eliana La Ferrara, Lata Gangadharan, Steffen Huck, Johanna Mollerstrom, Nikos Nikiforakis, Ricardo Perez-Ruglia, Bjorn-Atle Reme, Roberto Weber, Georg Weizsaecker, and David Yanagizawa-Drott for valuable comments and suggestions. We are also grateful to Caroline Bonn for her early role in the project, and to Johanna Croton for excellent research assistance. The project was financed by support from the Research Council of Norway through its Centres of Excellence Scheme, FAIR project No 262675 and Research Grant 236995, and administered by FAIR–The Choice Lab.

1 Introduction

Fairness considerations are highly important to how people view income inequality and shape attitudes towards redistributive polices (Alesina and Giuliano (2011); Almås et al. (2020); Stantcheva (2021)). The meritocratic fairness view, which considers income inequality due to differences in performance as fair but inequality due to luck as unfair, is prominent in modern society (Piketty (2020); Sandel (2020)). However, there are heated debates about the implications of this fairness view because most economic situations are characterized by limited information about the actual source of inequality, and few topics create a greater political divide than the role of luck in individual success (Frank (2016); Mankiw (2013); Moffitt (2015)). A key question for understanding inequality acceptance is therefore how people respond to limited information about the source of inequality when considering whether an inequality is fair or unfair.

We study, theoretically and experimentally, how limited information about the source of inequality affects inequality acceptance. In the theoretical analysis we consider economic environments in which an individual's earnings are determined by performance and luck, and a third–party spectator decides whether to redistribute income between two individuals without knowing the source of the inequality in their earnings. The spectator knows the distribution of performance and luck in society, and updates their beliefs about the performance and luck of the two individuals based on information about their earnings. We show that Bayesian spectators who give each individual their expected fair income, implement in expectation less inequality with limited information than with full information if and only if limited information causes fairness-ranking uncertainty, that is uncertainty about who deserves a higher income. This result applies to any fairness view that the spectator may hold and any earnings function. As a corollary, if the spectator is a Bayesian meritocrat, it follows that limited information causes less inequality acceptance when there is uncertainty about who has the better performance.

A large literature has shown that people often violate Bayes rule when updating their beliefs, and the second part of our theoretical analysis studies how irrational updating influences inequality acceptance with limited information. We focus on signal-neglecting spectators (who have posterior beliefs equal to their prior beliefs) and base–rate-neglecting spectators (who ignore the prior beliefs when updating), and show that these two types of irrationality have very different effects on inequality acceptance with limited information. In a large set of economic environments, we show that signal–neglecting meritocrats implement the egalitarian solution and base-rate-neglecting meriotcrats implement the libertarian solution. This analysis highlights how irrational belief updating may shape inequality acceptance, and shows that differences in belief updating may cause people with the same meritocratic fairness view and the same available information to accept very different levels of inequality.

In the final part of the paper, we report from an experimental study on how limited information affects inequality acceptance. In the experiment, we randomly vary in a between-individual design whether spectators have full information or limited information about the source of inequality when making redistributive decisions between two other individuals. The experimental environment is a special case of our theoretical framework and, in line with the theoretical analysis, we show that there is a strong positive relationship between the spectator's belief updating and how much inequality is implemented in the experiment. Spectators underreacting to the earnings signal implement less inequality than spectators overreacting to the earnings signal. Overall, we do not find that the spectators implement significantly more inequality with limited information than with full information, which is consistent with irrational updating creating counteracting effects on implemented inequality. We further estimate a structural behavioral model of spectator behavior that allows for heterogeneity in both fairness views and belief updating. The estimated behavioral model finds that most spectators have a meritocratic fairness view but do not update in a Bayesian manner. The behavioral model fits the experimental data better than an estimated model assuming that all spectators are Bayesian updaters. Finally, we show that the estimated distribution of updating strength from the behavioral model is strikingly similar to the distribution of updating strength established by an incentivized belief elicitation procedure. Taken together, the empirical analysis provides evidence that heterogeneity in belief updating is an important explanation for why people differ in their inequality acceptance when there is limited information.

The paper connects two hitherto unrelated literatures, the literature on fairness preferences and the literature on belief updating. It provides several insights on the nature of inequality acceptance and meritocracy. First, it provides a better understanding of how limited information may shape inequality acceptance. In particular, the theoretical analysis clarifies the key role of fairness-ranking uncertainty for people who hold the meritocratic fairness view. Limited information only leads to increased inequality acceptance among rational spectators in economic environments where limited information causes uncertainty about who deserves a higher income. Limited information that does not introduce uncertainty about who is more deserving, would not affect the inequality acceptance of rational spectators. Second, the theoretical analysis makes a novel link between fairness preferences and irrational belief updating by showing how people with the same fairness view may diverge in their inequality acceptance when there is limited information. In many economic environments, signal-neglecting meritocrats implement the egalitarian solution, whereas base-rate-neglecting meritocrats implement the libertarian solution. Hence, differences in belief updating may be as important as differences in fairness views in accounting for disagreements about inequality. Relatedly, the analysis shows that limited information may cause an increase in inequality acceptance if people are base-rate-neglecting meritocrats and interpret earnings as an indicator of expected performance. Third, the experiment shows that people indeed differ in how they interpret an earnings signal, and provides an empirical illustration of the importance of heterogeneity in belief updating in explaining distributive behavior. With full information, a large majority of the spectators agree on the meritocratic solution, whereas they disagree significantly with limited information because they update the earnings signal differently. Finally, the paper makes a methodological contributions by providing, to our knowledge, the first structural behavioral model that captures heterogeneity both in fairness preferences and belief updating, which also allows for a comparison of beliefs inferred from choice data and beliefs elicited using a scoring rule.

The paper advances both the theoretical and the empirical literature on social preferences (Andreoni and Miller (2002); Bartling et al. (2015); Bellemare et al. (2008); Bolton and Ockenfels (2000); Cappelen and Tungodden (2019); Charness and Rabin (2002); Engelmann and Strobel (2004); Exley and Kessler (2019); Fehr et al. (1993); Fehr and Schmidt (1999); Rabin (1993)), and provides new insights on how the source of inequality shapes inequality acceptance (Akbaş et al. (2019); Alesina et al. (2001); Almås et al. (2020); Andre (2021); Barr et al. (2021); Balafoutas et al. (2013); Fong (2001); Konow (1996); Konow (2000); Konow (2009); Cappelen et al. (2007); Cappelen et al. (2013); Cappelen et al. (2022); Cassar and Klein (2019); Durante et al. (2014); Krawczyk (2010); Mollerstrom et al. (2015); Müller and Renes (2021); Sugden and Wang (2020)). The vast majority of papers in this literature has focused on characterizing social preferences by studying experimentally how people make decisions with complete information. The present study focuses on how people handle limited information about the source of inequality when they act as a third-party spectator.[1] The third-party spectator approach may be seen as capturing an individual's moral view of inequality, and thus the present study provides new insights on how limited information shapes the moral acceptability of inequality (Konow, 2012). In particular, we establish theoretical and empirical results on how belief updating about the source of inequality influences inequality acceptance, and thereby integrate two main topics in behavioral economics: social preferences and bounded rationality. A key message of this paper is that heterogeneity in belief updating may be as

[1]Some earlier papers have examined how limited information may introduce a self-serving bias when people have a stake in the decision (Chavanne (2018); Cruces et al. (2013); Davidai and Gilovich (2016); Di Tella et al. (2015); Erkal et al. (2022); Fehr and Vollmann (2020); Langer (1975); Rodriguez-Lara and Moreno-Garrido (2012); Valero (2022)), and a related literature has examined theoretically and empirically how other types of limited information may shape redistributive preferences (Bortolotti et al. (2017); Gross et al. (2015); Cettolin and Riedl (2017); Cettolin et al. (2017); Exley (2016); Fehr and Vollmann (2020); Fudenberg and Levine (2012); Saito (2013)).

important as heterogeneity in fairness views for understanding disagreements about inequality. The paper advances the structural approach to behavioral economics by estimating a behavioral model that incorporates heterogeneity in both fairness views and belief updating (DellaVigna, 2018).

The paper also contributes to the literature on irrational beliefs (Benjamin (2019); Benjamin et al. (2019); Enke and Zimmermann (2019); Enke (2020)). This literature has provided evidence of both signal-neglect and base-rate neglect in belief updating (Benjamin, 2019), and the present paper studies how these types of irrational belief updating shape inequality acceptance. Our theoretical analysis considers updating of continuous distributions, and the corresponding experimental analysis studies belief updating in choices based on beliefs about distributions over many possible states. Bayesian spectators would update toward the signal in the experimental setting (Chambers and Healy, 2012), where the weight assigned to the signal is determined by the variance in the performance distribution relative to the variance in the random component. However, we find extensive evidence of irrational belief updating, with a significant share of signal neglecters and a significant share of base-rate neglecters.

Section 2 introduces the general framework and derives our main theoretical results. Sections 3 and 4 introduce the experimental design and establish theoretical predictions for the experiment. Section 5 reports descriptive analysis and treatment effects, and Section 6 estimates a structural model of spectator behavior. Section 7 concludes. Supporting theoretical and empirical analysis are provided in Appendixes A and B, and the experimental instructions are provided in Appendix C.

2 Theory

In this section, we provide some general results on how limited information affects inequality acceptance. First, we present formally the economic environment and characterize optimal spectator behavior in any given distributive situation. Second, we study the effects of Bayesian and non-Bayesian updating on expected inequality acceptance across situations, with a particular focus on the implications for spectators with the meritocratic fairness view.

2.1 The economic environment

Consider an economic environment where workers perform a task. Each worker earns $x_i = s(p_i, \varepsilon_i)$, where p_i is the performance of worker i and ε_i is a random factor, $x_i, p_i, \varepsilon_i \in \mathbb{R}$.[2] We assume that earnings are strictly increasing in both arguments, but do not impose any further restrictions on the earnings function. An impartial spectator has to decide how to distribute the total earnings, $X = x_i + x_j, X > 0$, between two workers, where y_i and y_j are the incomes assigned to worker i and worker j by the spectator. There is no cost of redistribution, $y_i + y_j = X$. The implemented inequality in income is given by $I = \frac{|y_i - y_j|}{y_i + y_j}$. Let $m_i \geqslant 0$ be what the spectator considers the fair income to worker i, $m_i + m_j = X$.

The spectator knows the earnings function, that the workers' performances are randomly drawn from a distribution of pairs of performances, $f^{prior}(\mathbf{p})$, and that the random components are drawn from a distribution of pairs of random components, $h^{prior}(\boldsymbol{\epsilon})$, which define the prior beliefs of the spectator. We allow for the possibility that the prior beliefs about the performance and the random component may differ for the two workers and that performances or the random components may be correlated. Let $f^{prior}(m_i)$ be the prior distribution of what the spectator considers the fair income to worker i.

When making the distributive decision, the spectator is either in a full information situation or a limited information situation. In the full information situation, the spectator receives a fully informative signal $S = (x_i, x_j, p_i, p_j, \varepsilon_i, \varepsilon_j)$. In the limited information situation, the spectator only receives a signal in

[2] The theoretical framework is very general and allows for alternative interpretation of the variables performance and luck. The variables may be interpreted as capturing the distinction between factors that people believe that others should be held responsible for (performance) and not responsible for (luck), or they may be interpreted as capturing factors that are within individual control (performance) and beyond individual control (luck). For a broader discussion of these issue, see Cappelen (2019). In the experimental part of the paper, we indeed find that people make a distinction between the performance and the luck of the participants in their distributive choices, in line with the theoretical framework.

terms of the earnings of the two workers, $\mathbf{x} = (x_i, x_j) = (s(p_i, \varepsilon_i), s(p_j, \varepsilon_j))$. Let $f^{posterior}(m_i|\mathbf{x})$ be the posterior belief distribution of the fair income to worker i and $f^{posterior}(\mathbf{p}|\mathbf{x})$ be the posterior belief distribution about the performance of the two workers after observing the earnings signal $\mathbf{x}$, where $g(\mathbf{x})$ is the distribution of all possible earnings signals. Generally, $f^{posterior}(m_i|\mathbf{x}) = \int f^{posterior}(\mathbf{p}|\mathbf{x}) \cdot 1_{m_i(\mathbf{p},\mathbf{x})=m_i}\, d\mathbf{p}$. In the analysis, we denote $dp_i dp_j$ as $d\mathbf{p}$ and $dx_i dx_j$ as $d\mathbf{x}$, and all integrals are over the full domain unless specified otherwise. The only constraint that we impose on the distributions is that they allow for the use of Bayes rule.

2.2 Optimal spectator behavior

We assume that spectators are motivated by fairness and dislike deviating from what they consider fair (Cappelen et al., 2013), as captured by the following utility function,

$$U_{spectator} = -(y_i - m_i)^2 \tag{1}$$

It follows straightforwardly that the optimal choice of the spectator in a full information situation is to give each worker their fair income,

$$y_i^{*FI} = m_i \tag{2}$$

Consequently, the optimal inequality in a full information situation is given by:

$$I^{*FI} = |\frac{m_i - m_j}{m_i + m_j}| = 2|\frac{m_i}{X} - \frac{1}{2}| \tag{3}$$

The optimal level of inequality depends on the spectator's fairness view, where three fairness views are particularly salient (Almås et al. (2020), Konow (2000)):

- *Egalitarian fairness view:* It is fair that the total earnings are divided equally between the two workers, $m_i = \frac{1}{2} \cdot X$.
- *Meritocratic fairness view:* It is fair that that the total earnings are

divided proportional to performance, $m_i = \frac{p_i}{p_i+p_j} \cdot X$.

- *Libertarian fairness view:* It is fair that the workers receive their earnings, $m_i = x_i$.

In a limited information situation, the expected utility of the spectator is given by:

$$EU_{spectator} = -E(y_i - m_i)^2 = -\int f^{posterior}(m_i|\mathbf{x}) \cdot (y_i - m_i)^2 \, dm_i \tag{4}$$

It follows that the optimal choice of the spectator in a limited information situation is to give worker i the expected fair income (see Appendix A.1):

$$y_i^{*LI} = E(m_i) = \int f^{posterior}(m_i|\mathbf{x}) \cdot m_i \, dm_i \tag{5}$$

To establish the optimal inequality in situations with limited information, we introduce the function $\zeta(m_i) = 2 \cdot \left(\frac{m_i}{X} - \frac{1}{2}\right)$. It gives the signed fair inequality (ranging from -1 to 1) for a given fair income m_i to worker i. We can now show that the optimal inequality implemented by the spectator in a limited information situation is given by (see Appendix A.1):

$$I^{*LI} = \left| \int f^{posterior}(m_i|\mathbf{x}) \cdot \zeta(m_i) \, dm_i \right| \tag{6}$$

Limited information may cause a spectator to hold posterior beliefs that give positive support both to states of the world in which the spectator considers it fair to give more to worker i and to states of the world in which the spectator considers it fair to give more to worker j. In such cases, we say that the posterior beliefs reflect fairness-ranking uncertainty.

Fairness-ranking uncertainty: Given a signal $\mathbf{x}$, the spectator's posterior beliefs reflect fairness-ranking uncertainty if and only if there exist both situations with $f^{posterior}(m_i > m_j|\mathbf{x}) > 0$ and $f^{posterior}(m_i < m_j|\mathbf{x}) > 0$.

Limited information causes fairness-ranking uncertainty for a spectator if there is positive support for a signal $\mathbf{x}$, $g(\mathbf{x}) > 0$, such that this signal causes fairness-ranking uncertainty in the posterior beliefs of the spectator.

It follows from (6) that a spectator, after observing an earnings signal, implements the expected optimal inequality given their posterior beliefs if there is no fairness-ranking uncertainty; if there is fairness-ranking uncertainty, the spectator implements a lower level of income inequality.

2.3 Bayesian spectators

In this section, we study how limited information affects the expected implemented inequality for a Bayesian spectator across situations.

By simple manipulation, we establish the following lemma:

Lemma 1: *Bayesian updating implies:*

$$\int f^{Bposterior}(m_i|\mathbf{x}) \cdot g(\mathbf{x})\ d\mathbf{x} = f^{prior}(m_i) \tag{7}$$

Lemma 1 shows that limited information does not make a Bayesian spectator distort their beliefs about the underlying distribution of situations in the economy.

We can now establish that for a Bayesian spectator, the effect of limited information on implemented inequality depends critically on whether limited information causes fairness-ranking uncertainty.

Proposition 1: *A Bayesian spectator implements in expectation the same level of inequality with limited information and full information if limited information does not cause fairness-ranking uncertainty, and strictly less inequality with limited information than with full information if limited information causes fairness-ranking uncertainty.*

Proof: (i) We first show that a Bayesian spectator implements in expectation the same level of inequality with limited and full information if limited information does not cause fairness-ranking uncertainty.

It follows from equation (6) that the expected level of inequality with limited information is given by:

$$E\left(I^{*LI}\right)=\int\left|\int f^{Bposterior}(m_i|\mathbf{x})\cdot\zeta\left(m_i\right)\ dm_i\right|\cdot g\left(\mathbf{x}\right)\ d\mathbf{x} \tag{8}$$

If limited information does not cause fairness-ranking uncertainty, then, without loss of generality, we can assume that $\zeta\left(m_i\right)\geq 0$ for all possible states of the world and all signals. Hence, it follows that:

$$\int\left|\int f^{Bposterior}(m_i|\mathbf{x})\cdot\zeta\left(m_i\right)\ dm_i\right|\cdot g\left(\mathbf{x}\right)d\mathbf{x}=\int\int f^{Bposterior}(m_i|\mathbf{x})\cdot\left|\zeta\left(m_i\right)\right|dm_i\cdot g\left(\mathbf{x}\right)d\mathbf{x} \tag{9}$$

By simple manipulation:

$$\int\int f^{Bposterior}(m_i|\mathbf{x})\cdot\left|\zeta\left(m_i\right)\right|dm_i\cdot g\left(\mathbf{x}\right)d\mathbf{x}=\int\int f^{Bposterior}(m_i|\mathbf{x})\cdot g\left(\mathbf{x}\right)d\mathbf{x}\cdot\left|\zeta\left(m_i\right)\right|dm_i \tag{10}$$

By Lemma 1:

$$\int\int f^{Bposterior}(m_i|\mathbf{x})\cdot g\left(\mathbf{x}\right)d\mathbf{x}\cdot\left|\zeta\left(m_i\right)\right|dm_i=\int f^{prior}(m_i)\left|\zeta\left(m_i\right)\right|dm_i=E\left(I^{*FI}\right) \tag{11}$$

(ii) We now show that a Bayesian spectator implements in expectation strictly less inequality with limited information than with full information if limited information causes fairness-ranking uncertainty.

In this cases, both $\zeta\left(m_i\right)>0$ and $\zeta\left(m_i\right)<0$ have positive support in the posterior belief distribution $f^{Bposterior}(m_i|\mathbf{x})$ for some signal $\mathbf{x}$. It follows that:

$$\int\left|\int f^{Bposterior}(m_i|\mathbf{x})\cdot\zeta\left(m_i\right)\ dm_i\right|\cdot g\left(\mathbf{x}\right)d\mathbf{x}<\int\int f^{Bposterior}(m_i|\mathbf{x})\cdot\left|\zeta\left(m_i\right)\right|dm_i\cdot g\left(\mathbf{x}\right)d\mathbf{x} \tag{12}$$

The result follows from applying the last two steps in part (i) of the proof.■

The proposition shows that limited information causes in expectation (weakly) less inequality acceptance among Bayesian spectators under very general conditions. It covers any fairness view, including the specific fairness views introduced in the previous section, and does not impose any restrictions on the prior distributions or the earnings function. The proposition also shows that fairness-ranking uncertainty is key to understanding how limited information affects inequality acceptance among Bayesian spectators. Limited information only leads to strictly less inequality acceptance among Bayesian spectators across situations when there is fairness-ranking uncertainty. The proof of the propositions rests on two main insights. First, with limited information and fairness-ranking uncertainty, a Bayesian spectator implements strictly less than the expected optimal inequality in any given situation (equation (6)); second, a Bayesian does not distort their beliefs about the underlying distribution of situations in the economy (Lemma 1). Taken together, this leads a Bayesian spectator to implement strictly less inequality with limited information than with full information across situations. See Appendix A.1 for a numerical example illustrating the intuition in the proof

Let us now consider the implications of Proposition 1 for egalitarians, libertarians, and meritocrats. Bayesian egalitarians and libertarians do not rely on the signal when considering what is fair, and hence their posterior beliefs are never characterized by fairness-ranking uncertainty. Therefore, in line with the first part of Proposition 1, they implementation in expectation the same level of inequality with full information and with limited information. For Bayesian meritocratic spectators, the signal is important because it makes them update their beliefs about the performance of the two workers. Limited information causes fairness-ranking uncertainty for Bayesian meritocratic spectators if it causes performance-ranking uncertainty.

Performance-ranking uncertainty: Given a signal $\mathbf{x}$, the spectator's posterior beliefs reflect performance-ranking uncertainty if and only if there exist both situations with $f^{posterior}(p_i > p_j|\mathbf{x}) > 0$ and $f(p_i < p_j|\mathbf{x}) > 0$.

Performance-ranking uncertainty implies fairness-ranking uncertainty for a Bayesian meritocratic spectator, who considers it fair to give more to worker i in the states of the world in which worker i has performed better and fair to give more to worker j in the states of the world in which worker j has performed better. Bayesian meritocratic spectators face fairness-ranking uncertainty if and only if they face performance-ranking uncertainty, and we can state the following corollary:

Corollary 1*: A Bayesian meritocratic spectator implements in expectation the same level of inequality with limited information and full information if limited information does not cause performance-ranking uncertainty, and strictly less inequality with limited information than with full information if limited information causes performance-ranking uncertainty.*

The corollary highlights that performance-ranking uncertainty is of fundamental importance for inequality acceptance among Bayesian meritocratic spectators, and that limited information about the performance of workers may lead them to implement less inequality compared with when they have full information.

2.4 Non-Bayesian spectators

A large literature has shown that people often violate Bayes rule when updating their beliefs, and we now turn to a study of how irrational updating influences inequality acceptance under limited information. We focus on the following two types of non-Bayesian spectators:[3]

Signal-neglecter: A signal-neglecting spectator holds posterior beliefs:

$$f^{SNposterior}(m_i|\mathbf{x}) = f^{prior}(m_i), \forall \mathbf{x}$$

Base-rate neglecter: A base-rate-neglecting spectator holds posterior beliefs:

$$f^{BRNposterior}(m_i|\mathbf{x}) = \frac{g(\mathbf{x}|m_i)}{\int g(\mathbf{x}|m_i)dm_i}$$

[3]The definitions of non-Bayesian spectators are in line with Gerther's (1980) reduced form model of updating: $f^{posterior}(m_i|\mathbf{x}) \propto g(\mathbf{x}|m_i)^c \cdot f^{prior}(m_i)^d$, with $c = 1$ defining signal neglecters and $d = 0$ defining base-rate neglecters, see also Benjamin (2019).

The posterior beliefs of signal neglecters are the same as their prior beliefs. Hence, as for Bayesian spectators, limited information does not make a signal-neglecting spectator distort their beliefs about the underlying distribution of situations in the economy, which means that a version of Lemma 1 also applies to signal neglecters. Hence, we can establish the following proposition:

Proposition *2: A signal-neglecting spectator implements in expectation the same level of inequality with limited information and full information when there is no fairness-ranking uncertainty, and strictly less inequality with limited information than with full information if limited information causes fairness-ranking uncertainty.*

Proof: The result follows from replacing Lemma 1 with $\int f^{SNposterior}(m_i|\mathbf{x}) \cdot g(\mathbf{x})\ d\mathbf{x} = f^{prior}(m_i)$ in the proof of Proposition 1.■

Signal neglecters face posterior fairness-ranking uncertainty when there is prior fairness-ranking uncertainty, which implies that there exist $f^{prior}(m_i > m_j) > 0$ and $f^{prior}(m_i < m_j) > 0$. If the signal neglecter is a meritocrat, it follows that there is posterior fairness-ranking uncertainty when the prior performance distribution $f^{prior}(\mathbf{p})$ has positive support for both $f^{prior}(p_i > p_j)$ and $f^{prior}(p_i < p_j)$. It follows from Proposition 2 that in such economic environments, limited information causes signal-neglecting meritocrats to implement less inequality than with full information.

Base-rate neglecters do not take the prior performance distribution into account when updating. As a result, base-rate neglecters may distort their beliefs about the underlying distribution of situations in the economy, which means that we cannot apply a version of Lemma 1. Hence, we need to impose further assumptions on the economic environment to establish results on how limited information affects the inequality acceptance of base-rate neglecters. In this analysis, we focus on spectators who hold the meritocratic fairness view.

We consider the class of economic environments in which the performance and the random factor enter additively in the earnings function and are drawn from the same normal distributions for both workers:

Additive normal economic environment: The earnings of worker i are given by $x_i{=}p_i{+}\varepsilon_i$, with p_i and ε_i being drawn independently from the normal distributions $f^{prior}(p)$ and $h^{prior}(\epsilon)$.

The additive normal framework implies that there are earnings signals that generate fairness-ranking uncertainty for meritocrats. We further assume that the economic environment satisfies the following property:

Expected performance inequality approximation: An economic environment satisfies expected performance inequality approximation if, $E(\frac{p_i-p_j}{p_i+p_j}|\mathbf{x}) = \int f^{posterior}(p_i, p_j|\mathbf{x}) \cdot \frac{p_i-p_j}{p_i+p_j}\, d\mathbf{p}$ can be approximated by the ratio $\frac{E(p_i|x_i)-E(p_j|x_j)}{E(p_i|x_i)+E(p_j|x_j)}$, with $E(p_i|x_i) = \int f^{posterior}(p_i|x_i) \cdot p_i\, dp_i$.

Expected performance inequality approximation holds as long as we do not consider performances that are close to 0, since the expected value of the ratio of a set of normally distributed random variables is well approximated by the ratio of the expected values when the denominator is not close to 0.

We can now establish the following result:

***Proposition** 3: A base-rate-neglecting meritocratic spectator implements in expectation strictly more inequality with limited information than with full information if the economic environment is additive normal and satisfies expected performance inequality approximation.*

Proof: (i) We first establish that $E\left(\left|\frac{p_i-p_j}{p_i+p_j}\right|\right) < E\left(\left|\frac{\mu^{BRN}_{p_i|x}-\mu^{BRN}_{p_j|x}}{\mu^{BRN}_{p_i|x}+\mu^{BRN}_{p_j|x}}\right|\right)$.

The normal distribution has the special property that the pdf of any normal distribution can be obtained from taking another normal distribution and implement a linear transformation. As a special case, for two normal distributions with an identical mean but different variances, we can obtain one from the other by multiplying each point's distance to the mean with a factor α, with $\alpha < 1$ if we project the higher variance distribution on the lower variance distribution.

This means we can construct a mapping $\omega(b_i) = \alpha \cdot (b_i - \mu^*) + \mu^* = \alpha{\cdot}b_i{+}(1-\alpha){\cdot}\mu^*$ from the higher variance distribution of b_i to the lower variance distribution of a_i such that the likelihood of a_i is equal to the likelihood of $\omega(b_i)$.

It follows that:

$$E\left(\left|\frac{a_i - a_j}{a_i + a_j}\right|\right) = 2 \cdot E\left(\left|\frac{a_i}{a_i + a_j} - \frac{1}{2}\right|\right) = 2 \cdot E\left(\left|\frac{\alpha \cdot b_i + (1-\alpha) \cdot \mu^*}{\alpha \cdot b_i + (1-\alpha) \cdot \mu^* + \alpha \cdot b_j + (1-\alpha) \cdot \mu^*} - \frac{1}{2}\right|\right)$$

$$< 2 \cdot E\left(\left|\frac{b_i}{b_i + b_j} - \frac{1}{2}\right|\right) = E\left(\left|\frac{b_i - b_j}{b_i + b_j}\right|\right)$$

By the additive normal assumption, $p_i \sim N\left(\mu_p, \sigma_p^2\right)$ and $\mu_{p_i|x}^{BRN} = x_i \sim N\left(\mu_p, \sigma_p^2 + \sigma_\varepsilon^2\right)$, with $\sigma_p^2 < \sigma_p^2 + \sigma_\varepsilon^2$. Hence, it follws that $E\left(\left|\frac{p_i - p_j}{p_i + p_j}\right|\right) < E\left(\left|\frac{\mu_{p_i|x}^{BRN} - \mu_{p_j|x}^{BRN}}{\mu_{p_i|x}^{BRN} + \mu_{p_j|x}^{BRN}}\right|\right)$.

(ii) The expected inequality implemented by a base-rate-neglecting meritocrat with full information is $E\left(\left|\frac{p_i - p_j}{p_i + p_j}\right|\right)$. Combining the expected performance inequality approximation with equation (6) and the transformation $f^{posterior}(m_i|\mathbf{x}) = \int f^{posterior}(\mathbf{p}|\mathbf{x}) \cdot 1_{m_i(\mathbf{p},\mathbf{x}) = m_i} \, d\mathbf{p}$, it follows that the expected inequality implemented with limited information is $E\left(\left|\frac{\mu_{p_i|x}^{BRN} - \mu_{p_j|x}^{BRN}}{\mu_{p_i|x}^{BRN} + \mu_{p_j|x}^{BRN}}\right|\right)$. The result follows from taking together (i) and (ii). $\blacksquare$

Proposition 3 shows that in an important class of economic environments, limited information causes base-rate-neglecting meritocrats to implement more inequality than they would with full information. The basic intuition of the proof is that when the signaling function is additive normal, then limited information causes base-rate neglecting meriotcrats to overestimate the performance inequality in the underlying distribution of situations. The distorted beliefs create a pull towards implementing more inequality with limited information than with full information, which counteracts the effect of fairness-ranking uncertainty. This pull towards accepting more inequality for base-rate-neglecting meritocrats also applies to many other economic environments, including environments with a multiplicative earnings function, but there exist economic environments in which limited information causes a pull towards less inequality acceptance among base-rate-neglecting meritocrats.[4]

Taking together the different parts of the analysis, we observe that limited information may cause significant divergence in the inequality acceptance by meritocratic spectators with the same available information, depending on how they update their beliefs. We summarize this insight in the following proposition:

Proposition *4: Limited information causes signal-neglecting meritocrats to implement the egalitarian solution and in expectation strictly less inequality than Bayesian meritocrats, and base-rate-neglecting meritocrats to implement the libertarian solution and in expectation strictly more inequality than Bayesian meritocrats, if the economic environment is additive normal and satisfies expected performance inequality approximation.*

Proof. Combining equation (6) and transformation, $f^{posterior}(m_i|\mathbf{x}) = \int f^{posterior}(\mathbf{p}|\mathbf{x}) \cdot 1_{m_i(\mathbf{p},\mathbf{x})=m_i}\, d\mathbf{p}$, it follows that a signal-neglecting meritocratic spectator implements the egalitarian solution under limited information if, $E^{SN}\left(I^{*LI}\right) = \left|\int f^{prior}(\mathbf{p}) \cdot \zeta(p_i, p_j)\, d\mathbf{p}\right| = 0$, which holds trivially in an additive normal economic environment (but also in many other economic environments). A Bayesian meritocrat would always implement some inequality in an additive normal economic environment and thus implements more inequality than a signal-neglecting meritocrat.[5] It follows from the proof of Proposition 3 that, in an additive normal economic environment that satisfies expected performance inequality approximation, base-rate-neglecting meritocrats implement the libertarian solution, and from combining Proposition 1 and Proposition 3 that they implement more inequality than the Bayesian meritocrat in such an economic environment.■

[4]To illustrate, imagine an almost fully uninformative signaling technology, with every performance in the domain $[0, 100]$ being approximately equally likely to emit each possible earnings signal. In such a case, limited information would make a base-rate neglecter believe that there is less performance inequality than what is actually the case.

[5]In Appendix A.2, we provide a proposition showing more generally when a signal-neglecting meritocratic spectator implements strictly less inequality than a Bayesian meritocratic spectator in situations with limited information.

3 Experimental design

The experiment consisted of three parts, see Appendix C for detailed instructions. In the first part, workers earned money in a real effort task; in the second part, spectators decided on how to divide the earnings between randomly paired workers, and, in the third part, the workers were paid according to the decisions of the spectators. The spectator decisions are the main focus of our analysis, with the workers recruited only to create real economic situations.

3.1 The workers

We recruited 800 workers via Amazon Mechanical Turk. They performed the real effort encryption task introduced by Benndorf et al. (2014). In this task, workers are shown three letters on a screen. At the bottom of the screen, the letters of the alphabet are presented in random order, with each letter assigned a random number between 100 and 1,000. The task is to enter the corresponding numbers for the three letters. The workers then proceed to the next screen where they are given three new letters. They had 15 minutes to work on the task, and they earned a point for each correctly filled screen.

The workers were told that they would receive 2 US dollars (USD) in base payment if they completed at least 20 correct encryptions and that they also could get a bonus income. They were informed about the procedure that determined this income:

- Each worker were assigned earnings points equal to their performance (the number of correct encryptions) plus a random number (an integer) drawn from a normal distribution with a mean 0, truncated between –60 and 60.
- Each worker would then be randomly matched with another worker, and a third-party spectator would decide how to divide the earnings between the two workers.

- The income of a worker from the task would be equal to the share of the earnings points assigned to this worker by the spectator.

The workers completed the task the week before the spectator part of the experiment.[6] Workers were only informed about their own performance. The random factor for each worker was drawn after all workers had completed the task, which allowed us to calibrate the variance in the distribution of the random factor to the variance in the worker performance. The workers were paid the income from the task according to the spectators' decisions in the following week, with a conversion rate of 0.05 USD for each point. They were given no further information. On average, the workers earned 5.43 USD (including the base payment).

3.2 The spectators

We recruited 425 first-year students from NHH Norwegian School of Economics to be spectators in the study.[7] The spectators were randomly allocated either to the Full Information treatment or the Limited Information treatment. In both treatments, the spectators were given the following information:

- A description of the encryption task that the workers had completed.
- The procedure determining earnings for the workers. The spectators were informed that the workers had been told that a third party would decide how total earnings would be divided between two randomly paired workers.
- The distribution of worker performance and the distribution of the random component.

The spectators were presented with 10 pairs of workers for which they made distributive decisions. In the Full Information treatment, for each pair, the

[6]Four workers completed less than 20 screens, with the lowest performance being 13 screens. They did not receive the base payment, but were matched with another worker and paid an income according to the assigned spectator choice.

[7]In line with Gächter et al. (2010), we consider students to be an excellent subject pool for this experiment because it aims to test theory with rather complex instructions.

spectators were informed about the earnings, performance, and random factor of each worker, and they then decided how to split the total earnings between the two workers. The spectators in the Limited Information treatment were only informed about the earnings of each worker. We implemented comprehension checks to ensure that the participants understood the instructions for the distributive decisions. In the Limited Information treatment, after the spectators had made the distributive decisions, we elicited incentivized posterior beliefs from the spectators about performance for each of the 10 pairs of workers. On the incentivized belief questions, the spectators earned extra points depending on how close their guesses were to the correct answer.[8]

In addition, the participants provided background information (age, gender, political preferences) and answered some further questions. In Table B1, we show that the sample is balanced between the the two treatments on the background characteristics. The participants were paid a fixed fee of 50 NOK (equivalent to 6.29 USD at the time of the experiment) and their earnings from the belief questions. On average, the spectators earned 89 NOK. The experiment was double-blind and the spectators where paid anonymously in sealed envelopes when they left the experiment.

4 Experiment - predictions

In this section, we provide predictions for the spectator behavior in the experiment. The experimental environment is a special case of the general theory, which covers the discretization of the state space and the truncated distributions used in the experiment.

4.1 The economic environment in the experiment

In the experiment, we implement an additive normal framework. Each worker earns $x_i = p_i + \varepsilon_i$, where p_i is the performance of worker i and ε_i is an indepen-

[8]We implemented a quadratic scoring rule (Savage, 1971). On each belief question, the spectator earned: $max(0, 10 - 0.04 \cdot (correct - guess)^2 * 5)$ NOK.

dent and identically distributed (i.i.d.) random component of earnings with a mean of 0 and a standard deviation of $\sigma_\varepsilon = 15$. The workers are randomly paired and a spectator is randomly allocated to a pair of workers, worker i and worker j. In Appendix B.4., we discuss the economic environment in the experiment in more detail. Taken together, it represents a special case of the general theory with performance ranking uncertainty, an additive earnings function, (approximately) normal distributions, the same prior belief distribution for both workers, independence between the worker performances, and independence between worker performance and the random factor.

4.2 Spectator behavior

In the limited information situation, it follows from the worker performance and the random component being normally distributed and expected performance inequality approximation that the posterior belief of a Bayesian spectator is given by:

$$E(p_i|x_i) = \frac{\sigma_\varepsilon^2 \mu_p + \sigma_p^2 x_i}{\sigma_\varepsilon^2 + \sigma_p^2} = (1 - \rho_B) \cdot \mu_p + \rho_B \cdot x, \rho_B = \frac{1}{1 + A} \tag{13}$$

Equation (13) shows that a Bayesian spectator updates toward the signal (Chambers and Healey, 2012), where the weight attached to the earnings signal is determined by the variance in the distribution of performance relative to the variance in the random component. The expected performance of worker i given the earnings signal is the average of the mean performance and the earnings if the two distributions have the same variance ($A = 1$). However, if the variance in the random component is smaller than the variance in performance, the expected performance is closer to earnings than to the mean performance, and vice versa.

We chose to have approximately the same variance in the random component and the performance ($\sigma_\varepsilon = 15$ versus $\sigma_p = 17.22$, $A = 1.14$), such that Bayesian spectators can be clearly distinguished from both base-rate-neglecting and signal-neglecting spectators in their updating behavior. The experimental design implies that the Bayesian updating strength is $\rho_B = 0.56$.

Irrational spectators can be characterized by having an updating strength ρ that deviates from the Bayesian updating strength ρ_B, with signal-neglecting spectators assigning no weight to the earnings signal (ρ= 0) and base-rate-neglecting spectators only assigning weight to the earnings signal (ρ= 1).

The spectators may also differ in their fairness views. In the experimental analysis, we focus on the three most salient fairness views: egalitarianism, meritocracy, and libertarianism. The behavior of the egalitarian and libertarian spectators is not affected by the treatment manipulation because their fairness views do not depend on the source of the inequality. However, the meritocratic fairness view is sensitive to the spectator's beliefs about the source of the inequality.

Given expected performance inequality approximation, it follows that the optimal choice of a Bayesian meritocrat with limited information is given by:

$$y_i^{*BM}(LI) = (\alpha \cdot \frac{1}{2} + (1-\alpha) \cdot x_i^s) \cdot X \tag{14}$$

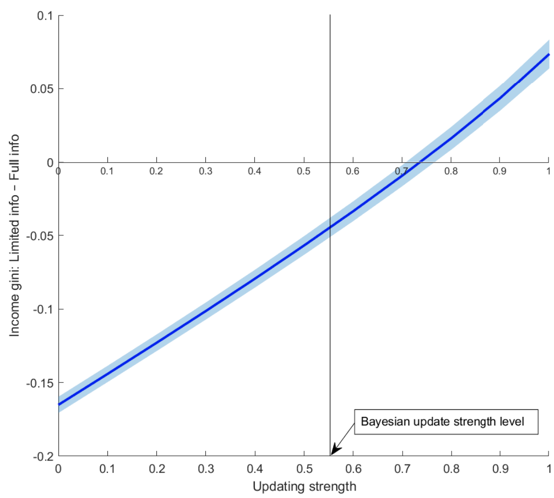
with $\alpha = \frac{2}{2+AB}$, $x_i^s = \frac{x_i}{X}$, and $B = \frac{X}{\mu_p}$.

Equation (14) shows that a Bayesian meritocrat gives to worker i a weighted average of the egalitarian solution and the libertarian solution. Signal-neglecting meritocratic spectators implement the egalitarian solution and base-rate-neglecting meritocratic spectators implement the libertarian solution. Hence, spectator behavior in the experiment depends on both the spectator's fairness view and their rationality in updating, which highlights that both sources of heterogeneity are potentially of great importance for inequality acceptance. Spectators with an egalitarian fairness view would not accept any inequality, whereas spectators with a libertarian fairness view would not redistribute at all. However, this difference in inequality acceptance may also arise among people who share the meritocrats fairness view, if some are signal neglecters and others are base-rate neglecters.

4.3 The treatment effect

We now consider how the treatment effect depends on the spectator's fairness views and their rationality. A large share of egalitarians and libertarians would pull toward no treatment difference in implemented inequality, whereas a large share of Bayesian meritocrats would pull toward less inequality being implemented with limited information. Irrational meritocratic spectators may introduce opposing forces on the treatment effect, depending on how they deviate from Bayesian updating.

Figure 1: Meritocrats: Updating strength and the treatment effect



Note: The figure shows the relationship between the updating strength ρ and the predicted treatment effect on implemented gini in the experiment. The light-blue shaded area indicates the 95 percent confidence intervals.

In Figure 1, we show the relationship between the updating strength ρ and the treatment effect on implemented inequality for meritocratic spectators in the experiment. We observe that the updating strength among mertiocratic spectators is of great importance for the treatment effect. The implemented gini would be 0.04 lower with limited information compared with full information if all spectators were Bayesian meritocrats, whereas it would be 0.07 greater if all spectators were base-rate-neglecting meritocratic spectators. The figure also shows that the difference in treatment effect for base-rate neglecting meritocrats and signal-neglecting meritocrats is very large (0.24 gini points).

In Figure B1 in Appendix B.1, we show that there is a corresponding positive relationship between updating strength and the treatment effect on implemented variance in inequality in the experiment, with Bayesian meritocratic spectators implementing less variance in inequality with limited information and base-rate-neglecting meritocratic spectators implementing more variance in inequality with limited information.

5 Experimental results

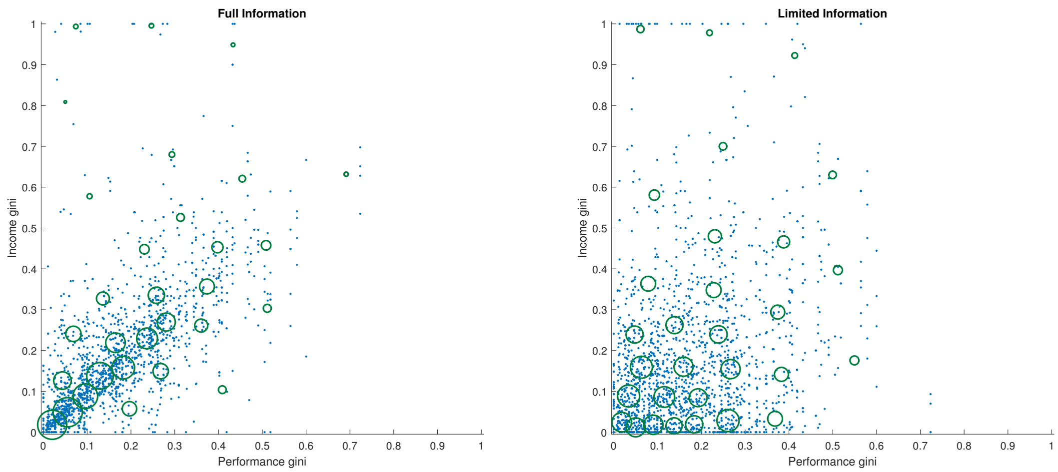
In this part we proceed in three steps. We first provide an overview of the spectator choices in the experiment, and then initial evidence on the role of beliefs in spectator behavior. Finally, we report regression analysis on how the treatment manipulation affected spectator behavior.

In Figure 2, we provide scatter plots showing how implemented income inequality relates to performance inequality in the worker pair (see Appendix B5 for additional descriptive statistics). Considering first the full information situations, we observe from the left panel that spectators divide income proportional to performance in the large majority of situations, in line with the theoretical framework assuming that meritocrats make a distinction between performance and luck. In the large majority of the full information situations, we observe that the spectators implement an income inequality that is strictly smaller than the earnings inequality. However, in 29.14 percent of the situations, the spectators implement an income inequality greater than the earnings inequality, which would be in line with the meritocratic fairness view if the worker with greater performance has been very unlucky in the random draw.

Moving to the limited information situations, we observe in the right panel of Figure 2 that there is a much weaker relationship between income inequality and performance inequality when the spectators have limited information and need to infer performance from the earnings signal.[9]

[9] In Appendix B.7 we also show that under limited information, a great majority of the

Figure 2: Income inequality on performance inequality



Note: The figure shows the relationship between income inequality and performance inequality, by treatment.

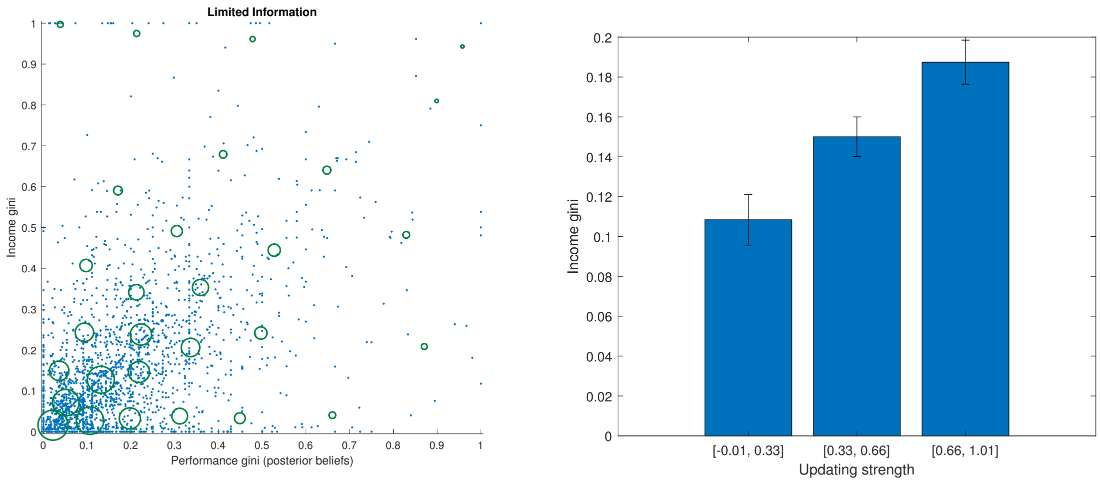
We now provide initial evidence on the role of beliefs in the spectator choices in the Limited Information treatment, using the posterior beliefs that we elicited at the end of the experiment. It follows from (13) that we can infer the updating strength from the elicited posterior beliefs as follows:

$$\varrho = \frac{E(p_i|x_i) - E(p_j|x_j)}{x_i - x_j} \tag{15}$$

In the left panel of Figure 3, we show the relationship between the believed performance inequality in the worker pair and the implemented income inequality. We observe that the spectators in the large majority of situations, conditional on their posterior beliefs, implement (approximately) the meritocratic distribution, which suggests that the meritocratic fairness view also is prominent when there is limited information. In the right panel of Figure 3, we consider the relationship between updating strength and implemented inequality for the spectators who consistently implement the meritocratic fairness view in the Limited Information treatment.

spectators (83.84 percent) implement an income inequality within the range of the egalitarian solution and the libertarian solution, as predicted by the theoretical framework.

Figure 3: Posterior beliefs and implemented income inequality



Note: The left panel shows a scatterplot of the spectators' believed performance inequality in the worker pair (based on the elicited posterior beliefs) and implemented income inequality, for the situations in the Limited Information treatment. The right panel shows the average implemented inequality for the meritocratic spectators in the Limited Information treatment, for three intervals of average updating strength. A spectator is classified as meritocrat in the right panel if $\left|\frac{y_i}{y_i+y_j} - \frac{E(p_i|x_i)}{E(p_i|x_i)+E(p_j|x_j)}\right| < 0.1$ *in at least six of the ten decisions, where* $E(p_i|x_i)$ *is inferred from the spectator's elicited posterior beliefs. 86.1 percent of the spectators are classified as meritocrats according to this procedure, with 26.8 percent having an average updating strength between -0.01 and 0.33, 38.1 percent between 0.33 and 0,66, and 38.1 percent between 0.66 to 1.01. Standard error bars are indicated.*

In line with the theoretical analysis, we observe that there is a strong positive relationship between updating strength and implemented inequality for meritocratic spectators: meritocratic spectators who underreact to the signal implement less inequality and meritocratic spectators who overreact to the signal implement more inequality.[10] We now turn to the regression analysis of how the treatment manipulation affected the spectator choices. In Table 1, we report the regression estimates from the following empirical specification:

$$I_{ij} = \alpha_i + \beta \cdot LI_i + \gamma \cdot X_i + \varepsilon_{ij} \tag{16}$$

[10]In Figure B3 in Appendix B.4, we show that this relationship does not reflect that there is a correlation between updating strength and the nature of the distributive situations. Bayesian meritocratic spectators would implement almost the same level of inequality across the distributive situations of the three groups shown in the right panel of Figure 3.

where I_{ij} is the inequality implemented by spectator i in situation j, LI_i is an indicator variable for whether spectator i was assigned to the Limited Information treatment, X_i is a vector of background characteristics for spectator i, and ε_{ij} is an i.i.d. error term. We report regression estimates both with and without the background characteristics, and for regressions where the dependent variable is an indicator variable for the spectator implementing the egalitarian solution or the libertarian solution.

In columns (1) and (2) of Table 1, we observe that there is no significant average treatment effect on implemented inequality, and the estimated treatment effect is significantly smaller than what would be the treatment effect if all spectators were Bayesian meritocrats ($p = 0.0025$). In columns (3) and (4), we observe that there is a significant increase in the share of spectators implementing the egalitarian and libertarian solutions, which, given the theoretical analysis, is suggestive of some spectators being signal-neglecting meritocrats and some spectators being base-rate neglecting meritocrats.

Table 1: Regression analysis of treatment effects

	Implemented gini	Implemented gini	Equal division	No redistribution
LI	-0.0107	-0.0107	0.0588	0.1020
	(0.0096)	(0.0096)	(0.0131)	(0.0303)
Female		0.0098	-0.0089	-0.0409
		(0.0093)	(0.0135)	(0.0327)
Age		0.0006	0.0006	0.0140
		(0.0017)	(0.0027)	(0.0082)
Right-wing		0.0020	0.0099	0.0520
		(0.0102)	(0.0131)	(0.0298)
	$N = 4,240$	$N = 4,240$	$N = 4,240$	$N = 4,240$

Standard errors clustered at the individual level (425 clusters)

Notes: The table reports ordinary least square estimates for the regression of implemented inequality on a treatment indicator (LI), without (column (1)) and with a set of background characteristics (column (2)). In columns (3) and (4), we report the corresponding regressions with the dependent variable being an indicator for whether the spectator implemented the egalitarian solution or the libertarian solution. Female is an indicator variable for whether the participant is a woman, Age is the age of the participant in years, and Right-wing is an indicator variable for whether the participants self-reported voting for the Conservative Party or the Progress Party.

Taken together, the descriptive evidence and the regression analysis suggest that a large majority of spectators are meritocrats but differ significantly in their belief updating. To investigate further how heterogeneity in fairness views and belief updating shape spectator behavior, we now turn to a structural analysis.

6 Structural model

We here provide structural estimates of the utility model introduced in the theoretical analysis. We use a random utility framework with a random component added to the decision utility V:

$$V(y_i, \cdot) = \beta U_{spectator}(y_i, \cdot) + \varepsilon_i \tag{17}$$

We assume that ε_i is i.i.d. extreme value distributed and that each spectator is characterized by a fairness view m_i (egalitarian, meritocrat, libertarian), a weight β assigned to the deterministic component of the utility function, and an updating strength ρ.

We estimate two types of models: a model where we assume that all the spectators are Bayesian updaters (rational model) and a model where we allow for non-Bayesian updating (behavioral model). In both models, we estimate a population distribution of the share λ^h of the different fairness types and a log-normal population distribution of β characterized by the parameter set $\theta_\beta = \left\{\zeta_\beta, \sigma_\beta^2\right\}$. In the rational model, all spectators are assigned the Bayesian updating strength $\rho_B = 0.56$; in the behavioral model, we estimate a normal population distribution of ρ characterized by the parameter set $\theta_\rho = \{\mu_\rho, \sigma_\rho, \theta_{signal\, neglect}, \theta_{base-rate\, neglect}\}$, which allows for mass points for signal neglecters ($\theta_{signal\, neglect}$) and base-rate neglecters ($\theta_{base-rate\, neglect}$). We do not restrict the updating parameter in the behavioral model to be between 0 and 1, but allow it to range from minus infinity to infinity.

Given these modeling assumptions, we can write the likelihood of the observed spectator behavior for a spectator with fairness view m_h as:

Table 2: Structural analysis

	Full Information treatment	Rational model	Behavioral model
$\lambda^{Meritocrats}$	81.05% (3.04%)	64.82% (2.58%)	81.22% (2.87%)
$\lambda^{Egalitarians}$	4.34% (1.76%)	11.18% (1.71%)	3.87% (1.28%)
$\lambda^{Libertarians}$	14.6% (2.66%)	24.00% (2.25%)	14.91% (2.68%)
ζ_{β}	-3.6351 (0.1064)	-3.6420 (0.0968)	-3.0636 (0.1093)
σ_{β}	1.8738 (0.0622)	2.2278 (0.0893)	2.8841 (0.0855)
μ_{ρ}			0.4678 (0.0234)
σ_{ρ}			0.1842 (0.0216)
$\theta_{signal\,neglect}$			0.0993 (0.0403)
$\theta_{base-rate\,neglect}$			0.2864 (0.0916)
Log likelihood		-11,956	-11,783
Log likelihood FI	-5,867	-5,891.1	-5,903
Log likelihood LI		-6,064.6	-5,879.8

Notes: The table shows the estimated parameters from the structural analysis. In the first column, the model is estimated on the Full Information treatment sample only; in the second and third column, the rational model and the behavioral model are estimated on the full sample.

$$L^h = \int_{-\infty}^{\infty} \int_{0}^{\infty} \left(\prod_{k=1}^{10} \frac{e^{V(y_{ik}, m_h, \beta, \rho)}}{\sum_{y_{jk} \in 0,1,..,X} e^{V(y_{jk}, m_h, \beta, \rho)}} \right) dF\left(\theta_\beta\right) dG\left(\theta_\rho\right) \tag{18}$$

with y_{ik} indicating the income given by the spectator to worker i in decision k. The total likelihood contribution of a spectator is now given by:

$$L = \sum_h \lambda^h \cdot L^h \tag{19}$$

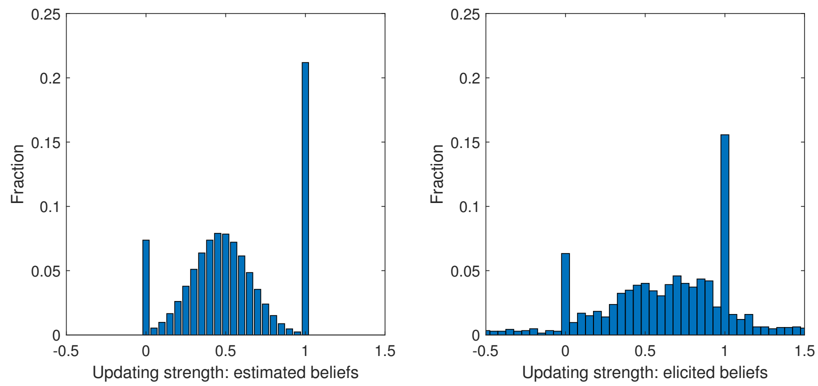
Both the rational model and the behavioral model are estimated on the full sample, where we assume that the fairness type m_h and weight β are independent of treatment. In the Full Information treatment, the spectators know the performance of the workers; in the Limited Information treatment, they update their beliefs about performance with the updating strength ρ given the earnings signal. We also estimate the model separately on the Full Information treatment sample; in this estimation, there is no difference between the rational model and the behavioral model.

In Table 2, we report the estimates from the structural analysis. In column (1), we observe that the estimated share of meritocrats is 81.1 percent when only using the sample of spectators in the Full Information treatment, and the estimated shares of egalitarians and libertarians are 4.3 percent and 14.6 percent. The estimated shares of fairness types differ significantly when estimating the rational model on the full sample (column (2)), with a much smaller share of meritiocrats and larger shares of egalitarians and libertarians. However, when estimating the behavioral model on the full sample (column (3)), the estimated shares are unchanged and in line with the initial evidence that we reported in the previous section. The behavioral model estimates less disagreement in fairness views than the rational model.

In terms of belief updating, the behavioral model estimates a median updating strength close to Bayesian updating, but also significant mass on signal neglecting and base-rate neglecting. In Figure 4, we report the estimated distribution of updating strength from the choice data using the behavioral model (left panel) and the inferred distribution of updating strength from the elicited posterior beliefs (right panel). We observe that the distributions are strikingly similar. Both distributions are largely between 0 and 1, have median updating strength close to Bayesian updating, and a significant mass at base-rate neglect and signal neglect, with the mass at base-rate neglect being particularly pronounced.

Taken together, the structural estimates highlight that the rational model fits the data by estimating significant heterogeneity in fairness views, whereas the behavioral model fits the data by estimating significant heterogeneity in belief updating. Consequently, the two models provide very different explanations for why we observe a smaller treatment effect than predicted if all spectators were Bayesian meritocrats.

Figure 4: Updating strength: estimated versus elicited updating strength

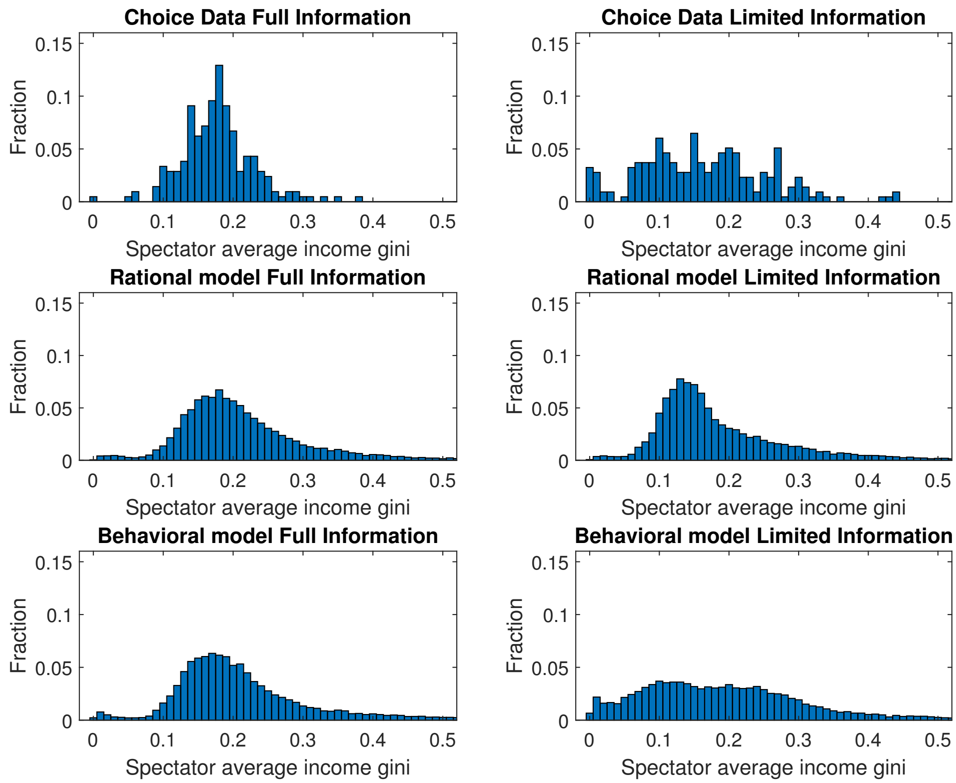


Notes: The figure reports the estimated distribution of updating strength ρ for the structural behavioral model (left panel) and the inferred updating strength for the posterior beliefs of the spectators across situations (right panel).

The rational model suggests that a large share of the spectators are not meritocrats and therefore not affected by the treatment manipulation; the behavioral model suggests that a large share of the spectators are meritocrats but not Bayesian, with base-rate neglecting meritocrats implementing more inequality with limited information than with full information, counteracting the effect of limited information on Bayesian (and signal-neglecting) meritocrats.

In Figure 5, we compare the rational model and the behavioral model using simulation results. In the upper two panels, we show the distribution of the average implemented gini per spectator in the Full Information treatment (left) and in the Limited Information treatment (right), and compare them with the simulated distributions for the rational model (middle panel) and the behavioral model (lower panel). We observe that both models fit the choice data from the Full Information treatment reasonably well, but the behavioral model fits the choice data much better than the rational model for the Limited Information treatment.

Figure 5: Choice data versus predicted data



Notes: The upper two panels show the distribution of the the average gini implemented per spectator in the experiment for the Full Information treatment (left) and the Limited Information treatment (right). The middle and lower two panels show the simulated distributions for the rational model and the behavioral model, in each case based on 100 iterations.

The two models are nested, with four extra degrees of freedom in the behavioral model. A likelihood ratio test accounting for this additional freedom clearly rejects that the rational model performs equally well as the behavioral model (p<0.0001). In Appendix B.3, we show that the findings from the structural analysis are robust to allowing the spectator's fairness view to vary between treatments and to restricting the updating strength to be between 0 and 1. The behavioral model estimated in Table 2 performs better than all the alternative model specifications.

In sum, the structural analysis provides evidence suggesting that heterogeneity in belief updating may be as important as heterogeneity in fairness views in explaining why people differ in their moral acceptability of inequality acceptance.

7 Conclusion

In this study, we have investigated theoretically and empirically how limited information about the source of inequality affects inequality acceptance. In particular, we have shown how heterogeneity in belief updating may cause people who share the same meritocratic fairness view to disagree strongly about whether an inequality is fair or unfair when there is limited information: signal-neglecting meritocrats may act as if they are egalitarians and base-rate-neglecting meritocrats may act as if they are libertarians.

We have also shown that the overall effect of limited information on inequality acceptance depends on the extent to which people update rationally. Limited information makes Bayesian meritocrats implement less inequality as long as there is uncertainty about who deserves a higher income (fairness-ranking uncertainty), but base-rate neglect may create a counteracting effect and imply greater inequality acceptance with limited information than with full information. These insights have been established in a general theoretical framework, and are shown to be important in explaining spectator behavior in a controlled experimental study.

The findings in the present paper are relevant to a core issue in the political economy of redistribution, where voters' beliefs about the relative importance of performance and luck in creating inequality in society are central (Piketty (1995); Alesina and Angeletos (2005)). A common assumption in much of this literature has been that people update rationally, but our experimental findings show that there is great heterogeneity in belief updating that may affect the support for redistribution. Hence, differences in beliefs about the source of inequality may reflect not only different experiences, or a self-serving bias (Bénabou and Tirole, 2006), but also that people differ in how they handle limited information. An interesting avenue for future research would be to incorporate heterogeneity in both belief updating and fairness views in political economy models of redistribution.

The insights are also relevant in a number of other areas in economics, including, for example, for understanding workplace inequality (Akerlof and Yellen (1990), Roberts and Milgrom (1992), and Abeler et al. (2011)). The extent to which workplace inequality is considered acceptable is likely to depend on whether the inequality is considered to reflect differences in luck or performance, and the present study shows that irrational updating has the potential to generate disagreements about the fairness of workplace inequality even if all parties share the same fairness views and have the same available information.

The meritocratic fairness view is prominent and powerful in society, but it is also a source of disagreement. It requires information about the relative importance of performance and luck in shaping people's lives (Moffitt, 2015), and we have shown that beliefs about the role of performance and luck may reflect irrational considerations. Thus, the present heated debate on inequality in many societies may not only reflect a fundamental disagreement about what constitutes a fair distribution, but also that people react differently when they have limited information about the nature of the inequality.

References

Abeler, J., A. Falk, L. Götte, and D. Huffman (2011). Reference points and effort provision. *American Economic Review 101*(2), 470–492.

Akbaş, M., D. Ariely, and S. Yuksel (2019). When is inequality fair? an experiment on the effect of procedural justice and agency. *Journal of Economic Behavior & Organization 161*, 114–127.

Akerlof, G. A. and J. L. Yellen (1990). The fair wage-effort hypothesis and unemployment. *The Quarterly Journal of Economics 105*(2), 255–283.

Alesina, A. and G.-M. Angeletos (2005). Fairness and redistribution. *American Economic Review 95*(4), 960–980.

Alesina, A. and P. Giuliano (2011). Preferences for redistribution. In J. Benhabib, A. Bisin, and M. O. Jackson (Eds.), *Handbook of Social Economics*, Volume 1, Chapter 4, pp. 99–131. Amsterdam, NL: Elsevier.

Alesina, A., E. L. Glaeser, and B. Sacerdote (2001). Why doesn't the United States have a European-style welfare state? *Brookings Papers on Economic Activity 2001*(1), 187–254.

Almås, I., A. W. Cappelen, and B. Tungodden (2020). Cutthroat capitalism versus cuddly socialism: Are Americans more meritocratic and efficiency-seeking than Scandinavians? *Journal of Political Economy 128*(5), 1753–1788.

Andre, P. (2021). Shallow meritocracy: An experiment on fairness views. *Working paper, available at SSRN 3916303*.

Andreoni, J. and J. Miller (2002). Giving according to GARP: An experimental test of the consistency of preferences for altruism. *Econometrica 70*(2), 737–753.

Balafoutas, L., M. G. Kocher, L. Putterman, and M. Sutter (2013). Equality, equity and incentives: An experiment. *European Economic Review 60*, 32–51.

Barr, A., L. Miller, and P. Ubeda (2021). Is the acknowledgment of earned entitlement effect robust across experimental modes and populations? *Sociological Methods & Research*, https://doi.org/10.1177/0049124120986194.

Bartling, B., R. A. Weber, and L. Yao (2015). Do markets erode social responsibility? *The Quarterly Journal of Economics 130*(1), 219–266.

Bellemare, C., S. Kröger, and A. Van Soest (2008). Measuring inequity aversion in a heterogeneous population using experimental decisions and subjective probabilities. *Econometrica 76*(4), 815–839.

Bénabou, R. and J. Tirole (2006). Belief in a just world and redistributive politics. *The Quarterly Journal of Economics 12*(2), 699–746.

Benjamin, D., A. Bodoh-Creed, and M. Rabin (2019). Base-rate neglect: Foundations and implications. Technical report.

Benjamin, D. J. (2019). Errors in probabilistic reasoning and judgment biases. In B. D. Bernheim, S. DellaVigna, and D. Laibson (Eds.), *Handbook of Behavioral Economics - Foundations and Applications 1*, Chapter 2, pp. 69 – 186. North-Holland.

Benndorf, V., H. A. Rau, and C. Sölch (2014). Minimizing learning behavior in experiments with repeated real-effort tasks. *Available at SSRN 2503029*.

Bolton, G. E. and A. Ockenfels (2000). ERC: A theory of equity, reciprocity, and competition. *American Economic Review 90*(1), 166–193.

Bortolotti, S., I. Soraperra, M. Sutter, and C. Zoller (2017). Too lucky to be true. Fairness views under the shadow of cheating. CESifo Working Paper Series No. 6563.

Cappelen, A. and B. Tungodden (2019). *The economics of fairness*. Edward Elgar Publishing.

Cappelen, A. W., A. Drange Hole, E. Ø. Sørensen, and B. Tungodden (2007). The pluralism of fairness ideals: An experimental approach. *American Economic Review 97*(3), 818–827.

Cappelen, A. W., J. Konow, E. Ø. Sørensen, and B. Tungodden (2013). Just luck: An experimental study of risk-taking and fairness. *American Economic Review 103*(4), 1398–1413.

Cappelen, A. W., J. Mollerstrom, B.-A. Reme, and B. Tungodden (2022). A meritocratic origin of egalitarian behaviour. *The Economic Journal 132*(646), 2101–2117.

Cassar, L. and A. H. Klein (2019). A matter of perspective: How failure shapes distributive preferences. *Management Science 65*(11), 5050–5064.

Cettolin, E. and A. Riedl (2017). Justice under uncertainty. *Management Science 63*(11), 3739–3759.

Cettolin, E., A. Riedl, and G. Tran (2017, December). Giving in The Face of Risk. *Journal of Risk and Uncertainty 55*(2-3), 95–118.

Chambers, C. P. and P. J. Healy (2012). Updating toward the signal. *Economic Theory 50*(3), 765–786.

Charness, G. and M. Rabin (2002). Understanding social preferences with simple tests. *The Quarterly Journal of Economics 117*(3), 817–869.

Chavanne, D. (2018). Headwinds, tailwinds, and preferences for income redistribution. *Social Science Quarterly 99*(3), 851–871.

Cruces, G., R. Perez-Truglia, and M. Tetaz (2013). Biased perceptions of income distribution and preferences for redistribution: Evidence from a survey experiment. *Journal of Public Economics 98*, 100–112.

Davidai, S. and T. Gilovich (2016). The headwinds/tailwinds asymmetry: An availability bias in assessments of barriers and blessings. *Journal of Personality and Social Psychology 111*(6), 835.

DellaVigna, S. (2018). Structural behavioral economics. In *Handbook of Behavioral Economics: Applications and Foundations 1*, Volume 1, pp. 613–723. Elsevier.

Di Tella, R., R. Perez-Truglia, A. Babino, and M. Sigman (2015). Conveniently upset: avoiding altruism by distorting beliefs about others' altruism. *American Economic Review 105*(11), 3416–42.

Durante, R., L. Putterman, and J. Weele (2014). Preferences for redistribution and perception of fairness: An experimental study. *Journal of the European Economic Association 12*(4), 1059–1086.

Engelmann, D. and M. Strobel (2004). Inequality aversion, efficiency, and maximin preferences in simple distribution experiments. *American Economic Review 94*(4), 857–869.

Enke, B. (2020). What you see is all there is. *The Quarterly Journal of Economics 135*(3), 1363–1398.

Enke, B. and F. Zimmermann (2019). Correlation neglect in belief formation. *The Review of Economic Studies 86*(1), 313–332.

Erkal, N., L. Gangadharan, and B. H. Koh (2022). By chance or by choice? Biased attribution of others outcomes when social preferences matter. *Experimental Economics 25*(2), 413–443.

Exley, C. and J. B. Kessler (2019). Equity concerns are narrowly framed. *Harvard Business School Research Paper Series* (18-040).

Exley, C. L. (2016). Excusing selfishness in charitable giving: The role of risk. *The Review of Economic Studies 83*(2), 587 – 628.

Fehr, D. and M. Vollmann (2020). Misperceiving economic success: Experimental evidence on meritocratic beliefs and inequality acceptance. Technical report, AWI Discussion Paper Series.

Fehr, E., G. Kirchsteiger, and A. Riedl (1993). Does fairness prevent market clearing? An experimental investigation. *The Quarterly Journal of Economics 108*(2), 437–459.

Fehr, E. and K. M. Schmidt (1999). A theory of fairness, competition and cooperation. *The Quarterly Journal of Economics 114*(3), 817–868.

Fong, C. (2001). Social preferences, self-interest, and the demand for redistribution. *Journal of Public Economics 82*(2), 225–246.

Frank, R. H. (2016). *Success and Luck: Good Fortune and the Myth of Meritocracy.* Princeton, NJ: Princeton University Press.

Fudenberg, D. and D. K. Levine (2012). Fairness, risk preferences and independence: Impossibility theorems. *Journal of Economic Behavior and Organization 81*(2), 606–612.

Gächter, S., B. Herrmann, and C. Thoeni (2010). Culture and cooperation. *Philosophical Transactions of the Royal Society B: Biological Sciences 365*(1553), 2651–2661.

Gross, T., C. Guo, and G. Charness (2015). Merit pay and wage compression with productivity differences and uncertainty. *Journal of Economic Behavior & Organization 117*, 233–247.

Konow, J. (1996). A positive theory of economic fairness. *Journal of Economic Behavior and Organization 31*(1), 13–35.

Konow, J. (2000). Fair shares: Accountability and cognitive dissonance in allocation decisions. *American Economic Review 90*(4), 1072–1091.

Konow, J. (2009). Is fairness in the eye of the beholder? An impartial spectator analysis of justice. *Social Choice and Welfare 33*(1), 101–127.

Konow, J. (2012). Adam Smith and the modern science of ethics. *Economics & Philosophy 28*(3), 333–362.

Krawczyk, M. (2010). A glimpse through the veil of ignorance: Equality of opportunity and support for redistribution. *Journal of Public Economics 94*(1-2), 131–141.

Langer, E. J. (1975). The illusion of control. *Journal of Personality and Social Psychology 32(2)*, 311–328.

Mankiw, N. G. (2013). Defending the one percent. *Journal of Economic Perspectives 27*(3), 21–34.

Moffitt, R. A. (2015). The deserving poor, the family, and the U.S. welfare system. *Demography 52*(3), 729–749.

Mollerstrom, J., B.-A. Reme, and E. Ø. Sørensen (2015). Luck, choice and responsibility: An experimental study of fairness views. *Journal of Public Economics 131*, 33–40.

Müller, D. and S. Renes (2021). Fairness views and political preferences: evidence from a large and heterogeneous sample. *Social Choice and Welfare 56*(4), 679–711.

Piketty, T. (1995). Social mobility and redistributive politics. *The Quarterly Journal of Economics 110*(3), 551–584.

Piketty, T. (2020). *Capital and ideology.* Harvard University Press.

Rabin, M. (1993). Incorporating fairness into game theory and economics. *American Economic Review 83*(5), 1281–1302.

Roberts, J. and P. Milgrom (1992). *Economics, organization and management.* Prentice-Hall Englewood Cliffs, NJ.

Rodriguez-Lara, I. and L. Moreno-Garrido (2012). Self-interest and fairness: Self-serving choices of justice principles. *Experimental Economics 15*(1), 158–175.

Saito, K. (2013). Social preferences under risk: Equality of opportunity versus equality of outcome. *American Economic Review 103*(7), 3084–3101.

Sandel, M. J. (2020). *The tyranny of merit: What's become of the common good?* Penguin UK.

Savage, L. J. (1971). Elicitation of personal probabilities and expectations. *Journal of the American Statistical Association 66*(336), 783–801.

Stantcheva, S. (2021). Understanding tax policy: How do people reason? *The Quarterly Journal of Economics 136*(4), 2309–2369.

Sugden, R. and M. Wang (2020). Equality of opportunity and the acceptability of outcome inequality. *European Economic Review 130.*

Valero, V. (2022). Redistribution and beliefs about the source of income inequality. *Experimental Economics 25*(3), 876–901.

ONLINE APPENDIX

A Appendix: Supplementary theoretical analysis

A.1 Optimal spectator behavior

We here derive the optimal spectator choice and optimal inequality in a limited information situation.

The spectator maximizes the utility function:

$$EU_{spectator} = -E(y_i - m_i)^2 = -\int_{m_{min}}^{m_{max}} f(m_i) \cdot (y_i - m_i)^2 dm_i$$

The first order condition is given by:

$$\frac{\partial U_i}{\partial y_i} = -\frac{\partial}{\partial y_i}\left(\int_{m_{min}}^{m_{max}} f(m_i) \cdot (y_i - m_i)^2 dm_i\right) = 0$$

From the Leibniz rule for differentiation of an integral, it follows that:

$$\frac{\partial U_i}{\partial y_i} = -\int_{m_{min}}^{m_{max}} \frac{\partial}{\partial y_i} f(m_i) \cdot (y_i - m_i)^2 dm_i = -\int_{m_{min}}^{m_{max}} 2 \cdot f(m_i) \cdot (y_i - m_i)\, dm_i = 2 \cdot (y_i - E(m_i)) = 0$$

Hence, it follows that the optimal income to worker i is given by:

$$y_i^{*LI} = E(m_i) = \int f^{posterior}(m_i|\mathbf{x}) m_i dm_i$$

It now follows that the optimal inequality is given by:

$$I_i^{*LI} = 2|\frac{E(m_i)}{X} - \frac{1}{2}|.$$

Taking into account that,

$$2\cdot\left(\frac{E\left(m_i\right)}{X}-\frac{1}{2}\right)=2\cdot\left(\frac{\int f^{posterior}(m_i|\mathbf{x})m_i\,dm_i}{X}-\frac{1}{2}\right)=\int f^{posterior}(m_i|\mathbf{x})\cdot\left(2\cdot\left(\frac{m_i}{X}-\frac{1}{2}\right)\right)dm_i$$

$$=\int f^{posterior}(m_i|\mathbf{x})\cdot\zeta\left(m_i\right)\,dm_i,$$

it follows that:

$$I_i^{*LI}=|\int f^{posterior}(m_i|\mathbf{x})\cdot\zeta\left(m_i\right)\,dm_i|$$

Proposition 1 can be illustrated by considering the behavior of a Bayesian meritocratic spectator. First, consider an economic environment where the underlying performance distribution contains two equally likely performance pairs, $(30, 10)$ and $(10, 30)$. In this case, with full information, the inequality implemented across situations by a Bayesian meritocratic spectator is given by: $E^{FI}\left(\left|\frac{y_i^*-y_j^*}{y_i^*+y_j^*}\right|\right)=E\left(\left|\frac{p_i-p_j}{p_i+p_j}\right|\right)=\frac{1}{2}\cdot\left|\frac{30-10}{30+10}\right|+\frac{1}{2}\cdot\left|\frac{10-30}{30+10}\right|=\frac{1}{2}$. Compare this to the limited information situation where the Bayesian meritocratic spectator, after observing the earnings signals, holds the posterior beliefs that the performance pairs $(30, 10)$ and $(10, 30)$ are equally likely for all worker pairs (which means that the posterior beliefs imply the same underlying performance distribution, in line with Lemma 1). In this case, limited information causes fairness-ranking uncertainty and the Bayesian meritocratic spectator implements less inequality in each situation than with full information: $E^{LI}\left(\left|\frac{y_i^*-y_j^*}{y_i^*+y_j^*}\right|\right)=E\left(\left|E\left(\frac{p_i-p_j}{p_i+p_j}|\mathbf{x}\right)\right|\right)=\frac{1}{2}\cdot\left|\frac{1}{2}\cdot\frac{30-10}{30+10}+\frac{1}{2}\cdot\frac{10-30}{30+10}\right|+\frac{1}{2}\cdot\left|\frac{1}{2}\cdot\frac{30-10}{30+10}+\frac{1}{2}\cdot\frac{10-30}{30+10}\right|=0$. Consequently, the Bayesian meritocratic spectator implements less inequality with limited information than with full information across situations. Second, consider the economic environment where the only difference to the example above is that the two equally likely performance pairs are $(30, 10)$ and $(90, 10)$, which implies that, with full information, the inequality implemented across situations by a Bayesian meritocratic spectator is given by: $E^{FI}\left(\left|\frac{y_i^*-y_j^*}{y_i^*+y_j^*}\right|\right)=E\left(\left|\frac{p_i-p_j}{p_i+p_j}\right|\right)=\frac{1}{2}\cdot\left|\frac{30-10}{30+10}\right|+\frac{1}{2}\cdot\left|\frac{90-10}{90+10}\right|=\frac{13}{20}$. Assume that the Bayesian meritocratic spectator with limited information, after observing the earnings signals, holds the posterior beliefs that the performance pairs $(30, 10)$ and $(90, 10)$ are equally likely for all worker pairs. In this case, there is no fairness-ranking uncertainty and the expected implemented inequality with limited information in each situation is equal to the expected optimal inequality across situations: $E^{LI}\left(\left|\frac{y_i^*-y_j^*}{y_i^*+y_j^*}\right|\right) = E\left(\left|E\left(\frac{p_i-p_j}{p_i+p_j}|\mathbf{x}\right)\right|\right) = \left|\frac{1}{2}\cdot\frac{30-10}{30+10}+\frac{1}{2}\cdot\frac{90-10}{90+10}\right| = \frac{13}{20}$. Consequently, the Bayesian meritocratic spectator implements the same level of inequality with limited information as with full information across situations.

A.2 Signal-neglecting spectators

We here show generally when a signal-neglecting meritocratic spectator implements less inequality than a Bayesian meritocratic spectator.

We first introduce the following definition of income-ranking uncertainty:

Income-ranking uncertainty: The spectator's posterior beliefs reflect income-ranking uncertainty if there exist two signals $\mathbf{x}$ and $\mathbf{x}'$ such that

$$\int f^{posterior}(m_i|\mathbf{x})\cdot\zeta\left(m_i\right)\ dm_i > 0$$

and

$$\int f^{posterior}(m_i|\mathbf{x}')\cdot\zeta\left(m_i\right)\ dm_i < 0$$

We can now state the following result:

Proposition: *Signal-neglecting meritocratic spectators implement strictly less inequality with limited information than a Bayesian meritocratic spectator if and only if the prior distribution f and signaling technology s together exhibit income-ranking uncertainty.*

Proof: It follows from (6) and the fact that $f^{SNposterior}(m_i|\mathbf{x}) = f^{prior}(m_i)$ that optimal inequality for a signal neglecter in any given situation with limited information is given by:

$$E^{SN}\left(I^{*LI}\right)=\left|\int f^{prior}(m_i)\cdot\zeta\left(m_i\right)\ dm_i\right| \tag{20}$$

Further, we know from Bayes rule that:

$$\int f^{Bposterior}(m_i|\mathbf{x})\cdot g\left(\mathbf{x}\right)d\mathbf{x}=f^{prior}(m_i) \tag{21}$$

Hence, it follows that that optimal inequality across situations is given by:

$$E^{SN}\left(I^{*LI}\right)=\left|\int\int f^{Bposterior}(m_i|\mathbf{x})\cdot\zeta\left(m_i\right)\ dm_i\cdot g\left(\mathbf{x}\right)d\mathbf{x}\right|$$

$$\leq\int\left|\int f^{Bposterior}(m_i|\mathbf{x})\cdot\zeta\left(m_i\right)\ dm_i\right|\cdot g\left(\mathbf{x}\right)d\mathbf{x}$$

$$=E^{Bayesian}\left(I^{*LI}\right)$$

The result follows from income-ranking uncertainty implying strict inequality (if-part) and the absence of income-ranking uncertainty implying equality (only-if part).■

A.3 Bayesian spectators and limited information: variance in inequality

We here establish the following result:

Proposition*: A Bayesian meritocrat implements in expectation lower variance in inequality with limited information then with full information if limited information does not cause performance-ranking uncertainty.*

Proof: (i) In a limited information situation, the variance in implemented inequality for a Bayesian meritocratic spectator is equal to:

$$Var^{LI}\left(\left|\frac{y_i^*-y_j^*}{y_i^*+y_j^*}\right|\right)=Var\left(\left|E\left(\zeta\left(p_i,p_j\right)|\mathbf{x}\right)\right|\right)=\int\left(\left|E\left(\zeta\left(p_i,p_j\right)\mathbf{x}\right)\right|-E\left(\left|E\left(\zeta\left(p_i,p_j\right)|\mathbf{x}\right)\right|\right)\right)^2\cdot g\left(\mathbf{x}\right)d\mathbf{x}$$

If there is no performance-ranking uncertainty, then it follows from Proposition 1 that $E\left(\left|E\left(\zeta\left(p_i,p_j\right)|\mathbf{x}\right)\right|\right)=E^{LI}\left(\left|\frac{y_i^*-y_j^*}{y_i^*+y_j^*}\right|\right)=E^{FI}\left(\left|\frac{y_i^*-y_j^*}{y_i^*+y_j^*}\right|\right)=I.$

This implies that:

$$\int\left(\left|E\left(\zeta\left(p_i,p_j\right)|\mathbf{x}\right)\right|-E\left(\left|E\left(\zeta\left(p_i,p_j\right)|\mathbf{x}\right)\right|\right)\right)^2\cdot g\left(\mathbf{x}\right)d\mathbf{x}=\int\left(\left|E\left(\zeta\left(p_i,p_j\right)|\mathbf{x}\right)\right|-I\right)^2\cdot g\left(\mathbf{x}\right)d\mathbf{x}$$

By rearranging, we have that:

$$\int\left(\left|E\left(\zeta\left(p_i,p_j\right)|\mathbf{x}\right)\right|-I\right)^2\cdot g\left(\mathbf{x}\right)d\mathbf{x}=\int\left(\left|\int f^{posterior}\left(\mathbf{p}|\mathbf{x}\right)\cdot\zeta\left(p_i,p_j\right)d\mathbf{p}\right|-I\right)^2\cdot g\left(\mathbf{x}\right)d\mathbf{x}$$

Given the assumption of no performance-ranking uncertainty, it follows that:

$$\int\left(\left|\int f^{posterior}\left(\mathbf{p}|\mathbf{x}\right)\cdot\zeta\left(p_i,p_j\right)d\mathbf{p}\right|-I\right)^2\cdot g\left(\mathbf{x}\right)d\mathbf{x}$$

$$=\int\left(\int f^{posterior}\left(\mathbf{p}|\mathbf{x}\right)\cdot\left|\zeta\left(p_i,p_j\right)\right|d\mathbf{p}-I\right)^2\cdot g\left(\mathbf{x}\right)d\mathbf{x}$$

Hence, we have that

$$Var^{LI}\left(\left|\frac{y_i^*-y_j^*}{y_i^*+y_j^*}\right|\right)=\int\left(\int f^{posterior}\left(\mathbf{p}|\mathbf{x}\right)\cdot\left|\zeta\left(p_i,p_j\right)\right|d\mathbf{p}-I\right)^2\cdot g\left(\mathbf{x}\right)d\mathbf{x}$$

(ii) In a full information situation, the variance in implemented inequality for a Bayesian meritocratic spectator is equal to:

$$Var^{FI}\left(\left|\frac{y_i^*-y_j^*}{y_i^*+y_j^*}\right|\right)=\int f^{prior}\left(\mathbf{p}\right)\cdot\left(\left|\zeta\left(p_i,p_j\right)\right|-I\right)^2d\mathbf{p}$$

Using Lemma 1, we can rewite such that for any possible signaling structure $s\left(p_i,\varepsilon_i\right)$ and $g\left(x_i,x_j\right)=g\left(\mathbf{x}\right)$,

$$\int f^{prior}\left(\mathbf{p}\right)\cdot\left(\left|\zeta\left(p_i,p_j\right)\right|-I\right)^2d\mathbf{p}$$

$$=\int\int f^{posterior}\left(\mathbf{p}|\mathbf{x}\right)\cdot g\left(\mathbf{x}\right)d\mathbf{x}\cdot\left(\left|\zeta\left(p_i,p_j\right)\right|-I\right)^2d\mathbf{p}$$

Rearranging the integration order gives:

$$\int \int f^{posterior}(\mathbf{p}|\mathbf{x}) \cdot g(\mathbf{x})\, d\mathbf{x} \cdot (|\zeta(p_i, p_j)| - I)^2\, d\mathbf{p}$$

$$= \int \int f^{posterior}(\mathbf{p}|\mathbf{x}) \cdot (|\zeta(p_i, p_j)| - I)^2\, d\mathbf{p} \cdot g(\mathbf{x})\, d\mathbf{x}$$

Hence, we have that

$$Var^{FI}\left(\left|\frac{y_i^* - y_j^*}{y_i^* + y_j^*}\right|\right) = \int \int f^{posterior}(\mathbf{p}|\mathbf{x}) \cdot (|\zeta(p_i, p_j)| - I)^2\, d\mathbf{p} \cdot g(\mathbf{x})\, d\mathbf{x}$$

(iii) We now want to show that $Var^{LI}\left(\left|\frac{y_i^*-y_j^*}{y_i^*+y_j^*}\right|\right) < Var^{FI}\left(\left|\frac{y_i^*-y_j^*}{y_i^*+y_j^*}\right|\right)$.
If there is no performance-ranking uncertainty:

$$Var^{LI}\left(\left|\frac{y_i^* - y_j^*}{y_i^* + y_j^*}\right|\right) = \int \left(\int f^{posterior}(\mathbf{p}|\mathbf{x}) \cdot |\zeta(p_i, p_j)|\, d\mathbf{p} - I\right)^2 \cdot g(\mathbf{x})\, d\mathbf{x}$$

This means that if in general

$$\left(\int f^{posterior}(\mathbf{p}|\mathbf{x}) \cdot |\zeta(p_i, p_j)|\, d\mathbf{p} - I\right)^2 < \int f^{posterior}(\mathbf{p}|\mathbf{x}) \cdot (|\zeta(p_i, p_j)| - I)^2\, d\mathbf{p},$$

then $Var^{LI}\left(\left|\frac{y_i^*-y_j^*}{y_i^*+y_j^*}\right|\right) < Var^{FI}\left(\left|\frac{y_i^*-y_j^*}{y_i^*+y_j^*}\right|\right)$.
(iv) We now show that in general:

$$\left(\int f^{posterior}(\mathbf{p}|\mathbf{x}) \cdot |\zeta(p_i, p_j)|\, d\mathbf{p} - I\right)^2 < \int f^{posterior}(\mathbf{p}|\mathbf{x}) \cdot (|\zeta(p_i, p_j)| - I)^2\, d\mathbf{p},$$

using the Jensen's inequality.

Jensen's inequality states that if X is a random variable and ϕ is a convex function, then $\phi(E(X)) < E(\phi(X))$. Let $\phi(y) = y^{*2}$ and $X = |\zeta(p_i, p_j)| - I$, and the result follows. Taking together (3) and (4), it follows that $Var^{LI}\left(\left|\frac{y_i^*-y_j^*}{y_i^*+y_j^*}\right|\right) < Var^{FI}\left(\left|\frac{y_i^*-y_j^*}{y_i^*+y_j^*}\right|\right)$.■

We now provide a counterexample showing that we cannot generally state that the variance in inequality implemented by a Bayesian meritocrat is lower in limited information compared to full information when there is performance-ranking uncertainty. Assume that the distribution of performance pairs (p_i, p_j) in the economy is $(30, 10)$ in $\frac{1}{4}$ of the situations, $(10, 30)$ in $\frac{1}{4}$ of the situations, and $(90, 10)$ in $\frac{1}{2}$ of the situations. With full information, the expected implemented inequality is $E^{FI}\left(\left|\frac{y_i - y_j}{y_i + y_j}\right|\right) = E\left(\left|\frac{p_i - p_j}{p_i + p_j}\right|\right) = \frac{1}{4} \cdot \left|\frac{30-10}{30+10}\right| + \frac{1}{4} \cdot \left|\frac{10-30}{30+10}\right| + \frac{1}{2} \cdot \left|\frac{90-10}{90+10}\right| = \frac{13}{20}$, and the variance in implemented inequality is $\frac{1}{4} \cdot \left(\frac{1}{2} - \frac{13}{20}\right)^2 + \frac{1}{4} \cdot \left(\frac{1}{2} - \frac{13}{20}\right)^2 + \frac{1}{2} \cdot \left(\frac{8}{10} - \frac{13}{20}\right)^2 = \frac{9}{400}$. Assume that with limited information, the signaling technology reveals the performance pair in the $(90, 10)$ situations, but makes it equally likely that the performance pair is $(30, 10)$ and $(10, 30)$ in the remaining situations. This implies that expected implemented inequality with limited information is $E^{LI}\left(\left|\frac{y_i - y_j}{y_i + y_j}\right|\right) = E\left(\left|E\left(\frac{p_i - p_j}{p_i + p_j}|\mathbf{x}\right)\right|\right) = \frac{1}{2} \cdot \left|\frac{1}{2} \cdot \frac{30-10}{30+10} + \frac{1}{2} \cdot \frac{10-30}{30+10}\right| + \frac{1}{2} \cdot \left|\frac{90-10}{90+10}\right| = \frac{8}{20}$, and the variance in implemented inequality is $\frac{1}{2} \cdot \left(0 - \frac{8}{20}\right)^2 + \frac{1}{2} \cdot \left(\frac{8}{10} - \frac{8}{20}\right)^2 = \frac{4}{25} > \frac{9}{400}$. Hence, the variance in implemented inequality can be higher with limited information than with full information with certain signaling technologies.

B Appendix: Supplementary empirical analysis

We here provide supplementary empirical analysis, as referred to in the main text.

B.1 Additional tables and figures

In Table B1, we show that the sample is balanced between treatments on the background characteristics.

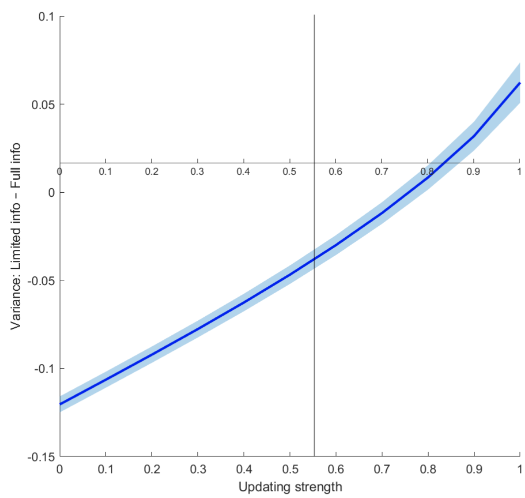
Table B1: Balance table

	FI	LI	FI vs LI Mann-Whitney
Female	65.07%	64.35%	$p = 0.8768$
Age	20.76 (3.17)	20.53 (3.30)	$p = 0.1687$
Right-wing	54.55%	60%	$p = 0.2404$
n	209	216	425

Notes: The table reports background characteristics by treatment and a Mann-Whitney test of whether the distributions are different. Female is an indicator variable for whether the participant is a woman, Age is the age of the participant in years, and Right-wing is an indicator variable for whether the participants self-reported voting for the Conservative Party or the Progress Party.

In Figure B1, we show the relationship between the updating strength and the treatment effect on the variance in implemented inequality.

Figure B1: Meritocrats: Updating strength and predicted treatment effects (variance)



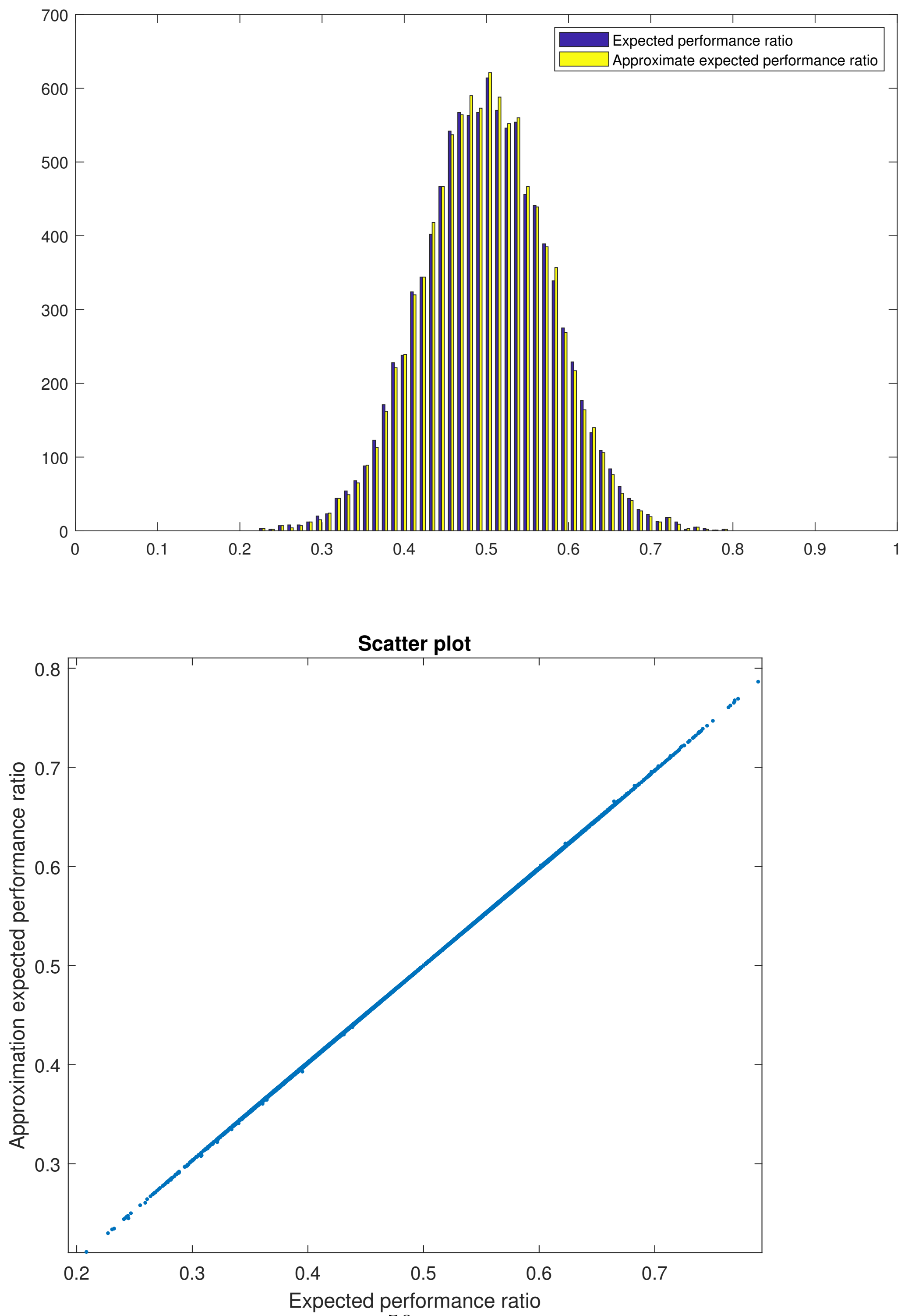
Note: The figure shows the relationship between the updating strength ρ and the predicted treatment effect on the variance in income gini implemented by the spectators. The light blue shaded area indicates the 95% confidence intervals.

We observe from Figure B1 that Bayesian meritocratic spectators implement lower variance in inequality in the Limited Information treatment than in the Full Information treatment. Signal-neglecting meritocrats also implement lower variance in the Limited Information treatment, whereas base-rate-neglecting meritocrats implement greater variance in the Limited Information treatment. In Appendix A.3, we show theoretically that Bayesian meritocrats always implement lower variance in inequality with limited information if there is no performance-ranking uncertainty. With performance-ranking uncertainty, we show that there exist counter-examples. However, as shown in Figure B1, in important economic environments, Bayesian meritocrats implement less variance in inequality also when there is performance-ranking uncertainty.

B.2 Expected performance inequality approximation

The expected value of the ratio of a set of normally distributed random variables is well approximated by the ratio of the expected values when the denominator is not close to 0. To show this for the experimental setting, where the distribution of worker performance is in the range $[10, 120]$, we simulate the performances by drawing from a normal random distribution with mean 56 and standard deviation 17, truncated between 10 and 120. The random factor is simulated by drawing from a normal distribution with mean 0 and standard deviation 15, truncated between -60 and 60. We simulate, with 10000 draws, both the distribution of $E(\frac{p_i}{p_i+p_j})$ (based on updating using the earnings signals x_i and x_j as information) and $\frac{E(p_i)}{E(p_i)+E(p_j)}$. In Figure B2, we show that the distributions of $E(\frac{p_i}{p_i+p_j})$ and $\frac{E(p_i)}{E(p_i)+E(p_j)}$ are almost identical in this environment, and thus it follows that expected performance inequality approximation holds.

Figure B2: Simulation of performance ratios



Notes: The upper panel shows the distribution of both the expected performance ratio and the approximated expected performance ratio based on the simulations. The lower panel shows a scatter plot of the approximated expected performance ratios against the expected performance ratios based on the same simulations as the upper panel.

B.3 Structural analysis

We estimate the following likelihood function:

$$L^h = \int_{-\infty}^{\infty} \int_{0}^{\infty} \left(\prod_{k=1}^{10} \frac{e^{V(y_{ik}, m_h, \beta, \rho)}}{\sum_{y_{jk} \in 0,1,..,X} e^{V(y_{jk}, m_h, \beta, \rho)}} \right) dF\left(\theta_\beta\right) dG\left(\theta_\rho\right)$$

with y_{ik} indicating the income given by the spectator to worker i in decision k. Note that we do not restrict the updating parameter in the behavioral model to be between 0 and 1, we allow it to range from minus infinity to infinity. The total likelihood contribution of a spectator is now given by:

$$L = \sum_h \lambda^h \cdot L^h$$

To numerically integrate over the log-normal distribution of β as a function of ζ and σ_β, the following approximation was used. For a given value of ζ and σ_β, the $[0, \infty)$ line was split in 20 equal-probability intervals, and for each interval, except the last, we calculated the contribution to the likelihood based on selecting the middle value of the interval for β. For the last interval, the lowest value of the interval was chosen. For the distribution of updating strength we assumed 41 bins with a width of 0.05 ranging from -1.05 to 2.05. There were 18 (out of 2160) situations in the Limited Information treatment where one of the worker's earnings turned out to be negative. For these cases, we interpreted the libertarian fairness view as giving all the income to the worker with positive earnings and 0 to the worker with negative earnings.

Robustness checks - structural estimates

In Table B2, we compare the structural estimates for the rational model (upper panel) and the behavioral model (lower panel) reported in Table 2 in the paper (first column, Table B2) to the structural estimates for a model that allows the distribution of fairness types to vary between the two treatment (right column, Table B2).

Table B2: Additional model estimations

Rational model	Types fixed across treatments	Types flexible across treatments
$\lambda^{Meritocrats}$	64.82% (2.58%)	80,91% (3,04%)
$\lambda^{Egalitarians}$	11,18% (1,71%)	4,16% (1,72%)
$\lambda^{Libertarians}$	24,00% (2,25%)	14,93% (2,68%)
$\lambda^{Meritocrats}LI$		48,44% (3,87%)
$\lambda^{Egalitarians}LI$		17,60% (2,83%)
$\lambda^{Libertarians}LI$		33,96% (3,55%)
ζ_β	-3,6420 (0,0968)	-3,6418 (0,0943)
σ_β	2,2278 (0,0893)	2.2288 (0,0886)
Log likelihood	-11956	-11935
Log likelihood FI	-5891,1	-5880,1
Log likelihood LI	-6064,6	-6055,0

Behavioral model	Types fixed across treatments	Types flexible across treatments
$\lambda^{Meritocrats}$	81.22% (2.87%)	81.04% (3.03%)
$\lambda^{Egalitarians}$	3,87% (1,28%)	4.08% (1.71%)
$\lambda^{Libertarians}$	14,91% (2,68%)	14.88% (2,67%)
$\lambda^{Meritocrats}LI$		81.36% (29,03%)
$\lambda^{Egalitarians}LI$		3,59% (1,92%)
$\lambda^{Libertarians}LI$		15,05% (42,99%)
ζ_β	-3,0636 (0,1093)	-3.0632 (0.1098)
σ_β	2,8841 (0,0855)	2.8846 (0.0865)
ρ_μ	0.4678 (0.0234)	0.4675 (0.0218)
ρ_σ	0.1842 (0.0216)	0.1845 (0.0217)
$\rho_{signal\,neglect}$	0.0993 (0.0403)	0.1010 (0.0417)
$\rho_{base-rate\,neglect}$	0.2864 (0.0916)	0.2840 (0.7280)
Log likelihood	-11783	-11783
Log likelihood FI	-5903	-5903
Log likelihood LI	-5879,8	-5879,8

Notes: This table reports structural estimates for the rational model (upper panel) and the behavioral model (lower panel). The left column reports the structural estimates reported in the main paper (Table 2); the right column reports the structural estimates for the model when we allow the distribution of fairness types to vary between treatments.

In a further robustness check, we have also estimated the behavioral model (reported in Table 2 in the main text) when the updating strength is restricted to be between 0 and 1. The estimated distribution of fairness types and the beta distribution are nearly identical to the structural estimates reported in

Table 2, but with a slightly lower log-likelihood score of 11790.

Updating strength and implemented inequality

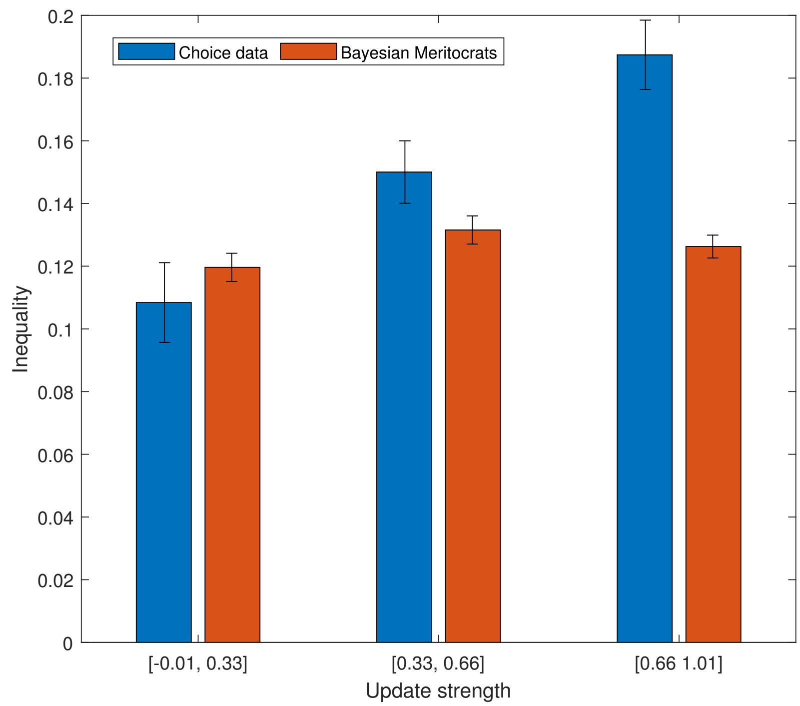
We here show that the relationship between the updating strength and implemented inequality reported in Figure 3 in the main text does not reflect that there is a correlation between updating strength and the nature of the distributive situations. In Figure B3, we show that a Bayesian meritocrat would implement almost the same level of inequality across the distributive situations of the three groups. The blue bars correspond to Figure 3, while each red bar shows what a Bayesian meritocrat would implement if making decisions in the situations captured in the corresponding blue bar. We observe that a Bayesian meritocrat would implemented almost the same inequality across the three groups.

B.4 The economic environment

We here provide some further description of the economic environment. In the upper part of Figure B4, we show the distributions of worker performance and earnings. We observe that worker performance is (approximately) normally distributed in the range of $[10, 120]$, with a mean of $\mu_p = 56.13$ and a standard deviation of $\sigma_p = A\sigma_\varepsilon = 17.22$. The range in worker performance and the range in the random component imply that there is performance-ranking uncertainty in the posterior beliefs of Bayesian spectators, and, consequently, fairness-ranking uncertainty for Bayesian spectators with a meritocratic fairness view. Expected performance inequality approximation holds in this environment because we do not have any performances close to 0,(see Appendix B.2.

In the lower part of figure B4 below, we show the distributions of performance inequality and earnings inequality in the worker pairs. We observe that there is large heterogeneity in both inequalities, there are worker pairs with large inequality in performance (earnings) and worker pairs with no inequality in performance (earnings). There is greater variance in the distribution of

Figure B3: Updating strength and implemented inequality

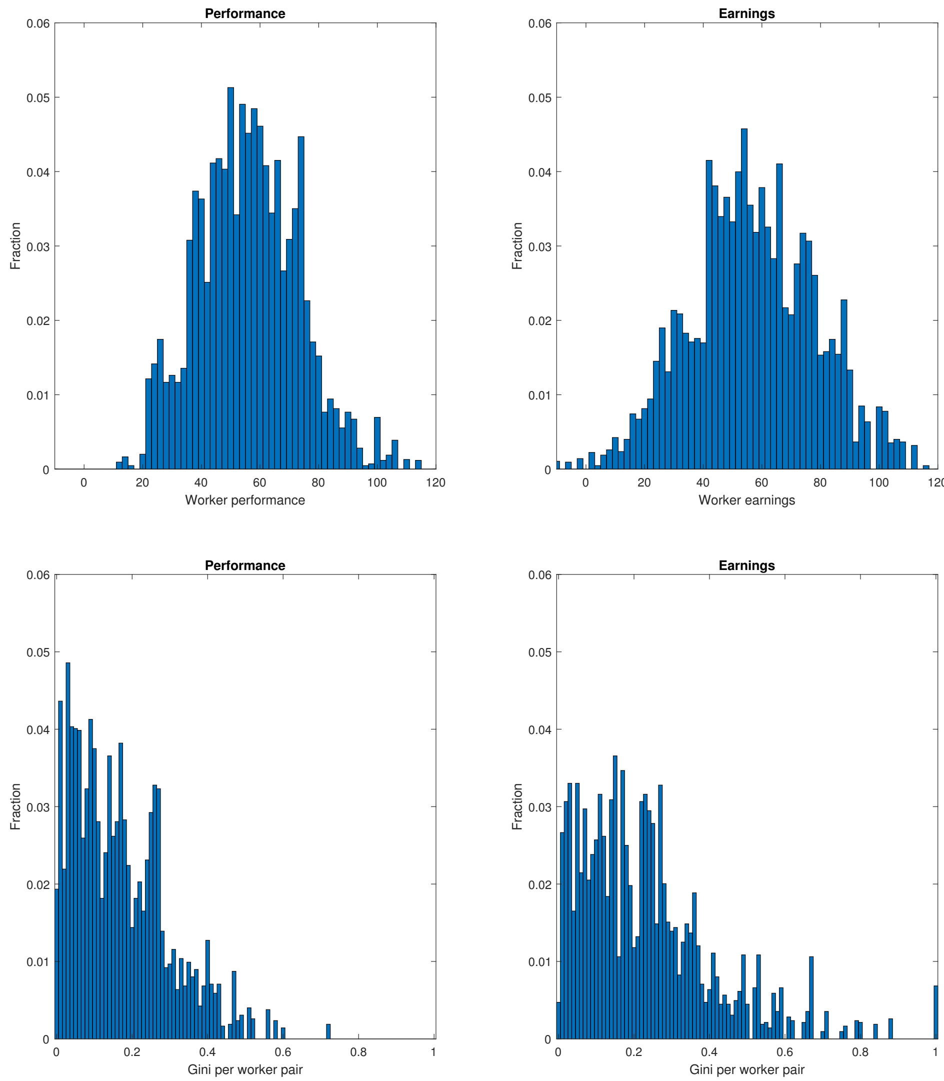


Notes: This figure shows the level of average implemented inequality by spectators classified as meritocrats based on their decisions in the Limited Information treatment, split into three groups based on their updating strength. The classification of spectators into meritocrats and non-meritocrats is as follows: We classify a decision as meritocratic if $\left|\frac{y_i}{y_i+y_j} - \frac{E(p_i|x_i)}{E(p_i|x_i)+E(p_j|x_j)}\right| < 0.1,$ *with* $E(p_i|x_i)$ *being inferred from the spectator's stated beliefs at the end of the experiment. A spectator is classified as a meritocratic if 6 or more of the decisions are classified as meritocratic. Of the 216 spectators in the Limited Information treatment, 186 are classifed as meritocrats according to this procedure. For each of the classified meritocratic spectators, we calculate the average implemented inequality over their 10 decisions, and their updating strength* ϱ *using equation (15). The blue bars in the figure report the average implemented inequality by the spectators with average updating strength between -0.01 and 0.33 (30 spectators), 0.33 and 0,66 (44 spectators), and 0.66 and 1.01 (71 spectators). The red bars indicate what would have been the average implemented inequality in each of the three groups if everyone was a Bayesian spectator. Standard error bars are indicated.*

earnings inequality than in the distribution of performance inequality, which follows from earnings partly being determined by the random component.

In sum, the economic environment in the experiment represents a special case of the general theory. It focuses on situations with performance-ranking uncertainty, an additive earnings function, (approximately) normal distributions, the same prior belief distribution for both workers, independence between the worker performances, and independence between worker performance and the random factor.

Figure B4: The economic environment in the experiment

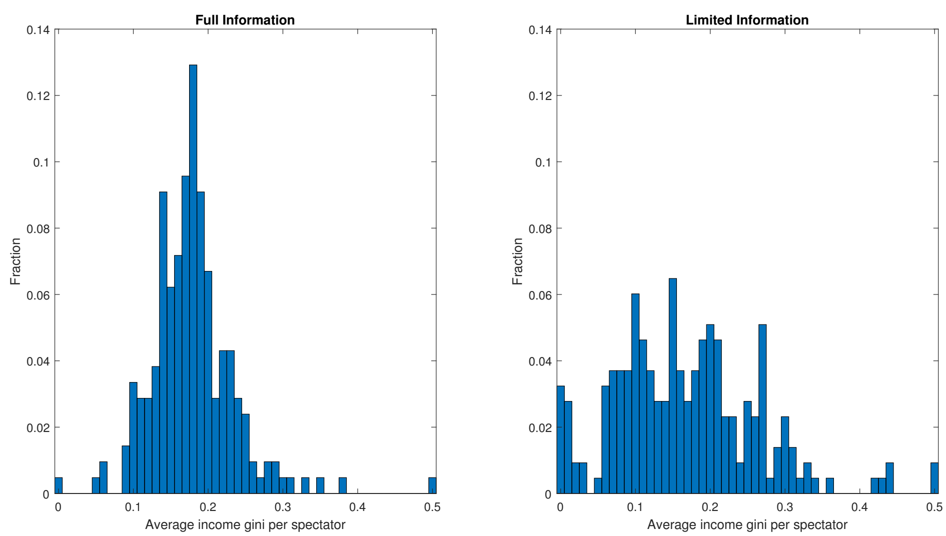


Note: The upper panels show the distributions of worker performance and worker earnings. The lower panels show the distributions of performance inequality and earnings inequality in the worker pairs.

B.5 Additional results

We here provide some additional descriptive statistics of the spectator behavior. In Figure B5, we observe that there is substantial variation in the average inequality implemented by the spectators in their 10 decisions when they have full information. The large majority of spectators implement an average inequality between 0.1 and 0.3, but we also observe some spectators dividing equally in all situations. Only three spectators implemented an average inequality larger than 0.5 gini points. The variation across spectators in this treatment is likely to reflect both the different situations to which they were assigned and differences in fairness views. In the upper-right panel, we observe that limited information increases the variation in spectator behavior (variance ratio test, $p < 0.001$), in contrast to what we would expect if all spectators were Bayesian meritocrats.

Figure B5: Histograms: implemented income inequality across spectators, by treatment



In figure B6 below, we show the distribution of implemented inequality across all situations by treatment. In the large majority of situations, the spectators implement a gini between 0 and 0.5; in about 2.8 percent of the situations, the spectators implement an equal division of the earnings, whereas in about 3.6 percent of situations, they implement an inequality above 0.5 gini points. We observe that limited information causes an increase in the share of situations where the spectators implement an equal division, but also an increase in the share of situations where the spectators implement maximal inequality. Overall, we observe that limited information causes a significant increase in the variance in implemented inequality across situations (variance ratio test, $p < 0.001$).

Figure B6: Histograms: implemented income inequality across situations, by treatment

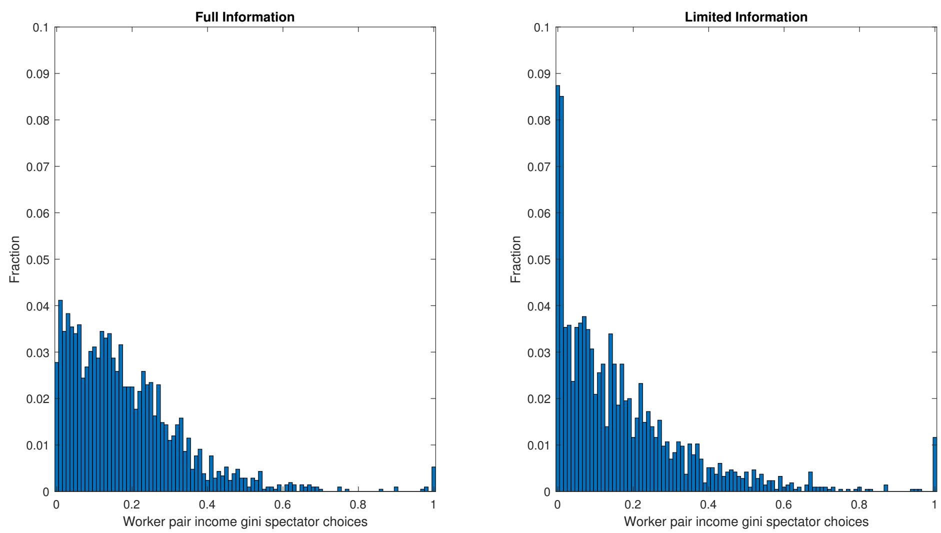
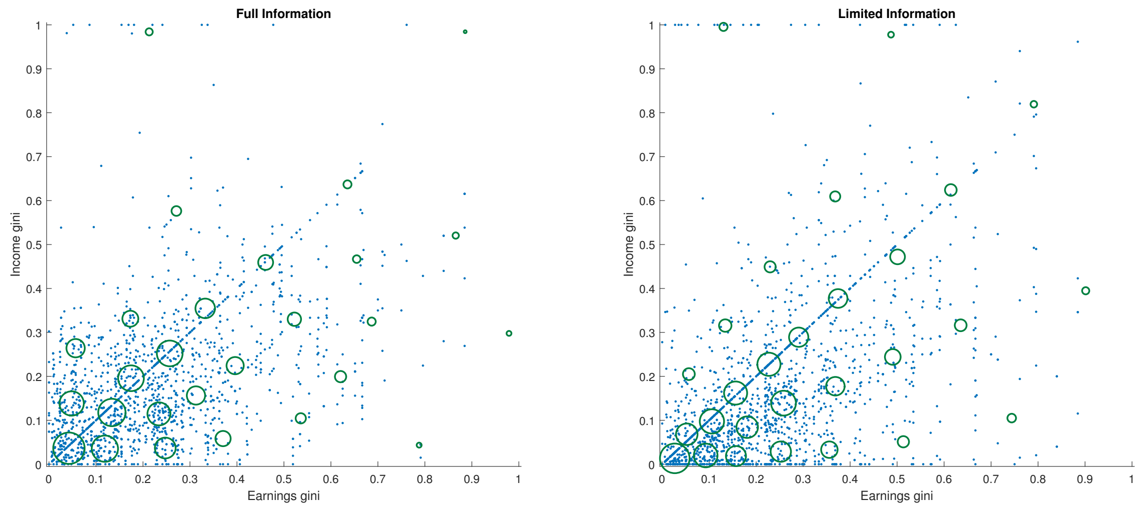


Figure B7: Income inequality on earnings inequality



Note: The figure shows the relationship between income inequality and earnings inequality, by treatment.

In the right panel of the above figure B7, we observe that the great majority of the spectators (83.84 percent) implement an income inequality within the range of the egalitarian solution and the libertarian solution, as predicted by the theoretical framework.

C Appendix: Instructions for spectators and Mturk workers

C.1 Spectator instructions

<u>Screen 1:</u>
Welcome to the experiment.

What is your desk number (ping-pong ball)?
What is your spectator number (piece of paper)?

<u>Screen 2:</u>
You have entered the following:
Desk Number (ping-pong ball) X
Spectator Number (piece of paper) X
If this does not match the numbers you were given, please stay on this page and call an experimenter!
Otherwise, click >> to begin the experiment.

<u>Screen 3:</u>
Introduction

Welcome to this experiment.

The results from this experiment will be used in a research project. It is therefore very important that you follow certain rules of conduct. You are not allowed to communicate with any of the other participants during the experiment. If you have a question, please raise your hand and an experimenter will come and assist you. All electronic devices must be turned off, and it is not permitted to access any programs on your computer other than the one that we use for this experiment. Everyone has received a copy of these rules of conduct. Those who violate the rules will be asked to leave the experiment.

You will be completely anonymous throughout the experiment. This means that you will not be asked to reveal your identity at any time during the experiment, and your decisions will only be linked to your table number so neither the experimenters nor any other participant will find out what decisions you have made.

At the end of the session you will be given an envelope corresponding to your computer/table number, which contains your payoff for the session. The person who has prepared the envelope will not be in the room when the envelopes are distributed, which ensures that no one can identify how much each of you have received in the session.

<u>Screen 4:</u>
You will receive a fixed amount of 50 NOK in cash as a participation fee. During the course of the experiment, you will also have the opportunity to receive additional money as a bonus.

Throughout the experiment you will be able to earn points. At the end of the experiment the points will be exchanged at the rate of:

1 point = 5 NOK.

We will ask you to make a number of decisions in this experiment. Before each decision situation, you will receive instructions, starting on the following page.

Screen 5:
A guessing question

On the following screen, we will ask you a question based on a short scenario description.

The closer your answer is to the correct value, the more extra points you will earn for this task. Please only enter a number into the text box and be aware that you are entering percentages (i.e. if you want to answer 50%, then just type in 50). Anything else (including if you leave it empty) will not be counted as a real answer and you will not earn any points for your answer.

You have 2 minutes for the question. After the timer has run out, you will auto-advance to the next screen.

Screen 7:
Question:

Imagine a test to detect a disease of which we know 1 in 1000 people are infected with the disease. The test correctly detects whether or not a person is infected with the disease 95% of the time.

If a random person takes the test, and the result of the test is positive, what is the chance that this person is infected with the disease?

Screen 8:
Thank you for your answer, it has been recorded and any points you earn will be added to your final payout.

Screen 9:
A number of distributive decisions

Please read the following instructions carefully. At the end of the instructions we will ask you to answer a number of comprehension questions regarding these instructions.

Last week, we had a number of people work on an online real effort task. In the following, you will be asked to make 10 distributive decisions regarding how to pay these individuals for their work. Please be aware that many of the distributive decisions made in this experiment will actually be implemented and determine the payout for a pair of workers

For each distributive decision that you have to make, you will be matched with two anonymous workers, who have both worked on the same information processing task to solve as many encryption problems as possible in 15 minutes. The task consisted of correctly decoding three encrypted letters. Below, you can see an example of what such a task looked like:
> to zoom in, press: Ctrl and +
> to zoom out, press: Ctrl and -

LETTER:	CODE:
N	
U	
Z	

G	I	T	V	C	E	H	K	P	A	L	J	X	N	D	Y	M	O	W	Q	Z	B	R	S	U	F
834	979	149	186	107	609	948	789	796	891	782	955	385	769	254	271	471	534	740	998	544	676	251	334	834	124

A worker's **performance** is equal to the number of encryptions that the worker correctly solved. Each worker's **earnings** is his or her performance plus a **random factor. The total earnings** for the pair is the sum of the individual earnings. In summary:

Worker 1's earnings = Worker 1's performance + Worker 1's random factor.
Worker 2's earnings = Worker 2's performance + Worker 2's random factor.

Total earnings = Earnings of Worker 1 + Earnings of Worker 2

Your decision
The workers have been told that a third person will determine how the total earnings will be allocated between them. You are the third person and you will determine how the total earnings will be allocated between the two workers. Again, your decision is completely anonymous. The workers will not know who you are and you will not know the identity of the workers.

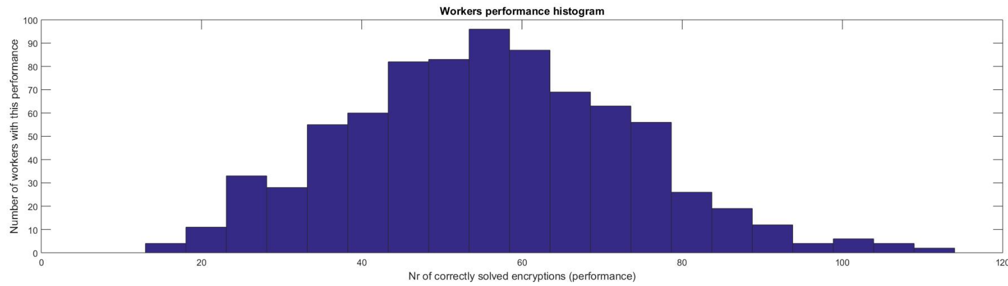
[FI: You will receive information regarding the performance, random factors and earnings of the two workers.]

[LI: **You will receive information <u>only</u> regarding the earnings of the two workers**.]

Please note that your allocation choice <u>will not</u> affect your payout in any way, it may only determine the payment to these two other workers.

<u>Screen 10:</u>
Worker performance

The worker's performance on the task is the number of correct 3 letter encryptions solved by the worker in the 15 minutes he or she worked on the task. Below, you can see the distribution of the performances recorded last week for all the workers:
> to zoom in, press Ctrl and +
> to zoom out, press Ctrl and -

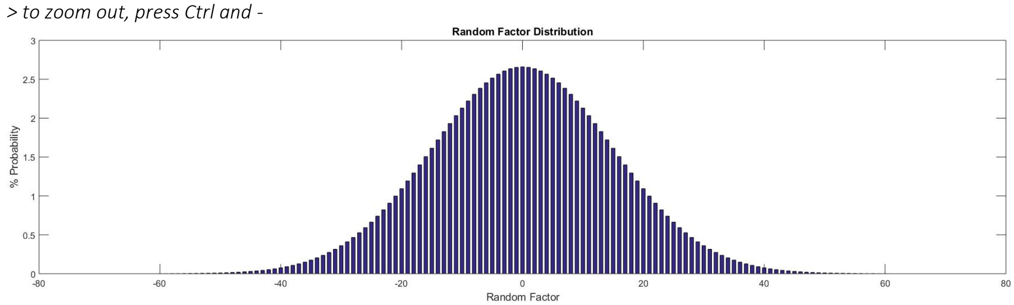


For each performance level (number of correct tasks), you can see the number of workers who completed the task with that performance level during the 15 minutes. For example, you can see that on average the workers completed 50 tasks, while most of the workers performed between 40 and 60 tasks.

The random factor

To determine a worker's earnings, we will draw a random factor to add to the production. This factor can be positive or negative. The random factor will be drawn from the following distribution:

> to zoom in, press Ctrl and +
> to zoom out, press Ctrl and -



To give you some more intuition about the distribution each worker's random factors will be drawn from, here are a few indicators:

- The probability that a drawn random factor is equal to -1, 0 or 1 is equal to 8%
- The probability that a drawn random factor lies between (and including) -5 and 5 is equal to 29%
- The probability that a drawn random factor lies between (and including) -10 and 10 is equal to 52%
- The probability that a drawn random factor lies between (and including) -20 and 20 is equal to 83%
- The probability that a drawn random factor is larger than 20 or smaller than -20 is equal to 17%

By clicking the next button, you will go on to a questionnaire to check your understanding of the instructions. You can only proceed to the experiment if you have answered all of the comprehension questions correctly. If you have a question, please raise your hand.

Screen 11: **Comprehension checks**

To make sure that you understand the instructions, please answer the following comprehension questions. Please state your answers to the following questions and click the next button to submit your answers. If all of your answers are correct, you can proceed to your choices.

- The performance of a worker is equal to the number of correct encryptions that he or she managed to do in 15 minutes. (TRUE)
- The earnings of Worker 1 (W1) are equal to: (A) the performance of W1 /(B) the random factor of W1 /(C) W1's performance + W1's random factor" (C)
- The probability that the random factor is equal to either -10 or 10 or any number in between is larger than 60%. (FALSE)
- The random factor is the same for all workers. (FALSE)
- Total earnings for a pair of workers are equal to the sum of the earnings of Worker 1 and the earnings of Worker 2. (TRUE)
- Your payout will depend on how you distribute the total earnings between the workers. (FALSE)

Screen 12: **Choices**

You are now ready to make your ten distribution choices. Please look at the displayed table before making your distributive decision. Below, you can then state how much of the total earnings you want to allocate to Worker 1 and how much of the total earnings you want to allocate to Worker 2.

Once you have made each of the 10 decisions, you will be shown a summary page. On this page, you can edit any of the choices. Once you have clicked the "Confirm & Submit" button on that page, your distribution choices will be final.

Click >> to advance to the choices.

[Choices 1-10]

Screen 13:
You now have the opportunity to review your choices before submitting them. You can either edit your decisions in the summary below, or leave them as they are.

Once you are ready, press the 'Confirm & Submit' button below to confirm your choices you would like to submit.

[Summary Choices 1-10 on one page]

Screen 14 (only for LI treatment)**: Posterior belief elicitation**

We would now like to ask you to make a guess about the performance of the workers you were matched with. Please be reminded that performance refers to the number of correctly solved encryptions, while the number displayed for each worker shows their earnings (= performance + random factor).

You will earn extra points depending on how good your guesses are, and these points will be paid out to you at the end of the experiment (1 point = 5 NOK). The closer your guesses are to the actual values, the more extra points you will earn for this task.

[Choices 1-10 on one page]

Performance = number of correctly solved encryptions
Earnings = Performance + Random factor

Earnings of Worker 1: X
Earnings of Worker 2: X
What is your best guess of the performance of Worker 1?
What is your best guess of the performance of Worker 2?

Screen 15:
Thank you for your input. The choice that will be randomly implemented is Choice ${e://Field/item}.

Screen 16:

Demographics

Your answers have been recorded.

Please answer the following questionnaire.

What is your age? [number input]

What is your gender?
-Male
-Female
-Other/prefer not to answer

Which political party would you vote for if there were an election tomorrow?
-Arbeiderpartiet
-Høyre
-Fremskrittspartiet
-Kristelig Folkeparti
-Senterpartiet
-Venstre
-Sosialistisk Venstreparti
-Miljøpartiet De Grønne
-Rødt
-Other

I generally see myself as a person who likes to take risks.
-Strongly agree
-Agree
-Somewhat agree

-Neither agree nor disagree
-Somewhat disagree
-Disagree
-Strongly disagree

Please indicate to what degree you agree with the following statements. [Scale 1-7]

I think basically the world is a just place.
I believe that, by and large, people get what they deserve.
I am confident that justice always prevails over injustice.
I am convinced that in the long run people will be compensated for injustice.
I firmly believe that injustices in all areas of life (e.g., professional, family, politics) are the exception rather than the rule.
I think people try to be fair when making important decisions.

What was the main motivation for the distribution decisions that you made in this experiment?

Payout

Thank you for completing this experiment. On the next page, you will find your payout information.

In the disease question, you earned: ${e://Field/base_points} exp. points

[only for LI treatment:]
The sum of your points from guessing the workers' performances is: ${e://Field/post_points} exp. points

These points have been added to your final payout.

Your complete payout (base pay + potential bonus): X NOK

When you click the next button below, you will be automatically redirected to a further thought experiment and two follow-up questions.

Please stay in your seat and work on this task (and when you have finished, wait quietly for further instructions). The experimenters will bring your payment envelope to your desk. Once you have received your envelope, you may exit without speaking to any of the other participants.

C.2 Worker Instructions

Screen 1:

Introduction
Welcome to this research project! We very much appreciate your participation. In this study, you will have the opportunity to earn money by solving simple encryption tasks. The full amount you can earn depends on how quickly and carefully you solve the tasks.

Payment
You will have **15 minutes** to work on the encryptions. During this time, you are asked to complete as many encryptions as you can. For completing this study (which includes solving a minimal number of 20 tasks), you will receive a base payment of $2. **If you do not fulfill this minimum requirement, you will not receive the base payment.**

Bonus
In addition to the base payment, all participants can earn a potential bonus. How high this bonus is will depend in part on how hard you worked on the task. This bonus can be quite substantial compared to the base payment. Note that you will still have the opportunity to earn this bonus even if you did not fulfill the requirements for the base payment. However, be aware that the bonus will depend in part on the number of tasks you completed.

Your performance will be equal to the number of encryptions that you solved. In addition, you will receive a random factor that is drawn randomly by the computer. Your experiment points will be the sum of your performance + your random factor. At the end of the experiment, these points, and in some cases the decisions of others, will be taken as the basis for the distribution of your bonus. If you are paired up with another worker who completed this same task, a third person will decide how much of the total points will be allocated as a bonus to you and how much of the total points will be allocated to the other participant.

Participation
Participation in this research study is completely voluntary. You have the right to withdraw at any time or refuse to participate entirely without jeopardy to future participation in other studies conducted by us.

Questions about the Research
If you have questions regarding this study, you may contact thechoicelab@nhh.no

Thank you for participating!

Screen 2:

Before you start the study, you will complete 2 trial (practice) encryptions. This will give you the opportunity to familiarize yourself with the task. After you have successfully completed the trial block, you will be able to work on the same type of tasks for money. The encryption task is explained in further detail on the next page.

Screen 3:

Trial Instructions

The task involves solving encrypted letters. Here is an example of what an encryption can look like:

LETTER:	CODE:
N	
U	
Z	

G	I	T	V	C	E	H	K	P	A	L	J	X	N	D	Y	M	O	W	Q	Z	B	R	S	U	F
834	979	149	186	107	609	948	789	796	891	782	955	385	769	254	271	471	534	740	998	544	676	251	334	834	124

In each encryption, there are three letters for you to code. These letters are indicated on the left side of the page, with an empty text entry field for your code entry. The corresponding 3-digit codes for each letter can be found at the bottom of the page. Your task is simply to type in the correct code for each letter, and hit the next button.

You will not be able to pass to the next encryption until you have solved the current task correctly. Once all three code answers are correct, you will advance to the next trial and your number completed encryptions will increase.

Screen 4:

[trial encryptions]

Screen 5:

Encryption Block Instructions

Now that you have completed the practice trials and are familiar with the task, you are ready to start the actual study.

You will now have 15 minutes to solve as many encryption tasks as you can. The timer starts when you click the next button on this page. On the top of the page, you will see your starting time as well as the time when you began the current trial (the time will remain static for the duration of each trial, and will update when you move to the next trial). As soon as you enter an answer for one task, a new encryption problem will appear until the 15 minutes have run out.

For each encryption that you solve correctly, you will receive 1 experimental point. As explained before, at the end of the experiment, your potential bonus will be derived in part from this sum of correct encryptions.

Once the timer has run out, a summary of your work will be displayed. **Remember that you must complete at least 20 correct encryptions to receive the base payment.**

Click >> to begin working on the task for money.

Screen 6:

[encryptions task]

Screen 7:

Time is up!

Your number of correct tasks has been recorded. Please click the next button to see your results.

Screen 8:

Your performance (number of correctly solved encryptions): **${e://Field/counter}**.

If you have fulfilled the minimal requirements for this survey, you will receive the base payment of $2 within the next days.

You will also receive your bonus payment within the next couple of weeks, after the experiment has ended.

Thank you for participating!

Screen 9:

Do you have any comments for the researchers?